Climate Policy at the Crossroads

Inequality, Trade, Uncertainty, and Incentives for Climate Change Mitigation

Thomas Bourany

Dissertation

in candidacy for the degree of

Doctor of Philosophy (Ph.D.) in Economics

The University of Chicago

Division of Social Sciences

Department of Economics

June 2025

Dissertation Committee

Professor Mikhaïl Golosov (Chair)
University of Chicago
Department of Economics

Professor Esteban Rossi-Hansberg
University of Chicago
Department of Economics

Professor Lars Hansen
University of Chicago
Department of Economics, Statistics and Booth School of Business

Professor Michael Greenstone
University of Chicago
Department of Economics and Energy Policy Institute of UChicago

1 Climate Change, Inequality, and Optimal Climate Policy

2 The Winners and Losers of Climate Policies: A Sufficient Statistics Approach

joint with Jordan Rosenthal-Kay

3 The Optimal Design of Climate Agreements Inequality, Trade, and Incentives for Climate Policy

Job Market Paper

Climate Change, Inequality and Optimal Climate Policy

Work in progress

Most recent version accessible here – Updated regularly

Thomas Bourany*
The University of Chicago

First draft: December 2022
This draft: April 5, 2025

Abstract

What is the optimal policy to fight climate change? Taxation of carbon and fossil fuels has strong redistributive effects across countries: (i) curbing energy demand is costly for developing economies, which are the most affected by climate change in the first place (ii) carbon taxation has strong general equilibrium effects through energy markets and fossil fuel rents. Through the lens of an Integrated Assessment Model (IAM) with heterogeneous countries, I show that the optimal taxation of carbon depends crucially on the availability of redistribution instruments. After characterizing the Social Cost of Carbon (SCC), I provide formulas for the Second-Best carbon tax in the presence of inequalities in incomes and climate damages, and redistributive and distortionary effects on energy markets. I show that a uniform carbon tax should be reduced by approximately 15% in the presence of inequality compared to First-Best where cross-country transfers are available. If country-specific carbon taxes are available, the distribution of carbon prices is proportionally related to the level of income: poor and hot countries should pay lower energy taxes than rich and cold countries. These qualitative results are general, and I propose a dynamic quantitative model to provide recommendations for the optimal path of carbon taxes.

*Thomas Bourany, *thomasbourany@uchicago.edu*.
I thank my advisors, Mikhail Golosov, Esteban Rossi-Hansberg, Lars Hansen, and Michael Greenstone, for their valuable guidance and advice. I also thank Aditya Bhandari, Jordan Rosenthal-Kay, and other seminar participants at UChicago, Booth, IMSI, NBER Summer Institute, Society of Economic Dynamics 2024, and European Economic Association Annual Meetings 2024, for comments and stimulating discussions. All errors are mine.

1 Introduction

Greenhouse gas emissions generated by economic activity are causing climate change, and global atmospheric temperatures have increased by almost $1.5°C$ since the Industrial Revolution. The sources of these emissions are unequally distributed: developed economies account for over 65% of cumulative greenhouse gas (GHG) emissions with $\sim$ 25% each for the European Union countries and the United States, while some developing countries have barely emitted anything compared to their population level. Moreover, carbon emissions and energy consumption tend to correlate highly with development and income (e.g. GDP per capita).

Moreover, the consequences of global warming are also unequal: the increase in temperatures disproportionately affects developing countries where the climate is already warm. Most emerging and low-income economies lie geographically closer to the tropics and the equator and tend to be most vulnerable to global warming, e.g. Burke et al. (2015), Carleton et al. (2022).

Finally, implementing climate policy in the form of Pigouvian carbon taxation has strong redistributive effects as well. Countries that consume a large share of their energy mix in oil, gas, and coal will be affected more by distortionary carbon taxation. Moreover, phasing out fossil fuels reduces energy prices, lowering energy rents and hurting disproportionally exporters of fossil fuels.

These three layers of inequalities raise the following question: what is the optimal carbon policy in the presence of climate externality and inequality? Should the optimal tax on carbon and fossil fuels account for these different dimensions of heterogeneity?

To answer this question, I develop a simple yet general Integrated Assessment Model with countries' heterogeneity. Individual countries are heterogeneous in many dimensions, including (i) income, (ii) damages from climate change, and (iii) exposure to energy markets through differences in energy mix and fossil-fuel exports. Since the quantitative framework is very general, I first provide an extremely simple model to provide the main theoretical intuitions, keeping the same features regarding climate externalities and energy markets.

In both models, I study the design of optimal taxation of carbon and the characterization of the Social Cost of Carbon (SCC), which summarizes the costs of climate change for one additional ton of carbon emitted. I show that the optimal policy depends crucially on the availability of redistributive instruments – such as lump-sum transfers – across countries. In the First-Best, the optimal tax follows the Pigouvian benchmark and equals the Social Cost of Carbon. However, when cross-country transfers are not allowed, the optimal tax needs to account for inequality and redistributive considerations and adjust the level of the uniform carbon tax. Moreover, when choosing country-specific carbon taxes, the optimal policy is to lower the carbon tax for poorer countries, and higher for more advanced economies.

I first show these two main results in a simple "toy model", where I summarize these lessons in a four-equations static model. Differences in TFP, and thus income, in the impacts of climate change, and in the costs of extracting fossil fuels provide a rationale for redistribution. The unconstrained planner – in the First-Best – uses transfers to offset the redistributive effects of the carbon tax itself. However, in the Second-Best, when these transfers are assumed away, we see that the uniform carbon should account for three effects: (i) the climate externality represented by the Social Cost of Carbon (SCC), (ii) a supply redistribution that summarizes the equilibrium effect on the price of fossil-fuel energy which redistribution between exporter and importers, scaled by the inverse elasticity of the energy supply, and (iii) a demand distortion term that scale with energy inputs choices, shares, and demand elasticity. Moreover, (iv) these three terms are the aggregation of these effects for each country, weighted by the "*social welfare weights*", which is the product of the marginal utility of consumption and the Pareto weights. The planner puts more weight on poorer countries than on more advanced ones. To summarize, the world's optimal carbon policy differs from the standard *Carbon tax = Social Cost of Carbon*, and the taxation should be adapted to the specific situation of each country.

Second, in this static model, I show how to choose country-specific carbon taxes. In that case, two motives – the Social Cost of Carbon and Supply Redistribution – remain, and Demand Distortion disappears thanks to the ability of the planner to act independently in each country. However, the carbon itself is now inversely proportional to the social welfare weights: the planner strongly reduces the carbon tax required for countries with high Pareto weights or high marginal utility of consumption – i.e. relatively poorer countries. I show in this simple setting how to implement such policies in emission trading systems or cap-and-trade. Moreover, I demonstrate that a direct mapping exists between price instruments – like carbon taxes or carbon prices – and quantity regulations – as discussed in Weitzman (2003), and Weitzman (2015).

I then build a quantitative dynamic model to provide policy recommendations to these questions. I consider an Integrated Assessment Model extending the standard Neoclassical Growth Model with multiple countries and many dimensions of heterogeneity. First, in each country, a representative firm produces a final good using capital, labor, and energy inputs. Countries differ in total productivity and energy efficiency, which implies differences in incomes or GDP/capita and total energy demand. Second, a representative household makes consumption, capital, and borrowing decisions over time. Third, there are three energy firms in each country: oil-gas, coal, and renewable, which are energy inputs used in the final good production. Countries differ in energy mix due to differences in costs of energy production. Moreover, differences in endowments in oil-gas – and dynamically depleting reserves – lead to countries being exporters or importers of fossil fuels. Finally, global fossil-fuel consumption – from oil-gas and coal – emits carbon into the atmosphere, which then feeds back into the climate system. This affects temperatures across regions and has heterogeneous damages across countries for firms and households. I calibrate the model to economic, climate, and energy data for a sample of 68 countries to match the dimensions of heterogeneity at the heart of countries' vulnerability to climate change and climate policies.

Despite the richness of the model and the many market forces and general equilibrium effects, the main result on the optimal carbon policy carries through. First, relying on the continuous time formulation of the model, I provide an analytical characterization of the Social Cost of Carbon (SCC) – which depends on the climate system and damage parameters but also the differences in social welfare weights. I show that in the First-Best allocation, in the absence of redistributive motives, the Social Cost of Carbon is the sum of the Local Costs of Carbon for each country, weighed by the planner's Pareto weights. This result aligns with models that can be perfectly aggregated where redistribution motives are absent. However, in the Second-Best, in cases where the planner cannot undermine preexisting inequality or offset redistributive effects, the Social Cost of Carbon is the sum of Local Costs of Carbon for each country, weighed by the *social welfare weights* which now integrate differences in marginal utility of consumption. If poorer countries are the most affected by climate change, the Social Cost of Carbon would be higher. However, in this class of models, the Local Costs of Carbon scale with consumption and income, which implies that the Social Cost of Carbon is lower than in the First-Best or the Representative Agent economy.

Then, I characterize the optimal carbon policy. In the First-Best allocation, we recover the Pigouvian benchmark, where the carbon tax equals the Social Cost of Carbon. This is because the planner redistributes across countries using lump-sum transfers, for example, taxing lump-sum European and American countries and transferring to South Asian and African economies. In the Second-Best Ramsey policy, when the planner is unable to redistribute freely across countries due to limitations on lump-sum transfers, the world optimal carbon tax needs to be adjusted. As before, it needs to account for Supply Redistribution – depending on energy supply curve elasticities – and Demand Distortion – which is a function of the energy demand elasticities. Finally, I also show that country-specific carbon taxes are scaled by the inverse of the social welfare weights, such that the poorer or warmer the country, the lower the carbon tax it needs to pay. Moreover, in the quantitative model, these terms are slightly more involved as they depend on the path of temperature – as the Social Cost of Carbon increases over time, as is common in carbon taxation in IAM – and the substitution patterns between energy sources and change dynamically. Similarly, the social welfare weights used for aggregating these effects across countries also change over time.

In addition, this framework is general, and I develop a method inspired by the Heterogeneous Agents literature and Mean-Field Games to solve this class of model globally in continuous time. This relies on the sequential formulation of the optimal control problem, which allows to follow the trajectories of each country/agent. As a result, it allows us to consider an arbitrary number of dimensions of ex-ante heterogeneity and a larger number of dimensions of time-varying states than what is typically covered in the literature using dynamic programming methods.

The main quantitative result is that accounting for inequality implies changing the optimal carbon tax in three ways. First, computing the Social Cost of Carbon with the social welfare weights results in a SCC of \$210/tCO_2 instead of \$430/tCO_2 when simply doing the simple sum of Local Cost of Carbon. Second, implementing the carbon mitigation does reduce the Social Cost of Carbon – which is an equilibrium object depending on climate damage. I show that in the First-Best, the planner would use large transfers to offset inequality. In the Second Best, when redistribution instruments are absent, the SCC is approximately 12% lower, from \$170 to \$150. This results from the fact that the carbon tax puts more weight on poorer countries that have a higher marginal value of wealth. Finally, it accounts for redistribution motives in the energy markets – supply redistribution and demand distortion. It implies, on the one hand, a lower carbon tax to avoid hurting fossil-fuel exporters and, on the other hand, a higher tax since the richest countries are the major consumers of fossil fuels. Quantitatively, these two effects cancel out despite the net effect being slightly negative. This implies a carbon tax slightly below 150%, which is in the range of estimates for the optimal carbon tax. Forthcoming results would show how these effects would change over time with climate change dynamics, the change in the valuation of fossil reserves, and the growth dynamics of developing economies.

Related literature

This paper stands at the intersection of several subfields of macroeconomics, climate economics, and computational and mathematical economics.

First, I develop an Integrated Assessment model (IAM) with heterogeneous countries, and this naturally relates to the classical approach of IAM by Nordhaus. I use a neoclassical model with a climate system and damage of temperatures, as in the DICE model, Nordhaus (1993, 2017), recently revisited in Barrage and Nordhaus (2024). In this representative agent framework, as in the rest of the literature, the standard Pigouvian result holds: the optimal taxation of carbon equals the Social Cost of Carbon (SCC). As a result, it is enough to measure the marginal cost of climate change to know the full path of the carbon tax. Golosov, Hassler, Krusell and Tsyvinski (2014) develop a complete study of the optimal taxation of fossil fuels in a class of models inspired by DICE models and derive the first-best policy and a closed-form formula as a function of the climate and economic parameters for the optimal carbon tax and Social Cost of Carbon.

Second, I extend this class of model to handle country heterogeneity. I build on the literature that started with the RICE – the multi-regions version of the DICE model – with Nordhaus and Yang (1996); Nordhaus (2011). As studied in Hillebrand and Hillebrand (2019), the optimal carbon tax should be the Social Cost of Carbon, i.e. the sum of Local Damages, and the heterogeneity across regions determines the optimal transfer policies. When transfers are unrestricted, there is no need to adjust the carbon tax or the Social Cost of Carbon for inequality. More recently, frameworks with more realistic heterogeneity have been developed to study the impact of climate change and design optimal policies such as Krusell and Smith (2022), Hassler, Krusell, Olovsson and Reiter (2020), Kotlikoff, Kubler, Polbin and Scheidegger (2021*b*), Kotlikoff, Kubler, Polbin, Sachs and Scheidegger (2021) or Belfiori (2018). I show how to solve for the optimal carbon taxation, which features of the heterogeneity matter, in this class of model.

Related, the spatial-economic geography literature has made important advances in studying the heterogeneous impact of climate change. Cruz and Rossi-Hansberg (2021), Cruz and Rossi-Hansberg (2022*a*), Rudik et al. (2021) or Bilal and Rossi-Hansberg (2023*b*) are rich frameworks that incorporate migration, agglomeration and congestion externality, and meaningful spatial heterogeneity. The design of optimal policies in such models is still being explored in the literature, and the approach in this article is well-suited for this body of work.

Moreover, a blooming literature has been developed to study the redistributive effects of carbon taxation within countries and the heterogeneous impacts of climate change across households. For example, Belfiori, Carroll and Hur (2024) provides similar theoretical and quantitative results for the optimal policy in the First-Best and Second-Best without transfers. In similar heterogeneous agents frameworks, Le Grand, Oswald, Ragot and Aurélien (2023); Wöhrmüller (2024); van der Ploeg, Rezai and Tovar (2024); Fried, Novan and Peterman (2024); Benmir and Roman (2022); Kuhn and Schlattmann (2024); Schlattmann (2024); Douenne, Hummel and Pedroni (2023); Douenne, Dyrda, Hummel and Pedroni (2024) have made significant contributions to understand the redistributive effects – due to differences in goods, car and durables, housing, or adaptation mechanisms along the wealth distribution.

I also relate to the literature closer to climate sciences, reexamining the empirical performances of Integrated Assessment Models, such as in Dietz, van der Ploeg, Rezai and Venmans (2021), Dietz and Venmans (2019), Ricke and Caldeira (2014) or Folini et al. (2021). Following this literature, I consider a simple climate system that allows me to both match larger IAMs and derive closed-form expressions for the social cost of carbon and optimal carbon tax.

Moreover, I also relate to a thriving literature that studies optimal policy design in Heterogeneous Agents models. Solving Ramsey policies, Le Grand et al. (2021), Bhandari et al. (2021*a*), Davila and Schaab (2023) or McKay and Wolf (2022) propose different approaches to conduct monetary and fiscal policy in HANK models. In my framework, I solve the Ramsey policy sequentially and solve climate externalities and Pigouvian taxation in the presence of heterogeneity rather than managing business cycle fluctuations.

The method developed here is flexible enough to handle aggregate uncertainty, such as climate risk and business cycle fluctuation, and results in this dimension are work-in-progress. The Stochastic DICE model of Cai and Lontzek (2019) and Lontzek, Cai, Judd and Lenton (2015) or the general approach to study model uncertainty and ambiguity aversion applied to climate change in Barnett, Brock and Hansen (2020, 2022) are particularly related. If the inclusion of aggregate risk is preliminary in the present paper, I provide intuitions in the toy model and will integrate this in forthcoming works.

Lastly, this work also relates to advances in the mathematical literature on Mean Field Games. Indeed, if the literature has leveraged approaches studying the PDE system, following Lasry-Lions' contribution, Cardaliaguet (2013/2018), or Achdou et al. (2022), they usually rely on dynamic programming methods. However, the Pontryagin maximum principle – used for solving the neoclassical model – extends to the stochastic case, as in Yong and Zhou (1999), or the case with a distribution of agents or – Mean-Field / McKean Vlasov dynamics – as in Carmona et al. (2015), Carmona and Delarue (2018) or Carmona and Laurière (2022). Using this approach in the deterministic case in large dimensions, I solve the model globally, compute the social cost of carbon analytically, and design optimal policy. For the case with aggregate risk, I borrow intuitions from Carmona et al. (2016), Bourany (2019), and Carmona and Delarue (2018) to solve the Stochastic FBSDE system in future work Bourany (2023).

The remainder of this paper is organized as follows. In Section 2, I study the optimal taxation carbon in a simple model to provide most of the intuitions. In Section 3, I lay out the Integrated Assessment Model that we study in the policy analysis. In Section 4 I derive the optimal carbon tax – First-Best and Second-Best Ramsey policy – in this context. In Section 5, I present how I match the model to the data. In Section 6, I present the main result of the quantitative analysis.

2 Toy model

In this section, I develop the simplest climate economy model to highlight the intuition behind the design of the optimal climate policy. The goal is to provide intuitions on the effects of heterogeneity across countries, the source of climate externality related to energy markets, and how it change the level of carbon taxation. In the next section, I develop a more general quantitative model that will be used for policy recommendation.

The model is static and all the decisions are taken in one period. Consider I countries $i \in \mathbb{I}$, heterogeneous in three dimensions that will be detailed below. A unique household in each country consumes the good c_i, produced by the representative firm with labor ℓ_i and energy e_i.[1] In each of these countries, an energy producer extracts energy and sells this input at price q^e on international markets. It earns profits and is owned by the household. Moreover, the countries are subject to climate damage on production. I describe each agent's problem in turn. Finally, a government, whose objective is specified in the next section, imposes a tax on emission $\mathrm{t}_i^{\varepsilon}$ and distributes lump-sum transfers t_i^{ls} in each country.

First, a representative household consumes their labor income $w_i\bar{\ell}_i$, where the exogenous labor supply is normalized to $\bar{\ell}_i = 1$, the profit of the energy firm of its country π_i^e and the lump-sum transfers given by the government t_i^{ls}.

$$\begin{aligned} \mathcal{V}_i &= U(c_i) \\ c_i &= w_i\bar{\ell}_i + \pi_i^e + \mathrm{t}_i^{ls} \ . \end{aligned}$$

Second, a representative firm produces a homogeneous good[2] using energy e_i and household labor ℓ_i with a constant return to scale technology. Since the labor supply is normalized to 1, e_i represents the energy use per capita. The production function $\widetilde{F}(\ell_i, e_i)$ is concave in (ℓ_i, e_i), and $\widetilde{F}_e(\ell, e) > 0$ and $\widetilde{F}_{ee}(\ell, e) < 0$ and features Inada conditions. This firm maximizes profits:

$$\max_{\ell_i, e_i} \ \mathcal{D}_i(\mathcal{S}) z_i \widetilde{F}(\ell_i, e_i) - (q^e + \mathrm{t}_i^{\varepsilon}) e_i - w_i \ell_i \quad , \tag{1}$$

where $\mathrm{t}_i^{\varepsilon}$ is a carbon tax paid per unit of energy.

Both countries are subject to climate damages $\mathcal{D}_i(\mathcal{S})$ caused by climate externalities related to the world fossil-fuel energy consumption that release greenhouse gas emissions in the atmosphere:

$$\mathcal{S} = \mathcal{S}_0 + \overbrace{\textstyle\sum_{i\in\mathbb{I}} e_i}^{\text{GHG emissions } \mathcal{E}} \quad ,$$

where energy use and emissions are measured in (metric) tons of Carbon or CO_2. This depends on the mix between fossil fuels and renewable energies, taken as given in this static model. The quantitative model introduces this endogenous channel of energy substitution.

[1] Generalization of this model, with differing population $\mathcal{P}_i$, endowments of inputs in the production function (e.g. capital k_i), do not change the qualitative implication of this framework, as we will see in the quantitative model.

[2] This good is traded costlessly across countries and its price is the numeraire, and hence normalized to 1.

The global carbon emission stock is not internalized by households in their energy consumption decision, leading to damage $\mathcal{D}_i(\mathcal{S})$ that affects country i's effective productivity, as in standard Integrated Assessment models, e.g. Nordhaus DICE models.

In each country, a competitive energy producer extracts energy e_i^x – for example oil, gas, or coal – maximizing its profit, subject to convex cost $\mathcal{C}(e^x)$, i.e. $\mathcal{C}'(e^x) > 0$ and $\mathcal{C}''(e^x) > 0$ that is paid in the homogenous good.

$$\pi_i^e = \max_{e_i^x} q^e e_i^x - \mathcal{C}_i(e_i^x) \quad ,$$

$$\Rightarrow \qquad q^e = \mathcal{C}_i'(e_i^x) \qquad \Rightarrow \qquad \begin{cases} e_i^x & = \mathcal{E}_i(q_i^e) = \mathcal{C}_i'^{-1}(q^e) \\ \pi_i^e & := q^e \mathcal{E}(q_i^e) - \mathcal{C}_i(\mathcal{E}_i(q_i^e)) \end{cases} ,$$

subject the energy price q^e. This corresponds to a decreasing-return-to-scale extraction technology, and implies positive profits $\pi_i^e > 0$. Since energy is traded without friction on international markets, this price is set to clear the supply and demand:

$$E = \sum_{i \in \mathbb{I}} e_i = \sum_{i \in \mathbb{I}} e_i^x \ .$$

Since the good firm's technology is constant return to scale (CRS), define $F(e_i) = \widetilde{F}(1, e_i)$ to aggregate firms and household budgets into a single constraint:

$$c_i + (q^e + \mathrm{t}_i^{\varepsilon})e_i = \mathcal{D}_i(\mathcal{S}) z_i F(e_i) + q_i^e e_i^x - \mathcal{C}_i(e_i^x) + \mathrm{t}_i^{ls} \qquad [\lambda_i] \ . \tag{2}$$

with λ_i the shadow value of that constraint, which plays an important role in redistribution motives.

Heterogeneity. The countries $i \in \mathbb{I}$ are symmetric in all regards, except for differences in three parameters. To fix ideas, consider two regions, *North* and *South*, to give qualitative predictions of the policy results. First, countries differ in terms of productivity z_i. Here, I consider a wide definition of z_i that accounts for technology, efficiency, market frictions, and institutions. This results in some countries – e.g. the North – producing more and being richer, leading to inequality in consumption.[3] Second, energy reserves endowments are unequally distributed, which results in differences in costs of extraction. I assume that northern countries, e.g. US, Canada, Russia, Norway, etc. have lower costs of extraction $\mathcal{C}_N'(e) < \mathcal{C}_S'(e)$, implying larger production and energy rents $\pi_N^e > \pi_S^e$. Third, some countries, e.g., Southern Hemisphere, are more vulnerable to the damages of climate, $\mathcal{D}_S(\mathcal{S}) < \mathcal{D}_N(\mathcal{S})$ for all $\mathcal{S}$ the stock of carbon. In this sense, the damage parameter $\gamma_i = -\frac{\mathcal{D}_i'(\mathcal{S})}{\mathcal{S}\mathcal{D}_i(\mathcal{S})}$ is higher in the South, $\gamma_S > \gamma_N$. All these differences yield heterogeneity in consumption in the competitive equilibrium and motives for redistribution, e.g. $c_N > c_S$.

[3] Indeed, assuming $F(e)$ is Cobb Douglas $F(e) = \bar{\ell}^{1-\alpha} e^{\alpha}$, with $\bar{\ell} = 1$, we obtain $\alpha \mathcal{D}_i(\mathcal{S}) z_i e_i^{\alpha-1} = q^e$ leading to

$$e_i = \big(\alpha \mathcal{D}_i(\mathcal{S}) z_i / q^e\big)^{1/(1-\alpha)} \qquad y_i - q^e e_i = \big(\mathcal{D}_i(\mathcal{S}) z_i\big)^{1/(1-\alpha)} (q^e)^{-\alpha/(1-\alpha)} \big[\alpha^{\alpha/(1-\alpha)} - \alpha^{1/(1-\alpha)}\big]$$

which is increasing in z_i and $\mathcal{D}_i(\mathcal{S})$

Definition 2.1 (**Competitive Equilibrium**). *Given a carbon policy* t^{ε}*, a competitive equilibrium (CE) is an allocation* $\{c_i, e_i, e_i^x\}_i$ *and energy price* q^e *such that (i) the good firm chooses input* e_i *maximizing profit, and (ii) the energy firm chooses extraction* e_i^x *maximizing profit, and both goods and energy markets clear:*

$$\sum_{i\in\mathbb{I}} c_i + \mathcal{C}_i(e_i^x) = \sum_{i\in\mathbb{I}} \mathcal{D}_i(\mathcal{S}) z_i F(e_i) \qquad\qquad E = \sum_{i\in\mathbb{I}} e_i = \sum_{i\in\mathbb{I}} e_i^x \quad .$$

This results in the following optimality conditions. First, for consumption, the multiplier λ_i represents the marginal value of wealth or the marginal utility of consumption.

$$\lambda_i = U'(c_i) \qquad \textit{with} \qquad c_i = \mathcal{D}_i(\mathcal{S}) z_i F(e_i) + q^e(e_i^x - e_i) - \mathcal{C}_i(e_i^x) + \mathrm{t}_i^{ls}$$

where consumption depends on production, energy cost, and net energy export.

The second and third optimality for energy use and energy extraction write as follow:

$$\begin{aligned} MPe_i &= q^e + \mathrm{t}_i^{\varepsilon} \qquad \textit{with} \qquad MPe_i := \mathcal{D}_i(\mathcal{S}) z_i F'(e_i) \ , \\ q^e &= \mathcal{C}'(e_i^x) \ , \end{aligned}$$

and this corresponds to the standard condition Marginal Product = Marginal Cost for Energy.

This competitive equilibrium is inefficient: climate damages $\mathcal{D}_i(\mathcal{S})$ are not internalized, and energy consumption is too high in view of the economic costs of global warming. Moreover, economic inequality results from the heterogeneity in productivity, energy endowment, and climate damage. In our two regions example, $c_N > c_S$ results in $\lambda_S > \lambda_N$. Redistribution from the North to the South could be desirable from a utilitarian point of view. This inequality in consumption and damages arises despite trade openness.[4] I explore how the social planner allocates consumption and energy in such an environment.

2.1 First-Best: Social planner allocation with full transfers

Consider a Social Planner choosing the agent's decisions, subject to the resource constraints in goods and energy as well as the climate externality.

$$\begin{aligned} \mathcal{W} = \max_{\{c_i, e_i, e_i^x\}_{i\in\mathbb{I}}} & \sum_{i\in\mathbb{I}} \omega_i U(c_i) \\ \sum_{i\in\mathbb{I}} c_i + \mathcal{C}_i(e_i^x) &= \sum_{i\in\mathbb{I}} \mathcal{D}_i(\mathcal{S}) z_i F(e_i) \qquad [\phi] \\ E = \sum_{i\in\mathbb{I}} e_i &= \sum_{i\in\mathbb{I}} e_i^x \qquad [\mu^e] \\ \mathcal{S} &:= \mathcal{S}_0 + \sum_{i\in\mathbb{I}} e_i \end{aligned} \tag{3}$$

[4] We could consider trade and financial autarky preventing production from being exported to other countries. This would strengthen heterogeneity and redistributive motives.

where ϕ is the shadow value of the good market clearing and μ^e the one of the energy market clearing. The welfare function is the weighted sum of countries' utilities, with Pareto weights ω_i. In the following, I denote the social planner allocation $\{\hat{c}_i, \hat{e}_i\}_{i\in\mathbb{I}}$ to distinguish it from the competitive equilibrium.

Choosing the consumption on behalf of the agents yields a redistribution motive:

$$[c_i] \qquad \phi = \omega_i U'(\hat{c}_i) \qquad \Rightarrow \qquad \omega_i U'(\hat{c}_i) = \omega_j U'(\hat{c}_j) \qquad \forall i,j \in \mathbb{I} \ .$$

Depending on the Pareto weights, there is a motive for transferring consumption across countries.

Regarding the choice of energy inputs:

$$[\hat{e}_i] \ \& \ [\hat{e}_i^x] \qquad c'(\hat{e}_i^x) = \frac{\mu^e}{\phi} = \mathcal{D}_i(\mathcal{S}) z_i F'(\hat{e}_i) + \underbrace{\sum_{j\in\mathbb{I}} \mathcal{D}_j'(\mathcal{S}) z_j F(\hat{e}_j)}_{=:\overline{SCC}}$$

with an additional term that represents the cost of emitting one ton of carbon in terms of forgone production. This term is the *Social Cost of Carbon (SCC)* in the social planner allocation and represents the marginal global damage of climate change. With $\gamma_i = -\frac{\mathcal{D}_i'(\mathcal{S})}{\mathcal{S}\mathcal{D}_i(\mathcal{S})}$, which is constant with Nordhaus' Damage function $\mathcal{D}_i(\mathcal{S}) = e^{-\frac{\gamma_i}{2}(\mathcal{S}-\mathcal{S}_0)^2}$, I redefine it as $\overline{SCC} = \sum_{j\in\mathbb{I}} \mathcal{S}\gamma_j y_j$.

I turn to the decentralization of such allocation. I consider a planner who has access to all instruments $\{\mathrm{t}_i^e, \mathrm{t}_i^{ls}\}_i$, and, in particular, lump-sum transfers t_i^{ls} across countries.

Proposition 1 (First-Best Policy and Decentralization).
The optimal policy decentralizing the First-Best allocation is a uniform carbon tax, $\mathrm{t}_i^\varepsilon = \mathrm{t}^\varepsilon$, equal to the Social Cost of Carbon $\mathrm{t}^\varepsilon = SCC$, and uses lump-sum transfers for redistribution. Indeed, optimality writes:

$$MPe_i := \mathcal{D}_i(\mathcal{S}) z_i F'(\hat{e}_i) = q^e + \mathrm{t}^\varepsilon \qquad q^e = c'(\hat{e}_i^x)$$

$$\textit{with} \qquad \mathrm{t}^\varepsilon = \overline{SCC} := -\sum_{j\in\mathbb{I}} \mathcal{D}_j'(\mathcal{S}) z_j F(\hat{e}_j) = \sum_{j\in\mathbb{I}} \mathcal{S}\gamma_j y_j = I\mathbb{E}_j\left[\mathcal{S}\gamma_j y_j\right]$$

and transfers $\{\mathrm{t}_i^{ls}\}_i$ which are implicitly defined by

$$\omega_i U'(\hat{c}_i) = \omega_j U'(\hat{c}_j) \ ,$$
$$\hat{c}_i = \mathcal{D}_i(\mathcal{S}) z_i F(\hat{e}_i) + q^e(\hat{e}_i^x - \hat{e}_i) - c_i(\hat{e}_i^x) + (\mathrm{t}_i^{ls} - \mathrm{t}^\varepsilon \hat{e}_i) \ .$$

For arbitrary Pareto weights $\boldsymbol{\omega}$, lump-sum transfers are redistributive: $\exists\, i,j$ s.t. $\mathrm{t}_i^{ls} > \mathrm{t}^\varepsilon \hat{e}_i, \mathrm{t}_j^{ls} < \mathrm{t}^\varepsilon \hat{e}_j$.

To see this last point, summing the budget constraints yields:

$$\sum_{i\in\mathbb{I}} \mathrm{t}_i^{ls} = \mathrm{t}^\varepsilon \sum_{i\in\mathbb{I}} \hat{e}_i$$

implying there is lump-sum redistribution from richer countries to poorer ones in the presence of heterogeneity across countries.[5] In our North-South example, with $z_S < z_N$, $c_N' < c_S'$ or $\mathcal{D}_S' > \mathcal{D}_N'$,

[5] Given that $\mathrm{t}_i^{ls} = \mathrm{t}^\varepsilon \hat{e}_i + U'^{-1}(\frac{\phi}{\omega_i}) - \mathcal{D}_i(\mathcal{S}) z_i F(\hat{e}_i) - q^e(\hat{e}_i^x - \hat{e}_i)$.

and for arbitrary Pareto weights[6], we obtain that $\mathrm{t}^{ls}_S > \mathrm{t}^{\varepsilon}\hat{e}_S$, and $\mathrm{t}^{ls}_N < \mathrm{t}^{\varepsilon}\hat{e}_N$. This implies that some funds are taxed from the North and redistributed lump-sum to the South. However, there exists a unique set of Pareto weights $\omega_i = 1/U'(\hat{c}_i)$ – the so-called Negishi weights – such that this motive disappears $\mathrm{t}^{ls}_S = \mathrm{t}^{\varepsilon}\hat{e}_S$ and $\mathrm{t}^{ls}_N = \mathrm{t}^{\varepsilon}\hat{e}_N$ and there are no transfers across countries.

In the following, I rule out this flexible lump-sum transfers assumption: if development aid exists, in practice, full redistribution with lump-sum taxes and transfers to cover the differences in technology, market frictions, and institutions is politically unfeasible.

2.2 Second Best: Ramsey problem, uniform carbon tax, and limited transfers

Consider now a social planner designing the optimal climate policy, taking into account the constraints preventing transfers across countries. Subject to the competitive equilibrium optimality conditions, the climate externality and the absence of financial instruments for full lump-sum redistribution, the planner takes the decisions of consumption and energy to maximize global welfare. I denote the Ramsey allocation $\{\tilde{c}_i, \tilde{e}_i, \tilde{e}^x_i\}_i$ to distinguish it from the competitive equilibrium $\{c_i, e_i, e^x_i\}$ and the First-best allocation $\{\hat{c}_i, \hat{e}_i, \hat{e}^x_i\}$.

$$\mathcal{W} = \max_{\{\tilde{c}_i, \tilde{e}_i, \tilde{e}^x_i\}_i, q^e, \mathrm{t}^{\varepsilon}} \sum_{i \in \mathbb{I}} \omega_i U(\tilde{c}_i) \tag{4}$$

I consider a uniform carbon tax for all countries $\mathrm{t}^{\varepsilon}_i = \mathrm{t}^{\varepsilon}$, and this, for several reasons. First, the goal is to provide a direct comparison to the standard Pigouvian framework, where the natural outcome is a global uniform tax on fossil energy. Second, following the arguments of Weitzman (2015), the uniform carbon tax or price of carbon serves as a "focal point", where the social-planner policy is a representation of the bargaining outcome in an agreement coming from all the countries in the world. Third, in the next section, I consider different tax rates for each country.

In both cases – uniform tax or country-specific taxes – I assume away cross-country transfers: as the revenue of the tax is redistributed lump-sum to the household $\tilde{\mathrm{t}}^{ls}_i = \mathrm{t}^{\varepsilon}\tilde{e}_i$. Moreover, in the Second-Best Ramsey policy, the planner internalizes the optimality conditions of the competitive equilibrium. Using the Primal Approach in public finance, the Ramsey problem accounts for the countries' budgets and the firms' optimality conditions and is written as:

$$\begin{aligned}
&\mathcal{W} = \max_{\{\tilde{c}_i, \tilde{e}_i, \tilde{e}^x_i\}_i, q^e, \mathrm{t}^{\varepsilon}} \sum_{i \in \mathbb{I}} \omega_i U(c_i) \\
&s.t \qquad \tilde{c}_i + (q^e + \mathrm{t}^{\varepsilon})\tilde{e}_i = \mathcal{D}_i(\mathcal{S}) z_i F(\tilde{e}_i) + q^e \tilde{e}^x_i - \mathcal{C}_i(\tilde{e}^x_i) + \mathrm{t}^{\varepsilon}\tilde{e}_i && [\phi_i] \\
&\qquad q^e = \mathcal{C}'_i(\tilde{e}^x_i) \qquad\qquad q^e + \mathrm{t}^{\varepsilon} = MPe_i && [\upsilon_i] \\
&\qquad \mathcal{S} := \mathcal{S}_0 + \sum_i e_i \qquad\qquad E = \sum_i e_i = \sum_i e^x_i && [\mu^e]
\end{aligned} \tag{5}$$

with the Lagrange Multipliers $\phi_i, \upsilon_i, \mu^e$, respectively for the budget constraint, the energy choice of the good firm, and the energy market clearing.

[6] In particular, this is the case if the Pareto weights are large enough, i.e. $\omega_S \geq \phi/U'(c_S)$ i.e. more than the weight imposed by the shadow value of good discounted by South' marginal utility

The Lagrangian of this problem, as well as the detailed optimality conditions, are derived in detail in Appendix A. They yields the following optimality conditions for consumption c_i, demand e_i and supply e_i^x for energy:

$$\omega_i U'(\tilde{c}_i) = \phi_i = \text{marginal value of income/wealth}$$

$$\phi_i \mathrm{t}^\varepsilon = \underbrace{\upsilon_i \, \mathcal{D}_i(\mathcal{S}) z_i F''(\tilde{e}_i)}_{=\text{demand distortion}} + \underbrace{\mu^e}_{=\text{supply redistribution}} \underbrace{- \textstyle\sum_j \phi_j \mathcal{D}_j'(\mathcal{S}) z_j F(\tilde{e}_j)}_{\propto \text{ Social Cost of Carbon}}$$

The planner chooses a single carbon policy instrument and thus accounts for several redistribution channels across countries through (i) the marginal value of income ϕ_i, (ii) the distortion of energy demand symbolized by the shadow value of energy choice υ_i, (iii) the redistributive effects on energy market with the market clearing multiplier μ^e, and (iv) the Social Cost of Carbon summarizing the marginal cost of climate change. There, before deriving the main formula for the optimal carbon tax t^ε and explaining the economic intuition behind it, let us introduce these key objects.

First, I define the "social welfare weight" $\widehat{\phi}_i = \phi_i / \overline{\phi}$ that represents the relative weight that the planner uses for global policy. I define it as a ratio, rescaling the multiplier ϕ_i, of the shadow value of relaxing the budget constraint for country i.

$$\widehat{\phi}_i = \frac{\phi_i}{\overline{\phi}} = \frac{\omega_i U'(c_i)}{\frac{1}{I} \sum_{i \in \mathbb{I}} \omega_i U'(c_i)} \lessgtr 1$$

where $\overline{\phi} = \partial \mathcal{W} / \partial c = \frac{1}{I} \big(\sum_{i \in \mathbb{I}} \omega_i U'(c_i) \big)$ is the average marginal utility. $\overline{\phi}$ is the "money↔welfare" conversion factor for the social planner. When there is no full redistribution, this factor $\widehat{\phi}_i$ is high for relatively poorer countries or countries with a high Pareto weight ω_i.

Second, climate change affects countries differently according to their marginal damages $\mathcal{D}_j'$. The social cost of carbon scales those damages by the social welfare weights/inequality factor $\widehat{\phi}_i \propto \omega_i U'(c_i)$ given that the planner does not have access to full redistribution. Rescaled in monetary unit, with the conversion factor $\overline{\phi}_i$, the SCC writes:

$$SCC := -\frac{\partial \mathcal{W} / \partial \mathcal{S}}{\partial \mathcal{W} / \partial c} = -\frac{1}{\overline{\phi}} \sum_j \phi_j \mathcal{D}_j'(\mathcal{S}) z_j F(e_j) = -\sum_j \widehat{\phi}_j \mathcal{D}_j'(\mathcal{S}) z_j F(e_j) = \sum_{j \in \mathbb{I}} \widehat{\phi}_j \, \mathcal{S} \gamma_j y_j$$

with the definition $\gamma_i = -\frac{\mathcal{D}_j'(\mathcal{S})}{\mathcal{D}_j(\mathcal{S}) \mathcal{S}}$, the slope of the climate damage function, as in Nordhaus' DICE model. In particular, in this heterogeneous countries model with limited redistribution, the Social Cost of Carbon integrates the distribution of consumption/income under the factor $\widehat{\phi}_i$:

$$SCC := \sum_{j \in \mathbb{I}} \widehat{\phi}_j \, \mathcal{S} \gamma_j y_j = I \mathbb{E}_j \left(\widehat{\phi}_j \, \mathcal{S} \gamma_j y_j \right)$$

$$SCC = I \, \mathbb{E}_j [\mathcal{S} \gamma_j y_j] + I \, \mathbb{C}\mathrm{ov}_j \left(\widehat{\phi}_j, \mathcal{S} \gamma_j y_j \right) \quad \lessgtr \quad I \, \mathbb{E}_j [\mathcal{S} \gamma_j y_j] =: \overline{SCC}$$

with the damage slope $\gamma_i = -\frac{\mathcal{D}_i'(\mathcal{S})}{\mathcal{D}_i(\mathcal{S}) \mathcal{S}}$, and country i's output $y_i = \mathcal{D}_i(\mathcal{S}) z_i F(e_i)$. The $\overline{SCC} =$

$I\mathbb{E}_j[-\mathcal{D}'_j(\mathcal{S})z_jF(e_j)] = I\mathbb{E}_j[\mathcal{S}\gamma_j y_j]$ is the Social Cost of Carbon when full redistribution is available, or equivalently in representative agent models where redistributive concerns are absent. Since $\mathbb{E}_j(\cdot)$ is a mean[7] over countries j, we need to multiply by the number of countries (I here) to obtain the sum of local damages.

Is the SCC higher in the model with inequality compared to the one-agent setting? Take the North-South economy as an example. First, low-income countries have a lower consumption and hence higher marginal utility of consumption, e.g. $c_S < c_N$ and $\widehat{\phi}_S > \widehat{\phi}_N$. Second, South is suffering from stronger damages $\mathcal{D}'_S > \mathcal{D}'_N$. However, third, productivity is higher in the North $z_N > z_S$ implying $F(e_N) > F(e_S)$ and income $\bar{y}_N{:=}z_N F(e_N) > \bar{y}_S$. Therefore, the covariance between $\widehat{\phi}_i$, $\bar{y}_i$ and $\gamma_i \propto \mathcal{D}'_i$ is ambiguous. Quantitatively, in a large class of Integrated Assessment models, the local cost of climate change $\mathcal{D}'_i(\mathcal{S})y_i$ is strongly correlated with income y_i, as there larger production loss of climate change in richer countries. In such cases, the covariance $\mathbb{C}\mathrm{ov}_j\Big(\hat{\phi}_j, \gamma_j y_i\Big)$ is negative, and as a result:

$$SCC = I\,\mathbb{E}_j[\mathcal{S}\gamma_i y_i] + I\,\mathbb{C}\mathrm{ov}_j\Big(\widehat{\phi}_j, \mathcal{S}\gamma_j y_j\Big) < I\mathbb{E}_j[\gamma_j y_j] = \overline{SCC} \qquad \text{if} \quad \mathbb{C}\mathrm{ov}_j\Big(\widehat{\phi}_j, \gamma_j y_j\Big) < 0$$

Third, I explore the redistributive effect of carbon taxation on the energy supply. Changing the price affects the market clearing, with shadow value μ^e. I formulate this supply side channel as a redistribution between energy importers and exporters, weighted by a factor representing the curvature of aggregate energy supply:

$$\begin{aligned}
\textit{Supply Redistribution} \;&= \frac{\mu^e}{\bar{\phi}} = \mathcal{C}_{EE}\,\frac{1}{I}\sum_j \frac{\phi_j}{\bar{\phi}}(e_j - e^x_j) \qquad \text{with} \qquad \mathcal{C}_{EE} = \big(\sum_j c''_j(e^x_j)^{-1}\big)^{-1} \\
&= \mathcal{C}_{EE}\;\mathbb{E}_j\Big(\widehat{\phi}_j(e_j - e^x_j)\Big) \\
&= \mathcal{C}_{EE}\;\mathbb{C}\mathrm{ov}_j\Big(\widehat{\phi}_j,\, e_j - e^x_j\Big) \lesseqgtr 0
\end{aligned}$$

What is the sign of this covariance? In our two regions example, we assumed that the North had a larger endowment in energy resources and hence higher net energy exports $e_N - e^x_N < e_S - e^x_S$. Therefore, since the net import of energy correlates with lower consumption, and hence a higher marginal value of consumption $U'(c_i)$, the covariance term is positive. Moreover, the magnitude of this terms-of-trade redistribution ultimately depends on the aggregate supply elasticity:

$$\mathcal{C}_{EE} = \big(\sum_j c''_j(e^x_j)^{-1}\big)^{-1} = q^e\,\frac{\bar{\nu}}{E} \qquad \text{with} \qquad \bar{\nu} = \big(\textstyle\sum_j \lambda^x_j \nu_j^{-1}\big)^{-1}$$

with ν_j the inverse supply elasticity, constant in the iso-elastic case $q^e = c'_i(e) = \bar{\nu}_i e^{\nu_i}$ and the share of country i in energy production $\lambda^x_j = e^x_i/E$. As a result, this Social "Supply Redistribution" is positive. It is larger when the energy supply is inelastic – price and profits vary a lot for small changes in quantity produced. It is null when the energy production is Constant Return to Scale (CRS) when $\nu_j = 0$, and therefore, no energy rents are redistributed.

[7]The formula of the expectation of a product writes $\mathbb{E}_i[x_i y_i] = \mathbb{E}_i[x_i]\mathbb{E}_i[y_i] + \mathbb{C}\mathrm{ov}_i[x_i y_i]$

Third, carbon taxation distorts energy choice across users and changes the equilibrium energy price along the demand curve. We derive the Social "Demand Distortion" term as:

$$\textit{Demand Distortion} = \frac{1}{I}\sum_j \frac{\upsilon_j}{\bar{\phi}} \mathcal{D}_j(\mathcal{S}) z_j F''(e_j) = \mathbb{E}_j\left(\widehat{\upsilon}_j\, \mathcal{D}_j(\mathcal{S}) z_j F''(e_j)\right)$$
$$= \mathbb{C}\mathrm{ov}_j\Big(\frac{\upsilon_j}{\bar{\phi}}, \mathcal{D}_j(\mathcal{S}) z_j F''(e_j)\Big) \lesseqgtr 0$$

with υ_j is the multiplier on the energy demand optimality condition: positive value implies that the planner would like to relax the constraint, increase the quantity e_i, lower the MPe_i, and conversely for negative values. I define $\widehat{\upsilon}_j = \upsilon_j/\bar{\phi}$ as the rescaled shadow value of country j's energy demand.

There is no aggregate distortion, only redistributive distortions across countries. This comes from the optimality of the tax t^{ε}, as $\mathbb{E}_j(\widehat{\upsilon}_j) = 0$, and this yields the last line as a covariance. How to determine its sign? In our North-South example, lower-income economies have energy demand more sensitive to price distortions since $z_i F''(e_i)$ relates to the energy share and demand elasticity:

$$\mathcal{D}_i(\mathcal{S}) z_i F''(e_i) = -\frac{q^e}{e_i \sigma_i^e}(1 - s_i^e) \qquad \Rightarrow \qquad \mathcal{D}_S(\mathcal{S}) z_S F''(e_S) < \mathcal{D}_N(\mathcal{S}) z_N F''(e_N)$$

where $s_i^e = \frac{e_i q^e}{y_i} < 1$ is the energy share in production and σ_i^e is country i's energy demand elasticity. The North relies "more" on energy – since $z_N > z_S$ implies that $e_N > e_S$: more productive countries have higher energy demand ceteris paribus. It would also be the case if energy is more substitutable in richer countries, i.e. large σ^e, and demand varies a lot with price. The covariance would then be negative: if the planner values more the production, high υ_j, of the most inelastic countries, lower $F''(e_i)$ then the tax would be lower. Moreover, that term is null if the energy demand/production function is constant return in energy such that $s_i^e = 1$, or if energy is perfectly substitutable $\sigma^e \to \infty$, or if we are in a representative agent economy $\mathcal{D}_N(\mathcal{S}) z_N F''(e_N) = \mathcal{D}_S(\mathcal{S}) z_S F''(e_S)$ and there is no heterogeneity in demand across countries.

Proposition 2 (**Second-Best Ramsey Policy with limited transfers**).
The optimal Second-Best carbon tax accounts for three distributional motives when setting a single uniform level*: (i) climate damage in the* ***Social Cost of Carbon (SCC)****, (ii)* ***Supply Redistribution*** *in energy markets through terms-of-trade and energy rents and (iii)* ***Demand Distortion*** *through distorted firms' energy choices. This includes redistribution motives due to the presence of inequality through the* ***social welfare weights*** $\widehat{\phi}_j = \phi_j/\bar{\phi} = \omega_j U'(c_j)/\frac{1}{I}\sum_i \omega_i U'(c_i)$*. The optimal energy tax writes:*

$$\mathrm{t}^{\varepsilon} = SCC \; + \; \textit{Supply Redistribution} \; + \; \textit{Demand Distortion}$$

$$\mathrm{t}^{\varepsilon} = -\sum_j \widehat{\phi}_j \mathcal{D}'_j(\mathcal{S}) z_j F(e_j) \; + \mathcal{C}_{EE}\, \frac{1}{I}\sum_j \widehat{\phi}_j (e_j - e_j^x) \; + \; \frac{1}{I}\sum_j \widehat{\upsilon}_j \mathcal{D}_j(\mathcal{S}) z_j F''(e_j)$$

$$\mathrm{t}^{\varepsilon} = I\mathbb{E}_j\Big(\mathcal{S}\gamma_j y_j\Big) + I\mathbb{C}\mathrm{ov}_j\Big(\widehat{\phi}_j, \mathcal{S}\gamma_j y_j\Big) + q^e \frac{\bar{\nu}}{E}\mathbb{C}\mathrm{ov}_j\Big(\widehat{\phi}_j, e_j - e_j^x\Big) - q^e\, \mathbb{C}\mathrm{ov}_j\Big(\widehat{\upsilon}_j, \tfrac{1 - s_i^e}{\sigma_i^e e_i}\Big) \; ,$$

for the carbon tax such that $MPe_i = c'(e^x) + \mathrm{t}^\varepsilon$, where $\gamma_i = -\frac{\mathcal{D}_i'(\mathcal{S})}{\mathcal{D}_i(\mathcal{S})\mathcal{S}}$ is the marginal damage of climate change[8]*, $y_i = \mathcal{D}_i(\mathcal{S})z_iF(e_i)$ is total production, $\bar{\nu} = \left(\sum_j \lambda_j^x \nu_j^{-1}\right)^{-1}$ the average inverse energy supply elasticity, s_i^e the energy cost shares, and σ_i^e the energy demand elasticity. We see these three motives matter with a single tax and lump-sum rebate. Compared to the economy with full redistribution, the carbon tax is **smaller** if (i) the cost of climate $\gamma_j y_j$ is concentrated in richer countries, with low $\widehat{\phi}_i$, (ii) the net energy imports are high, with higher $e_i - e_i^x$, in richer, low $\widehat{\phi}_i$ countries, (iii) the energy is more essential, with low demand elasticity and high $\frac{1-s_e^i}{\sigma_i^e e_i}$, in poorer, high distortion $\widehat{\upsilon}_i$, countries. For (ii) and (iii), carbon taxation is isomorphic to an energy terms-of-trade manipulation between the exporters and the importers in trade theory.*

In the next corollary, we reexpress the carbon tax as a function of observable sufficient statistics. Indeed, given its dependence on the multipliers for individual demand υ_i, the Demand Distortion term can be quite opaque.

Corollary 3 (**Second-Best Ramsey Policy with limited transfers, sufficient statistics**). *The optimal Second-Best uniform carbon tax can be rewritten as:*

$$\mathrm{t}^\varepsilon = \frac{1}{1 + \mathbb{C}\mathrm{ov}_j\left(\widehat{\phi}_j, \widehat{e}_j^{s,\sigma}\right)}\Big[SCC + \mathit{Supply\ Redistribution}\Big] \qquad \text{with} \qquad \widehat{e}_j^{s,\sigma} := \frac{\frac{\sigma_j^e e_j}{1-s_j^e}}{\sum_i \frac{\sigma_i^e e_i}{1-s_i^e}}$$

$$\mathrm{t}^\varepsilon = \frac{1}{1 + \mathbb{C}\mathrm{ov}_j\left(\widehat{\phi}_j, \widehat{e}_j^{s,\sigma}\right)}\left[I\mathbb{E}_j\left(\mathcal{S}\gamma_j y_j\right) + I\mathbb{C}\mathrm{ov}_j\left(\widehat{\phi}_j, \mathcal{S}\gamma_j y_j\right) + q^e\frac{\bar{\nu}}{E}\mathbb{C}\mathrm{ov}_j\left(\widehat{\phi}_j, e_j - e_j^x\right)\right]$$

for a uniform carbon tax t^ε, with marginal climate damage γ_i, output y_i, average inverse energy supply elasticity $\bar{\nu}$, energy cost shares s_i^e, energy demand elasticity σ_i^e. Note that $\widehat{e}_i^{s,\sigma} = \widehat{e}_i = e_i/E$ if the production function has a Cobb-Douglas form.[9] *Demand distortion amplifies or dampens the carbon taxation motives, i.e. the Social Cost of Carbon (SCC) and Supply Redistribution. The carbon tax is **lower** if the largest energy consumers e_i, with high energy share s_i^e and demand elasticity σ_i^e have low consumption c_i and thus high social welfare weights $\widehat{\phi}_j = \omega_j U'(c_j)/\frac{1}{I}\sum_i \omega_i U'(c_i)$.*

Note, in representative agent models, as in Nordhaus (2017), or Golosov et al. (2014), there is no heterogeneity across countries, making all the covariances trivially null. Hence, we obtain the standard Pigouvian result $\mathrm{t}^\varepsilon = SCC = I\mathbb{E}_j[\mathcal{S}\gamma_j y_j] = \overline{SCC}$. Moreover, models with unconstrained transfers, which can be aggregated, yield $\widehat{\phi}_j = \widehat{\phi}_i = 1$, also reducing these covariances to zero. Models with heterogeneity in income, climate damage, or country size, like Nordhaus and Yang (1996) or Krusell and Smith (2022), but no heterogeneity in energy demand nor energy rent redistribution also yield the Pigouvian result $\mathrm{t}^\varepsilon = SCC$, where the Social Cost of Carbon is adjusted for inequality.

These Second-Best carbon tax formulas hold for a single uniform carbon tax. If the planner has access to a distribution of carbon tax rates (or carbon prices), the *distribution* of the tax changes with the presence of inequality, as we see in the next section.

[8] The parameter γ_i is constant in the damage function used in the DICE model $\mathcal{D}_i(\mathcal{S}) = e^{-\gamma_i(\mathcal{S}-\mathcal{S}_0)^2}$.

[9] If all countries i have the same Cobb-Douglas production of the form $F(\ell_i, e_i) = e_i^\alpha \ell_i^{1-\alpha}$, we get $e_j^{s,\sigma} := \frac{\sigma_i^e e_i}{1-s_i^e} = \frac{1 \times e_i}{1-\alpha}$ and hence $\widehat{e}_j^{s,\sigma} := e_j^{s,\sigma}/\sum_i e_i^{s,\sigma} = e_i/E$

2.3 Ramsey Problem with heterogeneous carbon tax & limited transfers

We consider a case where the Social Planner implements a distribution of *country-specific* carbon taxes t_i^ε. I again assume away cross-country transfers, and the revenue of the carbon tax is rebated lump-sum $\mathrm{t}_i^{ls} = \mathrm{t}_i^\varepsilon e_i$. The welfare objective and the constraints internalized by the planner are the same as before: (i) the budget $\tilde{c}_i = \mathcal{D}_i(\mathcal{S}) z_i F(\tilde{e}_i) + q^e(\tilde{e}_i^x - \tilde{e}_i) - \mathcal{C}_i(\tilde{e}_i^x)$ with multiplier ϕ_i, (ii) the energy firm optimality $q^e = \mathcal{C}'(e_i^x)$, (iii) the energy demand optimality $q^e + \mathrm{t}_i^\varepsilon = MPe_i$ with multiplier υ_j, and (iv) the energy market clearing $E = \sum_i e_i = \sum_i e_i^x$. The only exception is that the carbon tax is country-specific t_i^ε. The planner optimality conditions become:

$$\omega_i U'(c_i) = \phi_i \qquad\qquad \widehat{\phi}_i = \frac{\phi_i}{\bar{\phi}} = \frac{\omega_i U'(c_i)}{\frac{1}{I}\sum_{i\in\mathbb{I}} \omega_i U'(c_i)} \lessgtr 1$$

the same as before, as the planner keeps the same motive for redistribution. The social welfare weights, or inequality factor, come for heterogeneity in the marginal value of income ϕ_i. However, when the planner can choose one instrument per country, the distortion of demand is absent:

$$\upsilon_j = 0 \qquad \Rightarrow \qquad \textit{Demand Distortion} = 0$$

Proposition 4 (Second-Best Ramsey Policy, heterogeneous taxes & limited transfers). *The optimal Second-Best energy taxation policy with heterogeneous taxes when transfers are absent becomes:*

$$\mathrm{t}_i^\varepsilon = \frac{1}{\widehat{\phi}_i}\underbrace{\sum_j \widehat{\phi}_j\big(-\mathcal{D}_j'(\mathcal{S}) z_j F(\tilde{e}_j)\big)}_{\propto\ \text{SCC}} + \frac{1}{\widehat{\phi}_i}\underbrace{\mathcal{C}_{EE}\,\frac{1}{I}\sum_j \widehat{\phi}_j(e_j - e_j^x)}_{=\textit{Supply Redistribution}}$$

$$\mathrm{t}_i^\varepsilon = \frac{1}{\widehat{\phi}_i}\Big[SCC + \textit{Supply Redistribution}\Big]$$

$$\mathrm{t}_i^\varepsilon = \frac{1}{\widehat{\phi}_i}\Big[I\mathbb{E}_j\Big(\mathcal{S}\gamma_j y_j\Big) + I\mathbb{C}\mathrm{ov}_j\Big(\widehat{\phi}_j, \mathcal{S}\gamma_j y_j\Big) + q^e\frac{\bar{\nu}}{E}\,\mathbb{C}\mathrm{ov}_j\Big(\widehat{\phi}_j, e_i - e_i^x\Big)\Big]$$

where γ_i is the marginal damage of climate change, y_i is total output and $\bar{\nu}_i$ the average inverse energy supply elasticity, and $\widehat{\phi}_j = \phi_j/\bar{\phi} \propto \omega_j U(c_j)$ are the social welfare weights

The planner accommodates country-specific levels of inequality for the distribution of carbon prices. Indeed, for a given – potentially arbitrary – distribution of Pareto weights ω_i, the optimal carbon tax is relatively lower for poorer countries. The two motives for carbon taxation, (i) the Pigouvian Social Cost of Carbon and (ii) the supply redistribution, changing terms-of-trade and energy rent in general equilibrium, both need to be *discounted* by the country level of inequality $\widehat{\phi}_i \propto \omega_i U'(c_i)$. The tax is reduced for countries with low consumption – due to inherently low income (due to TFP) or climate damage – or high Pareto weight in the global welfare. Lastly, the energy demand is not affected by this country-specific tax.

These main findings – that the *level* and the *distribution* of carbon taxes change with inequality – are general and hold in a dynamic quantitative model that I develop in the next sections.

2.4 From carbon taxation to carbon pricing in emissions markets

In this section, I investigate how to implement the optimal climate policy when the planner chooses to design a "cap-and-trade" emission market, such as the European Union's "emission trading system (ETS)". Emissions markets are a privileged policy solution as they simultaneously provide the strict regulation of a cap and the market efficiency of emission trading.

Consider a social planner designing the cap-and-trade system, choosing the number of allowances, or "quotas" or "permits", $\overline{\mathcal{E}}$ to be auctioned in a world market. Moreover, it also chooses the number of permits that are given "for free" to each country $\overline{\varepsilon}_i$. As a result, it controls the total supply of permits $\overline{\mathcal{E}} + \sum_i \overline{\varepsilon}_i$, which are traded on a global market at price q^ε. Finally, the representative firm in each country chooses how many permits to purchase ε_i to cover its emissions from energy use. The country i's representative household/firm faces the following problem:

$$\max_{c_i, e_i, \varepsilon_i} U(c_i) \qquad s.t \qquad \begin{cases} c_i &= \mathcal{D}_i(\mathcal{S}) z_i F(e_i) + q^e(e_i^x - e_i) - \mathcal{C}_i(e_i^x) + q^\varepsilon(\overline{\varepsilon}_i - \varepsilon_i) + \mathrm{t}_i^{ls} \\ e_i &\leq \varepsilon_i \end{cases}$$

where the firm pay for ε_i carbon permits at price q^ε. This yields the optimality condition for energy e_i and carbon permits ε_i:

$$MPe_i = q^e + q^\varepsilon$$

which implies that the implicit carbon tax is $MPe_i - q^e = \tilde{\mathrm{t}}^\varepsilon = q^\varepsilon$. The complete treatment of this example is detailed in Appendix A.3. Below, I summarize several lessons on the optimal design of this type of policy.

First, the unconstrained distribution of "free carbon permits" $\overline{\varepsilon}_i$ acts as implicit money transfers across countries. Indeed, if the planner "offers" the ownership of carbon permits, countries/firms can sell them on the market to get $\overline{\varepsilon}_i q^\varepsilon$ and redistribute that money to households. Moreover, due to the government budget constraints, the revenue and cost of those permits are redistributed lump-sum to the household $\sum_i \mathrm{t}_i^{ls} = q^\varepsilon \sum_i (\varepsilon_i - \overline{\varepsilon}_i)$. If these implicit transfers are allowed, the planner can achieve full redistribution $\omega_i U'(c_i) = \omega_j U'(c_j)$. As a result, we can recover the First-Best allocation with $q^\varepsilon = SCC$, exactly as in Proposition 1 and Section 2.1.

Second, if we prevent both explicit and implicit transfers, i.e. $\bar{\varepsilon}_i = 0$, the planner again needs to adjust the carbon price for the presence of inequality, as represented by the social welfare weights $\widehat{\phi}_i = \omega_j U(c_j) / \frac{1}{I} \sum_i \omega_i U(c_i)$. This implies that the total quantity $\bar{\mathcal{E}}$ to be auctioned on a global market should be chosen to target a carbon price of q^ε as in the Second-Best in Section 2.2.

Corollary 5 (Global carbon price, cap-and-trade system with limited transfers). *The carbon price target on a global cap-and-trade emission market needs to target the following level, accounting for redistribution motives, exactly as in Proposition 2:*

$$q^\varepsilon = SCC + \textit{Supply Redistribution} + \textit{Demand Distortion}$$

$$q^\varepsilon = -\sum_j \widehat{\phi}_j \mathcal{D}_j'(\mathcal{S}) z_j F(e_j) + \mathcal{C}_{EE} \frac{1}{I} \sum_j \widehat{\phi}_j (e_j - e_j^x) + \frac{1}{I} \sum_j \widehat{v}_j \mathcal{D}_j(\mathcal{S}) z_j F''(e_j)$$

$$q^\varepsilon = I\mathbb{E}_j\Big(\mathcal{S}\gamma_j y_j\Big) + I\mathbb{C}\mathrm{ov}_j\Big(\widehat{\phi}_j, \mathcal{S}\gamma_j y_j\Big) + q^e \frac{\bar{\nu}}{E} \mathbb{C}\mathrm{ov}_j\Big(\widehat{\phi}_j, e_j - e_j^x\Big) - q^e\, \mathrm{Cov}_j\Big(\widehat{v}_j, \tfrac{1-s_i^e}{\sigma_i^e e_i}\Big)\,.$$

Third, if the planner wants to achieve additional redistribution, it needs to design *segmented markets* with $I = \#\mathbb{I}$ different prices q_i^{ε}. In each market, the planner needs to set the supply $\bar{\mathcal{E}}_i = e_i$ to achieve the target prices $q_i^{\varepsilon} \equiv \mathrm{t}_i^{\varepsilon} = (1/\widehat{\phi}_i)(SCC + \textit{Supply redistribution})$, which are the same level as Proposition 4 in Section 2.3. In that context, the carbon permits/allowances are not tradeable across countries unless against the exchange rate $q_i^{\varepsilon}/q_j^{\varepsilon}$ for one ton of CO_2.

We see that the question of the number of instruments prevails over the *nature* of the instruments – price or quantity – as I explain in the next section.

2.5 Prices vs. Quantity

I now discuss the pros and cons of a global carbon tax or country-specific carbon taxes, or *price instruments*, in comparison to quantity targets in a cap-and-trade system, either with a global carbon budget or country-specific carbon targets. In that sense, I explore the arguments made in Weitzman (2003) and Weitzman (2015) on these two types of instruments.

Weitzman (2015) argues strongly against quantity targets and emissions cap-and-trade systems. He defends that a universal carbon tax ideally possesses (1) cost-effectiveness, (2) a natural one-dimensional *focal point* (as in Schelling), and (3) a built-in self-enforcement mechanism that internalizes the externality, balancing out countervailing forces and heterogeneous interests. According to him, n different quantity assignments, as in the top-down approach of the Kyoto Protocol, fail on the second and third points and are harder to bargain upon in international agreements. It spurs free-riding as each nation has a self-interest to negotiate a low cap on their own carbon emissions – lower than socially optimal and "disagreements over the subdivision of an aggregate world cap into n national quantity targets" prevent the achievement of an efficient solution.

I showed above that (i) the availability of transfers and (ii) the choice of instruments – whether a global policy or country-specific ones, are the most important determinants of the policy effectiveness, rather than the choice of quantity or price instruments. Indeed, as noted in Section 2.4, there is a direct mapping between a quantity of carbon permits $\bar{\mathcal{E}}$ along with a distribution of "free allowances" $\{\bar{\varepsilon}_i\}_i$ and a carbon tax t^{ε} along with cross-countries transfers t_i^{ls}, and both replicate the First-Best allocation. If, like Weitzman, we note that allocating carbon permits appears as "visible transfer payments across national borders", we could instead advocate for $\bar{\varepsilon}_i = 0$ to refrain from transfers that are undermined by free-riding incentives.

In that context, there is also a direct link between choosing a total quantity of permits $\bar{\mathcal{E}}$ and a single carbon price $q^{\varepsilon} = \mathrm{t}^{\varepsilon}$ that replicates Weitzman's idea of an "internationally harmonized but nationally retained carbon tax" exactly as detailed in Section 2.2, Proposition 2 and corollary 5. Regarding policy uncertainty and the volatility of carbon price in cap-and-trade systems for a given quantity target $\mathcal{E}$, one could note that policymakers could control the quantity of permits supplied and perform "open-market operations" to reach a stable carbon price target. Finally, country-specific quantity $\bar{\mathcal{E}}_i = e_i$ and country-specific carbon taxes $\{\mathrm{t}_i^{\varepsilon}\}$ both suffer from the same pitfalls because of free-riding, bargaining frictions, and transaction costs in international negotiations. These considerations are examined in Bourany (2024*a*).

2.6 Extensions and implication for carbon taxation

In the previous sections, I showed how the carbon tax should differ from the representative agent Pigouvian framework in the simplest model with heterogeneous countries. In Bourany (2024*a*), I analyze the case with two important extensions. First, international trade has redistributive "carbon leakage" effects" that dampen the effectiveness of carbon policy. For example, imposing a carbon tax on energy in country i reallocates production toward other countries j that now may export toward i. In Bourany (2024*a*), I show that the presence of inequality and trade create a fourth motive that differentiates the optimal carbon tax from the social cost of carbon.

Second, the carbon policy is chosen by a social planner at the world level. However, free-riding – when individual countries deviate without implementing the socially optimal policies – create participation constraints that limit what can be achieved by the planner. In Bourany (2024*a*), I study how to optimally design climate agreements accounting for such free-riding incentives.

Finally, the present analysis is deterministic: both the future gains and losses of climate policies and climate change are known to all agents. In Bourany (2023), I study how uncertainty about future climate damage and future economic opportunities create insurance motives for climate policy, as the Social Cost of Carbon would be heightened due to climate risk.

3 Quantitative model

I develop a neoclassical framework with rich heterogeneity across regions. Time is continuous $t \in [t_0, \infty)$, and the countries are indexed by $i \in \mathbb{I}$. They can be heterogenous in an arbitrary number of dimensions as summarized in the state variables s_{it}[10].

In each country, we consider five representative agents: (i) a household making consumption and saving decisions, (ii) a firm using capital, labor, and different energy sources to produce the final good, and three energy firms that (iii) extract fossil fuels (oil and gas), (iv) produce coal, and (v) produce renewable/non-carbon energy. Finally, each country has a government that collects taxes and redistributes lump-sum rebates.

3.1 Household

Each region $i \in \mathbb{I}$ is populated by a representative household of size $\mathcal{P}_{it}$ at time t. This population is increasing at an exogenous constant growth rate n_i, and $\dot{\mathcal{P}}_{it} = n_i \mathcal{P}_{it}$. As a result, the population is given as $\mathcal{P}_{it} = \mathcal{P}_{i0} e^{n_i t}$. The household owns all the different firms, including the representative firm that produces goods with total factor productivity z_{it}, which also grows with growth rate $\bar{g}_i$, implying $z_{it} = z_{i0} e^{\bar{g}_i t}$. In the tradition of the Neoclassical model, I normalize all the economic variables of the model per "effective capita", dividing by the trend $e^{(n_i + \bar{g}_i)t}$.

The household consumes the homogeneous final good c_{it} and is affected by the region's temperature τ_{it}. They can save and borrow in a liquid financial asset b_{it} at a world interest rate

[10] More precisely, state variables of heterogeneity can be split in two, $s_{it} = \{\underline{s}_i, \overline{s}_{it}\}$, where $\underline{s}_i$ represents ex-ante heterogeneity and states variables $\overline{s}_{it}$ represent ex-post heterogeneity that changes over time. In practice, the dimensionality of $\underline{s}$ can be arbitrarily large, as I explain in the computational section below.

$r_t^\star$. Moreover, they can invest and hold that wealth in capital k_{it} to be rented to the homogeneous good producer at rate r_{it}^k. Households supply locally their inelastic labor $\bar{\ell}_i = 1$ to the final good firm, receiving the wage income $\bar{\ell}_i \mathrm{w}_{it}$. Moreover, the household receives the profit that the fossil firm generates π_i^f, as will be detailed below. They maximize the per-capita present discounted utility with discount rate ρ, and solve the following intertemporal problem.

$$\mathcal{V}_{it_0} = \max_{\{c_{it}, b_{it}, k_{it}\}} \int_{t_0}^{\infty} e^{-(\rho - n_i)t} \; u_i(c_{it}, \tau_{it}) \, dt \tag{6}$$

The utility that households receive from consumption is also scaled by a damage function $\mathcal{D}_i^u(\tau)$, which represents the direct impact of temperature τ_{it}. I consider standard CRRA preference with the intertemporal rate of substitution $1/\eta$.

$$u_i(c_{it}, \tau_{it}) = u\Big(\mathcal{D}_i^u(\tau_{it}) c_{it}\Big) \qquad\qquad u(\mathcal{D}\, c) = \frac{(\mathcal{D}c)^{1-\eta}}{1-\eta} \; . \tag{7}$$

We aggregate the bond and capital of the individual country as a single wealth variable $w_{it} = k_{it} + b_{it}$ and rescale income and wealth per effective unit of labor, accounting for TFP and population growth $\bar{g}_i + n_i$, it yields the dynamics of wealth w_{it} accumulation:

$$\dot{w}_{it} = (r_t^\star - (n_i + \bar{g}_i)) w_{it} + \mathrm{w}_{it}\bar{\ell}_i + \pi_{it}^f + \mathrm{t}_{it}^{ls} \; , \tag{8}$$

starting from initial condition $w_{t_0} = k_0 + b_0$. The return on capital is r_{it}^k is equalized to the bond return $r_{it}^k = r_t^\star$ in the absence of other financial market frictions. Furthermore, the household receives the energy sector profits, and, more specifically, the profit from the oil-gas firm π_{it}^f that generates meaningful energy rents. Finally, the household also receives lump-sum transfers t_i^{ls} from the government. Wealth w_{it} is the first dimension of ex-post heterogeneity.

3.2 Final good firm

In each country $i \in \mathbb{I}$, a representative firm produces a homogeneous final good using a constant-return-to-scale technology $F(\cdot)$ and different inputs: labor $\bar{\ell}_i$ at wage w_{it}, capital per effective capita k_{it} at the rental rate $r_t^\star$, and energy per effective capita e_{it} at price q_{it} as detailed below.[11] The firm maximizing profit, i.e. output per capita $y = \mathcal{D}^y(\tau) z F(\cdot)$ net of input costs:

$$\max_{k_{it}, e_{it}} \; \mathcal{D}_i^y(\tau_{it}) z_i F_i(k_{it}, \bar{\ell}_i, e_{it}) - \mathrm{w}_{it}\bar{\ell}_i - q_{it}^e e_{it} - (r_t^\star + \delta) k_{it} \; . \tag{9}$$

The firm's productivity first differs across countries due to institutional and efficiency time-invariant factors summarized in z_i. Second, temperatures τ_{it} affect output through climate damages $\mathcal{D}_i^y(\tau_{it})$,

[11] The original – unnormalized – production function:

$$Y_t = \widetilde{F}(K_t, L_t, E_t) = \mathcal{D}(\tau_t) z_t \Big[(1-\varepsilon)^{\frac{1}{\sigma}} \left(K_t^\alpha L_t^{1-\alpha} \right)^{\frac{\sigma-1}{\sigma}} + \varepsilon^{\frac{1}{\sigma}} \left(z_t^e E_t \right)^{\frac{\sigma-1}{\sigma}} \Big]^{\frac{\sigma}{\sigma-1}}$$

I normalize output Y_t by the trends in TFP $e^{\bar{g}_i t}$ and population $L_t \equiv \mathcal{P}_0 e^{n_i t}$ to obtain output per effective capita.

which is the source of climate externality which will be detailed below. The production function has a constant elasticity of substitution between the capital-labor bundle $k^\alpha \ell^{1-\alpha}$ and energy e:

$$F_i(k_{it}, \ell_i, e_{it}) = \Big[(1-\varepsilon)^{\frac{1}{\sigma}} (k_{it}^\alpha \, \ell_i^{1-\alpha})^{\frac{\sigma-1}{\sigma}} + \varepsilon^{\frac{1}{\sigma}} (z_{it}^e e_{it})^{\frac{\sigma-1}{\sigma}}\Big]^{\frac{\sigma}{\sigma-1}}$$

with $\sigma < 1$, such as energy is complementary in production[12] and where directed technical change z_t^e is exogenous and deterministic. This directed – energy augmenting – technical change increases in output for a given energy consumption mix.[13] This exogenous trend implies $z_{it}^e = \bar{z}_i^e \, e^{g_e t}$.

The firm optimal inputs decision for capital, k_{it}, labor ℓ_i, and energy e_{it} is such that the marginal product of the inputs equals its price. With the marginal product of capital, labor and energy are defined as: $MPx_{it} = \partial_x[\mathcal{D}_i^y(\tau_{it}) z_i F_i(k_{it}, \ell_i, e_{it})]$ for $x \in \{k, \ell, e\}$, we obtain the first-order conditions:

$$r_{it}^k = MPk_{it} - \delta = r_t^\star \qquad \mathrm{w}_{it} = MP\ell_{it} \qquad q_{it}^e = MPe_{it} \tag{10}$$

Energy demand

Given the demand for energy inputs e_t in each country, the firm has the choice among three sources of energy: fossil fuels e_{it}^f, coal e_{it}^c and low-carbon/renewables e_{it}^r. These three sources are substitutable, and total energy e_{it} has constant elasticity of substitution σ_e.

$$e_{it} = \Big(\omega_f^{\frac{1}{\sigma_e}} (e_{it}^f)^{\frac{\sigma_e-1}{\sigma_e}} + \omega_c^{\frac{1}{\sigma_e}} (e_{it}^c)^{\frac{\sigma_e-1}{\sigma_e}} + \omega_r^{\frac{1}{\sigma_e}} (e_{it}^r)^{\frac{\sigma_e-1}{\sigma_e}}\Big)^{\frac{\sigma_e}{\sigma_e-1}}$$

subject to the budget for energy expenditures, which implies the price of the energy bundle q_{it}, for $\sigma_e \in (0, \infty)$:

$$q_{it}^e e_{it} = e_{it}^f (q_t^f + \xi^f \mathrm{t}_{it}^\varepsilon) + e_{it}^c (q_{it}^c + \xi^c \mathrm{t}_{it}^\varepsilon) + e_{it}^r q_{it}^r$$

$$q_{it}^e = \Big(\omega_f (q_t^f + \xi^f \mathrm{t}_{it}^\varepsilon)^{1-\sigma_e} + \omega_c (q_{it}^c + \xi^c \mathrm{t}_{it}^\varepsilon)^{1-\sigma_e} + \omega_r (q_{it}^r)^{1-\sigma_e}\Big)^{\frac{1}{1-\sigma_e}}$$

where q_t^f is the international price of oil and gas, q_{it}^c the local price of coal energy, and q_{it}^r the local price of low-carbon energy. Similarly, the representative final good firm choose the energy inputs according to the first-order condition:

$$q_t^f + \xi^f \mathrm{t}_{it}^\varepsilon = MPe_{it}^f = q_{it}^e \, \omega_f^{\frac{1}{\sigma_e}} \Big(\frac{e_{it}^f}{e_{it}}\Big)^{-\frac{1}{\sigma_e}} \tag{11}$$

and similarly for the other energy inputs $q_{it}^c + \xi^c \mathrm{t}_{it}^\varepsilon = MPe_{it}^c$ and $q_{it}^r = MPe_{it}^c$.

Energy from oil-gas, e_i^f, and coal, e_i^c, differ from renewable in the sense that they emit greenhouse gases, with respective carbon concentration ξ^f and ξ^c, as we will see in Section 3.4. As a result, there is a motive for taxing oil, gas, and coal energy with the carbon tax $\mathrm{t}_{it}^\varepsilon$, which is a tax per ton of CO_2. I discuss the choice of this tax in the next sections.

[12] If $\sigma = 1$, we have the Cobb Douglas : $F(k, \ell, e) = \bar{\varepsilon} z_t^{e\,\varepsilon} (k^\alpha \ell^{1-\alpha})^{1-\varepsilon} e^\varepsilon$

[13] An upward trend in such technology is sometimes argued to be behind the "relative decoupling" of developed economies: an increase in production and value-added simultaneous to a decline in energy consumption.

3.3 Energy markets

The final good firm consumes three kinds of energy sources – oil and gas, coal, or renewable (non-carbon) – supplied by three representative energy firms in each country $i \in \mathbb{I}$. Oil and gas sources are traded internationally, while coal and renewable sources are both traded locally.

3.3.1 Fossil firm

A competitive energy producer extracts fossil fuels oil and gas e_i^x from its pool of resources $\mathcal{R}_{it}$. The energy is extracted with convex production cost $\nu_i^f(e_{it}^x, \mathcal{R}_{it})$, where these costs are paid in units of the final good, and the oil and gas are sold in international markets at a price q^f.

The fossil-fuel reserves $\mathcal{R}_{it}$ are depleted with extraction e_{it}^x such that $\dot{\mathcal{R}}_{it} = -e_{it}^x$. We assume that neither the fossil firms nor the social planners internalize the scarcity of these resources. Internalizing the resource depletion would imply a more involved dynamic Hotelling problem – with stock effects – with the Hotelling rent rising over time, dampening the extraction rate e_i^x. I suggest an extension in appendix XX to see how these motives would change the taxation of fossil fuel and carbon theoretically. However, in the interest of keeping the framework simple, I refrain from considering this extension in the quantitative analysis. Richer models developed in Bornstein, Krusell and Rebelo (2023), Heal and Schlenker (2019), and Asker, Collard-Wexler, De Canniere, De Loecker and Knittel (2024) study the dynamic aspects of the oil market and the considerations for carbon emissions.

As a result, the static maximization problem of the fossil firm is given by :

$$\begin{aligned} \pi_{it}^f &= \max_{e_{it}^x} q_t^f e_{it}^x - \nu_i^f(\mathcal{P}_{it} e_{it}^x, \mathcal{R}_{it})/\mathcal{P}_{it} \ , \\ \dot{\mathcal{R}}_{it} &= -\mathcal{P}_{it} e_{it}^x \end{aligned} \tag{12}$$

where $\nu_i^f(\mathcal{P}_{it} e_{it}^x, \mathcal{R}_{it})/\mathcal{P}_{it}$ is the extraction cost per capita, which is convex in e_{it}^x, and $\mathcal{R}_{it_0} = \mathcal{R}_{i0}$ the initial condition for reserves. Since the extraction costs are convex, the production function has decreasing return to scale.[14] As a result, a positive energy rent π_{it}^f exists, even if the competitive firm takes the fossil price q_t^f as given. Moreover, I abstract from market power in the oil market, for example with the OPEC as a cartel – even though this framework could easily allow for such an extension. Any sources of misallocation – in the sense of Hsieh and Klenow (2009) – are accounted for in the calibration of the cost function $\nu_i^f(\cdot)$ as we will see in the quantification Section 5. I consider a functional form for cost that yields isoelastic supply curves for fossil energy extraction.

$$\nu_i^f(e_{it}^x, \mathcal{R}_{it}) = \frac{\bar{\nu}_i}{1+\nu_i}\Big(\frac{e_{it}^x}{\mathcal{R}_{it}}\Big)^{1+\nu_i} \mathcal{R}_{it} \tag{13}$$

which is homogeneous of degree one in $(e_i^x, \mathcal{R}_i)$ and where the elasticity $\nu_i = \frac{\nu_i^{f\,\prime\prime}(e^x,\mathcal{R})}{\nu_i^{f\,\prime}(e^x,\mathcal{R})e^x}$ is constant.

[14] We can also define a fossil production function with inputs x_i^f such that $e^x = g(x_i^f)$ and profit $\pi = q^f g(x) - x$ instead of $\pi = q^f e^x - \nu(e^x)$, in which case $g(x) = \nu^{-1}(x)$

Naturally, the optimal extraction decision for the fossil firm follows from the optimality condition:

$$q_t^f = \nu^f_{i\,e^x}(\mathcal{P}_{it} e^x_{it}, \mathcal{R}_{it}) = \bar{\nu}_i \Big(\frac{\mathcal{P}_{it} e^x_{it}}{\mathcal{R}_{it}}\Big)^{\nu_i} \tag{14}$$

which yields the implicit function $e^{x\star}_{it} = e^x_i(q^f_t) = \nu_i^{f\prime\,-1}(q^f_t) = \mathcal{R}_{it}\ (q_t^f/\bar{\nu}_i)^{1/\nu_i}$ for the optimal oil and gas extraction.

Finally, energy rent comes from fossil firms' profits $\pi^f(q^f, \mathbb{P}_i) = q^f e^x(q^f) - \nu_i^f(e^x(q^f), \mathcal{R}) > 0$, and the per-capita profit function writes:

$$\pi_i^f(q_t^f, \mathcal{R}_{it}) = q_t^f e^x_{it} - \nu_i^f(\mathcal{P}_{it} e^x_{it}, \mathcal{R}_{it})/\mathcal{P}_{it} = \frac{\nu_i \bar{\nu}_i}{1+\nu_i}\Big(\frac{\mathcal{P}_{it} e^x_i}{\mathcal{R}_i}\Big)^{1+\nu_i} \frac{\mathcal{R}_{it}}{\mathcal{P}_{it}} = \frac{\nu_i \bar{\nu}_i^{-1/\nu_i}}{1+\nu_i} \mathcal{R}_{it} (q_t^f)^{1+\frac{1}{\nu_i}} \ . \tag{15}$$

As we will see below, the profit $\pi_i^f(q^f, \mathcal{R})$ and its share in income $\eta_{it}^{\pi f} = \frac{\pi^f_{it}}{y_i + \pi^f_{it}}$ are key to determine the exposure of a country to carbon taxation. Indeed, reducing carbon emissions by phasing out of fossil fuels reduces energy demand and its price q^f and hence affects energy profit π_i^f and the welfare of large oil and gas exporters.

3.3.2 International fossil energy markets

Oil and gas are traded frictionlessly in international markets.[15] The market clears such that

$$E^f_{it} = \sum_{i \in \mathbb{I}} \mathcal{P}_i e^{(n_i + \bar{g}_i)t}\, e^f_{it} = \sum_{i \in \mathbb{I}} e^{(n_i + \bar{g}_i)t} e^x_{it} \ . \tag{16}$$

Countries have different exposure to this fossil energy market. As country i consumes fossil fuels in total quantity $\mathcal{P}_{it} e^f_{it} = \mathcal{P}_i e^{(n_i+\bar{g}_i)t} e^f_{it}$, and produces total quantity $\mathcal{P}_{it} e^x_{it} = \mathcal{P}_i e^{(n_i+\bar{g}_i)t} e^x_{it}$, its net exports of oil and gas per effective capita are $e^x_i - e^f_i \lesseqgtr 0$.

3.3.3 Coal firm

A representative firm produces coal that is consumed by the final good firm. I differentiate coal from other fossil fuels like oil and gas because coal production typically does not generate large energy rents for producing countries as a share of GDP. Moreover, large coal producers also consume a large fraction of that coal locally, as trade costs for coal transportation are larger. Hence, I make this empirically grounded assumption that coal is not traded internationally. Moreover, I assume that production is not subject to the finiteness of the stock of reserves. Indeed, the scarcity of coal sources is not a concern since the ratio of reserve/production is above a hundred for most large world producers.

The production $\bar{e}^c_{it}$ has constant returns to scale and uses final good inputs. The profit

[15] For the sake of simplicity, I make the simplifying assumption that fossil fuels produced in different countries are not distinguishable – crude oil or natural gas from Nigeria, Saudi Arabia, or Russia are not differentiated varieties.

maximization problem is analogous to the fossil problem:

$$\pi_{it}^c = \max_{\bar{e}_{it}^c} q_i^c \bar{e}_{it}^c - \kappa_i^c \bar{e}_{it}^c \ ,$$

where the marginal cost κ_i^c is constant. This implies that there is no coal profit[16] in equilibrium, i.e. $\pi_{it}^c = 0$. The price for coal and the market clearing condition are given by:

$$q_i^c = \kappa_i^c \ , \qquad\qquad \bar{e}_i^c = e_i^c \ . \tag{17}$$

This implies a perfectly elastic supply curve for coal energy, something we observe in practice as coal production is easily scalable in response to oil and gas price fluctuations.

3.3.4 Low-carbon, renewable, firm

The final good firm also uses renewable and other low-carbon energy sources, such as solar, wind, or nuclear electricity. This provides a way of substituting away from fossil fuels in the production function $F(\cdot)$.

A representative firm produces renewable or non-carbon energy, and this supply, $\bar{e}_{it}^r$, is not traded. This assumption is verified by the fact that electricity is rarely traded across countries – and when it is, it is only the result of temporary differences in electricity production due to intermittency rather than large structural imbalances. The production $\bar{e}_{it}^r$ also has constant returns to scale, and this input is paid in units of the final good. Hence, the renewable firm maximization problem is:

$$\pi_{it}^r = \max_{\bar{e}_{it}^r} q_{it}^r \bar{e}_{it}^r - \kappa_{it}^r \bar{e}_i^r \ ,$$

where κ_{it}^r is the marginal cost of producing renewables, resulting in zero profits $\pi_i^r = 0$. The price of renewable and the market clearing are given by:

$$q_{it}^r = \kappa_{it} = \bar{\kappa}_i e^{-g_r t} \ , \qquad\qquad e_{it}^r = \bar{e}_{it}^r \tag{18}$$

where I assume that the marginal cost κ_{it} decreases exogenously at rate g_r such that $\kappa_{it} =$ is lower over time. Given those marginal costs, this returns a perfectly elastic supply curve. This is a slightly stronger assumption in the context of renewable energy: In the short run, renewable energy requires investments in capacity, implying a fairly inelastic supply curve. This is especially true considering the intermittency problems of wind and solar energy, c.f. Gentile (2024). I take the conservative assumption that the supply curve is flat in the medium run.

However, in the long run, technological progress and learning-by-doing create positive externalities, substantially decreasing the cost of clean energy, resulting in a decreasing supply curve. To allow this learning-by-doing effect, as in Arkolakis and Walsh (2023), I consider an extension

[16] This is motivated by evidence that even the largest coal producers do not have coal rents above 1% of GDP.

where the marginal cost depends on power capacity:

$$q_{it}^r = \kappa_{it} = \bar{\kappa}_i \left(\mathcal{C}_{it}^r\right)^{-\kappa^r} \qquad \text{where} \qquad \dot{\mathcal{C}}_{it}^r = \max\{\dot{\bar{e}}_{it}^r, 0\}^{\gamma_r} \tag{19}$$

where the renewable capacity $\mathcal{C}_{it}^r$ increases for each additional renewable production $\bar{e}_{it}^r$. The capacity of production, in kW, increases for each additional kWh of production needed $\bar{e}_{it}^r$ at each point in time, and this by a factor γ_r, and decreases production costs by a factor κ^r. The parameter $\gamma_r \in (0, \infty)$ represents a *learning-by-doing* factor. Indeed, $\mathcal{C}_{it}^r = \int_{t_0}^t |\dot{\bar{e}}_{it}^r|^{\gamma_r} dt$, and if $\gamma_r > 1$ capacity increases more than one for one due to increasing return to scale. Note that when $\gamma_r = 1$, and $\dot{\bar{e}}_{it}^r > 0, \forall t$, we simply get that $\mathcal{C}_{it}^r = \bar{e}_{it}^r$, i.e. the capacity is exactly what is needed for production. This would imply a downward-sloping supply curve with (negative) supply elasticity $-\kappa^r$. I assume that neither the firm nor the social planner internalizes this learning-by-doing externality. The study of the optimal policy in the presence of both negative climate externalities and positive Schumperian externalities is beyond the scope of this paper.

3.4 The climate system

The economic activity and fossil fuel consumption of each country create a climate externality by emiting carbon in the atmosphere. This feeds back in the climate system, which increases the temperatures and causes heterogeneous damages over different regions. As in standard Integrated Assessment Models (IAM), it creates a Pigouvian motive for carbon taxation as summarized by the Social Cost of Carbon.

3.4.1 Emissions

The consumption of fossil fuels are emitting carbon dioxide (CO_2) and other greenhouse gas emissions in the atmosphere. Due to oil and gas e_{it}^f and coal e_{it}^c consumptions – expressed in unit per effective capita, subject to growth rate of population n_i and TFP $\bar{g}_i$– we obtain that each country $i \in \mathbb{I}$ releases total CO_2 emissions:

$$\epsilon_{it} = \mathcal{P}_i e^{(n_i + \bar{g}_i)t} \left(\xi^f e_{it}^f + \xi^c e_{it}^c\right) ,$$

where ξ^f and ξ^c denote the carbon content of respectively oil-gas and coal energy. As a result, global emissions aggregate to:

$$\mathcal{E}_t = \bar{\xi}_t \sum_{i \in \mathbb{I}} \epsilon_{it} \ .$$

I consider that emissions are non-exploding and I follow Krusell and Smith (2022) by assuming that part of emissions $\mathcal{E}_t$ is abated via carbon capture and storage (CCS) modeled by the exogenous parameter $\bar{\xi}_t$, with $\bar{\xi}_{t_0} = 1$. The share of emissions abated grows to 100% in the long-run, implying that $\bar{\xi}_t \to_{t \to \infty} 0$. Increasing CCS allows the system to reach net-zero in several centuries, stabilizing cumulative carbon emissions and temperatures.

3.4.2 Climate system and temperature

Moreover, these emissions are released in the atmosphere, adding up to the cumulative stock of greenhouse gas $\mathcal{S}_t$ – or atmospheric carbon concentration:

$$\dot{\mathcal{S}}_t = \mathcal{E}_t - \delta_s \mathcal{S}_t \ . \tag{20}$$

A part of these emissions exit the atmosphere and is stored in oceans or the biosphere, discounting the current stock by an amount δ_s. Moreover, these cumulative emissions push the global atmospheric temperature $\mathcal{T}_t$ upward linearly with climate sensitivity χ, with some inertia and delay represented by the parameter ζ.

$$\dot{\mathcal{T}}_t = \zeta(\chi \mathcal{S}_t - (\mathcal{T}_t - \mathcal{T}_{t_0})) \ . \tag{21}$$

More particularly, the inertia ζ is the inverse of persistence, and modern calibrations set $\zeta \approx 0.5$ is such that the pick of emissions happens after $10 - 15$ years. Dietz et al. (2021) show that classical IAM models such at Nordhaus' DICE tend to generate a too large climate system inertia, as shown in the Figure 2. Conversely, if $\zeta \to \infty$, temperature reacts immediately and we obtain a linear model – which is a good long-run approximation:

$$\mathcal{T}_t = \bar{\mathcal{T}}_{t_0} + \chi \mathcal{S}_t = \bar{\mathcal{T}}_{t_0} + \chi \int_{t_0}^{t} e^{-\delta_s s} \bar{\xi}_s \mathcal{E}_s \, ds$$

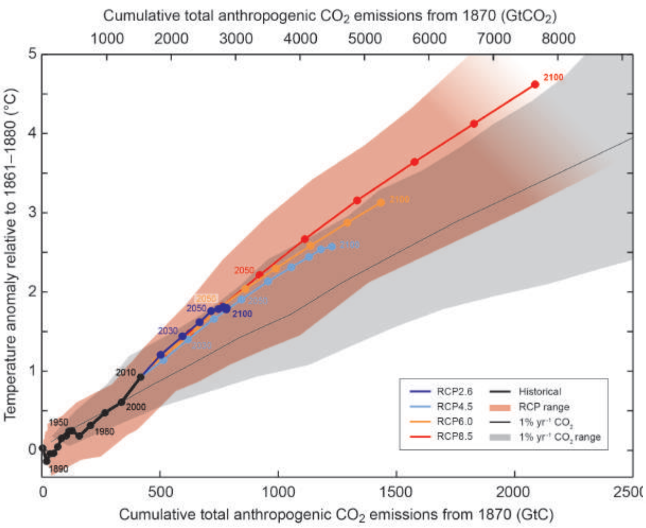


Figure 1: Linear temperature model - IPCC report

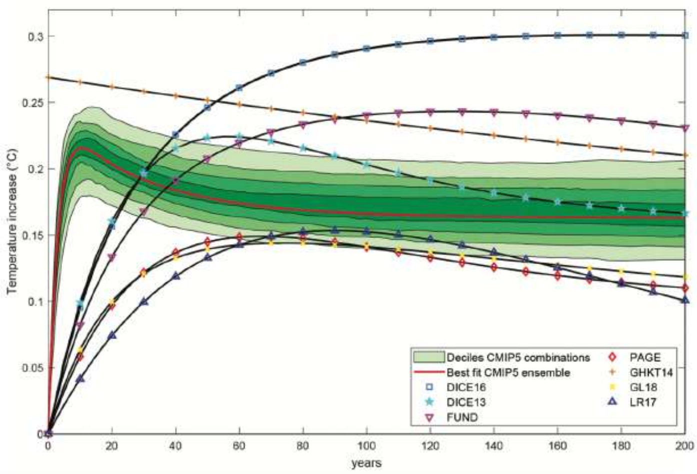


Figure 2: Pulse experiment

This simple two-equations climate system is a good approximation of large-scale climate models.[17] Indeed, with the appropriate calibration of parameters δ_s for carbon exit, ζ for climate system inertia, and χ for the climate sensitivity, we can match these larger models as represented by the pulse experiment as shown in Figure 2.

[17] These climate models have a much more complex climate block, adding 3 to 4 additional state variables, e.g.

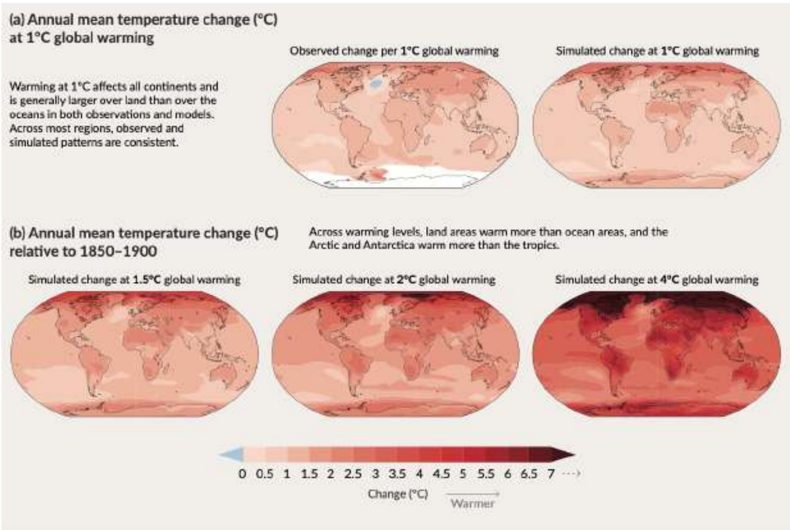
3.5 Damage and externality

Climate damages are related to local temperatures: warmer regions are more affected and vulnerable to extreme events and impacts.

The temperature in country i is affected by global warming of the atmosphere $\mathcal{T}_t$ with linear pattern scaling Δ_i

$$\begin{aligned} \dot{\tau}_{it} &= \Delta_i \, \dot{\mathcal{T}}_t \\ \tau_{it} &= \tau_{i0} + \Delta_i(\mathcal{T}_t - \mathcal{T}_{t_0}) \end{aligned} \tag{22}$$

Atmospheric temperature $\mathcal{T}_t$ translates into local temperature τ_{it} via the sensitivity Δ_i that depends on the geographic properties of country i – like temperature, latitude, longitude, elevation, distance from coasts and water bodies, vegetation, and albedo (sunlight reflexivity due to ice, vegetation and soil properties). Evidence of this temperature scaling is displayed in Section 3.5 from the IPCC report. As shown in Section 5, I estimate this pattern scaling by regressing local temperatures on global temperature.



Finally, I consider a period damage function $\mathcal{D}^y_i(\tau_{it}) := \mathcal{D}^y(\tau_{it}-\tau^\star_i)$ for productivity and $\mathcal{D}^u(\tau_{it}-\tau^\star_i)$ for damage to utility. The target $\tau^\star_i$ is the "optimal" temperature for country i. The function $\mathcal{D}^y(\hat{\tau})$ is a reduced-form representation of the economic damage. In the baseline quantification, I assume damages are quadratic, as in standard Integrated Assessment Models such as the DICE framework. This methodology follows Krusell and Smith (2022), Kotlikoff, Kubler, Polbin, Sachs and Scheidegger (2021) and Burke et al. (2015). Moreover, those damages

with $\mathbf{J}$ the vector of carbon "boxes": layers of the atmosphere and sinks such as layers of oceans:

$$\dot{\mathbf{J}}_t = \Phi^J \mathbf{J}_t + \rho^S \sum_{\mathbb{I}} \epsilon_{it}$$
$$F_t = \mathcal{F}\big(\rho^F \mathbf{J}_t\big) \qquad \dot{\mathcal{T}}_t = \Phi^T \mathcal{T} + \rho^F F_t$$

with F_t carbon forcing with $\mathcal{F}(\cdot) \sim \log(\cdot)$, ρ^S, ρ^F, and ρ^T are vectors are Φ^J and Φ^T Markovian transition matrices.

are such that productivity decays to zero when temperatures are extremely cold or hot.

$$\mathcal{D}_i^y(\tau) = \mathcal{D}^y(\tau - \tau_i^\star) = \exp\left(- \gamma^y \mathbb{1}_{\{\tau>\tau_i^\star\}}(\tau-\tau_i^\star)^2 - \alpha^\gamma\gamma^y\mathbb{1}_{\{\tau<\tau_i^\star\}}(\tau-\tau_i^\star)^2\right) , \tag{23}$$

where γ_y represents the damage parameter on output for warm temperatures, with an asymmetric impact $\alpha^\gamma\gamma^y < \gamma^y$ for cold temperatures following the quantification in Rudik et al. (2021), who show that productivity impact is much weaker for cold than for hot temperatures. The damage function for utility $\mathcal{D}_i^u(\tau)$ has the same functional form with damage γ^u.

This creates winners and losers: countries warmer than their target temperature $\tau_i^\star$ are extremely affected by global warming. In contrast, regions with negative $\tau_{it}-\tau_i^\star$ benefit – at least in the short-run – from a warmer climate. I deviate from the above articles by assuming that the target temperature $\tau_i^\star$ differs across countries: an already warm regions have different adaptation costs compared to a country which is historically cold. The target temperature $\tau_i^\star = \alpha^\tau\tau^\star + (1-\alpha^\tau)\bar{\tau}_{it_0}$ is more or less tilted toward historical baseline. I discuss this quantification in Section 5.

3.6 Competitive Equilibrium

The final good is freely traded, and, with output y_i, the market clearing holds:

$$\sum_{i\in\mathbb{I}} \mathcal{P}_i e^{(n_i+\bar{g}_i)t}\Big[c_{it}+(\dot{k}_{it}+(n_i+\bar{g}_i+\delta)k_{it})+\nu_i^f(e_{it}^x,\mathcal{R}_{it})+\kappa_i^c e_{it}^c+\kappa_{it}^r e_{it}^r\Big] = \sum_{i\in\mathbb{I}} \mathcal{P}_i e^{(n_i+\bar{g}_i)t}\mathcal{D}_i(\tau_{it})z_{it}F(k_{it},\ell_i,e_{it}) \,. \tag{24}$$

Similarly, the bond market, in zero net supply, clears[18] such that $\sum_i \mathcal{P}_i e^{(n_i+\bar{g}_i)t}b_{it} = 0$

Definition. Competitive equilibrium (C.E.):
For a set of policies $\{\mathrm{t}_{it}^\varepsilon, \mathrm{t}_{it}^{ls}\}_{it}$ across countries, a C.E. is a set $\{c_{it}, b_{it}, k_{it}, e_{it}^f, e_{it}^c, e_{it}^r, e_{it}^x, \bar{e}_{it}^c, \bar{e}_{it}^r\}_{it}$ of decisions, and prices $\{r_t^\star, q_t^f, \mathrm{w}_{it}, q_{it}^c, q_{it}^r\}_{it}$, and states $\{w_{it}, \tau_{it}, \mathcal{R}_{it}, \mathcal{S}_t, \tau_{it}\}_{it}$ such that:

(i) Households choose consumption, saving and investment, $\{c_{it}, k_{it}, b_{it}\}_{it}$ maximizing utility as in equation (6) subject to the budget constraint and wealth dynamics equation (8).

(ii) Final good firms choose inputs $\{k_{it}, \ell_i, e_{it}, e_{it}^f, e_{it}^c, e_{it}^r\}_{it}$ to maximize profits, resulting in input choices following equation (11) and equation (10).

(iii) Fossil energy firms maximize profits as in equation (12) and extract/produce $\{e_i^x\}_i$ given by equation (14)

(iv) Renewable and coal energy firms maximize profits, and supplies $\{\bar{e}_i^c, \bar{e}_i^r\}$ are given respectively by equation (17) and equation (18)

(v) Energy markets clears for fossils as in equation (16) and for coal and renewable in equation (17) and equation (18)

(vi) The emissions $\mathcal{E}_t = \sum_i \epsilon_{it}$ affects the climate system $\{\mathcal{S}_t, \mathcal{T}_t, \tau_{it}\}_{it}$, following equation (20), equations (21) and (22).

(vii) Good markets clear for final good for each country as in equation (24), and bond market clear by Walras law.

[18]This is also equivalent to $\sum_i \mathcal{P}_i e^{(n_i+\bar{g}_i)t}w_{it} = \sum_i \mathcal{P}_i e^{(n_i+\bar{g}_i)t}k_{it}$.

Heterogeneity. This model features many dimensions of heterogeneity, that can be summarized by the state variable $s_{it} = \{\underline{s}_i, \overline{s}_{it}\}$, describing the ex-ante dimensions of heterogeneity $\underline{s}_i$ – differences across countries that do not change over time – and ex-post heterogeneity $\overline{s}_{it}$ that change endogeneously over time. The states are: $\underline{s}_i = \{\mathcal{P}_i, n_i, \bar{g}_i, z_i, z^e_i, \bar{\nu}_i, \nu_i, \bar{\kappa}^c_i, \bar{\kappa}^r_i, \Delta_i, \tau^\star_i, \tau_{i0}, w_{i0}, \mathcal{R}_{i0}\}$ and $\bar{s}_{it} = \{w_{it}, \tau_{it}, \mathcal{R}_{it}\}$. For solving the model, we need to keep track of this dynamical system with heterogeneity, as we see next.

3.7 Sequential formulation and household decisions

First, since the Household owns the four firms – final good, fossil, coal, and renewable energy – we can aggregate profits and household budget constraint, which gives:

$$\begin{aligned} \dot{w}_{it} = (r^\star_t - (n_i{+}\bar{g}_i))w_{it} + \pi^f_i(q^f_t, \mathcal{R}_{it}) + \mathcal{D}^y_i(\tau_{it})z_{it}F(k_{it}, \ell_i, e^f_{it}, e^c_{it}, e^r_{it}) - (r^\star + \delta)k_{it} \\ - (q^f_t + \xi^f \mathrm{t}^\varepsilon_{it})e^f_{it} - (q^c_{it} + \xi^c \mathrm{t}^\varepsilon_{it})e^c_{it} - q^r_{it}e^r_{it} - c_{it} + \mathrm{t}^{ls}_{it} \ , \end{aligned} \tag{25}$$

and yields a single optimal control problem. The consumption/saving relates to the path of wealth w_{it}, given that the firms decisions, given by the optimality conditions

To solve for the competitive equilibrium and the optimal decision of the Household, we solve this class of Integrated Assessment Model with the sequential formulation of optimal control problem. This relies on the Pontryagin Maximum Principle, which can be applied in heterogeneous agents settings – with discrete agents in our case, or continuous agents/Mean-Field Games, as in Carmona and Delarue (2018). The household in each country has individual states $\mathbf{s} = \{\underline{s}_i, \overline{s}_{it}\}_{it} = \{\underline{s}_i, w_{it}, \tau_{it}, \mathcal{R}_{it}\}_{it}$, individual controls, $\mathbf{c} = \{c_{it}, b_{it}, k_{it}, e^f_{it}, e^c_{it}, e^r_{it}, e^x_{it}\}_{it}$, take prices as given $\mathbf{q} = \{r^\star_t, q^f_t, \mathrm{w}_{it}, q^c_{it}, q^r_{it}\}_{it}$, and has costates or Lagrange multipliers, $\boldsymbol{\lambda} = \{\lambda^w_{it}, \lambda^\tau_{it}, \lambda^S_{it}\}$, each of which represents shadow value of the respective states dynamics. The Hamiltonian of the individual country can be written as follow:

$$\mathcal{H}(\mathbf{s}, \mathbf{c}, \mathbf{q}, \boldsymbol{\lambda}) = u(c, \tau) + \lambda^w \dot{w} + \lambda^\tau \dot{\tau} + \lambda^S \dot{\mathcal{S}}$$

for the dynamics $\dot{w}_{it}$ given in equation (25), $\dot{\tau}_{it}$ given by equation (22) and $\dot{\mathcal{S}}_t$ given by equation (20).

As a result, the equilibrium relations for the household consumption/saving problem boil down to the standard neoclassical model dynamics and, for each country $i \in \mathbb{I}$, we obtain a system of coupled ODEs:

$$\begin{cases} \dot{\lambda}^w_{it} &= \lambda^w_{it}(\rho + \eta\bar{g}_i - r^\star_t) \\ \lambda^w_{it} &= u_c(c_{it}, \tau_{it}) \end{cases} \tag{26}$$

where λ^w_{it} is the costate, or "marginal value" of wealth w_{it}, equal to marginal utility $u_c(c_{it}, \tau_{it}) = \mathcal{D}^u(\tau_{it})u'(\mathcal{D}^u(\tau_{it})c_{it})$. Using the optimality for c, we obtain the Euler equation:

$$\frac{\dot{c}_{it}}{c_{it}} = \frac{1}{\eta}(r^\star_t + \eta\bar{g}_i - \rho) + \gamma^c(\tau_{it} - \tau^\star_i)\dot{\tau}_{it}$$

The dynamics of local temperature appear in the Euler equation. Indeed, because the marginal utility of consumption is affected directly by changes in temperature, an increase in temperature in the future triggers substitution from present to future consumption through saving.

To close the control problem, note that the household income is determined by the firms decisions. There, the capital and energy choices simply result from static optimization between the price or cost and the marginal return of those inputs in production. However, the climate variables affect damages, and the country i household internalizes that under the "local social cost of carbon" as we will see now.

3.8 Social and Local Cost of Carbon

The Social Cost of Carbon (SCC) is a measure used by climate scientists and economists to summarize the marginal welfare cost of climate change in monetary terms. The cost of carbon is an equilibrium concept: it depends on the path of temperatures but also on economic variables and policies. In the competitive equilibrium, the climate externality is not internalized and households and firms do not take climate damages into account for choosing consumption, production, and energy decisions. Still, forward-looking agents anticipate perfectly the evolution of climate.

The Local Cost of Carbon (LCC) represents such a welfare valuation, for the cost incurred by country i of one additional ton of CO_2 released in the atmosphere. In continuous time, and using our the Pontryagin Maximum Principle sequential approach, the Local Cost of Carbon (LCC) can be written easily as the ratio of the two costates:

$$LCC_{it} := -\frac{\frac{\partial \mathcal{V}_{it}}{\partial \mathcal{S}_t}}{\frac{\partial \mathcal{V}_{it}}{\partial c_{it}}} = -\frac{\lambda_{it}^S}{\lambda_{it}^w} \,. \tag{27}$$

The welfare cost of carbon λ_{it}^S represents the marginal welfare change from an additional ton of carbon $\mathcal{S}_t$. This is normalized in monetary units with the marginal value of wealth, as indeed the monetary value of welfare differs across regions i, $\frac{\partial \mathcal{V}_{it}}{\partial c_{it}} = \lambda_{it}^w = u_c(c_{it}, \tau_{it}) \neq u_c(c_{jt}, \tau_{jt}) = \lambda_{jt}^w$, due to inequality in consumption. This notion is exactly analogous to the Local Cost of Carbon concept developed in Cruz and Rossi-Hansberg (2022*a*), among many others.

As a result, following the dynamics of the LCC amounts to solve for the dynamics of both costates λ_{it}^w and λ_{it}^S. Recalling the dynamics of the climate system:

$$\begin{cases} \mathcal{E}_t &= \sum_{i\in\mathbb{I}} \epsilon_{it} = \sum_{\mathbb{I}} \mathcal{P}_i e^{(n_i+\bar{g}_i)t}(\xi^f e_{it}^f + \xi^c e_{it}^c) \\ \dot{\mathcal{S}}_t &= \mathcal{E}_t - \delta_s \mathcal{S}_t \\ \dot{\tau}_{it} &= \zeta\big(\Delta_i \chi \mathcal{S}_t - (\tau_{it} - \tau_{it_0})\big) \end{cases}$$

we can use the Pontryagin principle to pin down the dynamics of the local cost of carbon. First, the shadow value of increasing temperatures is affected by the cost of climate on both the productivity effect $\mathcal{D}^y(\tau)zF(k,e)$ and the utility effect $u(\mathcal{D}^u(\tau)c)$.

$$\dot{\lambda}_{it}^\tau = \lambda_{it}^\tau(\rho - n_i - (1-\eta)\bar{g}_i + \zeta) + \underbrace{\gamma^y(\tau_{it} - \tau_i^\star)\mathcal{D}_i^y(\tau_{it})}_{-\partial_\tau \mathcal{D}^y} z_i F(k_{it}, e_{it})\lambda_{it}^w + \underbrace{\gamma^u(\tau_{it} - \tau_i^\star)\mathcal{D}_i^u(\tau_{it})}_{-\partial_\tau \mathcal{D}^u} u'(\mathcal{D}^u(\tau_{it})c_{it})c_{it} \,. \tag{28}$$

Indeed, this shadow value increases with marginal damages, scaled by both marginal utility of wealth λ^w_{it} and consumption $u'(\mathcal{D}^u(\tau)c_{it})$. This change in the marginal value of temperature affects directly the shadow value of adding carbon in the atmosphere according to the dynamics of λ^S_{it}:

$$\dot{\lambda}^S_{it} = \lambda^S_{it}(\rho - n_i - (1-\eta)\bar{g}_i + \delta_s) - \zeta\,\chi\,\Delta_i\,\lambda^\tau_{it}\ . \tag{29}$$

Emitting carbon in the atmosphere has a differential marginal impacts across regions due to heterogeneous costs of temperature and vulnerability to climate synthesized by the pattern scaling Δ_i and marginal damages $\gamma^y(\tau_{it} - \tau^\star_i)$ parameters. Solving the differential equations analytically, we can obtain the general formula – as found in Appendix B. When climate inertia is null $\zeta \to \infty$, this rewrites:

$$\lambda^S_{it} \xrightarrow[\zeta\to\infty]{} -\int_t^\infty e^{-(\rho - n_i - (1-\eta)\bar{g}_i + \delta_s)(s-t)}\chi\Delta_i(\tau_{is} - \tau^\star_i)\big[\gamma^y y_{is} + \gamma^u c_{is}\big]\lambda^w_{is}ds\ ,$$

where output is $y_{it} = z_i\mathcal{D}^y_i(\tau_{it})F(k_{it}, e_{it})$ and $\lambda^w_{it} = \mathcal{D}^u_i(\tau_{it})\,u'(\mathcal{D}^u_i(\tau_{it})c_{it})$. Using the Euler equation, or costate dynamics of equation (26), we get $\lambda^w_{it} = \lambda^w_{is}e^{-\int_t^s(\rho+\eta\bar{g}_i - r^\star_s)du}$ for $s > t$, which gives the Local Cost of Carbon:

$$LCC_{it} \to \int_t^\infty e^{-\delta_s(s-t) - \int_t^s(r^\star_u - n_i - \bar{g}_i)du}\chi\Delta_i(\tau_{is} - \tau^\star_i)\big[\gamma^y y_{is} + \gamma^u c_{is}\big]ds\ . \tag{30}$$

Note that the Local Cost Carbon is discounted with the interest rate $r^\star_t$ of the global bond market, and how it compares to growth of population n_i and TFP $\bar{g}_i$.

The Social Cost of Carbon is an aggregate measure that summarizes the global effects of climate change. It depends on the global welfare metrics and aggregates the Local Costs of Carbon from different regions. We introduce it in the next section.

4 Optimal policy

I consider the social planner that design the optimal climate policy to maximize global welfare, following three benchmarks: (i) the First-Best allocation, where the social planner has access to all the instruments, including cross-countries lump-sum transfers, (ii) the Second-Best policy where the social planner only chooses a single carbon tax t^ε_t, and (iii) the Second-Best allocation, choosing country-specific carbon taxes $\mathrm{t}^\varepsilon_{it}$. Despite the model being richer than Section 2's model, the theoretical results are analogous: I show how inequality and the lack of redistributive instruments change the path of the carbon tax.

4.1 First-Best

We consider the optimal policy of a social planner can choose the allocation decision of each country, subject to the resource constraints of the economy. They maximize global welfare, which is the weighted sum of household utilities, with Pareto weights[19] ω_i:

$$\mathcal{W}_{t_0} = \max_{\{c,k,e^f,e^c e^r,e^x,\bar{e}^c,\bar{e}^r\}} \sum_{\mathbb{I}} \mathcal{P}_i\, \omega_i \int_{t_0}^{\infty} e^{-(\rho - n_i - (1-\eta)\bar{g}_i)t}\, u\big(\mathcal{D}_i^u(\tau_{it})\, c_{it}\big)\; dt \tag{31}$$

subject to the good and energy resource constraints and the climate system:

$$\sum_{i\in\mathbb{I}} \mathcal{P}_i e^{(n_i+\bar{g}_i)t}\Big[c_{it} + (\dot{k}_{it}+(n_i+\bar{g}_i+\delta)k_{it}) + \nu_i^f(e_{it}^x, \mathcal{R}_{it}) + \kappa_i^c e_{it}^c + \kappa_{it}^r e_{it}^r\Big] = \sum_{i\in\mathbb{I}} \mathcal{P}_i e^{(n_i+\bar{g}_i)t} \mathcal{D}_i(\tau_{it}) z_{it} F(k_{it}, \ell_i, e_{it}) \qquad [\phi_t^w]$$

$$E_{it}^f = \sum_{i\in\mathbb{I}} \mathcal{P}_i e^{(n_i+\bar{g}_i)t}\, e_{it}^f = \sum_{i\in\mathbb{I}} e^{(n_i+\bar{g}_i)t} e_{it}^x \qquad [\mu_t^f] \qquad \bar{e}_i^c = e_i^c \qquad [\mu_{it}^c] \qquad \bar{e}_i^r = e_i^r \qquad [\mu_{it}^r]$$

$$\dot{\mathcal{S}}_t = \mathcal{E}_t - \delta_s \mathcal{S}_t \qquad \mathcal{E}_t := \sum_{\mathbb{I}} \mathcal{P}_i e^{(n_i+\bar{g}_i)t} (\xi^f e_{it}^f + \xi^c e_{it}^c) \qquad [\phi_t^S]$$

$$\dot{\tau}_{it} = \zeta\big(\Delta_i \chi \mathcal{S}_t - (\tau_{it} - \tau_{it_0})\big) \qquad [\phi_{it}^\tau]$$

I associate the Lagrange multipliers ϕ_t^w for the good market clearing, $\mu_t^f, \mu_{it}^c, \mu_{it}^r$ for the market clearing respectively of fossil (oil-gas), coal and renewable (low-carbon) and ϕ_t^S and ϕ_{it}^τ for the shadow value of additional carbon emissions and temperatures increase.

The result is similar to the toy model example, and the complete description is available in Appendix C. The optimality condition for consumption shows a redistribution motive: the planner equalizes marginal utility subject to Pareto weights:

$$\omega_i u_c(c_{it}, \tau_{it}) = \phi_t^w = \omega_j u_c(c_{jt}, \tau_{jt})$$

with marginal utility $u_c(c_{it}, \tau_{it}) = \mathcal{D}^u(\tau_{it}) u'(\mathcal{D}^u(\tau_{it}) c_{it}) = \mathcal{D}^u(\tau_{it})^{1-\eta} c_{it}^{-\eta}$, with the CRRA functional form. Despite the possibility, in the competitive equilibrium, to trade in goods, bonds, and energy, strong inequality exists due to differences in productivity, energy rents or climate damage. As a consequence, in the First-Best, the social planner redistributes consumption using lump-sum transfers in the decentralized equilibrium.

4.1.1 First-Best – Social Cost of Carbon

Given the welfare function in Appendix C, the marginal cost of adding one unit of carbon in the atmosphere $\mathcal{S}_t$ can be summarized by the Social Cost of Carbon. It represents the

$$SCC_t^{fb} := -\frac{\frac{\partial \mathcal{W}_t}{\partial \mathcal{S}_t}}{\frac{\partial \mathcal{W}_t}{\partial c_{it}}} = -\frac{\phi_t^S}{\phi_t^w} \tag{32}$$

where the aggregate welfare change ϕ_t^S is normalized in monetary units with the aggregate marginal value of wealth $\partial \mathcal{W}_t / \partial c_{it} = \omega_i u_c(c_{it}, \tau_{it}) = \phi_t$. The welfare cost of carbon evolves again with the marginal damage of temperature, ϕ_{it}^τ which is the costate of the temperature dynamics, and the marginal value of carbon ϕ_t^S, the costate of carbon concentration dynamics, which aggregates these

[19]Pareto weights should sum to one, s.t. $\frac{1}{\mathcal{P}} \sum_{\mathbb{I}} \mathcal{P}_i \omega_i = 1$

different costs across all countries. Solving the differential equations for $\phi^{\tau}_{it}, \phi^{S}_{t}$ and ϕ^{w}_{t}, following the same approach as in Section 3.8, we obtain the Social Cost of Carbon:

$$SCC^{fb}_t \to \int_t^\infty \sum_{i\in\mathbb{I}} \omega_i \mathcal{P}_i e^{-\delta_s(t-t_0)-\int_{t_0}^t (r^k_{iu}-n_i-\bar{g}_i)du} \chi\Delta_i(\tau_{is}-\tau^\star_i)[\gamma^y y_{is}+\gamma^u c_{is}]dt \ ,$$

and this implies the following proposition.

Proposition 6 (**Social Cost of Carbon – First-Best Allocation**).
In the First-Best allocation, marginal utilities are equalized $\phi_i = \omega_i u_c(c_{it},\tau_{it})$, $\forall i \in \mathbb{I}$, and as a result, we obtain that the Social Cost of Carbon is the sum of Local Costs of Carbon of the different locations $i \in \mathbb{I}$.

$$SCC^{fb}_t = \sum_{i\in\mathbb{I}} \omega_i \mathcal{P}_i LCC_{it}$$

where the Local Cost of Carbon LCC_{it} is defined as in equation (30), i.e.

$$LCC_{it} = \int_t^\infty e^{-\delta_s(s-t)-\int_t^s (r^k_u-n_i-\bar{g}_i)du}\chi\Delta_i(\tau_{is}-\tau^\star_i)[\gamma^y y_{is}+\gamma^u c_{is}]ds \ .$$

Despite the model being substantially richer, the main result of the Section 2 holds. This aligns with models without heterogeneity in marginal value of wealth, which can be aggregated. Models where differences in climate damages do not have redistributive effects on production, consumption and welfare would imply that the Social Cost of Carbon is simply the sum, weighted by population, $\mathcal{P}_i$, and Pareto weights ω_i, of the Local Cost of Carbon LCC_{it}. Such models include Nordhaus and Yang (1996), and Krusell and Smith (2022), where in the later the free mobility of capital and the absence of heterogeneity in TFP or production allows to aggregate the economy.

4.1.2 Decentralization of the First-Best: transfers and carbon tax

I show how this allocation can be decentralized with carbon taxation in the competitive equilibrium. The detailed treatment of this result is in Appendix A.

Proposition 7 (**First-Best Allocation – Carbon taxation and transfers**).
To decentralize the First-Best allocation, we need two set of instruments: (i) a carbon tax, (ii) cross-country lump-sum transfers. First, the optimal First-Best carbon tax equals the Social Cost of Carbon, as defined in Proposition 6:

$$MPe^f_{it} = q^f_t + \xi^f \mathrm{t}^\varepsilon_t \qquad\qquad \mathrm{t}^\varepsilon_t = SCC^{fb}_t$$

A similar optimality condition holds for coal $MPe^c_{it} = q^r_{it} + \xi^c \mathrm{t}^\varepsilon_t$. In particular, the optimal carbon tax is equal across countries. This results from the equalization marginal value of wealth. To achieve such equalization, the planner uses lump-sum transfers:

$$\omega_i u_c(c_{it},\tau_{it}) = \phi_t = \omega_j u_c(c_{jt},\tau_{jt}) \qquad \Rightarrow \qquad c^\star_{it} = u_c^{-1}(\phi_t|\tau_{it})$$

which pins down the lump-sum transfers needed given the budget constraint:

$$c^{\star}_{it} = (r^{\star}_t - (n_i + \bar{g}_i))w_{it} + \pi^f_{it} + y_{it} - (r^{\star} + \delta)k_{it} - (q^f_t + \xi^f \mathrm{t}^{\varepsilon}_t)e^f_{it} - (q^c_{it} + \xi^c \mathrm{t}^{\varepsilon}_t)e^c_{it} - q^r_{it}e^r_{it} - \dot{w}_{it} + \mathrm{t}^{ls}_{it} \ ,$$

A uniform carbon tax result aligns with the standard policy recommendations in representative agent models – which encompass most Integrated Assessment Models, like DICE, Nordhaus (2017), Barrage and Nordhaus (2024), FUND, PAGE, MERGE, and others – or the optimal carbon taxation result of Golosov, Hassler, Krusell and Tsyvinski (2014). It also reminiscent of models with unrestricted redistribution such as Hillebrand and Hillebrand (2019).

Lump-sum transfers redistribute across countries and across time:

$$\int_{t_0}^{\infty} \mathcal{P}_i e^{(n_i + \bar{g}_i)t} \sum_{\mathbb{I}} \mathrm{t}^{ls}_{it} \, dt = 0$$

Situations where income y_{it}, energy rents π^f_{it}, climate damage and temperature τ_{it}, or Pareto weights ω_i are very heterogeneous such that consumption differentials in the equilibrium without policy intervention are large, the First-Best implies that some countries receive positive lump-sum transfers $\exists j, s.t.\ \mathrm{t}^{ls}_j > 0$ and others pay lump-sum taxes $\exists j', s.t.\ \mathrm{t}^{ls}_{j'} < 0$. Therefore, such decentralized allocation features direct lump-sum transfers across countries.

The question is whether such lump-sum transfers are politically feasible. Can a world central planner impose lump-sum transfers to solve world inequality, for example taxing North America and Europe and rebating it lump-sum to Africa or South Asia? In the next sections, I analyze a family of policies where transfers are not allowed.

4.2 Ramsey problem and optimal uniform carbon taxation

I consider the optimal carbon taxation, where the planner is prevented from achieving redistribution. Lump-sum transfers are not available instruments, for political, governance, or economic reasons. This implies to solve a Ramsey taxation problem with imperfect instruments, where the planner internalize the constraints that are impose by the Competitive Equilibrium.

In particular, it maximizes global welfare, choosing a uniform carbon tax for the world $\{\mathrm{t}^{\varepsilon}_t\}_t$ and rebates the revenue of that tax to the household of the country that pays it $\mathrm{t}^{ls} = \mathrm{t}^{\varepsilon}_t \varepsilon_{it}$.

$$\mathcal{W}_{t_0} = \max_{\substack{\{c,b,k,e^f,e^c e^r,e^x,\bar{e}^c,\bar{e}^r\} \\ \{\mathrm{t}^{\varepsilon},r,q^f,q^c,q^r\}}} \sum_{\mathbb{I}} \mathcal{P}_i \, \omega_i \int_{t_0}^{\infty} e^{-(\rho+n_i)t} \, u\big(\mathcal{D}^u_i(\tau_{it})\, c_{it}\big) \, dt \ . \tag{33}$$

The planner takes the agent decisions, $\{c_{it}, b_{it}, k_{it}, e^f_{it}, e^c_{it}, e^r_{it}, e^x_{it}, \bar{e}^c_{it}, \bar{e}^r_{it}\}_{it}$, prices $\{r^{\star}_t, q^f_t, \mathrm{w}_{it}, q^c_{it}, q^r_{it}\}_{it}$, and states $\{w_{it}, \tau_{it}, \mathcal{S}_t\}_{it}$ subject to the constraints of the Competitive equilibrium, which are (i) the budget constraint equation (25), with multiplier ψ^w_{it}, (ii) the climate dynamics for carbon $\mathcal{S}_t$, equation (21), with multiplier ψ^S_t, and temperatures τ_{it}, equation (22), with multiplier ψ^{τ}_{it}, (iii) the firms optimality conditions for energy $e^f_{it}, e^c_{it}, e^r_{it}$, from equation (11), and capital k_{it}, respectively with multipliers $\upsilon^f_{it}, \upsilon^c_{it}, \upsilon^r_{it}, \upsilon^k_{it}$, (iv) the energy firms optimality for extraction and production, from equations (14), (17) and (18), and finally (v) the market clearing for goods equation (24), bonds, fossil, coal and renewable energy, from equations (16) to (18).

We apply the Pontryagin Maximum Principle and the details of the entire system are found in Appendix D. Note that the social planner has full commitment, in the sense that decisions taken in the initial period t_0 binds $\forall t \in (t_0, \infty)$ and there is no time inconsistency. We provide some intuitions of the most important results.

First, before of lack of redistribution, we can define the normalized social welfare weights – as some inequality index – using the marginal value of wealth ψ^w_{it}:

$$\begin{aligned} &\psi^w_{it} = u_c(c_i, \tau_{it}) \qquad\qquad \Rightarrow \qquad \omega_i u_c(c_{it}, \tau_{it}) \neq \omega_j u_c(c_{jt}, \tau_{jt}) \\ &\widehat{\psi}^w_{it} := \frac{\omega_i \mathcal{P}_i \psi^w_{it}}{\overline{\psi}^w_t} = \frac{\omega_i \mathcal{P}_i u_c(c_{it}, \tau_{it})}{\frac{1}{\mathcal{P}} \sum_{i \in \mathbb{I}} \omega_i \mathcal{P}_i u_c(c_{it}, \tau_{it})} \qquad\qquad \overline{\psi}^w_t = \frac{1}{\mathcal{P}} \sum_{\mathbb{I}} \omega_i \mathcal{P}_i \psi^w_{it} \ , \end{aligned} \tag{34}$$

with $\overline{\psi}^w_t$ the marginal value of wealth of the "average" agent. If the ratio $\widehat{\psi}^w_{it}$ is higher than 1, the planner sees country i at time t with a lower welfare than the average household.

4.2.1 Second-Best, Social Cost of Carbon

Again, as in the previous section, the Social Cost of Carbon is measured as the ratio of the marginal value of carbon and the marginal value of wealth:

$$SCC^{sb}_t := -\frac{\frac{\partial \mathcal{W}_t}{\partial \mathcal{S}_t}}{\frac{\partial \mathcal{W}_t}{\partial c_t}} = -\frac{\psi^S_t}{\overline{\psi}^w_t} \tag{35}$$

it differs slightly from the definition in equation (32) of the First-Best, because now the planner consider the "average agent" marginal utility $\overline{\psi}^w_t$ to convert the Social Cost of Carbon in monetary unit.

To obtain the Social Cost of Carbon, we again solve the differential equation for the marginal damage to temperature and the marginal value of carbon.

$$SCC^{sb}_t = \int_t^\infty e^{-\delta_s(s-t) - \int_t^s r^\star_u du} \chi \sum_{i \in \mathbb{I}} e^{(n_i + \bar{g}_i)(s-t)} \Delta_i(\tau_{is} - \tau^\star_i)\left[\gamma^y y_{is} + \gamma^u c_{is}\right] \widehat{\psi}^w_{is} ds$$

This now differs from the First-Best because the temperature damages $(\tau_{it} - \tau^\star_i)\left[\gamma^y y_{it} + \gamma^u c_{it}\right]$ are weighted by welfare ψ^w_{it}. We thus obtain an important result in Second-Best economies:

Proposition 8 (Social Cost of Carbon – Second-Best Allocations).
In the Second-Best allocation, inequalities across countries persist as measured by the social welfare weights $\widehat{\psi}^w_{it} \neq 1, \forall i \in \mathbb{I}$. When redistribution is constrained, the Social Cost of Carbon is the weighted sum of Local Costs of Carbon of the different locations $i \in \mathbb{I}$, weighted by the social welfare weights $\widehat{\psi}^w_{it}$ as given in equation (34):

$$SCC^{sb}_t = \sum_{i \in \mathbb{I}} \widehat{\psi}^w_{it} \, LCC_{it} \quad \propto \quad \sum_{i \in \mathbb{I}} \mathcal{P}_i \, \omega_i \, u_c(c_{it}, \tau_{it}) \, LCC_{it}$$

where the Local Cost of Carbon LCC_{it} is defined as in equation (30), i.e.

$$LCC_{it} = \int_t^{\infty} e^{-\delta_s(s-t)-\int_t^s (r_u^{\star}-n_i-\bar{g}_i)du} \chi \Delta_i(\tau_{is}-\tau_i^{\star})\left[\gamma^y y_{is} + \gamma^u c_{is}\right] ds \ .$$

As a result, we can express the Social Cost of Carbon as:

$$\begin{aligned} SCC_t^{sb} &= \sum_{\mathbb{I}} \widehat{\psi}_{it}^w \, LCC_{it} \\ &= \mathcal{P}\,\mathbb{E}^{\mathbb{I}}\left[\omega_i LCC_{it}\right] + \mathcal{P}\,\mathbb{C}\text{ov}^{\mathbb{I}}\Big(\widehat{\psi}_{it}^w, LCC_{it}\Big) \qquad \lesseqgtr \quad \mathcal{P}\,\mathbb{E}^{\mathbb{I}}\left[\omega_i LCC_{it}\right] =: SCC_t^{fb} \end{aligned}$$

for the mean $\mathbb{E}^{\mathbb{I}}[\cdot]$ and covariance $\mathbb{C}ov^{\mathbb{I}}(\cdot)$ over locations.[20] *Since the Social Cost of Carbon is a sum – and not a mean – one needs to rescale by world population $\mathcal{P}$.*

To summarize, the presence of heterogeneity and the correlation between local damage and income change the Social Cost of Carbon from the Social Planner's perspective.

4.2.2 Second-Best, Uniform Carbon Taxation

As shown in model of Section 2, taxation of carbon and fossil fuels have strong redistributive general equilibrium effects through energy markets.

Fossil energy supply redistribution. First, the implementation of carbon taxation reduces demand for fossil fuels, which has strong redistributive effects on the energy rent and along the supply curve. Recall that υ_{it}^f is the Lagrange multiplier for the optimality condition on fossil fuel demand, and we denote $\widehat{\upsilon}_{it}^f = \upsilon_{it}^f / \overline{\psi}_{it}^w$ the rescaled value by the marginal value of wealth – it would represent the monetary value of marginally relaxing that optimality condition. Moreover, a change in equilibrium quantity of oil-gas relaxes the market clearing, leading to the following mechanism:

$$\textit{Supply Redistribution}_t^f = \Big(\sum_i \nu_{i\,e^x e^x(e_{it}^x,\mathcal{R}_{it})}^{f\ -1}\Big)^{-1}\Big(\sum_i \widehat{\psi}_{it}^w[e_{it}^f - e_{it}^x] - \sum_i \widehat{\upsilon}_{it}^f\Big) \ ,$$

where $\Big(\sum_i \nu_{i\,e^x e^x(e_{it}^x,\mathcal{R}_{it})}^{f\ -1}\Big)^{-1}$ is the aggregate supply elasticity for oil and gas, which depends on the equilbrium extraction e_{it}^x and reserves $\mathcal{R}_{it}$ of each country. Moreover, as we saw above, carbon taxation lowers the energy price q_t^f, leading to a term-of-trade effect. This redistributes energy rents from exporters to importers, and hence scale with $e_{it}^f - e_{it}^x$. This change is weighted by the marginal valuation of wealth $\widehat{\psi}_{it}^w$, and it is higher when the fossil-fuel producers are also relatively poorer or weighted more by the planner. This effect rewrites $\sum_i \widehat{\psi}_{it}^w[e_{it}^x - e_{it}^f] = \mathbb{C}\text{ov}^{\mathbb{I}}(\widehat{\psi}_{it}^w(e_{it}^x - e_{it}^f))$, and we can measure this covariance empirically as a sufficient statistics. Moreover, the change in energy price also affects the oil and gas price, affecting the demand, as denoted by the term $\sum_i \widehat{\upsilon}_{it}^f$. This term comes from the fact that carbon polity is affecting differentially oil-gas and coal. Recall that, if $\widehat{\upsilon}_{it}^f > 0$, the planner would increase energy prices, moving up along the supply curve, which results in a lower the optimal carbon tax $\mathrm{t}_t^{\varepsilon}$.

[20] We define them as $\mathbb{E}^{\mathbb{I}}[x_{it}] = \frac{1}{\mathcal{P}}\sum_{i\in\mathbb{I}} \mathcal{P}_i x_{it}$ and $\mathbb{C}\text{ov}^{\mathbb{I}}(x_{it}, y_{it}) = \frac{1}{\mathcal{P}}\sum_{i\in\mathbb{I}} \mathcal{P}_i\big(x_{it} - \mathbb{E}^{\mathbb{I}}[x_{it}]\big)\ \big(y_{it} - \mathbb{E}^{\mathbb{I}}[y_{it}]\big)it.$

Fossil energy demand distortion. Second, carbon taxation is a distortionary tax that create a wedge for the demand of energy. This demand distortion differs across countries due to differences in energy mix, productivity and the endowmnent and cost of energy sources. From the planner's optimality for energy, we can define the following term:

$$\begin{aligned}
\textit{Demand Distortion}^f_{it} := \widehat{\upsilon}^f_{it} &= \left[\widehat{\upsilon}^f_{it}\,\partial_{e^f} MPe^f_{it} + \widehat{\upsilon}^c_{it}\,\partial_{e^f} MPe^c_{it} + \upsilon^r_{it}\,\partial_{e^f} MPe^r_{it} + \widehat{\upsilon}^k_{it}\,\partial_{e^f} MPk_{it}\right] \\
&= z_i \mathcal{D}_i(\tau_{it})\Big[\widehat{\upsilon}^f_{it}\,F_{e^f e^f} + \widehat{\upsilon}^c_{it}\,F_{e^f e^c} + \widehat{\upsilon}^r_{it}\,F_{e^f e^r} + \widehat{\upsilon}^k_{it}\,F_{e^f k}\Big] \\
&= \frac{1}{e^f_{it}}\Big[-\widehat{\upsilon}^f_{it}(q^f_t + \xi^f \mathrm{t}^\varepsilon_t)\Big(\frac{1-s^f_{it}}{\sigma^e} + s^f_{it}\frac{1-s^e_{it}}{\sigma^y}\Big) + \widehat{\upsilon}^c_{it}(q^c_{it} + \xi^c \mathrm{t}^\varepsilon_t)s^f_{it}\Big(\frac{1}{\sigma^e} - \frac{1-s^e_{it}}{\sigma^y}\Big) \\
&\qquad + \widehat{\upsilon}^r_{it} q^r_{it} s^f_{it}\Big(\frac{1}{\sigma^e} - \frac{1-s^e_{it}}{\sigma^y}\Big) + \widehat{\upsilon}^k_{it}(r^\star_t + \delta)\frac{s^f_{it}s^e_{it}}{\sigma^y}\Big],
\end{aligned}$$

with energy share in production, $s^e_{it} = e_{it}q^e_i/y_i$, fossil share in energy mix $s^f_{it} = e^f_{it}q^f_t/e_{it}q^e_{it}$ and similarly $s^c_{it} = e^c_{it}q^c_{it}/e_{it}q^e_{it}$ and $s^r_{it} = e^r_{it}q^r_{it}/e_{it}q^e_{it}$. Moreover, σ^e is the elasticity of substitution between energy sources, σ^y the one between energy and the capital/labor bundle, and $\widehat{\upsilon}^f_{it} := \omega_i \mathcal{P}_i \upsilon^f_{it}/\overline{\psi}^w_{it}, \widehat{\upsilon}^r_{it}$ and $\widehat{\upsilon}^k_{it}$ are rescaled versions of the Lagrange multipliers for fossil energy, renewables, and capital respectively. Note, we can find a similar expression for coal energy, *Demand Distortion*$^c_{it}$, as function of $F_{e^c e^f}, F_{e^c e^c}$, etc.

To understand the intuition behind this term, take the first term, $-\frac{1-s^f_{it}}{\sigma^e} - s^f_{it}\frac{1-s^e_{it}}{\sigma^y}$, as an example. We see that this demand channel of taxation has two effects: the first channel lowers fossil consumption due to direct substitution effect between the three energy inputs, lowering the marginal product of fossil MPe^f_{it} with elasticity σ^e. The second effect is indirect through the total energy use e_{it}, proportionally to the fossil share s^f_{it}. Weighting these different distortions with the shadow values υ^x_{it} for input x and scaling it for the input prices q^f_t, q^c_{it}, etc. we obtain the total distortion caused by taxation of fossil fuels. This term is more involved than in the simpler model of Section 2 and Proposition 2, because of the general substitution patterns across energy and heterogeneity across countries. Moreover, the aggregate level of the carbon tax balances out all these distortions across countries and energies, and the optimality condition for t^ε_t gives:

$$\sum_{i\in\mathbb{I}} (\xi^f \widehat{\upsilon}^f_{it} \; + \; \xi^c \widehat{\upsilon}^c_{it}) = 0$$

This minimization of total distortion relates to the standard principles of Ramsey taxation, as in Diamond and Mirrlees (1971); Diamond (1973); Atkinson and Stiglitz (1976).

Optimal carbon tax. We now present our main result for the optimal Second-Best uniform carbon tax. Details on how this formula is derived from the optimality condition for energy choices can be found in Appendix D. The optimal level of carbon taxation integrates the different redistribution motives that we detailed above. As a result, the Ramsey planner accounts for these general equilibrium effects as symbolized by the curvature of demand and supply of energy.

Proposition 9 (**Uniform Carbon Taxation – Second-Best Allocation**).
The optimal Second-Best carbon tax accounts for three distributional motives when setting a single uniform level t^{ε}*, in the absence of cross-countries transfers, when revenues of the carbon are rebated locally* $\mathrm{t}^{ls}_{it} = \epsilon_{it}\mathrm{t}^{\varepsilon}_t$*: (i) climate damage in the* ***Social Cost of Carbon (SCC)****, (ii)* ***Supply Redistribution*** *in energy markets through terms-of-trade and energy rents and (iii)* ***Demand Distortion*** *through distorted firms' energy choices. This includes redistribution motives due to the presence of inequality through the* ***social welfare weights*** $\widehat{\psi}^w_{it} = \omega_i \mathcal{P}_i \psi^w_{it}/\overline{\psi}^w_{it} \propto \omega_i \mathcal{P}_i u_c(c_{it}, \tau_{it})$*. The optimal carbon tax writes:*

$$\xi^f \mathrm{t}^{\varepsilon}_t = \xi^f\, SCC^{sb}_t \,+\, \textit{Supply Redistribution}^f_t \,+\, \sum_{i\in\mathbb{I}} \textit{Demand Distortion}^f_{it}$$

$$\begin{aligned} SCC_t &= \mathcal{P}\,\mathbb{E}^{\mathbb{I}}\big[\omega_i LCC_{it}\big] + \mathcal{P}\,\mathbb{C}\mathrm{ov}^{\mathbb{I}}\Big(\widehat{\psi}^w_{it}, LCC_{it}\Big) \\ \textit{Supply Redistribution}^f_t &= \Big(\sum_i (\nu^f_{i\,e^x e^x})^{-1}\Big)^{-1}\Big[\mathcal{P}\mathbb{C}\mathrm{ov}^{\mathbb{I}}\big(\widehat{\psi}^w_{it}, e^x_{it} - e^f_{it}\big) - \mathcal{P}\mathbb{E}^{\mathbb{I}}\big[\widehat{\upsilon}^f_{it}\big]\Big] \\ \textit{Demand Distortion}^f_{it} &= \sum_{x\in\{e^f,e^c,e^r,k\}} \widehat{\upsilon}^x_{it} \quad \partial_{e^f} MPx_{it} \end{aligned} \tag{36}$$

where $\widehat{\upsilon}^x_{it}$ *are the rescaled multipliers for the optimality condition* $MPx_{it} = q^x_{it}$ *for the choice of input* x*, i.e.* $\widehat{\upsilon}^f_{it} := \omega_i \mathcal{P}_i \upsilon^x_{it}/\overline{\psi}^w_{it}$*, for* x *being fossil (oil-gas)* e^f*, coal* e^c*, renewable (low-carbon)* e^r *and capital* k*. Moreover, the supply redistribution depends on the curvature of the oil-gas supply with* $\nu^f_{i\,e^x e^x}$ *the supply elasticity, and the demand distortion depends on* $\partial_{e^f} MPx_{it}$ *the curvature of the production function, i.e. the own-elasticity in oil-gas* $F_{e^f e^f}$ *and cross-elasticity, e.g.* $F_{e^f e^r}$*.*

In addition, even without climate externality $SCC_t = 0$, the carbon tax $\mathrm{t}^{\varepsilon}_t$ could be positive, accounting for energy terms-of-trade manipulations, for example with wealthy exporters and relatively poorer importers, or when richer countries are consuming more fossil-fuels with higher elasticity, following the intuitions of Section 2.

Such a result holds as long as different agents – countries, firms, or households – have different marginal utilities of consumption, i.e. different $\widehat{\psi}^w_{it}$. However, these motives would be absent in models like Golosov et al. (2014) for two reasons: First if the supply curve for energy is perfectly elastic, because of constant return to scale, which yields $\nu_{e^x e^x} = 0$. Second, when the agent/firm is "representative" or multiple agents can be aggregated – at least in the inputs demand decisions – and a single energy tax instrument in the First-Best, the social planner is not "distorting" the energy demand across agents: the planner and the agents would achieve the same optimality condition for fossil fuel demand.

4.3 Country-specific carbon taxation

We now consider an experiment where the social planner design country-specific taxes that would allow to correct some of these redistributive concerns. In that case, not only the *level* but also the *distribution* of the fossil fuel/carbon tax is affected by redistribution motives. The planner maximizes global welfare as in equation (33), choosing a countries-specific carbon taxes $\{\mathrm{t}^{\varepsilon}_{it}\}_t$ over time and rebates the carbon tax revenue to the household of the country that pays it $\mathrm{t}^{ls} = \mathrm{t}^{\varepsilon}_{t}\varepsilon_{it}$. We solve the planner's problem choosing the agent decisions, prices, and states, subject to the constraints of the Competitive equilibrium.

Proposition 10 (**Country-specific Carbon Taxation – Second-Best Allocation**).
The optimal Second-Best country-specific carbon taxation, when transfers are absent and revenues rebated locally, can be written as:

$$\mathrm{t}^{\varepsilon}_{it} = \frac{1}{\widehat{\psi}^{w}_{it}}\Big[SCC_t \;+\; \textit{Supply Redistribution}^{f}_{t} \;+\; \textit{Demand Distortion}^{f}_{it}\Big]$$

$$SCC_t = \mathcal{P}\,\mathbb{E}^{\mathbb{I}}\big[\omega_i LCC_{it}\big] + \mathcal{P}\,\mathbb{C}\mathrm{ov}^{\mathbb{I}}\Big(\widehat{\psi}^{w}_{it}, LCC_{it}\Big)$$

$$\textit{Supply Redistribution}^{f}_{t} = \Big(\sum_i (\nu^{f}_{i\,e^x e^x})^{-1}\Big)^{-1}\Big[\mathcal{P}\mathbb{C}\mathrm{ov}^{\mathbb{I}}(\widehat{\psi}^{w}_{it}, e^{x}_{it} - e^{f}_{it}) - \mathcal{P}\mathbb{E}^{\mathbb{I}}[\widehat{v}^{f}_{it}]\Big]$$

$$\textit{Demand Distortion}^{f}_{it} = \sum_{x\in\{e^f,e^c,e^r,k\}} \widehat{v}^{x}_{it} \quad \partial_{e^f} MPx_{it}$$

*The demand distortion is only local since the planner can choose country-specific taxes. However, this distortion is not null, since the planner does not have **energy-sources**-specific taxes. In particular, the carbon tax affects both oil-gas and coal,*

$$\xi^{f}\widehat{v}^{f}_{it} \;+\; \xi^{c}\widehat{v}^{c}_{it} = 0$$

These local distortive effects remain if energy demands e^{f}_{it}, e^{c}_{it}, energy shares, and elasticities differ.

Following the logic in Section 2, the optimal tax still depend on both the Social Cost of Carbon SCC_t and *Supply Redistribution*$^{f}_{t}$ as developed above. These two energy taxation motives – climate externality correction and energy rents redistribution – remains the same. However, we observe that the demand distortion is reduced considerably. We observe that because the planner has access to n different instruments

At the difference with Section 2, we now have only the local distortions, represented by the Lagrange multipliers $\widehat{v}^{f}_{it}, \widehat{v}^{c}_{it}, \widehat{v}^{r}_{it}, \widehat{v}^{k}_{it}$ for the input choices. These inputs distortion are no longer null in equilibrium as in the Proposition 2, despite the planner choosing a country-specific tax level. The reason is that now there is a distortion and a reallocation between oil-gas, coal, and other inputs. The optimality for $\mathrm{t}^{\varepsilon}_{i}$ gives:

$$\xi^{f}\widehat{v}^{f}_{it} \;+\; \xi^{c}\widehat{v}^{c}_{it} = 0 \qquad\qquad \not\Rightarrow \qquad \widehat{v}^{c}_{it} = 0 \qquad \& \qquad \textit{Demand Distortion}^{f}_{it} = 0$$

Nevertheless, the tax is country i specific and depends on redistribution motives. Indeed the ratio $1/\widehat{\psi}^{w}_{it}$ is the inverse of the social welfare weight – the inequality index developed earlier in this section in equation (34). It implies that richer/colder countries, which have higher consumption and lower marginal utilities are charged higher carbon taxes, and conversely poorer countries should be charged a lower tax:

$$\text{low} \quad c_{it} \quad \text{high} \quad \tau_{it} \qquad \Rightarrow \qquad \text{high} \quad \widehat{\psi}^{w}_{it} \propto u_c(c_{it}, \tau_{it}) \qquad \Rightarrow \qquad \text{low} \quad \mathrm{t}^{f}_{it}$$

everything else being constant, in particular SCC_t, *Supply Redistribution*^f_t and *Demand Distortion*$^f_{it}$.

In particular, this result provides a justification for the use of a "tiered" carbon taxation – where carbon tax would be lower for developing economies and higher for advanced economies. This proposal, that tend to be promoted for moral reasons, is shown here to have an explanation based on efficiency and welfare maximization.

5 Quantification and calibration

The model is calibrated to a sample of 68 countries to provide realistic predictions on the impact of optimal carbon policy. I first describe the data used. I then provide details on the quantification, and how the parameters are calibrated to match the data. I summarize in Table 2 the dimensions of heterogeneity of the model. Table 1 contains the summary table for the calibration described in this section.

5.1 Data

First, I describe briefly the data used to calibrate the model. I use data for the year 2018-2023, taking the average over that period to smooth out the effect of the COVID-19 recession on energy and macroeconomic data.

I use data for GDP per capita, in Purchasing Power Parity (PPP, in 2016 USD) from the World Bank, as collected and processed by the Maddison Project, Bolt and van Zanden (2023). For the energy variables, I use the comprehensive data collected and processed in the Statistical Review of Energy Energy Institute (2024), that includes the production and consumption of various energy sources, including Oil, Gas, Coal and Renewable. It also includes proven reserves of those fossil fuels. For energy rent, I use the World Development Indicators that use national accounts to measure the share of GDP coming from energy (oil, gas and coal) and natural resource rents. Finally, for temperature, I use the same time series as Burke et al. (2015), which use the temperature at country level, averaged over the year and weighted by population across locations.

5.2 Welfare and Pareto weights

The welfare function that the social planner maximize, in Appendix C is the weighted sum of individual utilities in all countries, with $\mathcal{P}_i$ the population size and ω_i the Pareto weights.

I consider two sets of Pareto weights. First, I consider the utilitarian benchmark, where the planner weight every individual in the world equally: $\omega_i = 1$. Second, following the discussion in Anthoff et al. (2009), Nordhaus (2011) and Nordhaus and Yang (1996), one would like to choose Pareto weights that eliminate redistributive effects that are orthogonal to climate change and carbon policy. To that purpose, the "Negishi" Pareto weights make the preexisting competitive equilibrium efficient under that welfare metric. This implies that:

$$\omega_i = \frac{1}{u_c(\bar{c}_{it_0}, \tau_{it_0})} \qquad \Leftrightarrow \qquad C.E.(\bar{c}_i) \in \underset{\bar{c}_i}{\operatorname{argmax}} \sum_i \mathcal{P}_i \omega_i u(\bar{c}_{it_0}, \tau_{it_0})$$

$$\omega_i u_c(\bar{c}_{it_0}, \tau_{it_0}) = \omega_j u_c(\bar{c}_{jt_0}, \tau_{jt_0}) \qquad \forall i, j \ \in \mathbb{I}$$

where $\bar{c}_i$ is the consumption level in the present competitive equilibrium – the period 2018-2023 – absent future climate damage. This implies that the carbon policy do not have redistributive motives through energy general equilibrium effects. However, global warming, and the carbon taxation itself have redistributive effects, as they change the distribution of c_{it}. These effects would thus be taken into account in the choice of policies as shown in Proposition 9.

5.3 Macroeconomy, trade and production

For the macroeconomic part of the framework, I consider standard utility and production functions. For utility, as in the equation (7), I calibrate the CRRA/IES parameter to be $\eta = 1.5$, taken from Barrage and Nordhaus (2024).

For production, I use a nested CES framework. The firm combines a Cobb-Douglas bundle of capital k_i and labor ℓ_i[21] with a composite of energy e_i, with elasticity σ^y. Second, the energy e_i aggregates the different energy sources: oil and gas e^f, coal e_i^c, and renewable/non-carbon e_i^r, with elasticity σ^e. To calibrate these functions, I set the capital-labor ratio $\alpha = 0.35$ to match the cost share of capital. For the energy, I set $\varepsilon = 0.10$ to match the world average energy cost share $\frac{q_i^e e_i}{y_i} = 6\%$, as measured in Kotlikoff, Kubler, Polbin and Scheidegger (2021*b*) and used in Krusell and Smith (2022). For the elasticity between energy and other inputs, I set $\sigma^y = 0.3$ for all countries, which is in the range of estimates in Papageorgiou et al. (2017), among others.[22] Therefore, capital/labor and energy are complementary in production: an increase in the price of energy has a strong impact on output as it is less productive to "substitute away" to other inputs – capital, labor here. This aligns with other empirical and structural evidence on the impact of energy shocks, e.g. Hassler et al. (2021*a*). Then, I calibrate the energy mix for oil-gas, with $\omega^f = 0.56$, coal $\omega^c = 0.27$, and non-carbon $\omega^r = 0.17$, to match the aggregate shares of each of these energy sources in the data. In the next section, I document how I match the individual countries' energy mix using energy prices/costs. Finally, for the elasticity between energy inputs, I use the value $\sigma_e = 2$, following the rest of the literature, i.e. Papageorgiou et al. (2017), Kotlikoff, Kubler, Polbin and Scheidegger (2021*b*), and Hillebrand and Hillebrand (2019), among others.

I calibrate the productivity z_i of the production function $y_{it_0} = \mathcal{D}_i^y(\tau_{it_0})\, z_i\, \bar{y}_{it_0}$ to match exactly the GDP, y_{it_0}, across countries. This parameter z_i, represents productivity residuals as well as institutional/efficiency differences across countries. In Figure 3, I show the GDP levels, as they replicated with this model.

[21] Labor is inelastically supplied $\ell_i = \bar{\ell}_i$ in each country and normalized to 1 – since the country size $\mathcal{P}_i$ is already taken into account. As a result, all the variables can be seen as input per capita.

[22] It also aligns with my own estimation in Bourany (2022).

Table 1: Baseline calibration

Technology & Energy markets			
α	0.35	Capital share in $F(\cdot)$	Capital/Output ratio
ϵ	0.12	Energy share in $F(\cdot)$	Energy cost share (8.5%)
σ^y	0.3	Elasticity capital-labor vs. energy	Complementarity in production (c.f. Bourany 2022)
ω^f	0.56	Fossil energy share in $e(\cdot)$	Oil-gas/Energy ratio
ω^c	0.27	Coal energy share in $e(\cdot)$	Coal/Energy ratio
ω^r	0.17	Non-carbon energy share in $e(\cdot)$	Non-carbon/Energy ratio
σ^e	2.0	Elasticity fossil-coal-non-carbon	Slight substitutability & Study by Stern
δ	0.06	Depreciation rate	Investment/Output ratio
$\bar{g}$	$0.01^\star$	Long run TFP growth	Conservative estimate for growth
Preferences & Time horizon			
ρ	0.03	HH Discount factor	Long term interest rate & usual calib. in IAMs
η	1.5	IES / Risk aversion	Standard calibration
n	0.0035	Long run population growth	Conservative estimate for growth
ω_i	1	Pareto weights	Uniforms / Utilitarian Social Planner
ω_i	$1/u'(c_i)$	Pareto weights	Negishi / Status-quo Social Planner
T	400	Time horizon	Time for climate system to stabilize
Climate parameters			
ξ^f	2.761	Emission factor – Oil & natural gas	Conversion 1 $MTOE$ $\Rightarrow$ 1 $MT\ CO_2$
ξ^c	3.961	Emission factor – Coal	Conversion 1 $MTOE$ $\Rightarrow$ 1 $MT\ CO_2$
χ	$2.3/1e6$	Climate sensitivity	Pulse experiment: $100\,GtC \equiv 0.23°C$ medium-term warming
δ_s	0.0004	Carbon exit from atmosphere	Pulse experiment: $100\,GtC \equiv 0.15°C$ long-term warming
ζ	0.027	Growth rate, Carbon Capture and Storage	Starting after 2100, Follows Krusell Smith (2022)
$\gamma^\oplus$	0.003406	Damage sensitivity	Nordhaus' DICE
$\gamma^\ominus$	$0.3\times\gamma^\oplus$	Damage sensitivity	Nordhaus' DICE & Rudik et al (2022)
α^T	0.5	Weight historical climate for optimal temp.	Marginal damage correlated with initial temp.
$T^\star$	14.5	Optimal yearly temperature	Average spring temperature / Developed economies

5.4 Energy markets

For the energy market, I match the energy mix of different countries, using the CES framework displayed above, as well as differences in costs of production. For the supply side, we use iso-elastic fossil extraction cost to replicate the oil-gas supply of fossil producers.

First, in this model, oil and gas are traded on international markets, with demand $\mathcal{P}_i e^f_{it_0}$ from the final good firm and supply $\mathcal{P}_i e^x_{it_0}$ from the fossil energy firm, extracting oil and gas from its own reserves. We use the extraction function ν^f_i to have the following isoelastic form of equation (13), which is homogeneous of degree one in $(e^x_i, \mathcal{R}_i)$. The inputs are paid in the price of the consumption good – normalized to one.[23] This implies the profit function in equation (15). I calibrate the three parameters $\mathcal{R}_i$, ν_i and $\bar{\nu}_i$ to match the three country-level variables $\mathcal{R}_i$, e^x_i and π^f_i. The reserves $\mathcal{R}_i$ are taken directly from the data on oil and gas reserves documented by Energy Institute (2024). I calibrate the slope of this cost function $\bar{\nu}_i$ to match exactly the production of oil and gas e^x_i, as informed by that same data source. This is displayed in Figure 4. I then calibrate the curvature of the cost function to match the share $\eta^\pi_i = \frac{\pi^f_i}{y_i \mathrm{p}_i + \pi^f_i}$ of fossil energy profit as share of GDP. I choose

[23] I express the oil-gas extraction with a cost function $x^f_i = \nu^f_i(\cdot)$. We can also express analogously with the production function:

$$e^x_i = g(x^f_i) = \left(\tfrac{1+\nu_i}{\bar{\nu}_i}\right)^{\frac{1}{1+\nu}} \mathcal{R}_i^{\frac{\nu_i}{1+\nu_i}} (x^f_i)^{\frac{1}{1+\nu_i}}$$

with inputs x^f_i paid in the final good price. This production has constant returns to scale in $(x^f_i, \mathcal{R}_i)$.

ν_i to minimize the distance – mean squared error – between the model share η_i^π and the data, successfully matching the share within 5−10 percentage points. Differences in oil and gas energy rent across countries are not only determined by differences in cost and technology, but also in differences in trade costs and market power – by the existence of OPEC which control more than 28% of oil supply and around 15% of natural gas supply. This explains why it is difficult to match exactly the value η_i^π. However, to keep the simplicity and tractability of the model, I refrain from adding an additional Armington structure over energy sources, or oligopoly power over oil and gas as discussed in Bornstein et al. (2023) and Hassler et al. (2010).

Second, I match the energy mix of the different countries by relying on the two assumptions made in the model: (i) coal and renewable are only traded at the country level: $\bar{e}_i^c = e_i^c$ and $\bar{e}_i^r = e_i^r$ and (ii) the cost function is linear in goods, i.e. the production is Constant Returns to Scale, implying $q_i^c = \kappa_i^c$ and $q_i^r = \kappa_i^r$. This allows me to match the energy mix of each country by calibrating the energy costs parameters κ_i^c and κ_i^r for each country to match the data on coal share $\frac{e_i^c}{e_i^f+e_i^c+e_i^r}$ and non-carbon energy share $\frac{e_i^r}{e_i^f+e_i^c+e_i^r}$. Using the CES framework above, I match exactly the energy shares, successfully identifying countries that are more reliant on coal vs. oil and gas vs. non-carbon/renewable: for example, China and India are highly coal-dependent, and Russia, Middle-East and United-States/Canada are the biggest consumers of oil and gas.

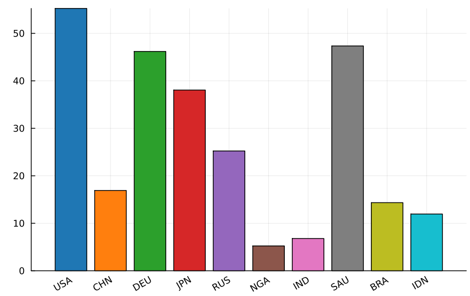


Figure 3: GDP per capita
Thsds 2011-USD PPP, avg. 2018-2023

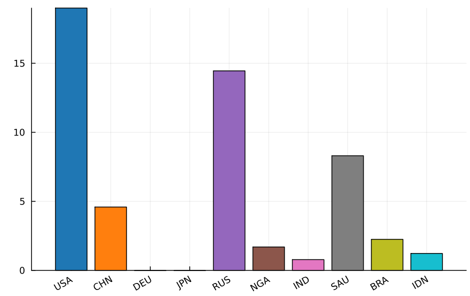


Figure 4: Oil and gas production
GTOE (gigatons oil equiv.), avg. 2018-2023

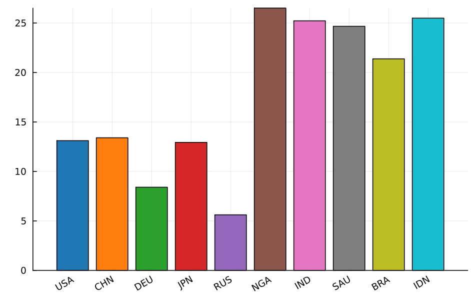


Figure 5: Temperatures
Avg., population-weighted, 2015

5.5 Climate system

Finally, I calibrate the climate model described in Section 3.4 to match important features of the relationship between carbon emissions, temperatures and climate damages.

First, I calibrate two parameters of the global climate system: the climate sensitivity χ, i.e. the reaction of global temperature, $\mathcal{T}_t$, to the atmospheric concentration of CO_2, $\mathcal{S}_t$, and the carbon decay rate, δ_s, representing the exit of carbon of the atmosphere into carbon sinks – oceans, biosphere – and out of the higher atmosphere. To this end, as is standard in Integrated Assessment models, I match the pulse experiment dynamics of larger IAMs – CMIP5 in this case: for a "pulse" of $100GT$ of carbon released – corresponding to 10 years of emissions – the global temperature reaches its peak between $0.20°C$ and $0.25°C$ after 10 years and then decreases slightly to stabilize around $0.17°C$ after 200 years. I follow Dietz et al. (2021), and calibrate $\chi = 0.23$ and $\delta = 0.0004$ to match these two moments, as seen in Figure 2.

Moreover, this climate system is inherently unstable for emissions, $\mathcal{E}_t = \bar{\xi}_t \sum_{i\in\mathbb{I}} \epsilon_{it}$, with trends in population growth, n, and long-term TFP growth $\bar{g}$, where $n = 0.0035$ and $\bar{g} = 0.01$ are the long-term growth rates according to forecast by the UN and World-Bank. To counteract the non-stationarity of the climate system, I follow Krusell and Smith (2022) and assume that part of the emissions $\mathcal{E}_t$ are captured and stored, under the variable $\bar{\xi}_t$. I assume the exponential form, $\bar{\xi}_t = e^{-\xi t}$ and calibrate ζ to match the moment suggested in Krusell and Smith (2022): 50% is captured by 2125, and 100% by 2300 – which is $> 99.9\%$ in our model. This implies that in the Business-as-Usual scenario, global temperatures reach $\sim 4.5°$ by 2100 and are stabilized around 7° by 2400. More optimistic scenarios for Carbon Capture and Storage (CCS) could be imagined without affecting the main result since most of the damages are discounted heavily after 2100.

Second, I calibrate initial temperatures τ_{it_0} with data from Burke et al. (2015), and I display selected countries in Figure 5. Furthermore, I assume the linear pattern scaling $\dot{\tau}_{it} = \Delta_i \mathcal{T}_t$. I identify the scaling parameter in reduced-form by estimating this linear regression over the period $t = 1950{-}2015$ for each country and then aggregating by region i.[24] This procedure does not require extensive and granular data such at geographical characteristics, albedo, etc.

Third, to calibrate the damages, I use a quadratic function as in the DICE - IAM, and seen in equation (23), with the damage parameter $\gamma = 0.00340$. This value is intermediary between the value $\gamma = 0.00311$ in Krusell and Smith (2022), calibrated to match Nordhaus' DICE calibration of 6.6% of loss of global GDP when temperature anomaly $\mathcal{T}_t = 5$, and the updated calibration in Barrage and Nordhaus (2024) which calibrate it at $\gamma = 0.003467$. For small values, I consider $\gamma^- = 0.3\gamma$, following the quantification in Rudik et al. (2021), who show that the negative productivity impact of cold temperatures is much weaker than for hot temperatures.

Finally, to calibrate $\tau_i^\star$, I use also an intermediary assumption between the following two cases: (i) the representative agent economy, like Barrage and Nordhaus (2024), would assume $\tau_i^\star = \tau_{it_0}$, which implies that $\tau_{it} - \tau_i^\star = \Delta_i \mathcal{T}_t$: differences in damages only comes from increases in aggregate temperature. The analysis by Bilal and Känzig (2024) shows that climate damage on GDP comes in large part from the increase in global temperature, causing extreme events. In contrast, (ii) a different view in heterogeneous countries economies would set $\tau_i^\star = \tau^\star$ the same for all regions, at an "ideal" temperature, as in Krusell and Smith (2022) and Kotlikoff, Kubler, Polbin, Sachs and Scheidegger (2021). In this case, differences in climate damages come essentially from differences in initial temperatures. I take the intermediary step and assume $\tau_i^\star = \alpha^T \tau^\star + (1-\alpha^T)\tau_{it_0}$, where $\alpha^\tau = 0.5$ and $\tau^\star = 14.5$ is the average spring temperature of developed economies – and around the yearly average of places like California or Spain.

5.6 Heterogeneity

In this section, I summarize the different dimensions of heterogeneity included in the model and aggregate the parameters of the calibration in Table 1.

[24] To control for the fact that country j has an influence on world temperature $\mathcal{T}_t = \sum_i g_i \tau_{it}$, I estimate the jackknife linear equation with $\mathcal{T}_{t,\neq j} = \sum_{i\neq j} g_i \tau_{it}$ for each j, i.e. $\tau_{jt} = \Delta_j \mathcal{T}_{t,\neq j}$.

Table 2: Heterogeneity across countries

Dimension of heterogeneity	Model parameter	Matched variable from the data	Source of the data
Population	Country size $\mathcal{P}_i$	Population $\mathcal{P}_i$	UN Population Prospect
TFP/technology/institutions	Firm productivity z_i	GDP per capita (2016-PPP) y_i	World Bank/Maddison project
Productivity in energy	Energy-augmenting productivity z_i^e	Energy cost share s_i^e	SRE Energy Institute (2024)
Cost of coal energy	Cost of coal production κ_i^c	Energy mix/coal share e_i^c/e_i	SRE Energy Institute (2024)
Cost of non-carbon energy	Cost of non-carbon production κ_i^r	Energy mix/coal share e_i^r/e_i	SRE Energy Institute (2024)
Local temperature	Initial temperature τ_{it_0}	Pop-weighted yearly temperature	Burke et al. (2015)
Pattern scaling	Pattern scaling Δ_i	Sensitivity of τ_{it} to world $\mathcal{T}_t$	Burke et al. (2015)
Oil-gas reserves	Reserves $\mathcal{R}_i$	Proved Oil-gas reserves $\mathcal{R}_i$	SRE Energy Institute (2024)
Cost of oil-gas extraction	Slope of extraction cost $\bar{\nu}_i$	Oil-gas extracted/produced e_i^x	SRE Energy Institute (2024)
Cost of oil-gas extraction	Curvature of extraction cost ν_i	Profit π_i^f / energy rent	World Bank / WDI

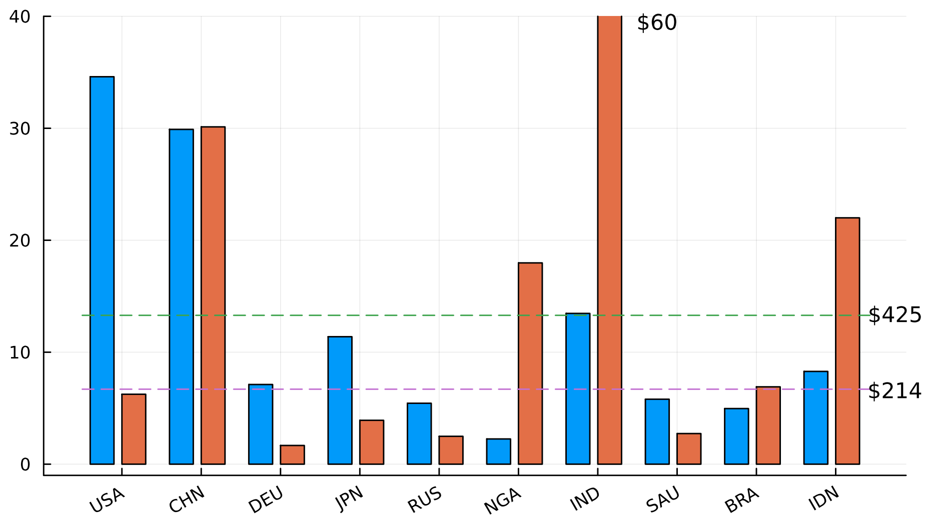
6 Quantitative model results

The results of this section are preliminary. I simulate the model with 68 countries. However, to display the quantitative results in the graphs, I show examples of ten large "example countries": the US, China, Germany, Japan, Russia, Nigeria, India, Saudi Arabia, Brazil, and Indonesia.

6.1 Local and Social Cost of Carbon

In the following graph, I first display the Social Cost of Carbon and Local Cost of Carbon derived above. Recall that in the Second-Best, we have $SCC = \sum_{\mathbb{I}} \widehat{\psi}_i^w \, LCC_i$.

In the next graph we plot the difference between $LCC_i = \frac{\psi_i^{\mathcal{S}}}{\lambda_i^w}$ (blue, left bars) and the carbon-tax relevant terms: $\widehat{\psi}_i^w LCC_i = \frac{\psi_i^{\mathcal{S}}}{\bar{\psi}^w}$ (red, right bars). We see that the US and China have the largest Local Costs of Carbon, since, as argued above, the LCC_i scales proportionally to population $\mathcal{P}_i$ and GDP_i per capital y_i. However, when we account for inequality and the social welfare weight – with marginal utility of consumption – we see that now India, Indonesia, and Nigeria are the countries with the higher "welfare/policy-relevant" Local Cost of Carbon.



In the competitive equilibrium – i.e. without mitigation policies implemented, the Social Cost of Carbon can be written in two ways: one without accounting for redistributive effects $\mathbb{E}^{\mathbb{I}}[\omega_i LCC_i]$ and one that does integrate inequality $\mathbb{E}^{\mathbb{I}}[\widehat{\psi}_i^w LCC_i]$. In our numerical exercise, we obtain:

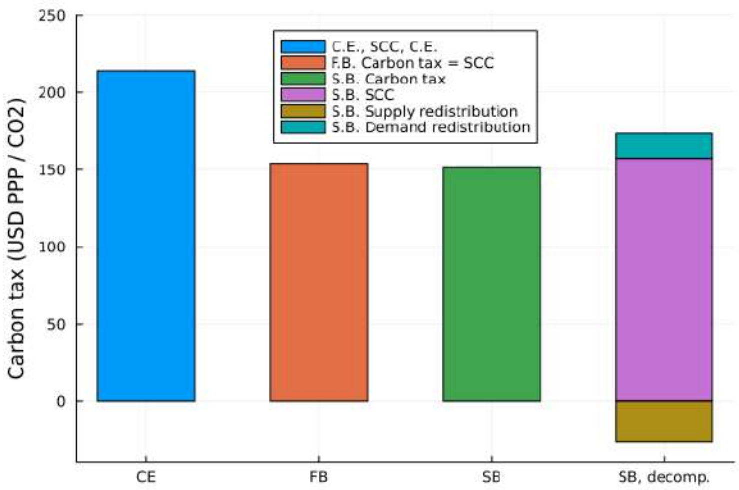
$$\mathbb{E}^{\mathbb{I}}[\omega_i LCC_i] = \$425/tCO_2 \qquad \mathbb{E}^{\mathbb{I}}[\widehat{\psi}_i^w LCC_i] = \$214/tCO_2$$

The result is displayed in Section 6.1.

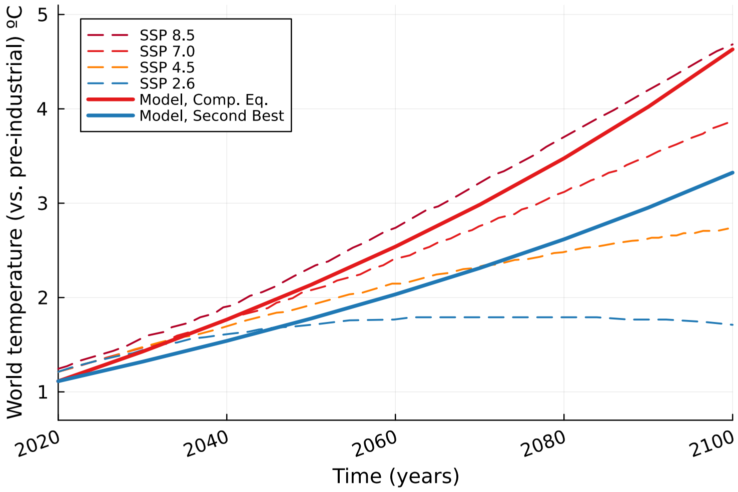
The optimal carbon tax with heterogeneity and redistribution motives in the Second Best allocation is given as:

$$\mathrm{t}^{\varepsilon} = SCC + Supply\ Redistribution + Demand\ Distortion$$

Correcting for the two additional redistributive terms, the optimal carbon tax is 5% lower than the SCC, because the Supply Redistribution and Demand Distortion term largely offset each others. The result with Negishi weights are displayed in **??**. With utilitarian weights, i.e. $\omega_i = 1, \forall i$, the carbon tax much higher, as both the Supply Redistribution and Demand Distortion. The planner would use carbon tax as a tool for redistribution using energy price and terms-of-trade manipulation.



The dynamic setting allows us to compute the temperature path. In Section 6.1, I display the path of global temperature in the competitive equilibrium (i.e. the Business as Usual scenario) vs. Second-Best with the optimal carbon tax. I show that the second-best allocation aligns closely with the Shared Socioeconomic Pathway (SSP) 4.5, i.e., the "Middle of the Road" Scenario. This is due to the fact that global emissions are only reduced by 35%, as the marginal costs of climate change equate the marginal cost of mitigation and the energy transition. As a result, it would not be optimal to reduce carbon emissions to net-zero in this class of Integrated Assessment Models.



7 Conclusion

In this paper, I examine how to design the optimal carbon policy in a world marked by multiple layers of inequality. Through both theoretical and quantitative analysis, I demonstrate that the traditional approach of setting a global carbon tax equal to the Social Cost of Carbon needs to be reconsidered when accounting for global inequalities in emissions, income, climate vulnerability, and policy impacts.

The key theoretical insights emerge from both a simplified model and a rich, dynamic framework. In the First-Best scenario with available redistributive instruments, the optimal policy follows the Pigouvian principle, where the carbon tax equals the Social Cost of Carbon. However, in the more realistic Second-Best scenario without cross-country transfers, the optimal policy must balance climate externalities with redistributive considerations. This leads to two main findings: First, the uniform global carbon tax needs to be adjusted for both supply redistribution and demand distortion effects. Second, when country-specific carbon taxes are possible, they should be inversely proportional to social welfare weights – the product of the planners' Pareto weights and the marginal utility of consumption – resulting in lower taxes for developing economies.

The quantitative analysis, based on a calibrated model covering 68 countries, reveals that accounting for inequality reduces the Social Cost of Carbon by approximately 12% (from \$170 to \$150). This reduction reflects the higher marginal value of wealth in poorer countries. The model also captures competing effects on the optimal carbon tax from energy markets equilibrium: downward pressure to protect fossil-fuel exporters versus upward pressure to minimize energy use distortions. These effects largely offset each other, resulting in an optimal carbon tax slightly below \$150, aligned with existing estimates.

These findings have important implications for international climate policy. They suggest that a one-size-fits-all approach to carbon taxation could be suboptimal when considering global inequalities. Instead, policymakers should consider differentiated carbon pricing schemes that account for countries' economic development levels, heterogeneous climate damages, and energy market exposure. This research also highlights the importance of developing complementary redistributive mechanisms in international climate policy to achieve both environmental and equity objectives.

Future research would extend this analysis by examining the dynamic evolution of these effects as developing economies grow, climate impacts intensify, and fossil fuel reserves deplete. Moreover, it would assess the role of uncertainty and climate risk for the optimal policy as in Bourany (2023). Additionally, investigating the political economy constraints and implementation challenges of differentiated carbon pricing schemes matters considerably for practical policy design, as I analyzed in Bourany (2024*a*).

References

Acemoglu, Daron, Philippe Aghion and David Hémous (2014), 'The Environment and Directed Technical Change in a North–South model', *Oxford Review of Economic Policy* **30**(3), 513–530.

Acemoglu, Daron, Philippe Aghion, Leonardo Bursztyn and David Hemous (2012), 'The Environment and Directed Technical Change', *American economic review* **102**(1), 131–166.

Acemoglu, Daron, Ufuk Akcigit, Douglas Hanley and William Kerr (2016), 'Transition to Clean Technology', *Journal of Political Economy* **124**(1), 52–104.

Achdou, Yves, Jiequn Han, Jean-Michel Lasry, Pierre-Louis Lions and Benjamin Moll (2022), 'Income and wealth distribution in macroeconomics: A continuous-time approach', *The review of economic studies* **89**(1), 45–86.

Anderson, Soren T, Ryan Kellogg and Stephen W Salant (2018), 'Hotelling under pressure', *Journal of Political Economy* **126**(3), 984–1026.

Anthoff, David, Cameron Hepburn and Richard SJ Tol (2009), 'Equity weighting and the marginal damage costs of climate change', *Ecological Economics* **68**(3), 836–849.

Anthoff, David and Johannes Emmerling (2019), 'Inequality and the social cost of carbon', *Journal of the Association of Environmental and Resource Economists* **6**(2), 243–273.

Anthoff, David and Richard SJ Tol (2014), The income elasticity of the impact of climate change, *in* 'Is the Environment a Luxury?', Routledge, pp. 34–47.

Arkolakis, Costas and Conor Walsh (2023), 'Clean Growth'.

Asker, John, Allan Collard-Wexler, Charlotte De Canniere, Jan De Loecker and Christopher R Knittel (2024), Two wrongs can sometimes make a right: The environmental benefits of market power in oil, Technical report, National Bureau of Economic Research.

Asker, John, Allan Collard-Wexler and Jan De Loecker (2019), '(mis) allocation, market power, and global oil extraction', *American Economic Review* **109**(4), 1568–1615.

Atkeson, Andrew and Patrick J Kehoe (1999), 'Models of energy use: Putty-putty versus putty-clay', *American Economic Review* **89**(4), 1028–1043.

Atkinson, Anthony Barnes and Joseph E Stiglitz (1976), 'The design of tax structure: direct versus indirect taxation', *Journal of public Economics* **6**(1-2), 55–75.

Bardi, Ugo (2011), *The limits to growth revisited*, Springer Science & Business Media.

Barnett, Michael, William Brock and Lars Peter Hansen (2020), 'Pricing uncertainty induced by climate change', *The Review of Financial Studies* **33**(3), 1024–1066.

Barnett, Michael, William Brock and Lars Peter Hansen (2022), 'Climate change uncertainty spillover in the macroeconomy', *NBER Macroeconomics Annual* **36**(1), 253–320.

Barrage, Lint and William Nordhaus (2024), 'Policies, projections, and the social cost of carbon: Results from the DICE-2023 model', *Proceedings of the National Academy of Sciences* **121**(13), e2312030121. Publisher: Proceedings of the National Academy of Sciences.

Behmer, Scott (2023), 'Sticks vs carrots: Climate policy under government turnover', *Available at SSRN 5066195* .

Belfiori, Elisa, Daniel Carroll and Sewon Hur (2024), 'Unequal Climate Policy in an Unequal World'.

Belfiori, Maria Elisa (2018), 'Climate change and intergenerational equity: Revisiting the uniform taxation principle on carbon energy inputs', *Energy Policy* **121**, 292–299.

Benmir, Ghassane and Josselin Roman (2022), The distributional costs of net-zero: A hank perspective, Technical report, Working paper.

Bhandari, Anmol, David Evans, Mikhail Golosov and Thomas J Sargent (2021*a*), 'Inequality, business cycles, and monetary-fiscal policy', *Econometrica* **89**(6), 2559–2599.

Bhandari, Anmol, David Evans, Mikhail Golosov and Thomas Sargent (2021*b*), Efficiency, insurance, and redistribution effects of government policies, Technical report, Working paper.

Bhandari, Anmol, Thomas Bourany, David Evans and Mikhail Golosov (2023), 'A Perturbational Approach for Approximating Heterogeneous Agent Models', National Bureau of Economic Research.

Bilal, Adrien (2021), Solving heterogeneous agent models with the master equation, Technical report, Technical report, University of Chicago.

Bilal, Adrien and Diego R Känzig (2024), 'The macroeconomic impact of climate change: Global vs. local temperature'.

Bilal, Adrien and Esteban Rossi-Hansberg (2023*a*), 'Anticipating Climate Change Across the United States'.

Bilal, Adrien and Esteban Rossi-Hansberg (2023*b*), Anticipating climate change risk across the united states, Technical report, Technical report, University of Chicago.

Bolt, Jutta and Jan Luiten van Zanden (2023), 'Maddison-style estimates of the evolution of the world economy: A new 2023 update', *Journal of Economic Surveys* **n/a**(n/a).

Bornstein, Gideon, Per Krusell and Sergio Rebelo (2023), 'A world equilibrium model of the oil market', *The Review of Economic Studies* **90**(1), 132–164.

Boucekkine, Raouf, Carmen Camacho and Benteng Zou (2009), 'Bridging the gap between growth theory and the new economic geography: The spatial ramsey model', *Macroeconomic Dynamics* **13**(1), 20–45.

Bourany, Thomas (2019), 'The Wealth Distribution over of Business Cycle: A Mean-Field Game with Common Noise', Master Thesis, Paris Diderot / Sorbonne University.

Bourany, Thomas (2022), 'Energy shocks and aggregate fluctuations'.

Bourany, Thomas (2023), 'The Distributional Consequences of Climate Uncertainty', Slides, University of Chicago.

Bourany, Thomas (2024*a*), 'The Optimal Design of Climate Agreements: Inequality, Trade and Incentives for Climate Policy', Job Market Paper, University of Chicago.

Bourany, Thomas (2024*b*), 'When is Aggregation Enough? Aggregation and Projection with the Master Equation', Slides, University of Chicago.

Burke, Marshall, Solomon M Hsiang and Edward Miguel (2015), 'Global non-linear effect of temperature on economic production', *Nature* **527**(7577), 235–239.

Cai, Yongyang, Kenneth L Judd and Thomas S Lontzek (2012*a*), Continuous-time methods for integrated assessment models, Technical report, National Bureau of Economic Research.

Cai, Yongyang, Kenneth L Judd and Thomas S Lontzek (2012*b*), 'DSICE: A dynamic stochastic integrated model of climate and economy'.

Cai, Yongyang and Thomas S Lontzek (2019), 'The social cost of carbon with economic and climate risks', *Journal of Political Economy* **127**(6), 2684–2734.

Cardaliaguet, Pierre (2013/2018), 'Notes on mean field games.', *Lecture notes from P.L. Lions' lectures at College de France and P. Cardaliaguet at Paris Dauphine* .

Cardaliaguet, Pierre, François Delarue, Jean-Michel Lasry and Pierre-Louis Lions (2019), *The Master Equation and the Convergence Problem in Mean Field Games*, Princeton University Press.

Carleton, Tamma, Amir Jina, Michael Delgado, Michael Greenstone, Trevor Houser, Solomon Hsiang, Andrew Hultgren, Robert E Kopp, Kelly E McCusker, Ishan Nath, James Rising, Ashwin Rode, Hee Kwon Seo, Arvid Viaene, Jiacan Yuan and Alice Tianbo Zhang (2022), 'Valuing the Global Mortality Consequences of Climate Change Accounting for Adaptation Costs and Benefits*', *The Quarterly Journal of Economics* **137**(4), 2037–2105.

Carleton, Tamma and Michael Greenstone (2021), 'Updating the united states government's social cost of carbon', *University of Chicago, Becker Friedman Institute for Economics Working Paper* (2021-04).

Carmona, Rene and François Delarue (2018), *Probabilistic Theory of Mean Field Games with Applications I-II*, Springer.

Carmona, René, François Delarue and Aimé Lachapelle (2013), 'Control of mckean–vlasov dynamics versus mean field games', *Mathematics and Financial Economics* **7**(2), 131–166.

Carmona, René, François Delarue and Daniel Lacker (2016), 'Mean field games with common noise'.

Carmona, René, François Delarue et al. (2015), 'Forward–backward stochastic differential equations and controlled mckean–vlasov dynamics', *The Annals of Probability* **43**(5), 2647–2700.

Carmona, René, Gökçe Dayanıklı and Mathieu Laurière (2022), 'Mean field models to regulate carbon emissions in electricity production', *Dynamic Games and Applications* **12**(3), 897–928.

Carmona, René and Mathieu Laurière (2022), 'Convergence analysis of machine learning algorithms for the numerical solution of mean field control and games: Ii – the finite horizon case', *The Annals of Applied Probability* **32**(6), 4065–4105.

Chari, V. V., Juan Pablo Nicolini and Pedro Teles (2023), 'Optimal Cooperative Taxation in the Global Economy', *Journal of Political Economy* **131**(1), 95–130. Publisher: The University of Chicago Press.

Conte, Maddalena, Pierre Cotterlaz, Thierry Mayer et al. (2022), 'The cepii gravity database'.

Costinot, Arnaud and Iván Werning (2023), 'Robots, trade, and luddism: A sufficient statistic approach to optimal technology regulation', *The Review of Economic Studies* **90**(5), 2261–2291.

Cruz, José-Luis and Esteban Rossi-Hansberg (2021), The Economic Geography of Global Warming, Technical report, National Bureau of Economic Research.

Cruz, José Luis and Esteban Rossi-Hansberg (2022*a*), 'Local carbon policy', *NBER Working Paper* (w30027).

Cruz, Jose-Luis and Esteban Rossi-Hansberg (2022*b*), 'Local Carbon Policy'.

Cruz, Jose-Luis and Esteban Rossi-Hansberg (2024), 'The Economic Geography of Global Warming', *The Review of Economic Studies* **91**(2), 899–939.

Davila, Eduardo and Andreas Schaab (2023), 'Optimal Monetary Policy with Heterogeneous Agents: Discretion, Commitment, and Timeless Policy'.

Davila, Eduardo and Ansgar Walther (2022), 'Corrective Regulation with Imperfect Instruments'.

Diamond, Peter A (1973), 'Consumption externalities and imperfect corrective pricing', *The Bell Journal of Economics and Management Science* pp. 526–538.

Diamond, Peter A and James A Mirrlees (1971), 'Optimal taxation and public production i: Production efficiency', *The American economic review* **61**(1), 8–27.

Dietz, Simon and Frank Venmans (2019), 'Cumulative carbon emissions and economic policy: in search of general principles', *Journal of Environmental Economics and Management* **96**, 108–129.

Dietz, Simon, Frederick van der Ploeg, Armon Rezai and Frank Venmans (2021), 'Are economists getting climate dynamics right and does it matter?', *Journal of the Association of Environmental and Resource Economists* **8**(5), 895–921.

Dissou, Yazid and Muhammad Shahid Siddiqui (2014), 'Can carbon taxes be progressive?', *Energy Economics* **42**, 88–100.

Douenne, Thomas and Adrien Fabre (2022), 'Yellow vests, pessimistic beliefs, and carbon tax aversion', *American Economic Journal: Economic Policy* **14**(1), 81–110.

Douenne, Thomas, Albert Jan Hummel and Marcelo Pedroni (2023), Optimal fiscal policy in a climate-economy model with heterogeneous households, Technical report, SSRN Working Paper.

Douenne, Thomas, Sebastian Dyrda, Albert Jan Hummel and Marcelo Pedroni (2024), 'Optimal climate policy with incomplete markets'.

Energy Institute (2024), 'Statistical review of world energy'.

Farrokhi, Farid and Ahmad Lashkaripour (2024), 'Can Trade Policy Mitigate Climate Change', *Econometrica* .

Folini, Doris, Felix Kübler, Aleksandra Malova and Simon Scheidegger (2021), 'The climate in climate economics', *arXiv preprint arXiv:2107.06162* .

Fried, Stephie, Kevin Novan and William B Peterman (2024), 'Understanding the inequality and welfare impacts of carbon tax policies', *Journal of the Association of Environmental and Resource Economists* **11**(S1), S231–S260.

Gentile, Claudia (2024), 'Relying on intermittency: Clean energy, storage, and innovation in a macro climate model'.

Golosov, Mikhail, John Hassler, Per Krusell and Aleh Tsyvinski (2014), 'Optimal Taxes on Fossil Fuel in General Equilibrium', *Econometrica* **82**(1), 41–88.

Grossman, Gene M, Elhanan Helpman, Ezra Oberfield and Thomas Sampson (2017), 'Balanced growth despite uzawa', *American Economic Review* **107**(4), 1293–1312.

Hansen, Lars Peter and Thomas J Sargent (2001), 'Robust control and model uncertainty', *American Economic Review* **91**(2), 60–66.

Hassler, John, Per Krusell and Anthony A Smith Jr (2016), Environmental macroeconomics, *in* 'Handbook of macroeconomics', Vol. 2, Elsevier, pp. 1893–2008.

Hassler, John, Per Krusell and Conny Olovsson (2010), 'Oil monopoly and the climate', *American Economic Review* **100**(2), 460–64.

Hassler, John, Per Krusell and Conny Olovsson (2021*a*), 'Directed technical change as a response to natural resource scarcity', *Journal of Political Economy* **129**(11), 3039–3072.

Hassler, John, Per Krusell and Conny Olovsson (2021*b*), 'Presidential Address 2020 Suboptimal Climate Policy', *Journal of the European Economic Association* **19**(6), 2895–2928.

Hassler, John, Per Krusell, Conny Olovsson and Michael Reiter (2020), 'On the effectiveness of climate policies', *IIES WP* **53**, 54.

Heal, Geoffrey and Wolfram Schlenker (2019), Coase, hotelling and pigou: The incidence of a carbon tax and co 2 emissions, Technical report, National Bureau of Economic Research.

Hillebrand, Elmar and Marten Hillebrand (2019), 'Optimal climate policies in a dynamic multi-country equilibrium model', *Journal of Economic Theory* **179**, 200–239.

Hotelling, Harold (1931), 'The economics of exhaustible resources', *Journal of political Economy* **39**(2), 137–175.

IPCC, Pörtner, H.-O., D.C. Roberts, H. Adams, I. Adelekan, C. Adler, R. Adrian, P. Aldunce, E. Ali, R. Ara Begum, B. Bednar Friedl, R. Bezner Kerr, R. Biesbroek, J. Birkmann, K. Bowen, M.A. Caretta, J. Carnicer, E. Castellanos, T.S. Cheong, W. Chow, G. Cisse G. Cisse and Z. Zaiton Ibrahim (2022), *Climate Change 2022: Impacts, Adaptation and Vulnerability*, Technical Summary, Cambridge University Press, Cambridge, UK and New York, USA.

Kellogg, Ryan (2014), 'The effect of uncertainty on investment: Evidence from texas oil drilling', *American Economic Review* **104**(6), 1698–1734.

Kilian, Lutz (2009), 'Not all oil price shocks are alike: Disentangling demand and supply shocks in the crude oil market', *American Economic Review* **99**(3), 1053–69.

Koeste, Mark J., Henri L.F. de Groot and Raymond J.G.M. Florax (2008), 'Capital-energy substitution and shifts in factor demand: A meta-analysis', *Energy Economics* **30**(5), 2236–2251.

Köppl, Angela and Margit Schratzenstaller (2023), 'Carbon taxation: A review of the empirical literature', *Journal of Economic Surveys* **37**(4), 1353–1388.

Kortum, Samuel S. and David A. Weisbach (2021), 'Optimal Unilateral Carbon Policy'.

Kotlikoff, Laurence, Felix Kubler, Andrey Polbin, Jeffrey Sachs and Simon Scheidegger (2021), 'Making Carbon Taxation a Generational Win Win', *International Economic Review* **62**(1), 3–46. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/iere.12483.

Kotlikoff, Laurence, Felix Kubler, Andrey Polbin and Simon Scheidegger (2021*a*), 'Pareto-improving carbon-risk taxation', *Economic Policy* **36**(107), 551–589.

Kotlikoff, Laurence J., Felix Kubler, Andrey Polbin and Simon Scheidegger (2021*b*), 'Can Today's and Tomorrow's World Uniformly Gain from Carbon Taxation?'.

Krusell, Per and Anthony A Smith (2022), Climate change around the world, Technical report, National Bureau of Economic Research.

Kuhn, Moritz and Lennard Schlattmann (2024), Distributional consequences of climate policies, Technical report, Working Paper, Universitäten Mannheim und Bonn.

Le Grand, François, Alaïs Martin-Baillon and Xavier Ragot (2021), Should monetary policy care about redistribution? optimal fiscal and monetary policy with heterogeneous agents, Technical report, Working Paper, SciencesPo.

Le Grand, François, Florian Oswald, Xavier Ragot and Aurélien Saussay (2023), Fiscal policy for climate change, Technical report, Working paper.

Le Grand, François, Florian Oswald, Xavier Ragot and Saussay Aurélien (2023), Fiscal policy for climate change, Technical report, Working Paper, SciencesPo.

LeGrand, François and Xavier Ragot (2022), Optimal policies with heterogeneous agents: Truncation and transitions, Technical report, Working Paper, SciencesPo.

Lemoine, Derek and Ivan Rudik (2017*a*), 'Managing climate change under uncertainty: Recursive integrated assessment at an inflection point', *Annual Review of Resource Economics* **9**, 117–142.

Lemoine, Derek and Ivan Rudik (2017*b*), 'Steering the climate system: using inertia to lower the cost of policy', *American Economic Review* **107**(10), 2947–2957.

Lenton, Timothy M (2011), 'Early warning of climate tipping points', *Nature climate change* **1**(4), 201–209.

Lontzek, Thomas S, Yongyang Cai, Kenneth L Judd and Timothy M Lenton (2015), 'Stochastic integrated assessment of climate tipping points indicates the need for strict climate policy', *Nature Climate Change* **5**(5), 441–444.

McKay, Alisdair and Christian K Wolf (2022), Optimal policy rules in HANK, Technical report, Working Paper, FRB Minneapolis.

Meadows, Donella H, Dennis L Meadows, Jorgen Randers and William W Behrens (1972), 'The limits to growth', *New York* **102**(1972), 27.

Nakov, Anton and Galo Nuno (2013), 'Saudi arabia and the oil market', *The Economic Journal* **123**(573), 1333–1362.

Nordhaus, William (2015), 'Climate Clubs: Overcoming Free-riding in International Climate Policy', *American Economic Review* **105**(4), 1339–1370.

Nordhaus, William D (1993), 'Optimal greenhouse-gas reductions and tax policy in the" dice" model', *The American Economic Review* **83**(2), 313–317.

Nordhaus, William D (2007), 'A review of the stern review on the economics of climate change', *Journal of economic literature* **45**(3), 686–702.

Nordhaus, William D (2011), 'Estimates of the social cost of carbon: background and results from the rice-2011 model'.

Nordhaus, William D (2017), 'Revisiting the social cost of carbon', *Proceedings of the National Academy of Sciences* **114**(7), 1518–1523.

Nordhaus, William D and Joseph Boyer (2000), *Warming the world: economic models of global warming*, MIT press.

Nordhaus, William D. and Zili Yang (1996), 'A Regional Dynamic General-Equilibrium Model of Alternative Climate-Change Strategies', *The American Economic Review* **86**(4), 741–765. Publisher: American Economic Association.

Papageorgiou, Chris, Marianne Saam and Patrick Schulte (2017), 'Substitution between Clean and Dirty Energy Inputs: A Macroeconomic Perspective', *The Review of Economics and Statistics* **99**(2), 281–290.

Pham, Huyên and Xiaoli Wei (2017), 'Dynamic programming for optimal control of stochastic mckean–vlasov dynamics', *SIAM Journal on Control and Optimization* **55**(2), 1069–1101.

Pindyck, Robert S (2017), 'The use and misuse of models for climate policy', *Review of Environmental Economics and Policy* .

Ricke, Katharine L and Ken Caldeira (2014), ‘Maximum warming occurs about one decade after a carbon dioxide emission’, *Environmental Research Letters* **9**(12), 124002.

Ricke, Katharine, Laurent Drouet, Ken Caldeira and Massimo Tavoni (2018), ‘Country-level social cost of carbon’, *Nature Climate Change* **8**(10), 895–900.

Rode, Ashwin, Tamma Carleton, Michael Delgado, Michael Greenstone, Trevor Houser, Solomon Hsiang, Andrew Hultgren, Amir Jina, Robert E Kopp, Kelly E McCusker et al. (2021), ‘Estimating a social cost of carbon for global energy consumption’, *Nature* **598**(7880), 308–314.

Rudik, Ivan, Gary Lyn, Weiliang Tan and Ariel Ortiz-Bobea (2021), ‘The Economic Effects of Climate Change in Dynamic Spatial Equilibrium’.

Schaab, Andreas and Stacy Yingqi Tan (2023), ‘Monetary and fiscal policy according to hank-io’, *Unpublished Manuscript* .

Schlattmann, Lennard (2024), Spatial redistribution of carbon taxes, Technical report, ECONtribute Discussion Paper.

Stern, Nicholas and Nicholas Herbert Stern (2007), *The economics of climate change: the Stern review*, cambridge University press.

Stoddard, Isak, Kevin Anderson, Stuart Capstick, Wim Carton, Joanna Depledge, Keri Facer, Clair Gough, Frederic Hache, Claire Hoolohan, Martin Hultman et al. (2021), ‘Three decades of climate mitigation: why haven’t we bent the global emissions curve?’, *Annual Review of Environment and Resources* **46**, 653–689.

Van den Bremer, Ton S and Frederick Van der Ploeg (2021), ‘The risk-adjusted carbon price’, *American Economic Review* **111**(9), 2782–2810.

van der Ploeg, Frederick, Armon Rezai and Miguel Tovar (2024), ‘Third-best carbon taxation: Trading off emission cuts, equity, and efficiency’, *Journal of the Association of Environmental and Resource Economists* **0**(ja), null.

Weitzman, Martin L. (2003), Prices vs. Quantities 1 , 2, *in* ‘The Theory and Practice of Command and Control in Environmental Policy’, Routledge. Num Pages: 15.

Weitzman, Martin L (2012), ‘Ghg targets as insurance against catastrophic climate damages’, *Journal of Public Economic Theory* **14**(2), 221–244.

Weitzman, Martin L. (2015), ‘Internalizing the Climate Externality: Can a Uniform Price Commitment Help?’, *Economics of Energy & Environmental Policy* **4**(2), 37–50. Publisher: International Association for Energy Economics.

Wöhrmüller, Stefan (2024), Carbon taxation and precautionary savings, Technical report, SSRN Working Paper.

Yong, Jiongmin and Xun Yu Zhou (1999), *Stochastic controls: Hamiltonian systems and HJB equations*, Vol. 43, Springer Science & Business Media.

A Toy model – Theoretical results

The first two sections are forthcoming. For a similar model – with multi-country trade – the proofs of the main theorems are analogous to the ones provided in the appendix of Bourany (2024*a*).

A.1 First Best

A.2 Second-Best

A.3 Cap and Trade

Competitive equilibrium

$$\max_{c_i, e_i, \varepsilon_i} U(c_i) \qquad s.t \qquad \begin{cases} c_i & = \mathcal{D}_i(\mathcal{S}) z_i F(e_i) + q^e(e_i^x - e_i) - \mathcal{C}_i(e_i^x) + q^\varepsilon(\bar{\varepsilon}_i - \varepsilon_i) \quad [\lambda_i] \\ \xi e_i & \leq \varepsilon_i \qquad [\eta_i q^\varepsilon] \end{cases}$$

where η_i is the Lagrange multiplier/shadow for each dollar of carbon permits – hence the normalization by q^ε. The carbon intensity of energy is ξ: it requires ξ allowance to get enough carbon permit ε_i for one unit (e.g. ton of oil equivalent) of energy e_i. This yields the optimality condition of the firm for energy e_i and carbon permits ε_i:

$$MPe_i = q^e + \xi \eta_i q^\varepsilon \qquad [\varepsilon_i] \qquad \eta_i q^\varepsilon = q^\varepsilon$$

which implies that the implicit carbon tax is $MPe_i - q^e = \xi \tilde{\mathrm{t}}^\varepsilon = \xi q^\varepsilon$. Multiplier η_i for inequality constraint $e_i \geq \varepsilon_i$. Optimality of ε_i: $\eta_i = 1$.

The Lagrangian for the Second Best allocation in this context:

$$\begin{aligned}\mathcal{L}(c_i, e_i, \varepsilon_i, \bar{\varepsilon}_i, \overline{\mathcal{E}}, \boldsymbol{\lambda}) = & \sum_i \mathcal{P}_i \omega_i U(c_i) + \sum_i \omega_i \mathcal{P}_i \phi_i \Big(\mathcal{D}_i(\mathcal{S}) z_i F(e_i) + q^e(e_i^x - e_i) - \mathcal{C}_i(\mathcal{P}_i e_i^x)/\mathcal{P}_i + q^\varepsilon(\bar{\varepsilon}_i - \varepsilon_i) + \mathrm{t}_i^{ls} - c_i \Big) \\ & + \mu^e q^e \Big(\sum_j \mathcal{P}_i (e_j^x - e_j) \Big) + \mu^\varepsilon q^\varepsilon \Big(\overline{\mathcal{E}} + \sum_j \mathcal{P}_i (\bar{\varepsilon}_i - \varepsilon_i) \Big) + \mu^g \Big(\sum_i q^\varepsilon \mathcal{P}_i (\varepsilon_i - \bar{\varepsilon}_i) - \mathcal{P}_i \mathrm{t}_i^{ls} \Big) \\ & + \sum_i \mathcal{P}_i \omega_i \theta_i (q^e - \mathcal{C}_i'(\mathcal{P}_i e_i^x)) + \sum_i \mathcal{P}_i \omega_i \upsilon_i (MPe_i - q^e - \xi q^\varepsilon)\end{aligned}$$

Two options for rebate:
(i) Global rebate (government budget): $\sum_i \mathrm{t}_i^{ls} = \sum_i q^\varepsilon(\varepsilon_i - \bar{\varepsilon}_i)$ (multiplier μ^g), or:
(ii) Local lump-sum rebate $\mathrm{t}_i^{ls} = q^\varepsilon(\varepsilon_i - \bar{\varepsilon}_i)$ (in that case $\mu^g = 0$, we replace t_i^{ls} in budget constraint)

Planner's optimality conditions

- $[c_i]$

$$\phi_i = U'(c_i)$$

- $[e_i]$

$$\mathcal{P}_i \omega_i \phi_i [MPe_i - q^e] + \xi \sum_j \mathcal{P}_j \omega_j \phi_j \mathcal{D}_i'(\mathcal{S}) z_j F(e_j) - \mu^e + \mathcal{P}_i \omega_i \upsilon_i \mathcal{D}_i(\mathcal{S}) z_i F''(e_i) = 0$$

$$\mathcal{P}_i \omega_i \phi_i \xi q^\varepsilon = \xi \overline{\phi} SCC + \mu^e - \mathcal{P}_i \omega_i \upsilon_i \mathcal{D}_i(\mathcal{S}) z_i F''(e_i)$$

- $[e_i^x]$

$$\underbrace{\mathcal{P}_i\omega_i\phi_i[q^e - \mathcal{C}_i'(\mathcal{P}_i e_i^x)]}_{=0} + \mu^e\mathcal{P}_i - \mathcal{P}_i\omega_i\theta_i\mathcal{P}_i\mathcal{C}_i''(\mathcal{P}_i e_i^x)$$

$$\mathcal{P}_i\omega_i\theta_i = \mu^e/\big(\mathcal{C}_i''(\mathcal{P}_i e_i^x)\big)$$

- $[q^e]$

$$\sum_i \mathcal{P}_i\omega_i\phi_i(e_i^x - e_i) + \sum_i \mathcal{P}_i\omega_i\theta_i - \sum_i \mathcal{P}_i\omega_i\upsilon_i = 0$$

- $[\bar{\mathcal{E}}]$

$$\mu^\varepsilon q^\varepsilon = 0$$

Given that the planner controls directly the supply of carbon allowances, the market clearing is not a binding constraint for the optimal policy.

- $[\varepsilon_i]$ & $[\bar{\varepsilon}_i]$ & $[\mathrm{t}_i^{ls}]$ in the case of global rebate:

$$\begin{aligned}
[\varepsilon_i] \qquad & -\omega_i\mathcal{P}_i\phi_i q^\varepsilon - q^\varepsilon\mu^\varepsilon\mathcal{P}_i + \mu^g q^\varepsilon\mathcal{P}_i = 0 \\
& \mu^\varepsilon - \mu^g = -\omega_i\phi_i \\
[\bar{\varepsilon}_i] \qquad & \omega_i\mathcal{P}_i\phi_i q^\varepsilon + q^\varepsilon\mu^\varepsilon - \mu^g q^\varepsilon\mathcal{P}_i = 0 \\
& \mu^\varepsilon - \mu^g = -\omega_i\phi_i \\
[\mathrm{t}_i^{ls}] \qquad & \sum_i \mathcal{P}_i\omega_i\phi_i = \mu^g\sum_i \mathcal{P}_i
\end{aligned}$$

This implies that the planner implements full redistribution using the distribution of "free" carbon permits $\bar{\varepsilon}_i$ and lump-sum transfers t_i^{ls}. The exact mix $\{\mathrm{t}_i^{ls}, \bar{\varepsilon}_i\}$ is undetermined as long as the following condition holds:

$$\omega_i\phi_i = \omega_i U'(c_i) = \mu^g = \omega_j U'(c_j)$$

- $[\varepsilon_i]$ & $[\bar{\varepsilon}_i]$ Local rebate: no impact in the budget constraint

$$[\varepsilon_i] - q^\varepsilon\mu^\varepsilon\mathcal{P}_i = 0$$
$$[\bar{\varepsilon}_i] q^\varepsilon\mu^\varepsilon\mathcal{P}_i = 0$$

Given that the purchase of carbon permits, and hence the revenue of "free" carbon permits, are redistributed/taxed lump-sum, we obtain that $\bar{\varepsilon}_i$ is a redundant policy instrument. Moreover, the planner chooses the same policy as the agent, such that $\varepsilon_i = \xi e_i$.

- $[q^\varepsilon]$

$$\sum_i \mathcal{P}_i\omega_i\upsilon_i = 0$$

Again, as in the carbon taxation case, this implies no "aggregate distortion" at the global level.

B Quantitative model - Competitive equilibrium

Dynamics of the individual state variables $\underline{s}_{it} = (w_{it}, \tau_{it}, \mathcal{R}_{it}, \mathcal{S}_t)$:

$$
\begin{aligned}
\dot{w}_{it} &= (r^\star_t - n_i - \bar{g}_i)w_{it} + \mathcal{D}(\tau_{it})F(k_{it}, e_{it}) - (r^\star_t + \delta)k_{it} + \pi^f_i(q^f_t, \mathcal{R}_{it}) - q^e_{it}e_{it} - c_{it} \\
\mathcal{E}_t &= \bar{\xi}_t \sum_{i\in\mathbb{I}} \mathcal{P}_i e^{(n_i+\bar{g}_i)t}(\xi^f e^f_{it} + \xi^c_{it}) \\
\dot{\tau}_{it} &= \zeta(\Delta_i\,\chi\mathcal{S}_t - (\tau_{it} - \tau_{it_0})) \qquad\qquad \dot{\mathcal{S}}_t = \mathcal{E}_t - \delta_s\mathcal{S}_t \\
\dot{\mathcal{R}}_{it} &= -e^f_{it} \qquad\qquad q^f_t = \bar{\nu}_t(e^x_{it}/\mathcal{R}_{it})^{\nu_i}
\end{aligned}
$$

Household problem: Pontryagin Maximum Principle

$$
\begin{aligned}
\mathcal{H}^{hh}(s, \{c,k,e^f,e^c,e^r\}, \{\lambda\}) = {}& e^{-(\rho-n_i-(1-\eta)\bar{g}_i)t}u(c_i,\tau_i) + e^{-(\rho-n_i-(1-\eta)\bar{g}_i)t}\lambda^w_{it}\Big((r^\star_t-(n_i+\bar{g}_i))w_{it} + \mathcal{D}^y_i(\tau_{it})z_iF(k_{it}, e^f_{it}, e^c_{it}, e^r_{it}) \\
&+ \pi^f_i(q^f_t, \mathcal{R}_{it}) - (r^\star_t + \delta)k_{it} - q^f_t e^f_{it} - q^c_{it}e^c_{it} - q^r_{it}e^r_{it} - c_{it}\Big) \\
&+ e^{-(\rho-n_i-(1-\eta)\bar{g}_i)t}\lambda^S_{it}(\mathcal{E}_t - \delta_s\mathcal{S}_t) \\
&+ e^{-(\rho-n_i-(1-\eta)\bar{g}_i)t}\lambda^\tau_{it}(\zeta(\Delta_i\,\chi\mathcal{S}_t - (\tau_{it} - \tau_{it_0})))
\end{aligned}
$$

Choice of controls:

$$
\begin{aligned}
& & u_c(c_{it}, \tau_{it}) = \mathcal{D}^u(\tau_{it})u'(\mathcal{D}(\tau_{it})c_{it}) = \lambda^w_{it} \\
&[k_t] & MPk_{it} = r^\star_t + \delta \\
&[x_t] & MPe^x_{it} = \mathcal{D}^y_i(\tau_{it})z_i\ \partial_x F(k_{it}, e^f_{it}, e^c_{it}, e^r_{it}) = q^x_{it} \qquad \text{for } x \in \{f,c,r\} \\
&[e^x_t] & q^f_t = \nu^f_{i\,e^x}(e^x_{it}, \mathcal{R}_{it})
\end{aligned}
$$

The Pontryagin maximum principle for the states $\{w_{it}, \tau_{it}, \mathcal{S}_t\}$

$$
\begin{aligned}
&[w_t] & \dot{\lambda}^w_{it} &= \lambda^w_{it}(\rho - n_i - (1-\eta)\bar{g}_i) - \mathcal{H}_w(\cdot) = \lambda^w_{it}[(\rho - n_i - (1-\eta)\bar{g}_i) - (r^\star_t - n_i - \bar{g}_i)] \\
& & \Rightarrow\qquad & \dot{\lambda}^w_{it} = \lambda^w_{it}(\rho + \eta\bar{g}_i - r^\star_t) \\
&[\tau_{it}] & \dot{\lambda}^\tau_{it} &= \lambda^\tau_{it}(\rho-n_i-(1-\eta)\bar{g}_i + \zeta) + \underbrace{\gamma^y(\tau_{it}-\tau^\star_i)\mathcal{D}^y_i(\tau_{it})}_{-\partial_\tau\mathcal{D}^y} z_iF(k_{it}, e_{it})\lambda^w_{it} + \underbrace{\gamma^u(\tau_{it}-\tau^\star_i)\mathcal{D}^u_i(\tau_{it})}_{-\partial_\tau\mathcal{D}^u} u'\big(\mathcal{D}^u(\tau_{it})c_{it}\big)c_{it} \\
&[\mathcal{S}_t] & \dot{\lambda}^S_{it} &= \lambda^S_{it}(\rho-n_i-(1-\eta)\bar{g}_i + \delta_s) - \zeta\,\chi\,\Delta_i\,\lambda^\tau_{it}
\end{aligned}
$$

Recall if $\rho + \eta\bar{g}_i < r^\star_t$, then λ^w_t decreases (and consumption increases) over time.

Solving the ODE for the Local Cost of Carbon

$$
\begin{aligned}
& & \lambda^S_{it} &= -\int_t^\infty e^{-(\rho-n_i-(1-\eta)\bar{g}_i+\delta_s)(s-t)}\zeta\chi\Delta_i\lambda^\tau_{is}ds \\
&\text{with} & \lambda^\tau_{it} &= \int_t^\infty e^{-(\rho-n_i-(1-\eta)\bar{g}_i+\zeta)(s-t)}(\tau_{is}-\tau^\star_i)(1+(\alpha^\gamma-1)\mathbb{1}_{\{\tau_{is}<\tau^\star_i\}})\big[\gamma^y y_{is} + \gamma^u c_{is}\big]\lambda^w_{is}ds \\
& & \lambda^S_{it} &\xrightarrow[\zeta\to\infty]{} -\int_t^\infty e^{-(\rho-n_i-(1-\eta)\bar{g}_i+\delta_s)(s-t)}\chi\Delta_i(\tau_{is}-\tau^\star_i)\big[\gamma^y y_{is} + \gamma^u c_{is}\big]\lambda^w_{is}ds\ ,
\end{aligned}
$$

with output is $y_{it} = z_i \mathcal{D}^y_i(\tau_{it}) F(k_{it}, e_{it})$ and $\lambda^w_{it} = \mathcal{D}^u_i(\tau_{it}) u'(\mathcal{D}^u_i(\tau_{it}) c_{it})$. Use the Euler equation, or costate dynamics:

$$\dot{\lambda}^w_{it} = \lambda^w_{it}(\rho + \eta\bar{g}_i - r^\star_t) \qquad \Rightarrow \qquad \lambda^w_{it} = \lambda^w_{is} e^{-\int_t^s (\rho+\eta\bar{g}_i - r^\star_s)du}$$

for $s > t$, which gives the Local Cost of Carbon:

$$LCC_{it} = -\frac{\lambda^S_{it}}{\lambda^w_{it}} \to \int_t^\infty e^{-(\rho-n_i-(1-\eta)\bar{g}_i+\delta_s)(s-t)} \chi\Delta_i(\tau_{is}-\tau^\star_i)\big[\gamma^y y_{is} + \gamma^u c_{is}\big] e^{+\int_t^s(\rho+\eta\bar{g}_i - r^\star_s)du} ds \ ,$$
$$LCC_{it} = \int_t^\infty e^{-\delta_s(s-t) - \int_t^s (r^\star_u - n_i - \bar{g}_i)du} \chi\Delta_i(\tau_{is}-\tau^\star_i)\big[\gamma^y y_{is} + \gamma^u c_{is}\big] ds$$

This implies that the future damages are discounted faster if $r^\star_{it} > n_i + \bar{g}_i$. Conversely, if growth rate of population n_i and TFP $\bar{g}_i$ are high compared to the world interest rate – think of developing economies – then they would put more weights on future damages on output and consumption per capita.

C Quantitative model - First-Best

First-Best allocation results from the global welfare maximization of the planner, who has access to all the instruments:

$$\mathcal{W}_{t_0} = \max_{\{c,k,e^f,e^c e^r,e^x,\bar{e}^c,\bar{e}^r\}} \sum_{\mathbb{I}} \mathcal{P}_i\, \omega_i \int_{t_0}^\infty e^{-(\rho-n_i-(1-\eta)\bar{g}_i)t}\, u\big(\mathcal{D}^u_i(\tau_{it})\, c_{it}\big)\ dt$$

subject to the good and energy resource constraints and the climate system:

$$\sum_{i\in\mathbb{I}} \mathcal{P}_i e^{(n_i+\bar{g}_i)t}\Big[c_{it} + (\dot{k}_{it} + (n_i+\bar{g}_i+\delta)k_{it}) + \nu^f_i(e^x_{it}, \mathcal{R}_{it}) + \kappa^c_i e^c_{it} + \kappa^r_{it} e^r_{it}\Big] = \sum_{i\in\mathbb{I}} \mathcal{P}_i e^{(n_i+\bar{g}_i)t} \mathcal{D}_i(\tau_{it}) z_{it} F(k_{it}, e^f_{it}, e^c_{it}, e^r_{it}) \qquad [\phi^w_t]$$

$$E^f_{it} = \sum_{i\in\mathbb{I}} \mathcal{P}_i e^{(n_i+\bar{g}_i)t}\, e^f_{it} = \sum_{i\in\mathbb{I}} e^{(n_i+\bar{g}_i)t} e^x_{it} \qquad [\mu^f_t] \qquad \bar{e}^c_i = e^c_i \qquad [\mu^c_{it}] \qquad \bar{e}^r_i = e^r_i \qquad [\mu^r_{it}]$$

$$\dot{\mathcal{S}}_t = \mathcal{E}_t - \delta_s \mathcal{S}_t \qquad \mathcal{E}_t := \sum_{\mathbb{I}} \mathcal{P}_i e^{(n_i+\bar{g}_i)t}(\xi^f e^f_{it} + \xi^c e^c_{it}) \qquad [\phi^S_t]$$

$$\dot{\tau}_{it} = \zeta(\Delta_i \chi \mathcal{S}_t - (\tau_{it} - \tau_{it_0})) \qquad [\phi^\tau_{it}]$$

Let us define several objects: total population $\mathcal{P}_t = \sum_{i\in\mathbb{I}} \mathcal{P}_i e^{n_i t}$, $\mathcal{P} = \mathcal{P}_{t_0} = \sum_i \mathcal{P}_i$ and global population growth rate:

$$n_t = \tfrac{1}{\mathcal{P}_t} \sum_{i\in\mathbb{I}} n_i \mathcal{P}_i e^{n_i t}$$

and the welfare-relevant discount rate:

$$\mathcal{P} e^{-\int_{t_0}^t \bar{\rho}_s ds} = \sum_{i\in\mathbb{I}} \omega_i \mathcal{P}_i e^{-(\rho-n_i-(1-\eta)\bar{g}_i)t}$$
$$\bar{\rho}_t = \frac{1}{\mathcal{P}} \sum_{i\in\mathbb{I}} (\rho-n_i-(1-\eta)\bar{g}_i)\omega_i \mathcal{P}_i e^{-(\rho-n_i-(1-\eta)\bar{g}_i)t - \int_{t_0}^t \bar{\rho}_s ds} = \frac{\sum_{i\in\mathbb{I}}(\rho-n_i-(1-\eta)\bar{g}_i)\omega_i\mathcal{P}_i e^{-(\rho-n_i-(1-\eta)\bar{g}_i)t}}{\sum_{i\in\mathbb{I}} \omega_i \mathcal{P}_i e^{-(\rho-n_i-(1-\eta)\bar{g}_i)t}}$$

Note, all the analysis can accomodate time-varying population growth rate n_{it} and time-varying TFP growth $\bar{g}_{it}$.

First-Best Optimal Control Problem – Pontryagin Principle

The Hamiltonian for the Social planner is:

$$\mathcal{H}^{fb}(\boldsymbol{s},\boldsymbol{c},\boldsymbol{\phi}) = \sum_{\mathbb{I}} \omega_i\, \mathcal{P}_i\, e^{-(\rho-n_i-(1-\eta)\bar{g}_i)t}\, u\big(\mathcal{D}_i^u(\tau_{it})\, c_{it}\big)$$
$$+e^{-\int_{t_0}^t \bar{\rho}_s ds}\phi_t^w \sum_{i\in\mathbb{I}} \mathcal{P}_i e^{(n_i+\bar{g}_i)t}\Big(\mathcal{D}_i(\tau_{it}) z_{it} F(k_{it}, e_{it}^f, e_{it}^c, e_{it}^r) - (n_i+\bar{g}_i+\delta)k_{it} - e^{\nu_i(n_i+\bar{g}_i)t}\nu_i^f(e_{it}^x, \mathcal{R}_{it}) - \kappa_i^c e_{it}^c - \kappa_{it}^r e_{it}^r - c_{it}\Big)$$
$$+ e^{-\int_{t_0}^t \bar{\rho}_s ds}\mu_t^f \sum_{i\in\mathbb{I}} \mathcal{P}_i e^{(n_i+\bar{g}_i)t}\Big(e_{it}^x - e_{it}^f\Big)$$
$$+ e^{-\int_{t_0}^t \bar{\rho}_s ds}\phi_t^S \Big\{ \sum_{\mathbb{I}} \mathcal{P}_i e^{(n_i+\bar{g}_i)t}\big(\xi^f e_{it}^f + \xi^c e_{it}^c\big) - \delta_s \mathcal{S}_t\Big\}$$
$$+ \sum_{i\in\mathbb{I}} \omega_i\, \mathcal{P}_i\, e^{-(\rho-n_i-(1-\eta)\bar{g}_i)t}\phi_{it}^\tau\ \zeta\big(\Delta_i \chi \mathcal{S}_t - (\tau_{it} - \tau_{it_0})\big)$$

Pontryagin maximum principle, optimality conditions, first for controls $\{c, e^f, e^c, e^r, e^x\}$

- Consumption $[c_{it}]$

$$\omega_i\, \mathcal{P}_i\, e^{-(\rho-n_i-(1-\eta)\bar{g}_i)t}\, u_c\big(\mathcal{D}_i^u(\tau_{it})\, c_{it}\big) = e^{-\int_{t_0}^t \bar{\rho}_s ds}\phi_t^w \mathcal{P}_i e^{(n_i+\bar{g}_i)t}$$
$$\omega_i\, e^{-(\rho+\eta\bar{g}_i)t+\int_{t_0}^t \bar{\rho}_s ds}\, \mathcal{D}_i^u(\tau_{it}) u'\big(\mathcal{D}_i^u(\tau_{it})\, c_{it}\big) = \phi_t^w$$

- Energy sources $[e_{it}^f]$, $[e_{it}^c]$, $[e_{it}^r]$

$$\phi_t^w MPe_{it}^f = \mathcal{D}_i(\tau_{it}) z_{it} F_{e^f}(k_{it}, e_{it}^f, e_{it}^c, e_{it}^r) = \mu_t^f - \xi^f \phi_t^S$$
$$MPe_{it}^c = \mathcal{D}_i(\tau_{it}) z_{it} F_{e^c}(k_{it}, e_{it}^f, e_{it}^c, e_{it}^r) = \kappa_i^c - \xi^c \phi_t^S \qquad\qquad MPe_{it}^r \;\; = \kappa_i^r$$

- Energy extraction $[e_{it}^x]$

$$\phi_t^w e^{\nu_i(n_i+\bar{g}_i)t}\nu_{i\,e^x}^f(e_{it}^x, \mathcal{R}_{it}) = \mu_t^f$$
$$e^{\nu_i(n_i+\bar{g}_i)t}\bar{\nu}_i\Big(\frac{e_{it}^x}{\mathcal{R}_{it}}\Big)^{\nu_i} = \frac{\mu_t^f}{\phi_t^w}$$

  note that e^x is the extraction rate per effective capita: with population/TFP growth, the marginal cost become larger.
- Note that we simplify the problem by avoiding treating e_{it}^c and $\bar{e}_{it}^c$ and e_{it}^r and $\bar{e}_{it}^r$ as separate variables.

Pontryagin maximum principle, optimality conditions, second for states $\{k_{it}, \mathcal{S}_t, \tau_{it}\}_{it}$

- Capital $[k_i]$

$$\dot{\phi}_t^w = \phi_t^w \bar{\rho}_t - \mathcal{H}_k^{fb}(\boldsymbol{s},\boldsymbol{c},\boldsymbol{\lambda}) = \phi_t^w\bar{\rho}_t + \phi_t^w(n_i+\bar{g}_i) - \phi_t^w(MPk_{it} - \delta)$$
$$\dot{\phi}_t^w = \phi_t^w\big(\bar{\rho}_t + n_i + \bar{g}_i - (MPk_{it} - \delta)\big)$$

Note if that there is only one country, we get $\bar{\rho}_t = \rho - n_i - (1-\eta)\bar{g}_i$ and then we obtain the standard Euler equation $\dot{\phi}^w_t = \phi^w_t(\rho + \eta\bar{g}_i - r^k_{it})$, with $r^k_{it} = MPk_{it} - \delta$

- Carbon concentration in atmosphere $[\mathcal{S}_t]$

$$\dot{\phi}^S_t = \phi^S_t(\bar{\rho}_t + \delta_s) - \sum_{i\in\mathbb{I}} \omega_i \mathcal{P}_i e^{-(\rho-n_i-(1-\eta)\bar{g}_i)t+\int_{t_0}^t \bar{\rho}_s ds} \zeta\Delta_i\chi\ \phi^\tau_{it}$$

- Temperature $[\tau_{it}]$, normalized by $e^{-(\rho-n_i-(1-\eta)\bar{g}_i)t}$ local discounting

$$\dot{\phi}^\tau_{it} = \phi^\tau_{it}(\rho - n_i - (1-\eta)\bar{g}_i) - \mathcal{H}^{fb}_\tau(\boldsymbol{s}, \boldsymbol{c}, \boldsymbol{\lambda})$$

$$\dot{\phi}^\tau_{it} = \phi^\tau_{it}(\rho - n_i - (1-\eta)\bar{g}_i + \zeta) + \omega_i u'(\mathcal{D}^u_i(\tau_{it})\, c_{it})\mathcal{D}^u_i(\tau_{it})(\tau_{it} - \tau^\star_i)\gamma^u c_{it} + e^{-\int_{t_0}^t \bar{\rho}_s ds + (n_i+g_i)t + (\rho-n_i-(1-\eta)\bar{g}_i)t} \phi^w_t(\tau_{it} - \tau^\star_i)\gamma^y y_{it}$$

Proof of Proposition 6

Solving for the shadow value of temperature ϕ^τ_{it} and carbon ϕ^S_t

$$\begin{aligned}
\dot{\phi}^\tau_{it} &= \phi^\tau_{it}(\rho - n_i - (1-\eta)\bar{g}_i + \zeta) + e^{-\int_{t_0}^t \bar{\rho}_s ds + (\rho+\eta\bar{g}_i)t} \phi^w_t(\tau_{it} - \tau^\star_i)[\gamma^y y_{it} + \gamma^u c_{it}] \\
\phi^\tau_{it_0} &= \int_{t_0}^\infty e^{-(\rho-n_i-(1-\eta)\bar{g}_i+\zeta)(t-t_0)} e^{-\int_{t_0}^t \bar{\rho}_s ds + (\rho+\eta\bar{g}_i)t} \phi^w_t(\tau_{it} - \tau^\star_i)[\gamma^y y_{it} + \gamma^u c_{it}] \\
\phi^\tau_{it_0} &= \int_{t_0}^\infty e^{-\int_{t_0}^t \bar{\rho}_s ds - (\zeta - n_i - \bar{g}_i)(t-t_0)} \phi^w_t(\tau_{it} - \tau^\star_i)[\gamma^y y_{it} + \gamma^u c_{it}]
\end{aligned}$$

Solving for the SCC_i, we get:

$$\begin{aligned}
\phi^S_{t_0} &= -\int_{t_0}^\infty e^{-\int_{t_0}^t \bar{\rho}_s ds - \delta_s(t-t_0)} \sum_{i\in\mathbb{I}} \omega_i \mathcal{P}_i e^{-(\rho-n_i-(1-\eta)\bar{g}_i)(t-t_0)+\int_{t_0}^t \bar{\rho}_s ds} \zeta\Delta_i\chi\ \phi^\tau_{it} dt \\
\phi^S_{t_0} &= -\int_{t_0}^\infty e^{-\delta_s(t-t_0)} \sum_{i\in\mathbb{I}} \omega_i \mathcal{P}_i e^{-(\rho-n_i-(1-\eta)\bar{g}_i)(t-t_0)} \zeta\Delta_i\chi\ \phi^\tau_{it} dt \\
\phi^S_{t_0} &\xrightarrow[\zeta\to\infty]{} -\int_{t_0}^\infty e^{-\delta_s(t-t_0)} \sum_{i\in\mathbb{I}} \omega_i \mathcal{P}_i e^{-(\rho-n_i-(1-\eta)\bar{g}_i)(t-t_0)}\ \chi\Delta_i(\tau_{it} - \tau^\star_i)[\gamma^y y_{it} + \gamma^u c_{it}]\phi^w_t dt\ ,
\end{aligned}$$

Realizing that the marginal value of wealth:

$$\dot{\phi}^w_t = \phi^w_t(\bar{\rho}_t + n_i + \bar{g}_i - \underbrace{(MPk_{it} - \delta)}_{=r^k_{it}}) \qquad \Rightarrow \qquad \phi^w_{t_0} = \phi^w_t e^{-\int_{t_0}^t (\bar{\rho}_s + n_i + \bar{g}_i - r^k_{is})ds}$$

This implies that the Social Cost of Carbon defined as $SCC_{t_0} = -\frac{\phi^S_{t_0}}{\phi^w_{t_0}}$ can be rewritten as:

$$-\phi^S_{t_0} \xrightarrow[\zeta\to\infty]{} \int_{t_0}^\infty e^{-\delta_s(t-t_0)} \sum_{i\in\mathbb{I}} \omega_i \mathcal{P}_i e^{-(\rho-n_i-(1-\eta)\bar{g}_i)(t-t_0)}\ \chi\Delta_i(\tau_{it} - \tau^\star_i)[\gamma^y y_{it} + \gamma^u c_{it}]\phi^w_{t_0} e^{-\int_{t_0}^t (r^k_{is} - \bar{\rho}_s - n_i - \bar{g}_i)ds} dt\ ,$$

$$SCC_{t_0} \to \int_{t_0}^\infty \sum_{i\in\mathbb{I}} \omega_i \mathcal{P}_i e^{-[(\rho-n_i-(1-\eta)\bar{g}_i)(t-t_0) - \int_{t_0}^t \bar{\rho}_s ds]}\ e^{-\delta_s(t-t_0) - \int_{t_0}^t (r^k_{is} - n_i - \bar{g}_i)ds}\ \chi\Delta_i(\tau_{it} - \tau^\star_i)[\gamma^y y_{it} + \gamma^u c_{it}]dt\ ,$$

The aggregate discount factor is defined as $\sum_{i\in\mathbb{I}}\omega_i\mathcal{P}_i e^{-(\rho-n_i-(1-\eta)\bar{g}_i)(t-t_0)} = \mathcal{P}e^{-\int_{t_0}^{t}\bar{\rho}_s ds}$, and given that c_{it}, y_{it} and τ_{it} are bounded, we can simplify the expression. Moreover, changing the order of the sum and integral by Fubini's theorem, we obtain:

$$
\begin{aligned}
SCC_{t_0} &\to \int_{t_0}^{\infty}\sum_{i\in\mathbb{I}}\omega_i\mathcal{P}_i e^{-\delta_s(t-t_0)-\int_{t_0}^{t}(r^k_{is}-n_i-\bar{g}_i)ds}\,\chi\Delta_i(\tau_{it}-\tau_i^\star)\big[\gamma^y y_{it}+\gamma^u c_{it}\big]dt\ ,\\
SCC_{t_0} &\to \sum_{i\in\mathbb{I}}\omega_i\mathcal{P}_i\int_{t_0}^{\infty} e^{-\delta_s(t-t_0)-\int_{t_0}^{t}(r^k_{is}-n_i-\bar{g}_i)ds}\,\chi\Delta_i(\tau_{it}-\tau_i^\star)\big[\gamma^y y_{it}+\gamma^u c_{it}\big]dt\ ,\\
SCC_{t_0} &\to \sum_{i\in\mathbb{I}}\omega_i\mathcal{P}_i\, LCC_{it}
\end{aligned}
$$

where the Local Cost of Carbon LCC_{it} is given in Appendix B. Given that, in the competitive equilibrium, we have free capital flows and frictionless borrowing, it implies $r^k_{it} = MPk_{it}-\delta = r^\star_t$, which gives the results of Proposition 6. □

D Quantitative model - Second-Best

Second-Best allocation results from the global welfare maximization of the planner, subject to choice of a global carbon tax, t^ε, and local lump-sum rebate: $\mathrm{t}^{ls}_{it} = \mathrm{t}^\varepsilon(\xi^f e^f_{it} + \xi^c e^c_{it})$.

$$
\mathcal{W}_{t_0} = \max_{\{c,k,e^f,e^c e^r,e^x,\bar{e}^c,\bar{e}^r\}}\sum_{\mathbb{I}}\mathcal{P}_i\,\omega_i\int_{t_0}^{\infty} e^{-(\rho-n_i-(1-\eta)\bar{g}_i)t}\,u\big(\mathcal{D}^u_i(\tau_{it})\,c_{it}\big)\ dt
$$

subject to the good and energy resource market clearing and the climate system:

$$
\begin{aligned}
\dot{w}_{it} = (r^\star_t - (n_i{+}\bar{g}_i))w_{it} + \pi^f_i(q^f_t,\mathcal{R}_{it}) + \mathcal{D}^y_i(\tau_{it})z_i F(k_{it},e^f_{it},e^c_{it},e^r_{it}) - (r^\star_t+\delta)k_{it}&\\
- (q^f_t + \xi^f \mathrm{t}^\varepsilon_{it})e^f_{it} - (q^c_{it} + \xi^c\mathrm{t}^\varepsilon_{it})e^c_{it} - q^r_{it}e^r_{it} - c_{it} + \mathrm{t}^{ls}_{it}\ ,&\qquad [\psi^w_{it}]
\end{aligned}
$$

$$
E^f_{it} = \sum_{i\in\mathbb{I}}\mathcal{P}_i e^{(n_i+\bar{g}_i)t}\,e^f_{it} = \sum_{i\in\mathbb{I}} e^{(n_i+\bar{g}_i)t}e^x_{it}\qquad [\mu^f_t]\qquad\qquad B_t = \sum_{i\in\mathbb{I}}\mathcal{P}_i e^{(n_i+\bar{g}_i)t}(w_{it}-k_{it}) = 0\qquad [\mu^b_t]
$$

$$
\bar{e}^c_i = e^c_i\qquad [\mu^c_{it}]\qquad\qquad \bar{e}^r_i = e^r_i\qquad [\mu^r_{it}]
$$

$$
\dot{\mathcal{S}}_t = \mathcal{E}_t - \delta_s\mathcal{S}_t\qquad\qquad \mathcal{E}_t := \sum_{\mathbb{I}}\mathcal{P}_i e^{(n_i+\bar{g}_i)t}(\xi^f e^f_{it} + \xi^c e^c_{it})\qquad\qquad [\psi^S_t]
$$

$$
\dot{\tau}_{it} = \zeta\big(\Delta_i\chi\mathcal{S}_t - (\tau_{it}-\tau_{it_0})\big)\qquad\qquad [\psi^\tau_{it}]
$$

as well as the optimality conditions of the agents of the Competitive equilibrium

$$
\begin{aligned}
&[k_t] && MPk_{it} = r^\star_t + \delta && [\upsilon^k_{it}]\\
&[x_t] && MPe^x_{it} = \mathcal{D}^y_i(\tau_{it})z_i\ \partial_x F(k_{it},e^f_{it},e^c_{it},e^r_{it}) = q^x_{it} \qquad \text{for } x\in\{f,c,r\} && [\upsilon^x_{it}]\\
&[e^x_t] && q^f_t = \nu^f_{i\,e^x}(e^x_{it},\mathcal{R}_{it}) && [\theta^x_{it}]
\end{aligned}
$$

Using the Primal approach, we can write the Hamiltonian, where the states are $\{w_{it},\mathcal{S}_t,\tau_{it}\}_{it}$, and the controls are $\{c_{it},b_{it},k_{it},e^f_{it},e^c_{it},e^r_{it},e^x_{it},\bar{e}^c_{it},\bar{e}^r_{it}\}_{it}$, and prices $\{r^\star_t,q^f_t,\mathrm{w}_{it},q^c_{it},q^r_{it}\}_{it}$ and where each country i variable is discounted by $\rho - n_i - (1-\eta)\bar{g}_i$:

$$\begin{aligned}
\mathcal{H}^{sb}(\boldsymbol{s}, \boldsymbol{c}, \boldsymbol{\psi}) &= \sum_{\mathbb{I}} \omega_i \, \mathcal{P}_i \, e^{-(\rho - n_i - (1-\eta)\bar{g}_i)t} \, u\big(\mathcal{D}_i^u(\tau_{it}) \, c_{it}\big) \\
&+ \sum_{i\in\mathbb{I}} \psi_{it}^w \omega_i \, \mathcal{P}_i e^{-(\rho - n_i - (1-\eta)\bar{g}_i)t} \Big((r_t^\star - (n_i + \bar{g}_i)) w_{it} + \mathcal{D}_i(\tau_{it}) z_{it} F(k_{it}, e_{it}^f, e_{it}^c, e_{it}^r) \\
&\qquad + \pi_i^f(q_t^f, \mathcal{R}_{it}) - (r^\star + \delta) k_{it} - (q_t^f + \xi^f \mathrm{t}_{it}^\varepsilon) e_{it}^f - (q_{it}^c + \xi^c \mathrm{t}_{it}^\varepsilon) e_{it}^c - q_{it}^r e_{it}^r - c_{it} + \mathrm{t}_{it}^{ls} \Big) \\
&\qquad + \sum_{i\in\mathbb{I}} \mathcal{P}_i e^{-(\rho - n_i - (1-\eta)\bar{g}_i)t} [\mu_{it}^c (\bar{e}_{it}^c - e_{it}^c) + \mu_{it}^r (\bar{e}_{it}^r - e_{it}^r)] \\
&\qquad + e^{-\int_{t_0}^t \bar{\rho}_s ds} \mu_t^f \sum_{i\in\mathbb{I}} \mathcal{P}_i e^{(n_i + \bar{g}_i)t} \Big(e_{it}^x - e_{it}^f \Big) + e^{-\int_{t_0}^t \bar{\rho}_s ds} \mu_t^b \sum_{i\in\mathbb{I}} \mathcal{P}_i e^{(n_i + \bar{g}_i)t} \Big(w_{it} - k_{it} \Big) \\
&\qquad + e^{-\int_{t_0}^t \bar{\rho}_s ds} \psi_t^S \Big\{ \sum_{\mathbb{I}} \mathcal{P}_i e^{(n_i + \bar{g}_i)t} (\xi^f e_{it}^f + \xi^c e_{it}^c) - \delta_s \mathcal{S}_t \Big\} \\
&\qquad + \sum_{i\in\mathbb{I}} \omega_i \, \mathcal{P}_i \, e^{-(\rho - n_i - (1-\eta)\bar{g}_i)t} \psi_{it}^\tau \, \zeta \big(\Delta_i \chi \mathcal{S}_t - (\tau_{it} - \tau_{it_0}) \big) \\
&+ \sum_{i\in\mathbb{I}} \omega_i \, \mathcal{P}_i e^{-(\rho - n_i - (1-\eta)\bar{g}_i)t} \Big[\upsilon_{it}^f (q_t^f + \xi^f \mathrm{t}^\varepsilon - MPe_{it}^f) + \upsilon_{it}^c (q_{it}^c + \xi^c \mathrm{t}^\varepsilon - MPe_{it}^c) + \upsilon_{it}^r (q_{it}^r - MPe_{it}^r) + \upsilon_{it}^k (r_t^\star + \delta - MPk_{it}) \Big] \\
&+ \sum_{i\in\mathbb{I}} \omega_i \, \mathcal{P}_i e^{-(\rho - n_i - (1-\eta)\bar{g}_i)t} \Big[\theta_{it}^x (\nu_{i\, e^x}^f (e_{it}^x, \mathcal{R}_{it}) - q_t^f) + \theta_{it}^c (\kappa_{it}^c - q_{it}^c) + \theta_{it}^r (\kappa_{it}^r - q_{it}^r) \Big]
\end{aligned}$$

PMP: Optimality conditions for the controls $\{c_{it}, b_{it}, k_{it}, e_{it}^f, e_{it}^c, e_{it}^r, e_{it}^x, \bar{e}_{it}^c, \bar{e}_{it}^r\}_{it}$ are:

- Consumption:

$$\omega_i \, \mathcal{P}_i u\big(\mathcal{D}_i^u(\tau_{it}) \, c_{it}\big) = \psi_{it}^w \omega_i \, \mathcal{P}_i$$

- Capital choice:

$$\begin{aligned}
\omega_i \, \mathcal{P}_i \, e^{-(\rho - n_i - (1-\eta)\bar{g}_i)t} \psi_{it}^w [MPk_{it} - \delta - r_t^\star] - e^{-\int_{t_0}^t \bar{\rho}_s ds} \mu_t^b \sum_{i\in\mathbb{I}} \mathcal{P}_i e^{(n_i + \bar{g}_i)t} \\
-\omega_i \mathcal{P}_i e^{-(\rho - n_i - (1-\eta)\bar{g}_i)t} [\upsilon_{it}^f \partial_k MPe_{it}^f + \upsilon_{it}^c \partial_k MPe_{it}^c + \upsilon_{it}^r \partial_k MPe_{it}^r + \upsilon_{it}^k \partial_k MPk_{it}] = 0
\end{aligned}$$

$$\mu_t^b = -e^{-[(\rho + \eta \bar{g}_i)t - \int_{t_0}^t \bar{\rho}_s ds]} \omega_i [\upsilon_{it}^f \partial_k MPe_{it}^f + \upsilon_{it}^c \partial_k MPe_{it}^c + \upsilon_{it}^r \partial_k MPe_{it}^r + \upsilon_{it}^k \partial_k MPk_{it}]$$

  The multiplier μ_t^b represents the shadow value of liquidity of aggregate bonds. If we increased bond supply B_t, it would decrease the interest rate and improve the ability of firms to borrow and invest, decreasing the marginal value of capital. This redistributive effect has an impact on the firm inputs optimality conditions for input x, written with $\upsilon_{it}^x \partial_k MPx_{it}$. As a result, μ_t^b is the equilibrium value equalizing these different redistributive/distortive effects.

- Energy extraction – Oil-gas (Fossil) $[e_{it}^x]$

$$\begin{aligned}
\omega_i \, \mathcal{P}_i \, e^{-(\rho - n_i - (1-\eta)\bar{g}_i)t} \psi_{it}^w [q_t^f - \nu_{i\, e^x}^f (e_{it}^x, \mathcal{R}_{it})] + e^{-\int_{t_0}^t \bar{\rho}_s ds} \mu_t^f \mathcal{P}_i e^{(n_i + \bar{g}_i)t} \\
+ \omega_i \, \mathcal{P}_i \, e^{-(\rho - n_i - (1-\eta)\bar{g}_i)t} \theta_{it}^x \mathcal{P}_i \nu_{i\, e^x e^x}^f (e_{it}^x, \mathcal{R}_{it})
\end{aligned}$$

$$\mu_t^f = -e^{-[(\rho + \eta \bar{g}_i)t - \int_{t_0}^t \bar{\rho}_s ds]} \omega_i \, \theta_{it}^x \mathcal{P}_i e^{(n_i + \bar{g}_i)} \nu_{i\, e^x e^x}^f (e_{it}^x, \mathcal{R}_{it})$$

The multiplier μ_t^f is the shadow value of liquidity of aggregate oil-gas supply. If we increased supply E_t^f, it would decrease the oil-gas price rate q_t^f going down the supply curve, as denoted by the curvature $\nu^f_{i\,e^x e^x}$ and weighted by the shadow value of optimality of the fossil firm's extraction. Note also that we scaling the curvature of the cost by population $\mathcal{P}_i e^{(n_i+\bar{g}_i)t}$ – population growth push extraction along the supply curve.

- Energy production (Coal and renewable)

$$\begin{aligned} \omega_i\,\mathcal{P}_i\,e^{-(\rho-n_i-(1-\eta)\bar{g}_i)t}\mu_{it}^c &= 0 \\ \omega_i\,\mathcal{P}_i\,e^{-(\rho-n_i-(1-\eta)\bar{g}_i)t}\mu_{it}^r &= 0 \end{aligned}$$

  There's no redistribution effects across countries through the market clearing, due to the fact that (i) the coal (and renewable) are traded locally, and (ii) there are no profits from coal and renewable production.

- Energy consumption – Oil-gas (Fossil)

$$\begin{aligned} \omega_i\,\mathcal{P}_i\,e^{-(\rho-n_i-(1-\eta)\bar{g}_i)t}\psi_{it}^w[MPe_{it}^f-q_t^f] - e^{-\int_{t_0}^t\bar{\rho}_s ds}\mu_t^f\mathcal{P}_i e^{(n_i+\bar{g}_i)t} + \xi^f e^{-\int_{t_0}^t\bar{\rho}_s ds}\psi_t^S\mathcal{P}_i e^{(n_i+\bar{g}_i)t} + \\ -\omega_i\mathcal{P}_i e^{-(\rho-n_i-(1-\eta)\bar{g}_i)t}\left[\upsilon_{it}^f\partial_{e^f}MPe_{it}^f+\upsilon_{it}^c\partial_{e^f}MPe_{it}^c+\upsilon_{it}^r\partial_{e^f}MPe_{it}^r+\upsilon_{it}^k\partial_{e^f}MPk_{it}\right]=0 \end{aligned}$$

$$\omega_i\,\mathcal{P}_i\,e^{-(\rho+\eta\bar{g}_i)t+\int_{t_0}^t\bar{\rho}_s ds}\psi_{it}^w\xi^f\mathrm{t}_t^\varepsilon = \mathcal{P}_i\mu_t^f - \mathcal{P}_i\xi^f\psi_t^S + \omega_i\mathcal{P}_i e^{-(\rho+\eta\bar{g}_i)t+\int_{t_0}^t\bar{\rho}_s ds}\left[\upsilon_{it}^f\partial_{e^f}MPe_{it}^f+\upsilon_{it}^c\partial_{e^f}MPe_{it}^c+\upsilon_{it}^r\partial_{e^f}MPe_{it}^r+\upsilon_{it}^k\partial_{e^f}\right.$$

- Energy consumption – Coal

$$\begin{aligned} \omega_i\,\mathcal{P}_i\,e^{-(\rho-n_i-(1-\eta)\bar{g}_i)t}\psi_{it}^w[MPe_{it}^c-q_{it}^c] - \omega_i\,\mathcal{P}_i\,e^{-(\rho-n_i-(1-\eta)\bar{g}_i)t}\mu_{it}^c + \xi^c e^{-\int_{t_0}^t\bar{\rho}_s ds}\psi_t^S\mathcal{P}_i e^{(n_i+\bar{g}_i)t} + \\ -\omega_i\mathcal{P}_i e^{-(\rho-n_i-(1-\eta)\bar{g}_i)t}\left[\upsilon_{it}^f\partial_{e^c}MPe_{it}^f+\upsilon_{it}^c\partial_{e^c}MPe_{it}^c+\upsilon_{it}^r\partial_{e^c}MPe_{it}^r+\upsilon_{it}^k\partial_{e^f}MPk_{it}\right]=0 \end{aligned}$$

$$\omega_i\,\mathcal{P}_i\,e^{-(\rho+\eta\bar{g}_i)t+\int_{t_0}^t\bar{\rho}_s ds}\psi_{it}^w\xi^c\mathrm{t}_t^\varepsilon = -\mathcal{P}_i\xi^c\psi_t^S + \omega_i\mathcal{P}_i e^{-(\rho+\eta\bar{g}_i)t+\int_{t_0}^t\bar{\rho}_s ds}\left[\upsilon_{it}^f\partial_{e^c}MPe_{it}^f+\upsilon_{it}^c\partial_{e^c}MPe_{it}^c+\upsilon_{it}^r\partial_{e^c}MPe_{it}^r+\upsilon_{it}^k\partial_{e^c}MPk_i\right.$$

- Energy consumption – Renewable

$$\begin{aligned} \omega_i\,\mathcal{P}_i\,e^{-(\rho-n_i-(1-\eta)\bar{g}_i)t}\psi_{it}^w[MPe_{it}^c-q_{it}^r] - \omega_i\,\mathcal{P}_i\,e^{-(\rho-n_i-(1-\eta)\bar{g}_i)t}\mu_{it}^r + \\ -\omega_i\mathcal{P}_i e^{-(\rho-n_i-(1-\eta)\bar{g}_i)t}\left[\upsilon_{it}^f\partial_{e^r}MPe_{it}^f+\upsilon_{it}^c\partial_{e^r}MPe_{it}^c+\upsilon_{it}^r\partial_{e^r}MPe_{it}^r+\upsilon_{it}^k\partial_{e^f}MPk_{it}\right]=0 \\ \Rightarrow \quad \left[\upsilon_{it}^f\partial_{e^c}MPe_{it}^f+\upsilon_{it}^c\partial_{e^c}MPe_{it}^c+\upsilon_{it}^r\partial_{e^c}MPe_{it}^r+\upsilon_{it}^k\partial_{e^c}MPk_{it}\right]=0 \end{aligned}$$

PMP: Optimality conditions for the controls over prices $\{r_t^\star, q_t^f, \mathrm{w}_{it}, q_{it}^c, q_{it}^r\}_{it}$

- Interest rate $[r_t^\star]$

$$\sum_i \omega_i\,\mathcal{P}_i\,e^{-(\rho-n_i-(1-\eta)\bar{g}_i)t}\psi_{it}^w[w_{it}-k_{it}] + \sum_i \omega_i\,\mathcal{P}_i\,e^{-(\rho-n_i-(1-\eta)\bar{g}_i)t}\upsilon_{it}^k = 0$$

  The redistributive effect on agents' budget, weighted by ψ_{it}^w should compensate for the distortionary effect on firms' optimality of capital, weighted by shadow value υ_{it}^k.

- Fossil energy price: $[q_t^f]$

$$\sum_i \omega_i \, \mathcal{P}_i \, e^{-(\rho - n_i - (1-\eta)\bar{g}_i)t} \psi_{it}^w [e_{it}^x - e_{it}^f] + \sum_i \omega_i \, \mathcal{P}_i \, e^{-(\rho - n_i - (1-\eta)\bar{g}_i)t} [\upsilon_{it}^f - \theta_{it}^x] = 0$$

- Coal energy price: $[q_{it}^c]$

$$\omega_i \, \mathcal{P}_i \, e^{-(\rho - n_i - (1-\eta)\bar{g}_i)t} \psi_{it}^w [\bar{e}_i^c - e_{it}^c] + \omega_i \, \mathcal{P}_i \, e^{-(\rho - n_i - (1-\eta)\bar{g}_i)t} [\upsilon_{it}^c - \theta_{it}^c] = 0$$
$$\Rightarrow \qquad \upsilon_{it}^c = \theta_{it}^c$$

- Renewable energy price: $[q_{it}^r]$, similarly:

$$\Rightarrow \qquad \upsilon_{it}^r = \theta_{it}^r$$

  For coal and renewable, since the price is local, the "distortive/redistributive" effect on the supply equals the one of its demand.
- Wages w_{it} are determined directly by the Marginal Product of Labor $MP\ell_{it}$, since the labor supply is inelastic and normalized to 1.
- Carbon tax/Carbon price $[\mathrm{t}^\varepsilon]$

$$\sum_i \omega_i \, \mathcal{P}_i \, e^{-(\rho - n_i - (1-\eta)\bar{g}_i)t} [\xi^f \upsilon_{it}^f + \xi^c \upsilon_{it}^c] = 0$$

  The optimal carbon tax level balances out all the distortions for each country, for oil-gas and coal, according to shadow values υ_{it}^f and υ_{it}^c.

PMP: Optimality conditions for the states $\{w_{it}, \mathcal{S}_t, \tau_{it}\}_{it}$.

- Wealth $[w_{it}]$

$$\begin{aligned}
\dot{\psi}_{it}^w &= \psi_{it}^w (\rho - n_i - (1-\eta)\bar{g}_i) - \mathcal{H}_w^{sb}(\boldsymbol{s}, \boldsymbol{c}, \boldsymbol{\psi}) \\
\dot{\psi}_{it}^w &= \psi_{it}^w (\rho + \eta \bar{g}_i - r_t^\star) + \frac{1}{\omega_i \mathcal{P}_i} e^{-\int_{t_0}^t \bar{\rho}_s ds + (\rho + \eta \bar{g}_i)t} \mu_t^b \\
\mu_t^b &= e^{-[(\rho + \eta \bar{g}_i)t - \int_{t_0}^t \bar{\rho}_s ds]} \omega_i [\upsilon_{it}^f \partial_k MPe_{it}^f + \upsilon_{it}^c \partial_k MPe_{it}^c + \upsilon_{it}^r \partial_k MPe_{it}^r + \upsilon_{it}^k \partial_k MPk_{it}] \\
\dot{\psi}_{it}^w &= \psi_{it}^w (\rho + \eta \bar{g}_i - r_t^\star) + e^{-\int_{t_0}^t \bar{\rho}_s ds + (\rho + \eta \bar{g}_i)t} e^{-[(\rho + \eta \bar{g}_i)t - \int_{t_0}^t \bar{\rho}_s ds]} \frac{1}{\mathcal{P}_i} [\dots] \\
\dot{\psi}_{it}^w &= \psi_{it}^w (\rho + \eta \bar{g}_i - r_t^\star) + \frac{1}{\mathcal{P}_i} [\upsilon_{it}^f \partial_k MPe_{it}^f + \upsilon_{it}^c \partial_k MPe_{it}^c + \upsilon_{it}^r \partial_k MPe_{it}^r + \upsilon_{it}^k \partial_k MPk_{it}]
\end{aligned}$$

  This implies time-varying liquidity motives for the marginal value of wealth. Abstracting from discounting $\rho - n_i - (1-\eta)\bar{g}_i$ and $\bar{\rho}_t$, if μ_t^b is positive (the planner would like to increase the supply of bond, decreasing return), it needs to be compensated for higher marginal value of wealth ψ_{it}^w in the future ($\dot{\psi}_{it}^w$ is higher if $\mu_t^b > 0$) which implies higher consumption, today at time t.

- Temperature $[\tau_{it}]$

$$\dot{\psi}^{\tau}_{it} = \psi^{\tau}_{it}(\rho - n_i - (1-\eta)\bar{g}_i) - \mathcal{H}^{sb}_{\tau}(\boldsymbol{s}, \boldsymbol{c}, \boldsymbol{\psi})$$

$$\dot{\psi}^{\tau}_{it} = \psi^{\tau}_{it}(\rho - n_i - (1-\eta)\bar{g}_i \ + \zeta) + \underbrace{\gamma^y(\tau_{it} - \tau^{\star}_i)\mathcal{D}^y_i(\tau_{it})}_{-\partial_\tau \mathcal{D}^y} z_i F(k_{it}, e_{it})\psi^w_{it} + \underbrace{\gamma^u(\tau_{it} - \tau^{\star}_i)\mathcal{D}^u_i(\tau_{it})}_{-\partial_\tau \mathcal{D}^u} u'(\mathcal{D}^u(\tau_{it})c_{it})c_{it}$$

$$\dot{\psi}^{\tau}_{it} = \psi^{\tau}_{it}(\rho - n_i - (1-\eta)\bar{g}_i \ + \zeta) + (\tau_{it} - \tau^{\star}_i)[\gamma^y y_{it} + \gamma^u c_{it}]\psi^w_{it}$$

- Carbon concentration $[\mathcal{S}_t]$

$$\dot{\psi}^S_t = \psi^S_t \bar{\rho}_t - \mathcal{H}^{sb}_{\tau}(\boldsymbol{s}, \boldsymbol{c}, \boldsymbol{\psi})$$

$$\dot{\psi}^S_t = \psi^S_t(\bar{\rho}_t + \delta_s) - \sum_{i \in \mathbb{I}} \omega_i \, \mathcal{P}_i \, e^{-(\rho - n_i - (1-\eta)\bar{g}_i)t + \int_{t_0}^t \bar{\rho}_s ds} \zeta \Delta_i \chi \, \psi^{\tau}_{it}$$

Proof of Proposition 8

Solving for the differential equations for the marginal value of temperature ψ^{τ}_{it} and carbon ψ^S_{it}.

$$\psi^{\tau}_{it} = \int_t^{\infty} e^{-(\rho - n_i - (1-\eta)\bar{g}_i + \zeta)(s-t)}(\tau_{is} - \tau^{\star}_i)\left[\gamma^y y_{is} + \gamma^u c_{is}\right]\psi^w_{is} ds$$

$$\psi^S_t = -\int_t^{\infty} e^{-\delta_s(s-t) - \int_t^s \bar{\rho}_s ds} \sum_{i \in \mathbb{I}} \omega_i \, \mathcal{P}_i \, e^{-(\rho - n_i - (1-\eta)\bar{g}_i)(s-t) + \int_t^s \bar{\rho}_s ds} \zeta \Delta_i \chi \, \psi^{\tau}_{is} ds$$

$$\psi^S_t = -\int_t^{\infty} e^{-\delta_s(s-t)} \sum_{i \in \mathbb{I}} \omega_i \, \mathcal{P}_i \, e^{-(\rho - n_i - (1-\eta)\bar{g}_i)(s-t)} \zeta \Delta_i \chi \, \psi^{\tau}_{is} ds$$

$$\psi^S_t \xrightarrow[\zeta \to \infty]{} -\int_t^{\infty} e^{-\delta_s(s-t)} \sum_{i \in \mathbb{I}} \omega_i \, \mathcal{P}_i \, e^{-(\rho - n_i - (1-\eta)\bar{g}_i)(s-t)} \Delta_i \chi \, (\tau_{is} - \tau^{\star}_i)\left[\gamma^y y_{is} + \gamma^u c_{is}\right]\psi^w_{is} ds$$

Using the dynamics for the marginal value of wealth;

$$\psi^w_{it} = \int_t^T e^{-(\rho + \eta\bar{g}_i)(s-t) + \int_t^s r^{\star}_u du} \bar{\upsilon}^k_{is} ds + e^{-(\rho + \eta\bar{g}_i)(T-t) + \int_t^T r^{\star}_u du} \psi^w_{iT}$$

$$\bar{\upsilon}^k_{it} = \frac{1}{\mathcal{P}_i}\left[\upsilon^f_{it}\partial_k MPe^f_{it} + \upsilon^c_{it}\partial_k MPe^c_{it} + \upsilon^r_{it}\partial_k MPe^r_{it} + \upsilon^k_{it}\partial_k MPk_{it}\right]$$

First, let us assume that $\bar{\upsilon}^k_{it} \approx 0$ there are no liquidity effects, giving $\psi^w_{it} = e^{-(\rho + \eta\bar{g}_i)(T-t) + \int_t^T r^{\star}_u du} \psi^w_{iT}$.

I define the social welfare weights:

$$\overline{\psi}^w_t = \frac{1}{\mathcal{P}_t} \sum_{i \in \mathbb{I}} \omega_i \mathcal{P}_i e^{-[(\rho - n_i - (1-\eta)\bar{g}_i)t - \int_{t_0}^t \rho_s ds]} \psi^w_{it}$$

$$\widehat{\psi}^w_{it} = \frac{\omega_i \mathcal{P}_i \psi^w_{it}}{\overline{\psi}^w_t}$$

This allow to simplify the marginal value of carbon:

$$\psi_t^S \to -\int_t^\infty e^{-\delta_s(s-t)} \sum_{i\in\mathbb{I}} \omega_i\, \mathcal{P}_i\, e^{-(\rho-n_i-(1-\eta)\bar{g}_i)(s-t)} \Delta_i \chi\left(\tau_{is}-\tau_i^\star\right)\left[\gamma^y y_{is} + \gamma^u c_{is}\right]\psi_{is}^w ds$$

$$\psi_t^S \to -\int_t^\infty e^{-\delta_s(s-t)} \sum_{i\in\mathbb{I}} \omega_i\, \mathcal{P}_i\, e^{-(\rho-n_i-(1-\eta)\bar{g}_i)(s-t)} \Delta_i \chi\left(\tau_{is}-\tau_i^\star\right)\left[\gamma^y y_{is} + \gamma^u c_{is}\right] e^{+(\rho+\eta\bar{g}_i)(s-t)-\int_t^s r_u^\star du}\psi_{it}^w ds$$

This implies the Social Cost of Carbon:

$$SCC_t = -\frac{\psi_t^S}{\overline{\psi}_t^w} = \int_t^\infty e^{-\delta_s(s-t)-\int_t^s r_u^\star du} \sum_{i\in\mathbb{I}} e^{(n_i+\bar{g}_i)(s-t)} \Delta_i \chi\left(\tau_{is}-\tau_i^\star\right)\left[\gamma^y y_{is} + \gamma^u c_{is}\right] \frac{\omega_i\, \mathcal{P}_i\, \psi_{it}^w}{\overline{\psi}_t^w} ds$$

$$SCC_t = \sum_{i\in\mathbb{I}} \widehat{\psi}_{it}^w LCC_{it} \qquad\qquad \widehat{\psi}_{it}^w = \frac{\omega_i\, \mathcal{P}_i\, \psi_{it}^w}{\overline{\psi}_t^w}$$

$$LCC_{it} = \int_t^\infty e^{-\delta_s(s-t)-\int_t^s (r_u^\star - n_i - \bar{g}_i)du}\, \Delta_i \chi\left(\tau_{is}-\tau_i^\star\right)\left[\gamma^y y_{is} + \gamma^u c_{is}\right] ds$$

which gives the results of Proposition 8. □

Proof of Proposition 9

Solving for the optimal carbon tax involves for solving for the objects in the optimality condition for energy choices. Take the energy choice for fossil-fuels:

$$\omega_i\, \mathcal{P}_i\, e^{-(\rho+\eta\bar{g}_i)t}\psi_{it}^w \xi^f \mathrm{t}_t^\varepsilon = e^{-\int_{t_0}^t \bar{\rho}_s ds}\mathcal{P}_i \mu_t^f - e^{-\int_{t_0}^t \bar{\rho}_s ds}\mathcal{P}_i \xi^f \psi_t^S$$
$$+\omega_i \mathcal{P}_i e^{-(\rho+\eta\bar{g}_i)t}\left[\upsilon_{it}^f \partial_{e^f} MPe_{it}^f + \upsilon_{it}^c \partial_{e^f} MPe_{it}^c + \upsilon_{it}^r \partial_{e^f} MPe_{it}^r + \upsilon_{it}^k \partial_{e^f} MPk_{it}\right]$$

We need to solve each of the objects in turn: (i) the marginal value of carbon, ψ_t^S, related to the Social Cost of Carbon, as seen in the previous proposition, (ii) the marginal value of oil supply μ_t^f, related to the energy supply redistribution, (iii) the marginal value of firms' inputs optimality conditions $\upsilon_{it}^f, \upsilon_{it}^c, \upsilon_{it}^r, \upsilon_{it}^k$.

<u>*Supply redistribution*</u>

We want to solve for μ_t^f. First, take the optimality for $[q_t^f]$.

$$\sum_i \omega_i\, \mathcal{P}_i\, e^{-(\rho-n_i-(1-\eta)\bar{g}_i)t}\theta_{it}^x = \sum_i \omega_i\, \mathcal{P}_i\, e^{-(\rho-n_i-(1-\eta)\bar{g}_i)t}\psi_{it}^w[e_{it}^x - e_{it}^f] + \sum_i \omega_i\, \mathcal{P}_i\, e^{-(\rho-n_i-(1-\eta)\bar{g}_i)t}\upsilon_{it}^f$$

The redistribution effect on oil-gas quantity between exporter and importers $e_{it}^x - e_{it}^f$ exactly compensate – for the planner – the equilibrium effect on demand υ_{it}^f and supply θ_{it}^x.

Second, using the optimality of $[e_{it}^x]$:

$$\mu_t^f = -e^{-[(\rho+\eta\bar{g}_i)t-\int_{t_0}^t \bar{\rho}_s ds]}\omega_i\, \theta_{it}^x\, \mathcal{P}_i e^{(n_i+\bar{g}_i)t}\, \nu^f_{i\, e^x e^x}(e_{it}^x, \mathcal{R}_{it})$$

$$\omega_i\, \mathcal{P}_i \theta_{it}^x = -\frac{\mathcal{P}_i e^{-(n_i+\bar{g}_i)t}}{\nu^f_{i\, e^x e^x}(e_{it}^x, \mathcal{R}_{it})} e^{-[(\rho+\eta\bar{g}_i)t-\int_{t_0}^t \bar{\rho}_s ds]}\mathcal{P}_i \mu_t^f$$

As a result, we obtain:

$$\sum_i \omega_i \, \mathcal{P}_i \, e^{-(\rho-n_i-(1-\eta)\bar{g}_i)t}\theta^x_{it} = -\mu^f_t \sum_i \frac{\mathcal{P}_i e^{(n_i+\bar{g}_i)t}}{\nu^f_{i\,e^x e^x}(e^x_{it},\mathcal{R}_{it})} e^{(\rho+\eta\bar{g}_i)t-\int_{t_0}^t \bar{\rho}_s ds} e^{-(\rho-n_i-(1-\eta)\bar{g}_i)t}$$

$$\sum_i \omega_i \, \mathcal{P}_i \, e^{-(\rho-n_i-(1-\eta)\bar{g}_i)t}\theta^x_{it} = -e^{-\int_{t_0}^t \bar{\rho}_s ds}\mu^f_t \sum_i \frac{1}{\nu^f_{i\,e^x e^x}(e^x_{it},\mathcal{R}_{it})}$$

$$\Rightarrow \qquad e^{-\int_{t_0}^t \bar{\rho}_s ds}\mu^f_t = -\Big[\sum_i \nu^f_{i\,e^x e^x}(e^x_{it},\mathcal{R}_{it})^{-1}\Big]^{-1} \sum_i \omega_i \, \mathcal{P}_i \, e^{-(\rho-n_i-(1-\eta)\bar{g}_i)t}\theta^x_{it}$$

$$\Rightarrow \qquad e^{-\int_{t_0}^t \bar{\rho}_s ds}\mu^f_t = \Big[\sum_i \nu^f_{i\,e^x e^x}(e^x_{it},\mathcal{R}_{it})^{-1}\Big]^{-1} \Big\{\sum_i \omega_i \, \mathcal{P}_i \, e^{-(\rho-n_i-(1-\eta)\bar{g}_i)t}\psi^w_{it}[e^f_{it}-e^x_{it}] - \sum_i \omega_i \, \mathcal{P}_i \, e^{-(\rho-n_i-(1-\eta)\bar{g}_i)t}\upsilon^f_{it}\Big\}$$

<u>*Demand distortion*</u>

We use a Nested CES framework for production. I express the formula without time indices for simplicity.

$$\text{Energy} \qquad e_i = \Big(\sum_\ell (\omega^\ell)^{\frac{1}{\sigma_e}} (e^\ell_i)^{\frac{\sigma_e-1}{\sigma_e}}\Big)^{\frac{\sigma_e}{\sigma_e-1}} \qquad\qquad \text{Output} \qquad y_i = z_i\Big((1-\varepsilon)^{\frac{1}{\sigma}}(z^e_i\,e_i)^{\frac{\sigma-1}{\sigma}} + \varepsilon^{\frac{1}{\sigma}}(k^\alpha_i \ell^{1-\alpha}_i)^{\frac{\sigma-1}{\sigma}}\Big)^{\frac{\sigma}{\sigma-1}}$$

Optimality for fossil energy demand features this complicated term, which we can simplify using the CES structure. The curvature of production, represented by terms like $\partial_{e^f} MPe^f_i$ are related the the elasticity of energy use.

$$\bar{\upsilon}^f_i = [\upsilon^f_i \partial_{e^f} MPe^f_i + \upsilon^c_i \partial_{e^f} MPe^c_i + \upsilon^r_i \partial_{e^f} MPe^r_i + \upsilon^k_i \partial_{e^f} MPk_i]$$

$$= \frac{1}{e^f_i}\Big[-\upsilon^f_i(q^f+\xi^f \mathrm{t}^\varepsilon)[\frac{1-s^f}{\sigma^e} + s^f\frac{1-s^e}{\sigma^y}] + \upsilon^c_i(q^c_i+\xi^f\mathrm{t}^\varepsilon)s^f_i[\frac{1}{\sigma^e} - \frac{1-s^e}{\sigma^y}] + \upsilon^r_i q^r_i s^f_i[\frac{1}{\sigma^e} - \frac{1-s^e}{\sigma^y}] + \upsilon^k_i(r^\star+\bar{\delta})\frac{s_i^{e^r/y}}{\sigma^y}\Big]$$

with Energy share in production: $s^e_i = \frac{e_i q^e_i}{y_i}$, Fossil share in energy mix $s^f_i = \frac{e^f_i q^f}{e_i q^e_i}$ and similarly $s^c_i = \frac{e^c_i q^c_i}{e_i q^e_i}$ and $s^r_i = \frac{e^r_i q^r_i}{e_i q^e_i}$.

When we normalize by $\overline{\psi}$, we obtain:

$$\widehat{\bar{\upsilon}}^f_i = \frac{1}{e^f_i}\Big[-\widehat{\upsilon}^f_i(q^f+\xi^f \mathrm{t}^\varepsilon)[\frac{1-s^f}{\sigma^e} + s^f\frac{1-s^e}{\sigma^y}] + \widehat{\upsilon}^c_i(q^c_i+\xi^f\mathrm{t}^\varepsilon)s^f_i[\frac{1}{\sigma^e} - \frac{1-s^e}{\sigma^y}] + \widehat{\upsilon}^r_i q^r_i s^f_i[\frac{1}{\sigma^e} - \frac{1-s^e}{\sigma^y}] + \widehat{\upsilon}^k_i(r^\star+\delta)\frac{s_i^{e^r/y}}{\sigma^y}\Big]$$

We can obtain similar formulas for $\widehat{\bar{\upsilon}}^f_{it}$, $\widehat{\bar{\upsilon}}^c_{it}$, $\widehat{\bar{\upsilon}}^r_{it}$, and $\widehat{\bar{\upsilon}}^k_{it}$.

Rewriting the optimal carbon tax

Take the optimality condition for energy choice, where I replaced $\bar{\upsilon}^f_{it}$. Sum over countries i and normalize by world population $\mathcal{P}_t$

$$\omega_i\,\mathcal{P}_i\,e^{-(\rho-n_i-(1-\eta)\bar{g}_i)t}\psi^w_{it}\xi^f\mathrm{t}^\varepsilon_t = e^{-\int_{t_0}^t\bar{\rho}_s ds}\mathcal{P}_i e^{n_i+\bar{g}_i}\mu^f_t - e^{-\int_{t_0}^t\bar{\rho}_s ds}\mathcal{P}_i\xi^f\psi^S_t + \omega_i\mathcal{P}_i e^{-(\rho-n_i-(1-\eta)\bar{g}_i)t}\,\bar{\upsilon}^f_{it}$$

$$\xi^f\mathrm{t}^\varepsilon_t\sum_{i\in\mathbb{I}}\omega_i\,\mathcal{P}_i\,e^{-(\rho-n_i-(1-\eta)\bar{g}_i)t}\psi^w_{it} = e^{-\int_{t_0}^t\bar{\rho}_s ds}\underbrace{\sum_i\mathcal{P}_i e^{(n_i+\bar{g}_i)t}}_{=\mathcal{P}_t}[\mu^f_t - \xi^f\psi^S_t] + \sum_i\omega_i\mathcal{P}_i e^{-(\rho-n_i-(1-\eta)\bar{g}_i)t}\,\bar{\upsilon}^f_{it}$$

$$\xi^f\mathrm{t}^\varepsilon_t\overline{\psi}^w_t = [\mu^f_t - \xi^f\psi^S_t] + \sum_i\omega_i\mathcal{P}_i e^{-(\rho-n_i-(1-\eta)\bar{g}_i)t+\int_{t_0}^t\bar{\rho}_s ds}\,\bar{\upsilon}^f_{it}$$

We define:

$$\begin{aligned}
\textit{Supply Redistribution}_t = \frac{\mu^f_t}{\overline{\psi}^w_t} &= \Big[\sum_i\nu^f_{i\,e^xe^x}(e^x_{it},\mathcal{R}_{it})^{-1}\Big]^{-1}\Big\{\sum_i\omega_i\,\mathcal{P}_i\,e^{-(\rho-n_i-(1-\eta)\bar{g}_i)t+\int_{t_0}^t\bar{\rho}_s ds}\frac{\psi^w_{it}}{\overline{\psi}^w_t}[e^x_{it}-e^f_{it}] \\
&\quad + \sum_i\omega_i\,\mathcal{P}_i\,e^{-(\rho-n_i-(1-\eta)\bar{g}_i)t+\int_{t_0}^t\bar{\rho}_s ds}\frac{\upsilon^f_{it}}{\overline{\psi}^w_t}\Big\} \\
&= \Big[\sum_i(\nu^f_{i\,e^xe^x})^{-1}\Big]^{-1}\Big\{\sum_i\widehat{\psi}^w_{it}[e^x_{it}-e^f_{it}] + \sum_i\widehat{\upsilon}^f_{it}\Big\} \\
\textit{Demand Distortion}_t = \frac{1}{\overline{\psi}^w_t}\sum_i\omega_i\mathcal{P}_i e^{-(\rho-n_i-(1-\eta)\bar{g}_i)t+\int_{t_0}^t\bar{\rho}_s ds}\,\bar{\upsilon}^f_{it} &= \sum_i\widehat{\bar{\upsilon}}^f_{it}
\end{aligned}$$

We obtain:

$$\begin{aligned}
\xi^f\mathrm{t}^\varepsilon_t &= \xi^f SCC_t + \Big[\sum_i(\nu^f_{i\,e^xe^x})^{-1}\Big]^{-1}\Big\{\sum_i\widehat{\psi}^w_{it}[e^x_{it}-e^f_{it}] + \sum_i\widehat{\upsilon}^f_{it}\Big\} + \sum_i\widehat{\bar{\upsilon}}^f_{it} \\
\xi^f\mathrm{t}^\varepsilon_t &= \xi^f SCC_t + \textit{Supply Redistribution}_t + \textit{Demand Distortion}_t
\end{aligned}$$

This implies the formula in Proposition 9 □.

The Winners and Losers of Climate Policies:
A Sufficient Statistics Approach

Most recent version

Thomas Bourany, Jordan Rosenthal-Kay*
The University of Chicago

14 March 2025

Abstract

To combat global warming, climate policies like carbon taxes, renewable subsidies, and carbon tariffs need to be implemented to phase out fossil-fuel consumption and lower emissions. Who are the winners and losers of such policies? Through a simple Integrated Assessment Model with heterogeneous countries and international trade, we study the costs of climate change through local damages and trade spillovers in international goods and energy markets. We study both the costs of implementing those policies unilaterally, and the local costs and global gains of international policy cooperation. To do so, we express and decompose these welfare changes to first order as a function of *sufficient statistics*, depending on observables and identifiable elasticities, like nations' energy mix, energy rents, trade shares, supply and demand elasticities, and damage parameters. We show that climate change has non-trivial reallocation effects through international trade in goods and energy. Pursuing unilateral policies generates strong leakage effects in goods and energy markets that are an order of magnitude larger than the gains due to reduced emissions. Finally, global climate policy cooperation has a large impact on energy markets, affecting mostly countries reliant on coal and fossil-fuel producers, causing larger welfare losses for those countries than the original costs of climate change.

*Thomas Bourany, *thomasbourany@uchicago.edu*, Jordan Rosenthal-Kay, *jrosenthalkay@uchicago.edu*.
We thank our advisors, Esteban Rossi-Hansberg, Mikhail Golosov, and Michael Greenstone, for their valuable advice and guidance. We also thank Rodrigo Adão, Jonathan Dingel, and Wojciech Kopczuk for interesting discussions. All errors are our own.

1 Introduction

Climate policies must be implemented to phase out fossil fuel consumption and keep the world temperature under $2°C$ to avoid dramatic consequences of global warming (IPCC et al., 2022). To that aim, several policies have been proposed. First, carbon taxation or carbon pricing has been the preferred instrument of economists. It follows from the Pigouvian benchmark, where the externality – and the social cost of carbon emissions – can be internalized by taxation. Second, several countries have been promoting subsidies for renewable energy sources as an alternative to carbon policy, with, for example, the Inflation Reduction Act in the United States. However, carbon taxation may engender "carbon leakage", as economic activity reallocates to other trade partners unaffected by the policy. As a result, countries have also been advocating for international cooperation through climate agreements.

Who are the winners and losers of different climate policies? Can we quantify which countries have the largest losses from climate damages and corrective climate policies such as carbon taxes, renewable energy subsidies, or carbon tariffs? Moreover, how large are the gains from cooperation for the distribution of welfare gains and losses? This paper addresses these questions by quantifying the heterogeneous impacts of those policies and decomposing the welfare impacts for different countries and across different transmission channels.

We develop a framework that allows us to quantify these different margins using sufficient statistics, in the sense of Chetty (2009): a set of observable data moments and elasticities estimated using quasi-experimental variation. In our framework, observables like the energy mix (i.e., the share of oil, gas, coal, and renewables in energy use), the energy rent share of GDP, and trade shares, provide crucial information on whether a country 'wins' or 'loses' as a result of implementing those policies. Using the structure of a quantitative model, we can summarize and decompose the welfare effects of policy: changes in productivity due to climate damages, and welfare changes stemming from general equilibrium effects in energy and goods markets.

To that purpose, we use a climate-economy framework – or Integrated Assessment Model (IAM) – augmented with heterogeneous countries, energy markets, and international trade, closely following Bourany (2025). Individual countries differ in their vulnerability to climate change and temperature, their energy mix in oil, gas, coal, and non-carbon energy, their costs of producing fossil fuels, as well as trade costs in international trade in goods. We approximate the model using a first-order, log-linear decomposition of welfare changes around the current – i.e., status-quo – equilibrium. This allows us to linearly break down the various channels through which climate change and climate policies affect different nations.

These climate policies have unequal impacts across countries, determining their willingness to implement such policies. First, regions are differentially affected by climate change due to differences in local temperature, exposure to global warming, or trade linkages with vulnerable countries. Second, if carbon taxation reduces fossil-fuel consumption, it also has substantial impacts on energy markets: it affects disproportionately the countries that consume a large share of fossil fuels – oil, gas, and coal – or that export those energy sources. Third, countries are heterogeneously exposed to international trade and thus to 'carbon leakage', which reinforces or dampens the gains and losses from the climate policies, especially when implemented unilaterally.

Quantifying the winners and losers from trade policies with our framework – and understanding the underlying mechanisms – requires several key elasticities, in addition to readily available moments in international trade, energy and national accounts data. First, we require estimates of the marginal damages of temperature shocks in different countries on a structural primitive of our model: TFP in traded goods. To identify the parameters of our structural damage function, we implement an estimation strategy inspired by Rudik et al. (2022) that leverages variation in import penetration within bilateral trading partners and changes in local temperature. Second, we require energy supply elasticities in oil-, gas-, and coal-producing sectors. We use time-series variation in local fossil rents and international prices and a simple empirical Bayes shrinkage procedure to recover spatially heterogeneous energy supply elasticities.

Armed with these elasticities and data moments, we use our sufficient statistics formula to study three different experiments for a sample of 193 countries. First, we analyze the effects of increasing greenhouse gas emissions and, hence, global temperature. This has heterogeneous impacts across locations due to differences in temperatures. However, climate change has large spillovers: by changing TFP, it affects production and the endogenous choice of energy inputs. As a result, declines in productivity are the main driver of reduction in CO_2 emissions. Moreover, climate change also reallocates production unequally across international trade partners. Therefore, productivity spillovers through trade represent an important transmission mechanism of global warming across countries, as even cold but open regions can be affected significantly through trade channels.

Second, we study the impacts of four different climate policies. (i) First, we consider the case where each country increases unilaterally the carbon tax, and the revenue of the carbon tax is rebated to the household. As a result, different countries are differently affected through their energy consumption and exports, but the climate impact in terms of emission reduction is limited by the size of each country. Moreover, unilateral carbon taxation leads to carbon leakage effects: it reduces domestic fossil fuel demand through taxation and thus also lowers the global equilibrium price for oil and gas, which then increases the carbon emissions of the countries not affected by the carbon tax. Additionally, by increasing the marginal cost domestically, it also reallocates activity through international trade from regions not affected by the carbon policy. Combining and decomposing all these effects, we can measure the extent of the 'free-riding incentives' that deter individual governments' climate action.

Then, (ii) we consider renewable energy subsidies as an alternative to carbon taxation. Such subsidies lower the relative price of renewable – compared to fossil fuels – and are financed by lump-sum taxes. As a result, it has different general equilibrium implications on energy markets, as well as different welfare effects across countries. Moreover, (iii) to prevent the carbon leakage consequences of unilateral carbon pricing, trade instruments have been at the center of policy discussions, e.g. with the Carbon Border Adjustment Mechanism in the European Union. In a third policy, we study the implementation of carbon tariffs, where the tariff scales with the carbon intensity of the imports and carbon taxes for countries that form a climate club. Carbon tariffs divert trade flows away from high emissions countries and reduce carbon emissions, which we can quantify for a small increase in the carbon price for imports. We analyze this policy for two sets of countries: the European Union (EU), which is already implementing carbon pricing and carbon tariffs, and ASEAN, which gathers southeast Asian countries affected by climate change.

Finally, (iv), we study the distribution of welfare gains and losses from implementing internationally coordinated policies. When all the countries implement carbon taxation together, greenhouse gas emissions are lowered significantly, improving climate and global temperature. In addition, the demand and, hence, the price of fossil fuels change depending on the strength of the substitution between oil, gas, and coal. Such change in oil and gas prices also has strong redistributive effects, as it depletes the energy rents for fossil-fuel exporters. Finally, global climate policy reallocates economic activity and trade patterns, and all these effects attenuate the direct costs of carbon taxation.

This work relates to a lengthy literature on the macroeconomics economics of climate change, specifically the literature that uses large-scale IAMs to evaluate the cost of climate change and the effects of different policies (Nordhaus and Yang, 1996; Barrage and Nordhaus, 2024; Cruz and Rossi-Hansberg, 2024). Our main contribution is to show that, in macroeconomic IAM of Bourany (2025), the effects of many climate policy regimes can be decomposed to the first order to direct effects, effects that operate through the energy market (including changes to energy rents), and leakage effects in international goods markets, and that these effects can be estimated by a set of sufficient statistics readily computable with off-the-shelf macroeconomic data and estimable elasticities. Thus, our work is similar to Lashkaripour (2021), who uses a sufficient statistics approach to estimate the cost of a global trade war, Baqaee and Farhi (2024), who examine how changes to trade barriers reallocate economic activity in general equilibrium, and Kleinman et al. (2024) who derive sufficient statistics for how productivity shocks differentially affect trading partners in constant-elasticity trade models. In essence, we extend this approach to a broad set of climate policy instruments in an environment that also features detailed energy markets, like in Abuin (2024). Additionally, our framework also allows us to derive a local cost of carbon, which accounts for heterogeneity in both damages and the marginal utility of income, like in Cruz and Rossi-Hansberg (2022).

What we do not do is study optimal climate policy at a global level, as in Golosov et al. (2014), or unilaterally optimal policy in an open economy setting, like in Kortum and Weisbach (2021). Unlike Bourany (2025), who studies the optimal design of international climate agreements with carbon taxation and tariffs, or Farrokhi and Lashkaripour (2024), who study the optimal trade policy, either unilaterally or in the context of climate clubs, we exploit the tractability of our sufficient statistics formula to evaluate a large set of climate policies and decompose their effects that operate through different markets. However, our framework is static, which precludes us from studying dynamic policy environments, like in Bourany (2024) who analyze climate policy and redistribution concerns, Hsiao (2022), who studies climate coordination with commitment, or Kotlikoff et al. (2021a) who study the benefit of carbon taxation across generations.

The rest of this paper is organized as follows. In Section 2, we lay out our full macroeconomic IAM. Section 3 derives our first-order decomposition of climate policies in our model and details the policy experiments we have in mind. We describe our data, estimation, and quantification in Section 4. Section 5 details our results for multiple policy counterfactuals, and Section 6 concludes.

2 An integrated assessment model with heterogeneous regions and trade

This model follows the structure of Bourany (2025). We build a simple integrated assessment model (IAM) incorporating multiple dimensions of heterogeneity, climate externality, energy markets, and a realistic trade structure that reproduces the leakage effects of climate policies.

We study a static economy with I countries indexed by $i \in \mathbb{I}$, each with population $\mathcal{P}_i$. All the economic variables are expressed per capita.[1] Each country is composed of five representative agents: (i) a household that consumes the final goods, (ii) a final-good firm producing goods using labor and energy, (iii) a fossil energy firm extracting oil and gas, (iv) a producer of coal energy, and (v) a producer of renewable/non-carbon energy. Moreover, each country has a government that sets taxes, subsidy, and tariffs.

2.1 Household problem

The representative household in country i imports from all countries $j \in \mathbb{I}$ and consumes the aggregate quantity c_i. Each country produces its own good variety. Household preferences have constant elasticity of substitution θ over different varieties, following an Armington structure (Anderson, 1979; Arkolakis, Costinot and Rodriguez-Clare, 2012),

$$\mathcal{U}_i = \max_{\{c_{ij}\}} u\left(\{c_{ij}\}_j\right) = u\left(c_i\right) \ , \qquad c_i = \left(\sum_{j\in\mathbb{I}} a_{ij}^{\frac{1}{\theta}} c_{ij}^{\frac{\theta-1}{\theta}}\right)^{\frac{\theta}{\theta-1}} \ , \tag{1}$$

where a_{ij} are the preference shifters for country i on the good purchased from country j.[2] We consider standard constant relative risk aversion (CRRA) utility $u(c) = c^{1-\eta}/(1-\eta)$.[3]

Households earn labor income, energy rents, and transfers, and their budget constraints is given by:

$$\sum_{j\in\mathbb{I}} c_{ij}\,(1+\mathrm{t}_{ij}^{b})\tau_{ij}\mathrm{p}_j = w_i\ell_i + \pi_i^e + \ \mathrm{t}_i^{ls} \ , \tag{2}$$

[1] For example, y_i or e_i^f are final output and fossil energy use, respectively, and $\mathcal{P}_i y_i$ and $\mathcal{P}_i e_i^f$ represent the total quantities produced/consumed in the country.

[2] We assume that preferences $\{a_{ij}\}$ and iceberg trade costs $\{\tau_{ij}\}$ are policy-invariant, in particular, they are not sensitive to price changes and tariffs.

[3] We do not include direct effects of climate change on utility, which could proxy for changes to local amenities, or the mortality effects of climate change. It is easy to augment the framework to include direct utility damages by assuming the climate externality affects consumption through a factor $\mathcal{D}_i^u(\mathcal{E})$ which summarizes climate damages, given world emissions $\mathcal{E}$. As labor is internationally immobile, utility damages have no general equilibrium effect, and so we omit them from this analysis. Including utility damages would simply amplify gains and losses stemming from changes in local temperature.

where w_i is the wage rate, ℓ_i the exogenous labor supply is normalized to 1, π_i^e the profit earned from the ownership of the energy firms, and t_i^{ls} the lump-sum transfer received from the government. Households in i imports quantities c_{ij} from country j, purchased at price p_j, and subject to iceberg cost τ_{ij} and to trade tariffs $1+\mathrm{t}_{ij}^b$.

The optimal consumption choice of the household yields the following quantities and trade shares given by:

$$c_{ij} = a_{ij} c_i \left(\frac{(1+\mathrm{t}_{ij}^b)\tau_{ij}\mathrm{p}_j}{\mathbb{P}_i} \right)^{-\theta} , \qquad s_{ij} \equiv \frac{c_{ij} p_{ij}}{c_i \mathbb{P}_i} = a_{ij} \frac{((1+\mathrm{t}_{ij}^b)\tau_{ij}\mathrm{p}_j)^{1-\theta}}{\sum_k a_{ik}((1+\mathrm{t}_{ik})\tau_{ik}\mathrm{p}_k)^{1-\theta}} , \tag{3}$$

where $p_{ij} = (1+\mathrm{t}_{ij}^b)\tau_{ij}\mathrm{p}_j$ is the effective price for a variety from country j sold in country i, and $\mathbb{P}_i$ is the price index of country i:

$$\mathbb{P}_i = \left(\sum_{k \in \mathbb{I}} a_{ik} \big((1+\mathrm{t}_{ik}^b)\tau_{ik}\mathrm{p}_k\big)^{1-\theta} \right)^{\frac{1}{1-\theta}} .$$

As a result, we summarize the budget constraint as $c_i \mathbb{P}_i = \sum_{j\in\mathbb{I}} c_{ij}\,(1+\mathrm{t}_{ij}^b)\tau_{ij}\mathrm{p}_j$, and the per-capita welfare of country i is then summarized by the indirect utility as the utility of income discounted by the price level and climate damages, namely:

$$\mathcal{U}_i = u\,(c_i) = \frac{1}{1-\eta} \left(\frac{w_i \ell_i + \pi_i^e + \mathrm{t}_i^{ls}}{\mathbb{P}_i} \right)^{1-\eta} . \tag{4}$$

2.2 Final good firm problem

The representative final good producer in country i is producing the domestic variety at price p_i. The firm's profit maximization is:

$$\max_{\ell_i, e_i^f, e_i^c, e_i^r} \; \mathrm{p}_i\, \mathcal{D}_i^y(\mathcal{E})\, z_i\, F(\ell_i, e_i^f, e_i^c, e_i^r) - w_i \ell_i - q^f(1+\xi^f \mathrm{t}_i^\varepsilon) e_i^f - q_i^c(1+\xi^c \mathrm{t}_i^\varepsilon) e_i^c - q_i^c - q_i^r(1-\mathrm{s}_i^\varepsilon) e_i^r \tag{5}$$

where the production function $\bar{y}_i = F(\ell_i, e_i^f, e_i^c, e_i^r)$ has constant returns to scale and is concave in all inputs. It uses labor, ℓ_i, at wage w_i, fossil energy, e_i^f, purchased at price, q^f, coal, e^c, at price, q_i^c, and renewable energy, e_i^r, at price, q_i^r. Energy from oil-gas, e_i^f, and coal, e_i^c, differ from renewable in the sense that they emit greenhouse gases, with respective carbon concentration ξ^f and ξ^c, as we will see in 2.4. As a result, there is a motive for taxing oil, gas, and coal energy with the carbon tax t_i^ε. Similarly, as an alternative, we consider renewable energy subsidy, which reduces the price of renewable subsidies by a factor s^ε.

The productivity of the domestic good firm, $y_i = \mathcal{D}_i^y(\mathcal{E})\, z_i\, \bar{y}_i$, can be decomposed in two terms. First, the TFP, z_i, represents productivity as well as institutional/efficiency differences between countries. This technology wedge accounts for income inequality across countries. These differences in TFP translate into differences in consumption that create redistribution motives for tax policy. The second difference in productivity comes from the climate externality. This is summarized by the net-of-damage function $\mathcal{D}^y_i(\mathcal{E})$, given world emissions $\mathcal{E}$, which is also a reduced-form representation of the climate system from temperatures. It decreases in $\mathcal{E}$ and is country-specific due to differences in costs of climate change, as detailed in Section 4.

The firm input decisions solve the optimality conditions, where we define the marginal product of an input x as $MPx_i \equiv \mathcal{D}^y_i(\mathcal{E})\, z_i F_x(\ell_i, e^f_i, e^c_i, e^r_i)$ for $x \in \{\ell_i, e^f_i, e^c_i, e^r_i\}$. For example, in the case of oil and gas e^f_i, the first-order condition can be written as:

$$\mathrm{p}_i \mathcal{D}^y_i(\mathcal{E})\, z_i\, F_{e^f}(\ell_i, e^f_i, e^c_i, e^r_i) =: \mathrm{p}_i MPe^f_i = q^f (1+\xi^f \mathrm{t}^\varepsilon_i)\ , \tag{6}$$

and similarly for other inputs ℓ_i, e^c_i, e^r_i. Crucially, the private decisions of firms do not internalize climate externalities of their own fossil-fuel energy use and only respond to price, carbon tax t^ε_i, and subsidy s^ε_i.

We consider a nested CES production function. The firm combines labor ℓ_i with a composite of energy e_i, with elasticity σ^y.[4] Second, energy e_i aggregates the different energy sources: oil and gas e^f, coal e^c_i, and renewable/non-carbon e^r_i, with elasticity σ^e.

$$\text{Output:} \qquad y_i = \mathcal{D}^y(\mathcal{E})\, z_i \left(\varepsilon^{\frac{1}{\sigma^y}} (e_i)^{\frac{\sigma^y - 1}{\sigma^y}} + (1-\varepsilon)^{\frac{1}{\sigma^y}} (\ell_i)^{\frac{\sigma^y-1}{\sigma^y}} \right)^{\frac{\sigma^y}{\sigma^y - 1}}\ ,$$

$$\text{Energy:} \qquad e_i = \left((\omega^f)^{\frac{1}{\sigma^e}} (e^f_i)^{\frac{\sigma^e-1}{\sigma^e}} + (\omega^c)^{\frac{1}{\sigma^e}} (e^c_i)^{\frac{\sigma^e-1}{\sigma^e}} + (\omega^r)^{\frac{1}{\sigma^e}} (e^r_i)^{\frac{\sigma^e-1}{\sigma^e}} \right)^{\frac{\sigma^e}{\sigma^e-1}}\ .$$

This allows us to distinguish between the substitution across energy sources and between energy and other inputs like labor, due to climate policies.

2.3 Energy markets

The final-good firm consumes three kinds of energy sources – oil-gas, coal, or renewable (non-carbon) energy – which are supplied by three representative energy firms in each country. Oil-gas sources are traded internationally, and countries can be exporters or importers. Coal and renewable sources are both traded locally, an empirically relevant assumption given the substantial trade costs in coal shipping or electricity transfers. The profits all the energy firms $\pi^e_i = \pi^f_i + \pi^c_i + \pi^r_i$ are redistributed lump-sum to the household.

2.3.1 Fossil production

In each country $i \in \mathbb{I}$, a competitive energy producer extracts fossil fuels – oil and gas – e^x_i and sells it to the international market at price q^f. The energy is extracted at convex cost $\mathcal{C}^f_i(e^x_i)$, where the convex costs are paid in the price of the good of country i.[5] The energy firm's profit

[4] Labor is inelastically supplied $\ell_i = \bar{\ell}_i$ in each country and normalized to 1 – since the country size $\mathcal{P}_i$ is already taken into account. As a result, all the variables can be seen as input per capita.

[5] Alternatively, one could assume that the energy firms use labor inputs, which is equivalent given that there is a direct mapping between prices and wages in Armington models.

maximization problem is given by:

$$\pi_i^f = \max_{e_i^x} q^f e_i^x - \mathcal{C}_i^f(e_i^x)\mathrm{p}_i \ , \tag{7}$$

where π_i^f is the fossil energy rent per capita in country i. Since the extraction costs are convex, the production function has decreasing return to scale.[6] Thus, a positive energy rent exists even with competitive firms taking the fossil price as given. For the sake of simplicity, we do not consider that energy firms have market power in the setting of global energy prices as in Bornstein et al. (2023), even though this framework could easily allow for such an extension. We account for misallocation (in the sense of Hsieh and Klenow, 2009) arising due to existing policy distortions that take the form of fossil taxes or subsidies as embedded in extraction productivity in $\mathcal{C}_i^f(\cdot)$, while we capture endogenous misallocation from market power (which can attenuate output elasticities) in our estimates of energy supply elasticities.

Naturally, the optimal extraction decision follows from the optimality condition,

$$q^f = \mathcal{C}_i^{f\,\prime}(e_i^x)\mathrm{p}_i \ , \tag{8}$$

which yields the implicit function $e_i^{x\star} = e^x(q^f/\mathrm{p}_i) = \mathcal{C}_i^{f\prime\,-1}(q^f/\mathrm{p}_i)$. Finally, the energy rent comes from fossil firms' profits $\pi^f(q^f, \mathrm{p}_i) = q^f e^x(q^f/\mathrm{p}_i) - \mathcal{C}_i^f\left(e^x(q^f/\mathrm{p}_i)\right)\mathrm{p}_i > 0$ and depends on the marginal costs as well as the inverse supply elasticity $\nu_i^f = \frac{\mathcal{C}_i^{f\,\prime\prime}(e^x)}{\mathcal{C}_i^{f\,\prime}(e^x)e^x}$. We use the isoelastic extraction function $\mathcal{C}_i^f$,

$$\mathcal{C}_i^f(e_i^x)\mathrm{p}_i = \frac{\bar{\nu}_i^f}{1+\nu_i^f}\left(\frac{e_i^x}{\mathcal{R}_i}\right)^{1+\nu_i^f}\mathcal{R}_i\mathrm{p}_i \ .$$

with $\mathcal{R}_i$ the oil-gas reserves, a fixed factor. This is homogeneous of degree one in $(e_i^x, \mathcal{R}_i)$ and implies a constant elasticity supply function. In this formulation, we think of market power as affecting the supply elasticity by affecting the intensity of reserves in the production of fossil energy. We can write the profit function as,

$$\pi_i^f = \frac{\nu_i^f \bar{\nu}_i^f}{1+\nu_i^f}\left(\frac{e_i^x}{\mathcal{R}_i}\right)^{1+\nu_i^f}\mathcal{R}_i\mathrm{p}_i = \frac{\nu_i^f(\bar{\nu}_i^f)^{-1/\nu_i^f}}{1+\nu_i^f}(q^f)^{1+\frac{1}{\nu_i^f}} \ .$$

As we will see below, the profit π_i^f and its share in income $\eta_i^{\pi f} = \frac{\pi_i{}^f}{y_i\mathrm{p}_i+\pi_i{}^e}$ are key to determine the exposure of a country to carbon taxation. Indeed, reducing carbon emissions by phasing out of fossil fuels reduces energy demand and its price q^f and hence affects energy profit π_i^f and the welfare of large oil and gas exporters.

[6] We can also define a fossil production function with inputs x_i^f such that $e^x = g(x_i^f)$ and profit $\pi = q^f g(x) - x\mathbb{P}_i$ instead of $\pi = q^f e^x - \mathcal{C}(e^x)\mathrm{p}_i$, in which case $g(x) = \mathcal{C}^{-1}(x)$.

2.3.2 International fossil energy markets

We assume that oil and gas are traded frictionlessly in international markets.[7] The market clears such that

$$E^f = \sum_{i\in\mathbb{I}} \mathcal{P}_i e_i^f = \sum_{i\in\mathbb{I}} \mathcal{P}_i e_i^x \ . \tag{9}$$

Countries have different exposure to this fossil energy market. As country i consumes total quantity of fossil fuels $\mathcal{P}_i e_i^f$, produces $\mathcal{P}_i e_i^x$, its net exports of oil-gas are $\mathcal{P}_i(e_i^x - e_i^f) \lesseqgtr 0$.

2.3.3 Coal production

A representative firm in each country produces coal, which is consumed by the final good firm. We differentiate coal from other fossil fuels like oil and gas because coal production typically generates smaller energy rents for producing countries as a share of GDP. Moreover, large coal producers also consume a large fraction of that coal locally, as trade costs for coal transportation are larger. Hence, we make this empirically grounded assumption that coal is not traded.

We again assume production of $\bar{e}_i^c$ is decreasing returns to scale (owing to convex extraction costs) and uses country i final good input. Coal producers' the profit maximization problem is,

$$\pi_i^c = \max_{e_i^c} q_i^c \bar{e}_i^c - \mathcal{C}_i^c(\bar{e}_i^c)\mathrm{p}_i \ ,$$

with the cost function $\mathcal{C}_i^c$, with inverse supply elasticity $\nu_i^c = \frac{\mathcal{C}_i^{c\prime\prime}(e^c)}{\mathcal{C}_i^{c\prime}(e^c)e^c}$. As before, the price for coal and the market clearing condition are given by:

$$q_i^c = \mathcal{C}_i^{c\prime}(\bar{e}_i^c)\mathrm{p}_i \quad , \tag{10}$$

where in equilibrium, $\bar{e}_i^c = e_i^c$. We consider the same isoelastic cost function as for the oil-gas production, with constant inverse elasticity ν_i^c.

2.3.4 Renewable and non-carbon energy production

The final good firm also uses renewable and other low-carbon energy sources, such as solar, wind, or nuclear electricity. A representative firm produces renewable or non-carbon energy, and this supply, $\bar{e}_i^r$, is not traded internationally. This assumption is verified by the fact that electricity is rarely traded across countries – and when it is, it is explained by intermittency rather than structural imbalances. The cost function $\mathcal{C}_i^r(\bar{e}_i^r) = x_i^r$ is paid in country i good at price p_i. Hence,

[7] We refrain from considering a general Armington structure in which each country produces unique, imperfectly substitutable energy varieties. We make the simplifying assumption that fossil fuels produced in different countries are not distinguishable and traded without cost in international markets. That is, crude oil or natural gas from Nigeria, Saudi Arabia, or Russia are perfect substitutes and their movement across borders is costless. In reality, fossil energy has less than an infinite elasticity of substitution due to quality grade differences like sulfur content, and there are trade costs in shipping oil, despite the considerable scale economies in transport on large crude carriers.

the renewable firm maximization problem is:

$$\pi_i^r = \max_{\bar{e}_i^r} q_i^r \bar{e}_i^r - \mathcal{C}_i^r(\bar{e}_i^r)\mathrm{p}_i \ ,$$

with inverse supply elasticity $\nu_i^r = \frac{\mathcal{C}_i^{r\prime\prime}(e^r)}{\mathcal{C}_i^{r\prime}(e^r)e^r}$. As a result, the price of renewable and the market clearing are given by:

$$q_i^r = \mathcal{C}_i^{r\prime}(\bar{e}_i^r)\mathrm{p}_i \ , \qquad \bar{e}_i^r = e^r \ . \tag{11}$$

Again, we consider the same isoelastic cost function as for the oil-gas and coal production, with constant inverse elasticity ν_i^r. When $\nu_i^r = 0$, we have constant return to scale, $\mathcal{C}_i^{r\prime}$ is a constant and zero profits $\pi_i^r = 0$. This, once again, would return a perfectly elastic supply curve, which is a slightly stronger assumption in the context of renewable energy.

2.4 The climate system

Carbon emissions released from the burning of fossil fuels create an externality as they feed back into the atmosphere, increase temperatures and affect damages. Despite the model being static, we incorporate a simple climate system as in standard Integrated Assessment Models.

We model the climate damage affecting country i's productivity with the structural damage function $\mathcal{D}_i^y(\mathcal{E})$, a reduced-form representation of how rising temperatures (and other correlated weather changes) affect a nation's productivity z_i. The structural damage function maps emissions $\mathcal{E}$ to atmospheric carbon concentration $\mathcal{S}$, which fosters a rise in global temperature, and, in turn, local temperatures T_i, which affect output.

The energy choices yield yearly emissions from fossil fuels, which sum up to,

$$\mathcal{E} = \sum_{i \in \mathbb{I}} \mathcal{P}_i(\xi^f e_i^f + \xi^c e_i^c) \ ,$$

where ξ^f and ξ^c represent the carbon concentration of oil-gas and coal, respectively. These emissions increase the carbon concentration in the atmosphere. We use the scalar $\mathbb{T}$ to convert a static – one year – model to a long-term / dynamic representation of climate.

$$\mathcal{S} = \mathcal{S}_0 + \mathbb{T}\mathcal{E}$$

with $\mathcal{S}_0$ the initial carbon concentration before all the policy decisions are made at the beginning of the 21st century. We assume a linear relationship between the cumulative CO_2 emissions $\mathcal{S}$ and the global temperature anomaly $\mathcal{T}$ compared to preindustrial levels.

$$\mathcal{T} = \chi\mathcal{S} = \chi\left(\mathcal{S}_0 + \mathbb{T}\mathcal{E}\right) \ ,$$

where χ is the climate sensitivity parameter, i.e. how much warming a ton of CO_2 causes, and where $\mathcal{E}$ and $\mathcal{S}$ are measured in carbon units. This specification is rationalized by the climate-sciences literature. For example, Dietz et al. (2021) shows an approximately linear relationship

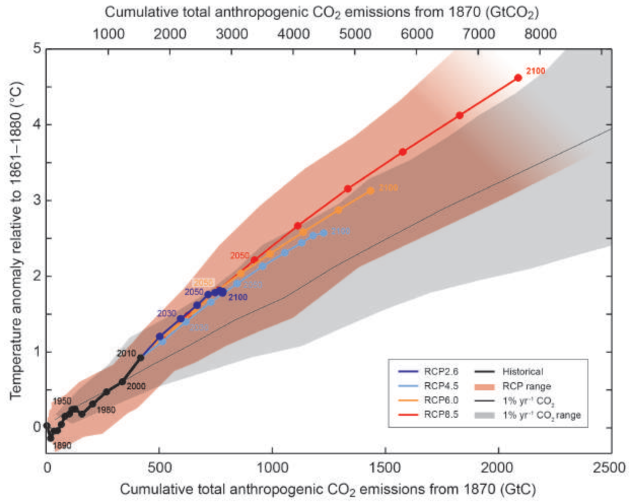


Figure 1: Cumulative emissions and temperature, IPCC et al. (2022)

between $\mathcal{S}$ and $\mathcal{T}$, as shown in Figure 1. It displays the relationship between temperature anomaly and cumulative CO_2 emissions over time, both for historical data in black and a large class of climate models in different Representative Concentration Pathways (RCP).

Moreover, we consider a linear relationship between global and local temperatures, namely,

$$T_i = \Delta_i \mathcal{T} = \Delta_i \chi \mathcal{S} \ ,$$

where Δ_i is a linear pattern scaling parameter that depends on geographical factors such as albedo or latitude.

Finally we consider a period damage function $\hat{\mathcal{D}}(T_{it} - T_i^\star)$ where $T_i^\star$ is the optimal' temperature for country i. The function $\hat{\mathcal{D}}(\hat{T})$ is a reduced-form representation of the economic damage to productivity, with curvature δ

$$\mathcal{D}_i^y(T_i - T_i^\star) = e^{-\gamma^y (T_i - T_i{}^\star)^{1+\delta}} \tag{12}$$

In our baseline quantification, we assume damages are quadratic, i.e. $\delta = 1$, as in the Integrated Assessment Models of Nordhaus' DICE-RICE, Krusell and Smith (2022), Kotlikoff et al. (2021a) and Burke et al. (2015). Such damage creates winners and losers: the countries warmer than the target temperature $T^\star$ are more affected by global warming. In contrast, regions with negative $T_i - T_i^\star$ benefit from a warmer climate. We discuss this quantification in Section 4.

To conclude, the reduced-form static damage functions $\mathcal{D}_i^y(\mathcal{E})$ and $\mathcal{D}_i^u(\mathcal{E})$, for productivity and utility, respectively, we summarize the future costs of climate change in present-discounted value,

$$\mathcal{D}_i^y(\mathcal{E}) = \mathcal{D}_i^y(T_i - T_i^\star) = \mathcal{D}_i^y\left(\Delta_i \chi (\mathcal{S}_0 + \mathbb{T}\mathcal{E}) - T_i^\star\right).$$

The damage function $\mathcal{D}_i^y(\mathcal{E})$ changes endogenously with climate policy choices, as they affect $\mathcal{E}$.

2.5 Equilibrium

To close the model, we need to determine the final good prices for each country p_i. To do so, we consider market clearing for each good i, which happens for the total quantity of goods, and not on a per capita basis, so we must adjust by total population, $\mathcal{P}_i$,

$$\mathcal{P}_i \underbrace{y_i}_{=\mathcal{D}_i^y(\mathcal{E})z_i F(\cdot)} = \sum_{k\in\mathbb{I}} \mathcal{P}_k \tau_{ki} c_{ki} + \mathcal{P}_i(x_i^f + x_i^c + x_i^r) \tag{13}$$

where x_i^f, x_i^c and x_i^r are the good inputs used in country i to produce fossil and renewable energy, respectively. To summarize, the competitive equilibrium of this economy is defined as follows:

Definition. Competitive equilibrium (C.E.):
For a set of policies $\{\mathrm{t}_i^{\varepsilon}, \mathrm{t}_{ij}^{b}, \mathrm{t}_i^{ls}\}_i$ across countries, a C.E. is a set of decisions $\{c_{ij}, e_i^f, e_i^c, e_i^r, e_i^x, \bar{e}_i^c, \bar{e}_i^r\}_{ij}$, and prices $q^f, \{\mathrm{p}_i, w_i, q_i^c, q_i^r\}_i$ such that:

(i) Households choose consumption $\{c_{ij}\}_{ij}$ maximizing utility equation (1) s.t. the budget constraint equation (2), which yield trade shares equation (3)
(ii) Final good firms choose inputs $\{\ell_i, e_i^f, e_i^r\}_i$ to maximize profits, resulting in equation (6)
(iii) Fossil energy firms maximize profits equation (7) and extract/produce $\{e_i^x\}_i$ given by equation (8)
(iv) Renewable and coal energy firms maximize profits and supplies $\{\bar{e}_i^c, \bar{e}_i^r\}$ are given respectively by equation (10) and equation (11)
(v) Energy markets clears for fossils as in equation (9) and for coal and renewable in equation (10) and equation (11)
(vi) Good markets clear for each country as in equation (13), and trade is balanced by Walras Law.

3 Welfare decomposition and climate policies experiments

Different countries have unequal exposure to both climate change and different climate policies. This international heterogeneity depends on how local temperature shocks impact countries' production, countries' exposure to international energy markets, and countries' position in the international goods trade network. In this section, we log-linearize the model to the first order to describe analytically the different transmission channels of several climate policy experiments.

3.1 Summary of the different experiments

In these experiments, we compute the marginal change in welfare $d\mathcal{U}_i$ for different policies, e.g., changes in energy taxes $d\mathrm{t}$. Our preferred welfare measure is the consumption-equivalent change in welfare measured as,

$$\Delta_{\mathrm{t}} \mathcal{U}_i = \frac{d\mathcal{U}_i}{d\mathrm{t}} \frac{1}{u'(c_i)\, c_i}$$

for policy change dt. Note that this welfare metric – in percentage terms – measures the *total derivative* $d\mathcal{U}_i/d$t and summarizes all the general equilibrium effects of the policy. While country i's welfare is simply the indirect utility $\mathcal{U}_i$, we compute the world welfare for several experiments,

$$\mathcal{W} = \sum_{i \in \mathbb{I}} \mathcal{P}_i \omega_i \mathcal{U}_i \qquad , \tag{14}$$

with Pareto weights ω_i and population $\mathcal{P}_i$. Similarly, the welfare change – measured in consumption-equivalent units is:

$$\Delta_{\mathrm{t}} \mathcal{W} = \frac{d\mathcal{W}/d\mathrm{t}}{\sum_i \mathcal{P}_i \omega_i \, u'(c_i) \, c_i} = \sum_i \mathcal{P}_i \widehat{\omega}_i \, \Delta_{\mathrm{t}} \mathcal{U}_i \tag{15}$$

with the consumption equivalent weights $\widehat{\omega}_i$,

$$\mathcal{P}_i \widehat{\omega}_i = \frac{\mathcal{P}_i \omega_i u'(c_i) \, c_i}{\sum_j \mathcal{P}_j \omega_j \, u'(\widetilde{c}_j) \, \widetilde{c}_j} \; .$$

Change in carbon emissions and global warming

In a first experiment, we want to understand the impact of climate change on different countries. For that, we write the response of agents welfare $d\mathcal{U}_i$ and decisions dx_i a first-order log-linear change in global emissions, $d \log \mathcal{E} = \frac{d\mathcal{E}}{\mathcal{E}}$, which will impact global temperature $\mathcal{T}$, by an amount $d \log \mathcal{T}$ and hence local temperature T_i. This changes productivity through damage as well as utility. However, the goal of this exercise is to quantify and decompose the welfare gains between (i) the direct effect in terms of TFP and the indirect effects due to (ii) the endogenous choice of inputs, in particular energy sources, and (iii) the reallocation of production through international trade. The total impact of climate is measured in consumption-equivalent percentage change,

$$\Delta_{\mathcal{E}} \mathcal{U}_i = \frac{1}{u'(c_i) \, c_i} \frac{d\mathcal{U}_i}{d\mathcal{E}} \; .$$

As we know from the climate economics literature, summarizing all these effects takes the form of the Social Cost of Carbon (SCC) – accounting for global welfare – and the Local Social Cost of Carbon (LCC) for the local welfare of country i. These represent the monetary cost of extra emissions. The Local Cost of Carbon is,

$$LCC_i = -\frac{\partial \mathcal{U}_i}{\partial \mathcal{E}} \Big/ \frac{\partial \mathcal{U}_i}{\partial c_i} = -\frac{d\mathcal{U}_i}{d\mathcal{E}} \frac{1}{u'(c_i)} \tag{16}$$

where $d\mathcal{U}_i$ is the change in i's welfare due to an 1% increase in emissions, and $u'(c_i)$ is the marginal utility of consumption.

Interestingly, at the first-order and small change in $d\mathcal{E}$, the Local Cost of Carbon is related to the consumption equivalent welfare change:

$$LCC_i = -\frac{d\mathcal{U}_i}{d\mathcal{E}} \frac{1}{u'(c_i)} = -c_i \, \Delta_{\mathcal{E}} \mathcal{U}_i$$

where LCC_i is measured in monetary units and $\Delta_c \mathcal{U}_i$ in percentage change.

Similarly, the global welfare changes for additional carbon emission relate to the Social Cost of Carbon for the welfare objective $\mathcal{W}$, which gives:

$$\begin{aligned} SCC = -\frac{\frac{\partial \mathcal{W}}{\partial \mathcal{E}}}{\frac{\partial \mathcal{W}}{\partial c}} &= \sum_i \mathcal{P}_i \omega_i u'(c_i)\, LCC_i \\ &= -\sum_i \mathcal{P}_i \omega_i u'(c_i) c_i\ \Delta_{\mathcal{E}} \mathcal{U}_i \\ &= -\left(\sum_i \mathcal{P}_i \omega_i u'(c_i) c_i\right)\ \Delta_{\mathcal{E}} \mathcal{W} \end{aligned}$$

As a result, the Social Cost of Carbon is proportional to the consumption equivalent welfare change, where the conversion factor is the household consumption weights by the social welfare weights $\omega_i u'(c_i)$. Details on this calculation in more general Integrated Assessment Models can be found in Bourany (2024).

Unilateral carbon taxation

We now analyze how welfare changes when each country implements a unilateral carbon tax t_i^ε on its own emissions. Such a tax increases the cost of oil-gas by $\xi^f d\log \mathrm{t}_i^\varepsilon \approx \xi^f d\mathrm{t}_i^\varepsilon$ and the cost of coal by $\xi^c d\mathrm{t}_i^\varepsilon$. The consumption-equivalent welfare measure for country i of implementing unilaterally a marginal increase in the carbon tax t_i^ε when other countries are passive $\mathrm{t}_j^\varepsilon = 0, \forall j \neq i$,

$$\Delta_{\mathrm{t}_i^\varepsilon} \mathcal{U}_i = \frac{1}{u'(c_i)\, c_i} \frac{d\mathcal{U}_i}{d\mathrm{t}_i^\varepsilon}\ .$$

Part of these welfare effects can be decomposed between (i) the direct effect on climate – which is limited given the unilateral implementation of the policy – and (ii) the general equilibrium effect through energy prices, oil-gas rents, and good prices. In particular, the leakage effects play a significant role that we can quantify. Naturally, such leakage effects generate welfare gains for country i when country j implements these carbon policies, $\Delta_{\mathrm{t}_j^\varepsilon} \mathcal{U}_i$. As a result, the leakage spillovers for all countries can be summarized by the Jacobian matrix:

$$\boldsymbol{\partial_{\mathrm{t}}} \mathcal{U} = \left\{ \Delta_{\mathrm{t}_j^\varepsilon} \mathcal{U}_i \right\}_{ij}$$

withwelfare change $\Delta_{\mathrm{t}_j^\varepsilon} \mathcal{U}_i$ for country i (row) from country j (column) unilateral carbon policy t^ε.

Unilateral renewable subsidies

Carbon taxation creates an additional cost for fossil fuels and is particularly detrimental to coal energy. Moreover, due to the leakage effect in terms of energy prices and international trade in goods, unilateral carbon taxation induces negative welfare losses, which give rise to policy inaction. We consider an alternative policy where each government unilaterally implements a subsidy on renewable energy $d\log \mathrm{s}_i^\varepsilon \approx d\mathrm{s}_i^\varepsilon$, which reduces the cost of renewable inputs e_i^r for final good firms. This type of industrial policy is financed through lump-sum taxation – maintaining symmetry with our carbon taxation experiments. The welfare changes can be written as,

$$\Delta_{\mathrm{s}_i^\varepsilon} \mathcal{U}_i$$

and similarly for the cross-country spillover Jacobian $\boldsymbol{\partial}_{\mathbf{s}} \mathcal{U} = \left\{ \Delta_{\mathrm{s}_j^\varepsilon} \mathcal{U}_i \right\}_{ij}$.

Unilateral carbon tariffs

To dampen the leakage effects through international trade in goods, carbon tariffs have been suggested. We consider the unilateral policy where each country i imposes tariffs that scale with the carbon intensity of the country j it imports from:

$$d \log \mathrm{t}_{ij}^b \approx d\mathrm{t}_{ij}^b = \xi_j^\varepsilon d\mathrm{t}_i^{b,\varepsilon} \qquad\qquad \xi_j^\varepsilon = \frac{\epsilon_i}{y_i \mathrm{p}_i}$$

where ξ_j^ε is the carbon intensity of country i with total emissions ϵ_i, and $d\mathrm{t}_i^{b,\varepsilon}$ is the the marginal increase in carbon price. Unilateral carbon tariffs change the relative prices of goods and has stronger terms-of-trade effects for carbon-intensive countries. The welfare change is denoted $\Delta_{\mathrm{t}_i^b} \mathcal{U}_i$.

Unilateral carbon taxation with carbon tariffs

Our fourth policy experiment is the combination of two unilateral policies : (i) a carbon tax $d\mathrm{t}_i^\varepsilon$ on fossil fuel consumption combined with (ii) a carbon tariff $d\mathrm{t}_i^{b,\varepsilon}$ on the carbon content of the imports, both of them the same size. Such a policy gives rise to a welfare change for country i,

$$\Delta_{\mathrm{t}_i^{b,\varepsilon}} \mathcal{U}_i$$

where the spillovers from j's policy on country i write in matrix form $\boldsymbol{\partial}_{\mathbf{t}^{b,\varepsilon}} \mathcal{U} = \left\{ \Delta_{\mathrm{t}_j^{b,\varepsilon}} \mathcal{U}_i \right\}_{ij}$.

Coordinated carbon policy

We now turn to coordinated climate policy. Unilateral policies are limited in their impact on climate due to the small size of countries relative to the world economy, which creates free-riding incentives. We consider the implementation of a marginal increase of the carbon tax in a large set J of countries – roughly, a climate agreement. We consider both an agreement with all the countries $\mathrm{J} = \mathbb{I}$, and agreements within the EU and ASEAN member states, in which nonmembers face a carbon tariff. The optimal design of climate agreements with carbon taxation and tariffs is studied and solved for in Bourany (2025).

The welfare of country i depends on whether it belongs to the agreements $i \in \mathrm{J}$ or not,

$$\Delta_{\mathrm{t}^\varepsilon | \mathrm{J}} \mathcal{U}_i,$$

for each agreement coalition J. In our experiments, we measure changes in welfare to the first-order given increase in the carbon tax. This implies that the cumulative impact of coordination scales linearly with the number of additional countries such that:

$$\begin{aligned}\Delta_{\mathrm{t}^{\varepsilon}|\mathrm{J}}\,\mathcal{U}_i &= \frac{1}{u'(c_i)\,c_i}\,\frac{d\mathcal{U}_i}{d\mathrm{t}^{\varepsilon}}\Big|_{\mathrm{J}} \\ &= \frac{1}{u'(c_i)\,c_i}\sum_{j\in\mathrm{J}}\alpha_{ij}\,\frac{d\mathcal{U}_i}{d\mathrm{t}_j^{\varepsilon}}\,\underbrace{\frac{d\mathrm{t}_j^{\varepsilon}}{d\mathrm{t}^{\varepsilon}}}_{=1} = \sum_{j\in\mathrm{J}}\alpha_{ij}\;\Delta_{\mathrm{t}_j^{\varepsilon}}\,\mathcal{U}_i,\end{aligned}$$

where α_{ij} are parameters to be determined as a function of the observables, like trade and energy market shares, energy elasticities, and damage parameters, as we will see below. Coordination gains may be nonlinear in terms of welfare – both in the number of countries J and the size of the carbon tax t^{ε} – as discussed in Bourany (2025).

We now derive the model's welfare decomposition for general experiments. In the next section, we turn to the result for specific policy experiments.

3.2 Observables and sufficient statistics

We now derive the impact of changes in prices and quantities on welfare through the budget constraint. First, we define several objects that are relevant for the decomposition:

- Trade share: $s_{ij} = \frac{c_{ij}p_{ij}}{c_i\mathbb{P}_i}$
- Energy share in production: $s_i^e = \frac{e_i q_i^e}{y_i \mathrm{p}_i}$
- Fossil share in energy mix $s_i^f = \frac{e_i^f q^f}{e_i q_i^e}$ and similarly $s_i^c = \frac{e_i^c q_i^c}{e_i q_i^e}$ and $s_i^r = \frac{e_i^r q_i^r}{e_i q_i^e}$
- Fossil energy share: $\eta_i^{\pi f} = \frac{\pi_i^f}{y_i\mathrm{p}_i + \pi_i^e}$ and similarly for $\eta_i^{\pi c}$, and $\eta_i^{\pi r}$.
- Production share/rent share in GDP: $\eta_i^y = \frac{y_i\mathrm{p}_i}{y_i\mathrm{p}_i+\pi_i^e} = 1-\eta_i^{\pi f}-\eta_i^{\pi c}-\eta_i^{\pi r}$
- Consumption expenditure: $x_i = c_i\mathbb{P}_i$
- Consumption share in GDP: $\eta_i^c = \frac{x_i}{y_i\mathrm{p}_i+\pi_i^e}$
- Consumption as a ratio of output: $s_i^{c/y} = \frac{c_i\mathbb{P}_i}{y_i\mathrm{p}_i} = \frac{x_i}{y_i\mathrm{p}_i+\pi_i^f}\frac{y_i\mathrm{p}_i+\pi_i^f}{y_i\mathrm{p}_i} = \frac{\eta_i^c}{1-\eta_i^{\pi}} = \frac{\eta_i^c}{\eta_i^y}$,
- Energy share as a ratio of consumption: $\frac{e_i q_i^e}{x_i} = \frac{e_i q_i^e}{y_i\mathrm{p}_i}\frac{y_i\mathrm{p}_i}{y_i\mathrm{p}_i+\pi_i^e}\frac{y_i\mathrm{p}_i+\pi_i^f}{x_i} = s_i^e\frac{\eta_i^y}{\eta_i^c}$
- Profit share as a ratio of consumption: $\frac{\pi_i^f}{x_i} = \frac{\pi_i^f}{y_i\mathrm{p}_i+\pi_i^e}\frac{y_i\mathrm{p}_i+\pi_i^f}{x_i} = \frac{\eta_i^{\pi f}}{\eta_i^c}$ and similarly for π_i^c and π_i^r.
- The share of GDP of energy imports and exports, with $v_i = \mathrm{p}_i y_i + q^f(e_i^x - e_i^f)$ and $v^y = \frac{\mathrm{p}_i y_i}{v_i}$, $v^{e^x} = \frac{q^f e_i^x}{v_i}$, $v^{e^f} = \frac{q^f e_i^f}{v_i}$ and $v^{ne} = \frac{q^f(e_i^x - e_i^f)}{v_k}$.

All these variables are measured in the current equilibrium, before implementing climate policies. Moreover, all these variables are observable in international trade and national accounts data.

3.3 Welfare, budget constraint and expenditure

We now compute the welfare of individual country i, $\mathcal{U}_i$, in consumption equivalent terms, accounting for the changes in consumption and climate damages,

$$\frac{d\mathcal{U}_i}{u'(c_i)c_i} = \frac{dx_i}{x_i} - \frac{d\mathbb{P}_i}{\mathbb{P}_i}\ ,$$

with $x_i = c_i \mathbb{P}_i$ the consumption expenditure c_i. To save on notation, we denote $d\log x_i = \frac{dx_i}{x_i}$ – with a slight abuse of notation.[8] As a result, we expand the budget constraint from equation (2). The carbon tax and subsidies are rebated/taxed lump-sum to the households, and hence, climate policies do not have direct income effects and only act through reallocation and equilibrium effects through price changes.

$$\Delta_t \mathcal{U}_i dt = \frac{\eta_i^y}{\eta_i^c}\Big(\underbrace{d\log \mathcal{D}_i^y}_{\substack{\text{climate}\\ \text{damage}}} + \underbrace{d\log \mathrm{p}_i}_{\substack{\text{exports terms}\\ \text{of trade}}} \Big) - \underbrace{d\log \mathbb{P}_i}_{\substack{\text{import terms}\\ \text{of trade}}} \quad - \frac{\eta_i^y}{\eta_i^c} s_i^e \Big(\underbrace{s_i^f d\log q^f + s_i^c d\log q_i^c + s_i^r d\log q_i^r}_{\text{energy price effects (demand)}} \Big)$$
$$+ \underbrace{\frac{\eta_i^{\pi f}}{\eta_i^c} d\log \pi_i^f + \frac{\eta_i^{\pi c}}{\eta_i^c} d\log \pi_i^c + \frac{\eta_i^{\pi r}}{\eta_i^c} d\log \pi_i^r}_{\text{energy rent effect (supply)}} \tag{17}$$

We see that the welfare gains and losses can be decomposed in five terms: (i) the direct impacts of climate damages production $\mathcal{D}_i^y$, (ii-iii) two terms of trade effects for p_i the sales of good i, and $\mathbb{P}_i$ the price index of the goods purchased by the household from countries j, and (iv-v) the effects of the energy prices changes – for oil-gas, q^f, coal, q_i^c, and renewable q_i^r – through the purchasing cost for production and through the revenue through energy rents. Note that we replace incomes from output y_i at the first-order using the standard Solow decomposition:

$$d\log y_i = d\log \mathcal{D}_i^y + \frac{MPe_i e_i}{y_i} d\log e_i = d\log \mathcal{D}_i^y + \; s_i^e \big[s_i^f \, d\log e_i^f + s_i^c d\log e_i^c + s_i^r d\log e_i^r \big]$$

where we use the labor supply being inelastic in this Armington trade structure $\ell_i = \bar{\ell}_i$. While each of these terms depends on how equilibrium prices change in a counterfactual, we do not need to simulate the model to recover how log prices change to the first order. Instead, each one of these terms can be computed using observable moments in the data and estimated or calibrated elasticities.

3.4 Climate system and damages

We log-linearize our simple climate system for small changes in emissions $\mathcal{E}$ and energy consumption in oil-gas E^f and in coal E^c. Given the linearity of the climate system $T_i = \Delta_i \mathcal{T} = \Delta_i \chi \mathcal{S}$, we naturally obtain that the percentage increase in temperature reflects the percentage change in carbon concentration,

$$d\log T_i = d\log \mathcal{S} \ .$$

The change in carbon concentration depends on the time horizon $\mathbb{T}$ where $\mathcal{S} = \mathcal{S}_0 + \mathbb{T}\mathcal{E}$, and $s^{\mathcal{E}/\mathcal{S}} = \mathbb{T}\mathcal{E}/\mathcal{S}$ which converges to 1, the longer the horizon we consider,[9]

$$d\log \mathcal{S} = s^{\mathcal{E}/\mathcal{S}} \; d\log \mathcal{E} \xrightarrow[\mathbb{T}\to\infty]{} d\log \mathcal{E} \ .$$

[8]This is not the case, for example, when $x_i < 0$ or changes sign.

[9]If $\mathcal{E}$ represents annual emissions $\sim 35GTCO_2$/year, given the calibration on 2015 temperature, a horizon $\mathbb{T}$ corresponding to ~ 85 years until 2100 would imply $\mathbb{T}\mathcal{E}/\mathcal{S} \sim 0.7$.

Now, using the damage function with slope γ and curvature δ as in equation (12), we obtain the linearization:

$$\begin{aligned} d\log \mathcal{D}_i^y &= -\gamma^y(1+\delta)\big(T_i - T^\star\big)^\delta T_i \; d\log T_i \\ d\log \mathcal{D}_i^y &= -\bar{\gamma}_i^y \; d\log \mathcal{E} \ . \end{aligned} \tag{18}$$

with $\bar{\gamma}_i^y = (1+\delta)(T_i - T_i^\star)^\delta T_i$. Therefore, the larger the curvature δ, the more significant the heterogeneity in damages $(T_i - T_i^\star)^\delta$, despite log-linearizing the model around the current equilibrium. Finally, total emissions reacts to changes in aggregate fossil fuels (oil-gas) and coal consumption,

$$d\log \mathcal{E} = s^{f/E} d\log E^f + s^{c/E} d\log E^c \qquad s^{f/E} = \frac{\xi^f E^f}{\mathcal{E}} \qquad s^{c/E} = \frac{\xi^c E^c}{\mathcal{E}}$$

To understand the determination of the equilibrium quantities of oil-gas E^f and E^c we now turn to energy markets.

3.5 Energy markets – Profits, prices, and supply

Using the energy firm problems, we log-linearize the profit change as a function of energy and input prices. For fossil energy (oil and gas), this yields,

$$d\log \pi_i^f = \left(1+\frac{1}{\nu_i^f}\right) d\log q^f - \frac{1}{\nu_i^f} d\log \mathrm{p}_i \tag{19}$$

with ν_i^f the fossil energy inverse supply elasticity. We obtain similar formulas for π_i^c and π_i^r as functions of ν_i^c, ν_i^r and q_i^c and q_i^r respectively.

Given that coal and renewables are assumed to be only traded locally (i.e., $\bar{e}_i^c = e_i^c$), we can write the supply curves for coal and renewable energy,

$$d\log q^c = \nu_i^c \, d\log e_i^c + d\log \mathrm{p}_i \tag{20}$$

and similarly for q^r with inverse elasticity ν_i^c. As a result, the price of coal and renewable energy rises both with the input cost p_i and with the quantity demanded e_i^c.

In contrast, oil-gas are traded internationally, where the price q^f clears the world market. Denoting E^f the total oil-gas quantity, changes to the aggregate supply curve are,

$$d\log E^f = \sum_i \lambda_i^x d\log e_i^x = \sum_i \lambda_i^f d\log e_i^f \ , \tag{21}$$

with $\lambda_i^f = \mathcal{P}_i e_i^f / E^f$ the demand share from country i and $\lambda_i^x = \mathcal{P}_i e_i^x / E^f$ the supply share (or market share) of country i as an exporter,

$$d\log q^f = \bar{\nu}^f d\log E^f + \sum_i \lambda_i^x \frac{\bar{\nu}}{\nu_i} d\log \mathrm{p}_i \ . \tag{22}$$

with the aggregate inverse supply elasticity $\bar{\nu}^f = \left(\sum_i \lambda_i^x \nu_i^{-1}\right)^{-1}$.

3.6 Energy markets – Taxation and demand

The production function equation (7) allows for different, finite elasticities of substitution between both labor and energy (with elasticity σ^y) and between the different energy sources (with elasticity σ^e). In our formulation, we can write the individual country demand curve – for example, for oil and gas – as,

$$\begin{aligned} d\log e_i^f = & \underbrace{-\left(\frac{\sigma^y}{1-s_i^e}s_i^f+(1-s_i^f)\sigma^e\right)\left[d\log q^f+\xi^f dt_i^\varepsilon\right]}_{\text{substitution away from oil-gas}} \\ & + \underbrace{\left(\sigma^e-\frac{\sigma^y}{1-s_i^e}\right)\left(s_i^c[d\log q_i^c+\xi^c dt_i^\varepsilon]+s_i^r[d\log q_i^r-ds_i^\varepsilon]\right)}_{\text{substitution away from coal and renewable toward oil-gas}} \\ & + \underbrace{\frac{\sigma^y}{1-s^e}(d\log\mathcal{D}_i+d\log p_i)}_{\substack{\text{energy demand changes due to climate} \\ \text{damage or price changes}}} \quad . \end{aligned}$$

We see that a price surge $d\log q^f$ not only decreases fossil demand e^f through substitution across energies σ^e but also away from energy σ^y scaling with the fossil share s_i^f. In comparison, a cost increase for coal also raises the demand for oil-gas e^f through substitution σ^e, but similarly pushes firms away from the energy bundle with elasticity σ^y – the difference of the two elasticities yielding the net effect. Moreover, an increase in productivity $\mathcal{D}_i^y$ or price p_i both increase input demand, as shown in the last two terms.

Finally, replacing input demand, we also summarize the impact on output of changes in prices and policies,

$$\begin{aligned} d\log y_i = (1+\alpha^y)d\log\mathcal{D}_i^y + \alpha^y d\log p_i - \alpha^y s_i^f\left[d\log q^f+\xi^f dt_i^\varepsilon\right] \\ - \alpha^y s_i^c[d\log q_i^c+\xi^c dt_i^\varepsilon] - \alpha^y s_i^r[d\log q_i^r - ds_i^\varepsilon], \end{aligned}$$

with the factor $\alpha^y = \frac{s_i^e\sigma^y}{1-s_i^e}$. We learn two lessons: first, climate damage $d\log\mathcal{D}_i^y$ reduces output more than one for one due to input reallocation away from energy. This multiplier effect $(1+\alpha^y)$ is even stronger when damage reduces when the energy supply curves are flat, i.e. $\nu^c=\nu^r=0$. However, this effect is dampened when those curves are less elastic $\nu^c\to\infty$: declines in TFP lower the energy price, which facilitates the purchase of additional inputs and improves production, providing some form of adaptation to climate change.

Second, we learn about the sensitivity of output to carbon taxation and other climate policies. The direct exposure of production y_i to the carbon tax t_i^ε is summarized by the factor $\alpha^y\left(\xi^f s_i^f+\xi^c s_i^c\right)=\frac{s_i^e\sigma^y}{1-s^e}(\xi^f s_i^f+\xi^c s_i^c)$, which represent the substitution effect of oil-gas and coal.

To decompose the effect further, we replace the input demand equation (23), with the supply curve equation (20) to find the equilibrium demand for coal and renewables – where the prices q_i^c and q_i^r are solved for. Due to the complex reallocations between the three inputs e_i^f, e_i^c, e_i^r, we keep the general formula for the appendix. The general intuition is that climate policies like carbon taxes t_i^ε and renewable subsidies reduce demand for oil and coal. However, this effect is dampened by the decline along the supply curve, which reduces the effectiveness of the policy.[10] As a result, the more elastic the supply for coal and renewable, the stronger the quantity – and emissions – response to carbon taxation and climate policies.

Oil-gas energy price

We now combine the individual countries' fossil energy demand, as in equation (23), with the country i share in global demand $\lambda_i^f = \frac{\mathcal{P}_i e_i^f}{E^f}$ as in equation (21), and the supply curve equation (22). We obtain, in general equilibrium, the total demand for oil,

$$
\begin{aligned}
d\log E^f &= \sum_i \lambda_i^f \, d\log e_i^f \\
&= \frac{1}{1+\bar{\nu}\overline{\lambda}_{\mathbb{I}}^{\sigma,f}} \sum_i \underbrace{\lambda_i^f \left[-\xi^f \Big(\frac{\sigma^y}{1-s_i^e} s_i^f + (1-s_i^f)\sigma^e \Big) + \xi^c s_i^c \Big(\sigma^e - \frac{\sigma^y}{1-s_i^e} \Big) \right]}_{\text{carbon taxation (oil-gas + coal) effect}} \mathrm{J}_i \, d\mathrm{t}^\varepsilon \\
&\quad - \frac{1}{1+\bar{\nu}\overline{\lambda}_{\mathbb{I}}^{\sigma,f}} \sum_i \underbrace{\lambda_i^f \, s_i^r \Big(\sigma^e - \frac{\sigma^y}{1-s_i^e} \Big)}_{\text{renewable subsidy effect}} \mathrm{J}_i \, d\mathrm{s}_i^\varepsilon + \frac{1}{1+\bar{\nu}\overline{\lambda}_{\mathbb{I}}^{\sigma,f}} \sum_i \underbrace{\lambda_i^f \frac{\sigma^y}{1-s_i^e} d\log \mathcal{D}_i^y}_{\text{climate damage effect}} \\
&\quad + \frac{1}{1+\bar{\nu}\overline{\lambda}_{\mathbb{I}}^{\sigma,f}} \sum_i \lambda_i^f \left(\underbrace{\frac{\sigma^y}{1-s_i^e} d\log \mathrm{p}_i}_{\text{terms-of-trade effect}} + \underbrace{\Big(\sigma^e - \frac{\sigma^y}{1-s_i^e} \Big)\big(s_i^c \, d\log q_i^c + s_i^r \, d\log q_i^r \big)}_{\text{substitution from coal and renewable}} \right)
\end{aligned}
\tag{23}
$$

with $\overline{\lambda}_{\mathbb{I}}^{\sigma,f} = \sum_{i\in\mathbb{I}} \lambda_i^f \left(\frac{\sigma^y}{1-s_i^e} s_i^f + (1-s_i^f) \right) \xi^f$ the aggregate demand elasticity for oil, and the aggregate inverse supply elasticity $\bar{\nu}^f = \left(\sum_i \lambda_i^x \nu_i^{-1} \right)^{-1}$.

This decomposition first reveals that the carbon tax affects oil-gas demand through two channels: the direct substitution away from oil-gas and the indirect substitution away from coal into oil-gas. Comparing the relative strength, the first effect dominates if:

$$
\overline{\lambda}_{\mathcal{J}}^{\sigma,f} := \sum_{i\in\mathcal{J}} \lambda_i^f \Big(\frac{\sigma^y}{1-s_i^e} s_i^f + (1-s_i^f)\sigma^e \Big) \, \xi^f > \sum_{i\in\mathcal{J}} \lambda_i^f \Big(\sigma^e - \frac{\sigma^y}{1-s_i^e} \Big) s_i^c \xi^c =: \overline{\lambda}_{\mathcal{J}}^{\sigma,c}
$$

where $\mathcal{J}$ is the set of countries imposing the carbon tax.[11] If the largest oil-gas consumers are reducing their demand for oil faster than they substitute away from coal, the net effect on oil-

[10] The formula in the appendix for coal and renewable demand have the following general form. Take an arbitrary energy demand curve $d\log e = -\sigma[d\log q + dt] + \tilde{\sigma} d\log b$, and the supply curve $d\log q = \nu d\log e + \tilde{\sigma} d\log c$, where b and c are demand and supply shifters respectively, we obtain the general demand:

$$
d\log e = -\frac{\sigma}{1+\nu\sigma} \mathrm{dt} + \frac{\tilde{\sigma}}{1+\nu\sigma} d\log b - \frac{\sigma}{1+\nu\sigma} d\log b
$$

[11] Reorganizing the terms, we can write it as follows – verified in most empirically relevant cases where large oil consumers λ_i^f also have high oil energy share s_i^f:

$$
\sum_{i\in\mathcal{J}} \lambda_i^f \big[\frac{\sigma^y}{1-s_i^e} (s_i^f \xi^f + s_i^c \xi^c) + \sigma^e \left((1-s_i^f)\xi^f - s_i^c \xi^c \right) \big] > 0.
$$

gas demand declines. Second, we also observe that climate policies, like carbon taxation, t^{ε}, and renewable subsidy, s^{ε}, are stronger when coordinated: oil-gas demand decline more for larger sets of countries implementing the policy J. However, it also generates a *energy price leakage effect*. When country i does not impose carbon policies, the price of oil declines for larger sets of climate agreement participants J – which is beneficial for country i's production and welfare. We quantify this leakage effect below.

To study the direct effect of carbon taxation and the gain for coordination with climate agreement J, we simplify the model further to obtain an analytical formula for the fossil price. In the following, we assume that the energy mix is concentrated on oil and gas $s_i^f = 1, s_i^c = s_i^r = 0$. This implies:

$$d\log E^f = -\frac{1}{1+\bar{\gamma}+\mathbb{C}\mathrm{ov}_i(\widetilde{\lambda}_i^f, \bar{\gamma}_i) + \bar{\nu}\overline{\lambda}^{\sigma,f}} \sum_i \widetilde{\lambda}_i^f \mathrm{J}_i d\mathrm{t}_i^{\varepsilon} + \sum_i \widehat{\lambda}_i^f d\log \mathrm{p}_i$$

with energy market shares $\lambda_i^f = \frac{\mathcal{P}_i e_i^f}{E^f}$, and weighted by elasticity $\widetilde{\lambda}_i^f = \lambda_i^f \frac{\sigma^y}{1-s_i^e}$. The average elasticity becomes $\overline{\lambda}^{\sigma,f} = \sum_i \lambda_i^f \frac{\sigma^y}{1-s_i^e}$, the price impact $\widehat{\lambda}_i^f = \widetilde{\lambda}_i^f/(1 + \sum_i \widetilde{\lambda}_i^f \gamma_i + \bar{\nu}\overline{\lambda}^{\sigma,f})$, and the covariance is the empirical analog, $\mathbb{C}\mathrm{ov}_i(x_i, y_i) = \sum_i x_i y_i - \sum_i x_i \sum_i y_i$.

The higher the carbon tax $d\mathrm{t}_i^{\varepsilon}$ – at the intensive margin – or the size of a climate policy agreement J – at the extensive margin – the lower the fossil-fuel demand. However, the strength of carbon taxation is muted for three reasons: (i) the more inelastic the supply – with higher curvature ν_i^f and $\bar{\nu}^f$ – the larger the price decline along the supply curve, which then in turn reinforce fossil demand and emissions. This contrasts with the analysis of the oil market in Asker et al. (2024): we claim that the more inelastic the oil supply – due to curvature of costs, concentration, or market power – is detrimental to the effectiveness of coordinated carbon taxation. Moreover, (ii) the energy curve q^f is affected by climate change: higher emissions damage the climate $\bar{\gamma} = \sum_i \bar{\gamma}_i$, which in turn reduces productivity, energy demand and hence emissions. Thus, with damages, the price impact of taxation is lower as it improves both the climate and energy demand in consequence. Additionally, (iii) this vicious effect is strengthened with the distribution of demand $\widetilde{\lambda}_i^f$ and climate damage $\bar{\gamma}_i$. The covariance term $\mathbb{C}\mathrm{ov}_i(\widetilde{\lambda}_i^f, \bar{\gamma}_i)$ is positive if large energy consumers – with larger market shares λ_i^f and high elasticity σ – are also the most affected by climate change – with higher costs $\bar{\gamma}_i$. The demand effect of taxation is thus muted in those circumstances.

3.7 International trade in goods

Recall the Armington trade block of the model where the good demand is constant elasticity of substitution θ, purchased at a price p_i and subject to tariffs t_{ki}^b for goods from country i. Using the market clearing in equation (13) – reformulated such as countries i sales equal countries k expenditures, coming from incomes in good sales as well as net-exports of fossil energy – the price index, and the trade shares as in equation (3), we obtain the log-linearization of trade patterns:

$$\mathcal{P}_i \mathrm{p}_i y_i = \sum_{k \in \mathbb{I}} \mathcal{P}_k s_{ki} \frac{1}{1+\mathrm{t}_{ki}} [\mathrm{p}_k y_k + q^f (e^x_k - e^f_k)]$$

$$\mathbb{P}_i = \left(\sum_j a_{ij} (\tau_{ij} (1+\mathrm{t}^b_{ij}) \mathrm{p}_j)^{1-\theta} \right)^{\frac{1}{1-\theta}} \qquad s_{ij} = \left(\frac{(1+\mathrm{t}^b_{ij}) \tau_{ij} \mathrm{p}_j}{\mathbb{P}_i} \right)^{1-\theta}$$

$$d \log s_{ij} = (\theta - 1) \left(\sum_k s_{ik} \left(d \log \mathrm{p}_k + d\mathrm{t}^b_{ik} \right) - \left(d \log \mathrm{p}_j + d\mathrm{t}^b_{ij} \right) \right)$$

The linearization of the market clearing is more complex as it also integrates the income effects of the changes in imports and exports of energy $e^f_i - e^x_i$ and energy prices q^f.

$$\underbrace{(d \log \mathrm{p}_i + d \log y_i)}_{\text{sales from } i} = \sum_k \boldsymbol{t}_{ik} \Big[\underbrace{\left(\frac{\mathrm{p}_k y_k}{v_k} \right) (d \log \mathrm{p}_k + d \log y_k)}_{\text{production income of } k} + \underbrace{\frac{q^f e^x_k}{v_k} d \log e^x_k - \frac{q^f e^f_k}{v_k} d \log e^f_k + \frac{q^f (e^x_k - e^f_k)}{v_k} d \log q^f}_{\text{oil-gas income of } k}$$
$$+ \underbrace{\theta \sum_h \left(s_{kh} d\mathrm{t}^b_{kh} - (1+s_{ki}) d\mathrm{t}^b_{ki} \right) + (\theta - 1) \sum_h (s_{kh} d \log \mathrm{p}_h - d \log \mathrm{p}_i)}_{\text{change in trade share in goods purchased by } k \text{ from } h \text{ relative to } i} \Big]$$

with $\mathbf{S} = \{s_{ij}\}_{ij} = \frac{c_{ij} p_{ij}}{c_i \mathbb{P}_i}$ the trade share matrix, $\mathbf{T} = \{\boldsymbol{t}_{ij}\}_{ij} = \{\frac{\mathcal{P}_j v_j}{\mathcal{P}_i v_i} s_{ji}\}$ income flow matrix – which represents the share of income v_i from i that is coming from country j – which is identical to the trade/income matrix in Kleinman, Liu and Redding (2024), and the Armington CES θ. This implies that rewritten in matrix notation, we get the price changes as a function of climate damage, carbon tax t^ε and renewable subsidies s^ε, and finally, the trade tariffs policies t^b where the matrix $\mathbf{J}$ represent which countries k impose tariffs on country i.

As for the reaction of oil-gas prices to climate policy, one general lesson from this decomposition relates to the gains from coordination: the larger the coalition of countries implementing carbon taxation and green subsidies, the stronger the reallocation effects. Climate damages lower the productivity $\mathcal{D}^y_k$ and income, which lowers the demand from the most affected countries and redirects the trade patterns. More policy coordination improves global climate, which mitigates those trade disruptions, even for countries that free-ride without implementing carbon policies. However, this free-riding is a double-edged sword: countries i outside climate agreements benefit more from the leakage effect caused by the climate policy in countries k, but when those policies are costly for k's income, that might lower the total demand for country i. All these channels are represented here, and the complete formula can be found in the appendix in Appendix B.

$$
\begin{aligned}
d\log \mathrm{p} = &\underbrace{\mathbf{A}\Big[(\mathbf{T}v^{y}{-}\mathbf{I})\delta^{y,z} - \mathbf{T}v^{e^f}\beta^{e,d,f}\Big]}_{\substack{\text{climate impact on output } \delta^{y,z}\\ \text{and energy demand } \beta^{e,d,f}}} d\log \mathcal{D}^{y} \;+\; \underbrace{\mathbf{A}\Big[-(\mathbf{T}v^{y}{-}\mathbf{I})\delta^{y,qf} + \mathbf{T}\left(v^{e^x}\tfrac{1}{\nu^f} - v^{e^f}\beta^{e,q}_{f,f} + v^{ne}\right)\Big]}_{\substack{\text{oil-gas price impact on energy output } \delta^{y,qf}\\ \text{and net energy bill } e^x - e^f}} d\log q^f \\
&+ \underbrace{\mathbf{A}\Big[-\mathbf{T}v^{e^f}\left(-\beta^{e,q}_{f,f}\xi^f + \beta^{e,t}_{f,c}\,\xi^c\right) - (\mathbf{T}v^{y}{-}\mathbf{I})\delta^{y,t\varepsilon}\Big]}_{\substack{\text{carbon tax impact on output } \delta^{y,t} \text{ and energy}\\ \text{demand away from oil-gas } \beta^{e,q}_{f,f} \text{ and from coal } \beta^{e,q}_{f,c}}} \mathrm{J}\, d\mathrm{t}^{\varepsilon} + \underbrace{\mathbf{A}\Big[\mathbf{T}v^{e^f}\beta^{e,t}_{f,r} + (\mathbf{T}v^{y}{-}\mathbf{I})\delta^{y,s\varepsilon}\Big]}_{\substack{\text{renewable subsidy impact on oil-gas}\\ \text{demand } \beta^{e,t}_{f,r} \text{ and output } \delta^{y,s\varepsilon}}} \mathrm{J}\, ds^{\varepsilon} \\
&+ \underbrace{\theta\mathbf{A}\left(\mathbf{TSJ}dt^{b} - \mathbf{T}(\mathbb{1}{+}\mathbf{S}')(\mathbf{J}dt^{b})'\right)}_{\text{trade reallocation due to tariffs}}
\end{aligned}
\tag{24}
$$

with $\mathbf{J}$ a tariff direction matrix – whether country i imposes tariffs on country j, with tariff increasing in the carbon intensity of j in the case of carbon tariffs. Moreover, the general equilibrium (and leakage) effects are summarized in a complicated matrix $\mathbf{A}$ that summarizes the fact that the price p_i also affects energy demand, oil-gas extraction, energy trade balance, and output. Further description can be found in the Appendix B.

3.8 Back to welfare and decomposition

In summary, we use the following welfare decomposition – following equation (17). We decompose the effects of the climate policies into the following four channels: (i) the climate damages or *direct productivity* impacts, (ii) the terms-of-trade effects, or *trade* effects, (iii) the effect on profit from the energy sector or *energy rents* and (iv) the impact on energy prices or *energy cost*.

$$
\begin{aligned}
\Delta_{\mathrm{t}}\mathcal{U}_i dt = &\underbrace{\frac{\eta_i^y}{\eta_i^c}\, d\log \mathcal{D}_i^y}_{\substack{\text{direct}\\ \text{productivity}}} \;+\; \underbrace{\frac{\eta_i^y}{\eta_i^c} d\log \mathrm{p}_i \;-\; d\log \mathbb{P}_i}_{\text{trade effects}} \;-\; \underbrace{\frac{\eta_i^y}{\eta_i^c} s_i^e\left(s_i^f\, d\log q^f + s_i^c d\log q_i^c + s_i^r d\log q_i^r\right)}_{\text{energy cost effects}} \\
&+ \underbrace{\frac{\eta_i^{\pi f}}{\eta_i^c} d\log \pi_i^f + \frac{\eta_i^{\pi c}}{\eta_i^c} d\log \pi_i^c + \frac{\eta_i^{\pi r}}{\eta_i^c} d\log \pi_i^r}_{\text{energy rents effects}}
\end{aligned}
\tag{25}
$$

We decompose the welfare impacts through these four different channels, given the other general equilibrium effects of the model: the climate damage as in equation (18), the individual countries' demand for energy as in equation (23), the global price for oil-gas in equation (22) which depends on aggregate quantity of energy consumed in equation (23), the profit of energy firms as in equation (19), and finally the general equilibrium of the good markets driving the terms of trade effects and prices as in equation (24). Before turning to the result of our model, we explain how we estimate the key parameters, such as climate damages and energy supply elasticities.

4 Estimation and quantification

Our main data sources are the World Development Indicators (WDIs), energy data compiled by Our World in Data from the Energy Institute (OWID, Energy Institute, 2024), and the International Trade and Production Database estimation sample (ITPD-E, Borchert et al., 2021) for international trade flows at the sector level, which allows us to remove energy trade from the flow of all other goods internationally. Additional data sources are described in the text. Our estimation covers years 2000-2016 for the trade data, and 1985-2019 for the energy data.

4.1 Estimating the structural damage function

We estimate the structural damage function $\mathcal{D}_i^y$ using temperature shocks and data on trade flows. Typical damage function estimation recovers damage parameters from regressions of GDP/capita on temperature shocks (see, e.g., Burke et al., 2015). In this context, a regression of log GDP per capita on temperature would fail to identify the parameters of the damage function, as GDP contains effects of temperature elsewhere (via their propagation through trade network), as well as general equilibrium effects on energy prices, rents and wages. Subsequently, in our context, an off-the-shelf damage function estimated on GDP recovers parameters that are subject to the Lucas critique. Our estimates of the structural adaptation function are robust to this critique because they net out endogenous adaptation and climate change effects that operate through energy markets.

To estimate T^* and γ, we use the model's gravity regression to estimate the parameters of the damage function, similar to Rudik et al. (2022). To do so, we leverage trade flow data from ITPD-E sample combined with local temperature data. For temperature data, we use Berkeley Earth near-surface temperature data (Rohde et al. (2013), available in $1° \times 1°$ cells), which we aggregate to the country level, population-weighing using population from the Global Human Settlement Layer in 2015 (Pesaresi et al., 2024).

The ITPD-E trade data measures trade flows X_{ij} at the industry level between most country pairs in 2000 through 2016. We use the ITPD-E data as it allows us to drop bilateral trade flows in the energy sector. Through the lens of the model, trade flows from exporter j to importer i are,

$$X_{ij} = p_{ij} c_{ij} = \left(\frac{(1+\mathrm{t}_{ij}^b)\tau_{ij}\mathrm{p}_j}{\mathbb{P}_i} \right)^{1-\theta} \frac{c_i}{\mathbb{P}_i} a_{ij}, \tag{26}$$

where,

$$\mathrm{p}_i = (\mathcal{D}_i^y(T_i) z_i)^{-1} \left(\epsilon (q_i^e)^{1-\sigma^y} + (1-\epsilon)(w_i)^{1-\sigma^y} \right)^{\frac{1}{1-\sigma^y}}$$

where q_i^e is a CES price index for energy prices. We use equation 26 as a basis for estimating $\mathcal{D}_i^y$. Dividing domestic trade, and treating each year t in the data as an equilibrium of the model, equation 26 becomes,

$$\frac{X_{ijt}}{X_{iit}} = (\mathrm{p}_{jt} \,/\, \mathrm{p}_{it})^{1-\theta} \varsigma_{ij} \varsigma_t \tilde{a}_{ijt} \tilde{\tau}_{ijt}$$

where ς_{ij} are all time-invariant bilateral preference or cost shifters, and $\tilde{a}_{ijt}\tilde{\tau}_{ijt}$ represents any time-varying components of bilateral preferences or trade costs. Taking logs and setting $\delta = 1$ generates the estimating equation,

$$\log \mathbb{E}[X_{ij}/X_{ii}] = \beta_0(T_{jt} - T_{it}) + \beta_1(T_{jt}^2 - T_{it}^2) + \varsigma_{ij} + \varsigma_t + \Gamma' W_{it} + \Omega' W_{jt}. \tag{27}$$

In this specification, ς_{ij} is a country-pair fixed effect, while ς_t is a time fixed effect. W_{it} and W_{jt} are importer- and exporter-year controls. These controls are a second order polynomial in log GDP/capita, the share of GDP that is attributed from oil rents, and renewable energy consumption as a percent of total final energy consumption, and are all taken from the WDIs. These controls proxy for the component of factory gate prices p_j driven by energy prices q_i^e or wages w_i. With these controls, the coefficients on temperature β_0 and β_1 identify parameters of the damage function (when $\delta = 1$), namely,

$$-\frac{1}{2}\frac{\beta_0}{\beta_1} = T^*, \quad \frac{-2\beta_1}{\theta - 1} = \gamma.$$

We estimate equation 27 with a Poisson pseudo-maximum likelihood estimator with high dimensional fixed effects to maintain zeros in the trade matrix (Silva and Tenreyro, 2006; Correia et al., 2020). We limit the estimation sample to countries for which domestic trade X_{ii} is not missing in the ITPD-E data and for which W_{it} and W_{jt} is not missing in the WDIs, as well as entities that are present in the ITPD but not other common trade datasets, like CEPII (Conte et al., 2022). This retains 169 countries, primarily dropping territories, dependencies, small island states, historical entities, or special jurisdictions, as well as North Korea.

Identification of β_0 and β_1 comes from within-trading-partner pair variation, tracing out how temperature shocks to bilateral terms-of-trade affect the import penetration ratio, net of year effects common to all pairs and importer- and exporter-year controls (which control for the time-varying component of factor prices). In short, the identification assumption is that time-varying shocks to preferences $\tilde{a}_{ijt}$ and shocks to bilateral cost shifters $\tilde{\tau}_{ijt}$ are such that $\mathbb{E}[\tilde{a}_{ijt}\tilde{\tau}_{ijt} \mid T_{it}, T_{it}^2, T_{jt}, T_{jt}^2, \varsigma_{ij}, \varsigma_t, W_{it}, W_{jt}] = 1$; i.e., conditional on the fixed effects, temperature (and its square), and controls, temperature shocks are uncorrelated with the error term.

Appendix Table 1 reports the results of our estimation. All specifications two-way cluster standard errors at the importer and exporter year to account for serial correlation of temperature within countries. Column (4) reports the results of our preferred specification, which uses differences in temperature $(T_{jt} - T_{it})$ and its square as a regressor, providing more efficient estimators of β_0 and β_1. We estimate $\hat{T}^* = 14.02$ and $\hat{\gamma} = 0.012$. Column (3) reports the results by separately estimating coefficients on T_{it}, T_{jt}, and their squares. Reassuringly, separate estimates have the correct sign and magnitude and are statistically indistinguishable from each other, though estimates on importer temperature are noisier. To estimate T^* and γ with separate coefficients on importer and exporter temperatures, we form precision-weighted averages of the estimands, and recover estimates of T^* and γ that are statistically indistinguishable to those estimated in Column (4).[12]

[12]Columns (1) and (2) use an OLS estimator. With the OLS estimator, the effects of importer temperature on

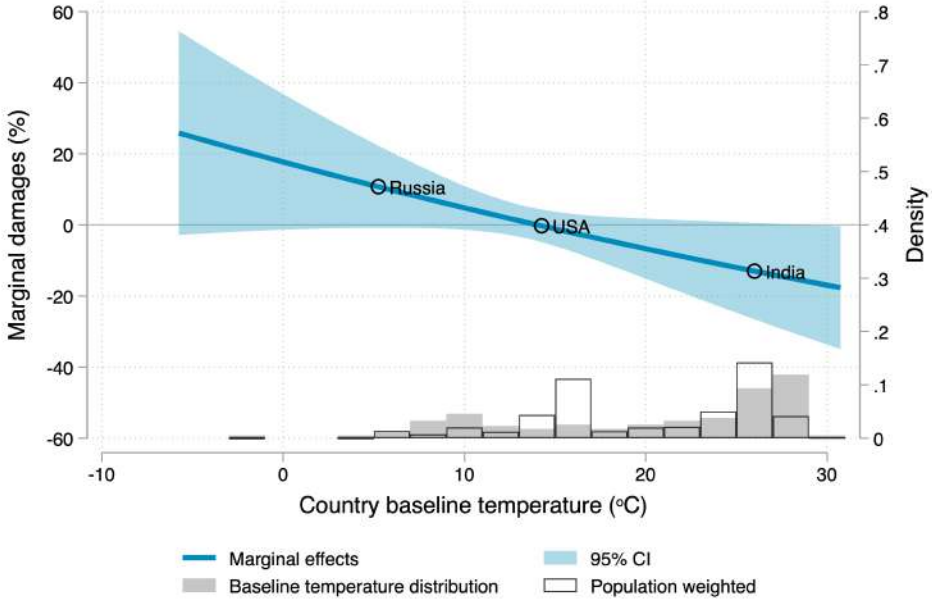


Figure 2: Marginal damages on productivity from a 1 degree change in local temperature versus country baseline temperature, using estimates of equation 27 available in Appendix Table 1, Column (4). 95% CIs computed using standard errors two-way clustered by importer and exporter. Also pictured: the histogram of countries across the baseline temperature distribution, weighted and unweighted by population. The coldest countries in the data are Greenland and Mongolia.

Figure 2 represents the damage function by using the estimates of β_0 and β_1 from Appendix Table 1, Column (4) to compute the marginal damages at each point along the baseline temperature distribution, i.e., the derivative of $\mathcal{D}_i^y$. Russia, which starts from a baseline cold temperature, experiences gains from local warming, while India, a baseline hot country, experiences large losses. This representation of the damage function is common in the literature, and our damage function resembles those estimated using GDP (Burke et al., 2015; Cruz and Rossi-Hansberg, 2024), despite our identification strategy that leverages panel variation in non-energy import penetration and temperature differences.

Finally, to calibrate $T_i^\star$, we also use an intermediary assumption that nests us between the following cases: (i) that T^* represents a global peak temperature for goods production TFP (as in Burke et al., 2015; Kotlikoff et al., 2021a; Krusell and Smith, 2022), (ii) or that global warming affects all the locations symmetrically, where deviation from the local baseline temperature $T_i^\star = T_{it_0}$ damage productivity, as in the representative agent economy of Barrage and Nordhaus (2024). A variant of (ii) is the assumption of full local adaptation to a changing climate, in which local weather shocks (i.e., temperature deviations from a moving average of local temperature) are the only source of damages, an assumption maintained in Kahn et al. (2019). Bilal and Känzig (2024) suggest, in contrast, that climate damages on GDP come in large part from shocks to global

import penetration are substantially noisier. With separate coefficients, we find a peak of around 16 and a flatter damage function $\gamma \approx 0.001$. The effects of temperature differences are indistinguishable from zero in, Column (2).

(rather than local) temperature, as they are associated with extreme events. Our intermediate step and assumes partial local adaptation by assuming,

$$T_i^\star = \alpha^T T^\star + (1-\alpha^T)T_{it_0}$$

where $\alpha^T = 0.5$ and $T^\star = 14.02$, as estimated. We hold γ fixed across countries, implying that all local adaptation affects the peak but not the shape of the damage function.

4.2 Estimating energy supply elasticities

Our goal is to recover the supply elasticities of fossil (oil and gas) and coal production for each country. Our estimation strategy begins with the model-implied relationship that,

$$d\log \pi_i^f = \left(1 + 1/\nu_i^f\right) d\log q^f - \left(1/\nu_i^f\right) d\log \mathrm{p}_i$$

As the world price of fossil is taken as exogenous to producers, a regression of oil rents on international oil-gas prices would recover the oil-gas supply elasticity, provided changes to international oil-gas prices are uncorrelated with changes in traded goods prices. For oil rents, we use data on the GDP share of oil and gas rents from the WDIs, η_i^f. For each country, we construct the effective price of fossil by taking an average of international oil and natural gas prices (from OWID), weighted by the share of oil and gas in the total fossil rents share of GDP. Treating each year as an equilibrium of the model, we leverage the time series to estimate the fossil supply elasticity by estimating,

$$\Delta \log \eta_{it}^f = \rho_i \Delta \log q_{it}^f + \Omega \Delta \log GDP_{it} + \varsigma_i + u_{it}$$

by OLS country-by-county, where here Δ indicates first-differencing and ς_i indicates a country fixed effect, which controls for country-specific time trends intended to capture potentially confounding secular trends in oil rents and international prices within each country. First-differencing implicitly nets out country fixed effects, which absorb all time-invariant shifters of p_i, while controlling for changes to GDP controls for year-over-year changes to p_i, as well as controlling for the denominator of η_{it}^f, so variation in the lefthandside reflects variation in π_i^f.

Estimating equation 4.2 results in very noisy estimates of ρ_i for some countries. Some estimates of ρ_i, while positive, fall below 1 (implying a negative oil-gas supply elasticity), while other estimates even negative, inconsistent with the model and incompatible with the quantification. To ameliorate this, we estimate a pooled estimate of ρ across countries and use an empirical Bayes shrinkage estimator where we impose a truncated normal prior on ρ_i with the truncation beginning at 1. This is tractable, as a normal likelihood (for the coefficient estimates, ρ_i) is conjugate with a truncated normal prior, which allows for easy recovery of the posterior mean. Imposing coefficient sign restrictions in estimation while jointly shrinking away noise is similar to the approach in estimating millions of retail product demand elasticities in Rosenthal-Kay et al. (2024). Histograms of the OLS and empirical Bayes estimates are available in Appendix Figure 11.

To recover coal supply elasticities, we repeat the analysis using the coal rent share of GDP

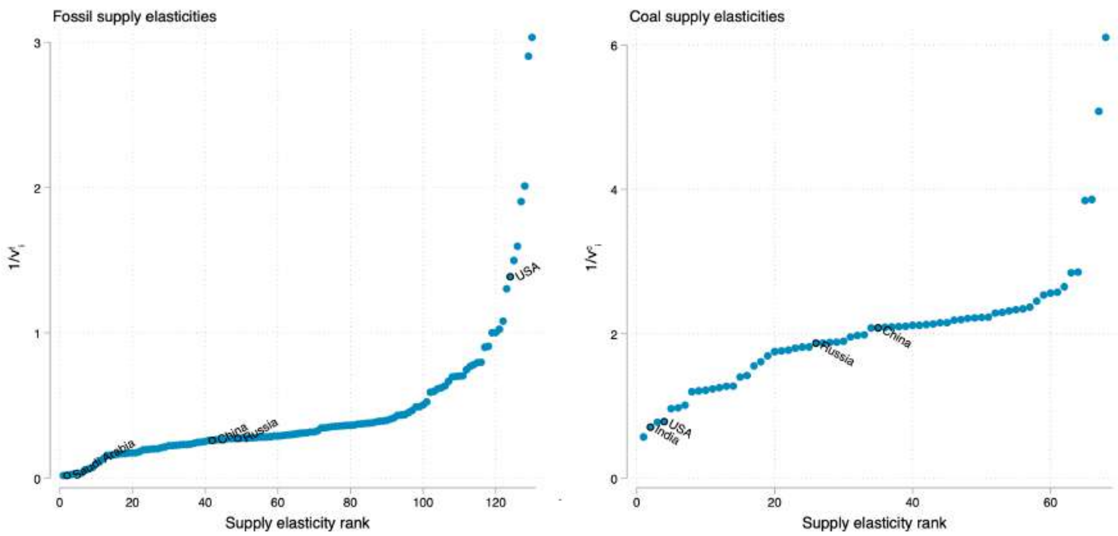


Figure 3: Empircal Bayes estimates of oil-gas and coal energy supply elasticities. Left: Hydrocarbon fossil (oil-gas) elasticities. Saudi Arabia, China, Russia, and the United States are labeled on the plot. Oil-gas energy is inelastically supplied in Saudi Arabia, potentially due to their market power through OPEC. Right: Coal energy supply elasticities.

as a regressand and the international price of coal (from OWID) as a regressor and apply the same empirical Bayes shrinkage routine. While in our model, coal is traded only locally, in reality, there does appear to be an international price of coal: differences in coal prices across countries are small, and movements in coal prices across countries are strongly correlated (Appendix Figure 10), as coal is traded to some extent, and as a commodity, there is significant and sophisticated arbitrage in global markets. With this in mind, we treat the global price of coal as exogenous for estimation.

Figure 3 plots the results of our estimation. Coal supply is substantially more elastic than fossil supply. There is large spatial heterogeneity in supply elasticity estimates: for example, fossil supply is nearly inelastic in OPEC nations like Saudi Arabia as well as Russia and China. Our estimates of energy supply elasticities do not uncover the true resource intensity of extraction technology that, in theory, could be uncovered by a production function estimation routine that accounts for market power. Instead, we view our estimates as 'reduced form' supply elasticities that combine both technology and market power. Market power can attenuate the effective supply elasticity as producers endogenously adjust quantities to move up along their perceived energy demand curve, consistent with low supply elasticities for OPEC nations. In contrast, we find fairly elastically supplied oil-gas in the U.S.

However, coal supply elasticities follow an opposite pattern, in which coal is fairly inelastically supplied in the U.S. and India and more elastically supplied in Russia and China. On average, coal is far more elastically supplied than fossil, generating a flatter supply curve and low coal rents in equilibrium. This is consistent with EPA estimates of the shape of the U.S. coal supply curve (U.S. Environmental Protection Agency, 2023), and the fact that even the largest coal producers do not have coal rents above 1% of GDP.

For countries with no fossil or coal rents in the data, we assign supply elasticities equal to the global pooled estimate (such countries are absent in the figure).

4.3 Externally calibrated parameters

For shares of energy rents in GDP and the other observable shares – as listed in Section 3.2 – and required by our sufficient statistics formula, we use 2000-2016 averages as the baseline, relying on data from the WDIs, OWID, and ITPD. For trade shares, we use data from the ITPD-S dataset for 2019, which fills in missing entries in the trade matrix with temporal smoothing and other extrapolation techniques to compute trade shares. For domestic trade, some entries in small nations with poor data, like Cuba, have implausibly low domestic trade shares. We replace reported domestic trade shares of 5% or less with these predicted domestic trade shares from a logistic regression of trade shares on log GDP/capita, population size, absolute latitude, temperature, and log bilateral flows with the United States (which has good reporting quality). We then renormalize the data so that trade shares sum to one, allowing us to construct $\mathbf{S}$. In practice, the trade data is not balanced, so we renormalize the columns of the $\mathbf{T}$ to mechanically enforce balanced trade in the data.[13]

For the household, we calibrate the CRRA/IES parameter to be $\eta = 1.5$, taken from Barrage and Nordhaus (2024).[14] and set the elasticity of substitution $\theta = 5$, consistent with a trade elasticity of 4, which accords with the estimates of Simonovska and Waugh (2014).

For the production function for goods, we use the average energy cost share $\frac{q_i^e e_i}{\mathrm{p}_i y_i}$ to 10%, as documented by OECD reports and used in Kotlikoff et al. (2021b) and Krusell and Smith (2022). For the elasticity between energy and other inputs, we set $\sigma^y = 0.6$ for all countries, which is in the range of estimates in Papageorgiou et al. (2017), among others.[15] This implies that labor and energy are complementary in production: an increase in the price of energy has a strong impact on output as it is less productive to substitute away toward other inputs – here, labor. This aligns with other empirical and structural evidence on the impact of energy shocks, e.g., Hassler et al. (2019). For energy sources, we calibrate the individual countries' energy mix for oil-gas, coal, and non-carbon (nuclear, hydroelectricity, solar, wind, etc.) to match the energy mix documented in Energy Institute (2024). This allows us to successfully identify countries that are more reliant on specific energy sources: for example, China and India are highly coal-dependent, while Russia, the Middle East, and the United States/Canada are the biggest consumers of oil and gas. Finally, for the elasticity between energy inputs, we use the value $\sigma_e = 2$, following the rest of the literature, i.e. Papageorgiou et al. (2017), Kotlikoff et al. (2021b), and Cruz and Rossi-Hansberg (2024), among others. As we are unable to estimate renewable energy supply elasticities at the country level,

[13]This ensures that the sum of the shares of income spent on traded goods sums to income. Visualizations of these trade matrices for 25 large economies are available in Appendix Figure 12.

[14]This is slightly lower than the standard value $\eta = 2$, for the reason that higher curvature would imply more unequal weights, ω_i, across different countries.

[15]It also aligns with the estimation in Bourany (2022).

we take $1/\nu^r = 2.7$, based on the estimate in Johnson (2014). A summary of these parameters is available in Appendix Table 2.

In addition, we calibrate the climate model described in Section 2.4 to match important features of the relationship between carbon emissions, temperatures, and climate damages. We consider linear models for the relationships between carbon emissions $\mathcal{E}$, carbon concentration $\mathcal{S}$, and global and local temperature $\mathcal{T}$ and T_i, and this implies that we do not require to parametrize the climate sensitivity χ, or the pattern scaling Δ_i.[16] However, the economic model being static, we consider the horizon $\mathbb{T}$ to be 2100 for performing policy experiments. In this context, we analyze the long-term policy impact of climate and trade policies through different channels.

Finally, when we consider welfare, as in equation (14), we consider the weighted sum of individual utilities, with Utilitarian weights $\omega_i = 1$. This implies that the aggregation of the consumption-equivalent welfare change in equation (15) can be aggregated with weights $\mathcal{P}_i\widehat{\omega}_i \propto \mathcal{P}_i\omega_i u'(c_i)c_i = \mathcal{P}_i c_i^{1-\eta}$ for η the inverse of the IES, which control "inequality aversion" in this type of models, as discussed in Anthoff et al. (2009). This, therefore, puts additional weight on the welfare costs of emerging and low-income countries with low consumption and we report his welfare measure in our policy experiments. Alternatively, we also consider Negishi weights $\omega_i = 1/u'(c_i)$. These would undermine these redistributive considerations by putting more weight on advanced economies, yielding $\widehat{\omega}_i = c_i \propto y_i$.

5 Results

In this section, we report the results of our main experiments, using our sufficient statistics formulas, data moments, and estimated damage functions and energy supply elasticities.

5.1 The welfare effects of global warming

We use our sufficient statistics formula to compute the welfare effects of global warming from an emissions impulse that generates 3°C of warming by 2100. The results of this exercise are displayed in Figure 4. In the left panel, we display the spatial distribution of welfare changes around the world. Our results accord with the vast majority of the literature. For example, just as in Cruz and Rossi-Hansberg (2024), the losers of climate change are predominately concentrated in the global south: Africa, Latin America, and South East Asia – all hot countries at baseline – lose, while cooler countries like Canada and Russia win, and the effects of global warming are small in the United States and China.

The right panel of Figure 4 decomposes the welfare changes into those driven by the direct effects of climate change, those driven by change terms-of-trade, changes in energy rents, and exposure to changes in the price of energy. Global warming, by making the world poorer, reduces global demand for energy, which lowers the equilibrium price q^f and provides relief for oil and

[16] Indeed, the log-linearization of the linear climate system yield $d\log T_i = d\log\mathcal{S} \propto d\log\mathcal{E}$ where there is no requirement for climate sensitivity or pattern scaling.

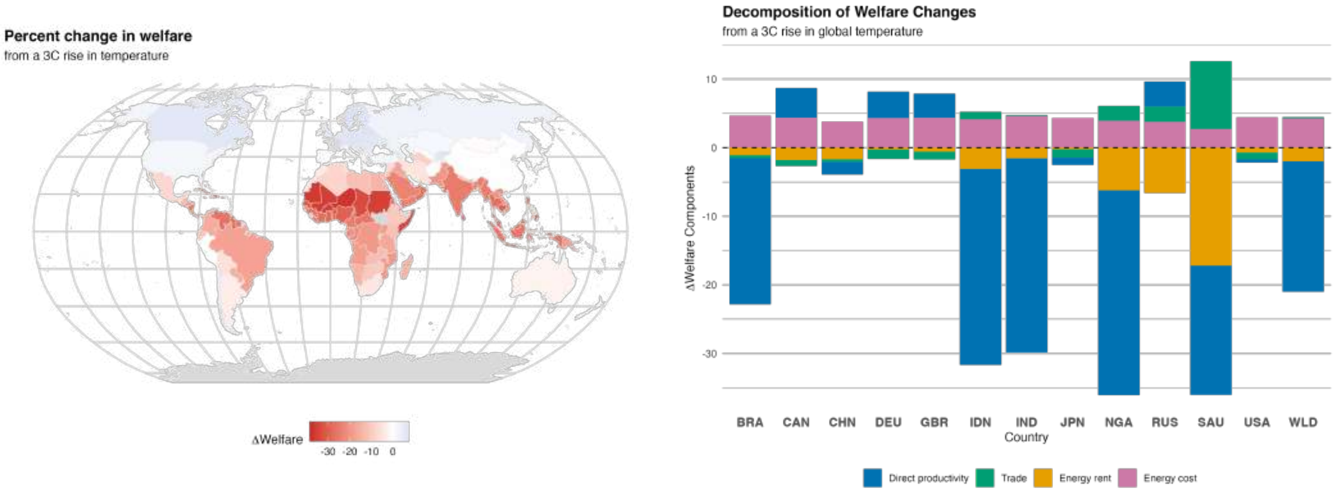


Figure 4: The welfare effects of a 3°C rise in global temperature. Left: map of global welfare changes, in % terms. Red countries lose, while blue countries win. Right: decomposition of welfare changes for several major economies.

gas importing countries. However, it also deteriorates the energy rents of fossil exporters such as Saudi Arabia. However, this baseline hot country loses due to the direct effect on productivity, this loss is partially offset by improving terms-of-trade, as Saudi Arabia is well-connected in the trade matrix to countries that gain, like Japan and Russia.

5.2 Unilateral carbon taxation

We now consider the welfare effects of unilateral carbon taxation. As a case study, we begin with a unilateral carbon tax imposed by China of \$50 USD/ton. While this is potentially an unrealistic scenario, this illustrates model mechanisms and the value of including a rich energy sector in a macroeconomic model of climate change. Despite being a large polluter, this moderate carbon tax has virtually no effect on global emissions, which fall by less than 0.1%. Yet, the policy nonetheless creates winners and losers, visible in the left panel of Figure 5. The winners of the policy include Gulf and North African nations, as well as Russia, and interestingly, China.

The reason for the gains in these nations is visible in the right panel. At baseline, China is heavily reliant on coal, reflected in their large share of coal in their energy mix (one of the data moments key to our sufficient statistics exercise). By imposing a carbon tax, China substitutes away from coal toward oil-gas fossil sources, as coal is dirtier than oil-gas ($\xi_c/\xi_f \approx 1.44$), putting upwards pressure on the price of these energy commodities. As a result, the global price of oil-gas rises by approximately 5%, and energy rents rise for fossil exporters as a result of carbon taxation in China. If oil-gas emissions rise by around 5% as well, emissions from coal fall by -7.6%. To sum up, utilitarian welfare falls by 0.14% (0.23% Negishi-weighted).

While the rise in the global price of oil-gas improve the rents of fossil exporters, energy costs inflate in energy importers. For example, European nations 'lose' as their energy costs rise. In China, the overall welfare effect of changing energy prices is positive: the direct effect of taxation drops out in the first-order decomposition as tax revenues are rebated to the household; only the

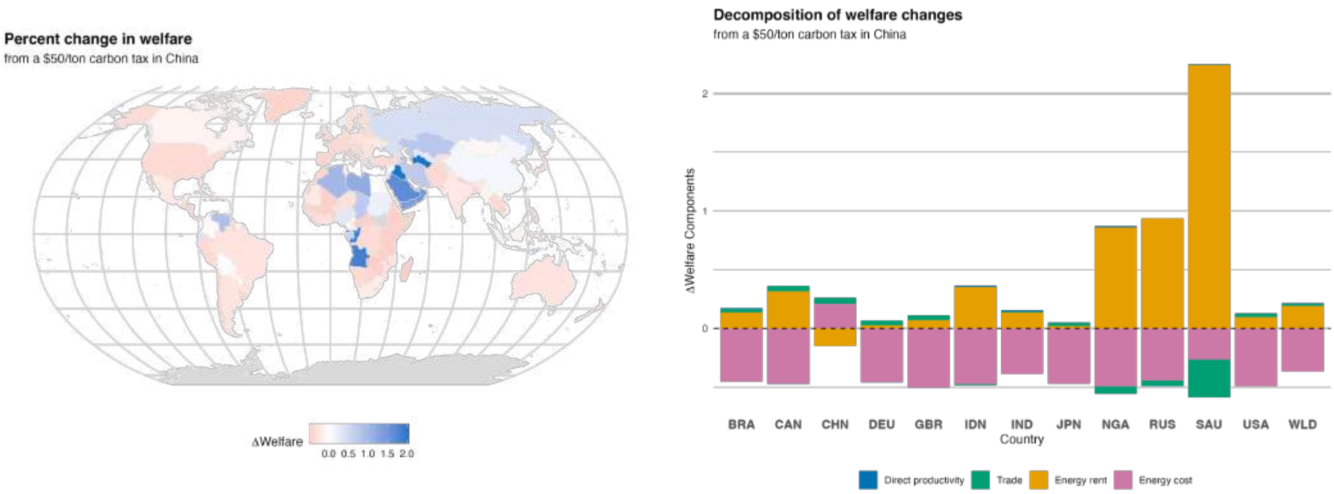


Figure 5: The welfare effects of a \$50/ton carbon tax imposed unilaterally by China. Left: map of global welfare changes, in % terms. Red countries lose, while blue countries win. Right: decomposition of welfare changes for several major economies.

general equilibrium of prices net of carbon taxes affects welfare. Coal prices fall as demand shifts to fossil, and as China has a large share of coal in their energy mix, they benefit.

In Appendix Figure 13, we display the results of a \$50 unilateral carbon tax imposed in the United States. This policy fosters a 0.8% decline in global emissions. This reduction in atmospheric CO_2 has positive welfare effects on nations in the global south and damages the 'winners' of climate change. Interestingly, though Saudi Arabia's energy rents appreciate, as was the case with China, their terms of trade deteriorate as their Middle-Eastern and South Asian trading partners suffer productivity losses. Welfare effects in the U.S. are small but positive: energy prices rise, more than totally offsetting the gains associated with reducing climate damages, but are balanced by improved terms of trade with Canada and European nations.

5.3 Unilateral renewable energy subsidies

In contrast to carbon taxation, we consider renewable energy subsidies. We compare the effect of subsidizing renewables with carbon taxation by plotting each country's welfare change from unilateral policy for a \$50 carbon tax and a 42.6% renewable energy subsidy. We choose this renewable energy subsidy so that the relative price change from policy between the oil-gas and coal bundle and renewables stays the same on average. The main reason why subsidizing renewable energy differs from taxing carbon is that it does not directly cause a reallocation from carbon-intensive coal to oil-gas within the dirty energy bundle, as the tax does not directly alter the oil-gas and coal price ratio. This affects both the aggregate change in emissions, as well as dirty energy exporters' energy rents, depending on their relative dirty energy supply elasticities.

In Figure 6, we plot both each country's consumption equivalent welfare change from pursuing a \$50 dollar carbon tax (left panel) and a 42.6% renewable energy subsidy (right).[17] There are

[17]In Appendix Figure 14, left panel, we plot for each country the global welfare gain – of a Utilitarian planner –

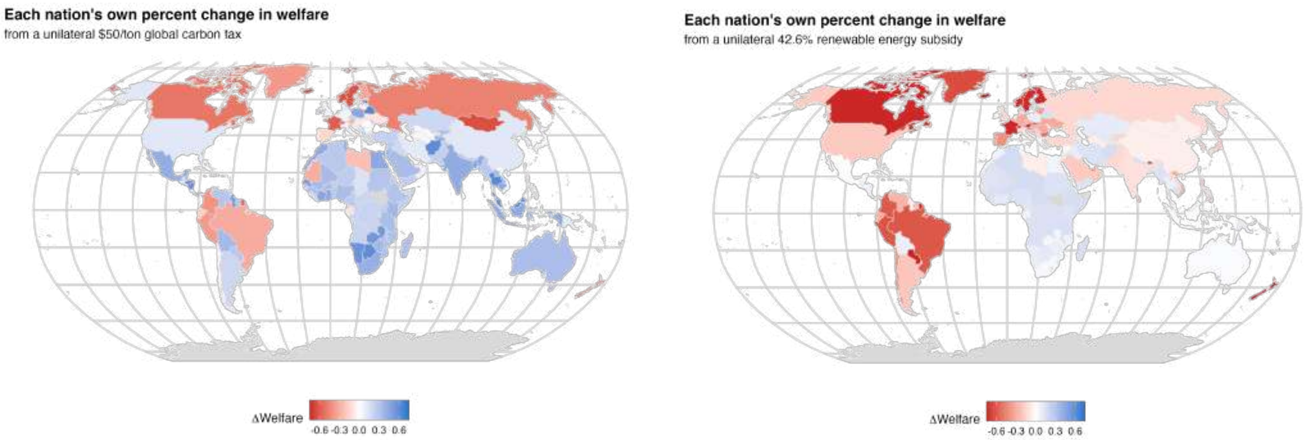


Figure 6: Left: each country's individual welfare change from pursuing unilateral carbon taxation. Right: each country's individual welfare change from pursuing a unilateral renewable subsidy.

large differences across the world in the effects of these two policies. On average, renewable energy subsidies are considerably less effective at raising welfare and cause more harm. For example, a renewable subsidy in France is much worse than a carbon tax, as it disproportionally induces France to move up its renewable energy supply curve. As France has a high share of renewable energy at baseline (over 40%), this effect on energy costs is considerable, as more resources are wasted on renewable 'extraction.' Simply put, the marginal nuclear power plant or solar farm has a high price in France. Likewise, in China, a renewable subsidy generates a large movement up its domestic renewable supply curve, rather than the small movement up the global oil-gas supply curve induced by carbon taxation.

5.4 Coordinated carbon policy

We now consider coordinated climate policy, in which blocs of nations jointly implement carbon taxes and tariffs that take the form of a carbon border adjustment mechanism (CBAMs). CBAMs levy a tariff on the carbon content of imports to each nation in the bloc, but do not place tariffs on bloc members – a 'climate club' as suggested by Nordhaus (2015) and studied in other work (Clausing and Wolfram, 2023; Ernst et al., 2023; Bourany, 2025). The carbon intensity of any nation's exports is observable by knowing the energy mix in production and the carbon intensity of those energy sources, and is readily observable simply by knowing the carbon emissions of any country. Exporting countries only respond to CBAM through the general equilibrium impact on energy prices as well as through terms-of-trade adjustments in the international goods market.

First, we examine the effects of a European Union-wide climate club with a carbon tax and tariff of $50. Figure 7 displays the results of this exercise. The EU is a loser from its own climate club, with only Spain and Portugal benefitting. Aggregate emissions decline by 2.4%, cooling

associated with unilateral carbon policy. While Russia loses from imposing a carbon tax on themselves, the world still benefits, suggesting a considerable gap between governmental and global incentives to internalize the carbon externality. Likewise, in the right panel, we can see similar effects for the renewable energy subsidy.

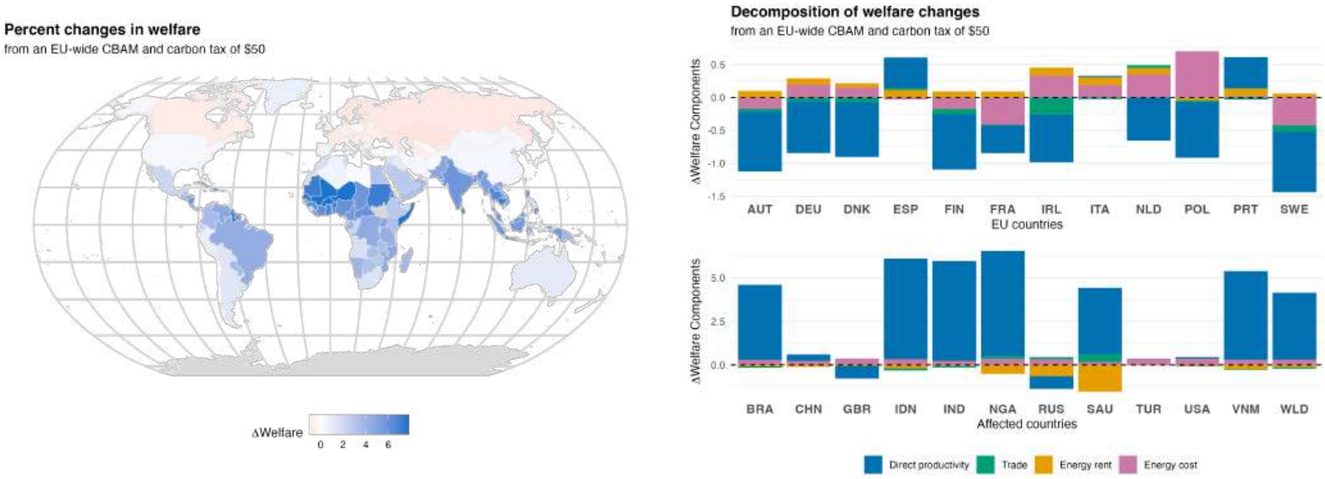


Figure 7: Left: changes in welfare from a EU climate club. Right: Decomposition of welfare changes within the EU (top) and for major trade partners and losers of climate change (bottom).

global temperatures and harming EU nations that benefit from climate change, alongside other cool nations like Canada and Russia. However, the global welfare effect of such a policy is positive, with global utilitarian-weighted welfare rising by 3.9% (a 0.8% with Negishi weights).

In the top right panel, we plot the welfare decomposition for EU member states. Energy cost effects are heterogeneous across the EU, rising in France and falling in Poland. Ireland is particular hurt through international trade, as their major trading partner, Great Britain, faces productivity declines. Countries in the global south benefit from the reduction in world temperature (bottom panel), and most economies benefit from a reduction in energy costs, as demand pressures from Europe in the international fossil hydrocarbon market lessen. This generates positive welfare effects in countries that are mostly unaffected by climate change directly, like the United States and China. Major oil-gas exporters like Saudi Arabia, Nigeria, and Russia lose energy rents as a result, though only Russia is a net loser due to the direct, negative productivity effect brought about by fewer carbon emissions.

In contrast, we consider a climate club composed of ASEAN members with the same policies as the EU climate club. These southeastern Asian countries are losers from climate change at baseline, so internalizing the carbon externality benefits them from the reduction in carbon emissions alone. Figure 8 plots the results of this exercise. The ASEAN climate club reduces global emissions by 0.6% and raises global welfare by 1% (0.2%, Negishi-weighted). By reducing global emissions, the climate club benefits the losers of climate change and harms the winners.

ASEAN members broadly benefit from the policy, owing to the reduction in world temperature. However, gains are heterogeneous not only because of exposure to climate change, but also because of trade in goods and energy, as seen in the top left panel. Energy exports like Brunei have smaller gains as they lose energy rents. Goods trade within the club reallocates, with Indonesia and Myanmar enjoying improved terms of trade at the expense of Malaysia, Singapore, Thailand, and Vietnam. Elsewhere, oil-gas exporters lose energy rents as the fossil fuels price falls by 0.8%.

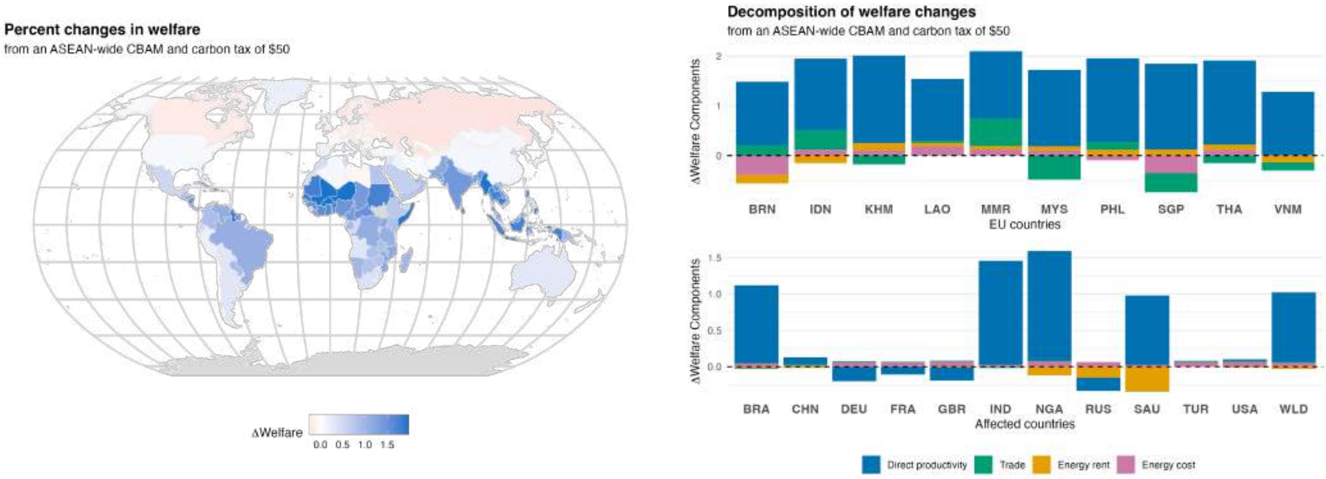


Figure 8: Left: changes in welfare from a ASEAN climate club. Right: Decomposition of welfare changes within ASEAN member states (top) and for major economies and losers of climate change (bottom).

5.5 Global carbon policy

Finally, we examine the case when all nations impose a $50 carbon tax in Figure 9. When all nations participate in carbon taxation, carbon border adjustment mechanisms are not needed due to the targeting principle; the carbon externality is internalized at its source. When we impose this policy, the global price of hydrocarbons rises by 1%. Indeed, this results from the mechanism explained in equation (23), where carbon taxation has a strong impact on coal consumption, and countries substitute toward oil and gas. In net, carbon emissions decline by 4%, which is one order of magnitude larger than for unilateral policies or carbon taxation implemented in small climate clubs. Such global climate policy results in a 6% increase in utilitarian-weighted global welfare. Were we to evaluate welfare using Negishi weights, the global change in welfare is 1%, as rich, cool countries receive a higher weight. Consequently, a large part of the measured welfare gain from implementing global carbon taxation, when evaluating welfare changes from a utilitarian perspective, stems from reducing international inequality.

We plot the results of this exercise in Figure 4. They are the reverse of the effects of climate change: nations in the global south benefit from global carbon taxation while baseline cold nations lose. These losses are offset in Europe, as European nations like Great Britain and Germany have improved terms of trade, while gains are attenuated in Saudi Arabia for the same reason. Saudi energy rents grow as the world substitutes from coal to oil-gas, which puts upward pressure on the international price of hydrocarbon fossil.

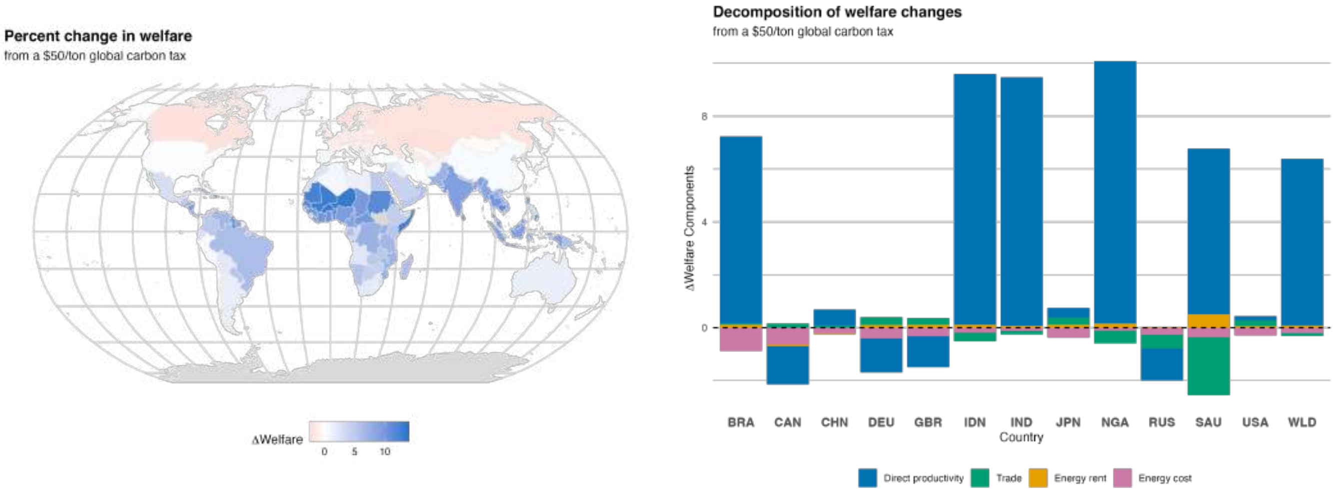


Figure 9: The welfare effects of a $50 carbon tax imposed in every country around the world. Left: map of global welfare changes, in % terms. Red countries lose, while blue countries win. Right: decomposition of welfare changes for several major economies.

6 Conclusion

In this paper, we use a first-order decomposition of the effects of climate policy derived from a macroeconomic IAM to study who wins and loses under various climate policy regimes. Our IAM features trade in both goods and energy markets, and our decomposition is computable with a modest set of sufficient statistics available using freely available trade and national accounts data and estimable elasticities.

We estimate heterogenous damages on productivity from rises in local temperature using bilateral trade data, and estimate a large set of heterogeneous energy supply elasticities for both hydrocarbon fossil and coal producers. We find that this heterogeneity is important not only in capturing heterogeneity in productivity damages across space from climate change but also in the response of nations' energy sectors to climate policy.

We use our estimates to consider a large set of climate policies. First, in agreement with the literature, we find spatially heterogeneous winners and losers from climate change, with baseline hot nations suffering as a result of a hotter planet, and cold nations gaining. These welfare changes are amplified by changes in international goods and energy markets: dirty energy exporters lose energy rents, and the pattern of trade adjusts, improving or deteriorating different nations' terms of trade.

We find that pursuing unilateral policy is often ineffective in combatting climate change, despite often creating positive welfare gains for the nations that pursue such policies. For example, carbon taxation in China does little to affect global emissions, and redirects their energy mix towards oil and gas imports, improving the extraction rents of energy exporters and indirectly generating an improved terms-of-trade for China in international goods markets. As nations' energy mixes differ, carbon taxation and subsidizing renewable energy generate different welfare responses. Broadly, we find that subsidizing renewables is substantially less effective than taxing carbon.

Coordinated climate policy through climate clubs, in which member nations impose a domestic carbon tax and carbon tariffs on imports from non-member nations, better addresses the climate externality. Climate policy in the EU broadly harms EU members but delivers sizable global gains. A climate club of ASEAN members both improves global welfare and the welfare of club members. However, this policy has unequal effects on member nations, as it redirects trade, causing some nations to benefit more than others.

In short, our results suggest that implementing climate policy is difficult unilaterally, as leakage effects can be an order of magnitude larger than the gains from cooling global temperature. International agreements are necessary to combat climate change, but can generate unequal effects among members party to the agreement through these leakage channels. While our sufficient statistics and welfare formulas allow us to quickly compute and decompose the effects of many possible climate clubs, they do not take into account the nonlinear effects induced by imposing large carbon taxes and tariffs. Bourany (2025) moves beyond sufficient statistics and solves for the *optimal* design of these types of climate agreements.

References

Abuin, Constanza, "Power Decarbonization in a Global Energy Market: The Climate Effect of US LNG Exports," Technical Report, working paper 2024.

Adao, Rodrigo, Arnaud Costinot, and Dave Donaldson, "Putting Quantitative Models to the Test: An Application to Trump's Trade War," June 2023.

Allcott, Hunt, Reigner Kane, Maximilian S Maydanchik, Joseph S Shapiro, and Felix Tintelnot, "The Effects of "Buy American": Electric Vehicles and the Inflation Reduction Act," Technical Report, National Bureau of Economic Research 2024.

Anderson, James E., "A Theoretical Foundation for the Gravity Equation," *The American Economic Review*, 1979, *69* (1), 106–116.

Anthoff, David, Cameron Hepburn, and Richard SJ Tol, "Equity weighting and the marginal damage costs of climate change," *Ecological Economics*, 2009, *68* (3), 836–849.

Arkolakis, Costas and Conor Walsh, "Clean Growth," August 2023.

_ , Arnaud Costinot, and Andres Rodriguez-Clare, "New Trade Models, Same Old Gains?," *American Economic Review*, February 2012, *102* (1), 94–130.

Asker, John, Allan Collard-Wexler, Charlotte De Canniere, Jan De Loecker, and Christopher R Knittel, "Two Wrongs Can Sometimes Make a Right: The Environmental Benefits of Market Power in Oil," Technical Report, National Bureau of Economic Research 2024.

Baqaee, David Rezza and Emmanuel Farhi, "The macroeconomic impact of microeconomic shocks: Beyond Hulten's theorem," *Econometrica*, 2019, *87* (4), 1155–1203.

_ and _ , "Productivity and misallocation in general equilibrium," *The Quarterly Journal of Economics*, 2020, *135* (1), 105–163.

_ and _ , "Networks, barriers, and trade," *Econometrica*, 2024, *92* (2), 505–541.

Barrage, Lint and William Nordhaus, "Policies, projections, and the social cost of carbon: Results from the DICE-2023 model," *Proceedings of the National Academy of Sciences*, March 2024, *121* (13), e2312030121. Publisher: Proceedings of the National Academy of Sciences.

Bhattarai, Keshab, Sushanta K. Mallick, and Bo Yang, "Are global spillovers complementary or competitive? Need for international policy coordination," *Journal of International Money and Finance*, February 2021, *110*, 102291.

Bilal, Adrien and Diego R Känzig, "The macroeconomic impact of climate change: Global vs. local temperature," Technical Report, National Bureau of Economic Research 2024.

_ and Esteban Rossi-Hansberg, "Anticipating Climate Change Across the United States," June 2023.

Bohringer, Christoph, Jared C. Carbone, and Thomas F. Rutherford, "Unilateral climate policy design: Efficiency and equity implications of alternative instruments to reduce carbon leakage," *Energy Economics*, December 2012, *34*, S208–S217.

Bolt, Jutta and Jan Luiten van Zanden, "Maddison-style estimates of the evolution of the world economy: A new 2023 update," *Journal of Economic Surveys*, 2023, *n/a* (n/a).

Borchert, Ingo, Mario Larch, Serge Shikher, and Yevgeniy Yotov, "The International Trade and Production Database for Estimation (ITPD-E)," *International Economics*, 2021, *166*, 140–166.

Bornstein, Gideon, Per Krusell, and Sergio Rebelo, "A World Equilibrium Model of the Oil Market," *The Review of Economic Studies*, January 2023, *90* (1), 132–164.

Bourany, Thomas, "Energy shocks and aggregate fluctuations," 2022.

— , "Climate Change, Inequality, and Optimal Climate Policy," 2024.

— , "The Optimal Design of Climate Agreements: Inequality, Trade, and Incentives for Climate Policy," *Job Market Paper*, 2025.

Burke, Marshall, Solomon M Hsiang, and Edward Miguel, "Global non-linear effect of temperature on economic production," *Nature*, 2015, *527* (7577), 235–239.

Carleton, Tamma A, Amir Jina, Michael T Delgado, Michael Greenstone, Trevor Houser, Solomon M Hsiang, Andrew Hultgren, Robert E Kopp, Kelly E McCusker, Ishan B Nath, James Rising, Ashwin Rode, Hee Kwon Seo, Arvid Viaene, Jiacan Yuan, and Alice Tianbo Zhang, "Valuing the Global Mortality Consequences of Climate Change Accounting for Adaptation Costs and Benefits," Working Paper 27599, National Bureau of Economic Research July 2020. Series: Working Paper Series.

Chetty, Raj, "Sufficient statistics for welfare analysis: A bridge between structural and reduced-form methods," *Annu. Rev. Econ.*, 2009, *1* (1), 451–488.

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

Conte, Bruno, Klaus Desmet, Dávid Krisztián Nagy, and Esteban Rossi-Hansberg, "Local sectoral specialization in a warming world," *Journal of Economic Geography*, 2021, *21* (4), 493–530.

Conte, Matteo, Pierre Cotterlaz, and Thierry Mayer, "The CEPII Gravity database," CEPII Working Paper 2022-05, CEPII July 2022.

Copeland, Brian R and M Scott Taylor, "Trade, growth, and the environment," *Journal of Economic literature*, 2004, *42* (1), 7–71.

Correia, Sergio, Paulo Guimarães, and Tom Zylkin, "Fast Poisson estimation with high-dimensional fixed effects," *The Stata Journal*, 2020, *20* (1), 95–115.

Costinot, Arnaud and Iván Werning, "Robots, trade, and luddism: A sufficient statistic approach to optimal technology regulation," *The Review of Economic Studies*, 2023, *90* (5), 2261–2291.

— , Dave Donaldson, Jonathan Vogel, and Ivan Werning, "Comparative Advantage and Optimal Trade Policy," *The Quarterly Journal of Economics*, May 2015, *130* (2), 659–702.

Cruz, Jose-Luis and Esteban Rossi-Hansberg, "Local Carbon Policy," May 2022.

— **and** — , "The Economic Geography of Global Warming," *The Review of Economic Studies*, March 2024, *91* (2), 899–939.

Dietz, Simon, Frederick van der Ploeg, Armon Rezai, and Frank Venmans, "Are Economists Getting Climate Dynamics Right and Does It Matter?," *Journal of the Association of Environmental and Resource Economists*, September 2021, *8* (5), 895–921. Publisher: The University of Chicago Press.

Dingel, Jonathan I., Kyle C. Meng, and Solomon M. Hsiang, "Spatial Correlation, Trade, and Inequality: Evidence from the Global Climate," January 2019.

Energy Institute, "Statistical Review of World Energy," 2024.

Ernst, Anne, Natascha Hinterlang, Alexander Mahle, and Nikolai Stahler, "Carbon pricing, border adjustment and climate clubs: Options for international cooperation," *Journal of International Economics*, September 2023, *144*, 103772.

Farrokhi, Farid and Ahmad Lashkaripour, "Can Trade Policy Mitigate Climate Change," *Econometrica*, 2024.

Folini, Doris, Aleksandra Friedl, Felix Kübler, and Simon Scheidegger, "The climate in climate economics," *Review of Economic Studies*, 2024, p. rdae011.

Fontagne, Lionel and Katheline Schubert, "The Economics of Border Carbon Adjustment: Rationale and Impacts of Compensating for Carbon at the Border," *Annual Review of Economics*, September 2023, *15* (Volume 15, 2023), 389–424. Publisher: Annual Reviews.

Golosov, Mikhail, John Hassler, Per Krusell, and Aleh Tsyvinski, "Optimal Taxes on Fossil Fuel in General Equilibrium," *Econometrica*, 2014, *82* (1), 41–88.

Goulder, Lawrence H, "Environmental taxation and the double dividend: a reader's guide," *International tax and public finance*, 1995, *2*, 157–183.

Hassler, John, Per Krusell, and Anthony A Smith, "Environmental macroeconomics," in "Handbook of macroeconomics," Vol. 2, Elsevier, 2016, pp. 1893–2008.

_ , _ , and Conny Olovsson, "Directed technical change as a response to natural-resource scarcity," Working Paper 375, Sveriges Riksbank Working Paper Series 2019.

_ , _ , and _ , "Presidential Address 2020 Suboptimal Climate Policy," *Journal of the European Economic Association*, December 2021, *19* (6), 2895–2928.

Hillebrand, Elmar and Marten Hillebrand, "Optimal climate policies in a dynamic multi-country equilibrium model," *Journal of Economic Theory*, 2019, *179*, 200–239.

Hsiao, Allan, "Coordination and Commitment in International Climate Action," 2022.

Hsieh, Chang-Tai and Peter J Klenow, "Misallocation and manufacturing TFP in China and India," *The Quarterly journal of economics*, 2009, *124* (4), 1403–1448.

Johnson, Erik Paul, "The cost of carbon dioxide abatement from state renewable portfolio standards," *Resource and Energy Economics*, 2014, *36* (2), 332–350.

Kahn, Matthew E., Kamiar Mohaddes, Ryan N.C. Ng, M. Hashem Pesaran, Mehdi Raissi, and Jui-Chung Yang, "Long-term macroeconomic effects of climate change: A cross-country analysis," Working Paper 26167, National Bureau of Economic Research August 2019.

Kleinman, Benny, Ernest Liu, and Stephen J Redding, "International friends and enemies," *American Economic Journal: Macroeconomics*, 2024.

_ , _ , _ , and Motohiro Yogo, "Neoclassical growth in an interdependent world," 2023.

Kortum, Samuel S. and David A. Weisbach, "Optimal Unilateral Carbon Policy," November 2021.

Kotlikoff, Laurence, Felix Kubler, Andrey Polbin, Jeffrey Sachs, and Simon Scheidegger, "Making Carbon Taxation a Generational Win Win," *International Economic Review*, 2021, *62* (1), 3–46. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/iere.12483.

Kotlikoff, Laurence J., Felix Kubler, Andrey Polbin, and Simon Scheidegger, "Can Today's and Tomorrow's World Uniformly Gain from Carbon Taxation?," September 2021.

Krusell, Per and Anthony A. Smith, "Climate change around the world," 2022.

Lashkaripour, Ahmad, "The cost of a global tariff war: A sufficient statistics approach," *Journal of International Economics*, 2021, *131*, 103419.

_ **and Volodymyr Lugovskyy**, "Profits, scale economies, and the gains from trade and industrial policy," *American Economic Review*, 2023, *113* (10), 2759–2808.

Nordhaus, William, "Climate clubs: Overcoming free-riding in international climate policy," *American Economic Review*, 2015, *105* (4), 1339–1370.

Nordhaus, William D, "Estimates of the social cost of carbon: background and results from the RICE-2011 model," 2011.

Nordhaus, William D. and Zili Yang, "A Regional Dynamic General-Equilibrium Model of Alternative Climate-Change Strategies," *The American Economic Review*, 1996, *86* (4), 741–765. Publisher: American Economic Association.

Ossa, Ralph, "Trade Wars and Trade Talks with Data," *American Economic Review*, December 2014, *104* (12), 4104–4146.

Papageorgiou, Chris, Marianne Saam, and Patrick Schulte, "Substitution between Clean and Dirty Energy Inputs: A Macroeconomic Perspective," *The Review of Economics and Statistics*, 05 2017, *99* (2), 281–290.

Pesaresi, Martino, Marcello Schiavina, Panagiotis Politis, Sergio Freire, Katarzyna Krasnodębska, Johannes H Uhl, Alessandra Carioli, Christina Corbane, Lewis Dijkstra, Pietro Florio et al., "Advances on the Global Human Settlement Layer by joint assessment of Earth Observation and population survey data," *International Journal of Digital Earth*, 2024, *17* (1), 2390454.

Pörtner, H.-O. IPCC, D.C. Roberts, H. Adams, I. Adelekan, C. Adler, R. Adrian, P. Aldunce, E. Ali, R. Ara Begum, B. Bednar Friedl, R. Bezner Kerr, R. Biesbroek, J. Birkmann, K. Bowen, M.A. Caretta, J. Carnicer, E. Castellanos, T.S. Cheong, W. Chow, G. Cisse G. Cisse, and Z. Zaiton Ibrahim, *Climate Change 2022: Impacts, Adaptation and Vulnerability* Technical Summary, Cambridge, UK and New York, USA: Cambridge University Press, 2022.

Rohde, Robert A. and Zeke Hausfather, "The Berkeley Earth land/ocean temperature record," *Earth System Science Data Discussions*, 2020, *2020*, 1–16.

Rohde, Robert, Richard A Muller, Robert Jacobsen, Elizabeth Muller, Saul Perlmutter, Arthur Rosenfeld, Jonathan Wurtele, Donald Groom, and Charlotte Wickham, "A New Estimate of the Average Earth Surface Land Temperature Spanning 1753 to 2011," *Geoinformatics & Geostatistics: An Overview*, 2013, *1* (1), 1–7.

Rosenthal-Kay, Jordan, James Traina, and Uyen Tran, "Several Million Demand Elasticities," *Working paper*, 2024.

Rudik, Ivan, Gary Lyn, Weiliang Tan, and Ariel Ortiz-Bobea, “The Economic Effects of Climate Change in Dynamic Spatial Equilibrium,” June 2022.

Shapiro, Joseph S, “The Environmental Bias of Trade Policy,” *The Quarterly Journal of Economics*, May 2021, *136* (2), 831–886.

Silva, JMC Santos and Silvana Tenreyro, “The log of gravity,” *The Review of Economics and statistics*, 2006, pp. 641–658.

Simonovska, Ina and Michael E Waugh, “The elasticity of trade: Estimates and evidence,” *Journal of international Economics*, 2014, *92* (1), 34–50.

U.S. Environmental Protection Agency, “Chapter 7 – Coal,” 2023. Power Sector Modeling Platform v6 Post-IRA 2022 Reference Case.

Weisbach, David A., Samuel Kortum, Michael Wang, and Yujia Yao, “Trade, Leakage, and the Design of a Carbon Tax,” *Environmental and Energy Policy and the Economy*, January 2023, *4*, 43–90. Publisher: The University of Chicago Press.

Weitzman, Martin L., “Internalizing the Climate Externality: Can a Uniform Price Commitment Help?,” *Economics of Energy & Environmental Policy*, 2015, *4* (2), 37–50. Publisher: International Association for Energy Economics.

Appendix

A Additional tables and figures

	OLS		Poisson	
	(1)	(2)	(3)	(4)
Exporter Temperature (C)	0.086***		0.722***	
	(0.030)		(0.203)	
Exporter Temperature^2	-0.002		-0.027***	
	(0.001)		(0.007)	
Importer Temperature (C)	0.058		-0.607	
	(0.117)		(0.826)	
Importer Temperature^2	-0.002		0.020	
	(0.005)		(0.023)	
Exporter-Importer Temperature (C) difference		0.015		0.652**
		(0.061)		(0.331)
Exporter-Importer Temperature^2 difference		0.000		-0.023*
		(0.002)		(0.012)
T^*	15.807	-31.560	13.399	14.016
	(5.0629)	(421.9431)	(0.9911)	(1.9916)
γ	0.001	-0.000	0.010	0.012
	(0.0006)	(0.0012)	(0.0027)	(0.0060)
Importer-Exporter pair FE	✓	✓	✓	✓
Origin GDP/cap control	✓	✓	✓	✓
Origin energy controls	✓	✓	✓	✓
R^2	0.865	0.865		
Pseudo-R^2			0.854	0.854
Observations	366,384	366,384	463,614	463,614

Table 1: Estimates of equation 27. $^{***}p < 0.1,^{**} p < 0.05,^{*} p < 0.01$. Standard errors are two-way clustered at the importer and exporter level in parentheses. Dependent variable: importer penetration ration X_{ij}/X_{ii}. Columns (1)-(2) use an OLS estimator with the log of the import penetration ratio on the lefthandside, while (3)-(4) use a Poisson estimator, which retains zeros in the trade matrix. Specifications differ by whether importer and exporter temperature are allowed to have different coefficients, or if the model-implied coefficient restriction is imposed. Computation of T^* and γ assumes $\theta = 5$, as in our quantification.

Parameter	Value	Description	Source
Household preferences			
η	1.5	Coefficient of relative risk aversion	Barrage and Nordhaus (2024)
$\theta - 1$	4	Trade elasticity	Simonovska and Waugh (2014)
Goods production			
T^*	14.02	Global peak temperature	Estimated
γ	0.012	Shape parameter of $\mathcal{D}_i^y$	–
σ^y	0.3	Elasticity of substitution between energy and labor	Papageorgiou et al. (2017)
ω^f	–	Share of oil-gas in production	Calibrated to match energy mix
ω^c	–	Coal share	–
ω^r	–	Renewables share	–
σ_e	2	Elasticity of substitution between energy sources	Papageorgiou et al. (2017); Kotlikoff et al. (2021b)
Energy production			
$1/\nu_i^f$	–	Supply elasticity of oil-gas	Estimated for each country
$1/\nu_i^c$	–	Supply elasticity of coal	–
$1/\nu^r$	2.7	Supply elasticity of renewable energy	Johnson (2014)

Table 2: Summary of estimated and externally calibrated parameters

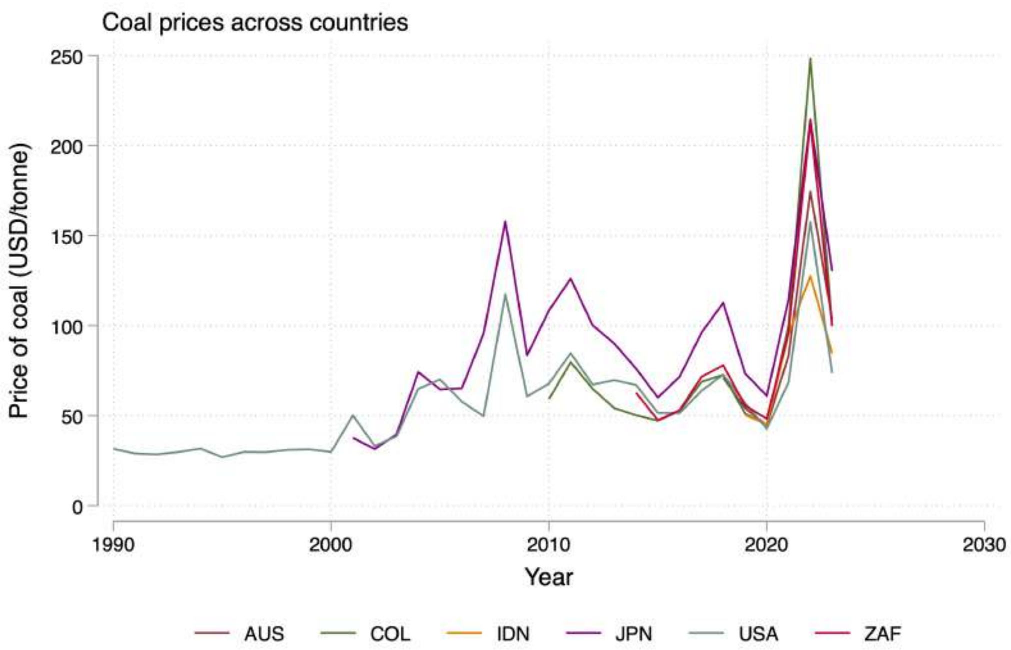


Figure 10: Coal prices in several countries. Source: Our World in Data.

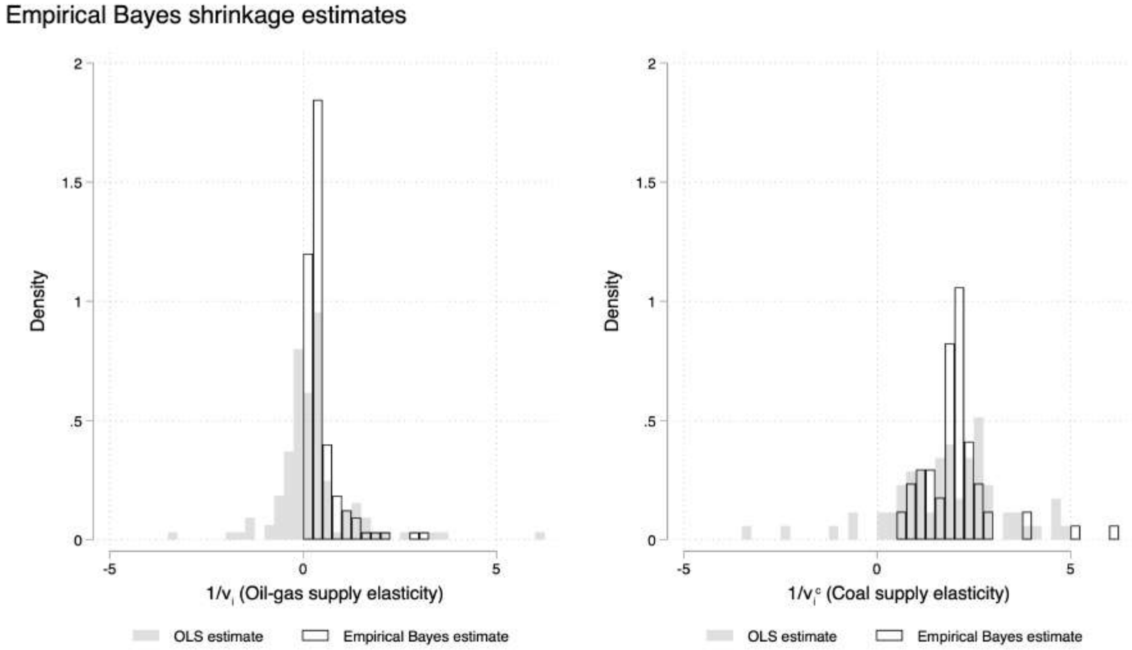


Figure 11: Histograms of the distribution of OLS country-specific energy supply elasticities and the empirical Bayes estimates. Left: oil-gas (hydrocarbon fossil) energy supply elasticities, $1/\nu_i^f$. Right: coal energy supply elasticities $1/\nu_i^c$. See main text for details.

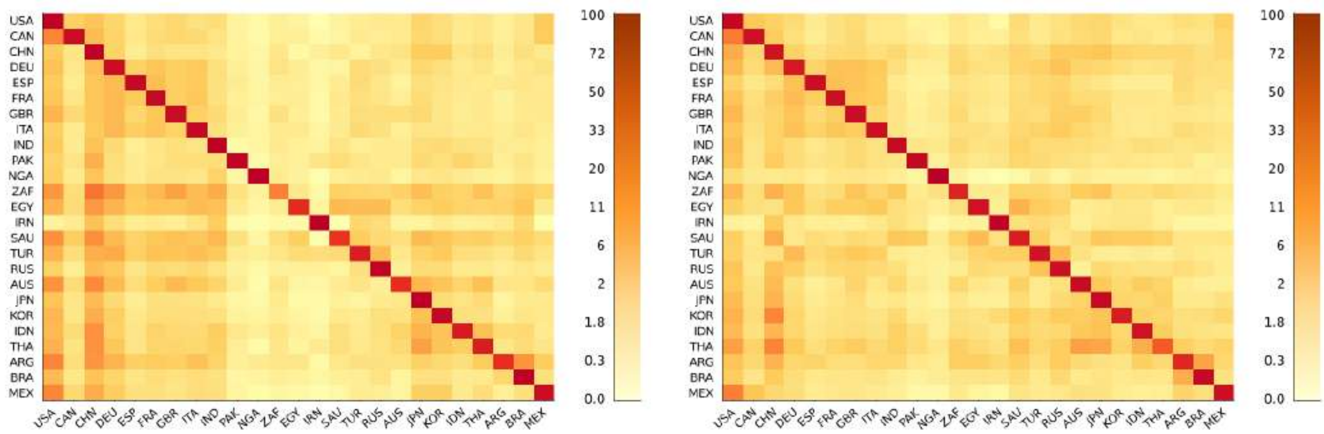


Figure 12: Left: Trade shares matrix, $\mathbf{S}$. Right: Income shares matrix $\mathbf{T}$. See main text for details.

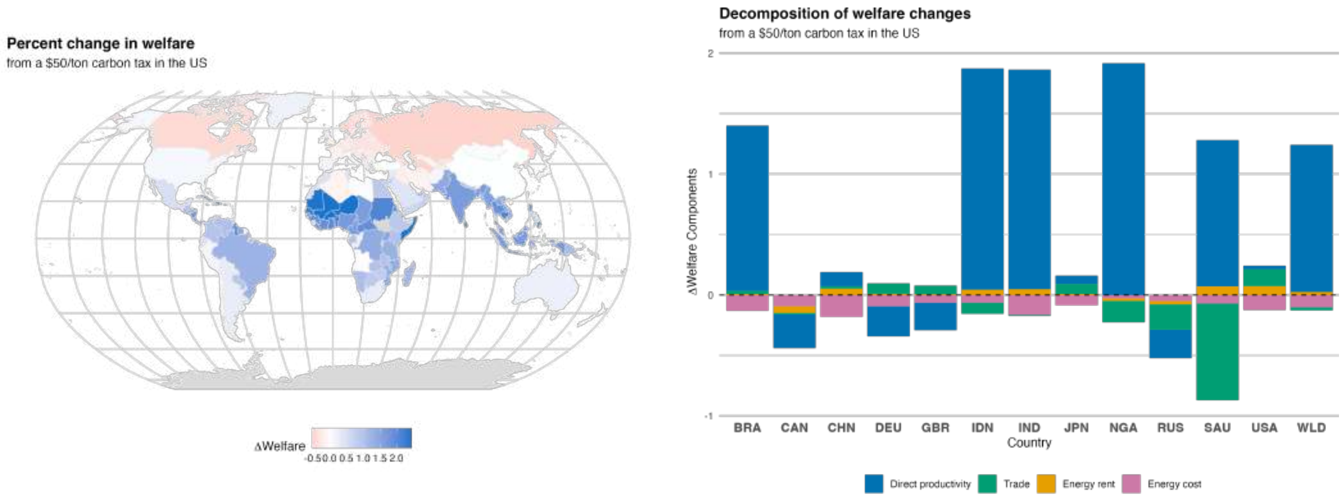


Figure 13: The welfare effects of a $50 carbon tax imposed unilaterally by the United States. Left: map of global welfare changes, in % terms. Red countries lose, while blue countries win. Right: decomposition of welfare changes for several major economies.

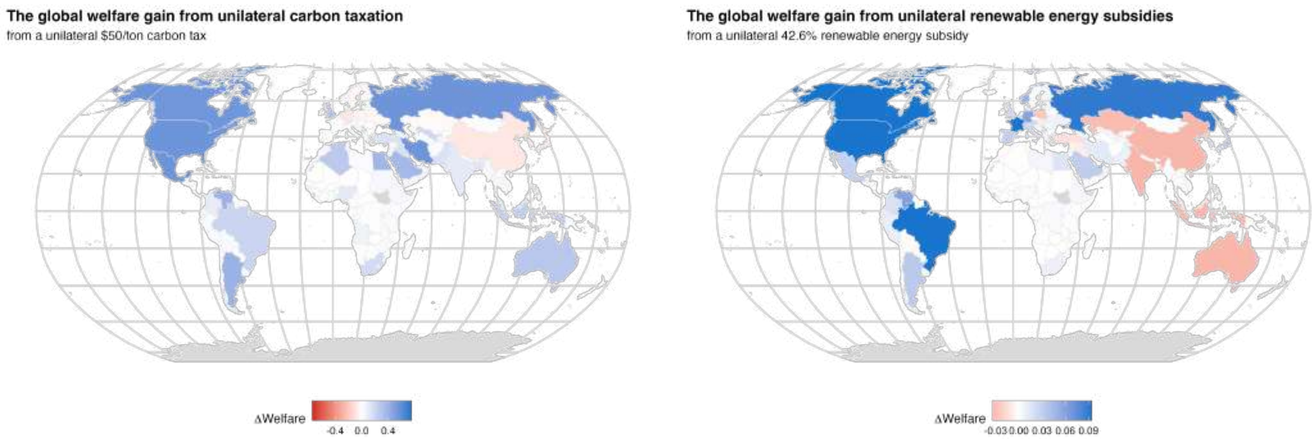


Figure 14: Left: Map of the global utilitarian welfare change associated with each country's unilateral $50 carbon tax. Right: the same, for a 42.6% renewable energy subsidy.

B Welfare decomposition

B.1 Model summary

First let us summarize the model, as presented above

$$
\begin{aligned}
c_i \mathbb{P}_i = x_i = w_i \ell_i + \pi_i^x + \mathrm{t}_i^{ls} &= \mathrm{p}_i z_i \mathcal{D}_i(T_i) F(e_i, \ell_i) - q_t^e e_i + \frac{1}{\mathcal{P}_i}\Big(q^e e_i^x - \mathrm{p}_i \mathcal{C}^f(\mathcal{P}_i e_i^x, \mathcal{R}_i)\Big) + \mathrm{t}_i^{ls} \\
\mathcal{P}_i \mathrm{p}_i y_i &= \sum_{k \in \mathbb{I}} \mathcal{P}_k s_{ki} \frac{v_k}{1 + \mathrm{t}_{ki}} \\
v_i &= \mathrm{p}_i y_i + q^f(e_i^x - e_i^f) + \mathrm{t}_i^{ls} \\
\pi_i^f &= \frac{1}{\mathcal{P}_i} \frac{\nu_i^f \bar{\nu}^{-1/\nu_i^f}}{1+\nu_i^f} \mathcal{R}_i (q^f)^{1+\frac{1}{\nu_i^f}} \mathrm{p}_i^{-1/\nu_i^f} \\
\pi_i^c &= \frac{1}{\mathcal{P}_i} \frac{\nu_i^c \bar{\nu}^{-1/\nu_i^c}}{1+\nu_i^c} (q_i^c)^{1+\frac{1}{\nu_i^c}} \mathrm{p}_i^{-1/\nu_i^c} \\
\pi_i^r &= \frac{1}{\mathcal{P}_i} \frac{\nu_i^r \bar{\nu}^{-1/\nu_i^r}}{1+\nu_i^r} (q_i^r)^{1+\frac{1}{\nu_i^r}} \mathrm{p}_i^{-1/\nu_i^r} \\
\sum_k \mathcal{P}_i e_i^f &= \sum_k e_i^x = (q^f)^{1/\nu} \sum_k \mathcal{R}_i \bar{\nu}_i^{-1/\nu} \mathrm{p}_i^{-1/\nu} \\
F_i(\varepsilon(e^f, e^c, e^r), \ell) &= \Big[(1-\epsilon)^{\frac{1}{\sigma_y}} \big(\bar{k}^{\alpha} \ell^{1-\alpha}\big)^{\frac{\sigma_y - 1}{\sigma_y}} + \epsilon^{\frac{1}{\sigma_y}} \big(z_i^e \;\varepsilon_i(e^f, e^c, e^r)\big)^{\frac{\sigma_y-1}{\sigma_y}}\Big]^{\frac{\sigma_y}{\sigma_y - 1}} \\
\varepsilon(e^f, e^c, e^r) &= \Big[(\omega_i^f)^{\frac{1}{\sigma_e}} (e^f)^{\frac{\sigma_e - 1}{\sigma_e}} + (\omega_i^c)^{\frac{1}{\sigma_e}} (e^c)^{\frac{\sigma_e - 1}{\sigma_e}} + (\omega_i^r)^{\frac{1}{\sigma_e}} (e^r)^{\frac{\sigma_e - 1}{\sigma_e}}\Big]^{\frac{\sigma_e}{\sigma_e - 1}}
\end{aligned}
$$

B.2 Change in welfare – experiments

We compute the change in the welfare of each country for different experiments. The model is linearized around an equilibrium where climate change is not realized yet $\mathcal{T} = 0$, and where the policies are identical to the "status quo": $\mathrm{t}^\varepsilon = \bar{\mathrm{t}}^\varepsilon = 0$ and $\mathrm{t}_{ij}^b = \bar{\mathrm{t}}_{ij}^b$. As a result, this corresponds to the competitive equilibrium.

We consider first-order deviations where we increase either (i) the impact of climate change and (ii) climate policy instruments by a small amount. To save on notation, we denote $d \ln x_i = \frac{dx_i}{x_i}$ – with a slight abuse of notation[18]

Effects of climate change

I consider a first-order change in global warming, which will impact global temperature $\mathcal{T}$, by an amount $d \ln \mathcal{T}$ and hence local temperature T_i and local productivity $d \ln z_i = \frac{dz_i}{z_i}$.

Unilateral climate policies – Carbon tax and Renewable subsidy

In addition, I consider a first-order change in local climate policies: carbon tax t_i^ε and renewable subsidies s_i^ε. As a result, the policy change we consider is $d \ln \mathrm{t}_i^\varepsilon = \frac{d\mathrm{t}_i^\varepsilon}{1+\mathrm{t}_i^\varepsilon} = d\mathrm{t}_i^\varepsilon$ where we consider a multiplicative carbon tax on fossil fuel $q_i^f(1+\mathrm{t}_i^\varepsilon)$. Similalry, for small renewable subsidy: $d \ln \mathrm{s}_i^\varepsilon = d\mathrm{s}_i^\varepsilon$.

[18] This is the case, for example, when $x_i < 0$ or change sign.

I consider the case where those policies are implemented unilaterally, i.e., for country i but not for country $j \neq i$, and compare the cost of such implementation in the presence of trade leakage.

Coordinated climate policies

Then, we consider the case of coordinated climate policies, where a large set of countries implement the policy jointly. I consider a set $\mathcal{J}$ of J countries that are linked by a coordination mechanism, e.g. a climate agreement. In matrix notation, these changes in carbon tax are noted:

$$\mathbf{J} d\mathbf{t}^{\varepsilon} = \left\{ \mathbb{1}_{\{i \in \mathcal{J}\}}\, d \ln \mathrm{t}_i^{\varepsilon} \right\}_i$$

with $\mathbf{J} = \mathrm{J}_i = \mathbb{1}_{\{i \in \mathcal{J}\}}$ the column vector that is one if $i \in \mathcal{J}$ and zero otherwise.

Carbon Border Adjustment Mechanism / Carbon tariffs

I also consider a first-order change in tariffs t_{ij}^b, imposed by country i on the goods from country j. As a result, the policy change we consider is $d \ln \mathrm{t}_{ij}^b = \frac{d\mathrm{t}_{ij}^b}{1+\mathrm{t}_{ij}^b} = d\mathrm{t}_{ij}^b$ for multiplicative tariffs. The tariff scale with carbon intensity of the country one imports from: $\mathrm{t}_{ij}^b = \xi_j^{\varepsilon} \mathrm{t}_i^{\varepsilon}$ with the carbon intensity of country j, $\xi_j^{\varepsilon} = \frac{\varepsilon_j}{y_j \mathrm{p}_j}$ with ε_j the per capita carbon emissions.

I consider three cases: First, this policy is implemented unilaterally for country i but not for country $j \neq i$. Second, the policy is implemented in coordination with the carbon tax at the same carbon price t^{ε}, again unilaterally. Third, this carbon tax + carbon tariff policy is coordinated among countries within a club $\mathcal{J}$, e.g. European Union or OECD countries against non-member countries, $\overline{\mathbf{J}} \equiv \mathrm{J}_{ij} = \mathbb{1}_{\{i \in \mathcal{J}, j \notin \mathcal{J}\}}$.

Welfare change

I now compute the welfare of individual country i, defined as the indirect utility, accounting for change in consumption and climate damages: $\mathcal{U}_i = u(\{c_{ij}\}_j, T_i) = u\big(c_i \mathcal{D}_i^u(T_i)\big)$. This changes writes as:

$$d\mathcal{U}_i = du\Big(c_i \mathcal{D}_i^u\Big) = u'(c_i \mathcal{D}_i^u)(c_i \mathcal{D}_i^u)\Big(\frac{dc_i}{c_i} + \frac{d\mathcal{D}_i^u}{\mathcal{D}_i^u}\Big) = u'(\widetilde{c}_i)\widetilde{c}_i\Big(\frac{dx_i}{x_i} - \frac{d\mathbb{P}_i}{\mathbb{P}_i} + \frac{d\mathcal{D}_i^u}{\mathcal{D}_i^u}\Big)$$

with $x_i = c_i \mathbb{P}_i$ the consumption expenditure and $\widetilde{c}_i = c_i \mathcal{D}_i^u$. As a result, in the main text, we display the result in consumption equivalent:

$$\frac{d\mathcal{U}_i}{u'(\widetilde{c}_i)\widetilde{c}_i} = \Big(\frac{dx_i}{x_i} - \frac{d\mathbb{P}_i}{\mathbb{P}_i} + \frac{d\mathcal{D}_i^u}{\mathcal{D}_i^u}\Big)$$

B.3 Climate externality

To see the effects of a change in emissions and carbon policies on climate, we unpack the damage $d\mathcal{D}_i$. In this section, we consider the following simplified climate system:

$$T_i = \Delta_i \mathcal{T} = \Delta_i \chi \mathcal{S}$$
$$\mathcal{S} = \mathcal{S}_0 + \mathbb{T}\mathcal{E} = \mathcal{S}_0 + \xi^f E^f + \xi^c E^c$$

where

$$\mathcal{E} = \sum_i \mathcal{P}_i(\xi^f e_i^f + \xi^c e_i^c)$$

is a representation of yearly emissions $\mathcal{E}$ due to oil-gas and coal. We scale those yearly emissions by a factor $\mathbb{T}$ to represent a specific horizon – say 50 or 100 years. Moreover, the parameters ξ^f and ξ^c represent the carbon contents for different fossil fuels per unit of energy supplied.

We use a damage function, conventionally used in Integrated Assessment models, with curvature δ and slope γ:

$$\mathcal{D}_i^y(T_i) = e^{-\frac{\gamma^y}{1+\delta}(T_i - T_i^\star)^{1+\delta}}$$

and similarly for $\mathcal{D}_i^u(T_i)$. The linear approximation of this climate system implies:

$$\frac{d\mathcal{D}_i^y}{\mathcal{D}_i^y} = -\gamma^y(T_i - T_i^\star)^\delta dT_i = -\gamma(T_i - T_i^\star)^\delta T_i \frac{dT_i}{T_i}$$

we notice that despite the approximation being linear – and hence abstracting from the curvature δ of damages – we still have that a higher curvature imply more heterogeneous damages between warm and cold regions based on $(T_i - T_i^\star)^\delta$.

Regarding the change in temperature caused by emissions, we get:

$$\begin{aligned}
& dT_i = \Delta_i \chi d\mathcal{S} = \Delta_i \chi(\xi^f dE^f + \xi^c dE^c) \\
\Rightarrow \quad & \frac{dT_i}{T_i} = \frac{\Delta_i \chi d\mathcal{E}}{\Delta_i \chi(\mathcal{S}_0 + \mathbb{T}\mathcal{E})} = s^{E/S}\Big(s^{f/E}\frac{dE^f}{E^f} + s^{c/E}\frac{dE^c}{E^c}\Big) \\
& \qquad \text{with } s^{E/S} = \frac{\mathbb{T}\mathcal{E}}{\mathcal{S}_0 + \mathbb{T}\mathcal{E}} \qquad s^{f/E} = \frac{\xi^f E^f}{\mathcal{E}} \qquad s^{c/E} = \frac{\xi^c E^c}{\mathcal{E}}
\end{aligned}$$

As a result, to summarize, the change in damage depends on the total energy used in fossil (oil-gas) and coal.

$$d\ln \mathcal{D}_i^y = -\bar{\gamma}_i^y\big(s^{f/E} d\ln E^f + s^{c/E} d\ln E^c\big) \qquad \bar{\gamma}_i^y = \gamma(T_i - T_i^\star)T_i\, s^{E/S}$$

and similarly for $d\ln \mathcal{D}_i^y$ where $\bar{\gamma}_i^y$ and $\bar{\gamma}_i^y$ summarize in simple parameters – as sufficient statistics – the heterogeneous impacts of climate change on output and utility.

B.4 Production

We now derive the impact of changes in prices and quantities on welfare through the budget constraint. First, we define several objects – like shares – that are relevant for the decomposition:

- Energy share in production: $s_i^e = \frac{e_i q_i^e}{y_i \mathrm{p}_i}$
- Fossil share in energy mix $s_i^f = \frac{e_i^f q^f}{e_i q_i^e}$ and similarly $s_i^c = \frac{e_i^c q_i^c}{e_i q_i^e}$ and $s_i^r = \frac{e_i^r q_i^r}{e_i q_i^e}$
- Production share/rent share in GDP: $\eta_i^y = \frac{y_i \mathrm{p}_i}{y_i \mathrm{p}_i + \pi_i^f} = 1 - \eta_i^\pi$
- Consumption share in GDP: $\eta_i^c = \frac{x_i}{y_i \mathrm{p}_i + \pi_i^f}$
- Consumption as a ratio of output: $s_i^{c/y} = \frac{c_i \mathbb{P}_i}{y_i \mathrm{p}_i} = \frac{x_i}{y_i \mathrm{p}_i + \pi_i^f} \frac{y_i \mathrm{p}_i + \pi_i^f}{y_i \mathrm{p}_i} = \frac{\eta_i^c}{1 - \eta_i^\pi} = \frac{\eta_i^c}{\eta_i^y}$,

- Energy share as a ratio of consumption: $\frac{e_i q_i^e}{x_i} = \frac{e_i q_i^e}{y_i \mathrm{p}_i} \frac{y_i \mathrm{p}_i}{y_i \mathrm{p}_i + \pi_i^f} \frac{y_i \mathrm{p}_i + \pi_i^f}{x_i} = s_i^e \frac{\eta_i^y}{\eta_i^c}$
- Profit share as a ratio of consumption: $\frac{\pi_i^f}{x_i} = \frac{\pi_i^f}{y_i \mathrm{p}_i + \pi_i^f} \frac{y_i \mathrm{p}_i + \pi_i^f}{x_i} = \frac{\eta_i^\pi}{\eta_i^c}$
- The share of GDP of energy imports and exports, with $v_i = \mathrm{p}_i y_i + q^f(e_i^x - e_i^f)$ and $v^y = \frac{\mathrm{p}_i y_i}{v_i}$, $v^{e^x} = \frac{q^f e_i^x}{v_i}$, $v^{ef} = \frac{q^f e_i^f}{v_i}$ and $v^{ne} = \frac{q^f(e_i^x - e_i^f)}{v_k}$.

Returning to our decomposition, we start from the budget constraint:

$$
\begin{aligned}
c_i \mathbb{P}_i = x_i &= \mathrm{p}_i z_i \mathcal{D}_i(T_i) F(e_i, \ell_i) - q_i^e e_i + \frac{1}{\mathcal{P}_i} \sum_\ell \Big(q^\ell \bar{e}_i^\ell - \mathrm{p}_i \mathcal{C}^\ell(e_i^\ell) \Big) + \mathrm{t}_i^{ls} \\
&= \mathrm{p}_i z_i \mathcal{D}_i(T_i) F(e_i, \ell_i) - \big(q^f(1{+}\mathrm{t}_i^\varepsilon) e_i^f + q_i^c(1{+}\mathrm{t}_i^\varepsilon) e_i^c + q_i^r(1{-}\mathrm{s}_i^\varepsilon) e_i^r\big) + \\
&\qquad \frac{1}{\mathcal{P}_i} \sum_\ell \Big(q^\ell \bar{e}_i^\ell - \mathrm{p}_i \mathcal{C}^\ell(e_i^\ell) \Big) + \widetilde{\mathrm{t}}_i^{ls} + q^f \mathrm{t}_i^\varepsilon e_i^f + q_i^c \mathrm{t}_i^\varepsilon e_i^c - q_i^r \mathrm{s}_i^\varepsilon e_i^r
\end{aligned}
$$

Since the revenues of the carbon tax and the renewable subsidy are redistributed/taxed lump-sum to the Household, we do not see any direct redistributive effect of carbon taxation, e.g. as the terms $q^f \mathrm{t}_i^\varepsilon e_i^f$ cancel out.

Taking the first-order expansion of the budget constraint, we obtain:

$$
\begin{aligned}
\frac{dc_i}{c_i} = \frac{dx_i}{x_i} - \frac{d\mathbb{P}_i}{\mathbb{P}_i} &= \frac{\mathrm{p}_i y_i}{x_i}\Big(\frac{d\mathrm{p}_i}{\mathrm{p}_i} + \frac{dy_i}{y_i}\Big) - \frac{e_i q_i^e}{x_i}\Big(\frac{e_i^f q^f}{e_i q_i^e}\Big(\frac{de^f}{e^f} + \frac{dq^f}{q^f}\Big) + \frac{e_i^c q^c}{e_i q_i^e}\Big(\frac{de^c}{e^c} + \frac{dq^c}{q^c}\Big) + \frac{e_i^r q^r}{e_i q_i^e}\Big(\frac{de^r}{e^r} + \frac{dq^r}{q^r}\Big)\Big) \\
&\quad + \frac{\pi^f}{x_i}\frac{d\pi_i^f}{\pi^f} + \frac{\widetilde{\mathrm{t}}_i^{ls}}{x_i}\Big(\frac{d\widetilde{\mathrm{t}}_i^{ls}}{\widetilde{\mathrm{t}}_i^{ls}}\Big) \quad - \frac{d\mathbb{P}_i}{\mathbb{P}_i}
\end{aligned}
$$

$$
\begin{aligned}
\frac{dc_i}{c_i} &= \frac{\eta_i^y}{\eta_i^c}\Big(\frac{d\mathrm{p}_i}{\mathrm{p}_i} + \frac{dy_i}{y_i}\Big) - s_i^e \frac{\eta_i^y}{\eta_i^c}\Big(s_i^f\Big(\frac{de^f}{e^f} + \frac{dq^f}{q^f}\Big) + s_i^c\Big(\frac{de^c}{e^c} + \frac{dq^c}{q^c}\Big) + s_i^r\Big(\frac{de^r}{e^r} + \frac{dq^r}{q^r}\Big)\Big) \\
&\quad + \frac{\eta_i^{\pi f}}{\eta_i^c}\frac{d\pi_i^f}{\pi_i^f} + \frac{\eta_i^{\pi c}}{\eta_i^c}\frac{d\pi_i^c}{\pi_i^c} + \frac{\eta_i^{\pi r}}{\eta_i^c}\frac{d\pi_i^r}{\pi_i^r} + \frac{\widetilde{\mathrm{t}}_i^{ls}}{x_i}\Big(\frac{d\widetilde{\mathrm{t}}_i^{ls}}{\widetilde{\mathrm{t}}_i^{ls}}\Big) \quad - \frac{d\mathbb{P}_i}{\mathbb{P}_i}
\end{aligned}
$$

First, using output changes, approximating the production function:

$$
\frac{dy_i}{y_i} = \frac{d\mathcal{D}_i^y}{\mathcal{D}_i} + \frac{MPe_i e_i}{y_i}\frac{de_i}{e_i} = \frac{d\mathcal{D}_i^y}{\mathcal{D}_i^y} + \ s_i^e\Big[s_i^f \frac{de_i^f}{e_i^f} + s_i^c \frac{de_i^c}{e_i^c} + s_i^r \frac{de_i^r}{e_i^r}\Big]
$$

we see that Hulten's theorem implies a first-order impact of a change in energy price that scales with the share of energy in production s_i^e, which is typically around $5-10\%$ and the share of fossils in the energy mix s_i^f, s_i^c, which sum to above 85%.

B.5 Energy markets – Profits, and prices for Coal and Renewable

Using the fossil-energy firm problem, we get the profit change as a function of the price:

$$
\frac{d\pi_i^f}{\pi_i^f} = \Big((1{+}\frac{1}{\nu_i^f})\frac{dq^f}{q^f} - \frac{1}{\nu_i^f}\frac{d\mathrm{p}_i}{\mathrm{p}_i}\Big)
$$

The energy rent is affected by changes in the aggregate fossil energy price dq^f. Since the cost also depends on imported inputs, the prices of goods $\mathbb{P}_i$ also matter for profit and welfare.

Similarly, for coal and renewable, we obtain the same formulation for profit:

$$\frac{d\pi_i^c}{\pi_i^c} = (1+\tfrac{1}{\nu_i^c})\frac{dq_i^c}{q_i^c} - \frac{1}{\nu_i^c}\frac{d\mathrm{p}_i}{\mathrm{p}_i}$$

and similarly for $d\ln\pi_i^r$ as a function of $d\ln q_i^r$.

Now, we use the production function for coal and renewable, which implies the simple supply curve $q_i^c = \mathcal{C}_i^{c\prime}(\bar{e}_i^c)\mathrm{p}_i = (e_i^c)^{\nu_i^c}\mathrm{p}_i$ and $q_i^r = \upsilon_i^r\mathbb{P}_i$, we get

$$\frac{dq^r}{q^r} = \nu_i^r\frac{de_i^r}{e_i^r} + \frac{d\mathrm{p}_i}{\mathrm{p}_i} \qquad \text{and} \qquad \frac{dq^c}{q^c} = \nu_i^c\frac{de_i^c}{e_i^c} + \frac{d\mathrm{p}_i}{\mathrm{p}_i}$$

the price of both coal and renewable energy are directly exposed to changes in the price of the domestic good used in production.

As a result, the profits from coal and renewable can also be rewritten in function of quantities:

$$\frac{d\pi_i^c}{\pi_i^c} = (1+\tfrac{1}{\nu_i^c})\Big[\nu_i^r\frac{de_i^r}{e_i^r} + \frac{d\mathrm{p}_i}{\mathrm{p}_i}\Big] - \frac{1}{\nu_i^c}\frac{d\mathrm{p}_i}{\mathrm{p}_i} = (1+\nu_i^c)\frac{de_i^r}{e_i^r} + \frac{d\mathrm{p}_i}{\mathrm{p}_i}$$

B.6 Returning to the budget/expenditure

Accounting for these different effects dramatically simplifies the change in consumption:

$$\begin{aligned}
\frac{dc_i}{c_i} &= \frac{\eta_i^y}{\eta_i^c}\Big(\frac{d\mathrm{p}_i}{\mathrm{p}_i} + \frac{dy_i}{y_i}\Big) - s_i^e\frac{\eta_i^y}{\eta_i^c}\Big(s_i^f(\frac{de^f}{e^f}+\frac{dq^f}{q^f})+s_i^c(\frac{de^c}{e^c}+\frac{dq^c}{q^c})+s_i^r(\frac{de^r}{e^r}+\frac{dq^r}{q^r})\Big) \\
&\qquad + \sum_\ell \frac{\eta_i^{\pi\ell}}{\eta_i^c}\Big((1+\frac{1}{\nu_i^\ell})\frac{dq^\ell}{q^\ell} - \frac{1}{\nu_i^\ell}\frac{d\mathrm{p}_i}{\mathrm{p}_i}\Big) + \frac{d\tilde{\mathrm{t}}_i^{ls}}{x_i}) \quad - \frac{d\mathbb{P}_i}{\mathbb{P}_i} \\
&= \frac{\eta_i^y}{\eta_i^c}\Big(\frac{d\mathrm{p}_i}{\mathrm{p}_i} + \frac{d\mathcal{D}_i}{\mathcal{D}_i}\Big) - \frac{\eta_i^y}{\eta_i^c}s_i^e\Big[s_i^f\frac{dq^f}{q^f} + s_i^c\frac{dq^c}{q^c} + s_i^r\frac{dq^r}{q^r}\Big] + \sum_\ell \frac{\eta_i^{\pi\ell}}{\eta_i^c}\Big((1+\frac{1}{\nu_i^\ell})\frac{dq^\ell}{q^\ell} - \frac{1}{\nu_i^\ell}\frac{d\mathrm{p}_i}{\mathrm{p}_i}\Big) + \frac{d\tilde{\mathrm{t}}_i^{ls}}{x_i} \quad - \frac{d\mathbb{P}_i}{\mathbb{P}_i}
\end{aligned}$$

$$\frac{dc_i}{c_i} = \Big[\frac{\eta_i^y}{\eta_i^c} - \sum_\ell \frac{1}{\nu_i^\ell}\frac{\eta_i^{\pi\ell}}{\eta_i^c}\Big]\frac{d\mathrm{p}_i}{\mathrm{p}_i} + \frac{\eta_i^y}{\eta_i^c}\frac{d\mathcal{D}_i^y}{\mathcal{D}_i^y} + \sum_\ell\Big[\frac{\eta_i^{\pi\ell}}{\eta_i^c}\Big(1+\frac{1}{\nu_i^\ell}\Big) - \frac{\eta_i^y}{\eta_i^c}s_i^e s_i^\ell\Big]\frac{dq^\ell}{q^\ell} - \frac{d\mathbb{P}_i}{\mathbb{P}_i} + \frac{d\tilde{\mathrm{t}}_i^{ls}}{x_i}$$

B.7 Energy markets

We now turn to energy where the demand and equilibrium effect on prices will be of first-oder importance for our welfare decomposition.

Energy demand

To examine the demand side of the market, we compute the elasticities of demand for each energy source, which are determined jointly by the firm First-Order Conditions. Thanks to our nested CES formulation, we can compute the elasticity $\varepsilon^{\ell}_{q^k} = \frac{\partial e^{\ell}_i}{\partial q^k}\frac{q^k}{e^{\ell}_i}$ as :

$$\begin{bmatrix} \varepsilon^f_{q^f} & \varepsilon^f_{q^c} & \varepsilon^f_{q^r} \\ \varepsilon^c_{q^f} & \varepsilon^c_{q^c} & \varepsilon^c_{q^r} \\ \varepsilon^r_{q^f} & \varepsilon^r_{q^c} & \varepsilon^r_{q^r} \end{bmatrix} = (\widetilde{H}^e)^{-1} = -\frac{\sigma^y}{1-s^e}\begin{bmatrix} s^f & s^c & s^r \\ s^f & s^c & s^r \\ s^f & s^c & s^r \end{bmatrix} + \sigma^e \begin{bmatrix} -(1-s^f) & s^c & s^r \\ s^f & -(1-s^c) & s^r \\ s^f & s^c & -(1-s^r) \end{bmatrix}$$

where the first part correspond to the change in aggregate price of energy q^e, since $\frac{\partial q^e_i}{\partial q^k}\frac{q^k}{q^e_i} = s^k_i$, which reduces demands for overall energy, according to elasticity $\frac{\sigma^y}{1-s^e_i}$ where s^e_i is the cost share of energy and σ^y the elasticity between energy and other inputs. Second, the later part summarizes the substitution effect across energy sources, negative along the diagonal and positive out of diagonal, due to positive cross-elasticity in the CES framework.

Moreover, the energy demand also depends on aggregate TFP (and hence climate damage), and the price level at which the final good is sold. As a result, the productivity elasticities and the final good price elasticity write:

$$\begin{bmatrix} \varepsilon^f_z \\ \varepsilon^c_z \\ \varepsilon^r_z \end{bmatrix} = \frac{\sigma^y}{1-s^e}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \qquad \begin{bmatrix} \varepsilon^f_p \\ \varepsilon^c_p \\ \varepsilon^r_p \end{bmatrix} = \frac{\sigma^y}{1-s^e}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

which again is standard in the Nested CES framework.

As a result, we can express the energy demand as a function of the other endogenous variables:

$$\begin{aligned}
d\ln e^f_i &= -(\frac{\sigma^y}{1-s^e_i}s^f_i + (1-s^f_i)\sigma^e)[d\ln q^f + \xi^f \mathrm{J}_i d\ln \mathrm{t}^{\varepsilon}] + (\sigma^e - \frac{\sigma^y}{1-s^e_i})s^c_i[d\ln q^c_i + \xi^c \mathrm{J}_i d\ln \mathrm{t}^{\varepsilon}] + (\sigma^e - \frac{\sigma^y}{1-s^e_i})s^r_i[d\ln q^r_i - \mathrm{J}_i d\ln \mathrm{s}^{\varepsilon}_i] \\
&\quad + \frac{\sigma^y}{1-s^e}d\ln \mathcal{D}_i + \frac{\sigma^y}{1-s^e}d\ln \mathrm{p}_i \\
d\ln e^c_i &= (\sigma^e - \frac{\sigma^y}{1-s^e_i})s^f_i[d\ln q^f + \xi^f \mathrm{J}_i d\ln \mathrm{t}^{\varepsilon}] - (\frac{\sigma^y}{1-s^e_i}s^c_i + (1-s^c_i)\sigma^e)[d\ln q^c_i + \xi^c \mathrm{J}_i d\ln \mathrm{t}^{\varepsilon}] + (\sigma^e - \frac{\sigma^y}{1-s^e_i})s^r_i[d\ln q^r_i - \mathrm{J}_i d\ln \mathrm{s}^{\varepsilon}_i] \\
&\quad + \frac{\sigma^y}{1-s^e}d\ln \mathcal{D}_i + \frac{\sigma^y}{1-s^e}d\ln \mathrm{p}_i \\
d\ln e^r_i &= (\sigma^e - \frac{\sigma^y}{1-s^e_i})s^f_i[d\ln q^f + \xi^f \mathrm{J}_i d\ln \mathrm{t}^{\varepsilon}] + (\sigma^e - \frac{\sigma^y}{1-s^e_i})s^c_i[d\ln q^c_i + \xi^c \mathrm{J}_i d\ln \mathrm{t}^{\varepsilon}] - (\frac{\sigma^y}{1-s^e_i}s^r_i + (1-s^r_i)\sigma^e)[d\ln q^r_i - \mathrm{J}_i d\ln \mathrm{s}^{\varepsilon}_i] \\
&\quad + \frac{\sigma^y}{1-s^e}d\ln \mathcal{D}_i + \frac{\sigma^y}{1-s^e}d\ln \mathrm{p}_i
\end{aligned}$$

Those endogenous energy demands can be reintegrated into the production function to obtain the change in output as a function of good prices, energy prices, and productivity:

$$
\begin{aligned}
d\ln y_i &= d\ln \mathcal{D}_i + s_i^e\big[s_i^f d\ln e_i^f + s_i^r d\ln e_i^c + s_i^r d\ln e_i^r\big] \\
&= (1{+}\frac{s_i^e\sigma^y}{1{-}s_i^e})d\ln \mathcal{D}_i + \frac{s_i^e\sigma^y}{1{-}s_i^e} d\ln \mathrm{p}_i - s_i^e\frac{\sigma^y}{1{-}s_i^e}s_i^f[\xi^f d\ln q^f + d\ln \mathrm{t}_i^{\varepsilon}] \\
&\qquad - s_i^e\frac{\sigma^y}{1{-}s_i^e}s_i^r[d\ln q_i^c + \xi^c d\ln \mathrm{t}_i^{\varepsilon}] - s_i^e\frac{\sigma^y}{1{-}s_i^e}s_i^r[d\ln q_i^r - \mathrm{J}_i d\ln \mathrm{s}_i^{\varepsilon}]
\end{aligned}
$$

$$
\begin{aligned}
d\ln y_i &= \alpha^{y,z} d\ln z_i + \alpha^{y,p} d\ln \mathrm{p}_i - \alpha^{y,qf}\big[d\ln q^f + d\ln \mathrm{t}_i^{\varepsilon}\big] - \alpha^{y,qc}\big[d\ln q_i^c + d\ln \mathrm{t}_i^{\varepsilon}\big] - \alpha^{y,qr}\big[d\ln q_i^r - d\ln \mathrm{s}_i^{\varepsilon}\big] \\
\alpha_i^{y,z} &= 1{+}\frac{s_i^e\sigma^y}{1{-}s_i^e} \qquad\qquad \alpha_i^{y,p} = \frac{s_i^e\sigma^y}{1{-}s_i^e} \\
\alpha_i^{y,qf} &= s_i^e\frac{\sigma^y}{1{-}s^e}s_i^f \qquad \alpha_i^{y,qc} = s_i^e\frac{\sigma^y}{1{-}s^e}s_i^c \qquad \alpha_i^{y,qr} = s_i^e\frac{\sigma^y}{1{-}s_i^e}s_i^r
\end{aligned}
$$

$$
\begin{aligned}
d\ln y_i &= \alpha^{y,z} d\ln z_i + \alpha^{y,p} d\ln \mathrm{p}_i - \alpha^{y,qf} d\ln q^f \\
&\quad - (\xi^f\alpha^{y,qf}{+}\xi^c\alpha^{y,qc})d\ln \mathrm{t}_i^{\varepsilon} + \alpha^{y,qr} d\ln \mathrm{s}_i^{\varepsilon} - \alpha^{y,qc} d\ln q_i^c - \alpha^{y,qr} d\ln q_i^r
\end{aligned}
$$

where this last equation uses the supply curve of coal and renewable. We can see the exposure of country i's output of carbon tax: $\xi^f\alpha^{y,qf}{+}\xi^c\alpha^{y,qc} = s_i^e\frac{\sigma^y}{1-s^e}(\xi^f s_i^f + \xi^c s_i^c)$, through the price and substitution effect of oil, gas and coal.

Coal and renewable energy markets

We write the demand curves in matrix forms:

$$
dq_i^{c,r} = \begin{bmatrix} d\ln q_i^c \\ d\ln q_i^r \end{bmatrix} = \begin{bmatrix} \nu_i^c & 0 \\ 0 & \nu_i^r \end{bmatrix}\begin{bmatrix} d\ln e_i^c \\ d\ln e_i^r \end{bmatrix} + d\ln \mathrm{p}_i
$$

$$
de_i^{c,r} = \begin{bmatrix} d\ln e_i^c \\ d\ln e_i^r \end{bmatrix} = A\, dq_i^{c,r} + \mathrm{J}_i A\begin{bmatrix} \xi^c d\ln \mathrm{t}_i \\ -d\ln \mathrm{s}_i \end{bmatrix} + (\sigma^e{-}\frac{\sigma^y}{1{-}s_i^e})s_i^f[d\ln q^f + \xi^f \mathrm{J}_i d\ln \mathrm{t}^{\varepsilon}] + \frac{\sigma^y}{1{-}s^e}d\ln \mathcal{D}_i + \frac{\sigma^y}{1{-}s^e}d\ln \mathrm{p}_i
$$

with

$$
A = \begin{bmatrix} -(\frac{\sigma^y}{1-s_i^e}s_i^c + (1{-}s_i^c)\sigma^e) & (\sigma^e - \frac{\sigma^y}{1-s_i^e})s_i^r \\ (\sigma^e - \frac{\sigma^y}{1-s_i^e})s_i^c & -(\frac{\sigma^y}{1-s_i^e}s_i^r + (1{-}s_i^r)\sigma^e) \end{bmatrix}
$$

We can summarize and solve the system, where the vector b compiles the other terms (productivity and oil price effects):

$$
\begin{aligned}
dq_i^{c,r} &= \boldsymbol{\nu} de_i^{c,r} + d\ln \mathrm{p}_i \\
de_i^{c,r} &= A\, dq_i^{c,r} + A d\ln \mathrm{t}_i^{c,r} + \frac{\sigma^y}{1{-}s^e}d\ln \mathrm{p}_i + b_i \\
de_i^{c,r} &= [\mathbb{I} - A\boldsymbol{\nu}]^{-1}A d\ln \mathrm{t}_i^{c,r} + [\mathbb{I} - A\boldsymbol{\nu}]^{-1}\big[A + \frac{\sigma^y}{1{-}s^e}\mathbb{1}\big]d\ln \mathrm{p}_i + [\mathbb{I} - A\boldsymbol{\nu}]^{-1}b_i
\end{aligned}
$$

Solving in matrix form yields the price and quantity expression for coal and renewables.

$$
dq_i^{c,r} = [\mathbb{I} - A\boldsymbol{\nu}]^{-1}\boldsymbol{\nu}A\; d\ln \mathrm{t}_i^{c,r} + [\mathbb{I} - A\boldsymbol{\nu}]^{-1}\boldsymbol{\nu}b_i + [\mathbb{I} - A\boldsymbol{\nu}]^{-1}\big[\frac{\sigma^y}{1{-}s^e}\boldsymbol{\nu} + 1\big]\mathbb{1}d\ln \mathrm{p}_i
$$

with $b_i = (\sigma^e - \frac{\sigma^y}{1-s_i^e}) s_i^f [d\ln q^f + \xi^f \mathrm{J}_i d\ln \mathrm{t}^\varepsilon] + \frac{\sigma^y}{1-s^e} d\ln \mathcal{D}_i$.

We can write the energy and price as follows:

$$
\begin{aligned}
dq_i^c &= -\beta_{c,c}^{q,t}\ \xi^c d\ln \mathrm{t}_i^\varepsilon - \beta_{c,r}^{q,t}\ d\ln \mathrm{s}_i^\varepsilon + \beta^{q,b,c}[(\sigma^e - \frac{\sigma^y}{1-s_i^e}) s_i^f [d\ln q^f + \xi^f \mathrm{J}_i d\ln \mathrm{t}^\varepsilon] + \frac{\sigma^y}{1-s^e} d\ln \mathcal{D}_i] + \beta^{q,p,c} d\ln \mathrm{p}_i \\
dq_i^r &= +\beta_{r,c}^{q,t}\ \xi^c d\ln \mathrm{t}_i^\varepsilon + \beta_{r,r}^{q,t}\ d\ln \mathrm{s}_i^\varepsilon + \beta^{q,b,r}\Big[(\sigma^e - \frac{\sigma^y}{1-s_i^e}) s_i^f [d\ln q^f + \xi^f \mathrm{J}_i d\ln \mathrm{t}^\varepsilon] + \frac{\sigma^y}{1-s^e} d\ln \mathcal{D}_i\Big] + \beta^{q,p,r} d\ln \mathrm{p}_i
\end{aligned}
$$

Where β^q are complicated parameters function of A and $\boldsymbol{\nu}$. For $\nu_i^r = \nu_i^r = 0$, we can simplify to have $\beta_{\ell,\ell'}^{q,t} = \beta^{q,b,\ell} = 0$ and $\beta^{q,p,\ell} = 1$, where we simply obtain $d\ln q_i^c = d\ln q_i^r = d\ln \mathrm{p}_i$. For quantities, we get the demand curve as a function of the policies:

$$
\begin{aligned}
de_i^c &= -\beta_{c,c}^{e,t}\ \xi^c d\ln \mathrm{t}_i^\varepsilon - \beta_{c,r}^{e,t}\ d\ln \mathrm{s}_i^\varepsilon + \beta^{e,p,c} d\ln \mathrm{p}_i + \beta^{e,b,c}\Big[(\sigma^e - \frac{\sigma^y}{1-s_i^e}) s_i^f [d\ln q^f + \xi^f \mathrm{J}_i d\ln \mathrm{t}^\varepsilon] + \frac{\sigma^y}{1-s^e} d\ln \mathcal{D}_i\Big] \\
de_i^r &= \beta_{r,c}^{e,t}\ \xi^c d\ln \mathrm{t}_i^\varepsilon + \beta_{r,r}^{e,t}\ d\ln \mathrm{s}_i^\varepsilon + \beta^{e,p,r} d\ln \mathrm{p}_i + \beta^{e,b,r}\Big[(\sigma^e - \frac{\sigma^y}{1-s_i^e}) s_i^f [d\ln q^f + \xi^f \mathrm{J}_i d\ln \mathrm{t}^\varepsilon] + \frac{\sigma^y}{1-s^e} d\ln \mathcal{D}_i\Big]
\end{aligned}
$$

where, β^e are again complicated parameters function of A and $\boldsymbol{\nu}$. Again, with $\nu_i^r = \nu_i^r = 0$, we obtain $\beta_{\ell,\ell'}^{e,t} = A_{\ell,\ell'}$ and $\beta^{e,p,\ell} = \sum_{\ell'} A_{\ell,\ell'} + \frac{\sigma^y}{1-s^e}, \forall \ell$ and $\beta^{e,b,\ell} = 1, \forall \ell$

As a result, the energy demand for fossil can also be rewritten:

$$
\begin{aligned}
d\ln e_i^f &= -(\frac{\sigma^y}{1-s_i^e} s_i^f + (1-s_i^f)\sigma^e)[d\ln q^f + \xi^f \mathrm{J}_i d\ln \mathrm{t}^\varepsilon] + (\sigma^e - \frac{\sigma^y}{1-s_i^e}) s_i^c [d\ln q_i^c + \xi^c \mathrm{J}_i d\ln \mathrm{t}^\varepsilon] + (\sigma^e - \frac{\sigma^y}{1-s_i^e}) s_i^r [d\ln q_i^r - \mathrm{J} \\
&\qquad + \frac{\sigma^y}{1-s^e} d\ln \mathcal{D}_i + \frac{\sigma^y}{1-s^e} d\ln \mathrm{p}_i \\
dq_i^c &= -\beta_{c,c}^{q,t}\ \xi^c d\ln \mathrm{t}_i^\varepsilon - \beta_{c,r}^{q,t}\ d\ln \mathrm{s}_i^\varepsilon + \beta^{q,b,c}[(\sigma^e - \frac{\sigma^y}{1-s_i^e}) s_i^f [d\ln q^f + \xi^f \mathrm{J}_i d\ln \mathrm{t}^\varepsilon] + \frac{\sigma^y}{1-s^e} d\ln \mathcal{D}_i] + \beta^{q,p,c} d\ln \mathrm{p}_i \\
dq_i^r &= \beta_{r,c}^{q,t}\ \xi^c d\ln \mathrm{t}_i^\varepsilon + \beta_{r,r}^{q,t}\ d\ln \mathrm{s}_i^\varepsilon + \beta^{q,b,r}\Big[(\sigma^e - \frac{\sigma^y}{1-s_i^e}) s_i^f [d\ln q^f + \xi^f \mathrm{J}_i d\ln \mathrm{t}^\varepsilon] + \frac{\sigma^y}{1-s^e} d\ln \mathcal{D}_i\Big] + \beta^{q,p,r} d\ln \mathrm{p}_i
\end{aligned}
$$

$$
\begin{aligned}
d\ln e_i^f &= \Big\{ -(\frac{\sigma^y}{1-s_i^e} s_i^f + (1-s_i^f)\sigma^e) + \Big[s_i^c \beta^{q,b,c} + s_i^r \beta^{q,b,r}\Big](\sigma^e - \frac{\sigma^y}{1-s_i^e})^2 s_i^f \Big\}[d\ln q^f + \xi^f \mathrm{J}_i d\ln \mathrm{t}^\varepsilon] \\
&\quad + \Big\{ s_i^c - s_i^c \beta_{c,c}^{q,t} + s_i^r \beta_{r,c}^{q,t} \Big\}(\sigma^e - \frac{\sigma^y}{1-s_i^e})\ \xi^c \mathrm{J}_i d\ln \mathrm{t}_i^\varepsilon + \Big\{ -s_i^r - s_i^c \beta_{c,r}^{q,t} + s_i^r \beta_{r,r}^{q,t} \Big\}(\sigma^e - \frac{\sigma^y}{1-s_i^e})\ \mathrm{J}_i d\ln \mathrm{s}_i^\varepsilon \\
&\quad \frac{\sigma^y}{1-s^e}\Big\{1 + (\sigma^e - \frac{\sigma^y}{1-s_i^e})\Big[s_i^c \beta^{q,b,c} + s_i^r \beta^{q,b,r}\Big]\Big\} d\ln \mathcal{D}_i \ + \ \Big\{\frac{\sigma^y}{1-s^e} + (\sigma^e - \frac{\sigma^y}{1-s_i^e})\Big[s_i^c \beta^{q,p,c} + s_i^r \beta^{q,p,r}\Big]\Big\} d\ln \mathrm{p}_i \\
d\ln e_i^f &= -\beta_{f,f,i}^{e,q}\,[d\ln q^f + \xi^f \mathrm{J}_i d\ln \mathrm{t}^\varepsilon] + \beta_{f,c,i}^{e,t}\ \xi^c \mathrm{J}_i d\ln \mathrm{t}_i^\varepsilon - \beta_{f,r,i}^{e,t}\ \mathrm{J}_i d\ln \mathrm{s}_i^\varepsilon + \beta_i^{e,d,f}\ d\ln \mathcal{D}_i \ + \beta_i^{e,d,f}\ d\ln \mathrm{p}_i
\end{aligned}
$$

This equation shows the layers of general equilibrium effects happening in the three energy markets. An increase in the price of fossil (oil-gas) decreases directly the oil-gas quantity consumed. However, it also creates substitution effects, as it now also increases both the demand and hence the price of coal and renewable – with magnitude $\beta^{q,b,c}$ and $\beta^{q,b,r}$ which then triggers a substitution effect away from those sources, and toward fossil, which then mitigates the drop in oil-gas demand. Similar effects arise for the carbon tax on coal $\xi^c \mathrm{t}_i^\varepsilon$ which increases demand for oil-gas, or subsidy for renewables s_i^ε that decreases it, accounting for all reallocation channels.

If we assume $\nu_i^c = \nu_i^r = 0$, the coal and renewable supply curves are perfectly elastic, and then, we obtain a simplified formula as a function of primitives:

$$
\begin{aligned}
d\ln e_i^f = & - \underbrace{(\frac{\sigma^y}{1-s_i^e} s_i^f + (1-s_i^f)\sigma^e)}_{=\beta^{e,q}_{f,f}} [d\ln q^f + \xi^f \mathrm{J}_i d\ln \mathrm{t}^\varepsilon] + \underbrace{s_i^c(\sigma^e - \frac{\sigma^y}{1-s_i^e})}_{=\beta^{e,t}_{f,c}} \xi^c \mathrm{J}_i d\ln \mathrm{t}_i^\varepsilon - \underbrace{s_i^r\,(\sigma^e - \frac{\sigma^y}{1-s_i^e})}_{=\beta^{e,t}_{f,r}} \mathrm{J}_i d\ln \mathrm{s}_i^\varepsilon \\
& + \underbrace{\frac{\sigma^y}{1-s^e}}_{=\beta^{e,d,f}} d\ln \mathcal{D}_i + \underbrace{\Big[\frac{\sigma^y}{1-s^e} + (\sigma^e - \frac{\sigma^y}{1-s_i^e})(s_i^c + s_i^r)\Big]}_{=\beta^{e,d,f}} d\ln \mathrm{p}_i
\end{aligned}
$$

Going back to the general case, we can rewrite output – substituting the price of coal and renewable – as a function of policies:

$$
\begin{aligned}
d\ln y_i &= \alpha^{y,z} d\ln z_i + \alpha^{y,p} d\ln \mathrm{p}_i - \alpha^{y,qf} d\ln q^f \\
&\quad - (\xi^f \alpha^{y,qf} + \xi^c \alpha^{y,qc}) d\ln \mathrm{t}_i^\varepsilon + \alpha^{y,qr} d\ln \mathrm{s}_i^\varepsilon - \alpha^{y,qc} d\ln q_i^c - \alpha^{y,qr} d\ln q_i^r \\
dq_i^c &= -\beta^{q,t}_{c,c}\ \xi^c d\ln \mathrm{t}_i^\varepsilon - \beta^{q,t}_{c,r}\ d\ln \mathrm{s}_i^\varepsilon + \beta^{q,b,c}[-(\sigma^e - \frac{\sigma^y}{1-s_i^e}) s_i^f [d\ln q^f + \xi^f \mathrm{J}_i d\ln \mathrm{t}^\varepsilon] + \frac{\sigma^y}{1-s^e} d\ln \mathcal{D}_i] + \beta^{q,p,c} d\ln \mathrm{p}_i \\
dq_i^r &= \beta^{q,t}_{r,c}\ \xi^c d\ln \mathrm{t}_i^\varepsilon + \beta^{q,t}_{r,r}\ d\ln \mathrm{s}_i^\varepsilon + \beta^{q,b,r}\Big[(\sigma^e - \frac{\sigma^y}{1-s_i^e}) s_i^f [d\ln q^f + \xi^f \mathrm{J}_i d\ln \mathrm{t}^\varepsilon] + \frac{\sigma^y}{1-s^e} d\ln \mathcal{D}_i\Big] + \beta^{q,p,r} d\ln \mathrm{p}_i \\
\Rightarrow\ d\ln y_i &= \Big[\alpha^{y,z} - \frac{\sigma^y}{1-s^e}(\alpha^{y,qc}\beta^{q,b,c} + \alpha^{y,qr}\beta^{q,b,r})\Big] d\ln \mathcal{D}_i + \Big[\alpha^{y,p} - (\alpha^{y,qc}\beta^{q,p,c} + \alpha^{y,qr}\beta^{q,p,r})\Big] d\ln \mathrm{p}_i \\
&\quad + \Big[-\alpha^{y,qf} + (\alpha^{y,qc}\beta^{q,b,c} + \alpha^{y,qr}\beta^{q,b,r})(\sigma^e - \frac{\sigma^y}{1-s_i^e}) s_i^f\Big] d\ln q^f \\
&\quad + \Big[-(\xi^f \alpha^{y,qf} + \xi^c \alpha^{y,qc}) + (\alpha^{y,qc}\beta^{q,t}_{c,c} - \alpha^{y,qr}\beta^{q,t}_{r,c})\xi^c + (\alpha^{y,qc}\beta^{q,b,c} + \alpha^{y,qr}\beta^{q,b,r})(\sigma^e - \frac{\sigma^y}{1-s_i^e}) s_i^f \xi^f\Big] d\ln \mathrm{t}_i^\varepsilon \\
&\quad + \Big[\alpha^{y,qr} + (\alpha^{y,qc}\beta^{q,t}_{c,r} - \alpha^{y,qr}\beta^{q,t}_{r,r})\Big] d\ln \mathrm{s}_i^\varepsilon
\end{aligned}
$$

this combines all the demand and supply substitution patterns arising in the three energy markets. To give an idea for the mechanism, take the parameter for the carbon tax t^ε, it shows different effects: (i) the direct impact of carbon taxation on the consumption of oil-gas and coal $(\xi^f \alpha^{y,qf} + \xi^c \alpha^{y,qc})$, (ii) the indirect impact of the decline in coal demand due to this tax $\xi^c \mathrm{t}^\varepsilon$ on respectively the price of coal q_i^c and the price of renewable q_i^r, hence change the input choices: $(\alpha^{y,qc}\beta^{q,t}_{c,c} - \alpha^{y,qr}\beta^{q,t}_{r,c})\xi^c$, and (iii) the indirect impact of the decline in *oil-gas* demand due to the tax $\xi^d \mathrm{t}^\varepsilon$ on the price of coal q_i^c and the price of renewable q_i^r respectively, i.e. $(\alpha^{y,qc}\beta^{q,b,c} + \alpha^{y,qr}\beta^{q,b,r})(\sigma^e - \frac{\sigma^y}{1-s_i^e}) s_i^f \xi^f$, the price of energies c and r affecting output with magniture $\alpha^{y,qc}$ and $\alpha^{y,qr}$ respectively.

To save on notation, I compile these effects under the new parameters δ

$$
d\ln y_i = \delta^{y,z} d\ln \mathcal{D}_i + \delta^{y,p} d\ln \mathrm{p}_i - \delta^{y,qf} d\ln q^f - \delta^{y,t\varepsilon} d\ln \mathrm{t}_i^\varepsilon + \delta^{y,s\varepsilon} d\ln \mathrm{s}_i^\varepsilon
$$

Fossil energy market – demand

The energy demand in fossil is the sum of individual countries demand, where we denote the share of country i in global production $\lambda_i^f = \frac{\mathcal{P}_i e_i^f}{E^f}$

$$
\begin{aligned}
dE^f &= \sum_i \mathcal{P}_i de_i^f \\
d\ln E^f &= \sum_i \lambda_i^f d\ln e_i^f \\
&= -\sum_i \lambda_i^f \big(\frac{\sigma^y}{1-s_i^e} s_i^f + (1-s_i^f)\sigma^e\big)[d\ln q^f + \xi^f \mathrm{J}_i d\ln \mathrm{t}^\varepsilon] + \sum_i \lambda_i^f \big(\sigma^e - \frac{\sigma^y}{1-s_i^e}\big) s_i^c [d\ln q_i^c + \xi^c \mathrm{J}_i d\ln \mathrm{t}^\varepsilon] \\
&\quad + \sum_i \lambda_i^f \big(\sigma^e - \frac{\sigma^y}{1-s_i^e}\big) s_i^r [d\ln q_i^r - \mathrm{J}_i d\ln \mathrm{s}^\varepsilon] + \sum_i \lambda_i^f \frac{\sigma^y}{1-s_i^e} d\ln \mathcal{D}_i + \sum_i \lambda_i^f \frac{\sigma^y}{1-s_i^e} d\ln \mathrm{p}_i
\end{aligned}
$$

We see that carbon taxation decreases demand for oil and gas by direct substitution but can also increase it if the substitution away for coal is strong enough. The first effect dominates the second – up to the first order – if:

$$
\overline{\lambda}_{\mathcal{J}}^{\sigma,f} := \sum_{i\in\mathcal{J}} \lambda_i^f \big(\frac{\sigma^y}{1-s_i^e} s_i^f + (1-s_i^f)\sigma^e\big)\ \xi^f > \sum_{i\in\mathcal{J}} \lambda_i^f \big(\sigma^e - \frac{\sigma^y}{1-s_i^e}\big) s_i^c \xi^c =: \overline{\lambda}_{\mathcal{J}}^{\sigma,c}
$$

which depend, among others, on the covariance $\mathbb{C}\mathrm{ov}_i(\lambda_i^f, 1-s_i^f)$ and $\mathbb{C}\mathrm{ov}_i(\lambda_i^f, s_i^c)$, since the substitution effect is stronger than the income effect $\sigma^e > \sigma^y/(1-s_i^e)$, in most empirically-relevant cases.

However, that simple condition only summarizes the direct effects. When we consider the indirect effects, accounting for the changes in prices for coal and renewable $d\ln q^c$ and $d\ln q^r$.

$$
\begin{aligned}
d\ln E^f &= \sum_i \lambda_i^f d\ln e_i^f \\
&= -\sum_i \lambda_i^f \beta_{f,f,i}^{e,q}\, [d\ln q^f + \xi^f \mathrm{J}_i d\ln \mathrm{t}_i^\varepsilon] + \sum_i \lambda_i^f \beta_{f,c,i}^{e,t}\ \xi^c \mathrm{J}_i d\ln \mathrm{t}_i^\varepsilon \\
&\quad + \sum_i \lambda_i^f \beta_{f,r,i}^{e,t}\ \mathrm{J}_i d\ln \mathrm{s}_i^\varepsilon + \sum_i \lambda_i^f \beta_i^{e,d,f}\ d\ln \mathcal{D}_i\ + \sum_i \lambda_i^f \beta_i^{e,d,f}\ d\ln \mathrm{p}_i
\end{aligned}
$$

The condition becomes:

$$
\begin{aligned}
\overline{\lambda}_{\mathcal{J}}^{\sigma,f} &:= \sum_{i\in\mathcal{J}} \lambda_i^f \Big\{ \big(\frac{\sigma^y}{1-s_i^e} s_i^f + (1-s_i^f)\sigma^e\big) + \Big[s_i^c \beta_i^{q,b,c} + s_i^r \beta_i^{q,b,r}\Big]\big(\sigma^e - \frac{\sigma^y}{1-s_i^e}\big)^2 s_i^f \Big\}\ \xi^f \\
&\qquad > \sum_{i\in\mathcal{J}} \lambda_i^f \Big\{ s_i^c - s_i^c \beta_{c,c}^{q,t} + s_i^r \beta_{r,c}^{q,t} \Big\} \big(\sigma^e - \frac{\sigma^y}{1-s_i^e}\big)\ \xi^c =: \overline{\lambda}_{\mathcal{J}}^{\sigma,c}
\end{aligned}
$$

We obtain that, if $\overline{\lambda}_{\mathcal{J}}^{\sigma,f} > \overline{\lambda}_{\mathcal{J}}^{\sigma,c}$, the direct effect of a carbon tax $\xi^f \mathrm{t}_i^\varepsilon$ on oil-gas outweighs the reallocation from coal to oil-gas due to the larger increase $\xi^c \mathrm{t}_i^\varepsilon$.

Fossil energy market – supply

Now, the energy supply curve can also be recast as the sum of individual extraction $E^f = \sum_i \mathcal{P}_i e_i^f = \sum_i \mathcal{P}_i e_i^x$, and, with the share of fossil production $\lambda_i^x = \mathcal{P}_i e_i^x / E^f$, it hence derives as follow:

$$
\begin{aligned}
e_i^x &= (q^f)^{1/\nu_i} \mathcal{R}_i \bar{\nu}_i^{-1/\nu_i} \mathrm{p}_i^{-1/\nu_i} \\
d\ln E^f &= \sum_i \lambda_i^x d\ln e_i^x = \sum_i \lambda_i^x \frac{1}{\nu_i} [d\ln q^f - d\ln \mathrm{p}_i] \\
\Rightarrow \qquad d\ln q^f &= \bar{\nu} d\ln E^f + \sum_i \lambda_i^x \frac{\bar{\nu}}{\nu_i} d\ln \mathrm{p}_i
\end{aligned}
$$

with the aggregate supply elasticity $\bar{\nu} = \big(\sum_i \lambda_i^x \nu_i^{-1}\big)^{-1}$, that we already encountered in the second best optimal Ramsey policy.

Now, replacing the energy demand quantity $d\ln E^f$ into the energy supply/price curve, we obtain:

$$
\begin{aligned}
d\ln q^f &= \bar{\nu} d\ln E^f + \sum_i \lambda_i^x \frac{\bar{\nu}}{\nu_i} d\ln \mathrm{p}_i \\
&= -\bar{\nu}\overline{\lambda}^{\sigma,f} d\ln q^f + \bar{\nu} \sum_i \lambda_i^{\sigma,f}\, \xi^f \mathrm{J}_i d\ln \mathrm{t}^\varepsilon + \bar{\nu} \sum_i \lambda_i^f \, (\sigma^e - \frac{\sigma^y}{1-s_i^e}) s_i^c [d\ln q_i^c + \xi^c \mathrm{J}_i d\ln \mathrm{t}^\varepsilon] \\
&\qquad + \bar{\nu} \sum_i \lambda_i^f \, (\sigma^e - \frac{\sigma^y}{1-s_i^e}) s_i^r [d\ln q_i^r - d\ln \mathrm{s}_i^\varepsilon] + \bar{\nu} \sum_i \lambda_i^f \frac{\sigma^y}{1-s_i^e} [d\ln \mathcal{D}_i + d\ln \mathrm{p}_i] + \sum_i \lambda_i^x \frac{\bar{\nu}}{\nu_i} d\ln \mathrm{p}_i \\
d\ln q^f &= -\bar{\nu} \underbrace{\sum_i \lambda_i^f \beta_{f,f,i}^{e,q}}_{=\overline{\lambda}^{\sigma,f}} [d\ln q^f + \xi^f \mathrm{J}_i d\ln \mathrm{t}^\varepsilon] + \bar{\nu} \sum_i \lambda_i^f \beta_{f,c,i}^{e,t} \; \xi^c \mathrm{J}_i d\ln \mathrm{t}_i^\varepsilon \\
&\qquad + \bar{\nu} \sum_i \lambda_i^f \beta_{f,r,i}^{e,t} \; \mathrm{J}_i d\ln \mathrm{s}_i^\varepsilon + \bar{\nu} \sum_i \lambda_i^f \beta_i^{e,d,f} \, d\ln \mathcal{D}_i \; + \bar{\nu} \sum_i \lambda_i^f \beta_i^{e,d,f} \, d\ln \mathrm{p}_i + \sum_i \lambda_i^x \frac{\bar{\nu}}{\nu_i} d\ln \mathrm{p}_i \\
d\ln q^f &= \frac{\bar{\nu}}{1+\bar{\nu}\overline{\lambda}^{\sigma,f}} \sum_i \lambda_i^f \mathrm{J}_i \, [-\xi^f \beta_{f,f,i}^{e,q} + \xi^c \beta_{f,c,i}^{e,t}] d\ln \mathrm{t}^\varepsilon + \frac{\bar{\nu}}{1+\bar{\nu}\overline{\lambda}^{\sigma,f}} \sum_i \lambda_i^f \; \beta_{f,r,i}^{e,t} \; \mathrm{J}_i d\ln \mathrm{s}_i^\varepsilon \\
&\qquad + \frac{\bar{\nu}}{1+\bar{\nu}\overline{\lambda}^{\sigma,f}} \sum_i \lambda_i^f \beta_i^{e,d,f} \, d\ln \mathcal{D}_i \; + \frac{\bar{\nu}}{1+\bar{\nu}\overline{\lambda}^{\sigma,f}} \sum_i [\lambda_i^f \beta_i^{e,p,f} \; + \lambda_i^x \frac{\bar{\nu}}{\nu_i}] d\ln \mathrm{p}_i
\end{aligned}
$$

where $\overline{\lambda}^{\sigma,f} = \overline{\lambda}_{\mathbb{I}}^{\sigma,f}$, for $\mathbb{I}$ the whole world. As before, we see that carbon taxation decreases the oil-gas energy price if $\overline{\lambda}_{\mathcal{J}}^{\sigma,f} > \overline{\lambda}_{\mathcal{J}}^{\sigma,c}$. Moreover, we see that a change in the good price $d\ln \mathrm{p}_i$ of all the countries change the aggregate price of oil and gas because it both increases the price of renewable and coal, increases demand for oil-gas by substitutions – the terms $\bar{\nu}\lambda_i^{\sigma,c}$ and $\bar{\nu}\lambda_i^{\sigma,r}$ – and it also increases the price of the input – through the term $\lambda_i^x \frac{\bar{\nu}}{\nu_i}$.

If we assume that again $\nu^c = \nu^r = 0$, this implies:

$$
\begin{aligned}
d\ln q^f = & \frac{\bar{\nu}}{1+\bar{\nu}\overline{\lambda}^{\sigma,f}} \sum_i \lambda_i^f \mathrm{J}_i \Big[-\xi^f (\frac{\sigma^y}{1-s_i^e} s_i^f + (1-s_i^f)\sigma^e) + \xi^c s_i^c (\sigma^e - \frac{\sigma^y}{1-s_i^e}) \Big] d\ln \mathrm{t}^\varepsilon + \frac{\bar{\nu}}{1+\bar{\nu}\overline{\lambda}^{\sigma,f}} \sum_i \lambda_i^f \; s_i^r \; (\sigma^e - \frac{\sigma^y}{1-s_i^e}) \; \mathrm{J}_i d\ln \mathrm{s}_i^\varepsilon \\
& + \frac{\bar{\nu}}{1+\bar{\nu}\overline{\lambda}^{\sigma,f}} \sum_i \lambda_i^f \frac{\sigma^y}{1-s^e} d\ln \mathcal{D}_i \; + \frac{\bar{\nu}}{1+\bar{\nu}\overline{\lambda}^{\sigma,f}} \sum_i \Big[\lambda_i^f \Big(\frac{\sigma^y}{1-s^e} + (\sigma^e - \frac{\sigma^y}{1-s_i^e})(s_i^c + s_i^r)\Big) \; + \lambda_i^x \frac{\bar{\nu}}{\nu_i}\Big] d\ln \mathrm{p}_i
\end{aligned}
$$

Similalry, we can write total energy demand as:

$$d\ln E^f = \frac{1}{1+\bar{\nu}\overline{\lambda}^{\sigma,f}}\sum_i \lambda_i^f \mathrm{J}_i\Big[-\xi^f(\frac{\sigma^y}{1-s_i^e}s_i^f+(1-s_i^f)\sigma^e)+\xi^c s_i^c(\sigma^e-\frac{\sigma^y}{1-s_i^e})\Big]d\ln \mathrm{t}^\varepsilon+\frac{1}{1+\bar{\nu}\overline{\lambda}^{\sigma,f}}\sum_i \lambda_i^f\, s_i^r\,(\sigma^e-\frac{\sigma^y}{1-s_i^e})\ \mathrm{J}_i d\ln \mathrm{s}_i^\varepsilon$$
$$+\frac{1}{1+\bar{\nu}\overline{\lambda}^{\sigma,f}}\sum_i \lambda_i^f\frac{\sigma^y}{1-s^e}d\ln\mathcal{D}_i\ +\frac{1}{1+\bar{\nu}\overline{\lambda}^{\sigma,f}}\sum_i\Big[\lambda_i^f\Big(\frac{\sigma^y}{1-s^e}+(\sigma^e-\frac{\sigma^y}{1-s_i^e})(s_i^c+s_i^r)\Big)\Big]d\ln\mathrm{p}_i$$

B.8 Trade *à la* Armington

To investigate how the price indices $\mathbb{P}_i$ and the good price p_i are determined, we should now consider the market for goods.

$$\mathcal{P}_i\mathrm{p}_i y_i=\sum_{k\in\mathbb{I}}\mathcal{P}_k s_{ki}\frac{v_k}{1+\mathrm{t}_{ki}}$$

Using the CES framework, we obtain that:

$$\mathbb{P}_i=\Big(\sum_j a_{ij}(\tau_{ij}(1+\mathrm{t}_{ij}^b)\mathrm{p}_j)^{1-\theta}\Big)^{\frac{1}{1-\theta}}$$
$$\frac{d\mathbb{P}_i}{\mathbb{P}_i}=\sum_j s_{ij}(\frac{d\mathrm{p}_j}{\mathrm{p}_j}+\frac{d\mathrm{t}_{ij}^b}{1+\mathrm{t}_{ij}^b})$$
$$s_{ij}=\frac{c_{ij}(1+\mathrm{t}_{ij})\tau_{ij}\mathrm{p}_j}{\sum_k c_{ik}(1+\mathrm{t}_{ik})\tau_{ik}\mathrm{p}_k}=a_{ij}\frac{\big((1+\mathrm{t}_{ij})\tau_{ij}\mathrm{p}_j\big)^{1-\theta}}{\sum_k\big((1+\mathrm{t}_{ik})\tau_{ik}\mathrm{p}_k\big)^{1-\theta}}=\Big(\frac{(1+\mathrm{t}_{ij})\tau_{ij}\mathrm{p}_j}{\mathbb{P}_i}\Big)^{1-\theta}$$
$$\frac{ds_{ij}}{s_{ij}}=(\theta-1)\Big(\frac{d\mathbb{P}_i}{\mathbb{P}_i}-(\frac{d\mathrm{p}_j}{\mathrm{p}_j}+\frac{d\mathrm{t}_{ij}^b}{1+\mathrm{t}_{ij}^b})\Big)$$
$$\frac{ds_{ij}}{s_{ij}}=(\theta-1)\Big(\sum_k s_{ik}(\frac{d\mathrm{p}_k}{\mathrm{p}_k}+\frac{d\mathrm{t}_{ik}^b}{1+\mathrm{t}_{ik}^b})-(\frac{d\mathrm{p}_j}{\mathrm{p}_j}+\frac{d\mathrm{t}_{ij}^b}{1+\mathrm{t}_{ij}^b})\Big)$$

Using those formulas, the market clearing linearizes as follows:

$$\frac{d[\frac{s_{ij}v_i}{1+t_{ij}}]}{\frac{s_{ij}v_i}{1+\mathrm{t}_{ij}}}=\Big[d\ln v_i+\theta\sum_k\big(s_{ik}d\ln\mathrm{t}_{ik}-(1+s_{ij})d\ln\mathrm{t}_{ij}\big)+(\theta-1)\sum_{k\neq j}\big(s_{ik}d\ln\mathrm{p}_k-d\ln\mathrm{p}_j\big)\Big]$$
$$\text{with}\qquad v_i=\mathrm{p}_i y_i+q^f(e_i^x-e_i^f)$$

This implies:

$$\mathcal{P}_i\widetilde{v}_i(\frac{d\mathrm{p}_i}{d\mathrm{p}_i}+\frac{dy_i}{y_i})=\sum_k\mathcal{P}_k\frac{s_{ki}v_k}{1+\mathrm{t}_{ki}}d\ln[\frac{s_{ki}v_k}{1+t_{ki}}]$$
$$(\frac{d\mathrm{p}_i}{d\mathrm{p}_i}+\frac{dy_i}{y_i})=\sum_k\frac{\mathcal{P}_k v_k}{\mathcal{P}_i v_i}s_{ki}\Big[d\ln v_k+\theta\sum_h\big(s_{kh}d\ln\mathrm{t}_{kh}-(1+s_{ki})d\ln\mathrm{t}_{ki}\big)+(\theta-1)\sum_h\big(s_{kh}d\ln\mathrm{p}_h-d\ln\mathrm{p}_i\big)\Big]$$
$$(\frac{d\mathrm{p}_i}{d\mathrm{p}_i}+\frac{dy_i}{y_i})=\sum_k\boldsymbol{t}_{ik}\Big[(\frac{\mathrm{p}_k y_k}{v_k})(d\ln\mathrm{p}_k+d\ln y_k)+\frac{q^f e_k^x}{v_k}d\ln e_k^x-\frac{q^f e_k^f}{v_k}d\ln e_k^f+\frac{q^f(e_k^x-e_k^f)}{v_k}d\ln q^f$$
$$+\theta\sum_h\big(s_{kh}d\ln\mathrm{t}_{kh}-(1+s_{ki})d\ln\mathrm{t}_{ki}\big)+(\theta-1)\sum_h\big(s_{kh}d\ln\mathrm{p}_h-d\ln\mathrm{p}_i\big)\Big]$$

with $\boldsymbol{t}_{ik} = \frac{\mathcal{P}_k v_k}{\mathcal{P}_i v_i} s_{ki}$, which is analogous to the same matrix in Kleinman et al. (2024). Using the fact that $\sum_k \boldsymbol{t}_{ik} = 1$ we factorize the p_i. This implies, rewritten in matrix notation:

$$
\begin{aligned}
(\theta d\ln \mathrm{p}_i + d\ln y_i) &= \sum_k \boldsymbol{t}_{ik}\Big[(\frac{\mathrm{p}_k y_k}{v_k})(d\ln \mathrm{p}_k + d\ln y_k) + \frac{q^f e_k^x}{v_k} d\ln e_k^x - \frac{q^f e_k^f}{v_k} d\ln e_k^f + \frac{q^f(e_k^x - e_k^f)}{v_k} d\ln q^f\Big] \\
&\quad + \theta \sum_k \boldsymbol{t}_{ik} \sum_h \big(s_{kh} d\ln \mathrm{t}_{kh} - (1{+}s_{ki}) d\ln \mathrm{t}_{ki}\big) + (\theta{-}1)\sum_k \boldsymbol{t}_{ik} \sum_h s_{kh} d\ln \mathrm{p}_h \\
\theta d\ln \mathrm{p} + d\ln y &= \mathbf{T}v^y[d\ln \mathrm{p} + d\ln y] + \mathbf{T}v^{e^x} d\ln e^x - \mathbf{T}v^{e^f} d\ln e^f + \mathbf{T}v^{ne} d\ln q^f + (\theta{-}1)\mathbf{T}\mathbf{S} d\ln \mathrm{p} \\
&\quad + \theta\big(\mathbf{T}(\mathbf{S}\odot\mathbf{J}\odot d\mathrm{t}^b)\mathbb{1} - \mathrm{diag}[\mathbf{T}(\mathbb{1}{+}\mathbf{S}')\odot(\mathbf{J}\odot d\mathrm{t}^b)']\big) \\
\Big[\theta\mathbf{I}{-}\mathbf{T}\odot v^y - (\theta{-}1)(\mathbf{T}\mathbf{S})\Big] d\ln \mathrm{p} &= (\mathbf{T}\odot v^y{-}\mathbf{I}) d\ln y + \mathbf{T}v^{e^x} d\ln e_k^x - \mathbf{T}v^{e^f} d\ln e_k^f + \mathbf{T}v^{ne} d\ln q^f \\
&\quad + \theta\big(\mathbf{T}(\mathbf{S}\odot\mathbf{J}\odot d\mathrm{t}^b)\mathbb{1} - \mathrm{diag}[\mathbf{T}(\mathbb{1}{+}\mathbf{S}')\odot(\mathbf{J}\odot d\mathrm{t}^b)']\big)
\end{aligned}
$$

with $v^y = \frac{\mathrm{p}_i y_i}{v_i}$, $v^{e^x} = \frac{q^f e_i^x}{v_i}$, $v^{e^f} = \frac{q^f e_i^f}{v_i}$ and $v^{ne} = \frac{q^f(e_i^x - e_i^f)}{v_k}$.

Recall fossil demand and supply and production of goods y_i,

$$
\begin{aligned}
d\ln e_i^f &= -\gamma_{f,f,i}^{e,q}\,[d\ln q^f + \xi^f \mathrm{J}_i d\ln \mathrm{t}^\varepsilon] + \gamma_{f,c,i}^{e,t}\ \xi^c \mathrm{J}_i d\ln \mathrm{t}_i^\varepsilon + \gamma_{f,r,i}^{e,t}\ \mathrm{J}_i d\ln \mathrm{s}_i^\varepsilon + \gamma_i^{e,d,f}\ d\ln \mathcal{D}_i\ + \gamma_i^{e,d,f}\ d\ln \mathrm{p}_i \\
d\ln e_i^x &= \frac{1}{\nu_i}[d\ln q^f - d\ln \mathrm{p}_i] \\
d\ln y_i &= \alpha^{y,z} d\ln z_i + \alpha^{y,p} d\ln \mathrm{p}_i - \alpha^{y,qf} d\ln q^f - (\xi^f \alpha^{y,qf}{+}\xi^c\alpha^{y,qc}) d\ln \mathrm{t}_i^\varepsilon - (\alpha^{y,qc}{+}\alpha^{y,qr}) d\ln \mathrm{p}_i
\end{aligned}
$$

We can replace that in the general equilibrium for prices:

$$
\begin{aligned}
\Big[\theta\mathbf{I}{-}\mathbf{T}\odot v^y - (\theta{-}1)(\mathbf{T}\mathbf{S})\Big] d\ln \mathrm{p} &= (\mathbf{T}v^y - \mathbb{I}) d\ln y + \mathbf{T}v^{e^x}\frac{1}{\nu^f}[d\ln q^f - d\ln \mathrm{p}] \\
&\quad - \mathbf{T}v^{e^f}\big[-\beta_{f,f}^{e,q}\, d\ln q^f + \big(-\beta_{f,f}^{e,q}\xi^f + \beta_{f,c}^{e,t}\ \xi^c\big)\mathrm{J} d\ln \mathrm{t}^\varepsilon - \beta_{f,r}^{e,t}\ \mathrm{J} d\ln \mathrm{s}^\varepsilon + \beta^{e,d,f}\, d\ln \mathcal{D}^y\ + \\
&\quad + \mathbf{T}v^{ne} d\ln q^f
\end{aligned}
$$

Rewriting and rearranging:

$$
\begin{aligned}
\Big[\theta\mathbf{I}{-}\mathbf{T}\odot v^y - (\theta{-}1)(\mathbf{T}\mathbf{S}){+}\mathbf{T}\odot(v^{e^f}\odot\beta^{e,d,f}{+}v^{e^x}\odot\tfrac{1}{\nu^f})\Big] d\ln \mathrm{p} &= (\mathbf{T}\odot v^y{-}\mathbf{I}) d\ln y \\
&\quad + \mathbf{T}\odot[v^{e^x}\frac{1}{\nu^f} - v^{e^f}\beta_{f,f}^{e,q}\ + v^{ne}] d\ln q^f - \mathbf{T}v^{e^f}\beta^{e,d,f}\ d\ln \mathcal{D}^y \\
&\quad - \mathbf{T}\odot v^{e^x}\big[\big(-\beta_{f,f}^{e,q}\xi^f + \beta_{f,c}^{e,t}\ \xi^c\big)\mathrm{J} d\ln \mathrm{t}^\varepsilon + \beta_{f,r}^{e,t}\ \mathrm{J} d\ln \mathrm{s}^\varepsilon\big] \\
&\quad + \theta\big(\mathbf{T}(\mathbf{S}\odot\mathbf{J}\odot d\mathrm{t}^b)\mathbb{1} - \mathrm{diag}[\mathbf{T}(\mathbb{1}{+}\mathbf{S}')\odot(\mathbf{J}\odot d\mathrm{t}^b)']\big)
\end{aligned}
$$

Replacing output y

$$\Big[\theta\mathbf{I}-\mathbf{T}\odot v^{y}-(\theta-1)(\mathbf{TS})+\mathbf{T}\odot(v^{e^{f}}\odot\beta^{e,d,f}+v^{e^{x}}\odot\tfrac{1}{\nu^{f}})\Big]d\ln\mathrm{p}=$$
$$(\mathbf{T}\odot v^{y}-\mathbf{I})\Big[\delta^{y,z}d\ln\mathcal{D}^{y}_{i}+\delta^{y,p}d\ln\mathrm{p}_{i}-\delta^{y,qf}d\ln q^{f}-\delta^{y,t\varepsilon}d\ln\mathrm{t}^{\varepsilon}_{i}+\delta^{y,s\varepsilon}d\ln\mathrm{s}^{\varepsilon}_{i}\Big]$$
$$+\mathbf{T}\odot[v^{e^{x}}\frac{1}{\nu^{f}}-v^{e^{f}}\beta^{e,q}_{f,f}+v^{ne}]d\ln q^{f}-\mathbf{T}v^{e^{f}}\beta^{e,d,f}\,d\ln\mathcal{D}^{y}$$
$$-\mathbf{T}\odot v^{e^{f}}[(-\beta^{e,q}_{f,f}\xi^{f}+\beta^{e,t}_{f,c}\,\xi^{c})\mathrm{J}d\mathrm{t}^{\varepsilon}+\beta^{e,t}_{f,r}\,\mathrm{J}d\mathrm{s}^{\varepsilon}]$$
$$+\theta\big(\mathbf{T}(\mathbf{S}\odot\mathbf{J}\odot dt^{b})\mathbb{1}-\mathrm{diag}[\mathbf{T}(\mathbb{1}+\mathbf{S}')\odot(\mathbf{J}\odot dt^{b})']\big)$$

As a result, the general equilibrium effects on the goods markets yield the following change in prices:

$$\Big[(\theta+\delta^{y,p})\mathbf{I}-\mathbf{T}\odot v^{y}(1+\delta^{y,p})-(\theta-1)(\mathbf{TS})+\mathbf{T}\odot(v^{e^{f}}\odot\beta^{e,d,f}+v^{e^{x}}\odot\tfrac{1}{\nu^{f}})\Big]d\ln\mathrm{p}=\Big[(\mathbf{T}\odot v^{y}-\mathbf{I})\,\delta^{y,z}-\mathbf{T}v^{e^{f}}\beta^{e,d,f}\Big]d\ln\mathcal{D}^{y}_{i}$$
$$+\Big[-(\mathbf{T}\odot v^{y}-\mathbf{I})\delta^{y,qf}+\mathbf{T}\odot(v^{e^{x}}\frac{1}{\nu^{f}}-v^{e^{f}}\beta^{e,q}_{f,f}+v^{ne})\Big]d\ln q^{f}$$
$$+\Big[-\mathbf{T}\odot v^{e^{f}}((-\beta^{e,q}_{f,f}\xi^{f}+\beta^{e,t}_{f,c}\,\xi^{c}))-(\mathbf{T}\odot v^{y}-\mathbf{I})\delta^{y,t\varepsilon}\Big]\mathrm{J}d\mathrm{t}^{\varepsilon}$$
$$+\Big[\mathbf{T}\odot v^{e^{f}}\beta^{e,t}_{f,r}+(\mathbf{I}-\mathbf{T}\odot v^{y})\delta^{y,s\varepsilon}\Big]\mathrm{J}ds^{\varepsilon}$$
$$+\theta\big(\mathbf{T}(\mathbf{S}\odot\mathbf{J}\odot dt^{b})\mathbb{1}-\mathrm{diag}[\mathbf{T}(\mathbb{1}+\mathbf{S}')\odot(\mathbf{J}\odot dt^{b})']\big) \tag{28}$$

Again, this compile all the different channels of transmissions that arise in our model. For example, taking the parameter for the carbon tax t^{ε} we see that it changes the price in multiple ways: (i) first, it lowers the oil-gas expenditure for country k, by a factor $\beta^{e,q}_{f,f}\xi^{f}$, (ii) however, it increases the oil-gas bill as a substitution away from coal $\beta^{e,t}_{f,c}\,\xi^{c}$. In addition, (iii) taxing carbon reduces output by a factor $\delta^{y,t\varepsilon}$, which then reduces the revenues in net by a factor $(\mathbf{T}\odot v^{y}-\mathbf{I})$.

B.9 Back to welfare and climate damage

From the budget/consumption expenditure, we saw that welfare is written as:

$$\frac{d\mathcal{U}_{i}}{u'(c_{i})c_{i}}=(d\ln c_{i}+d\ln\mathcal{D}^{u}_{i})$$
$$\frac{d\mathcal{U}_{i}}{u'(c_{i})c_{i}}=\Big[\frac{\eta^{y}_{i}}{\eta^{c}_{i}}-\sum_{\ell}\frac{1}{\nu^{\ell}_{i}}\frac{\eta^{\pi\ell}_{i}}{\eta^{c}_{i}}\Big]\frac{d\mathrm{p}_{i}}{\mathrm{p}_{i}}+\frac{\eta^{y}_{i}}{\eta^{c}_{i}}d\ln\mathcal{D}^{y}_{i}+\sum_{\ell}\Big[\frac{\eta^{\pi\ell}_{i}}{\eta^{c}_{i}}\Big(1+\frac{1}{\nu^{\ell}_{i}}\Big)-\frac{\eta^{y}_{i}}{\eta^{c}_{i}}s^{e}_{i}s^{\ell}_{i}\Big]\frac{dq^{\ell}}{q^{\ell}}-\frac{d\mathbb{P}_{i}}{\mathbb{P}_{i}}+\frac{d\widetilde{\mathrm{t}}^{ls}_{i}}{x_{i}}+d\ln\mathcal{D}^{u}_{i}$$

with the damage

$$d\ln\mathcal{D}^{y}_{i}=-\bar{\gamma}^{y}_{i}(s^{f/E}d\ln E^{f}+s^{c/E}d\ln E^{c})$$

and similarly with $\bar{\gamma}^{u}_{i}$ for $\mathcal{D}^{u}_{i}$

The oil-gas energy price is central for summarizing all the general equilibrium forces for fossil-fuel demand. As a result, since the aggregate supply curve for oil and gas is upward sloping, a higher price implies a higher demand, and hence higher quantity consumed and greenhouse gas emitted.

$$
\begin{aligned}
d\ln q^f &= \frac{\bar{\nu}}{1+\bar{\nu}\overline{\lambda}^{\sigma,f}} \sum_i \lambda_i^f \mathrm{J}_i \, [-\xi^f \gamma^{e,q}_{f,f,i} + \xi^c \gamma^{e,t}_{f,c,i}] d\ln \mathrm{t}^\varepsilon + \frac{\bar{\nu}}{1+\bar{\nu}\overline{\lambda}^{\sigma,f}} \sum_i \lambda_i^f \, \gamma^{e,t}_{f,r,i} \ \mathrm{J}_i d\ln \mathrm{s}_i^\varepsilon \\
&\quad + \frac{\bar{\nu}}{1+\bar{\nu}\overline{\lambda}^{\sigma,f}} \sum_i \lambda_i^f \gamma_i^{e,d,f} \, d\ln \mathcal{D}_i \ + \frac{\bar{\nu}}{1+\bar{\nu}\overline{\lambda}^{\sigma,f}} \sum_i [\lambda_i^f \gamma_i^{e,p,f} + \lambda_i^x \frac{\bar{\nu}}{\nu_i}] d\ln \mathrm{p}_i \\
d\ln E^f &= \frac{1}{\bar{\nu}} \big[d\ln q^f - \sum_i \lambda_i^x \frac{\bar{\nu}}{\nu_i} d\ln \mathrm{p}_i\big]
\end{aligned}
$$

Similarly, for total coal consumption, we can aggregate, with weights $\lambda_i^c = \frac{\mathcal{P}_i e_i^c}{E_i^c}$

$$
\begin{aligned}
d\ln E^c = \sum_i \lambda_i^c d\ln e_i^c &= \sum_i \lambda_i^c \big[- \beta^{e,t}_{c,c,i} \ \xi^c + \beta_i^{e,b,c} (\sigma^e - \frac{\sigma^y}{1-s_i^e}) s_i^f \xi^f \big] \mathrm{J}_i d\ln \mathrm{t}^\varepsilon - \sum_i \lambda_i^c \beta^{e,t}_{c,r,i} \ \mathrm{J}_i d\ln \mathrm{s}_i^\varepsilon + \sum_i \beta_i^{e,p,c} d\ln \mathrm{p}_i \\
&\quad + \sum_i \lambda_i^c \beta_i^{e,b,c} \Big[(\sigma^e - \frac{\sigma^y}{1-s_i^e}) s_i^f d\ln q^f \ + \frac{\sigma^y}{1-s^e} d\ln \mathcal{D}_i \Big]
\end{aligned}
$$

As a result, we can write damage depending on a complicated combination of supply and demand effects:

$$
\begin{aligned}
d\ln \mathcal{D}_i^y &= -\bar{\gamma}_i^y \big(s^{f/E} d\ln E^f + s^{c/E} d\ln E^c\big) \\
&= -\bar{\gamma}_i^y \Big\{ \frac{s^{f/E}}{1+\bar{\nu}\overline{\lambda}^{\sigma,f}} \sum_i \lambda_i^f \, [-\xi^f \gamma^{e,q}_{f,f,i} + \xi^c \gamma^{e,t}_{f,c,i}] \mathrm{J}_i + s^{c/E} \sum_i \lambda_i^c \big[- \beta^{e,t}_{c,c,i} \ \xi^c + \beta_i^{e,b,c} (\sigma^e - \frac{\sigma^y}{1-s_i^e}) s_i^f \xi^f \big] \mathrm{J}_i \Big\} d\ln \mathrm{t}^\varepsilon \\
&\quad - \bar{\gamma}_i^y \sum_i \Big\{ \frac{s^{f/E}}{1+\bar{\nu}\overline{\lambda}^{\sigma,f}} \lambda_i^f \, \gamma^{e,t}_{f,r,i} \ + s^{c/E} \lambda_i^c \beta^{e,t}_{c,r,i} \Big\} \mathrm{J}_i d\ln \mathrm{s}_i^\varepsilon \\
&\quad - \bar{\gamma}_i^y \sum_i \Big\{ \frac{s^{f/E}}{1+\bar{\nu}\overline{\lambda}^{\sigma,f}} \lambda_i^f \gamma_i^{e,d,f} \ + \ s^{c/E} \lambda_i^c \beta_i^{e,b,c} \frac{\sigma^y}{1-s^e} \Big\} d\ln \mathcal{D}_i^y \ + s^{c/E} \sum_i \lambda_i^c \beta_i^{e,b,c} (\sigma^e - \frac{\sigma^y}{1-s_i^e}) s_i^f d\ln q^f \\
&\quad - \bar{\gamma}_i^y \sum_i \Big\{ \frac{s^{f/E}}{1+\bar{\nu}\overline{\lambda}^{\sigma,f}} [\lambda_i^f \gamma_i^{e,p,f} \ - \bar{\nu}\overline{\lambda}^{\sigma,f} \lambda_i^x \frac{\bar{\nu}}{\nu_i}] + s^{c/E} \lambda_i^c \beta_i^{e,p,c} \Big\} d\ln \mathrm{p}_i
\end{aligned}
$$

with $\widetilde{\lambda}_i^f = \lambda_i^f + \lambda_i^c \beta_i^{e,b,c} (\sigma^e - \frac{\sigma^y}{1-s_i^e}) s_i^f \bar{\nu} \frac{s^{c/E}}{s^{f/E}}$, we can rewrite without the price of oil-gas q^f

$$
\begin{aligned}
d\ln \mathcal{D}_i^y &= -\bar{\gamma}_i^y \big(s^{f/E} d\ln E^f + s^{c/E} d\ln E^c\big) \\
\Big[1 + \bar{\gamma}_i^y \sum_i \Big\{ \tfrac{s^{f/E}}{1+\bar{\nu}\overline{\lambda}^{\sigma,f}} \widetilde{\lambda}_i^f \gamma_i^{e,d,f} + s^{c/E} \lambda_i^c \beta_i^{e,b,c} \tfrac{\sigma^y}{1-s^e} \Big\} \Big] d\ln \mathcal{D}_i^y &= \\
& - \bar{\gamma}_i^y \Big\{ \frac{s^{f/E}}{1+\bar{\nu}\overline{\lambda}^{\sigma,f}} \sum_i \widetilde{\lambda}_i^f \, [-\xi^f \gamma^{e,q}_{f,f,i} + \xi^c \gamma^{e,t}_{f,c,i}] \mathrm{J}_i + s^{c/E} \sum_i \lambda_i^c \big[- \beta^{e,t}_{c,c,i} \ \xi^c + \beta_i^{e,b,c} (\sigma^e - \frac{\sigma^y}{1-s_i^e}) s_i^f \xi^f \big] \mathrm{J}_i \Big\} d\ln \mathrm{t}^\varepsilon \\
& - \bar{\gamma}_i^y \sum_i \Big\{ \frac{s^{f/E}}{1+\bar{\nu}\overline{\lambda}^{\sigma,f}} \widetilde{\lambda}_i^f \, \gamma^{e,t}_{f,r,i} \ + s^{c/E} \lambda_i^c \beta^{e,t}_{c,r,i} \Big\} \mathrm{J}_i d\ln \mathrm{s}_i^\varepsilon \\
& - \bar{\gamma}_i^y \sum_i \Big\{ \frac{s^{f/E}}{1+\bar{\nu}\overline{\lambda}^{\sigma,f}} [\lambda_i^f \gamma_i^{e,p,f} \ - \bar{\nu}\overline{\lambda}^{\sigma,f} \lambda_i^x \frac{\bar{\nu}}{\nu_i}] + s^{c/E} \Big(\lambda_i^c \beta_i^{e,p,c} + \big(\sum_k \beta_k^{e,b,c} (\sigma^e - \frac{\sigma^y}{1-s_k^e}) s_k^f \big) \lambda_i^x \frac{\bar{\nu}}{\nu_i} \Big) \Big\} d\ln \mathrm{p}_i
\end{aligned}
$$

Damages change with the aggregate consumption of oil, gas, and coal, which each depends on various general equilibrium effects. The carbon tax and renewable subsidies create substitution effects away from fossil and toward renewable. Moreover, the price level p_i increases the terms of trade for countries that consume oil-gas, and reduces production for exporters, which then affects the equilibrium prices of oil. All these effects are accounted for in this formula, which then enter in the welfare calibration above.

B.10 Further simplification

To simplify the welfare formula even further, in the following we consider that energy is only composed of oil-gas. In practice, oil and gas compose the largest share of energy, with oil representing close to 35% of energy use and natural gas close to 20% at the world level.

We consider that $s_i^f = 1$ and $s_i^r = s_i^c = 0$ in all the formulas above.

This assumption simplify our setting dramatically. The previous welfare decomposition reduces to :

$$\frac{d\mathcal{U}_i}{u'(c_i)c_i} = \frac{dc_i}{c_i} = [\frac{\eta_i^y}{\eta_i^c} - \frac{\eta_i^\pi}{\eta_i^c}\frac{1}{\nu_i}]\frac{d\mathrm{p}_i}{\mathrm{p}_i} + \frac{\eta_i^y}{\eta_i^c}\frac{d\mathcal{D}_i}{\mathcal{D}_i} - \frac{\eta_i^y}{\eta_i^c}s_i^e\frac{dq^f}{q^f} + \frac{\eta_i^\pi}{\eta_i^c}\Big(1+\frac{1}{\nu}\Big)\frac{dq^f}{q^f} - \frac{d\mathbb{P}_i}{\mathbb{P}_i} + \frac{d\widetilde{\mathrm{t}}_i^{ls}}{x_i}$$

where the damage rewrite:

$$d\ln\mathcal{D}_i = -\bar{\gamma}_i d\ln E^f$$

with the average damage is defined as $\bar{\gamma} = \sum_i \bar{\gamma}_i$. And the oil-gas demand curve write:

$$\begin{aligned}
d\ln E^f &= \sum_i \lambda_i^f d\ln e_i^f \\
&= -\sum_i \lambda_i^f \frac{\sigma^y}{1-s_i^e}[d\ln q^f + \mathrm{J}_i d\ln \mathrm{t}^\varepsilon] + \sum_i \lambda_i^f \frac{\sigma^y}{1-s_i^e} d\ln\mathcal{D}_i + \sum_i \lambda_i^f \frac{\sigma^y}{1-s_i^e} d\ln\mathrm{p}_i \\
&= -\sum_i \widetilde{\lambda}_i^f [d\ln q^f + \mathrm{J}_i d\ln \mathrm{t}^\varepsilon] + \sum_i \widetilde{\lambda}_i^f d\ln\mathcal{D}_i + \sum_i \widetilde{\lambda}_i^f d\ln\mathrm{p}_i
\end{aligned}$$

where, to simplify notations, we denote $\widetilde{\lambda}_i^f = \lambda_i^f \frac{\sigma^y}{1-s_i^e}$, and it's average $\overline{\lambda}^{\sigma,f} = \sum_i \widetilde{\lambda}_i^f \frac{\sigma^y}{1-s_i^e}$. As a result, the demand now rewrites:

$$d\ln E^f = \frac{1}{1+\bar{\gamma}+\mathbb{C}\mathrm{ov}_i(\widetilde{\lambda}_i^f, \bar{\gamma}_i)}\Big[-\sum_i \widetilde{\lambda}_i^f [d\ln q^f + \mathrm{J}_i d\ln \mathrm{t}^\varepsilon] + \sum_i \widetilde{\lambda}_i^f d\ln\mathrm{p}_i\Big]$$

We can see that the energy demand curve is affected by climate change: more emission imply larger damage, which in turn reduce energy demand and hence emissions. Moreover, the covariance term indicates that if the large energy producers (with a larger share of the market, and high elasticity σ) are also the most affected by climate change, this effect is stronger and the demand curve is even steeper / more inelastic.

The Optimal Design of Climate Agreements

Inequality, Trade, and Incentives for Climate Policy

Most recent version accessible here

Thomas Bourany*

Job Market Paper

The University of Chicago

March 20, 2025

Abstract

Fighting climate change requires ambitious global policies, which are undermined by free-riding incentives. The heterogeneity in both the impacts of climate change and the costs of carbon taxation exacerbate non-cooperation, which makes the implementation of multilateral climate agreements difficult. This paper studies how to design an optimal climate club – in the spirit of Nordhaus (2015) – to maximize global welfare, incorporating strategic behavior when countries can exit climate agreements. In an Integrated Assessment Model with heterogeneous countries and international trade, I study the choice of countries in the agreement, the optimal level of carbon tax that members set on fossil fuels, and the tariffs they impose on non-members to incentivize participation. The decision balances an intensive margin – a club with few countries and large individual emission reductions – and an extensive margin – accommodating more countries at the cost of lowering the carbon tax. I find that the optimal climate club consists of all countries except several fossil producers – Russia, Saudi Arabia, Nigeria, and Iran – a \$110 tax per ton of CO_2 within the club, and a 50% tariff on goods from non-members. In contrast, the globally optimal carbon tax is \$130 when free-riding is absent. In several extensions, I study additional policy instruments, such as transfers (as in the COP29's NCQG on climate finance), carbon tariffs (e.g. UE's CBAM), or fossil-fuel-specific tariffs, and examine the effects of trade retaliation for the stability of climate agreements.

*Thomas Bourany, *thomasbourany@uchicago.edu*. I thank my advisors, Mikhail Golosov, Esteban Rossi-Hansberg, Lars Hansen, and Michael Greenstone, for their valuable guidance and advice. I also thank Rodrigo Adao, Jonathan Dingel, Jordan Rosenthal-Kay, Tom Hierons, Aditya Bhandari, Sasha Petrov, Elena Aguilar, and other seminar participants at UChicago, Booth, and the Chicago Fed for stimulating discussions. All errors are mine.

1 Introduction

Fighting climate change requires ambitious global policies. To avoid severe consequences of global warming, carbon emissions must reach net zero in the next decades, and our economies need to phase out fossil fuels in a concerted effort to keep the world temperature under $2°C$ (IPCC et al. (2022)). However, the world is currently facing climate inaction. One of the main reasons behind this lack of cooperation is the presence of the free-riding problem: the benefits of climate policies are global, while the costs of reducing emissions using carbon pricing are local. Individual countries have incentives to free-ride on the rest of the world's reduction in emissions without implementing costly carbon abatement themselves.

Taxation of carbon and fossil fuels has strong redistributive effects across countries, determining their willingness to implement climate policy. First, emerging economies may face challenges in reducing the fossil fuel consumption necessary to continue their economic development. Second, carbon taxation has a substantial impact on energy markets, affecting the surplus of fossil fuel exporters and importers. Finally, imposing a carbon tax in one country reallocates economic activity and carbon emissions toward other countries through international trade – or "carbon leakage". All these effects reinforce free-riding incentives and climate inaction.

Multilateral climate agreements have been the traditional answer to address climate inaction, with the United Nations Conference of the Parties (COP) as an example. Unfortunately, they have failed to achieve decisive binding policy agreements. More recently, trade instruments have been the focus of policy discussions as trade policy offers the potential to give incentives to other countries to reduce emissions. In particular, Nordhaus (2015) proposes the idea of a "climate club", a voluntary agreement where members implement common carbon taxation as well as retaliatory tariffs on countries that do not participate in the club. In this context, trade sanctions are necessary to foster participation in the club and reduce free-riding incentives.[1]

What should be the design of a climate agreement that accounts for free-riding incentives as well as redistributive effects? What is the optimal climate club, its composition and level of carbon tax? This paper addresses these questions by examining the conditions necessary to construct a universal climate agreement with globally optimal carbon tax and tariffs. I explore which factors incentivize countries to join such an agreement, and I investigate how carbon and trade policy needs to be implemented to promote participation, maximize welfare, and fight climate change.

I tackle these policy questions in a climate-economy framework – or Integrated Assessment Model (IAM) – augmented with heterogeneous countries and international trade. I build a multi-country IAM, extended with bilateral trade frictions and energy markets. Individual countries differ in their vulnerability to climate change, income levels, their energy mix in oil, gas, coal, and non-carbon energy, their costs of producing fossil fuels as well as trade costs in trade in goods.

[1] Another notable example is the European Union's Carbon Border Adjustment Mechanism (CBAM), which is proposed to address the carbon leakage problem. This policy is a "carbon tariff" – i.e. a tariff whose rate increases with the carbon content of imported goods. This also has the potential to generate incentives for trade partners to implement climate policy in order to lower the carbon footprint of their exports.

This framework allows me to account for the multifaceted redistribution and leakage effects that arise in general equilibrium as a result of climate change and climate policy. The model serves as a laboratory for evaluating the welfare effects of different agreement designs.

With endogenous participation, countries have differing incentives to join a climate agreement. As a result, the decisions on the optimal levels of carbon tax and trade tariffs, as well as the choice of participants in the club, should be made *jointly*. Indeed, the optimal design reveals a tradeoff between an intensive and an extensive margin. At the intensive margin, an agreement could gather a small set of countries that can individually implement large emissions reductions with high carbon taxes. However, this is not sufficient to reduce global emissions and combat climate change effectively. In contrast, building a more extensive climate club requires accommodating the participation of a larger number of countries, which can only be done at the cost of lowering the carbon tax.

In this context, I address the policy problem where a global social planner maximizes the world's welfare by designing a climate agreement, or "climate club", that consists of three elements: (1) a set of countries included in the agreement – also called "climate coalition" – that are subject to the climate and trade policies, (2) the level of the carbon tax that club members set on their oil, gas, and coal energy consumption and (3) the level of the uniform trade tariffs that members impose on imported goods from non-member countries, while club members benefit from free-trade among themselves.[2] This policy design follows closely Nordhaus (2015)'s climate club setting.

Countries make individual choices to join or leave the agreement, and such strategic participation needs to be accounted for in the design of the agreement. I consider Nash equilibria, where countries make participation decisions either unilaterally or in "sub-coalition deviations", i.e. when a subset of countries decide together to deviate and leave the agreement. The club design thus mirrors an optimal taxation and trade policy problem with limited instruments, together with a choice of countries. This problem is particularly challenging as one needs to simultaneously identify the countries willing to participate as well as the optimal policies. I propose an approach to separate this joint problem by fixing the policies and solving for the optimal climate coalition. This coalition choice resembles the type of combinatorial discrete choice problem (CDCP) that arises in trade economics, e.g. Arkolakis et al. (2023). I propose different numerical solution methods to tackle this problem in the presence of participation constraints. In that context, I consider a restricted set of instruments: a single carbon tax and a single uniform good tariff. This follows the idea that the Social Planner represents the bargaining outcome between club members. With bargaining frictions and transaction costs, negotiating over a small number of instruments – in particular, a single carbon tax – facilitates the finding of an agreement and undermines free-riding.[3]

[2]In the main club design, fossil fuels are still freely traded for all countries. Moreover, non-members are *passive* in the sense that they do not retaliate with additional trade tariffs against the clubs. These two assumptions are relaxed in the extensions.

[3]Weitzman (2015) argues that a single carbon price serves as a "focal point" and is superior to binding quantity targets. Indeed, with transaction and bargaining costs, the Coase theorem may fail, preventing the agreement from reaching international cooperation. He attributes the failure of the Kyoto Agreement to the fact that quantity quotas represent a subdivision of efforts over countries and are, hence, more subject to disagreements and free-riding.

I contrast this framework with policy benchmarks absent endogenous participation. First, I consider the optimal carbon policy when the coalition gathers the entire world without participation constraints. I show that the choice of the carbon tax depends crucially on the availability of redistribution instruments, such as lump-sum transfers, in the First-Best allocation. Without such transfer instruments, I show how the choice of the Second-Best carbon tax accounts for distributional motives. Indeed, the carbon tax accounts for income inequality and its effect on demand distortion, supply redistribution through fossil-fuel energy markets, as well as trade leakage. As a result, the optimal carbon tax is \$130 per ton of CO_2 in the Second-Best and is lower than the Social Cost of Carbon, i.e. the marginal cost of climate change, a result that contrasts with the conventional Pigouvian principle.[4]

Second, I also compare the "climate club" framework to the non-cooperative Nash equilibrium, in which each individual countries choose its "unilaterally optimal" carbon tax and trade tariffs. The unilateral carbon taxation policy can become a subsidy if the Local Cost of Carbon – the cost of climate change as internalized by an individual country – is lower than two terms-of-trade manipulation terms: one for the good market and one for the fossil-fuel energy market. Similarly, optimal tariffs are also used for terms-of-trade manipulation of goods, a logic that aligns with conventional results in Ossa (2014) or Farrokhi and Lashkaripour (2024).

In comparison, climate agreements provide an "issue linkage" (Maggi (2016)) by linking the implementation of carbon policy with a reduction in tariffs, as the club promotes free trade among coalition members. The countries' participation choice depends on a balance between two effects: the distortionary cost of carbon taxation against the cost of tariffs – related to gains from trade. To choose whether to exit the club, individual countries consider if the first outweighs the second.

I find that the optimal climate club consists of all countries with the exception of Russia, former Soviet countries, Saudi Arabia, Nigeria, and Iran. The agreement imposes a moderate carbon tax of \$110 per ton of CO_2 and a 50% tariff on traded goods of non-participants. The optimal climate agreement cannot achieve the world's optimal policy with complete participation – an agreement with a \$130 carbon tax and all the countries – despite full discretion on the choice of carbon tax and tariffs.

The reason is threefold. First, to increase participation, it is beneficial to reduce the carbon tax. Several Middle Eastern countries and developing economies in South Asia and Africa would not join an agreement with a high carbon tax, regardless of the level of the tariffs, since the gains from trade are bounded. Therefore, it is optimal to lower the tax from \$130 to $\$110/tCO_2$ to include those countries and share the "burden" of carbon abatement across more countries.

Second, it is beneficial to leave several fossil fuel producers like Russia, Iran, Saudi Arabia, and Nigeria outside of the climate agreement. Indeed, they suffer large welfare costs from carbon taxation, being relatively closed and exporters of oil and gas. They would never join an agreement

[4]The optimal policy problem with limited instruments is treated extensively in Bourany (2025) in a large class of climate-macroeconomic models. In the present paper, I draw a particular emphasis on international trade and leakage effects, a novel channel that needs to be accounted for in optimal carbon taxation.

unless the carbon tax was very small, which is not optimal from a global perspective.

Third, trade policy is a key strategic instrument to deter free-riding and incentivize countries to join the agreement. All the countries for which the cost of large tariffs outweighs the distortionary cost of carbon taxation are willing to participate in such climate clubs. That is especially the case for countries in Europe, East Asia, including China, and South-East Asia, which trade internationally a large share of goods production and have large gains from trade. Absent tariff retaliation, free-riding prevails over the cost of climate actions, as discussed in Nordhaus (2015). However, if moderate tariffs spur participation for a low carbon tax, this incentive effect vanishes quickly as the carbon tax increases and larger emissions reductions are required. The gains from trade are bounded – and small for some countries like the Middle East and Russia – and therefore, there is a limit to what carbon policy can achieve.

In extensions, I consider additional policy instruments: (i) transfers with a fund aimed at supporting mitigation and adaptation of developing countries, (ii) Carbon Border Adjustment Mechanisms replacing the club's uniform tariffs, and (iii) fossil-fuel-specific tariffs. First, transfers have been at the center of policy discussions at the UN's COP28 in Dubai and COP29 in Baku. I consider an experiment where the climate agreement redistributes part of revenues from the carbon tax to poorer economies like South Asia and Africa – which consume less fossil fuel per capita. It improves the welfare of these countries much more than the loss incurred by the richer economies of North America, Europe, and East Asia. However, the total scale of such policy is limited due to free-riding and redistributive concerns. The maximum transfers that the global club designer can collect is around \$350 millions. Interestingly, this amount aligns well with the amount of climate finance that was agreed in the COP29 in Baku.[5] I show that the free-riding problem limits the amount that be agreed upon in such climate agreements.

Second, I analyze the implementation of Carbon-Border-Adjustment Mechanisms (CBAM) in the context of these climate agreements. CBAM – and more generally "carbon tariffs" – have been discussed in the European Union as a tool to fight against carbon leakage and avoid the negative consequences of carbon pricing or fossil taxation in terms of trade competitiveness. I consider an experiment in the design of climate agreements where the carbon tariff impose an additional cost on the carbon content of the traded-goods, and would thus act as a substitute punishment for uniform tariffs. I show that, to replace the 50% tariff, the agreement need to impose a carbon price of \$1250 on imported goods. This price is much larger than the carbon tax ($\sim$ \$100) since the carbon content of production is usually very low ($\sim 0.3\,\text{kg}\,CO_2/\$$) and the punishment need to be strong enough to incentive participation of non-members to the club. However, if one would impose that the internal price of carbon should equal the external price used for carbon tariffs – for example by following the World Trade Organization rules – then it imposes an additional constraint that limit considerably the strategic power of this climate agreement.

[5] There, "developed nations have agreed to help channel "at least" \$300bn a year into developing countries by 2035 to support their efforts to deal with climate change." This amount fell short of the \$1.3tr/year initially proposed by developing countries, which raised objections from the delegations of India and Nigeria regarding the final text.

Third, I study the impacts of fossil-fuel-specific tariffs. Imposing tariffs directly on oil-gas exports has a strong effect on the energy rents of fossil-fuel-exporting countries. This targets directly the large fossil-fuel producers, like Middle-Eastern Gulf countries and Russia, which are typically the first ones to free-ride from the club. Such tariffs increases the penalty costs and hence the retaliatory power of the climate club and thus enforces the participation of these countries. As a result, this allows the club to increase the attainable carbon tax and global carbon abatement. With such instruments, we can reach the globally optimal allocation with complete participation and a more ambitious carbon policy.

Lastly, I compare how these optimal agreements results change depending on the strategic response of countries outside the club. In the baseline result, non-members are passive and do not impose tariffs on the climate club members. If, instead, we consider that non-members cooperate strategically to impose retaliatory tariffs on club members, this has important consequences on the stability of the club. First, it makes the intensive-extensive margin tradeoff more salient, as countries have even more incentive to deviate from the club for high carbon taxes and moderate tariffs. However, retaliation in the face of high tariffs from club members also raises the costs of a trade war for *both* club members and non-members. This implies that if the club is large enough, it can threaten to escalate a trade war to enforce the participation of the non-members and ultimately achieve the optimal policy with a very high carbon tax and complete participation.

Literature

This work relates to a large literature on the economics of climate change and bridges a gap with both the international trade policy and the game theoretical literature. First, I contribute to the debate on the formation of Climate Clubs, following the pioneering contribution of Nordhaus (2015). The implementation of climate policy suffers from a free-riding problem and Nordhaus proposed a simple framework to evaluate the principle of issue linkage, i.e. linking the enforcement of a climate policy with trade tariffs. He shows with the C-DICE model that for different – exogenously set – carbon prices and tariffs rates, we can achieve varying participation to a climate club. With a low carbon price – up to $25\$/tCO_2$ – and high tariffs – above 10%, the climate club can achieve a club with all the 15 regions he considers.

I depart from Nordhaus' Climate Club framework in three directions. First, I show that when a social planner chooses endogenously and optimally both the carbon tax, the tariffs, and the club members, we observe an intensive margin - extensive margin tradeoff. A lower carbon tax and higher tariffs increase participation. Second, I depart from the C-DICE model that uses ad-hoc functions for the carbon abatement – inspired by the DICE model – and the gain from trade and costs of tariffs – a quadratic approximation of the results of Ossa (2014). I show that modeling the energy market – both with heterogeneity in demand and supply of fossil energy – and trade in goods – accounting for leakage effects and terms-of-trade manipulation – highlights the tradeoff between the cost of carbon taxation and the cost of tariffs. In particular, in this micro-funded setting, gains from trade are bounded, which makes some countries unwilling to join an agreement if the loss from phasing out fossil fuels is too large, and this *for any level of tariffs*. Third, I model the cost of climate on production as endogenous to policy, which makes the optimal carbon tax account for redistributive effects through income inequality, trade leakage, and energy markets. Iverson (2024) take the same C-DICE model as Nordhaus and analyze how a "Tiered Climate Club" can achieve higher carbon abatement, where two different "tiers" imply different levels of carbon tax and tariff retaliations for different sets of countries. Hagen and Schneider (2021) also analyze how retaliation can undermine the stability of the club and results in potentially suboptimal climate clubs where the gains from climate policy are undermined by the costs of a trade war. In comparison to these three models, I provide a quantitative analysis where optimal policy is chosen strategically to enforce participation.

Farrokhi and Lashkaripour (2024) also study how climate policy can be conducted with trade instruments. They solve for the optimal trade policy in a rich multi-industry trade model, inspired by Copeland and Taylor (2004), and show that unilateral policy accounts for carbon leakage when setting tariffs. In this setting, they explore the sequential construction of a climate club, where European Union starts a coalition, implements the unilaterally optimal trade-climate policy, and iteratively grows the participation to the club. In contrast, I show how the club should *design* the trade-climate policy *strategically* to spur participation. My framework also incorporates several redistribution channels absent from their framework. Non-linear damage makes the cost of climate change endogenous to policy, and inequality across countries creates differences between policies maximizing output, reducing emissions, and improving welfare.

This project lies at the intersection of three bodies of literature: one on trade policy, one on the game-theoretical aspects of climate policy cooperation, and one on macroeconomic models of climate change.

First, the interdependence between climate, environmental, and trade policies is explored extensively in Kortum and Weisbach (2021), Barrett (2001), Bohringer et al. (2016), Bohringer et al. (2012) or Hsiao (2022). These articles explore the differences between unilateral policies implemented at the country level and the potential for climate cooperation using trade policies. Other articles in this trade literature explore the underpinnings of optimal trade policies, e.g. Costinot et al. (2015), Ossa (2014), Adao et al. (2023), Antras et al. (2024). More specifically they study the choice of trade tariffs with different objectives, like terms of trade manipulation or supply-chains considerations. I show how these policy instruments can be used for issue linkage and climate policy and study the optimal design of climate agreements in the presence of free-riding.

Moreover, I also borrow from the theoretical literature on climate cooperation, with classical references such as Barrett (1994, 2003, 2013), Carraro and Siniscalco (1993), Chander and Tulkens (1995, 1997), Harstad (2012), Dutta and Radner (2004, 2009), Nordhaus (2015) or the older literature collected and summarized in Batabyal (2000), Chander (2018) and Maggi (2016). There is also a large literature on dynamic games and coalition formation games, that focus on the building of agreements, either through coordination games or through bargaining procedures, and summarized in Ray and Vohra (2015), or Okada (2023) more recently. Similarly, Nordhaus (2021), Harstad (2023), or Maggi and Staiger (2022) study those questions as well as other dynamic features, such as technical change, the path of climate dynamics or intertemporal decision-making.

I draw inspiration from many of these references and provide a quantitative framework with many dimensions of heterogeneity to reveal the different factors that drive the countries' decisions to participate in climate clubs and provide policy recommendations.

Third, I also draw from the macroeconomic literature on the implications of climate change and carbon policy. Indeed, building on my own work Bourany (2025), I show that the optimal carbon policy accounts for several general equilibrium channels, as well as macroeconomic dynamics – in the spirit of Integrated Assessment Models. Starting from a static version of the DICE/RICE models as in Barrage and Nordhaus (2024) and Nordhaus and Yang (1996), I study the optimal fossil fuel taxation, as in Golosov et al. (2014) with heterogeneous countries/regions, as in Krusell and Smith (2022), Cruz and Rossi-Hansberg (2024, 2022), Kotlikoff, Kubler, Polbin, Sachs and Scheidegger (2021), Kotlikoff, Kubler, Polbin and Scheidegger (2021), as well as international trade. Similarly, I study a quantitative model with many dimensions of heterogeneity but keep it static to be able to study strategic interactions between large countries and a climate agreement design, which includes a combinatorial discrete choice joint with an optimal policy choice.

The remainder of this paper is structured as follows. In Section 2, I lay out the Integrated Assessment Model that we study in the policy analysis. The design of the climate agreement is exposed in Section 3. In Section 4, I present how I match the model to the data. In Section 5, I discuss the optimal policy benchmarks with or without cooperation in a setting without free-riding incentives. In Section 6, I present the main result of our analysis on the optimal climate agreement. In Section 7, I develop extensions that relax some of the assumptions made in the baseline results, such as the availability of additional instruments or retaliation.

2 An integrated assessment model with heterogeneous regions and trade

I build an integrated assessment model (IAM) that incorporates various dimensions of heterogeneity influencing individual countries' incentives to join climate agreements. This framework is the simplest model that includes both climate externality, a non-trivial energy market for fossil energy, and a realistic trade structure that reproduces the leakage effects of taxation.

I study a static economy[6] with I countries indexed by $i \in \mathbb{I}$, each with population $\mathcal{P}_i$. All the economic variables are expressed per capita.[7] Each country is composed of five representative agents: (i) a household that consumes the final goods, (ii) a final-good firm producing goods using labor and energy, (iii) a fossil energy firm extracting oil and gas, (iv) a producer of coal energy, and (v) a producer of renewable/non-carbon energy. Moreover, each country has a government that sets taxes and tariffs.

[6] More particularly, the static is a stationary representation of a dynamic model, as I describe in Appendix B.3. This allows keeping the framework simple enough to study the strategic interaction between countries as well as the joint design between a combinatorial discrete choice and optimal policy choice for the carbon tax and tariffs.

[7] For example, y_i or e_i^f are final output and fossil energy use respectively, and $\mathcal{P}_i y_i$ and $\mathcal{P}_i e_i^f$ represent the total quantities produced/consumed in the country. I allow for population growth n and TFP growth $\bar{g}$ in the dynamic model, and we display here the stationary version of the Balance Growth Path.

2.1 Household problem

The representative household in country i imports from all countries $j \in \mathbb{I}$ and consumes the aggregate quantity c_i. I consider an Armington structure, c.f. Anderson (1979), Arkolakis, Costinot and Rodriguez-Clare (2012), where each country produces its own variety. The household preferences have constant elasticity of substitution θ over goods from different countries.

$$\mathcal{U}_i = \max_{\{c_{ij}\}} u(\{c_{ij}\}_j) = u\Big(\mathcal{D}^u_i(\mathcal{E})c_i\Big) \ , \qquad\qquad c_i = \Big(\sum_{j\in\mathbb{I}} a_{ij}^{\frac{1}{\theta}} c_{ij}^{\frac{\theta-1}{\theta}}\Big)^{\frac{\theta}{\theta-1}} \ , \tag{1}$$

where a_{ij} are the preference shifters for country i on the good purchased from country j, which also include the home-bias a_{ii}.[8] The climate externality affects consumption which a factor $\mathcal{D}^u_i(\mathcal{E})$ which summarizes climate damages, given world emissions $\mathcal{E}$. It is a reduced-form representation of the climate system – and the path of temperatures – and decreases in $\mathcal{E}$ and is country-specific due to differences in the vulnerability and costs of climate change. In the quantification Section 4, we detail how we calibrate this function for each country i. Households earn labor income, energy rent, and transfers, and their budget constraints is given by:

$$\sum_{j\in\mathbb{I}} c_{ij}\,(1+\mathrm{t}^b_{ij})\tau_{ij}\mathrm{p}_j = w_i\ell_i + \pi^f_i + \ \mathrm{t}^{ls}_i \ , \tag{2}$$

where w_i is the wage rate, ℓ_i the exogenous labor supply is normalized to 1, π^f_i the profit earned from the ownership of the energy firms, and t^{ls}_i the lump-sum transfer received from the government. On the expenditure side, the household in i imports quantities c_{ij} from j, purchased at price p_j, and subject to iceberg cost τ_{ij} and to trade-tariffs $1+\mathrm{t}^b_{ij}$. The choice of trade policy will be made explicit below.

The optimal consumption choice of the household yields the following quantities and Armington trade shares given by:

$$\begin{aligned} c_{ij} &= a_{ij}c_i\Big(\frac{(1+\mathrm{t}^b_{ij})\tau_{ij}\mathrm{p}_j}{\mathbb{P}_i}\Big)^{-\theta} \ , \\ s_{ij} &\equiv \frac{c_{ij}p_{ij}}{c_i\mathbb{P}_i} = a_{ij}\frac{((1+\mathrm{t}_{ij})\tau_{ij}\mathrm{p}_j)^{1-\theta}}{\sum_k a_{ik}((1+\mathrm{t}_{ik})\tau_{ik}\mathrm{p}_k)^{1-\theta}} \ , \end{aligned} \tag{3}$$

where $p_{ij} = (1+\mathrm{t}^b_{ij})\tau_{ij}\mathrm{p}_j$ is the effective price for a variety from country j sold in country i, and $\mathbb{P}_i$ is the price index of country i:

$$\mathbb{P}_i = \Big(\sum_{k\in\mathbb{I}} a_{ik}((1+\mathrm{t}_{ik})\tau_{ik}\mathrm{p}_k)^{1-\theta}\Big)^{\frac{1}{1-\theta}} \ .$$

As a result, we summarize the budget constraint as $c_i\mathbb{P}_i = \sum_{j\in\mathbb{I}} c_{ij}\,(1+\mathrm{t}^b_{ij})\tau_{ij}\mathrm{p}_j$, and the per-capita welfare of country i is then summarized by the indirect utility as the utility of income discounted

[8] We assume that preferences $\{a_{ij}\}$ and iceberg trade costs $\{\tau_{ij}\}$ are policy-invariant, in particular, they are not sensitive to price changes and tariffs.

by the price level and climate damages, namely:

$$\mathcal{U}_i = u\Big(\mathcal{D}^u_i(\mathcal{E})c_i\Big) = u\Big(\mathcal{D}^u_i(\mathcal{E})\,\frac{w_i\ell_i+\pi^f_i+\mathrm{t}^{ls}_i}{\mathbb{P}_i}\Big)\ . \tag{4}$$

2.2 Final good firm problem

The representative final good producer in country i is producing the domestic variety at price p_i. The firm's profit maximization is:

$$\max_{\ell_i,e^f_i,e^c_i,e^r_i} \quad \mathrm{p}_i\,\mathcal{D}^y_i(\mathcal{E})\,z_i\,F(\ell_i,e^f_i,e^c_i,e^r_i) - w_i\ell_i - (q^f+\xi^f\mathrm{t}^\varepsilon_i)e^f_i - (q^c_i+\xi^c\mathrm{t}^\varepsilon_i)e^c_i - q^r_i e^r_i$$

where the production function $\bar{y}_i = F(\ell_i, e^f_i, e^c_i, e^r_i)$ is constant returns to scale and concave in all inputs. It uses labor, ℓ_i, at wage w_i, fossil energy, e^f_i, purchased at price, q^f, coal, e^c, at price, q^c_i, and renewable energy, e^r_i, at price, q^r_i. Energy from oil-gas, e^f_i, and coal, e^c_i, differ from renewable in the sense that they emit greenhouse gases, with respective carbon concentration ξ^f and ξ^c, as we will see in Section 2.4. As a result, there is a motive for taxing oil, gas, and coal energy with the carbon tax t^ε_i. We discuss the choice of this tax in the next sections.

The productivity of the domestic good firm, $y_i = \mathcal{D}^y_i(\mathcal{E})\,z_i\,\bar{y}_i$, can be decomposed in two terms. First, the TFP, z_i, represents productivity as well as institutional/efficiency differences between countries. Invariant to prices and policy, this technology wedge accounts for income inequality across countries. These differences in TFP translate into differences in consumption that create redistribution motives for tax policy.

The second difference in productivity comes from the climate externality summarized by the net-of-damage function $\mathcal{D}^y_i(\mathcal{E})$, given world emissions $\mathcal{E}$. This function is also a reduced-form representation of the climate system from future temperatures, decreases in $\mathcal{E}$, and is country-specific due to differences in costs of climate change, as we detail how we detail in the quantification Section 4.

The firm input decisions solve the optimality conditions, where we define the marginal product of an input x as $MPx_i \equiv \mathcal{D}^y_i(\mathcal{E})\,z_iF_x(\ell_i,e^f_i,e^c_i,e^r_i)$ for $x \in \{\ell_i, e^f_i, e^c_i, e^r_i\}$. For example, in the case of oil and gas e^f_i, the first-order condition can be written as:

$$\mathrm{p}_i\mathcal{D}^y_i(\mathcal{E})\,z_i\,F_{e^f}(\ell_i,e^f_i,e^c_i,e^r_i) =: \mathrm{p}_i MPe^f_i = q^f + \xi^f\mathrm{t}^\varepsilon_i\ , \tag{5}$$

and similarly for other inputs ℓ_i, e^c_i, e^r_i. Crucially, the private decision of firms do not internalize climate externalities of their own fossil-fuel energy use and only responds to carbon tax t^ε_i.

2.3 Energy markets

The final-good firm is consuming three kinds of energy sources – oil-gas, coal, or renewable (non-carbon) – which are supplied by three representative energy firms in each country. Oil-gas sources are traded internationally, and countries can be exporters or importers. Coal and renewable sources are both traded locally, an empirically relevant assumption given the substantial trade costs in coal shipping or electricity transfers.

2.3.1 Fossil firm

In each country $i \in \mathbb{I}$, a competitive energy producer extracts fossil fuels – oil and gas – e_i^x and sells it to the international market at price q^f. The energy is extracted at convex cost $\mathcal{C}_i^f(e_i^x)$, where the convex costs are paid in the unit of the consumption bundle of the household.[9] The energy firm's profit maximization problem is given by:

$$\mathcal{P}_i \pi_i^f = \max_{e_i^x} q^f e_i^x - \mathcal{C}_i^f(e_i^x)\mathbb{P}_i \ , \tag{6}$$

where $\mathcal{P}_i \pi_i^f$ is the total energy rent of country i. Since the extraction costs are convex, the production function has decreasing return to scales,[10] and hence, even with competitive firms, taking the fossil price as given, a positive energy rent exists. Moreover, for the sake of simplicity, we do not consider that energy firms have market power in the setting of energy prices – for example, in the case of OPEC – even though this framework could easily allow for such an extension. Any sources of misallocation – in the sense of Hsieh and Klenow (2009) – are accounted for in the calibration of the cost function $\mathcal{C}_i^f(\cdot)$ as we will see in the quantification Section 4.

Naturally, the optimal extraction decision follows from the optimality condition:

$$q^f = \mathcal{C}_i^{f\,\prime}(e_i^x)\mathbb{P}_i \ , \tag{7}$$

which yields the implicit function $e_i^{x\star} = e^x(q^f/\mathbb{P}_i) = \mathcal{C}_i^{f\prime\,-1}(q^f/\mathbb{P}_i)$. Finally, the energy rent comes from fossil firms' profits $\pi^f(q^f, \mathbb{P}_i) = q^f e^x(q^f/\mathbb{P}_i) - \mathcal{C}_i^f(e^x(q^f/\mathbb{P}_i))\mathbb{P}_i > 0$ and depends on the marginal costs as well as the elasticity $\nu_i = \frac{\mathcal{C}_i^{f\,\prime\prime}(e^x)}{\mathcal{C}_i^{f\,\prime}(e^x)e^x}$.

As we will see below, the profit $\pi^f(q^f, \mathbb{P}_i)$ and its share in income $\eta_i^{\pi f} = \frac{\pi_i{}^f}{y_i \mathrm{p}_i + \pi_i{}^f}$ are key to determine the exposure of a country to carbon taxation. Indeed, reducing carbon emissions by phasing out of fossil fuels reduces energy demand and its price q^f and hence affects energy profit π_i^f and the welfare of large oil and gas exporters.

2.3.2 International fossil energy markets

I assume that oil and gas are traded frictionlessly in international markets.[11] The market clears such that

$$E^f = \sum_{i \in \mathbb{I}} \mathcal{P}_i e_i^f = \sum_{i \in \mathbb{I}} e_i^x \ . \tag{8}$$

Countries have different exposure to this fossil energy market. As country i consumes fossil fuels in total quantity $\mathcal{P}_i e_i^f$, and produces total quantity e_i^x, its net exports of oil and gas are $e_i^x - \mathcal{P}_i e_i^f \lessgtr 0$.

[9] This allows to account for international inputs in goods and services for building capital for resource extraction.

[10] We can also define a fossil production function with inputs x_i^f such that $e^x = g(x_i^f)$ and profit $\pi = q^f g(x) - x\mathbb{P}_i$ instead of $\pi = q^f e^x - \mathcal{C}(e^x)\mathbb{P}_i$, in which case $g(x) = \mathcal{C}^{-1}(x)$

[11] For the sake of simplicity, I refrain from considering a general Armington structure, combining different fossil varieties with demand $e_i^f = \left(\sum_j (e_{ij}^f)^{\frac{\Theta-1}{\Theta}}\right)^{\frac{\Theta}{\Theta-1}}$. I make the simplifying assumption that fossil fuels produced in different countries are not distinguishable – crude oil or natural gas from Nigeria, Saudi Arabia, or Russia are not differentiated varieties – corresponding to the limiting case $\Theta \to \infty$

2.3.3 Coal firm

A representative firm produced coal, that is consumed by the final good firm. I differentiate coal from other fossil fuels like oil and gas because coal production typically does not generate large energy rents for producing countries as a share of GDP. Moreover, large coal producers also consume a large fraction of that coal locally, as trade costs for coal transportation are larger. Hence, I make this empirically grounded assumption that coal is not traded.

The production $\bar{e}_i^c$ is constant returns to scale and uses final good inputs. I assume the production function is of the form $\bar{e}_i^c = x_i^c/\mathcal{C}_i^c$, where x_i^c is a CES aggregator of exactly the same form as eq. (1). As a result, the inputs $x_i^c = \mathcal{C}_i^c \bar{e}_i^c$ is paid in the consumption bundle at price $\mathbb{P}_i$ and the profit maximization problem is:

$$\pi_i^c = \max_{e_i^c} q_i^c \bar{e}_i^c - \mathcal{C}_i^c \bar{e}_i^c \mathbb{P}_i \ ,$$

where the marginal cost $\mathcal{C}_i^c$ is a constant. This implies that there is no coal profit[12] in equilibrium, i.e. $\pi_i^c = 0$. The price for coal and the market clearing condition are given by:

$$q_i^c = \mathcal{C}_i^c \mathbb{P}_i \ , \qquad \bar{e}_i^c = e_i^c \ . \tag{9}$$

Hence, for a given price index of inputs $\mathbb{P}_i$, this implies a perfectly elastic supply curve for coal energy, something we observe in practice as coal production is easily scalable in response to oil and gas price fluctuations.

2.3.4 Renewable, non-carbon, firm

The final good firm also uses renewable and other low-carbon energy sources, such as solar, wind or nuclear electricity. This provides a way of substituting away from fossil fuel in the production function $F(\cdot)$.

A representative firm produces renewable or non-carbon energy, and this supply, $\bar{e}_i^r$, is not traded. This assumption is verified by the fact that electricity is rarely traded across countries – and when it is, it only is only the result of temporary differences in electricity production due to intermittency, rather than large structural imbalances. The production $\bar{e}_i^r$ is also constant returns to scale, with production $\bar{e}_i^c = x_i^r/\mathcal{C}_i^r$, and x_i^r a CES aggregator of the same form as eq. (1). This input is paid in units of the final good at price $\mathbb{P}_i$. Hence, the renewable firm maximization problem is:

$$\pi_i^r = \max_{\bar{e}_i^r} q_i^r \bar{e}_i^r - \mathcal{C}_i^r \bar{e}_i^r \mathbb{P}_i \ ,$$

where $\mathcal{C}_i^r$ is a constant and resulting in zero profits $\pi_i^r = 0$. As a result, the price of renewable and the market clearing are given by:

$$q_i^r = \mathcal{C}_i^r \mathbb{P}_i \ , \qquad \bar{e}_i^r = e^r \ . \tag{10}$$

[12]This is motivated by evidence that even the largest coal producers do not have coal rents above 1% of GDP.

This once again returns a perfectly elastic supply curve, which is a slightly stronger assumption in the context of renewable energy. In the short run, renewable energy requires investments in capacity, implying a fairly inelastic supply curve. This is especially true considering the intermittency problems of wind and solar energy, c.f. Gentile (2024). However, in the long run, technological progress and learning-by-doing create positive externalities, substantially decreasing the cost of clean energy, resulting in a decreasing supply curve, c.f. Arkolakis and Walsh (2023). I take the intermediary conservative assumption that the supply curve is flat and will explore the robustness of this assumption in future extensions.

2.4 The climate system

Carbon emissions released from the burning of fossil fuels create a climate externality as they feed back into the atmosphere, increasing temperatures and affecting damages. Despite the model being static, I incorporate climate system dynamics[13] as in standard Integrated Assessment Models. These future damages, summarized in the stationary equilibrium, affect the Social Cost of Carbon and the Pigouvian level of carbon taxation, as we will see in Section 5.1.

I model the damage functions affecting country i utility $\mathcal{D}_i^u(\mathcal{E})$ and productivity $\mathcal{D}_i^y(\mathcal{E})$ as a reduced-form summary of the impact of climate change. I develop a standard dynamic climate system that can be summarized in a static form in a simple way. It expresses the mapping from (i) emissions $\mathcal{E}$ to a path of atmospheric carbon concentration $\mathcal{S}_t$, (ii) from carbon concentration to a path of global and local temperatures T_{it}, and then from local temperatures to damage $\hat{\mathcal{D}}(T_{it}-T_i^\star)$, and (iv) finally summarizes it in present discounted value to obtain $\mathcal{D}_i^y(\mathcal{E})$ and $\mathcal{D}_i^u(\mathcal{E})$.

First, the static model represents stationary decisions on energy choices taken "once and for all". These yearly emissions from fossil fuels sum up to

$$\mathcal{E} = \sum_{i\in\mathbb{I}} \mathcal{P}_i(\xi^f e_i^f + \xi^c e_i^c) \ ,$$

where ξ^f and ξ^c represent the carbon concentration of oil-gas and coal, respectively. Accounting for population and TFP growth, this leads to a path of emissions given by $\mathcal{E}_t = e^{(\bar{g}+n)t}\mathcal{E}$, $\forall t$.[14] They represent trajectories of emissions given the emissions and policies decisions in the initial equilibrium, e.g. e_i^f and e_i^c.

Second, I consider a dynamic system – in continuous time – for carbon concentration in the atmosphere:

$$\dot{\mathcal{S}}_t = \zeta_t\mathcal{E}_t - \delta_s\mathcal{S}_t \qquad \text{with} \qquad \mathcal{S}_{t_0} = \mathcal{S}_{2024} \ ,$$

where δ_s is the exit rate of carbon out of the atmosphere, which is typically small for standard calibrations. To make the carbon concentration bounded and non-exploding – given the constant

[13] For simplicity, I refrain from using a larger scale climate system as in Dietz et al. (2021) or Folini et al. (2024).

[14] In the stationary model, we consider a balanced-growth path with global population and TFP growth and all the path of variables over time are expressed as level per "effective" capita. In future calibration, I will consider country-specific growth rates for TFP $\bar{g}_i$ and population n_i.

path of emissions – I follow Krusell and Smith (2022) by assuming that part of emissions $\mathcal{E}_t$ is abated via carbon capture and storage (CCS) modeled by the exogenous parameter ζ_t. The share of emissions abated grows to 100% in the long-run, implying that $\zeta_t \rightarrow_{t\rightarrow\infty} 0$. Increasing CCS allows the system to reach net-zero in several centuries, stabilizing cumulative carbon emissions and temperature.

Third, I assume a linear relationship between the cumulative CO_2 emissions $\mathcal{S}_t$ and the global temperature anomaly $\mathcal{T}_t$ compared to preindustrial levels.

$$\mathcal{T}_t = \chi \mathcal{S}_t = \chi\Big(\mathcal{S}_{t_0} + \int_{t_0}^{\infty} e^{-\delta_s(t-t_0)}\zeta_t \mathcal{E}_t dt\Big) ,$$

where χ is the climate sensitivity parameter, i.e. how much warming a ton of CO_2 causes, and where $\mathcal{E}_t$ is measured in carbon units, and $\mathcal{S}_{t_0}$ is the initial stock of carbon before all the policy decisions are made – i.e. in 2024. This specification is rationalized by a large climate-sciences literature, e.g. Dietz et al. (2021), that shows there exists an approximately linear relationship between $\mathcal{S}_t$ and $\mathcal{T}_t$, as is shown in the following Figure 1. It displays the relationship between temperature anomaly and cumulative CO_2 emissions over time, both for historical data in black and a large class of climate models in different Representative Concentration Pathways (RCP).

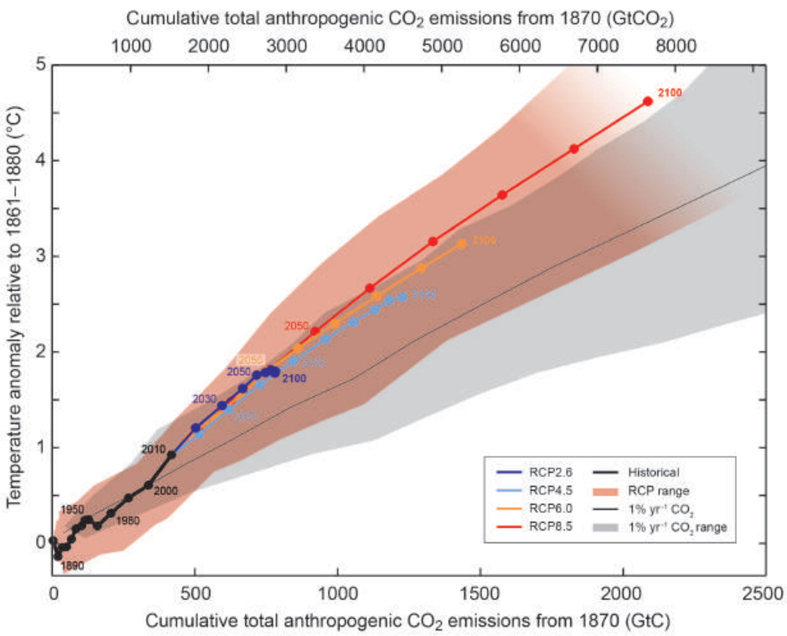


Figure 1: Linearity – Cumulative emissions and temperature, IPCC et al. (2022)

Fourth, I consider linear relationship between global and local temperatures, namely:

$$T_{it} = \bar{T}_{i0} + \Delta_i \mathcal{T}_t = \bar{T}_{i0} + \Delta_i \chi \mathcal{S}_t \ ,$$

where Δ_i is a linear pattern scaling parameter that depends on geographical factors such as albedo or latitude. In the quantification Section 4.5, I explain how I estimate this pattern scaling by regressing local temperatures on global temperature.

Fifth, I consider a period damage function $\hat{\mathcal{D}}(T_{it} - T_i^\star)$ where $T_i^\star$ is the "optimal" temperature for country i. The function $\mathcal{D}(\hat{T})$ is a reduced-form representation of the economic damage to productivity. In the baseline quantification, I assume damages are quadratic, as in standard Integrated Assessment Models. This methodology follows Krusell and Smith (2022), Kotlikoff, Kubler, Polbin, Sachs and Scheidegger (2021) and Burke et al. (2015). Such damage creates winners and losers: countries that are substantially warmer than a target temperature $T_i^\star$ are extremely affected by increases in temperature due to climate change. In contrast, regions with negative $T_{it} - T_i^\star$ benefit – at least in the short-run – from a warmer climate. I consider a slight deviation to the above articles by assuming that the target temperature $T_i^\star$ might be different across countries: an already warm regions have different adaptation costs compared to a country which is historically cold. As a result, the target temperature $T_i^\star = \alpha T^\star + (1-\alpha)\bar{T}_{i0}$ can be more or less tilted toward historical temperature. I discuss this quantification in Section 4.5.

Finally, to obtain a reduced-form static damage functions $\mathcal{D}_i^y(\mathcal{E})$ and $\mathcal{D}_i^u(\mathcal{E})$, for productivity and utility, respectively, I summarize the future costs of climate change in present-discounted value:

$$\mathcal{D}_i^y(\mathcal{E}) = \mathcal{D}_i^u(\mathcal{E}) = \bar{\rho} \int_{t_0}^{\infty} e^{-\bar{\rho} t}\, \hat{\mathcal{D}}(T_{it} - T_i^\star) dt \ ,$$

with $\bar{\rho} = \rho - n + \eta \bar{g}$ the "effective discount factor" and ρ is the household discount factor, n is the global population growth and $\bar{g}$ the global TFP growth. This net-of-damage function $\mathcal{D}_i^y(\mathcal{E})$ will be internalized by the Social Planner when making optimal climate policy choices.

2.5 Equilibrium

To close the model, we need to determine the final good prices for each country p_i, and we consider the market clearing for each good i

$$\begin{aligned} \mathcal{P}_i y_i &= \sum_{k \in \mathbb{I}} \mathcal{P}_k \tau_{ki} c_{ki} + \sum_{k \in \mathbb{I}} \mathcal{P}_k \tau_{ki} (x_{ki}^f + x_{ki}^c + x_{ki}^r) \\ \mathcal{P}_i \, \mathrm{p}_i \underbrace{y_i}_{= \mathcal{D}_i^y(\mathcal{E}) z_i F(\cdot)} &= \sum_{k \in \mathbb{I}} \frac{s_{ki}}{1 + \mathrm{t}_{ki}^b} \left(\mathcal{P}_k \mathrm{p}_k y_k + q^f (e_k^x - \mathcal{P}_k e_k^f) + \mathcal{P}_k \mathrm{t}_k^{ls} \right) \end{aligned} \tag{11}$$

where x_{ki}^f, x_{ki}^c and x_{ki}^r are the good inputs used by country k and imported from country i to produce fossil and renewable energy respectively. The second equation is a reformulation of the market clearing where the sales of countries i equals the expenditures from all countries k, coming from their incomes in good sales as well as net-exports of fossil energy.

To summarize, the competitive equilibrium of this economy is defined as follows:

Definition. Competitive equilibrium (C.E.):
For a set of policies $\{\mathrm{t}_i^\varepsilon, \mathrm{t}_{ij}^b, \mathrm{t}_i^{ls}\}_i$ across countries, a C.E. is a set of decisions $\{c_{ij}, e_i^f, e_i^c, e_i^r, e_i^x, \bar{e}_i^c, \bar{e}_i^r\}_{ij}$, and prices $q^f, \{\mathrm{p}_i, w_i, q_i^c, q_i^r\}_i$ such that:

(i) Households choose consumption $\{c_{ij}\}_{ij}$ maximizing utility eq. (1) s.t. the budget constraint eq. (2), which yield trade shares eq. (3)

(ii) Final good firms choose inputs $\{\ell_i, e_i^f, e_i^r\}_i$ to maximize profits, resulting in eq. (5)
(iii) Fossil energy firms maximize profits eq. (6) and extract/produce $\{e_i^x\}_i$ given by eq. (7)
(iv) Renewable and coal energy firms maximize profits, and supplies $\{\bar{e}_i^c, \bar{e}_i^r\}$ are given respectively by eq. (9) and eq. (10)
(v) Energy markets clears for fossils as in eq. (8) and for coal and renewable in eq. (9) and eq. (10)
(vi) Good markets clear for final good for each country as in eq. (11), and trade is balanced by Walras Law.

3 The optimal agreement design with endogenous participation

Because of unequal exposure to climate change and carbon policy, countries have different incentives to enforce climate policy, creating a free-riding problem. Therefore, the optimal carbon tax needs to account for endogenous participation. Designing a climate agreement reveals a trade-off between an intensive margin – associated with the choice of the policy instruments – and an extensive margin – related to the extent of participation in the agreement.

3.1 Agreement design and participation constraints

The social planner solves a Ramsey problem, choosing the optimal agreement, which boils down to a carbon tax, retaliatory tariffs on non-participants, and a set of countries participating in the agreement, subject to participation constraints. I first design the set of climate agreements considered and then define the planner's objective.

Definition. Climate Agreements
A climate agreement is a set $\{\mathbb{J}, \mathrm{t}^\varepsilon, \mathrm{t}^b\}$, with a coalition of countries $\mathbb{J} \subseteq \mathbb{I}$, a carbon tax t^ε, and a tariff t^b, such that in the competitive equilibrium,

- Countries $i \in \mathbb{J}$ are subject to a carbon tax t^ε on fossil energy e_i^f and coal e_i^c.
- If country j exits the agreement, club members $i \in \mathbb{J}$ charge uniform tariffs $\mathrm{t}_{ij}^b = \mathrm{t}^b$ on the final good imported from j.
- Countries in the club rebate the revenues of the carbon tax and tariffs to the household $\mathrm{t}_i^{ls} = \mathrm{t}^\varepsilon(\xi^f e_i^f + \xi^c e_i^c) + \sum_{j \notin \mathbb{J}} \mathrm{t}^b \tau_{ij} \mathrm{p}_j (c_{ij} + x_{ij}^f + x_{ij}^c + x_{ij}^r)$
- Countries inside the club, $i, j \in \mathbb{J}$, benefit from free-trade $\mathrm{t}_{ij}^b = 0$.
- Countries outside the club, $k \notin \mathbb{J}$, keep a passive trade policy,[15] $\mathrm{t}_{k\ell}^b = 0, \ \forall \ell \in \mathbb{I}$.
- All countries – members as well as non-members – still trade in fossil (oil-gas) energy at international price q^f.[16]

I keep the number of policy instruments $\mathbf{t} = \{\mathrm{t}^\varepsilon, \mathrm{t}^b\}$ considered in the agreement purposefully small for two reasons. First, this is consistent with the idea behind deviations of the Coase theorem: when bargaining over many policy instruments is associated with transaction costs, a negotiation between n parties – $\mathbb{J}$ countries here – "can be prevented from attaining a socially

[15] This assumption will be relaxed in the extension, Section 7.4.
[16] This assumption will be relaxed in Section 7.3.

desirable outcome", c.f. Weitzman (2015). For example, quantity targets ε_i and bilateral tariffs t^b_{ij} exacerbate free-riding incentives since an agreement requires all countries to accept the policies of all the other countries.[17] Second, carbon pricing is based on standard principles of Pigouvian taxation, where the optimal carbon tax equals the marginal cost of emitting one additional ton of carbon – the Social Cost of Carbon – common for all countries. This optimal uniform carbon price serves as a "focal point" of an international agreement, and the goal of this policy problem is to compare how it needs to be changed when accounting for free-riding.

Participation constraints

I define indirect utility $\mathcal{U}_i(\mathbb{J}, \mathrm{t}^\varepsilon, \mathrm{t}^b) \equiv u\big(c_{i(\mathbb{J},\mathrm{t}^\varepsilon,\mathrm{t}^b)}\mathcal{D}_i(\mathcal{E}_{(\mathbb{J},\mathrm{t}^\varepsilon,\mathrm{t}^b)})\big)$ as in Section 2.1. Then we can define participation constraints in two ways, depending on the type of deviations we consider.

1. *Unilateral deviation:* country i can choose to exit the agreement unilaterally. This does not affect the composition of the agreement or the decision of the other members. Country i in the agreement will participate if the value of staying is larger than the value of being outside the agreement:

$$\mathcal{U}_i(\mathbb{J}, \mathrm{t}^\varepsilon, \mathrm{t}^b) \geq \mathcal{U}_i(\mathbb{J}\backslash\{i\}, \mathrm{t}^\varepsilon, \mathrm{t}^b) \qquad \forall\, i \subseteq \mathbb{J}\,. \qquad \text{[Unilateral-Nash PC]} \tag{12}$$

2. *Sub-coalition deviation:* country i can choose to exit the agreement in cooperation with other members of a potential sub-coalition $\hat{\mathbb{J}}$. All these members leave the agreement. The decision of all those countries $i \in \hat{\mathbb{J}}$ to leave is made jointly: the value of being outside is above the value of staying for all $i \in \hat{\mathbb{J}}$. This makes the participation constraints more intricate and write as follow:

$$\mathcal{U}_i(\mathbb{J}, \mathrm{t}^\varepsilon, \mathrm{t}^b) \geq \mathcal{U}_i(\mathbb{J}\backslash\hat{\mathbb{J}}, \mathrm{t}^\varepsilon, \mathrm{t}^b) \qquad \forall\, i \in \hat{\mathbb{J}}\ \&\ \forall\, \hat{\mathbb{J}} \subseteq \mathbb{J}\,. \quad \text{[Coalition-Nash PC]} \tag{13}$$

The optimal agreement needs to account for these participation constraints, and be robust to unilateral or sub-coalition deviations.

Welfare criterion and planner's objective

We consider a global social planner maximizing the world welfare:

$$\max_{\mathbb{J},\mathrm{t}^\varepsilon,\mathrm{t}^b} \mathcal{W}(\mathbb{J}, \mathrm{t}^\varepsilon, \mathrm{t}^b) = \max_{\mathbb{J},\mathrm{t}^\varepsilon,\mathrm{t}^b} \sum_{i\in\mathbb{I}} \mathcal{P}_i\omega_i\, \mathcal{U}_i(\mathbb{J}, \mathrm{t}^\varepsilon, \mathrm{t}^b)\ , \tag{14}$$

subject to participation constraint – Unilateral Nash, robust to deviation

$$\mathbb{J} \in \mathbb{S}(\mathrm{t}^\varepsilon, \mathrm{t}^b) = \{\mathcal{J} \mid \mathcal{U}_i(\mathcal{J}, \mathrm{t}^\varepsilon, \mathrm{t}^b) \geq\ \mathcal{U}_i(\mathcal{J}\backslash\{i\}, \mathrm{t}^\varepsilon, \mathrm{t}^b), \quad \forall\, i \in \mathcal{J}\}\ , \tag{15}$$

[17] For this reason, I refrain from studying agreements that bargain over a set of country-specific taxes t^ε_i or emissions quantity targets ε_i, and bilateral tariffs t^b_{ij}, since the bargaining costs increase at least proportionally – if not exponentially – in the number of countries. The case of individual carbon taxes is analyzed in Bourany (2025) absent free-riding. Accounting for strategic interactions between I countries also makes the problem computationally more involved, as in theory all the instruments need to account jointly for the Lagrange multipliers of all the participation constraints, for any possible coalition.

or Coalitional-Nash, robust to sub-coalition $\hat{\mathbb{J}}$ deviations:

$$\mathbb{J} \in \mathbb{C}(\mathrm{t}^{\varepsilon}, \mathrm{t}^{b}) = \{\mathcal{J} \mid \mathcal{U}_i(\mathcal{J}, \mathrm{t}^{\varepsilon}, \mathrm{t}^{b}) \geq \mathcal{U}_i(\mathcal{J} \backslash \hat{\mathbb{J}}, \mathrm{t}^{\varepsilon}, \mathrm{t}^{b}), \quad \forall\, i \in \mathcal{J},\ \forall\, i \in \hat{\mathbb{J}}\ \&\ \forall\, \hat{\mathbb{J}} \subseteq \mathcal{J}\} \,. \tag{16}$$

where ω_i are the Pareto weights, and $\mathcal{P}_i$ the population size of country i. The social planner maximizes world welfare, in part with the goal of fighting the climate change externality. As a result, the planner maximizes the sum over $\mathbb{I}$ instead of $\mathbb{J}$. In the case where a planner only maximizes the coalition $\mathbb{J}$'s welfare, it would yield the unintended consequences that the optimal agreement could be restricted to a subset of rich, cold, high-value $\mathcal{U}_i$ countries which would manipulate terms-of-trade and potentially subsidize fossil fuels. I give intuitions for such results in Section 5.2 when countries choose their climate and trade policy unilaterally. However, these outcomes are unintended for global climate agreements, which aim at maximizing world welfare.

The set of agreements stable under Coalition-Nash resembles the concept of "core" in general equilibrium theory. Both of these sets $\mathbb{S}(\mathrm{t}^{\varepsilon}, \mathrm{t}^{b})$, $\mathbb{C}(\mathrm{t}^{\varepsilon}, \mathrm{t}^{b})$ could be empty: it is possible that no country finds it beneficial to be part of the agreement for a given policy $\mathrm{t}^{\varepsilon}, \mathrm{t}^{b}$.

Extensive margin vs. intensive margin tradeoff

The problem of a world planner determines jointly the policy instruments $\mathbf{t} = \{\mathrm{t}^{\varepsilon}, \mathrm{t}^{b}\}$ and the choice of the country subject to participation constraints $\mathbb{J} \in \mathbb{S}(\mathrm{t}^{\varepsilon}, \mathrm{t}^{b})$ or $\mathbb{C}(\mathrm{t}^{\varepsilon}, \mathrm{t}^{b})$. As a result, this reveals an extensive-intensive margin trade-off. For a given set of participants $\mathbb{J}$, higher carbon tax t^{ε} and lower tariffs t^{b} increase global welfare. That is, this planner would like to reduce carbon emissions and promote free trade at the intensive margin to maximize welfare. However, this choice of instruments also affects countries' participation: higher taxes t^{ε} and lower tariffs t^{b} reduce incentives for countries to participate. If a country deviates by exiting the agreement, it increases its emissions, and international trade is reduced, lowering welfare at the extensive margin. As a result, the planner would like to balance these two countervailing effects. This tradeoff is analyzed in detail in the context of this model in Section 6.3.

3.2 Optimal design and solution method

This design problem combines a choice of instruments and a choice of countries, making it difficult to solve. I provide two methods to handle the joint optimal policy/combinatorial discrete choice problem.

3.2.1 Framework for the optimal design

I formalize the policy problem under the two types of participation constraints – Unilateral deviations vs. Coalition deviations – subject to the allocation being a competitive equilibrium. I consider a general class of policy instruments $\mathbf{t}$ that encompass carbon tax $\mathrm{t}_i^{\varepsilon} = \mathrm{t}^{\varepsilon}\mathbb{1}_{\{i \in \mathbb{J}\}}$, uniform tariffs on non-members $\mathrm{t}_{ij}^{b} = \mathrm{t}^{b}\mathbb{1}_{\{i \in \mathbb{J}, j \notin \mathbb{J}\}}$ as well as potential additional instruments as analyzed in

Section 7. The design problem can be stated as

$$\max_{\mathbb{J},\mathbf{t}} \mathcal{W}(\mathbb{J},\mathbf{t}) = \max_{\mathbb{J},\mathbf{t}} \sum_{i\in\mathbb{I}} \mathcal{P}_i\, \omega_i\, \mathcal{U}_i(\mathbb{J},\mathbf{t}) \ ,$$

$$s.t. \qquad \mathbb{J} \in \mathbb{S}(\mathbf{t})\,, \quad or \quad \mathbb{J} \in \mathbb{C}(\mathbf{t}) \ .$$

Participation constraints $\mathbb{S}(\mathbf{t})$ or $\mathbb{C}(\mathbf{t})$, as defined in eq. (15) and eq. (16), make the problem intricate as they limit the instruments the planner can use for each set of countries in the agreements.

This design problem is particularly challenging: solving jointly for the optimal policies and the coalition of countries, subject to individual participation constraints, makes the problem intricate, especially in a rich quantitative model. I take the following approach: I split the problem into an inner problem and an outer problem. First, the planner chooses the policy instruments $\mathbf{t}$. Then, given $\mathbf{t}$, the optimal coalition is chosen subject to participation constraints $\mathbb{S}(\mathbf{t})$ or $\mathbb{C}(\mathbf{t})$. If no coalition is achievable, then the welfare for those instruments is $-\infty$. This choice of country in the inner problem is analogous to a combinatorial discrete choice problem (CDCP), which yields a optimal coalition $\mathbb{J}^\star(\mathbf{t})$,

$$\max_{\substack{\mathbb{J},\mathbf{t}\\ \mathbb{J}\in\mathbb{S}(\mathbf{t})}} \mathcal{W}(\mathbb{J},\mathbf{t}) = \max_{\mathbf{t}} \max_{\mathbb{J}\,|\,\mathbb{J}\in\mathbb{S}(\mathbf{t})} \mathcal{W}(\mathbb{J},\mathbf{t}) \ .$$

I explain in Appendix E why the opposite approach – solving for the coalition as an outer problem and for policy instruments in the inner problem – is intractable.[18]

The *outer problem* for the choice of instrument $\mathbf{t}$ is solved with a simple grid search[19] since the indirect welfare is now discontinuous and non-convex in the application:

$$\max_{\mathbf{t}} \widehat{\mathcal{W}}(\mathbf{t}) \ , \qquad \text{where} \qquad \widehat{\mathcal{W}}(\mathbf{t}) = \max_{\mathbb{J}\,|\,\mathbb{J}\in\mathbb{S}(\mathbf{t})} \mathcal{W}(\mathbb{J},\mathbf{t}) = \mathcal{W}(\mathbb{J}^\star(\mathbf{t}),\mathbf{t}) \ .$$

3.2.2 Solution methods

I propose two methods to solve the *inner problem* of the optimal choice of countries $\mathbb{J}^\star$, out of all the possible combinations $\mathcal{P}(\mathbb{I})$. This combinatorial discrete choice problem is prohibitive numerically for large numbers of countries $\#\mathbb{I}$. To handle this challenge, I first use an exhaustive search method (brute force method), and then I propose a squeezing procedure adapted from the trade literature, e.g. Arkolakis, Eckert and Shi (2023), as a more efficient alternative in the unilateral-Nash case. I introduce the combinatorial problem before presenting each method in turn. The combinatorial discrete choice problem for a given policy $\mathbf{t}$ is given by:

$$\max_{\mathbb{J}\in\mathcal{P}(\mathbb{I})} \mathcal{W}(\mathbb{J},\mathbf{t}) \ ,$$

$$s.t. \quad \mathbb{J} \in \mathbb{S}(\mathbf{t}) \ .$$

[18] In subgame perfect equilibria, it makes the Lagrange multipliers on the participation constraints ν_i depends not only on the coalition considered $\mathbb{J}$, but all the coalitions in every subgame, e.g. $\mathbb{J}\backslash\{i\}$ etc. These multipliers affect the policy choice and make the problem unsolvable.

[19] For this reason, I keep the number of instruments small $\mathbf{t}=\{t^\varepsilon, t^b\}$, to search over a low-dimensional space $\mathbf{t} \in \mathbb{R}^2_+$.

I express the Lagrangian of the constrained optimization, with multiplier ν_i for country i's participation, we obtain, with a slight abuse of notation[20], as follow:

$$\max_{\mathbb{J}\in\mathcal{P}(\mathbb{I})} \mathcal{W}(\mathbb{J},\mathbf{t}) + \sum_{i\in\mathbb{J}} \nu_{i,\mathbb{J}}\Big(\mathcal{U}_i(\mathbb{J},\mathbf{t}) - \mathcal{U}_i(\mathbb{J}\backslash\{i\},\mathbf{t})\Big) =: \max_{\mathbb{J}\in\mathcal{P}(\mathbb{I})} \widetilde{\mathcal{W}}(\mathbb{J},\mathbf{t}) \tag{17}$$

in the case where the participation constraints only account for unilateral deviations.[21]

First method: Exhaustive enumeration

First, when the number of countries $I = \#\mathbb{I}$ is small, one obvious yet costly solution is to perform an exhaustive search over $\mathcal{P}(\mathbb{I})$. The idea is to enumerate all the combinations $\mathbb{J} \in \mathcal{P}(\mathbb{I})$, and evaluate welfare $\mathcal{W}(\mathbb{J},\mathbf{t})$. This has evidently a computational cost proportional to $2^{\#\mathbb{I}}$, i.e. the number of potential combinations.

This solution has, however, the advantage of considering all the participation constraints – including the coalition-robust agreements – "for free". Indeed, we can assess if the coalition is stable both in the case of unilateral-Nash $\mathbb{J} \in \mathbb{S}(\mathbf{t})$ and coalitional-Nash $\mathbb{J} \in \mathbb{C}(\mathbf{t})$, for all sets $\mathbb{J}$. This is feasible because every possible deviation of sub-groups $\hat{\mathbb{J}}$ yields a new agreement $\mathbb{J}' = \mathbb{J}\backslash\hat{\mathbb{J}}$ which is already computed as another coalition $\mathbb{J}' \in \mathcal{P}(\mathbb{I})$. If one of the participation constraints is violated, the set considered $\mathbb{J}'$ is discarded, i.e. $\nu_{i,\mathbb{J}'} = \infty$, $\mathcal{W}(\mathbb{J}') = -\infty$. In practice, several coalitions can be stable for a given policy $\mathbf{t}$, and the exhaustive search selects the one that maximizes welfare. In practice, among all the stable coalitions the one that maximizes welfare is the largest one since for a given policy $\mathbf{t} = \{\mathrm{t}^{\varepsilon}, \mathrm{t}^{b}\}$, the larger the coalition, the higher the gains from trade and the gains from reducing emissions.[22]

Second method: Squeezing procedure for CDCP with Participation Constraints

Second, since full enumeration is costly, I provide an alternative algorithm inspired by methods used in the international trade literature to solve combinatorial discrete choice problems. The additional difficulty that needs to be considered is the presence of participation constraints. In this section, we only consider unilateral deviations. The idea behind this method is greatly inspired by Arkolakis, Eckert and Shi (2023) and Farrokhi and Lashkaripour (2024).

The idea is to build iteratively sets that are lower bound $\underline{\mathcal{J}}$ and upper bound $\overline{\mathcal{J}}$ for the optimal coalition $\mathbb{J}$: subset $\underline{\mathcal{J}}$ *includes* all the countries that are known to be part of the optimal set $\mathbb{J}$ and $\overline{\mathcal{J}}$ is a superset, which it *excludes* the countries that we know are not part of the optimal set. The set $\overline{\mathcal{J}}\backslash\underline{\mathcal{J}}$ is the set of potential countries. The natural starting point is $\underline{\mathcal{J}} = \varnothing$, $\overline{\mathcal{J}} = \mathbb{I}$.

The squeezing step in standard CDCP is a mapping from $\mathcal{J}$ to members that bring a positive marginal value to the objective $\mathcal{W}(\mathbb{J}) := \mathcal{W}(\mathbb{J},\mathbf{t})$. The modification needed in settings with participation constraints is that the country also needs to gain marginal *individual* value $\mathcal{U}_i(\mathbb{J}) := \mathcal{U}_i(\mathbb{J},\mathbf{t})$

[20] The objective function $\mathcal{W}(\mathbb{J},\mathbf{t})$ is not continuous, differentiable, or even convex. The handling of the inequality constraints could not, in theory, rely on the KKT theorem which applies in the $\mathcal{C}^1$ and convex case. However, with $\nu_{i,\mathbb{J}} = 0$ if $\mathbb{J} \in \mathbb{S}(\mathbf{t})$ and $\nu_{i,\mathbb{J}} = \infty$ if $\mathbb{J} \notin \mathbb{S}(\mathbf{t})$, the problem well defined.

[21] A longer list of constraints needs to be included if we consider coalition deviations.

[22] As long as t^{ε} is below the globally optimal carbon tax as we derive it in Section 5.1.2.

to be part of the coalition:

$$\Phi(\mathcal{J},\mathbf{t}) \equiv \{j \in \mathbb{I} \,|\, \Delta_j \mathcal{W}(\mathcal{J},\mathbf{t}) > 0 \text{ \& } \Delta_j \mathcal{U}_j(\mathcal{J},\mathbf{t}) > 0\} \tag{18}$$

where the marginal values for global welfare and individual welfare are

$$\Delta_j \mathcal{W}(\mathcal{J},\mathbf{t}) \equiv \mathcal{W}(\mathcal{J}\cup\{j\},\mathbf{t}) - \mathcal{W}(\mathcal{J}\backslash\{j\},\mathbf{t}) = \sum_{i\in\mathbb{I}} \mathcal{P}_i \omega_i \big(\mathcal{U}_i(\mathcal{J}\cup\{j\},\mathbf{t}) - \mathcal{U}_i(\mathcal{J}\backslash\{j\},\mathbf{t})\big)$$
$$\Delta_j \mathcal{U}_i(\mathcal{J},\mathbf{t}) \equiv \mathcal{U}_i(\mathcal{J}\cup\{j\},\mathbf{t}) - \mathcal{U}_i(\mathcal{J}\backslash\{j\},\mathbf{t})$$

The iterative procedure builds the lower bound $\underline{\mathcal{J}}$ and upper bound $\overline{\mathcal{J}}$ by successive application of the squeezing step.

$$\underline{\mathcal{J}}^{(k+1)} = \Phi(\underline{\mathcal{J}}^{(k)},\mathbf{t}) \qquad\qquad \overline{\mathcal{J}}^{(k+1)} = \Phi(\overline{\mathcal{J}}^{(k)},\mathbf{t}) \tag{19}$$

Under some conditions (complementarity, as defined next section and in the appendix) this sequential procedure yields two sets $\underline{\mathcal{J}}$ and $\overline{\mathcal{J}}$ such that $\underline{\mathcal{J}} \subseteq \mathbb{J} \subseteq \overline{\mathcal{J}}$. In some cases $\underline{\mathcal{J}} = \overline{\mathcal{J}} = \mathbb{J}$, yielding the optimal coalition. If not, with $\overline{\mathcal{J}}\backslash\underline{\mathcal{J}} = \mathcal{J}^{pot}$, we find the optimal coalition by searching exhaustively over all coalitions $\mathcal{J} = \underline{\mathcal{J}} \cup \hat{\mathcal{J}}$, with $\hat{\mathcal{J}} \in \mathcal{P}(\mathcal{J}^{pot})$.

Applicability of the squeezing procedure

From the combinatorial discrete choice literature, Arkolakis, Eckert and Shi (2023), we know that the squeezing procedure applies in cases where the model exhibits "complementarity" or single-crossing differences in choices. We detail how these conditions can be expressed in Appendix E.2.

Indeed, we say that the objective $\mathcal{W}(\mathcal{J})$ satisfies "complementarity", if the marginal gain $\Delta_j \mathcal{W}(\mathcal{J})$ of the objective is monotone in the set $\mathcal{J}$, i.e. $\Delta_j \mathcal{W}(\mathcal{J}) \leq \Delta_j \mathcal{W}(\mathcal{J}')$, for $\mathcal{J} \subseteq \mathcal{J}'$ & $j \in \mathbb{I}$. In the climate agreement setting, participation constraints and stability require to adjust the welfare objective, from $\mathcal{W}(\mathbb{J})$ to $\widetilde{\mathcal{W}}(\mathbb{J})$ as in eq. (17). In this context, the complementarity (or single crossing differences in choice for its weaker form), with participation constraints, takes an intricate form (SCD-C-PC) which we detail in Appendix E.2.

Theorem The SCD-C-PC from below is *sufficient* for the application of modified squeezing algorithm, i.e. successive application of eq. (18), starting from $\{\varnothing, \mathbb{I}\}$ and eq. (19), to yield bounding sets $\underline{\mathcal{J}} \subseteq \mathbb{J} \subseteq \overline{\mathcal{J}}$ in CDCPs with participation constraints.

One of the advantages is that, for a small number of countries $\#\mathbb{I} \approx 10$, we can evaluate numerically if the sufficient conditions mentioned above are satisfied. The fact that the model is rich, with many dimensions of heterogeneity and general equilibrium effects through energy markets and international trade, prevents the simple evaluation of those sufficient conditions analytically.

4 Quantification

The model is calibrated to a panel of thirty-two countries to provide realistic predictions on the impact of optimal carbon policy. I first describe the data used. I then provide details on the quantification, which functional forms are used, and how the parameters are calibrated to match the data. I summarize in Table 1 the dimensions of heterogeneity of the model. Table 2 in the appendix contains the summary table for the calibration described in this section.

4.1 Data

First, I describe briefly the data used to calibrate the model. I use data for the year 2018-2023, taking the average over that period to smooth out the effect of the COVID-19 recession on energy and macroeconomic data. I use a sample of twenty-five countries and seven regions: (i) United States, (ii) Canada, (iii) China, (iv) Germany, (v) France, (vi) Spain, (vii) Italy, (viii) Rest of EU, (ix) United Kingdom, (x) India, (xi) Pakistan, (xii) Rest of South Asia, (xiii) Nigeria, (xiv) South Africa, (xv) Rest of Africa, (xvi) Egypt, (xvii) Iran, (xviii) Saudi Arabia, (xix) Turkey, (xx) Rest of Middle-East+Maghreb, (xxi) Russia, (xxii) Rest of CIS, (xxiii) Australia, (xxiv) Japan, (xxv) Korea, (xxvi) Indonesia, (xxvii) Thailand, (xxviii) Rest of South-East Asia, (xxix) Argentina, (xxx) Brazil, (xxxi) Mexico, (xxxii) Rest of Latin America.[23]

I use data for GDP per capita, in Purchasing Power Parity (PPP, in 2011 USD) from the World Bank, as collected and processed by the Maddison Project (Bolt and van Zanden (2023)). For the energy variables, I use the comprehensive data collected and processed in the Statistical Review of Energy (Energy Institute (2024)) that includes the production and consumption of various energy sources, including Oil, Gas, and Coal. It also includes proven reserves of those fossil fuels. For energy rent, I use the World Development Indicators that use national accounts to measure the share of GDP coming from energy (oil, gas and coal) and natural resource rents. Finally, for temperature, I use the same time series as Burke et al. (2015), which use the temperature at country level, averaged over the year and weighted by population across locations. For trade variables, I take the trade flows and gravity variables compiled by the CEPII in Conte et al. (2022).

4.2 Welfare and Pareto weights

The welfare function that the climate agreement designer would maximize is the weighted sum of individual utilities in all countries:

$$\mathcal{W}(\mathbb{J}, \mathbf{t}) = \sum_{i \in \mathbb{I}} \mathcal{P}_i \, \omega_i \, \mathcal{U}_i(\mathbb{J}, \mathbf{t}) = \sum_{i \in \mathbb{I}} \mathcal{P}_i \, \omega_i \, u(c_i) \mathcal{D}_i^u(\mathcal{E})$$

with $\mathcal{P}_i$ the population size per country, ω_i the Pareto weights and $\mathcal{U}_i$ the country indirect utility per capita. Note that the climate agreement designer maximizes the *world* welfare.

[23] In a previous iteration of this project, I considered a panel of 10 regions, to compare the exhaustive enumeration with the algorithm for Combinatorial Discrete Choice.

Following the discussion in Anthoff et al. (2009), Nordhaus (2011) and Nordhaus and Yang (1996), one would like to choose Pareto weights that eliminate redistributive effects that are orthogonal to climate change and climate policy. To that purpose, I choose the "Negishi" Pareto weights that make the preexisting competitive equilibrium efficient under that welfare metric. This implies that:

$$\omega_i = \frac{1}{u'(\bar{c}_i)\mathcal{D}_i^u(\mathcal{E})} \qquad \Leftrightarrow \qquad C.E.(\bar{c}_i) \in \underset{\bar{c}_i}{\operatorname{argmax}} \sum_i \mathcal{P}_i \omega_i u(\bar{c}_i)\mathcal{D}_i^u(\mathcal{E})$$

$$\omega_i u'(\bar{c}_i)\mathcal{D}_i^u(\mathcal{E}) = \omega_j u'(\bar{c}_j)\mathcal{D}_j^u(\mathcal{E}) \qquad \forall i,j \ \in \mathbb{I}$$

where $\bar{c}_i$ is the consumption level in the present competitive equilibrium – the period 2018-2023 – absent future climate damage. This implies that the climate agreement and the carbon policy do not look for redistributing across countries through goods and energy general equilibrium effects. However, global warming, carbon taxation, and tariffs have redistributive effects, as they change the distribution of c_i. These effects are taken into account in the choice of policies, as we see in Section 5.1.2. In Figure 2, I display the weights ω_i, and $\omega_i\widehat{\mathcal{P}}_i$ adjusted for population $\widehat{\mathcal{P}}_i = \mathcal{P}_i/\mathcal{P}$.

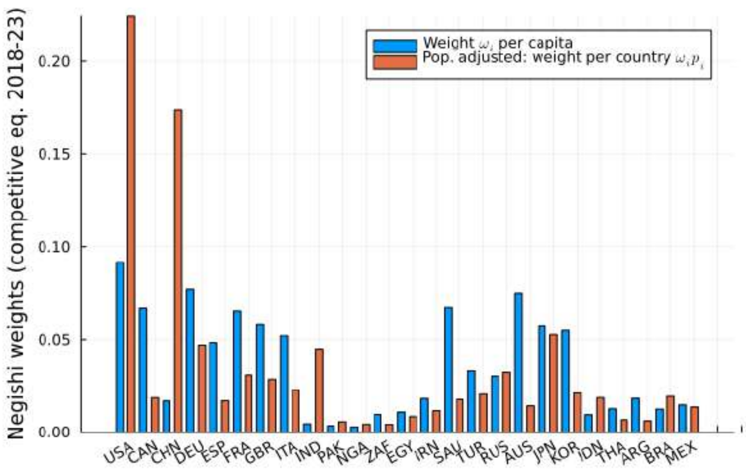


Figure 2: Pareto weights across regions
weights ω_i (blue, left), $\widehat{\mathcal{P}}_i\omega_i$ (red, right)

4.3 Macroeconomy, trade and production

For the macroeconomic part of the framework, I consider standard utility and production functions. First, I consider constant relative risk aversion (CRRA) utility over consumption, as well as climate damage in the utility function. This implies that countries that have a very low production / GDP per capita still suffer potential large losses due to climate damages:

$$\mathcal{U}_i = \frac{\left(c_i\widetilde{\mathcal{D}}_i^u(\mathcal{E})\right)^{1-\eta}}{1-\eta} = \frac{c_i^{1-\eta}}{1-\eta}\mathcal{D}_i^u(\mathcal{E})$$

I calibrate the CRRA/IES parameter to be $\eta = 1.5$, taken from Barrage and Nordhaus (2024).[24] Moreover, the damage $\mathcal{D}^u_i(\mathcal{E})$ is adjusted for the curvature in the utility function.

For production, I use a nested CES framework. The firm combines a Cobb-Douglas bundle of capital k_i and labor ℓ_i[25] with a composite of energy e_i, with elasticity σ^y. Second, the energy e_i aggregates the different energy sources: oil and gas e^f, coal e^c_i, and renewable/non-carbon e^r_i, with elasticity σ^e.

$$\text{Output} \qquad y_i = \mathcal{D}^y_i(\mathcal{E})\, z_i\, \bar{y}_i = \mathcal{D}^y(\mathcal{E})\, z_i \Big((1-\varepsilon)^{\frac{1}{\sigma^y}} (e_i)^{\frac{\sigma^y-1}{\sigma^y}} + \varepsilon^{\frac{1}{\sigma^y}} (k_i^\alpha \ell_i^{1-\alpha})^{\frac{\sigma^y-1}{\sigma^y}} \Big)^{\frac{\sigma^y}{\sigma^y-1}} ,$$

$$\text{Energy} \qquad e_i = \Big((\omega^f)^{\frac{1}{\sigma^e}} (e^f_i)^{\frac{\sigma^e-1}{\sigma^e}} + (\omega^c)^{\frac{1}{\sigma^e}} (e^c_i)^{\frac{\sigma^e-1}{\sigma^e}} + (\omega^r)^{\frac{1}{\sigma^e}} (e^r_i)^{\frac{\sigma^e-1}{\sigma^e}} \Big)^{\frac{\sigma^e}{\sigma^e-1}} .$$

To calibrate these functions, I set the capital-labor ratio to be $\alpha = 0.35$ to match the cost share of capital. For the energy share, I set $\varepsilon = 0.10$ to match the average energy cost share of $\frac{q^e_i e_i}{\mathrm{p}_i y_i} = 6\%$, as measured in Kotlikoff, Kubler, Polbin and Scheidegger (2021) and used in Krusell and Smith (2022). For the elasticity between energy and other inputs, I set $\sigma^y = 0.3$ for all countries, which is in the range of estimates in Papageorgiou et al. (2017), among others.[26] This implies that capital/labor and energy are complementary in production: an increase in the price of energy has a strong impact on output as it is less productive to "substitute away" to other inputs – capital, labor here. This aligns with other empirical and structural evidence on the impact of energy shocks, e.g. Hassler et al. (2019). For each energy source, I calibrate the energy mix for oil-gas, with $\omega^f = 0.56$, coal $\omega^c = 0.27$, and non-carbon $\omega^r = 0.17$, to match the aggregate shares in each of these energy sources in the data. In the next section, I document how to match the individual countries' energy mix using energy prices/costs. Finally, for the elasticity between energy inputs, I use the value $\sigma_e = 2$, following the rest of the literature, i.e. Papageorgiou et al. (2017), Kotlikoff, Kubler, Polbin and Scheidegger (2021), and Hillebrand and Hillebrand (2019), among others.

I calibrate the productivity z_i of the production function $y_i = \mathcal{D}^y_i(\mathcal{E})\, z_i\, \bar{y}_i$ to match exactly the GDP, $y_i \mathrm{p}_i$, across countries. This parameter z_i, represents productivity residuals as well as institutional/efficiency differences across countries. In Figure 3, I show the GDP levels, as they replicated with this model.

Finally, we use trade flow data as seen in Figure 23 to match the pattern of international trade in goods. First, I estimate a gravity regression between trade flow and geographical distance[27]– with fixed effects for importers and exporters – finding an elasticity with distance $\kappa = -1.45$. To rationalize it in the model, I project iceberg trade costs on this geographical distance $\tau_{ij} = d^\beta_{ij}$. All the residual differences in trade flows, not rationalized by trade costs, τ_{ij}, prices, p_j, or demand,

[24]This is slightly lower than the standard value $\eta = 2$, for the reason that higher curvature would imply more unequal weights, ω_i, across different countries.

[25]Labor is inelastically supplied $\ell_i = \bar{\ell}_i$ in each country and normalized to 1 – since the country size $\mathcal{P}_i$ is already taken into account. As a result, all the variables can be seen as input per capita.

[26]It also aligns with my own estimation in Bourany (2022).

[27]The gravity regression is standard: $\log x_{ij} = \kappa \log d_{ij} + \alpha_i + \gamma_j$. In the model, $\tau_{ij} = d^\beta_{ij}$, we get $\kappa = (1-\theta)\beta$.

y_i, are then explained by differences in preferences a_{ij}. We calibrate β to minimize dispersion in a_{ij} over countries j, which implies a trade elasticity $\theta = 5.0$. I calibrate those parameters a_{ij} to minimize the distance – mean squared error – between model-generated trade shares $s_{ij} = \frac{c_{ij}\tau_{ij}\mathrm{p}_j}{c_i\mathbb{P}_i} = a_{ij}\frac{(\tau_{ij}\mathrm{p}_j)^{1-\theta}}{\sum_k a_{ik}(\tau_{ik}\mathrm{p}_k)^{1-\theta}}$, and observed shares $\bar{s}_{ij}$ in the data. Since our model imposes a trade balance, those trade shares cannot be matched exactly, only approximately. It does match the fact that some countries are relying more on trade exports and imports – like China, East and South-East Asia, and Europe – compared to others – Middle East, Africa, and Russia.

4.4 Energy markets

For the energy market, I match the energy mix of different countries, using the CES framework displayed above, as well as differences in cost of production. For the supply side, we use iso-elastic fossil extraction cost, to replicate the oil-gas supply of fossil producers.

First, in this model, oil and gas are traded on international markets, with demand $\mathcal{P}_i e_i^f$ from the final good firm and supply e_i^x from the fossil energy firm, extracting oil and gas from its own reserves. We use the extraction function $\mathcal{C}_i^f$ to have the following isoelastic form

$$\mathcal{C}_i^f(e_i^x, \mathcal{R}_i)\mathbb{P}_i = \frac{\bar{\nu}_i}{1+\nu}\Big(\frac{e_i^x}{\mathcal{R}_i}\Big)^{1+\nu}\mathcal{R}_i\mathbb{P}_i\ .$$

which is homogeneous of degree one in $(e_i^x, \mathcal{R}_i)$. The inputs are paid in the price of the consumption bundle $\mathbb{P}_i$ since the input $x_i^f = \mathcal{C}_i^f(e_i^x, \mathcal{R}_i)$ takes the same CES form as the consumption demand c_i.[28] This implies the profit function

$$\mathcal{P}_i\pi_i^f = q^f e_i^x - \mathcal{C}_i^f(e_i^x, \mathcal{R}_i) = \frac{\nu_i\bar{\nu}_i}{1+\nu_i}\Big(\frac{e_i^x}{\mathcal{R}_i}\Big)^{1+\nu_i}\mathcal{R}_i\mathbb{P}_i\ .$$

I calibrate the three parameters $\mathcal{R}_i$, ν_i and $\bar{\nu}_i$ to match two country-level variables e_i^x and π_i^f. The reserve $\mathcal{R}_i$ is taken directly from the data on oil and gas reserves documented by Energy Institute (2024). I calibrate the slope of this cost function $\bar{\nu}_i$ to match exactly the production of oil and gas e_i^x, as informed by that same data source. This is displayed in Figure 4. I then calibrate the curvature of the cost function to match the share $\eta_i^\pi = \frac{\pi_i^f}{y_i\mathrm{p}_i + \pi_i^f}$ of fossil energy profit as share of GDP. I choose ν to minimize the distance – mean squared error – between the model share η_i^π and the data, successfully matching the share within $5-10$ percentage points. Differences in oil and gas energy rent across countries are not only determined by differences in cost and technology, but also in differences in trade costs and market power – by the existence of OPEC which control more than 28% of oil supply and around 15% of natural gas supply. This explains why it is difficult to match exactly the value η_i^π. However, to keep the simplicity and tractability of the model, I refrain from adding an additional Armington structure over energy sources, or oligopoly power over oil

[28] I express the oil-gas extraction with a cost function $x_i^f = \mathcal{C}_i^f$. We can also express analogously with the following production function:

$$e_i^x = g(x_i^f) = \big(\tfrac{1+\nu_i}{\bar{\nu}_i}\big)^{\frac{1}{1+\nu}}\mathcal{R}_i^{\frac{\nu_i}{1+\nu_i}}(x_i^f)^{\frac{1}{1+\nu_i}}$$

where the inputs x_i^f are paid in the final good bundle. This production has constant returns to scale in $(x_i^f, \mathcal{R}_i)$.

and gas as discussed in Bornstein et al. (2023) and Hassler et al. (2010).

Second, I match the energy mix of the different countries by relying on the two assumptions made in the model: (i) coal and renewable are only traded at the country level: $\bar{e}_i^c = e_i^c$ and $\bar{e}_i^r = e_i^r$ and (ii) the cost function is linear in goods, i.e. the production is Constant Returns to Scale, implying $q_i^c = \mathcal{C}_i^c \mathbb{P}_i$ and $q_i^r = \mathcal{C}_i^r \mathbb{P}_i$. This allows me to match the energy mix of each country by calibrating the energy costs parameters $\mathcal{C}_i^c$ and $\mathcal{C}_i^r$ for each country to match the data on coal share $\frac{e_i^c}{e_i^f + e_i^c + e_i^r}$ and non-carbon energy share $\frac{e_i^r}{e_i^f + e_i^c + e_i^r}$. Using the CES framework above, I match exactly the energy shares, successfully identifying countries that are more reliant on coal vs. oil and gas vs. non-carbon/renewable: for example, China and India are highly coal-dependent, and Russia, Middle-East and United-States/Canada are the biggest consumers of oil and gas.

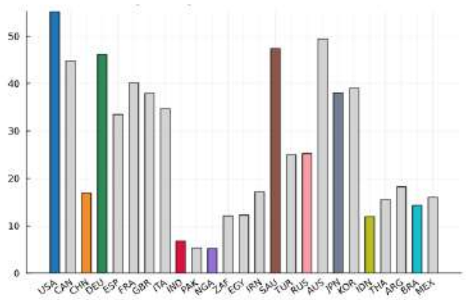


Figure 3: GDP per capita
Thsds 2011-USD PPP, avg. 2018-2023

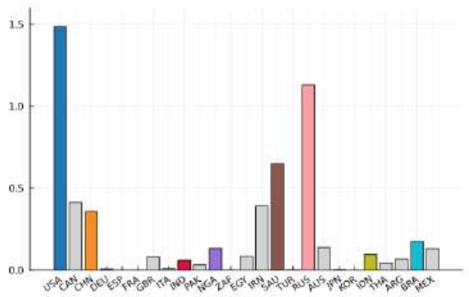


Figure 4: Oil and gas production
GTOE (gigatons oil equiv.), avg. 2018-2023

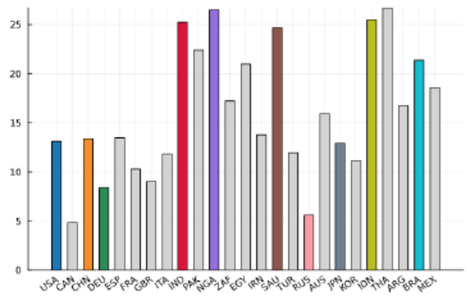


Figure 5: Temperatures
Avg., population-weighted, 2015

4.5 Climate system

Finally, I calibrate the climate model described in Section 2.4 to match important features of the relationship between carbon emissions, temperatures and climate damages.

First, I calibrate two parameters of the global climate system: the climate sensitivity χ, i.e. the reaction of global temperature, $\mathcal{T}_t$, to the atmospheric concentration of CO_2, $\mathcal{S}_t$, and the carbon decay rate, δ_s, representing the exit of carbon of the atmosphere into carbon sinks – oceans, biosphere – and out of the higher atmosphere. To this end, as is standard in Integrated Assessment models, I match the pulse experiment dynamics of larger IAMs – CMIP5 in this case: for a "pulse" of $100GT$ of carbon released – corresponding to 10 years of emissions – the global temperature reaches its peak between $0.20°C$ and $0.25°C$ after 10 years and then decreases slightly to stabilize around $0.17°C$ after 200 years. I follow Dietz et al. (2021), and calibrate $\chi = 0.23$ and $\delta = 0.0004$ to match these two moments, as seen in Figure 24 displayed in the appendix.

Moreover, this climate system is inherently unstable for a given trend of emissions – given once and for all by our static economic model $\mathcal{E}_t = \sum_i \mathcal{P}_i e^{(n+\bar{g})t}(\xi^f e_i^f + \xi^c e_i^c)$ with the population growth, n, and the long-term TFP growth $\bar{g}$, where $n = 0.0035$ and $\bar{g} = 0.01$ are the long-term growth rates of world population and world GDP according to forecast by the UN and World-Bank. To counteract the non-stationarity of the climate system, I follow Krusell and Smith (2022) and assume that part of the emissions $\mathcal{E}_t$ are captured and stored. I assume the exponential form:

$$\mathcal{T}_t = \chi \mathcal{S}_t \qquad \dot{\mathcal{S}}_t = \zeta_t \mathcal{E}_t - \delta_s \mathcal{S}_t \qquad \zeta_t = e^{-\zeta t}$$

and calibrate ζ to match the moment suggested in Krusell and Smith (2022): 50% is captured by 2125, and 100% by 2300 – which is $> 99.9\%$ in our model. This implies that in the Business-as-Usual scenario, global temperatures reach $\sim 5°$ by 2100 and are stabilized around 9° by 2400.[29] More optimistic scenarios for Carbon Capture and Storage (CCS) could be imagined without affecting the main result since most of the damages are discounted heavily after 2100.

Second, I calibrate the initial temperature T_{it_0} using data from Burke et al. (2015), and I display those differences across regions in Figure 5. Furthermore, we consider the linear pattern scaling $\dot{T}_{it} = \Delta_i \mathcal{T}_t$. I identify the scaling parameter in reduced-form by estimating this linear regression over the period $t = 1950-2015$ for each country and then aggregating by region i.[30] This procedure does not require extensive and granular data such at geographical characteristics, albedo, etc.

Third, to calibrate the damage function, I use the following quadratic function as in the DICE Integrated Assessment Model:

$$\hat{\mathcal{D}}^y(T_{it} - T_i^\star) = \exp\big(- \gamma^+ \mathbb{1}_{\{T_{it} > T_i^\star\}} (T_{it} - T_i^\star)^2 - \gamma^- \mathbb{1}_{\{T_{it} < T_i^\star\}} (T_{it} - T_i^\star)^2 \big)$$

with the damage parameter $\gamma^+ = 0.00340$. This value is intermediary between the value $\gamma^+ = 0.00311$ in Krusell and Smith (2022), calibrated to match Nordhaus' DICE calibration of 6.6% of loss of global GDP when temperature anomaly $\mathcal{T}_t = 5$, and the updated calibration in Barrage and Nordhaus (2024) which calibrate it at $\gamma^+ = 0.003467$. For small values, I consider $\gamma^- = 0.3\gamma^+$, following the quantification in Rudik et al. (2021), who show that the negative productivity impact of cold temperatures is much weaker than for hot temperatures.

Finally, to calibrate $T_i^\star$, I use also an intermediary assumption between the following two cases: (i) the representative agent economy, like Barrage and Nordhaus (2024), would assume $T_i^\star = T_{it_0}$, which implies that $T_{it} - T_i^\star = \Delta_i \mathcal{T}_t$: differences in damages only comes from increases in aggregate temperature. The analysis by Bilal and Känzig (2024) shows that climate damage on GDP comes in large part from the increase in global temperature, causing extreme events. In contrast, (ii) a different view in heterogeneous countries economies would set $T_i^\star = T^\star$ the same for all regions, at an "ideal" temperature, as in Krusell and Smith (2022) and Kotlikoff, Kubler, Polbin, Sachs and Scheidegger (2021). In this case, differences in climate damages come essentially

[29] Such high temperatures between 2100 and 2400 come from our static model assumption that the model and emission decisions are made once and for all. In a dynamic model, the damages over time decrease TFP and economic activity leading to an endogenous reduction in the path of emissions and temperature. In Bourany (2025), I simulate the dynamic model over time, which aligns with standard paths of future temperatures from IAMs.

[30] To control for the fact that country j has an influence on world temperature $\mathcal{T}_t = \sum_i g_i T_{it}$, I estimate the jackknife linear equation with $\mathcal{T}_{t,\neq j} = \sum_{i \neq j} g_i T_{it}$ for each j, i.e. $T_{jt} = \Delta_j \mathcal{T}_{t,\neq j}$.

from differences in initial temperatures. I take the intermediary step and assume:

$$T_i^\star = \alpha^T T^\star + (1-\alpha^T) T_{it_0}$$

where $\alpha^T = 0.5$ and $T^\star = 14.5$ is the average spring temperature of developed economies – and around the yearly average of places like California or Spain.

4.6 Heterogeneity

In this section, I summarize the different dimensions of heterogeneity included in the model, and aggregate the parameters of the calibration in appendix Table 2.

Table 1: Heterogeneity across countries

Dimension of heterogeneity	Model parameter	Matched variable from the data	Source of the data
Population	Country size $\mathcal{P}_i$	Population	UN Population Prospect
TFP/technology/institutions	Firm productivity z_i	GDP per capita (2011-PPP)	World Bank/Maddison project
Productivity in energy	Energy-augmenting productivity z_i^e	Energy cost share	SRE Energy Institute (2024)
Cost of coal energy	Cost of coal production $\mathcal{C}_i^c$	Energy mix/coal share e_i^c/e_i	SRE Energy Institute (2024)
Cost of non-carbon energy	Cost of non-carbon production $\mathcal{C}_i^r$	Energy mix/coal share e_i^r/e_i	SRE Energy Institute (2024)
Local temperature	Initial temperature T_{it_0}	Pop-weighted yearly temperature	Burke et al. (2015)
Pattern scaling	Pattern scaling Δ_i	Sensitivity of T_{it} to world $\mathcal{T}_t$	Burke et al. (2015)
Oil-gas reserves	Reserves $\mathcal{R}_i$	Proved Oil-gas reserves	SRE Energy Institute (2024)
Cost of oil-gas extraction	Slope of extraction cost $\bar{\nu}_i$	Oil-gas extracted/produced e_i^x	SRE Energy Institute (2024)
Cost of oil-gas extraction	Curvature of extraction cost ν_i	Profit π_i^f / energy rent	World Bank / WDI
Trade costs	Distance iceberg costs τ_{ij}	Geographical distance $\tau_{ij} = d_{ij}^\beta$	CEPII Conte et al. (2022)
Armington preferences	CES preferences a_{ij}	Trade flows	CEPII Conte et al. (2022)

5 Optimal policy benchmarks without participation constraints

I provide two benchmarks for optimal policy when participation is exogenous. First, I consider a global social planner policy that maximizes aggregate welfare, representing the cooperative allocation. Second, in the non-cooperative Nash-equilibrium, each country implements its unilaterally optimal policy.

5.1 Global Climate Policy with cooperation

The cooperative policy depends on the availability of redistribution instruments. In the First-Best, with unlimited instruments, and in particular lump-sum transfers, the optimal tax is the Social Cost of Carbon (SCC), a measure of the marginal cost of climate change. Without transfers, the optimal tax needs to account for inequality across countries and trade leakage effects. Accounting for inequality and lack of redistribution, the optimal tax is lower than the SCC. The main lessons from this analysis are also described in detail in Bourany (2025) where I develop this argument in a large class of Integrated Assessment Models.

5.1.1 First Best allocation with unlimited instruments

With unlimited instruments, the social planner uses lump-sum transfers to redistribute across countries and offset the negative effects of climate change and carbon taxation. In this context, the optimal tax is the standard Pigouvian tax, which is the Social Cost of Carbon.

Consider a planner that maximizes global welfare by choosing the allocation, i.e. consumption, inputs for energy production, and energy demand: $\boldsymbol{x} = \{c_{ij}, x^{\ell}_{ij}, e^{\ell}_i\}$, for $\ell \in \{f, c, r\}$

$$\mathcal{W} = \max_{\{c_i, e_i, e^x_i, \dots\}_i} \sum_{i \in \mathbb{I}} \mathcal{P}_i \omega_i \ u(c_j) = \textstyle\sum_{\mathbb{I}} \mathcal{P}_i \, \omega_i \, \mathcal{U}_i \ , \tag{20}$$

subject to the market clearing of energy and goods.

The derivation is detailed in Appendix C.1. Taking the First-Order Conditions, we learn several insights on the conduct of the optimal policy. The first lesson is that the planner equalizes marginal utilities through the following conditions:

$$\frac{\omega_i u'(c_i) \mathcal{D}^u_i(\mathcal{E})}{\mathbb{P}_i} = \frac{\omega_j u'(c_j) \mathcal{D}^u_j(\mathcal{E})}{\mathbb{P}_j} = \overline{\lambda} \qquad\qquad \forall \ i, j \in \mathbb{I} \ .$$

This implies arbitrary large redistribution, using lump-sum transfers, such that:

$$\begin{aligned} c_i &= u'^{-1}\left(\frac{\overline{\lambda} \, \mathbb{P}_i}{\omega_i \, \mathcal{D}^u_i(\mathcal{E})}\right), \ \forall i \in \mathbb{I} \ , \\ \mathrm{t}^{ls}_i &= w_i \ell_i + \pi^f_i - c_i \mathbb{P}_i \ . \end{aligned}$$

In this case, the planner would like to increase the consumption of countries with high Pareto weights ω_i and offset differences in $\mathbb{P}_i$ and $\mathcal{D}^u_i(\mathcal{E})$. Importantly, the transfers t^{ls}_i are designed such that the marginal utilities of consumption are equalized. This implies redistribution, as $\mathrm{t}^{ls}_i < 0$ for some countries and $\mathrm{t}^{ls}_i > 0$ for some other countries.

Second, the Social Cost of Carbon, defined as the ratio of the marginal value of emissions over the marginal utility of consumption, can then be reformulated as:

$$\begin{aligned} SCC = -\frac{\frac{\partial \mathcal{W}}{\partial \mathcal{E}}}{\frac{\partial \mathcal{W}}{c_i}} = \frac{\phi^{\mathcal{E}}}{\overline{\lambda}} &= -\sum_{\mathbb{I}} \mathcal{P}_i \omega_i \Big[\frac{u(c_i)}{u'(c_i)} \mathbb{P}_i \overline{\mathcal{D}}^{u\prime}(\mathcal{E}) + \mathcal{D}^{y\prime}_i(\mathcal{E}) z_i F(\ell_i, e_i) \mathrm{p}_i\Big] > 0 \\ &= \sum_{\mathbb{I}} \mathcal{P}_i \omega_i \Big[\gamma^c c_i \mathbb{P}_i + \gamma^y y_i \mathrm{p}_i\Big] \widetilde{\mathcal{D}}_i(\mathcal{E}) > 0 \end{aligned}$$

where $\phi^{\mathcal{E}}$ represents the welfare cost of one additional ton of carbon, which corresponds to the multiplier of the constraint $\mathcal{E} = \sum_i \xi^f e^f_i + \xi^c e^c_i$, and $\overline{\lambda}$ the average marginal utility of consumption – or marginal value of wealth. As usual in this class of integrated assessment model, e.g. Golosov et al. (2014), the marginal cost of climate change scales with the damage parameters γ^c, γ^y, output $y_i \mathrm{p}_i$ and consumption $c_i \mathbb{P}_i$, and a damage function $\widetilde{\mathcal{D}}_i(\mathcal{E})$, depending on the climate system.[31]

[31] Specifically, the function $\widetilde{\mathcal{D}}_i(\mathcal{E}) = \bar{\rho} \int_{t_0}^{\infty} \int_t^{\infty} e^{-\bar{\rho} s} e^{-\delta_s (s-t)} (T_{is} - T^{\star}_i) ds \, dt$ depends on the underlying climate system, and how a ton of CO_2 emitted in time t affects temperatures in all future periods s.

In that context, the optimal tax is simply the Social Cost of Carbon (SCC)

$$\mathrm{t}^{\varepsilon} = SCC = \frac{\phi^{\mathcal{E}}}{\overline{\lambda}} = \sum_{\mathbb{I}} \mathcal{P}_i \omega_i LCC_i$$

This recovers the result in Golosov et al. (2014) that prevails in representative agents economies: absent redistributive motive, the optimal tax is the Pigouvian tax.

In the next section, we see how this result changes when transfers and other instruments, like tariffs or subsidies, are constrained, preventing the planner from performing this redistribution.

5.1.2 Second-Best Ramsey problem without transfers

When the social planner does not have access to cross-country transfers, the optimal carbon tax needs to account for redistributive effects. The carbon tax needs to be corrected for (i) the heterogeneous effects of climate change and inequality across countries, (ii) the redistributive effects on energy markets, (iii) the distortion of demand, (iv) the leakage effects of trade.

Consider the welfare maximization problem, as in eq. (20), where the planner chooses a *single uniform carbon tax*, t^{ε}, rebates the revenue of the carbon tax locally $\mathrm{t}_i^{ls} = \mathrm{t}^{\varepsilon}(\xi^f e_i^f + \xi^c e_i^c)$. In this Ramsey problem, with restricted instruments – a single carbon tax – the planner chooses an allocation subject to the constraints of the competitive equilibrium. Given the policy, the agents (households and firms) optimize and market clear for a set of prices. As a result, the planner chooses $\boldsymbol{x} = \{c_{ij}, x_{ij}^{\ell}, e_i^{\ell}\}$, i.e. the traded good for consumption, c_{ij}, for energy inputs for the production of in fossil, x_{ij}^f, coal, x_{ij}^c, and non-carbon, x_{ij}^r, the energy demand, in fossil e_i^f, coal, e_i^c, and non-carbon, e_i^r, as well as the carbon tax, t^{ε}. However, the Ramsey allocation, $\boldsymbol{x}$, and prices $\boldsymbol{p} = \{\mathrm{p}_i, q^f, q_i^c, q_i^r\}_i$ should be a competitive equilibrium: in that case, the planner restricts controls to respect the individual optimality conditions.

In the Primal approach of Public Finance, these optimality conditions are internalized by the planner as additional constraints in its welfare maximization, through a large array of multipliers. The multiplier for each of these constraints summarizes a different effect of the taxation. For example, (i) the multiplier for the budget constraint eq. (2) of the household, λ_i, represents a redistribution motive for the planner. Moreover, (ii) the multipliers v_i^f for the fossil fuel demand e_i^f optimality condition represent the distortionary effect of the taxation on fossil-fuel choice, and similarly (iii) the multipliers for the fossil fuel production choice e_i^x and the multiplier for the energy market clearing cab be combined and represent the redistribution that arises in general equilibrium. It represents how carbon taxation affects the supply of oil and gas, which affects energy rent. Finally, (iv) the multiplier for the market clearing μ_i for the good produced from country i summarizes the equilibrium effects on good markets, which are related to the carbon trade leakage in the context of carbon policy. As a result, the optimal carbon tax t^{ε} corrects for the climate externality, but also need to account for these distributional equilibrium motives. The complete set of optimality conditions for the planner is technical, and I restrict its exposition to Appendix C.2.

The optimal tax formula for carbon – in the case of fossil for example – can be summarized as:

$$\xi^f \mathrm{t}^\varepsilon = \xi^f \underbrace{\textstyle\sum_i \widehat{\lambda}_i \, LCC_i}_{=SCC} + \textstyle\sum_i \widehat{\lambda}_i \, \text{Supply Redistrib}^\circ_i + \sum_i \widehat{\lambda}_i \, \text{Demand Distort}^\circ_i - \sum_i \text{Trade Redistrib}^\circ_i \ ,$$

$$\text{with} \qquad \widehat{\lambda}_i = \frac{\omega_i \mathcal{P}_i \lambda_i}{\sum_i \omega_i \mathcal{P}_i \lambda_i} = \frac{\omega_i \mathcal{P}_i u'(c_i)}{\sum_i \omega_i \mathcal{P}_i u'(c_i)} \ , \qquad LCC_i = \frac{\phi^\varepsilon_i}{\lambda_i} \ .$$

The "social welfare weights" $\widehat{\lambda}_i = \frac{\omega_i \mathcal{P}_i \lambda_i}{\bar{\lambda}} = \frac{\omega_i \mathcal{P}_i \lambda_i}{\sum_i \omega_i \mathcal{P}_i \lambda_i}$ are the rescaled multipliers for the budget constraint λ_i. As a result, the Social Cost of Carbon: $SCC = \sum_i \widehat{\lambda}_i \, LCC_i$ is the weighted sum of the Local Costs of Carbon LCC_i adjusted for inequalities across regions. It measures the marginal cost of climate cost as valued by country i, and – in the second-best – accounts for all the redistributive and distortive effects caused by global warming.[32] Moreover, the multipliers for the FOC demand $\widehat{\upsilon}_i = \frac{\omega_i \mathcal{P}_i \upsilon_i}{\bar{\lambda}}$, for market clearing for good: $\widehat{\mu}_i = \frac{\omega_i \mathcal{P}_i \mu_i}{\bar{\lambda}}$ are rescaled in the optimal tax formula. With the aggregate inverse supply elasticity for fossil $\bar{\nu} = \big(\sum_i \lambda^x_i \nu_i^{-1}\big)^{-1}$, the optimal carbon tax can be unpacked as:

$$\xi^f \mathrm{t}^\varepsilon = \xi^f \mathcal{P} \sum_i \widehat{\lambda}_i LCC_i + q^f \mathcal{P} \frac{\bar{\nu}}{E^f} \sum_i \widehat{\lambda}_i \big(e^f_i - \frac{e^x_i}{\mathcal{P}_i}\big) - (q^f + \xi^f \mathrm{t}^\varepsilon) \sum_i \widehat{\mu}_i - \sum_i \widehat{\upsilon}^f_i \quad ,$$

$$\xi^f \mathrm{t}^\varepsilon = \xi^f \underbrace{\mathcal{P}\mathbb{E}_i\big[LCC_i\big] + \mathcal{P}\mathbb{C}\mathrm{ov}_i\big(\widehat{\lambda}_i, LCC_i\big)}_{=\text{Social Cost of Carbon}} + \underbrace{q^f \mathcal{P} \frac{\bar{\nu}}{E^f} \mathbb{C}\mathrm{ov}_i\big(\widehat{\lambda}_i, e^f_i - \frac{e^x_i}{\mathcal{P}_i}\big)}_{\substack{E^f \text{ supply} \\ \text{redistribut}^\circ}} - \underbrace{\mathbb{E}_i\big[\widehat{\upsilon}^f_i\big]}_{\substack{e^{\ell'}_i \text{ demand} \\ \text{distort}^\circ}} - (q^f + \xi^f \mathrm{t}^\varepsilon) \underbrace{\mathbb{E}_i\big[\widehat{\mu}_i\big]}_{\substack{y_i \text{Trade} \\ \text{redistribut}^\circ}}$$

This implies that the carbon tax can be different from the Social Cost of Carbon when the planner has redistributive motives. In Appendix C.2, I further develop the distortion terms $\widehat{\upsilon}_i$ to understand how this distortion can be expressed as a function of the elasticities of demand σ^y and σ^e for fossil fuels. This general logic is described in more detail in the companion paper Bourany (2025). It matters quantitatively for our model, as I show in the following figure.

In Figure 6, I display the Social Cost of Carbon (SCC) and optimal carbon tax in the different equilibria we studied. In the competitive – Business-as-Usual – equilibrium, the SCC represents the endogenous cost of climate change and is thus very large due to climate inaction, i.e. around \$200 per ton of CO_2. However, in the First-Best, the planner lowers both the climate damages ϕ^ε, and, thanks to large redistribution, the marginal value of wealth $\bar{\lambda}$ for the "average household". Therefore, the Social Cost of Carbon $SCC = \phi^\varepsilon / \bar{\lambda}$ is also high, and the optimal carbon tax equals the SCC at $\mathrm{t}^\varepsilon = SCC = \$200/tCO_2$.

In the Second-Best allocation, the social planner cannot redistribute easily and therefore accounts for the redistributive effects in the choice of the carbon tax. The average household is

[32] The Local Cost of Carbon in the Second Best is slightly more involved than in the First Best as it now account for additional distortions and general equilibrium effects (see Appendix C.2 for details):

$$LCC_i = \frac{\phi^\varepsilon_i}{\lambda_i} = \frac{u(c_i)}{u'(c_i)} \mathbb{P}_i \mathcal{D}^{u\prime}(\mathcal{E}) + \mathcal{D}^{y\prime}_i(\mathcal{E}) z_i F(\ell_i, e_i) \mathrm{p}_i \big(1 + \frac{\mu_i}{\lambda_i}\big) + \mathcal{D}^{y\prime}_i(\mathcal{E}) \mathrm{p}_i \sum_{\ell'} \frac{\upsilon^{\ell'}_i}{\lambda_i} MPe^{\ell'}_i$$

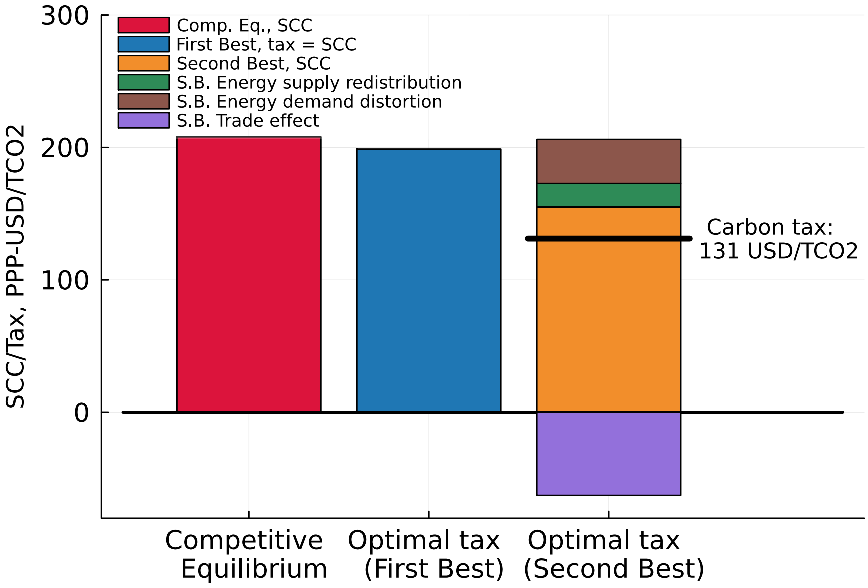


Figure 6: Social Cost of Carbon and optimal carbon taxation

now "poorer", increasing the aggregate weight $\bar{\lambda}$, lowering the Social Cost of Carbon from \$200 to approximately \$155 per ton of CO_2. Moreover, redistributive effects are accounted for in the level of the carbon tax: first, the countries that consumes the largest fossil-energy are the richest/least affected by climate change. Second, the oil and gas importers are also more inclined to have higher social welfare weights $\omega_i u'(c_i)$. For these two reasons, the carbon tax should be increased. However, the countries the most affected by climate change and poorest, are also experiencing large trade reallocation. This is accounted for in the optimal policy, which is now lower because of this leakage effects. For all these reasons, the optimal carbon tax is now set at \$131 per ton of CO_2.

Winners and losers from cooperative carbon taxation

This optimal second-best carbon tax, despite accounting for redistributive effects, still has heterogeneous impact across countries. In Figure 7, I display the welfare change from this policy $\mathcal{U}_i(\mathbb{I}, \mathrm{t}^{\varepsilon})/\mathcal{U}_i(\mathbb{I}, 0)$ in consumption equivalent, in comparison to the competitive equilibrium. We first observe that most regions benefit greatly from cooperation, with an aggregate effect of 7% on global welfare. However, this hides large heterogeneity across countries.

The biggest winners are, without contest, the countries that are the most affected by climate change: South-East Asia, Africa, South Asia, and warm Middle-Eastern and European countries, which gain between 7% and 20% of consumption equivalent. Many of them see a reduction in the cost of their energy imports due to the reverse leakage effects – as indeed, phasing out fossil-fuel cause a reduction in the fossil price q^f. However, countries that consume a large share of coal, like China, India, South-Africa. and Australia are not gaining as much due to the large distortive effect of carbon taxation. Finally, large fossil fuels – oil and gas – exporters like Russia, Canada, Nigeria, Saudi Arabia, Iran or other Gulf counries, all lose from carbon taxation because they see a reduction of their energy rents. In addition, Russia and former Soviet countries are also cold countries that do not gain anything from slowing down climate change.

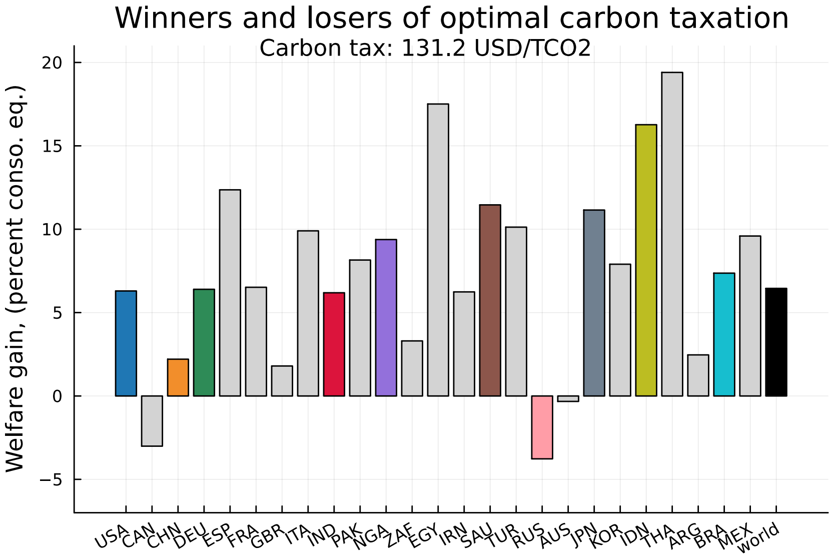


Figure 7: Welfare gains across countries – Second-Best equilibrium

5.2 Unilateral policies: Nash-equilibrium

In this section, I consider the case where countries act non-cooperatively. In such cases, the optimal carbon policy does not account for the externality imposed on other countries. Moreover, each country might use energy policy and trade tariffs strategically for terms-of-trade manipulation. As a result, for some countries, the carbon tax may become a subsidy. This exercise extends the approach of Ossa (2014), Farrokhi and Lashkaripour (2024) or Kortum and Weisbach (2021) who all solve for optimal unilateral tariffs. In the setting I consider, redistributive effects, general-equilibrium on energy markets and terms-of-trade manipulation yield results that contrast with the above references. I describe the complete setting in Appendix C.3

In this exercise, I consider a Local Social Planner who chooses a local carbon tax $\mathrm{t}_i^{\varepsilon}$ and trade tariffs t_{ij}^{b} to maximize household welfare:

$$\mathcal{V}_i = \max_{\mathrm{t}^{\varepsilon},\mathrm{t}^{b}} \mathcal{P}_i \mathcal{U}_i$$

subject to world equilibrium constraints and taking the policies of the other countries as given – as a Nash Equilibrium. It redistribute the carbon tax and tariffs revenues to the local household: $\mathrm{t}_i^{ls} = \mathrm{t}_i^{\varepsilon}(\xi^f e_i^f + \xi^c e_i^c) + \sum_k \mathrm{t}_{ik}^{b} c_{ik} \tau_{ik} \mathrm{p}_k$.

In the spirit of the previous section, using the Primal approach, we see that the planner i chose the policy instruments $\mathrm{t}_i^{\varepsilon}, \mathrm{t}_{ij}^{b}$ internalizing a wide array of market clearing and optimality conditions. More specifically, the planner accounts for the market clearing for good j as household and energy firms in country i imports this variety, e.g. as consumption c_{ij}. I denote this shadow value for good j eq. (11) as $\mu_j^{(i)}$. If $\mu_j^{(i)} > 0$, planner i would like to expand j's supply – relaxing its market clearing and lowering its price. In contrast, we usually obtain $\mu_i^{(i)} < 0$: the planner would like to reduces its own i-supply, to increase its price and manipulate terms of trade $T.o.T_i = \mathrm{p}_i/\mathrm{p}_j$, $\forall j \neq i$.

Similarly, the planner accounts for the fossil energy market clearing eq. (8), which is global, with multiplier $\mu^{f(i)}$. In addition, the planner internalizes the country-level constraints, such as budget with multiplier $\lambda_i = u'(c_i)\mathcal{D}_i^u(\mathcal{E})/\mathbb{P}_i$, or the optimality condition for energy demand, energy supply and good imports. For ease of exposition, I display the result where the only energy input is oil-gas $e_i = e_i^f$. The general case for three energy sources is detailed in Appendix C.3.

One key result in this setting is the absence of local distortions for unilaterally optimal policy. Several distortionary effects that appeared in the previous Second-Best framework disappeared. It is, for example, the case for the optimality conditions for good imports, c_{ij}, or energy demand e_i, which are not distorted by the planner's choice. As a result, the planner's decision aligns with the household and firms in country i, because of the existence of tariff and country-specific tax.

Optimal tariffs

The optimal tariffs are chosen to manipulate terms-of-trade, namely:

$$\mathrm{t}_{ij}^b = \frac{\mu_j^{(i)}}{\lambda_i}$$

where $\mu_j^{(i)}$ are the Lagrange multiplier for good j's market clearing as accounted by planner i, and λ_i the marginal utility of consumption. Tariffs increase when the planner seeks to decrease demand for good j. This is especially the case when the household in country i is richer: the redistributive effect of tariffs are amplified when $c_i \to \infty$, $\lambda_i \to 0$ and thus $\mathrm{t}_{ij}^b \to \infty$. Unfortunately, with the primal approach, these multipliers $\mu_j^{(i)}$ typically do not have closed forms expressions in the Armington trade model.

Local Social Cost of Carbon

For designing climate policy, we need to evaluate the impact of climate on country i welfare. For that, we can summarize this welfare cost as the Local Cost of Carbon, as in Cruz and Rossi-Hansberg (2022):

$$LCC_i = -\frac{\frac{\partial \mathcal{V}_i}{\partial \mathcal{E}}}{\frac{\partial \mathcal{V}_i}{\partial c_i}} = \frac{\phi_i^{\varepsilon}}{\lambda_i}$$

where ϕ_i^{ε} is the multiplier for carbon dynamics – or shadow cost of increasing emissions by 1 ton of CO_2. In the unilaterally optimal allocation, this cost of climate change is given by:

$$\begin{aligned} LCC_i &= \mathcal{P}_i \Big[\frac{u(c_i)}{u'(c_i)} \mathbb{P}_i \overline{\mathcal{D}}_i^{u\prime}(\mathcal{E}) + \mathcal{D}_i^{y\prime}(\mathcal{E}) z_i F(\ell_i, e_i) \mathrm{p}_i + \sum_j \frac{\mu_j^{(i)}}{\lambda_i} \mathcal{D}_j^{y\prime}(\mathcal{E}) z_j F(\ell_j, e_j) \mathrm{p}_j \Big] , \\ &= \mathcal{P}_i \big[\mathbb{P}_i c_i \gamma^c + y_i \mathrm{p}_i \gamma^y \big] \widetilde{\mathcal{D}}_i(\mathcal{E}) + \mathcal{P}_i \sum_j \frac{\mu_j^{(i)}}{\lambda_i} \gamma^y y_j \mathrm{p}_j \widetilde{\mathcal{D}}_j(\mathcal{E}) , \end{aligned}$$

where $\widetilde{\mathcal{D}}_i(\mathcal{E})$ are damage functions that only depend on the climate system as shown above – e.g. pattern scaling Δ_i and climate sensitivity and decay χ and δ – and the parameters γ^c and γ^y affect utility and production damages. We note that in addition to considering the direct damage on country i household utility and output, the planner in i also internalizes the damage that climate has on the production of trade partner y_j, through multiplier $\mu_j^{(i)}$ again, and this impact is even larger for rich, low λ_i-countries. This introduces a novel channel through which local costs of climate change can become correlated due to international trade linkages. Another argument made in Dingel et al. (2019) relates the spatial correlation of the costs of climate change, the gains from trade, and welfare change. Here, I also show how this policy relevant local costs of carbon can be correlated, even in non-cooperative Nash-equilibria, potentially providing a motive for coordinated climate actions.

Unilaterally optimal carbon tax – or subsidy

Finally, I derive the unilaterally optimal carbon tax – here exposed in the case where energy is entirely consumed in fossil-fuel (oil and gas).

$$\xi^f \mathrm{t}_i^\varepsilon = \underbrace{\xi^f LCC_i}_{\substack{\text{Pigouvian}\\ \text{motive}}} \underbrace{- q^f \frac{\mu_i^{(i)}}{\lambda_i}}_{\substack{\text{terms-of-trade}\\ \text{manipulation}}} \underbrace{+ q^f \nu_i \frac{\mathcal{P}_i e_i^f - e_i^x}{e_i^x}}_{\substack{\text{energy supply}\\ \text{redistribution}}} .$$

In the Nash equilibrium of this model, the carbon tax is the sum of three terms: First, the country i's planner internalizes the effect its emissions have on the welfare of its own country. However, it does not internalize the rest of the world: this is a symptom of free-riding as country i does not account for the impact of its emissions on other countries' welfare.

Second, it features a terms-of-trade manipulation term: if country i would like to reduce its demand, acting like a monopolist on its variety i, for example, in the case where $\mu_i^{(i)} < 0$, then it would tax carbon to lower production. In the opposite case, if country i looks after expanding its supply and lowering its prices, when $\mu_i^{(i)} > 0$, then the planner would lower taxation of energy, providing a motive for subsidy.

Third, the taxation of fossil fuels has an impact on the international energy market, creating a redistribution term. This is positive for net importers and negative for exporters. Energy exporters would like to subsidize energy to increase demand, in an attempt to raise the equilibrium price q^f and benefit from better terms-of-trade. This is weighted by the country i production e_i^x and inverse elasticity ν_i: a more inelastic supply, with large ν_i, would amplify this effect. Note that this is the same logic as the global social planner of Section 5.1.2, at a local level.

To sum up, if the terms-of-trade manipulation motive $\mu_i^{(i)} > 0$ is large enough or if the energy-supply redistribution term is negative, for example for oil-gas exporters, and if the local cost of carbon LCC_i is small, the optimal carbon tax can become a *subsidy.*

6 Optimal Climate agreement

I now turn to the main result of this paper. The optimal design of climate agreement is a climate club that consists of all the countries at the exception of major fossil fuel producers: Russia, Saudi Arabia, Nigeria and Iran. Members of the club impose a $114 carbon tax per ton of CO_2 and a 48% tariff on goods from non-members. The intuition behind this result can be summarized by the tradeoff between the distortionary effect of carbon taxes and the cost of tariffs, which relate to the gains from trade. For some countries, like developing economies with a fossil-intensive energy mix, the first outweighs the second, implying that they would not participate in a climate agreement unless the tax is decreased from \$131 to \$114. This encourages the participation of several Middle-Eastern and South Asian countries, but the optimal agreement does not include the entire world. Indeed, lowering the tax so low to incentivize fossil-fuel exporters to participate would compromise climate action and reduce world's welfare.

I first provide details on the trade-off behind country participation, both numerically and theoretically, using first-order decomposition of the model. I then present the main result of the optimal climate club.

6.1 Trade-off: distortionary effects of carbon taxation vs. gains from trade

The following two effects influence the design of the agreement. First, the distortionary effect of the carbon tax differs across countries, and some countries – poor, closed to trade or cold countries and fossil-fuel producers have very large gains from free-riding. Second, the cost of tariffs from trade partners, related to the gains from trade, also differ across countries.

In the Figure 8, I present an experiment where all countries $j \in \mathbb{I}$ implement the optimal level of fossil-fuel taxation $\mathrm{t}^{\varepsilon} = \$131/tCO_2$, except for country i, which deviates from that policy, setting the tax to zero. In this experiment, other countries j do not impose retaliatory tariffs on country i, and continue to implement the optimal policy. For each country i, I plot, in consumption-equivalent units, the welfare gain of such "deviation" compared to the case where the country i stays in this "agreement". This represents the gains from free-riding, while the Rest of the World, or equivalently, is the cost of the distortionary taxation of carbon and fossil fuels.

We first see that such gains from "deviating" range from 1.5% – for Europe, which uses more renewable energy and less coal – to close to 4% for Russia, former soviet countries, and Middle-Eastern countries, whose economies are relying on oil and gas both for good production and energy exports. These distortionary costs are also relatively high for developed economies like South Asia, Sub-Saharan Africa, and Latin America, for the reason that energy, and fossil and coal in particular, are necessary inputs in production and the welfare cost scale with marginal utility of consumption $u'(c_i)$. I provide a welfare decomposition in Section 6.2 to show the sources of these welfare costs and how it differ across regions.

Now, let us compare this cost of carbon taxation to the cost of trade tariffs. In Figure 9, I measure the welfare costs of tariffs in the following experiment: all countries $j \neq i$ impose a very large tariff – 500% – on country i. For each country i, I display the welfare loss, in consumption equivalent percentage changes in the figure. This is a good representation of the upper bound on welfare cost of tariffs – as those welfare costs are virtually identical for higher values of tariffs. This is closely related to the cost of autarky, or gains-from-trade which are bounded and relatively small in standard trade models, c.f. Arkolakis et al. (2012).[33]

[33] In this experiment, autarky is when both countries j impose large tariffs on i and i imposes tariffs on countries j as well. I consider one-sided tariffs, which is closer to the policy implemented in our climate club.

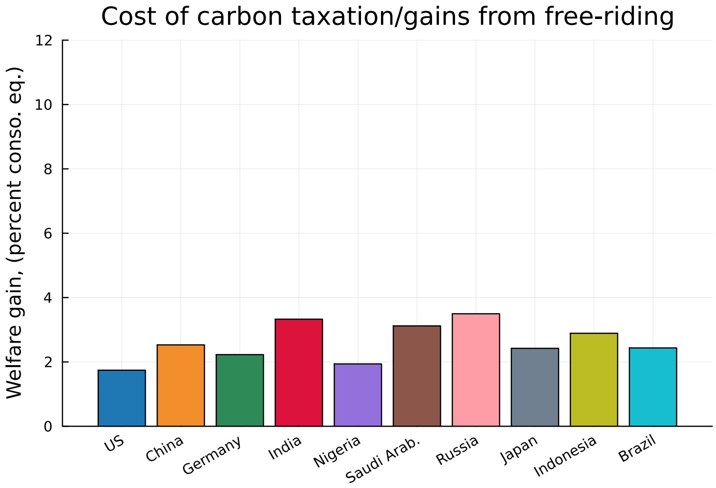


Figure 8: Welfare gain for country i of unilaterally deviating from a world agreement setting a \$131 carbon tax

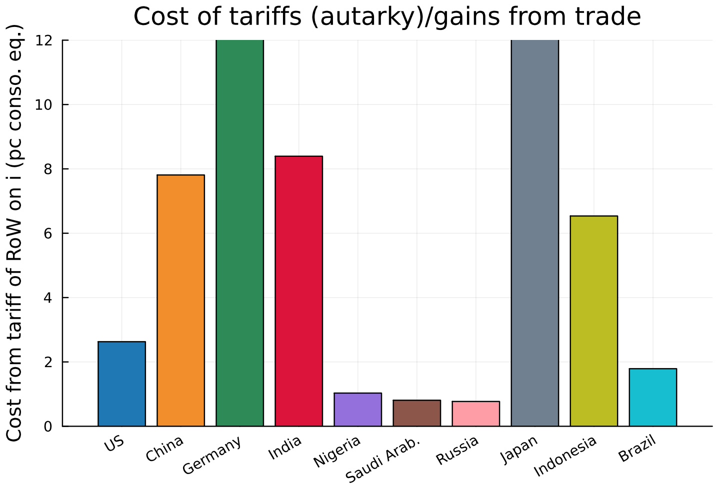


Figure 9: Losses for country i of country $j \neq i$ imposing a 500% tariff on country i

Asian countries – like Japan and Korea, China and South-East Asian economies – and Europe have the most to lose to be in subject to tariffs – above 5% welfare equivalent consumption losses. This is in part due to the large trade shares of these countries with each other and with Europe. In comparison, countries like the US, Middle-East and Russia, which are more closed to international trade, suffer less from tariffs, which change their willingness to join a climate club. This difference is also reinforced because the tariffs hurt the terms of trade of energy importers, which now have to pay a larger fossil-fuel bill in real terms. This explains

6.2 Welfare decomposition

To understand the mechanisms through which climate change, carbon taxation, and tariffs affect welfare, I provide a first-order approximation of welfare to shed light on different mechanisms. This welfare decomposition is described in thorough detail in the companion paper Bourany and Rosenthal-Kay (2025), and in Appendix D.

There we compute the change in welfare, linearizing the model around the competitive equilibrium where $\mathrm{t}^\varepsilon = \bar{\mathrm{t}}^\varepsilon = 0$ and $\mathrm{t}^b_{ij} = \bar{\mathrm{t}}^b_{ij} = 0$, where policies are identical to the "status-quo". I start from a climate agreement $\mathcal{J}$ of J countries. Those countries are indifferent between being in the club or not, since the policy $(\mathrm{t}^\varepsilon_i, \mathrm{t}^b_{ij}) = (0, 0)$ does not affect the equilibrium. I then consider a log-linear perturbation where those policy instruments are increased by a small amount, $d\mathrm{t}^\varepsilon_i$ and $d\mathrm{t}^b_{ij}$ respectively for the club members $i \in \mathcal{J}$. In that context, we can do a decomposition of welfare as a function of four channels: (i) the direct impact on climate damage – productivity $\mathcal{D}^y_i$ and utility $\mathcal{D}^u_i$ – (ii) the terms-of-trade effect through change in p_i and $\mathbb{P}_i$, (iii) the energy costs for the goods firms, through prices q^f, q^c_i and q^r_i, and (iv) the energy rents of fossil fuel producers π^f_i. We can map those welfare channels as function of "sufficient statistics" which only depends on observables data moments – like energy mix, e.g. e^f_i/e_i, energy rent π^f_i as share of GDP, and trade shares s_{ij} – and elasticities that can be estimated, like the energy supply elasticities, e.g. for fossil ν_i, or the climate damage parameter γ^y_i.

6.3 Optimal climate agreement

In this section, I describe the design of the optimal climate agreement. The climate club that maximizes the world's welfare is a large coalition with all the country at the exception of several fossil producers. Moreover, the carbon tax for members of the club is lowered below \$115, and tariffs are set at a moderate rate of 48% for non-members. This outcome balances the intensive margin-extensive margin tradeoff of this policy design.

At the intensive margin, increasing the carbon tax reduces fossil-fuel use and emissions for the countries participating in the climate agreement. As a result, aggregate welfare is increased until the optimal carbon tax $t^{\varepsilon} = \$131$ is reached. However, at the extensive margin, a higher tax reduces participation as free-riding incentives are reinforced with the cost of taxation. If the tax becomes too high, individual countries deviate and leave the agreement, which raises world's emissions. In Figure 10, I show this phenomenon, where I plot the maximum participation that can be achieved depending on the choice of the levels of carbon tax for club members on the y-axis and the tariffs that are imposed on non-members on the x-axis.

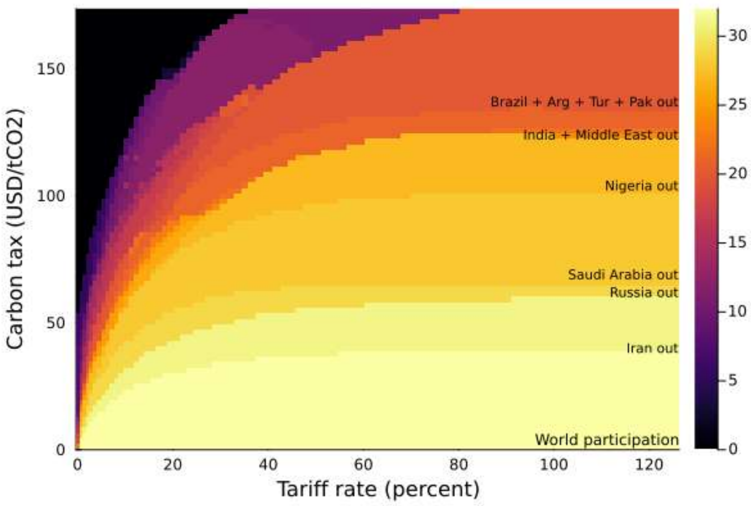


Figure 10: Participation: Intensive and extensive margin trade-off for agreement design: t^{ε} on y-axis, t^{b} on x-axis

For tax under \$50, the cost of carbon abatement is low, making it relatively costless for countries to participate in the agreement. For higher taxes, we observe that the first regions to deviate are Iran, Russia, and Saudi Arabia. For larger tax, other developing countries, like Nigeria, India, and South American economies would also exit the agreement. These decisions originate from the tradeoff explained in Section 6.1. Indeed, those countries have a high cost of distortionary carbon taxation, either because they are producers of oil and gas, like Russia and Gulf countries, or because they consume a significant part of their energy mix in coal or oil-gas, like India, Africa and South American countries. This compares to the cost of tariffs, which are relatively small for those regions for tariffs below 150%.

Another lesson from this analysis is that trade policy is a key strategic instrument to deter free-riding. Indeed, absent tariff retaliation, with $t^{b} < 5\%$, the gain from unilateral deviation prevail over the cost of climate action, and no carbon tax above $t^\varepsilon > \$30$ could be implemented for a large enough set of countries. This result is discussed in Nordhaus (2015) and I recover this effect in this quantitative model. However, if moderate tariffs spur participation for low carbon taxes, this incentive effect vanishes quickly as the carbon tax increases. Since the gains from trade are bounded – and small for some countries like Middle-East and Russia – there is a limit to what carbon policy can achieve.

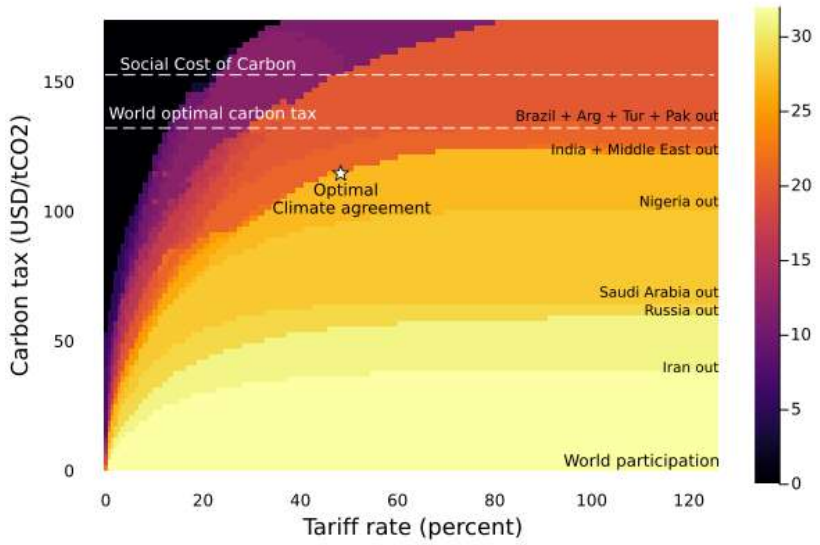


Figure 11: Optimal climate agreement:
participation as function of t^ε on y-axis and t^b on x-axis

We turn now to the design of optimal climate agreements, and the choice of carbon tax and tariffs. We note that the optimal Second-Best policy with a large carbon tax $t^\varepsilon = \$131$ and complete participation is not achievable in a climate club. As shown in the dotted lines in Figure 11, it corresponds to an area where many countries in South-Asia, Middle-East, and Russia would all exit the agreement. As a result, the optimal agreement that would maximize welfare is such that the carbon tax is lowered from \$131 to \$114: this incentivizes the participation of several South Asian and South-American countries.

It is optimal to leave Russia outside the agreement. Reducing the carbon tax to accommodate Russia's participation to the agreement necessitates a large fall in climate effort. A decrease of the tax from \$98 to around \$50 increases emissions of the entire world. This compromises the implementation of effective solutions for global warming, and would lower aggregate welfare.

This optimal climate agreement realizes close to 92% of the welfare gains attained in the optimal-policy without endogenous participation, as seen in Figure 7 of Section 5.1.2. In Figure 12, I plot global welfare for the different carbon taxes and tariffs (t^ε, t^b). in consumption equivalent relative to welfare in the competitive equilibrium $(t^\varepsilon, t^b) = (0, 0)$. Welfare change non-monotonically in the carbon tax as emissions and temperature are reduced. However, when participation declines and countries exit the club, the deviating countries go back to their status-quo policies, raising their emissions, which decreases discontinuously global welfare. The optimal agreement achieves almost

5% of consumption-equivalent welfare gains, close to 90% of the welfare gains of the Second-Best where all countries are participating exogenously absent free-riding.

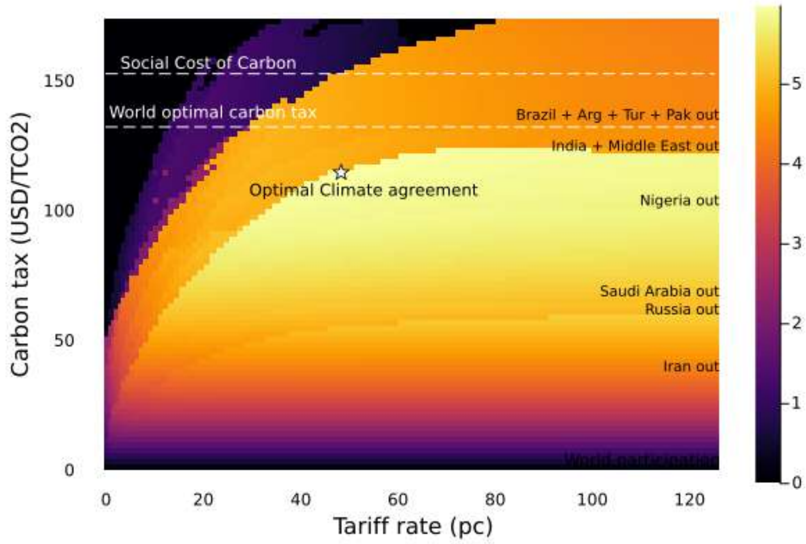


Figure 12: Welfare over different climate agreements
t^{ε} on y-axis, t^{b} on x-axis

Note that tariffs increase have a very moderate impact on welfare. In fact, they have a strong impact on individual countries' utility if they are outside the club. However, since such countries – e.g. Russia, Middle-East, and South-Asia – do not have much weight in the welfare criterion we used, especially considering the Negishi weights in Figure 2, this has a limited influence on global welfare. It mainly has a strong influence on participation through the impact on the country's outside options and welfare.

Changing the level of the carbon tax is fundamental for participation and the optimal design of the agreement, creating a *Laffer curve* for emissions and welfare. Figure 14 and Figure 13 plot the change in global welfare and carbon emissions varying the carbon tax, and keeping the tariffs fixed at the optimum, $t^{b} = 48\%$. Raising the carbon tax reduces emissions and improves welfare up to the point where participation declines. It is therefore optimal to "share the burden" of the carbon tax on a larger set of countries. In the optimal agreement, where all the countries in the world except Russia, Saudi Arabia, Nigeria, and Iran are included, one can reduce emissions from 35 to below $22\,GtCO_2$ per year, a decline of 38% compared to the competitive – Business-as-Usual – equilibrium.

Note that the optimal climate agreement achieves almost the same emission reduction as in the case with a higher carbon tax $t^{\varepsilon} = SCC = \$155/tCO_2$ – corresponding to the Social Cost of Carbon – and fewer countries. However, this implies that the very affected countries – Russia, Middle-East, South-Asia – all exit the agreements. The remaining countries, which are the developed economies, Europe, North America, and East Asia, all have to bear a much higher cost of taxation. This agreement is still stable due to the enforcement power of tariffs, but the negative welfare impact of taxation for those countries is now much larger. The difference in welfare between those two cases is sizable: the optimal climate agreement achieves a 6% welfare

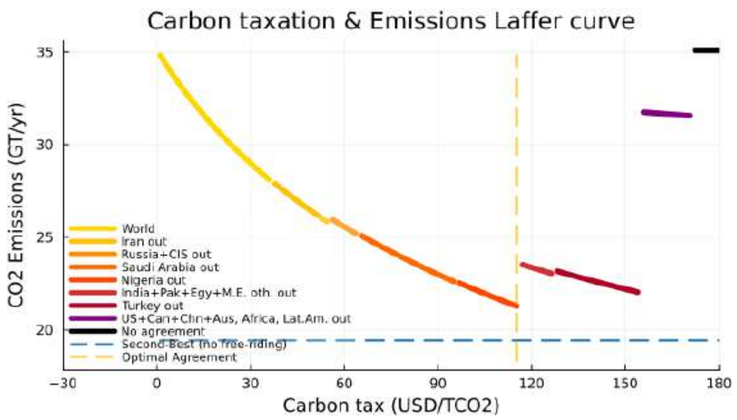


Figure 13: Global Emissions (yearly)
for different carbon taxes a given tariff $t^b = 50\%$

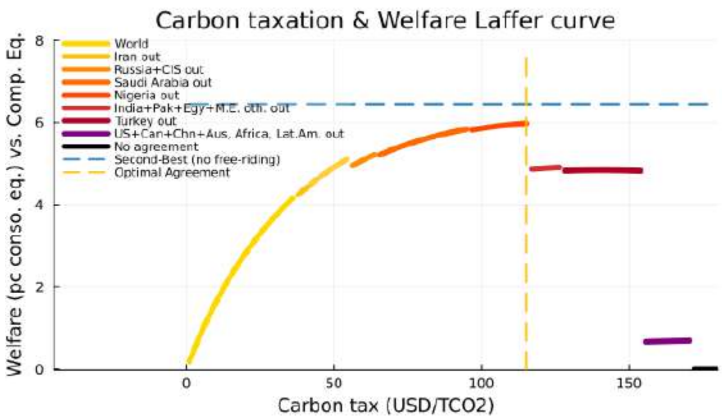


Figure 14: Global welfare (percent consumption eq.),
compared to the Competitive Equilibrium
for a given tariff $t^b = 48\%$

gain while a club with a more restricted set of countries and larger tax reaches only a 4.8% welfare gain. Making a smaller set of countries bear the cost of taxation is detrimental to their welfare: developed countries consume larger quantities of energy, and developing economies have a higher cost of distortion as their production and consumption are scarce – especially if they are affected by climate change. The agreement is stable because the cost of tariffs is larger, enforcing this cooperation. However, it is beneficial to work at the extensive margin to reduce the distortionary carbon tax, foster participation to share the costs of fighting climate change.

Because of endogenous participation, welfare and emissions are indeed different metrics that provide contrasting insights on what should be the optimal policy. In the next graphs, Figure 15 and Figure 16, I summarize, for the different equilibria we considered above, respectively the global emissions in Gigatons of CO_2 and welfare in consumption equivalent difference compared to the competitive equilibrium. Clearly, the First Best has the lowest emissions – $18.5GtCO_2$, a reduction of 47% relative to the Business-as-Usual scenario – and the maximum welfare – 13% of consumption equivalent change. In this case, the planner has access to unlimited instruments: it uses transfers to redistribute across countries, which offset the negative general equilibrium effects of taxation and allows the increase in carbon taxes and a further reduction in global carbon emissions. In contrast, in the Second-Best, these redistributive instruments are not available, which makes the welfare gains much smaller at 6.5%, although emissions are only slightly higher.

In these two benchmarks, we assume away the free-riding problem, which constrains the achievable policy and carbon reduction. I now compare the two equilibria considered above: the optimal agreement with all the countries except Russia, Saudi Arabia, Iran, and Nigeria reaches a reduction of 38% of carbon emissions. This reaches the welfare of 6% consumption equivalent while a club with larger taxes loses on welfare by increasing the tax burden on smaller set of participants.

The analysis of the potential welfare gains of the First-Best highlights that transfers can serve as a strong instrument to offset the negative effects of the uniform carbon taxation and tariffs, and I investigate if we can provide such welfare improvements with transfers Section 7.1 and with fossil-fuels specific tariffs in Section 7.3.

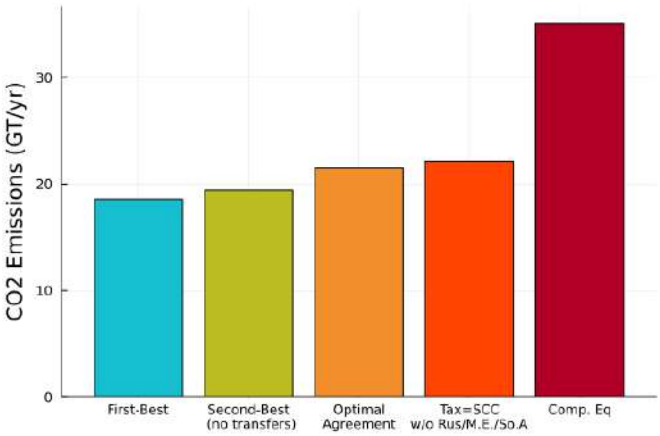


Figure 15: Global Emissions (yearly) comparison across equilibria

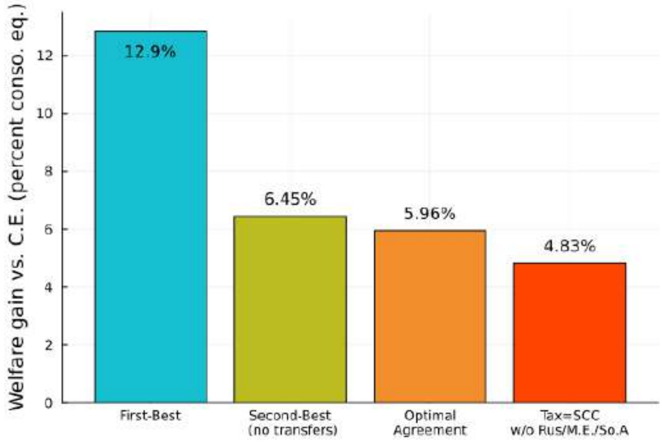


Figure 16: Global welfare (percent consumption eq.), comparison across equilibria

6.4 Coalition building

The social planner/designer chose the proposed optimal agreement. However, can this agreement be achieved by coalition building? Can a sequence of countries joining the climate agreement, in turn, reach this agreement? This relates to the question of which country has the most interest in joining such a club.

I investigate if this climate agreement can be constructed, with a sequence of "rounds" of our static equilibrium: At each round (n), each country decide to enter or not depending on the welfare gain:

$$\Delta_i \mathcal{U}_i(\mathbb{J}^{(n)}) = \mathcal{U}_i(\mathbb{J}^{(n)} \cup \{i\}, \mathrm{t}^{\varepsilon}, \mathrm{t}^{b}) - \mathcal{U}_i(\mathbb{J}^{(n)} \setminus \{i\}, \mathrm{t}^{\varepsilon}, \mathrm{t}^{b})$$

For now, the construction is evaluated at the optimal carbon tax $\mathrm{t}^{\varepsilon} = 98\$$, and tariff $\mathrm{t}^{b} = 50\%$ and perform this sequential procedure – which is a direct output in our CDCP algorithm / squeezing procedure in Section 3.2. This experiment is inspired by an analogous exercise in Farrokhi and Lashkaripour (2024).

The result of this exercise is a sequence building up to the optimal climate agreement:

- Round 1: European Union, i.e. Germany, France, Spain, Italy, Rest of EU
- Round 2: China, UK, Turkey, Rest of South and South-East Asia
- Round 3: USA, Japan, Korea, Australia, Thailand,
  Indonesia, Pakistan, Rest of Africa & Latin America
- Round 4: Canada, South-Africa, Mexico
- Round 5: India, Brazil, Egypt, Argentina, Rest of Middle-East

$\notin$ Stay out of the agreement: Russia, CIS, Saudi Arabia, Iran, Nigeria

The European Union has the best interest in reducing climate change, being positively affected by a decrease in the fossil-fuel price, consuming a small share of coal in their energy mix, and being wealthy enough to suffer less from the energy taxation cost in their production. In the second round, China, Turkey, important trading partners of European countries, and Southeast Asia, which has one of the highest gains in fighting global warming, in turn, join the climate agreement. In the third and fourth round, most other countries, which have large gains from trade, join the climate club to avoid retaliatory tariffs. Lastly, South-Asian and Middle Eastern countries also joins to be able to trade with the rest of the world.

7 Extensions: the impact of additional policy instruments and retaliation

In this section, I propose extensions to our baseline climate agreement by suggesting additional instruments to improve the allocation. By proposing simple policies that could be achievable in practice, I investigate if we can improve on the optimal climate agreement presented above.

7.1 Transfers and COP's "climate fund"

One of the major policy proposals of the COP28 in Dubai was the idea of *a loss and damage fund*, which translated into the COP29 in Baku in the *New Collective Quantified Goal (NCQG) on Climate Finance*. These ideas were, in substance, to give transfers to developing countries to accelerate the achievement of the Paris Agreement goals and compensate them for adaptation against climate change. I propose a simple implementation of such policies in my climate agreement framework. Given that the club is implementing a large carbon tax t^ε, one practical proposal is to redistribute a share of this tax revenue through lump-sum transfer across countries.

In the baseline agreement, the carbon tax revenues are redistributed to the household of the country paying the tax: $\mathrm{t}_i^{ls} = \mathrm{t}^\varepsilon(\xi^f e_i^f + \xi^c e_i^c)$. Here, the exercise allocates a share α^ε to a "climate fund" which then redistributes those revenues equally across countries, with a simple rule:

$$\mathrm{t}_i^{ls} = (1-\alpha^\varepsilon)\,\mathrm{t}^\varepsilon\,(\xi^f e_i^f + \xi^c e_i^c) + \alpha^\varepsilon \frac{1}{\mathcal{P}} \sum_{j\in\mathbb{J}} \mathcal{P}_j \mathrm{t}^\varepsilon(\xi^f e_j^f + \xi^c e_j^c)\ , \qquad \forall\ i \in \mathbb{J}$$

In practice, it transfers from large emitters – the developed economies – to low emitters – which are developing economies that tend to be more vulnerable to climate change.

I then choose the optimal share α^ε to maximize global welfare $\mathcal{W}(\alpha^\varepsilon)$, using a simple grid search. This results in the climate fund being optimal for $\alpha^{\varepsilon,\star} = 0\%$. With Negishi weights, the welfare-maximizing size of the climate fund is null: developed countries pollute more and do not want to transfer part of their carbon tax revenues.

I also simulate a climate agreement with a "climate fund" with a size $\alpha^\varepsilon \sum_{j\in\mathbb{J}} \mathcal{P}_j \mathrm{t}^\varepsilon(\xi^f e_j^f + \xi^c e_j^c)$ in line with amount that was agreed in the COP29 in Baku, through the NCQG on Climate Finance – \$300 billions per year. The results of that experiment are shown in Figure 17. This policy only redistributes lump-sum from high to low-energy users. The welfare costs are around $0.6 - 0.8\%$ for Europe, the United States, and advanced countries in Asia and Oceania. However, even a small "climate fund" is particularly welfare-improving for developing countries – like India, Pakistan, the rest of South Asia and the rest of Africa (not displayed), Brazil, and Indonesia – who gain around 1 to 3% welfare gains. Those regions are small contributors to the global climate externality, and such transfers would allow them to lower the cost of climate change through adaptation and dampen the redistributive cost of carbon taxation.

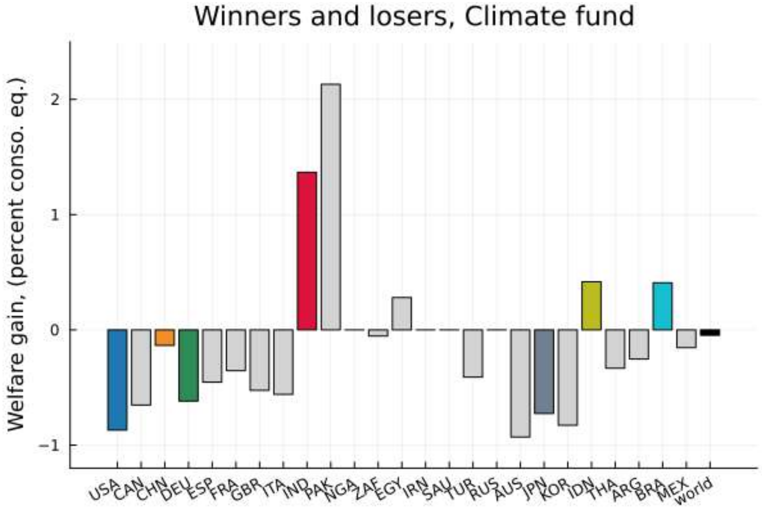


Figure 17: Welfare across countres
Optimal loss and damage fund: $\alpha^{\varepsilon,\star} = 15\%$

The amount agreed in COP29's final text was much lower than the \$1.3 trillion/year initially proposed by developing countries – raising objections from India and Nigeria delegations. Nevertheless, the framework I developed here provides explanations on why this amount can not be increased to more substantial amounts. First, more ambitious climate finance relying on direct transfers would imply large losses for the developed countries, which would contribute the most due to their large emissions. This would reinforce the free-riding problem by raising the cost of participating in the agreement. Second, the welfare function $\mathcal{W} = \sum_i \mathcal{P}_i \omega_i \mathcal{U}_i$ using the Negishi weights $\omega_i = 1/u'(c_i)$ is biased toward these advanced countries. It provides a representation of the balance of power in these international climate agreements. This explains why the effort for transfers – even as low as the one accepted in the COP29 – would be detrimental to the developed economies' welfare, which is not optimal for the planner.

7.2 Carbon Border Adjustment Mechanism (CBAM) and "carbon tariffs"

Carbon Border Adjustment Mechanism (CBAM) introduced in the European Union (and taking effect in 2026) and more generally "carbon tariffs" have been the main policy proposals to address carbon leakage. Indeed, implementing carbon pricing and fossil-fuel taxation reallocate production to regions not affected by the policy. To address this competitiveness effect, policy-makers have been suggesting implementing trade policy in a form of tariffs that scale with the carbon intensity, measured by the amount of carbon emission "embedded" in the production of the good imported. I analyze the implementation of such carbon tariffs in the climate agreement I studied above. In this context, this trade penalty would *replace* the uniform trade tariffs that act as a trade sanction for countries deviating.

$$\mathrm{t}^b_{ij} = \xi^y_j\, \mathrm{t}^{b,\varepsilon} = \frac{\xi^f e^f_j + \xi^c e^c_j}{y_j \mathrm{p}_j}\, \mathrm{t}^{b,\varepsilon} \qquad \text{if } i \in \mathbb{J}, j \notin \mathbb{J} \ ,$$

where $\mathrm{t}^{b,\varepsilon}$ is the carbon tax or carbon price (in $\$/t\,CO_2$) imposed on the good from country j and ξ^y_i is the carbon intensity (in $t\,CO_2/\$$) for the output y_j. The climate agreement need to decide on the design $\{\mathbb{J}, \mathrm{t}^\varepsilon, \mathrm{t}^{b,\varepsilon}\}$ and the result of this policy experiments are diplayed in Figure 18.

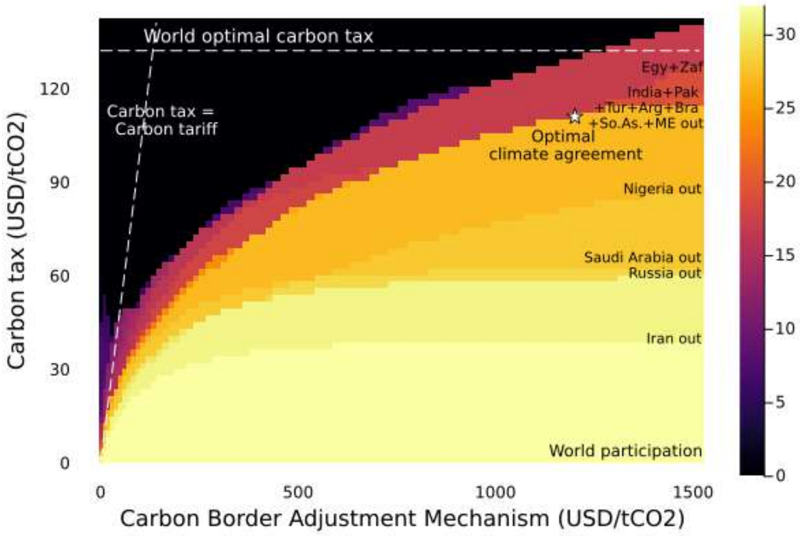


Figure 18: Agreement participation
Penalty with carbon tariffs: $\mathrm{t}^b_{ij} = \xi^y_j\, \mathrm{t}^{b,\varepsilon}$ (with $\mathrm{t}^{b,\varepsilon}$ on x-axis)

We see on that figure that the carbon price/carbon tariff that need to be implemented is much higher $> 500\$/tCO_2$ to have a chance of sustaining a club with a significant carbon tax internally and a large enough set of countries. In that case, the optimal climate agreement gathers again all the regions in the world at the exception of Russia, Iran, Nigeria, and Saudi Arabia. It needs to impose a carbon tax of $113\$/tCO_2$ and a carbon tariff of $1250\,\$/tCO_2$. The intuition for this result are the same as above: to navigate the intensive-extensive margins tradeoff, it is beneficial to lower the carbon tax for members of the club and leave several fossil-fuel producers outside of the club.

The external carbon price imposed as carbon tariffs is much higher than the internal carbon price. Indeed, to encourage participation, we need implement a strong penalty. To replace the 50% uniform tariff on goods, the club now need to impose a $1250\,\$/tCO_2$ as carbon tariff. However, suppose now that the climate agreement faces an additional policy constraint – for example from World Trade Organization rules – that the tariff set on carbon externally should equal the carbon tax set internally $\mathrm{t}^{b,\varepsilon} = \mathrm{t}^\varepsilon$. This is represented in Figure 18 by the white dashed 45° line. We see that this constraint prevents the agreement to achieve any meaningful carbon tax and large participation. The optimal agreement in that context would be much less ambitious: it would achieve policies $\{\mathrm{t}^\varepsilon, \mathrm{t}^{b,\varepsilon}\} = \{45, 45\}\,\$/tCO_2$ with a participation gathering the world without Russia and many Middle-Eastern, South-Asian and Latin American countries. The analysis of this exercise implies that such agreements would need to bypass some of these WTO policy constraints to achieve higher carbon mitigation.

7.3 Fossil-fuels specific tariffs

In this section, I relax the assumption of free trade on fossil-fuel energy. In the current climate club, the members of the club only impose penalty tariffs on the final goods traded by the firm, and not on energy imports. This is empirically relevant, c.f. Shapiro (2021) and Copeland et al. (2021): Inputs are more emission-intensives but trade policy is biased against downstream goods. Moreover, in the context of this model, fossil-fuel energy inputs are not carbon-intensive *per-se*; it is their use – i.e. the burning – of those fossil fuels in production that is carbon-intensive. As a result, carbon-border adjustment mechanisms only impose a tariff on the "scope-1/scope-2" carbon footprint of fossil fuel extraction – and not the "scope-3" of its use along the downstream supply chain. Imposing a carbon tax on both fossil-fuel imports and their use in production amounts to a double tax on carbon, which is, in general, not optimal nor implemented in practice.

In our climate club setting, these tariffs are also strategic to incentivize participation. Therefore, I propose an alternative mechanism where the club members impose a tariff on the fossil fuel exports of the countries outside the club. The tariff is an import tax on energy imports from non-participants, taking the form:

$$q^f_{\mathbb{J}} = (1+\mathrm{t}^{bf})q^f_{\mathbb{I}\setminus\mathbb{J}}$$

if non-members export fossil fuels to the club, i.e. $\sum_{i\in\mathbb{I}\setminus\mathbb{J}} e^x_i > \sum_{i\in\mathbb{I}\setminus\mathbb{J}} e^f_i$ and $\sum_{i\in\mathbb{J}} e^f_i > \sum_{i\in\mathbb{J}} e^x_i$. This tax imposed on oil-gas is redistributed lump-sum to the club households, scaled by the country oil-gas consumption. Such a policy implies a lower equilibrium price[34] for non-member $q^f_{\mathbb{I}\setminus\mathbb{J}} < q^f_{\mathbb{J}}$.

In the following graph Figure 19, I plot over different values of t^{bf} the welfare impact for members of the club and the non-members – which is Russia, Saudi Arabia, Nigeria, and Iran here – to see if this strategic tariff provides enough incentives for those fossil-fuels producers to join the club with high carbon taxes.

With very large fossil-fuel-specific tariffs of $\mathrm{t}^{bf\star} = 30\%$ of the price of oil and gas, the welfare of Russia, Saudi Arabia, as well as Iran and Nigeria (not plotted) can be lowered enough due to the drying out of their energy rents. In that case, they find it beneficial to participate in an agreement that replicates the second-best allocation instead of suffering such sanctions.

Note that such a strong tariff on oil-gas aligns well with the price cap on Russian oil that has been implemented after Russia's invasion of Ukraine. Indeed, the European Union implemented a price cap of \$60/barrel in Dec. 2022, compared to an oil price around \$100/barrel in the Winter 2022-23. The economic theory of such price caps is studied in Johnson, Rachel and Wolfram (2023).

[34] However, if non-member countries do not export fossil fuels to the club, then the two oil-gas clear separately and can have different prices – although this case does not occur in equilibrium.

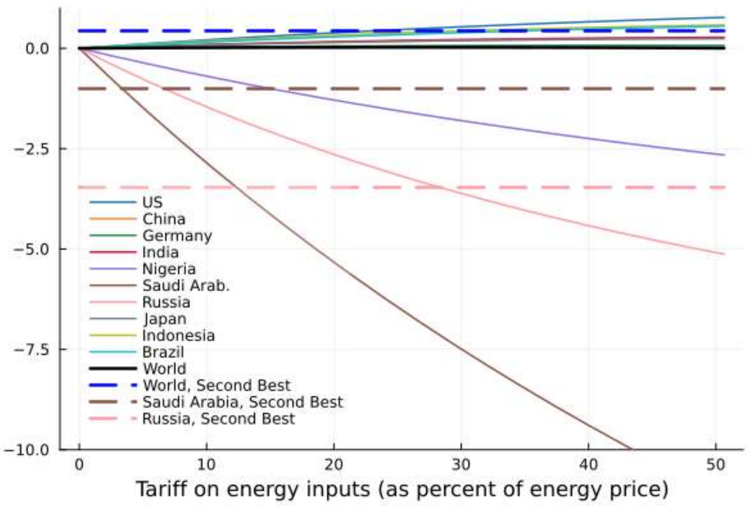


Figure 19: Welfare for world and for non-member (Russia)
Tariff on oil-gas energy: $\mathrm{t}^{bf\,\star} = 40\%$

7.4 Retaliation

In this section, I consider an extension of the game-theoretical setting where countries outside the climate club can also act strategically. Until now, the countries outside the agreement were passive, setting policy $\{\mathrm{t}_i^{\varepsilon}, \mathrm{t}_{ij}^{b}\} = \{0, 0\}_{i \notin \mathbb{J}}$.

I now relax this assumption by conducting the following exercise: All the countries outside the club $j \notin \mathbb{J}$ impose a tariffs on club members i as a retaliation from any trade policy they would be targeted with. That is,

$$\mathrm{t}_{ji}^{b} = \kappa \mathrm{t}_{ij}^{b} \ , \qquad\qquad \forall\, i \in \mathbb{J} \ ,$$

where $\kappa \in (0, 1]$ is an exogenous parameter that represents the extent of the retaliation. In future extensions, I plan to make this decision endogenous, creating a multiplayer game between a climate coalition and a fringe of non-members.

I perform this exercise for different values of κ. The results are displayed in Figure 20, Figure 21 and Figure 22, respectively for $\kappa = 10, 50\%, 100\%$.

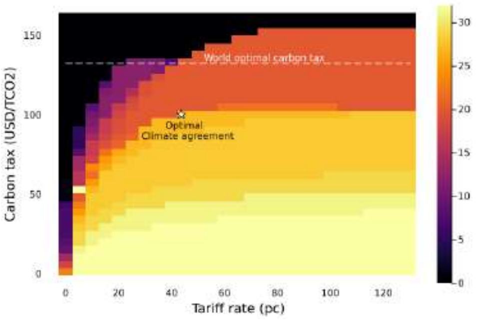


Figure 20: $\kappa = 0.1$

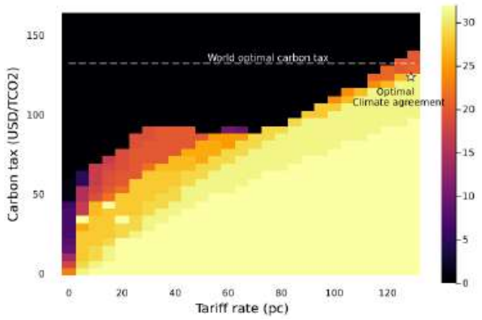


Figure 21: $\kappa = 0.5$

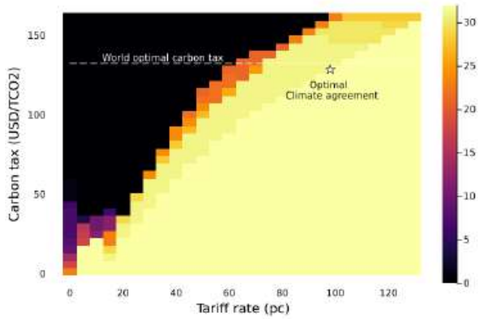


Figure 22: $\kappa = 1.0$

For moderate values of retaliation of non-members, i.e. $\kappa \in (0, 0.4)$, we note that the climate club is more constrained than before: the achievable carbon tax is slightly lower, and the tariff needed to enforce it needs to be raised. For example for $\kappa = 0.10$, the optimal carbon tax t^{ε} is less than $\$100/tCO_2$, and the tariff is still $t^b = 45\%$ – instead of \$115 carbon tax in the $\kappa = 0$ case.

However, the larger the retaliation, the larger the cost of trade disruption for both members *and* nonmembers. The countries with the largest gains from trade would still choose optimally to participate in the agreements, which makes the cost of being outside larger as κ grows above 0.5.

When κ becomes very large, it is optimal for the climate club to engage in an aggressive trade war, which pushes non-members to finally join the agreement. For example, when $\kappa = 1$, we see that the climate club recovers its enforcement ability. It can even incentivize complete participation, for a carbon tax up to the optimal level $t^{\varepsilon} = \$131/tCO_2$, for large tariffs of $t^b = 100\%$. Importantly, in equilibrium with complete participation, no country pays these tariffs.

These findings underscore that understanding the underpinnings of trade tariff strategic behavior – beyond the simple terms-of-trade manipulation motives – is key to designing climate agreements.

8 Conclusion

This paper examines the design of an optimal climate agreement in the presence of free-riding incentives and redistributive effects. I develop a multi-country Integrated Assessment Model (IAM) that incorporates international trade in goods and energy markets for fossil fuels. This model accounts for heterogeneity across countries in terms of their vulnerability to climate change, income levels, energy mix, and positions as exporters or importers of goods and energy.

The analysis focuses on a global social planner's problem of maximizing world welfare through a climate agreement comprising three key elements: (1) the set of countries included in the "climate club", (2) a carbon tax imposed on club members, and (3) a level of uniform trade tariffs imposed on non-member countries. I consider Nash equilibria where countries make strategic decisions about their participation, either unilaterally or through coalition deviations.

This study reveals a crucial trade-off between an intensive and an extensive margin in designing the optimal climate agreement. A small coalition of countries can implement high carbon taxes, achieving significant individual emissions reductions. However, a more extensive club with broader participation may be necessary for effectively combating global climate change, albeit at the cost of lower carbon taxes and higher tariffs.

The main findings is first that the optimal climate club includes all countries except Russia, with a moderate carbon tax of \$100 per ton of CO_2 and a 50% tariff on goods from non-participants. To increase participation, it is beneficial to reduce the carbon tax by 35% from the globally optimal level of \$150 per ton of CO_2. This allows for the inclusion of Middle Eastern countries and several developing economies in South Asia and Africa. Excluding fossil fuel producers like Russia from the agreement is optimal, as their welfare costs from carbon taxation are too high to justify inclusion at any reasonable tax rate. Trade policy, particularly the threat of tariffs, is a key strategic instrument for undermining free-riding and incentivizing participation. However, its effectiveness diminishes as carbon taxes increase. I analyze the benefits and limitations of additional policy instruments, such as transfers through a "climate fund", carbon border adjustement mechanisms and "carbon tariffs", or fossil-fuel-specific tariffs. I show how they can improve the climate agreement and push the carbon tax closer to the second-best allocation.

In conclusion, this research highlights the strategic considerations at play in the design of effective climate agreements in the presence of heterogeneous countries and diverging interests. It demonstrates that while a universal agreement with globally optimal carbon taxation may be unattainable due to free-riding incentives and redistributive effects, carefully designed climate clubs with strategic use of trade policy can achieve significant progress in global climate action.

In future research, I will explore the dynamics of climate clubs. As discussed in Nordhaus (2021), dynamics in technological change of non-carbon energy production offer opportunities for countries to adopt carbon mitigation policies. The intertemporal trade-off between trade sanctions, future costs of climate change, rising costs of fossil-energy prices due to depleting reserves, and future gains of the renewable energy transition should be tilted toward climate action. Analyzing this problem remains mathematically challenging as it involves solving a dynamic multilateral game between heterogeneous countries. However, the approach I have presented in this article can be adapted to study coalition-building and time-varying carbon tax and tariffs. Indeed, would it be better to build a broad-but-shallow club now and deepening it later with higher carbon tax, or is it optimal to start from a deep-but-narrow agreement today to broaden it later? I keep the analysis of this question for future research. The current paper shows that climate agreements, with climate policies such as carbon taxation and trade tariffs, can provide the right incentives and offer a pivotal step toward reducing global emissions and temperatures and a hope in our fight against climate change.

References

Adao, Rodrigo, Arnaud Costinot and Dave Donaldson (2023), 'Putting Quantitative Models to the Test: An Application to Trump's Trade War'.

Anderson, James E. (1979), 'A theoretical foundation for the gravity equation', *The American Economic Review* **69**(1), 106–116.

Anthoff, David, Cameron Hepburn and Richard SJ Tol (2009), 'Equity weighting and the marginal damage costs of climate change', *Ecological Economics* **68**(3), 836–849.

Antras, Pol, Teresa C. Fort, Agustin Gutierrez and Felix Tintelnot (2024), 'Trade Policy and Global Sourcing: An Efficiency Rationale for Tariff Escalation', *Journal of Political Economy Macroeconomics* **2**(1), 1–44. Publisher: The University of Chicago Press.

Arkolakis, Costas, Arnaud Costinot and Andres Rodriguez-Clare (2012), 'New Trade Models, Same Old Gains?', *American Economic Review* **102**(1), 94–130.

Arkolakis, Costas and Conor Walsh (2023), 'Clean Growth'.

Arkolakis, Costas, Fabian Eckert and Rowan Shi (2023), 'Combinatorial discrete choice: A quantitative model of multinational location decisions'.

Atkeson, Andrew and Patrick J Kehoe (1999), 'Models of energy use: Putty-putty versus putty-clay', *American Economic Review* **89**(4), 1028–1043.

Barrage, Lint and William Nordhaus (2024), 'Policies, projections, and the social cost of carbon: Results from the DICE-2023 model', *Proceedings of the National Academy of Sciences* **121**(13), e2312030121. Publisher: Proceedings of the National Academy of Sciences.

Barrett, Scott (1994), 'Self-Enforcing International Environmental Agreements', *Oxford Economic Papers* **46**(Supplement 1), 878–894.

Barrett, Scott (2001), Strategic Environmental Policy and International Trade, *in* 'The Economics of International Trade and the Environment', CRC Press. Num Pages: 11.

Barrett, Scott (2003), *Environment and Statecraft : The Strategy of Environmental Treaty-Making: The Strategy of Environmental Treaty-Making*, OUP Oxford. Google-Books-ID: uqrey86neSIC.

Barrett, Scott (2013), 'Climate treaties and approaching catastrophes', *Journal of Environmental Economics and Management* **66**(2), 235–250.

Batabyal, Amitrajeet, ed. (2000), *The Economics of International Environmental Agreements*, Routledge, London.

Belfiori, Elisa, Daniel Carroll and Sewon Hur (2024), 'Unequal Climate Policy in an Unequal World'.

Belfiori, Maria Elisa (2018), 'Climate change and intergenerational equity: Revisiting the uniform taxation principle on carbon energy inputs', *Energy Policy* **121**, 292–299.

Bhandari, Anmol, David Evans, Mikhail Golosov and Thomas J. Sargent (2021*a*), 'Inequality, Business Cycles, and Monetary-Fiscal Policy', *Econometrica* **89**(6), 2559–2599.

Bhandari, Anmol, David Evans, Mikhail Golosov and Tom Sargent (2021*b*), 'Efficiency, Insurance, and Redistribution Effects of Government Policies'.

Bhattarai, Keshab, Sushanta K. Mallick and Bo Yang (2021), 'Are global spillovers complementary or competitive? Need for international policy coordination', *Journal of International Money and Finance* **110**, 102291.

Bilal, Adrien and Diego R Känzig (2024), 'The macroeconomic impact of climate change: Global vs. local temperature'.

Bilal, Adrien and Esteban Rossi-Hansberg (2023), 'Anticipating Climate Change Across the United States'.

Bohringer, Christoph, Jared C. Carbone and Thomas F. Rutherford (2012), 'Unilateral climate policy design: Efficiency and equity implications of alternative instruments to reduce carbon leakage', *Energy Economics* **34**, S208–S217.

Bohringer, Christoph, Jared C. Carbone and Thomas F. Rutherford (2016), 'The Strategic Value of Carbon Tariffs', *American Economic Journal: Economic Policy* **8**(1), 28–51.

Bolt, Jutta and Jan Luiten van Zanden (2023), 'Maddison-style estimates of the evolution of the world economy: A new 2023 update', *Journal of Economic Surveys* **n/a**(n/a).

Bornstein, Gideon, Per Krusell and Sergio Rebelo (2023), 'A World Equilibrium Model of the Oil Market', *The Review of Economic Studies* **90**(1), 132–164.

Bourany, Thomas (2022), 'Energy shocks and aggregate fluctuations'.

Bourany, Thomas (2025), 'Climate Change, Inequality, and Optimal Climate Policy'.

Bourany, Thomas and Jordan Rosenthal-Kay (2025), 'The Winners and Losers of Climate Policies: A Sufficient Statistics Approach'.

Burke, Marshall, Solomon M Hsiang and Edward Miguel (2015), 'Global non-linear effect of temperature on economic production', *Nature* **527**(7577), 235–239.

Carleton, Tamma A, Amir Jina, Michael T Delgado, Michael Greenstone, Trevor Houser, Solomon M Hsiang, Andrew Hultgren, Robert E Kopp, Kelly E McCusker, Ishan B Nath, James Rising, Ashwin Rode, Hee Kwon Seo, Arvid Viaene, Jiacan Yuan and Alice Tianbo Zhang (2020), Valuing the Global Mortality Consequences of Climate Change Accounting for Adaptation Costs and Benefits, Working Paper 27599, National Bureau of Economic Research. Series: Working Paper Series.

Carraro, Carlo and Domenico Siniscalco (1993), 'Strategies for the international protection of the environment', *Journal of public Economics* **52**(3), 309–328.

Castro-Vincenzi, Juanma, Gaurav Khanna, Nicolas Morales and Nitya Pandalai-Nayar (2024), 'Weathering the Storm: Supply Chains and Climate Risk'.

Chander, Parkash (2018), *Game theory and climate change*, Columbia University Press.

Chander, Parkash and Henry Tulkens (1995), 'A core-theoretic solution for the design of cooperative agreements on transfrontier pollution', *International tax and public finance* **2**, 279–293.

Chander, Parkash and Henry Tulkens (1997), 'The core of an economy with multilateral environmental externalities', *International Journal of Game Theory* **26**, 379–401.

Chari, V. V., Juan Pablo Nicolini and Pedro Teles (2023), 'Optimal Cooperative Taxation in the Global Economy', *Journal of Political Economy* **131**(1), 95–130. Publisher: The University of Chicago Press.

Chen, Zhiqi (1997), 'Negotiating an Agreement on Global Warming: A Theoretical Analysis', *Journal of Environmental Economics and Management* **32**(2), 170–188.

Chichilnisky, Graciela and Geoffrey Heal (1994), 'Who should abate carbon emissions?: An international viewpoint', *Economics Letters* **44**(4), 443–449.

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

Conte, Bruno, Klaus Desmet, Dávid Krisztián Nagy and Esteban Rossi-Hansberg (2021), 'Local sectoral specialization in a warming world', *Journal of Economic Geography* **21**(4), 493–530.

Conte, Maddalena, Pierre Cotterlaz, Thierry Mayer et al. (2022), 'The cepii gravity database'.

Copeland, Brian R., Joseph S. Shapiro and M. Scott Taylor (2021), 'Globalization and the Environment'.

Copeland, Brian R and M Scott Taylor (2004), 'Trade, growth, and the environment', *Journal of Economic literature* **42**(1), 7–71.

Costinot, Arnaud, Dave Donaldson, Jonathan Vogel and Ivan Werning (2015), 'Comparative Advantage and Optimal Trade Policy', *The Quarterly Journal of Economics* **130**(2), 659–702.

Costinot, Arnaud and Iván Werning (2023), 'Robots, trade, and luddism: A sufficient statistic approach to optimal technology regulation', *The Review of Economic Studies* **90**(5), 2261–2291.

Covert, Thomas, Michael Greenstone and Christopher R Knittel (2016), 'Will we ever stop using fossil fuels?', *Journal of Economic Perspectives* **30**(1), 117–138.

Cruz, Jose-Luis and Esteban Rossi-Hansberg (2022), 'Local Carbon Policy'.

Cruz, Jose-Luis and Esteban Rossi-Hansberg (2024), 'The Economic Geography of Global Warming', *The Review of Economic Studies* **91**(2), 899–939.

Davila, Eduardo and Andreas Schaab (2023), 'Optimal Monetary Policy with Heterogeneous Agents: Discretion, Commitment, and Timeless Policy'.

Davila, Eduardo and Ansgar Walther (2022), 'Corrective Regulation with Imperfect Instruments'.

Dietz, Simon, Frederick van der Ploeg, Armon Rezai and Frank Venmans (2021), 'Are Economists Getting Climate Dynamics Right and Does It Matter?', *Journal of the Association of Environmental and Resource Economists* **8**(5), 895–921. Publisher: The University of Chicago Press.

Dingel, Jonathan I., Kyle C. Meng and Solomon M. Hsiang (2019), 'Spatial Correlation, Trade, and Inequality: Evidence from the Global Climate'.

Douenne, Thomas, Albert Jan Hummel and Marcelo Pedroni (2023), 'Optimal Fiscal Policy in a Climate-Economy Model with Heterogeneous Households'.

Dutta, Prajit K and Roy Radner (2004), 'Self-enforcing climate-change treaties', *Proceedings of the National Academy of Sciences* **101**(14), 5174–5179.

Dutta, Prajit K and Roy Radner (2009), 'A strategic analysis of global warming: Theory and some numbers', *Journal of Economic Behavior & Organization* **71**(2), 187–209.

Energy Institute (2024), 'Statistical review of world energy'.

Ernst, Anne, Natascha Hinterlang, Alexander Mahle and Nikolai StÃ¤hler (2023), 'Carbon pricing, border adjustment and climate clubs: Options for international cooperation', *Journal of International Economics* **144**, 103772.

Falkner, Robert (2016), 'The Paris Agreement and the new logic of international climate politics', *International' Affairs* **92**(5), 1107–1125.

Falkner, Robert, Naghmeh Nasiritousi and Gunilla Reischl (2022), 'Climate clubs: politically feasible and desirable?', *Climate Policy* **22**(4), 480–487.

Farrokhi, Farid and Ahmad Lashkaripour (2024), 'Can Trade Policy Mitigate Climate Change', *Econometrica* .

Folini, Doris, Aleksandra Friedl, Felix Kübler and Simon Scheidegger (2024), 'The climate in climate economics', *Review of Economic Studies* p. rdae011.

Fontagne, Lionel and Katheline Schubert (2023), 'The Economics of Border Carbon Adjustment: Rationale and Impacts of Compensating for Carbon at the Border', *Annual Review of Economics* **15**(Volume 15, 2023), 389–424. Publisher: Annual Reviews.

Gentile, Claudia (2024), 'Relying on intermittency: Clean energy, storage, and innovation in a macro climate model'.

Golosov, Mikhail, John Hassler, Per Krusell and Aleh Tsyvinski (2014), 'Optimal Taxes on Fossil Fuel in General Equilibrium', *Econometrica* **82**(1), 41–88.

Hagen, Achim and Jan Schneider (2021), 'Trade sanctions and the stability of climate coalitions', *Journal of Environmental Economics and Management* **109**, 102504.

Harstad, Bård (2012), 'Climate contracts: A game of emissions, investments, negotiations, and renegotiations', *Review of Economic Studies* **79**(4), 1527–1557.

Harstad, Bård (2023), 'Pledge-and-Review Bargaining: from Kyoto to Paris', *The Economic Journal* **133**(651), 1181–1216.

Harstad, Bård, Francesco Lancia and Alessia Russo (2022), 'Prices vs. quantities for self-enforcing agreements', *Journal of Environmental Economics and Management* **111**, 102595.

Hassler, John, Per Krusell and Anthony A Smith (2016), Environmental macroeconomics, *in* 'Handbook of macroeconomics', Vol. 2, Elsevier, pp. 1893–2008.

Hassler, John, Per Krusell and Conny Olovsson (2010), 'Oil monopoly and the climate', *American Economic Review* **100**(2), 460–64.

Hassler, John, Per Krusell and Conny Olovsson (2019), Directed technical change as a response to natural-resource scarcity, Working Paper 375, Sveriges Riksbank Working Paper Series.

Hassler, John, Per Krusell and Conny Olovsson (2021), 'Presidential Address 2020 Suboptimal Climate Policy', *Journal of the European Economic Association* **19**(6), 2895–2928.

Hillebrand, Elmar and Marten Hillebrand (2019), 'Optimal climate policies in a dynamic multi-country equilibrium model', *Journal of Economic Theory* **179**, 200–239.

Hoel, Michael (1991), 'Global environmental problems: The effects of unilateral actions taken by one country', *Journal of Environmental Economics and Management* **20**(1), 55–70.

Hoel, Michael (1994), 'Efficient Climate Policy in the Presence of Free Riders', *Journal of Environmental Economics and Management* **27**(3), 259–274.

Hsiao, Allan (2022), 'Coordination and Commitment in International Climate Action'.

IPCC, Pörtner, H.-O., D.C. Roberts, H. Adams, I. Adelekan, C. Adler, R. Adrian, P. Aldunce, E. Ali, R. Ara Begum, B. Bednar Friedl, R. Bezner Kerr, R. Biesbroek, J. Birkmann, K. Bowen, M.A. Caretta, J. Carnicer, E. Castellanos, T.S. Cheong, W. Chow, G. Cisse G. Cisse and Z. Zaiton Ibrahim (2022), *Climate Change 2022: Impacts, Adaptation and Vulnerability*, Technical Summary, Cambridge University Press, Cambridge, UK and New York, USA.

Iverson, Terrence (2024), 'Tiered Climate Clubs: Global Abatement Without Global Agreement'.

Jia, Panle (2008), 'What happens when wal-mart comes to town: An empirical analysis of the discount retailing industry', *Econometrica* **76**(6), 1263–1316.

Johnson, Simon, Lukasz Rachel and Catherine Wolfram (2023), 'Design and implementation of the price cap on russian oil exports', *Journal of Comparative Economics* **51**(4), 1244–1252.

Kleinman, Benny, Ernest Liu and Stephen J Redding (2020), 'International friends and enemies', *Forthcoming, American Economic Journal: Macroeconomics* .

Kleinman, Benny, Ernest Liu, Stephen J Redding and Motohiro Yogo (2023), 'Neoclassical growth in an interdependent world'.

Kortum, Samuel S. and David A. Weisbach (2021), 'Optimal Unilateral Carbon Policy'.

Kotlikoff, Laurence, Felix Kubler, Andrey Polbin, Jeffrey Sachs and Simon Scheidegger (2021), 'Making Carbon Taxation a Generational Win Win', *International Economic Review* **62**(1), 3–46. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/iere.12483.

Kotlikoff, Laurence J., Felix Kubler, Andrey Polbin and Simon Scheidegger (2021), 'Can Today's and Tomorrow's World Uniformly Gain from Carbon Taxation?'.

Krusell, Per and Anthony A. Smith (2022), 'Climate change around the world'.

Le Grand, François, Alaïs Martin-Baillon and Xavier Ragot (2024), 'Should monetary policy care about redistribution? optimal monetary and fiscal policy with heterogeneous agents'.

Maggi, G. (2016), Issue Linkage, *in* 'Handbook of Commercial Policy', Vol. 1, Elsevier, pp. 513–564.

Maggi, Giovanni and Robert W. Staiger (2022), 'International Climate Agreements and the Scream of Greta'.

Mahdiraji, Hannan Amoozad, Elham Razghandi and Adel Hatami-Marbini (2021), 'Overlapping coalition formation in game theory: A state-of-the-art review', *Expert Systems with Applications* **174**, 114752.

Nordhaus, William (2015), 'Climate Clubs: Overcoming Free-riding in International Climate Policy', *American Economic Review* **105**(4), 1339–1370.

Nordhaus, William (2021), 'Climate Club Futures: On the Effectiveness of Future Climate Clubs', *Cowles Foundation Discussion Papers* .

Nordhaus, William D (2011), 'Estimates of the social cost of carbon: background and results from the rice-2011 model'.

Nordhaus, William D. and Zili Yang (1996), 'A Regional Dynamic General-Equilibrium Model of Alternative Climate-Change Strategies', *The American Economic Review* **86**(4), 741–765. Publisher: American Economic Association.

Okada, Akira (2023), 'A dynamic climate negotiation game achieving full cooperation', *Journal of Economic Behavior & Organization* **214**, 657–669.

Ossa, Ralph (2014), 'Trade Wars and Trade Talks with Data', *American Economic Review* **104**(12), 4104–4146.

Papageorgiou, Chris, Marianne Saam and Patrick Schulte (2017), 'Substitution between Clean and Dirty Energy Inputs: A Macroeconomic Perspective', *The Review of Economics and Statistics* **99**(2), 281–290.

Pickering, Jonathan, Steve Vanderheiden and Seumas Miller (2012), '"If Equity's In, We're Out": Scope for Fairness in the Next Global Climate Agreement', *Ethics & International Affairs* **26**(4), 423–443.

Pindyck, Robert S (2017), 'The use and misuse of models for climate policy', *Review of Environmental Economics and Policy* .

Ray, Debraj and Rajiv Vohra (2015), Chapter 5 - Coalition Formation, *in* H. P.Young and S.Zamir, eds, 'Handbook of Game Theory with Economic Applications', Vol. 4, Elsevier, pp. 239–326.

Rudik, Ivan, Gary Lyn, Weiliang Tan and Ariel Ortiz-Bobea (2021), 'The Economic Effects of Climate Change in Dynamic Spatial Equilibrium'.

Shapiro, Joseph S (2021), 'The Environmental Bias of Trade Policy', *The Quarterly Journal of Economics* **136**(2), 831–886.

Tarr, David G., Dmitrii E. Kuznetsov, Indra Overland and Roman Vakulchuk (2023), 'Why carbon border adjustment mechanisms will not save the planet but a climate club and subsidies for transformative green technologies may', *Energy Economics* **122**, 106695.

Weisbach, David A., Samuel Kortum, Michael Wang and Yujia Yao (2023), 'Trade, Leakage, and the Design of a Carbon Tax', *Environmental and Energy Policy and the Economy* **4**, 43–90. Publisher: The University of Chicago Press.

Weitzman, Martin L. (2003), Prices vs. Quantities 1 , 2, *in* 'The Theory and Practice of Command and Control in Environmental Policy', Routledge. Num Pages: 15.

Weitzman, Martin L (2012), 'Ghg targets as insurance against catastrophic climate damages', *Journal of Public Economic Theory* **14**(2), 221–244.

Weitzman, Martin L. (2015), 'Internalizing the Climate Externality: Can a Uniform Price Commitment Help?', *Economics of Energy & Environmental Policy* **4**(2), 37–50. Publisher: International Association for Energy Economics.

Appendix

A Calibration

Table 2: Baseline calibration

Technology & Energy markets			
α	0.35	Capital share in $F(\cdot)$	Capital/Output ratio
ϵ	0.12	Energy share in $F(\cdot)$	Energy cost share (8.5%)
σ^y	0.3	Elasticity capital-labor vs. energy	Complementarity in production (c.f. Bourany 2022)
ω^f	0.56	Fossil energy share in $e(\cdot)$	Oil-gas/Energy ratio
ω^c	0.27	Coal energy share in $e(\cdot)$	Coal/Energy ratio
ω^r	0.17	Non-carbon energy share in $e(\cdot)$	Non-carbon/Energy ratio
σ^e	2.0	Elasticity fossil-coal-non-carbon	Slight substitutability & Study by Stern
δ	0.06	Depreciation rate	Investment/Output ratio
$\bar{g}$	0.01	Long run TFP growth	Conservative estimate for growth
θ	5.0	Trade elasticity (CES)	Gravity equation estimation
Preferences & Time horizon			
ρ	0.03	HH Discount factor	Long term interest rate & usual calib. in IAMs
η	1.5	IES / Risk aversion	Standard calibration
n	0.0035	Long run population growth	Conservative estimate for growth
ω_i	1	Pareto weights	Uniforms / Utilitarian Social Planner
ω_i	$1/u'(c_i)$	Pareto weights	Negishi / Status-quo Social Planner
T	400	Time horizon	Time for climate system to stabilize
Climate parameters			
ξ^f	2.761	Emission factor – Oil & natural gas	Conversion 1 $MTOE \Rightarrow$ 1 $MT\ CO_2$
ξ^c	3.961	Emission factor – Coal	Conversion 1 $MTOE \Rightarrow$ 1 $MT\ CO_2$
χ	$2.3/1e6$	Climate sensitivity	Pulse experiment: $100\,GtC \equiv 0.23°C$ medium-term warming
δ_s	0.0004	Carbon exit from atmosphere	Pulse experiment: $100\,GtC \equiv 0.15°C$ long-term warming
ζ	0.027	Growth rate, Carbon Capture and Storage	Starting after 2100, Follows Krusell Smith (2022)
$\gamma^{\oplus}$	0.003406	Damage sensitivity	Nordhaus' DICE
$\gamma^{\ominus}$	$0.3\times\gamma^{\oplus}$	Damage sensitivity	Nordhaus' DICE & Rudik et al (2022)
α^T	0.5	Weight historical climate for optimal temp.	Marginal damage correlated with initial temp.
$T^\star$	14.5	Optimal yearly temperature	Average spring temperature / Developed economies

A.1 Additional calibration graphs

A.1.1 Quantification – Trade shares

We displayed the trade share from the data in Figure 23 and how we calibrate the trade model.

Armington Trade model and trade shares:

$$s_{ij} \equiv \frac{c_{ij}p_{ij}}{c_i\mathbb{P}_i} = a_{ij}\frac{((1+\mathrm{t}_{ij})\tau_{ij}\mathrm{p}_j)^{1-\theta}}{\sum_k a_{ik}((1+\mathrm{t}_{ik})\tau_{ik}\mathrm{p}_k)^{1-\theta}}$$

We estimate a gravity regression, and CES $\theta = 5.63$. The Iceberg cost τ_{ij} are projection of geographical distance $\log\tau_{ij} = \beta\log d_{ij}$. The preference parameters a_{ij} identified as remaining variation in the trade share s_{ij}. As a results, both τ_{ij} and a_{ij} are policy invariant in our climate agreement setting. The description of the procedure is detailed in Section 4.3.

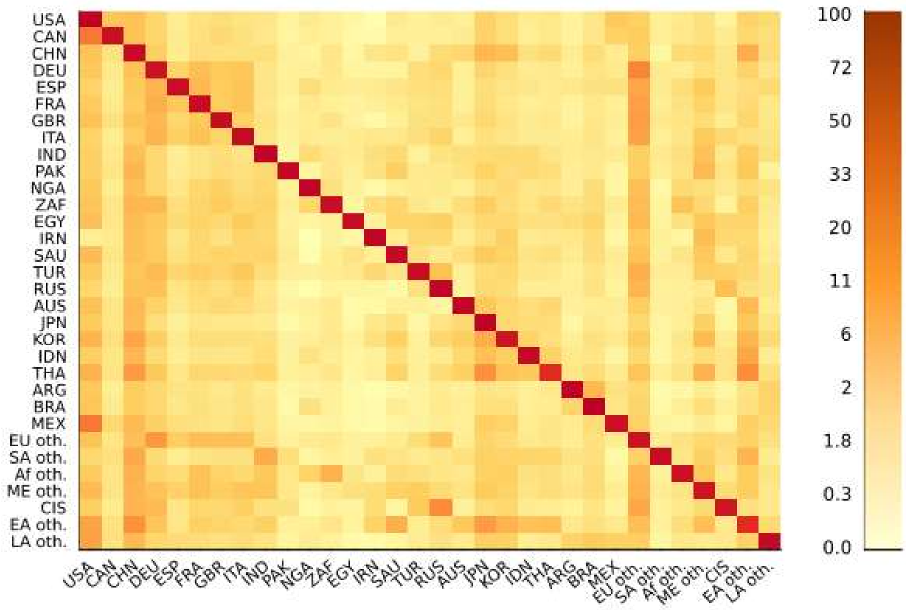


Figure 23: Trade shares as mesured in Conte et al. (2022)

A.1.2 Climate system and pulse experiment

This pulse experiment, from Dietz et al. (2021), summarizes how our climate model should be calibrated to replicate larger scale IAMs like CMIP5.

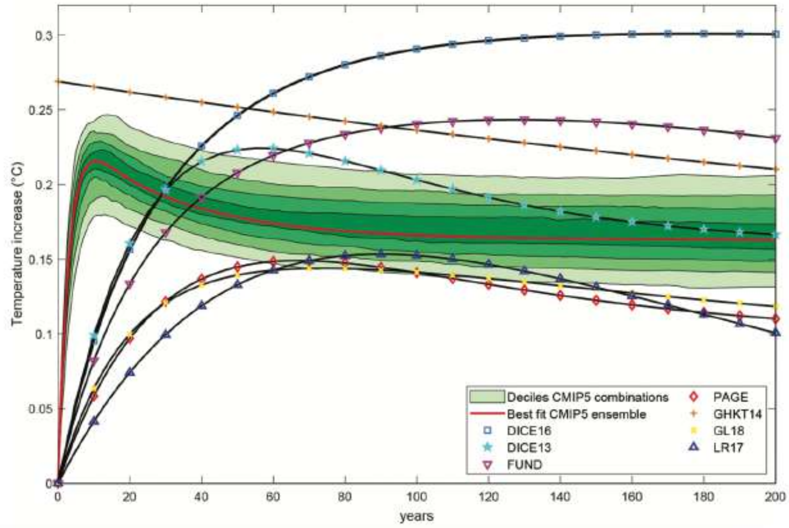


Figure 24: Pulse experiment: $\mathcal{E} = 100GtCO_2$ at $t = 0$
Comparison across different IAMs, Dietz et al. (2021)

B Model

B.1 Model environment

- Gov. policies $t_i = \{\mathrm{t}^{\varepsilon}_{it}, \mathrm{t}^{b}_{ijt}, \mathrm{t}^{ls}_{it}\}$
- State: $s_i = \{k_{it_0}, T_{it}, \mathcal{R}_{it}\}_t$,
- Agents (HH/firms) controls $c_i = \{c_{it}, c_{ijt}, k_{it}, e^f_{it}, e^c_{it}, e^r_{it}, e^x_{it}\}_t$
- Eq prices: $p = \mathrm{p}_{it}, w_{it}, q^f_t, q^c_{it}, q^r_{it}$
- Lagrange multipliers / costates: $\lambda_i = \{\lambda^w_{it}, \lambda^S_{it}, \lambda^T_{it}\}$
- Local welfare vs Global welfare

$$\mathcal{U}_i = \max_{c_i} \int_0^T e^{-\bar{\rho}_i t} \mathcal{P}_i u(c_{it}, T_{it}) dt$$
$$\mathcal{W} = \max_{\{c_i\}_i} \sum_{i \in \mathbb{I}} \omega_i \mathcal{P}_i \int_0^T e^{-\bar{\rho}_i t} u(c_{it}, T_{it}) dt$$

- Assumptions:
  - weak separability of utility

$$u(c_{it}, T_{it}) = \tilde{u}(c_i)\mathcal{D}^u(T_{it}) = \frac{c_{it}^{1-\eta}}{1-\eta}\big(\tilde{\mathcal{D}}^u(T_{it})\big)^{1-\eta}$$

  - DICE damage functions $\mathcal{D}^u_i(T)$ and $\mathcal{D}^y_i(T)$

$$\mathcal{D}^u_i(T) = e^{-\frac{\gamma^c}{2}(T-T^\star_i)^2} \qquad \Rightarrow \qquad \mathcal{D}^{u\prime}_i(T) = -\mathcal{D}^u_i(T)\gamma^c(T-T^\star_i)$$

  - CES demand / Armington structure, price of imports $p_{ij} = \tau_{ij}(1+\mathrm{t}^b_{ij})\mathrm{p}_j$

$$u(\{c_{ij}\}_j, T_i)) = u(c_i, T_i) \qquad c_i = \Big(\sum_j a_{ij}^{\frac{1}{\theta}} c_{ij}^{\frac{\theta-1}{\theta}}\Big)^{\frac{\theta}{\theta-1}}$$

- Heterogenous discount rate: $\bar{\rho}_i = \rho - n_i - (1-\eta)\bar{g}_i$
- Climate system:

$$\dot{\mathcal{S}}_t = \zeta_t \mathcal{E}_t - \delta_s \mathcal{S}_t = e^{-ot}[\xi^f e^f_{it} + \xi^c e^c_{it}] - \delta_s \mathcal{S}_t$$
$$\mathcal{S}_t = \mathcal{S}_0 e^{-\delta_s t} + \int_0^t e^{-\delta_s(t-u)} e^{(n+\bar{g}-o)u}(\xi^f e^f_{iu} + \xi^c e^c_{iu}) du$$
$$T_{it} = T_{it_0} + \Delta_i \chi \mathcal{S}_t$$

with $o = g_\zeta \mathbb{1}\{t > 2100\}$ the rate of growth of additional abatement due to CCS after 2100.

B.2 Summary, model setting

- Expenditure by household:

$$\sum_j c_{ijt}\tau_{ij}(1+\mathrm{t}^b_{ij})\mathrm{p}_{jt} = c_{it}\mathbb{P}_{it}$$

- Final good firm problem: pay for labor and capital and buys three energy inputs:

$$\pi^g_{it} = \mathrm{p}_i \mathcal{D}^y(T_{it}) z_{it} F(\ell_i, k_{it}, e^f_{it}, e^c_{it}, e^r_{it}) - w_{it}\ell_i - r_t k_{it} - (q^f_t + \xi^f \mathrm{t}^s_{it}) e^f_{it} - (q^c_t + \xi^c \mathrm{t}^s_{it}) e^c_{it} - q^r_{it} e^r_{it} = 0$$

- Fossil Energy firm profit:

$$\mathcal{P}_i \pi^f_{it} = q^f_t e^x_{it} - \mathcal{C}^f_i(e^x_{it}, \mathcal{R}_{it}) \mathcal{P}_i$$

- Budget constraint for the household:
  replace labor income, and divide by price index (analog of "real" quantities). Reminder that capital expenditure are made in the final consumption good bundle.

$$\begin{aligned} c_{it}\mathbb{P}_{it} + \big(\dot{k}_{it} + (n_i + \bar{g}_i + \delta)\big)\mathbb{P}_{it} &= w_{it}\ell_i + r_t k_{it} + \pi^f_t + \mathrm{t}^{ls}_i \\ 0 &= \frac{\mathrm{p}_{it}}{\mathbb{P}_{it}} \mathcal{D}^y(T_{it}) z_{it} F(k_{it}, e^f_{it}, e^c_{it}, e^r_{it}) - (n_i + \bar{g}_i + \delta) k_{it} \\ &\quad + \frac{1}{\mathcal{P}_i \mathbb{P}_{it}} \Big[q^f_t e^x_{it} - \nu(e^x_{it}, \mathcal{R}_{it}) \Big] - \frac{(q^f_t + \xi^f \mathrm{t}^s_{it})}{\mathbb{P}_{it}} e^f_{it} - \frac{(q^c_t + \xi^c \mathrm{t}^s_{it})}{\mathbb{P}_{it}} e^c_{it} - \frac{q^r_{it}}{\mathbb{P}_{it}} e^r_{it} - c_{it} + \frac{\mathrm{t}^{ls}_{it}}{\mathbb{P}_{it}} - \dot{k}_{it} \end{aligned}$$

- CES / Armington trade model, with price of imports $p_{ij} = \tau_{ij}(1+\mathrm{t}^b_{ij})\mathrm{p}_j$

$$u(\{c_{ij}\}_j, T_i)) = u(c_i)\, \mathcal{D}^u_i(T_i) \qquad c_i = \Big(\sum_j a_{ij}^{\frac{1}{\theta}} c_{ij}^{\frac{\theta-1}{\theta}} \Big)^{\frac{\theta}{\theta-1}}$$

$$FOC \qquad [c_{ij}] \qquad u'(c_i) \mathcal{D}^u_i(T_i)\, c_i^{\frac{1}{\theta}} a_{ij}^{\frac{1}{\theta}} c_{ij}^{-\frac{1}{\theta}} = p_{ij} \lambda^w_i$$

Price index:

$$\mathbb{P}_i = \Big(\sum_j a_{ij} p_{ij}^{1-\theta} \Big)^{\frac{1}{1-\theta}} = \Big(\sum_j a_{ij} (\tau_{ij}(1+\mathrm{t}^b_{ij})\mathrm{p}_j)^{1-\theta} \Big)^{\frac{1}{1-\theta}}$$

Demand system:

$$\Rightarrow \qquad \frac{c_{ij}}{c_i} = a_{ij} \big(\frac{\mathbb{P}_i}{p_{ij}}\big)^{\theta} \qquad \Rightarrow \qquad \frac{c_{ij} p_{ij}}{c_i \mathbb{P}_i} = a_{ij} \big(\frac{p_{ij}}{\mathbb{P}_i}\big)^{1-\theta}$$

$$s_{ij} = \frac{c_{ij} p_{ij}}{c_i \mathbb{P}_{it}} = a_{ij} \frac{p_{ij}^{1-\theta}}{\sum_k a_{ik} p_{ik}^{1-\theta}} = a_{ij} \frac{((1+\mathrm{t}_{ij})\tau_{ij}\mathrm{p}_j)^{1-\theta}}{\sum_k a_{ik} ((1+\mathrm{t}_{ik})\tau_{ik}\mathrm{p}_k)^{1-\theta}}$$

Aggregating, we obtain the marginal value of "wealth":

$$\lambda^w_i = \frac{u'(c_i) \mathcal{D}^u_i(T_i)}{\mathbb{P}_i}$$

B.2.1 Market clearing

We reexpress the market clearing, for good i in expenditure terms. The time subscripts are removed for conciseness.

$$\mathcal{P}_i y_i = \overline{\mathcal{D}}_i(\{T_{it}\}) z_i F(k_i, e_i) = \sum_{k \in \mathbb{I}} \tau_{ki} \mathcal{P}_k c_{ki} + \sum_{k \in \mathbb{I}} \mathcal{P}_k \tau_{ki} (x^f_{ki} + x^c_{ki} + x^r_{ki} + x^k_{ki})$$

Rewriting in expenditure, multiplying by p_j, and using the fact that the input choice is identical between

$$
\begin{aligned}
\mathcal{P}_i y_i \mathrm{p}_i &= \sum_{k\in\mathbb{I}} \mathcal{P}_k \tau_{ki}(x^f_{ki} + x^c_{ki} + x^r_{ki} + x^k_{ki})\mathrm{p}_i \\
&= \sum_{k\in\mathbb{I}} \mathcal{P}_k \frac{1}{1+\mathrm{t}^b_{ki}} \tau_{ki}(1+\mathrm{t}^b_{ki})\mathrm{p}_i(c_{ki} + x^f_{ki} + x^c_{ki} + x^r_{ki} + x^k_{ki}) \\
&= \sum_{k\in\mathbb{I}} \mathcal{P}_k \frac{1}{1+\mathrm{t}^b_{ki}} s_{ki}(c_i + \mathcal{C}^f_i(e^x_i) + z^c_i e^c_i + z^r_i e^r_i + (n_k + \bar{g}_k + \delta)k_k)\mathbb{P}_k
\end{aligned}
$$

Using the budget constraint to replace c_i

$$
\begin{aligned}
\mathcal{P}_i y_i \mathrm{p}_i = \sum_{k\in\mathbb{I}} \frac{\mathcal{P}_k s_{ki}}{1+\mathrm{t}^b_{ki}} \Big(& y_k \mathrm{p}_k - (q^f + \xi^f \mathrm{t}^\varepsilon_k) e^f_k - (q^c_k + \xi^c \mathrm{t}^\varepsilon) - q^r_k e^r_k - (n_i + \bar{g}_i + \delta)k_{it}\mathbb{P}_k + \Big[q^f e^x_k - \mathcal{C}^f_k(e^x_k)\mathbb{P}_k\Big] + \mathrm{t}^{ls}_k \\
& + \mathcal{C}^f_i(e^x_k)\mathbb{P}_k + z^c_k e^c_k \mathbb{P}_k + z^r_k e^r_k \mathbb{P}_k + (n_k + \bar{g}_k + \delta)k_k \mathbb{P}_k \Big) \\
\mathcal{P}_i y_i \mathrm{p}_i = \sum_{k\in\mathbb{I}} \frac{\mathcal{P}_k s_{ki}}{1+\mathrm{t}^b_{ki}} \Big(& y_k \mathrm{p}_k + q^f(e^x_k - e^f_k) + \widetilde{\mathrm{t}}^{ls}_k \Big) = \sum_{k\in\mathbb{I}} \frac{s_{ki}}{1+\mathrm{t}^b_{ki}} \mathcal{P}_k(\widetilde{v}_k + \widetilde{\mathrm{t}}^{ls}_k)
\end{aligned}
$$

where $\widetilde{v}_k = y_k \mathrm{p}_k + q^f(e^x_k - e^f_k)$ represent the revenues of country k in terms of production and energy export and the lump-sum transfers of the tariffs: $\widetilde{\mathrm{t}}^{ls}_k = \sum_j \mathrm{t}^b_{kj}\tau_{kj}(c_{kj} + x^f_{kj} + x^c_{kj} + x^r_{kj} + x^k_{kj})$.

We see that the lump-sum transfer also depends on the quantities. To be able to express the market in expenditure, we solve:

$$
\begin{aligned}
\mathcal{P}_i p_i y_i &= \sum_i \frac{s_{ki}}{1+\mathrm{t}^b_{ki}} \mathcal{P}_k[\widetilde{v}_k + \widetilde{\mathrm{t}}^{ls}_k] \\
v_k &:= \widetilde{v}_k + \widetilde{\mathrm{t}}^{ls}_k \\
\widetilde{\mathrm{t}}^{ls}_k &= \sum_j \mathrm{t}^b_{kj}\tau_{kj}\underbrace{(c_{kj} + x^f_{kj} + x^r_{kj} + x^c_{kj} + x^k_{kj})}_{=:x_{kj}}\mathrm{p}_k \qquad\qquad x_{kj} = \frac{s_{kj} v_k}{(1+\mathrm{t}_{kj})\tau_{kj}\mathrm{p}_k}
\end{aligned}
$$

As a result, we solve the "fixed point" for v_i as follow:

$$
\begin{aligned}
v_i &= \widetilde{v}_i + v_i \sum_j \frac{\mathrm{t}^b_{ij}}{1+\mathrm{t}^b_{ij}} s_{ij} \\
v_i &= \frac{1}{1 - \sum_j \frac{\mathrm{t}^b_{ij}}{1+\mathrm{t}^b_{ij}} s_{ij}} \widetilde{v}_i = m_i \widetilde{v}_i \qquad \text{with} \qquad m_i = \frac{1}{1 - \sum_j \frac{\mathrm{t}^b_{ij}}{1+\mathrm{t}^b_{ij}} s_{ij}}
\end{aligned}
$$

To conclude, the market clearing writes:

$$
\mathcal{P}_i p_i y_i = \sum_i \frac{s_{ki}}{1+\mathrm{t}^b_{ki}} \mathcal{P}_k m_k \Big[y_k \mathrm{p}_k + q^f(e^x_k - e^f_k)\Big]
$$

B.3 Making the dynamic model stationary

We solve the optimization problem of the household – who own the firms. This is a dynamic problem, since climate changes over time, with emissions $\mathcal{E}_t$. We express the Lagrangian for the problem as a finite-horizon problem, and we take the finite horizon $T \to \infty$.

$$\mathcal{L}(s_i, c_i, \lambda_i) = \int_0^T e^{-\bar{\rho}_i t} \mathcal{P}_i u(c_{it}, T_{it}) dt + \int_0^T e^{-\bar{\rho}_i t} \mathcal{P}_i \lambda^w_{it} \Big(\mathrm{p}_i \mathcal{D}^y(T_{it}) z_{it} F(k_{it}, e^f_{it}, e^c_{it}, e^r_{it}) - (n_i + \bar{g}_i + \delta) k_{it}$$
$$\frac{1}{\mathcal{P}_i} \Big[q^f_t e^x_{it} - \nu(e^x_{it}, \mathcal{R}_{it}) \Big] - (q^f_t + \xi^f \mathrm{t}^s_{it}) e^f_{it} - (q^c_t + \xi^c \mathrm{t}^s_{it}) e^c_{it} - q^r_{it} e^r_{it} - c_{it} \mathbb{P}_{it} + \mathrm{t}^{ls}_i - \dot{k}_{it} \Big) dt$$
$$+ \int_0^T e^{-\bar{\rho}_i t} \mathcal{P}_i \lambda^S_{it} \Big[\mathcal{S}_0 e^{-\delta_s t} + \int_0^t e^{-\delta_s(t-u)} e^{(n+\bar{g}-o)u} (\xi^f e^f_{iu} + \xi^c e^c_{iu}) du - \mathcal{S}_t \Big] dt$$
$$+ \int_0^T e^{-\bar{\rho}_i t} \mathcal{P}_i \lambda^T_{it} \Big(\Delta_i \chi \mathcal{S}_t - (T_{it} - T_{it_0}) \Big) dt$$

We impose the "constraint" that all the economic controls need to be constant over time, $c_{it} = c_i \ \forall i$. As a result, all the equilibrium prices are also constant over time $p_{it} = p_i \ \forall i$.

$$\mathcal{L}(s_i, c_i, \lambda_i) = \mathcal{P}_i u(c_i) \underbrace{\int_0^T e^{-\bar{\rho}_i t} \mathcal{D}^u(T_{it}) dt}_{=\overline{\mathcal{D}}^u(\{T_{it}\}_t)} + \mathcal{P}_i \Big(\int_0^T e^{-\bar{\rho}_i t} \lambda^w_{it} \mathcal{D}^y(T_{it}) dt \Big) \mathrm{p}_i z_{it} F(k_i, e^f_i, e^c_i, e^r_i) \ - \int_0^T e^{-\bar{\rho}_i t} \dot{k}_{it} \lambda^w_{it} dt$$
$$+ \frac{1}{\mathcal{P}_i} [q^f e^x_i - \nu(e^x_i, \mathcal{R}_i)] \underbrace{\int_0^T e^{-\bar{\rho}_i t} \lambda^w_{it} dt}_{=\overline{\lambda}^w_i} - \mathcal{P}_i \overline{\lambda^w_i} \Big((q^f + \xi^f \mathrm{t}^\varepsilon_i) e^f_i + (q^c + \xi^c \mathrm{t}^\varepsilon_i) e^c_i + q^r_i e^r_i + (n_i + \bar{g}_i + \delta) k_i + c_i \mathbb{P}_i + \mathrm{t}^{ls}_i \Big)$$
$$+ \mathcal{P}_i \int_0^T e^{-\bar{\rho}_i t} \lambda^S_{it} \mathcal{S}_t dt + \mathcal{P}_i \int_0^T e^{-\bar{\rho}_i t} \lambda^T_{it} \Big(\Delta_i \chi \mathcal{S}_t - (T_{it} - T_{it_0}) \Big) dt$$

with $\mathcal{S}_t = \mathcal{S}_0 e^{-\delta_s t} + \int_0^t e^{-\delta_s(t-u)} e^{(n+\bar{g}-o)u} (\xi^f e^f_{iu} + \xi^c e^c_{iu}) du$

Optimality conditions:

- Consumption $[c_i]$:

$$u'(c_i) \overline{\mathcal{D}}^u(\{T_{it}\}_t) = \overline{\lambda}^w_{it} \mathbb{P}_i$$

- Energy choices $[e^k_i]$, for $k \in \{f, c, r\}$:

$$\overline{\lambda}^w_i (\mathrm{p}_i MPe^f_i - q^f + \xi^f \mathrm{t}^\varepsilon_i)) = 0$$
$$\overline{\lambda}^w_i (\mathrm{p}_i MPe^c_i - q^c + \xi^c \mathrm{t}^\varepsilon_i) = 0$$
$$\overline{\lambda}^w_i (\mathrm{p}_i MPe^r_i - q^r) = 0$$

- Capital choice $[k_{it}]$

$$\dot{\lambda}^w_{it} = \lambda^w_{it} (\mathrm{p}_i MPk_i - \delta - \eta \bar{g}_i - \rho)$$
$$\mathrm{p}_i MPk_i - \delta = \bar{r} = \eta \bar{g}_i + \rho \qquad \Rightarrow \qquad \dot{\lambda}^w_{it} = 0 \quad \& \quad \lambda^w_{it} = \lambda^w_{it'} = \lambda^w_i$$
$$\overline{\lambda}^w_{iT} = \int_0^T e^{-\bar{\rho}_i t} \lambda^w_{it} dt = \frac{1}{\bar{\rho}_i} (1 - e^{-\bar{\rho}_i T}) \lambda^w_i$$

- Fossil production choice $[e^x_i]$:

$$\overline{\lambda}^w_i (q^f - \mathcal{C}_{e^x}(e^x_i, \mathcal{R}_i) \mathbb{P}_i) = 0$$

- Stock of carbon in the atmosphere $[\mathcal{S}_t]$

$$\lambda_{it}^{S} = \Delta_i \chi \lambda_{it}^{T}$$

- Local temperatures $[T_i]$

$$\begin{aligned}\lambda_{it}^{T} &= \mathcal{D}^{u'}(T_{it})u(c_i) + \mathcal{D}^{y'}(T_{it})z_i F(k_i, e_i)\lambda_i^w \\ &= -(T_{it} - T_i^{\star})(\gamma_i^c c_{it} + \gamma_i^y y_{it})\lambda_i^w \qquad [w/\text{DICE damage fcts + CRRA pref}]\end{aligned}$$

- Market clearing, energy

$$\sum_{i\in\mathbb{I}} e_i^x = \sum_{i\in\mathbb{I}} \mathcal{P}_i e_i^f \qquad e_i^c = \bar{e}_i^c \qquad e_i^r = \bar{e}_i^r$$

- Market clearing, good i

$$\begin{aligned}y_i = \overline{\mathcal{D}}_i(\{T_{it}\})z_i F(k_i, e_i) &= \sum_{k\in\mathbb{I}} \tau_{ki} c_{ki} + \sum_{k\in\mathbb{I}} \tau_{ki}(x_{ki}^f + x_{ki}^c + x_{ki}^r) \\ \mathrm{p}_i \underbrace{y_i}_{=\mathcal{D}(T_i)z_i F(\cdot)} &= \sum_{k\in\mathbb{I}} \frac{s_{ki}}{1+\mathrm{t}_{ki}^b}\left(\mathrm{p}_k y_k + q^f(e_k^x - e_k^f) + \mathrm{t}_k^{ls}\right)\end{aligned}$$

B.3.1 Social Cost of Carbon and present discounted value of damages

In Integrated Assessment, we want to measure the present-discounted value of damages – for country i – of one ton of carbon emitted in the atmosphere at time $\mathcal{S}_u$ over all its "lifetime" in the atmosphere $t \in [u, T]$

As a result, with discounting :

$$\begin{aligned}\lambda_{iu}^{S,pv} &= \int_u^T e^{-\bar{\rho}_i(t-u)} e^{-\delta_s(t-u)} \lambda_{it}^S dt = \int_u^T e^{-(\bar{\rho}_i+\delta_s)(t-u)} \Delta_i \chi \lambda_{it}^T dt \\ &= u(c_i)\Delta_i \chi \int_u^T e^{-(\bar{\rho}_i+\delta_s)(t-u)} \mathcal{D}^{u'}(T_{it})dt + \lambda_i^w z_i F(k_i, e_i)\Delta_i \chi \int_u^T e^{-(\bar{\rho}_i+\delta_s)(t-u)} \mathcal{D}^{y'}(T_{it})dt\end{aligned}$$

If the cost of carbon is constant (supposing that temperature is stable $T_{it} \to \bar{T}_i$) then the welfare cost of one ton of carbon writes:

$$\lim_{T_{it}\to\bar{T}_i} \lambda_{iu}^{S,pv} = u(c_i)\frac{\chi\Delta_i}{\bar{\rho}_i + \delta_s}(1 - e^{-(\bar{\rho}_i+\delta_s)(T-u)})\mathcal{D}^{u'}(\bar{T}_i) + \lambda_i^w z_i F(k_i, e_i)\frac{\chi\Delta_i}{\bar{\rho}_i + \delta_s}\mathcal{D}^{y'}(\bar{T}_i)(1 - e^{-(\bar{\rho}_i+\delta_s)(T-u)})$$

$$\lim_{T_{it}\to\bar{T}_i, T\to\infty} \lambda_{iu}^{S,pv} = \bar{\lambda}_i^S = \frac{\chi\Delta_i}{\bar{\rho}_i + \delta_s}\Big(u(c_i)\mathcal{D}^{u'}(\bar{T}_i) + \mathcal{D}^{y'}(\bar{T}_i)\lambda_i^w z_i F(k_i, e_i)\Big)$$

The local cost of carbon LCC_i of emitting one ton $[\varepsilon_{it}]$ summarizes the damages of a ton emitted – per effective capita unit! – at time t by country i. It accounts for the damages occurred between t and T. We measure it in monetary unit by dividing it by marginal value of wealth λ_{it}^w at time t

– the time of the emission.

$$
\begin{aligned}
LCC_{it} = -\frac{\frac{\partial \mathcal{V}_{it}}{\partial \varepsilon_{it}}}{\frac{\partial \mathcal{V}_{it}}{\partial c_{it}}} &= -\frac{1}{\lambda^w_{it}} e^{(n+\bar{g})t} \int_t^T e^{-(\bar{\rho}_i+\delta_s)(s-t)} \mathcal{P}_i \lambda^S_{is} ds \\
&= -e^{(n+\bar{g})t} \frac{\lambda^{S,pv}_{it}}{\lambda^w_{it}}
\end{aligned}
$$

Now, we were trying to measure the model in stationary form, by taking the present discounted value of the welfare costs and the marginal value of wealth. The stationary local cost of carbon LCC_i writes:

$$
LCC_i = -\frac{\overline{\lambda}^s_i}{\overline{\lambda}^w_i} = -\frac{\int_0^T e^{-\bar{\rho}_i t} e^{(n+\bar{g})t} \lambda^{S,pv}_{it} dt}{\int_0^T e^{-\bar{\rho}_i t} \lambda^w_{it} dt}
$$

The numerator can be rearranged:

$$
\begin{aligned}
\int_0^T e^{-\bar{\rho}_i t} \lambda^{S,pv}_{it} dt &= \int_0^T e^{-\bar{\rho}_i t} \int_t^T e^{-\bar{\rho}_i (s-t)} e^{-\delta_s (s-t)} \lambda^S_{is} ds \ dt \\
&= \int_0^T e^{-\bar{\rho}_i t} \int_t^T e^{-(\bar{\rho}_i+\delta_s)(s-t)} \Delta_i \chi \lambda^T_{is} ds \ dt \\
&= \Delta_i \chi \mathcal{P}_i \int_0^T \int_t^T e^{-\bar{\rho}_i s} e^{-\delta_s (s-t)} \lambda^T_{is} ds \ dt \\
&= \Delta_i \chi \mathcal{P}_i \Big[u(c_i) \int_0^T \int_t^T e^{-\bar{\rho}_i s} e^{-\delta_s (s-t)} \mathcal{D}^{u'}(T_{is}) ds \ dt + \lambda^w_i z_i F(k_i, e_i) \int_0^T \int_t^T e^{-\bar{\rho}_i s} e^{-\delta_s (s-t)} \mathcal{D}^{y'}(T_{is}) ds \ dt \Big]
\end{aligned}
$$

We see that the "dynamic marginal cost" can be isolated from the other economic variables $y_i, c_i, e_i, \lambda^w_i$. These are the objects we will use when considering optimal climate policy.

Example for policy

To give an example for policy, remember that the LCC_i summarize the future cost of climate change:

$$
LCC_i = -\frac{\overline{\lambda}^s_i}{\overline{\lambda}^w_i} = -\frac{\int_0^T e^{-\bar{\rho}_i t} e^{(n+\bar{g})t} \lambda^{S,pv}_{it} dt}{\int_0^T e^{-\bar{\rho}_i t} \lambda^w_{it} dt}
$$

Suppose one conduct the unilateral climate policy, choosing yearly oil consumption per (effective) capita $[e^f_i]$, internalizing the climate externality ε_u at every period u, and considering that the revenue of the carbon tax is redistributed lump-sum $\mathrm{t}^{ls}_i = \xi^f \mathrm{t}^s_i e^f_i$. The FOC for e^f_i becomes:

$$
\begin{aligned}
& \mathcal{P}_i \overline{\lambda}^w_i (\xi^f \mathrm{t}^s_i) + \mathcal{P}_i \int_0^T \int_u^T e^{-\bar{\rho}_i t} \lambda^S_{it} e^{-\delta_s (t-u)} e^{(n_i+\bar{g}_i)u} \xi^f dt \ du = 0 \\
\Rightarrow \qquad & \mathrm{t}^s_i = LCC_i
\end{aligned}
$$

the optimal unilateral carbon tax is the local cost of carbon for country i. This is the standard Pigouvian result and we will see how to conduct the policy at the global level and accounting for redistribution effects and endogenous participation.

B.3.2 Summary: Climate model

Here, we summarize the climate model, and express the present-discounted damages $\overline{\mathcal{D}}(\mathcal{E})$, normalized by discounting $\bar{\rho}_i = \rho - n - (1-\eta)\bar{g}_i$

$$\mathcal{S}_t = \mathcal{S}_0 e^{-\delta_s t} + \int_0^t e^{-\delta_s(t-u)} e^{(n+\bar{g})u} \mathcal{E} du \qquad T_{it} = T_{it_0} + \Delta_i \chi \mathcal{S}_t$$

$$\overline{\mathcal{D}}^u(\mathcal{E}) = \frac{\bar{\rho}_i}{1-e^{-\bar{\rho}_i T}} \int_0^T e^{-\bar{\rho}_i t} \mathcal{D}^u(T_{it}) dt \qquad \overline{\mathcal{D}}^y(\mathcal{E}) = \frac{\bar{\rho}_i}{1-e^{-\bar{\rho}_i T}} \int_0^T e^{-\bar{\rho}_i t} \mathcal{D}^y(T_{it}) dt$$

$$\lambda_i^w = \frac{\bar{\rho}_i}{1-e^{-\bar{\rho}_i T}} \overline{\lambda}_{iT}^w = \frac{\bar{\rho}_i}{1-e^{-\bar{\rho}_i T}} \int_0^T e^{-\bar{\rho}_i t} \lambda_{it}^w dt$$

$$\overline{\lambda}_{iT}^w LCC_i = \Delta_i \chi \mathcal{P}_i \Big[u(c_i) \int_0^T \int_t^T e^{-\bar{\rho}_i s} e^{-\delta_s(s-t)} \mathcal{D}^{u'}(T_{it}) ds\ dt + \lambda_i^w z_i F(k_i, e_i) \int_0^T \int_t^T e^{-\bar{\rho}_i s} e^{-\delta_s(s-t)} \mathcal{D}^{y'}(T_{it}) ds\ dt \Big]$$

$$\mathcal{E} = \sum_{i \in \mathbb{I}} \mathcal{P}_i (\xi^f e_i^f + \xi^c e_i^c)$$

B.3.3 Summary: Economic model

$$u'(c_i) \overline{\mathcal{D}}^u(\mathcal{E}) = \overline{\lambda}_{it}^w \mathbb{P}_i$$

$$\mathrm{p}_i MPe_i^f = q^f + \xi^f \mathrm{t}_i^{\varepsilon} \qquad \mathrm{p}_i MPe_i^c = q^c + \xi^c \mathrm{t}_i^{\varepsilon}$$

$$\mathrm{p}_i MPe_i^r = q^r \qquad \mathrm{p}_i MPk_i - \delta = \bar{r} = \eta \bar{g}_i + \rho$$

$$q^f = \mathcal{C}_{e^x}(e_i^x, \mathcal{R}_i) \mathbb{P}_i$$

$$\sum_{i \in \mathbb{I}} e_i^x = \sum_{i \in \mathbb{I}} \mathcal{P}_i e_i^f \qquad e_i^c = \bar{e}_i^c = z_i^c \mathbb{P}_i \qquad e_i^r = \bar{e}_i^r = z_i^r \mathbb{P}_i$$

$$\overline{\mathcal{D}}_i(\mathcal{E}) z_i F(k_i, e_i) = \sum_{k \in \mathbb{I}} \tau_{ki} c_{ki} + \sum_{k \in \mathbb{I}} \tau_{ki} (x_{ki}^f + x_{ki}^c + x_{ki}^r + x_{ki}^k)$$

$$\mathcal{E} = \sum_i \mathcal{P}_i (\xi^f e_i^f + \xi^c e_i^c)$$

C Policy

In this section, we provide details on the three policy benchmark, considered in section Section 5.1 and Section 5.2 of the main text. We cover first the optimal allocation when the planner only accounts for resources constraints – the First-Best. Then, we turn to the Ramsey allocation when the planner is constrained and is not allowed cross-countries transfers nor bilateral tariffs, and can only choose carbon taxation. In the last section, we consider unilateral policy, which is a benchmark policy in Nash equilibrium when countries do not cooperate and choose their carbon taxation and trade tariffs to maximize their country's utility.

C.1 First Best

In this allocation, the planner chooses $\boldsymbol{x} = \{c_{ij}, x_{ij}^{\ell}, e_i^{\ell}\}$, i.e. the traded good for consumption c_{ij}, for energy inputs for the production of in fossil x_{ij}^f, coal x_{ij}^c, non-carbon x_{ij}^r or capital x_{ij}^k, and the energy demand, in fossil e_i^f, coal e_i^c and non-carbon e_i^r.

The welfare criterion the planner maximizes is:

$$\mathcal{W} = \sum_i \omega_i \mathcal{P}_i u(\{c_{ij}\}_j) \overline{\mathcal{D}}^u(\mathcal{E})$$

The Planner Lagrangian – in the First-Best allocation – writes:

$$\begin{aligned}\mathcal{L}(\boldsymbol{x}, \boldsymbol{\lambda}) = &\sum_i \omega_i \mathcal{P}_i u(\{c_{ij}\}_j) \overline{\mathcal{D}}^u(\mathcal{E}) + \overline{\lambda}\mu^f \Big[\sum_{i\in\mathbb{I}} e_i^x - \mathcal{P}_i e_i^f\Big] + \sum_{\mathbb{I}} \overline{\lambda}\mu_i^c [\bar{e}_i^c - \mathcal{P}_i e_i^c] + \sum_{\mathbb{I}} \overline{\lambda}\mu_i^r [\bar{e}_i^r - \mathcal{P}_i e_i^r] \\ &+ \sum_{\mathbb{I}} \mathcal{P}_i \omega_i \phi_i^{\varepsilon} \Big(\mathcal{E} - \sum_{i\in\mathbb{I}} \mathcal{P}_i (\xi^f e_i^f + \xi^c e_i^c)\Big) + \sum_i \omega_i \mu_i \overline{\lambda} \Big[\mathcal{P}_i z_i \overline{\mathcal{D}}^y(\mathcal{E}) F(\ell_i, k_i, e_i) - \sum_{k\in\mathbb{I}} \mathcal{P}_k \tau_{ki} (c_{ki} + x_{ki}^f + x_{ki}^c + x_{ki}^r)\Big]\end{aligned}$$

where we rescale the multipliers for the market clearing for good $\overline{\lambda}\mu_i$, for fossil energy $\overline{\lambda}\mu_i^f$, coal energy $\overline{\lambda}\mu_i^c$ and non-carbon energy $\overline{\lambda}\mu_i^r$ by the constant $\overline{\lambda}$ to simplify the comparison with the decentralized equilibrium.

The problem being convex, we write the optimality conditions for each of the controls:

- Consumption

$$[c_{ij}] \qquad \omega_i \mathcal{P}_i u'(c_i) c_i^{1/\theta} a_{ij}^{1/\theta} c_{ij}^{-1/\theta} = \mathcal{P}_i \tau_{ij} \omega_j \mu_j \overline{\lambda}$$

$$c_{ij} = a_{ij} c_i \Big(\tau_{ij} \omega_j \mu_j \frac{\overline{\lambda}}{\omega_i u'(c_i)}\Big)^{-\theta}$$

To get the ideal "price" index, we aggregate:

$$\Rightarrow \qquad c_{ij}^{(\theta-1)/\theta} = [\omega_i u'(c_i)]^{\theta-1} c_i^{(\theta-1)/\theta} a_{ij}^{(\theta-1)/\theta} (\tau_{ij}\omega_j\mu_j\overline{\lambda})^{1-\theta}$$

$$\Rightarrow \qquad c_i^{\frac{\theta-1}{\theta}} = \sum_j a_{ij}^{1/\theta} c_{ij}^{(\theta-1)/\theta} = [\omega_i u'(c_i)]^{\theta-1} c_i^{(\theta-1)/\theta} \sum_j a_{ij} (\tau_{ij}\omega_j\mu_j\overline{\lambda})^{1-\theta}$$

$$\omega_i u'(c_i) = \overline{\lambda} \Big[\sum_j a_{ij} (\tau_{ij}\omega_j\mu_j)^{1-\theta}\Big]^{\frac{1}{1-\theta}}$$

- Energy inputs $\bar{e}_i^\ell$ and x_{ij}^ℓ

$$
\begin{aligned}
[x_{ij}] \qquad & \overline{\lambda}\mu_i^\ell g'(x_i^\ell) x_i^{1/\theta} a_{ij}^{1/\theta} x_{ij}^{-1/\theta} = \tau_{ij}\omega_j\mu_j\overline{\lambda} \\
\Rightarrow \qquad & x_{ij}^{(\theta-1)/\theta} = [\mu_i^\ell g'(x_i^\ell)]^{\theta-1}(x_i^\ell)^{(\theta-1)/\theta} a_{ij}^{(\theta-1)/\theta}(\tau_{ij}\omega_j\mu_j)^{1-\theta} \\
\Rightarrow \qquad (x_i^\ell)^{\frac{\theta-1}{\theta}} = \sum_j a_{ij}^{1/\theta}\,(x_{ij}^\ell)^{(\theta-1)/\theta} & = [\mu_i^\ell]^{\theta-1}(x_i^\ell)^{(\theta-1)/\theta}\sum_j a_{ij}(\tau_{ij}\omega_j\mu_j)^{1-\theta} \\
& \mu_i^\ell g'(x_i^\ell) = \big[\sum_j a_{ij}(\tau_{ij}\omega_j\mu_j)^{1-\theta}\big]^{\frac{1}{1-\theta}}
\end{aligned}
$$

- Energy demand e_i^ℓ

$$
\begin{aligned}
[e_i^f] \qquad & \omega_i \mathcal{P}_i \mu_i \overline{\lambda} MPe_i^f = \mathcal{P}_i\overline{\lambda}\mu^f + \mathcal{P}_i\xi^f\lambda^{\mathcal{S}} \\
\Rightarrow \qquad & \omega_i\mu_i MPe_i = \mu^f + \xi^\ell\frac{\phi^{\mathcal{E}}}{\overline{\lambda}} \\
[e_i^c] \qquad & \omega_i\mu_i MPe_i^f = \mu_i^c + \xi^c\frac{\phi^{\mathcal{E}}}{\overline{\lambda}} \\
[e_i^r] \qquad & \omega_i\mu_i MPe_i^f = \mu_i^r
\end{aligned}
$$

- Climate damage through carbon emissions $\mathcal{E}$

$$
[\mathcal{E}] \qquad \phi^\varepsilon = \sum_{\mathbb{I}}\mathcal{P}_i\omega_i\phi_i^\varepsilon = -\sum_{\mathbb{I}}\mathcal{P}_i\omega_i\Big[u(c_i)\overline{\mathcal{D}}^{u\prime}(\mathcal{E}) + \overline{\lambda}\mu_i\mathcal{D}_i^{y\prime}(\mathcal{E})z_iF(e_i,\ell_i)\Big]
$$

Decentralization

We now look at how this planner allocation can be decentralized in the competitive equilibrium.

First, we note the that social cost of carbon is formulated with the multipliers:

$$
SCC = -\frac{\frac{\partial\mathcal{W}}{\partial\mathcal{E}}}{\frac{\partial\mathcal{W}}{c_i}} = \frac{\phi^{\mathcal{E}}}{\overline{\lambda}}
$$

where we recognize that the multiplier $\phi^{\mathcal{E}}$ is the welfare value of one additional ton of carbon (the welfare cost comes from the minus sign), and $\overline{\lambda}$ the average marginal utility of consumption – or marginal value of wealth.

Indeed, the First-Best allocation equalizes marginal utilities through the condition:

$$
\overline{\lambda} = \frac{\omega_i u(c_i)\overline{\mathcal{D}}_i^u(\mathcal{E})}{\mathbb{P}_i} = \frac{\omega_j u(c_j)\overline{\mathcal{D}}_j^u(\mathcal{E})}{\mathbb{P}_j} \qquad \forall\ i,j\in\mathbb{I}
$$

This implies large redistribution, using lump-sum transfers, such that

$$
\begin{aligned}
c_i &= u'^{-1}\Big(\frac{\overline{\lambda}\mathbb{P}_i}{\omega_i\overline{\mathcal{D}}_i^u(\mathcal{E})}\Big), \forall i\in\mathbb{I} \\
c_i\mathbb{P}_i &= w_i\ell_i + \pi_i^f + \mathrm{t}_i^{ls}
\end{aligned}
$$

In that cases, the transfers t_i^{ls} are designed, such that the consumptions are equalized. This implies redistribution, as $\mathrm{t}_i^{ls} < 0$ for some countries and $\mathrm{t}_i^{ls} > 0$ for some other countries.

The price p_i, output subsidy t_i^y (or inputs subsidy) and tariffs t_{ij}^b in the allocation are determined such that the FOC in the goods demand (for consumption and energy inputs) are satisfied:

$$
\begin{aligned}
(1{+}\mathrm{t}_i^y)\mathrm{p}_i &= \omega_i \mu_i \\
(1{+}\mathrm{t}_{ij}^b)\mathrm{p}_j &= \omega_j \mu_j \\
\mathbb{P}_i &= \big[\sum_j a_{ij}(\tau_{ij}(1{+}\mathrm{t}_{ij}^b)\mathrm{p}_j)^{1-\theta}\big]^{\frac{1}{1-\theta}} \\
\mathbb{P}_i &= \big[\sum_j a_{ij}(\tau_{ij}\omega_j\mu_j)^{1-\theta}\big]^{\frac{1}{1-\theta}}
\end{aligned}
$$

A priori, there could be multiple sets of $\{\mathrm{t}_i^y, \mathrm{t}_{ij}^b\}$ such that these conditions are met. It can not be characterized further, because the prices p_i in the Armington model are endogenous objects that depend on the demand and market clearing of each good, and can not be expressed analytically. If the conditions above are satisfied, the energy prices are simply the multipliers :

$$
q^f = \mu^f \qquad\qquad q_i^c = \mu_i^c \qquad\qquad q_i^r = \mu_i^r
$$

Finally, the optimal tax is simply the Social Cost of Carbon (SCC)

$$
\begin{aligned}
\mathrm{t}^\varepsilon = \frac{\phi^{\mathcal{E}}}{\overline{\lambda}} &= -\sum_{\mathbb{I}} \mathcal{P}_i\omega_i\Big[\frac{u(c_i)}{u'(c_i)}\mathbb{P}_i\overline{\mathcal{D}}^{u\prime}(\mathcal{E}) + \mathcal{D}_i^{y\prime}(\mathcal{E})z_iF(e_i,\ell_i)\mathrm{p}_i\Big] > 0 \\
&= \sum_{\mathbb{I}} \mathcal{P}_i\omega_i\Big[\gamma^c c_i\mathbb{P}_i + \gamma^y y_i\mathrm{p}_i\Big]\widetilde{\mathcal{D}}_i(\mathcal{E}) > 0
\end{aligned}
$$

where the second line uses the functional form (CRRA) and our simple climate for the damage of future path of temperature with quadratic damage: $\widetilde{\mathcal{D}}(\mathcal{E}) = \bar{\rho}\int_{t_0}^{\infty}\int_t^{\infty} e^{-\bar{\rho}s}e^{-\delta_s(s-t)}(T_{is}{-}T_i^\star)dtds$. The optimal carbon tax is the Pigouvian level that summarizes the marginal cost of climate change for all countries i.

We will see now how that results changes when the transfers and other instruments (like tariffs or subsidies) are constrained and prevented to do redistribution.

C.2 Second best: Ramsey policy with constrained instruments

In this allocation, the Ramsey planner again chooses $\boldsymbol{x} = \{c_{ij}, x^\ell_{ij}, e^\ell_i\}$, i.e. the traded good for consumption c_{ij}, for energy inputs for the production of in fossil x^f_{ij}, coal x^c_{ij}, non-carbon x^r_{ij} or capital x^k_{ij}, and the energy demand, in fossil e^f_i, coal e^c_i and non-carbon e^r_i, as well as the carbon tax t^ε and the prices $\boldsymbol{p} = \{\mathrm{p}_i, q^f, q^c_i, q^r_i\}_i$. However, the allocation and prices are constrained to be a competitive equilibrium: in that case, the planner is restricted to choose controls that respect the individual optimality conditions.

We use the same multipliers: $\boldsymbol{\lambda} = \{\mu_i, \mu^c_i, \mu^r_i\}$ and μ^f, ϕ^ε for the market clearing clearing of the final goods, the coal, renewable and fossil energy, and the carbon emissions. We add the constraints that are satisfied in competitive equilibria: λ_i for the budget constraint, ϕ^c for the consumption decision, θ^ℓ_i for the production quantity (supply) choice of energy firms $\ell = f, c, r$ for fossil, coal and renewable of country i, υ^ℓ_i for the quantity (demand) of energy ℓ chosen by the good firm, η_{ij} for the consumption choice for imports j by the household in i, ϑ^ℓ_{ij} for the import choice for inputs from j for the energy firm j. Note that all the multipliers are normalized by ω_i, $\mathcal{P}_i$, and prices or quantity, to simplify optimal policies formulas.

As a result, the controls are $\boldsymbol{x} = \{c_{ij}, x^\ell_{ij}, e^\ell_i, \mathrm{p}_i, q^f, q^c_i, q^r_i, \mathrm{t}^\varepsilon\}_i$ and the multipliers are $\boldsymbol{\lambda} = \{\lambda_i, \mu_i, \mu^c_i, \mu^r_i, \mu^f, \phi^c_i, \theta^\ell_i, \upsilon^\ell_i, \phi^\varepsilon, \eta_{ij}, \vartheta^\ell_{ij}\}_{\ell,i,j}$.

We see that, the Ramsey planner, in choosing t^ε, with other instruments fixed at baseline value t^b_{ij} need to account for many redistributive effects through all the agents decisions.

$$
\begin{aligned}
\mathcal{L}(\boldsymbol{x}, \boldsymbol{\lambda}) = & \sum_{\mathbb{I}} \omega_i \mathcal{P}_i u(c_i) \overline{\mathcal{D}}^u_i(\mathcal{E}) + \sum_{\mathbb{I}} \omega_i \mathcal{P}_i \lambda_i \Big(\mathrm{p}_i \overline{\mathcal{D}}^y(\mathcal{E})_i F(\ell_i, k_i, e^f_i, e^c_i, e^r_i) + \frac{1}{\mathcal{P}_i} [q^f g^f(x^\ell_i) - \sum_j x^\ell_{ij} \tau_{ij} \mathrm{p}_j (1{+}\mathrm{t}^b_{ij})] \\
& + \sum_\ell \{ q^\ell g^\ell(x^\ell_i) - \sum_j x^\ell_{ij} \tau_{ij} \mathrm{p}_j (1{+}\mathrm{t}^b_{ij}) \} - ((q^f {+} \xi^f \mathrm{t}^\varepsilon_i) e^f_i + (q^c {+} \xi^c \mathrm{t}^\varepsilon_i) e^c_i + q^r_i e^r_i + (n_i {+} \bar{g}_i {+} \delta) k_i + c_i \mathbb{P}_i + \mathrm{t}^{ls}_i) \Big) \\
& + \sum_{\mathbb{I}} \omega_i \mathrm{p}_i \mu_i \Big(\mathcal{P}_i \overline{\mathcal{D}}^y(\mathcal{E}) z_i F(\ell_i, k_i, e^i) - \sum_{k \in \mathbb{I}} \mathcal{P}_k \tau_{ki} c_{ki} + \sum_{k \in \mathbb{I}} \tau_{ki} (x^f_{ki} + x^c_{ki} + x^r_{ki}) \Big) \\
& + \mu^f q^f [\sum_{i \in \mathbb{I}} e^x_i - \mathcal{P}_i e^f_i] + \sum_{\mathbb{I}} \omega_i \mu^c_i q^c_i (\bar{e}^c_i - \mathcal{P}_i e^c_i) + \sum_{\mathbb{I}} \omega_i \mu^r_i q^r_i (\bar{e}^r_i - \mathcal{P}_i e^r_i) \\
& + \sum_{i \in \mathbb{I}} \omega_i \mathcal{P}_i \phi^\varepsilon_i [\mathcal{E} - \sum_{i \in \mathbb{I}} \mathcal{P}_i (\xi^f e^f_i + \xi^c e^c_i)] + \sum_{\mathbb{I}} \omega_i \mathcal{P}_i \phi^c_i (\mathbb{P}_i \lambda^h_i - u'(c_i) \overline{\mathcal{D}}^u(\mathcal{E})) + \sum_{\ell \in \{f,c,r\}} \sum_{\mathbb{I}} \omega_i \theta^\ell_{it} (\mathbb{P}_{it} - q^\ell_{it} g'(x^\ell_{it})) \\
& + \sum_{i \in \mathbb{I}} \omega_i \mathcal{P}_i \Big(\upsilon^f_i [q^f {+} \xi^f \mathrm{t}^\varepsilon - \mathrm{p}_i MPe^f_i] + \upsilon^c_i [q^c_i {+} \xi^c \mathrm{t}^\varepsilon - \mathrm{p}_i MPe^c_i] + \upsilon^r_i [q^r - \mathrm{p}_i MPe^r_i] + \upsilon^k_i [\underbrace{\rho {+} \eta \bar{g}_i}_{r_i + \bar{g} + n} + \delta - \mathrm{p}_i MPk_i] \Big) \\
& + \sum_{i,j \in \mathbb{I}} \omega_i \mathcal{P}_i \eta_{ij} c_{ij} [(1{+}\mathrm{t}_{ij}) \tau_{ij} \mathrm{p}_j - \mathbb{P}_i c_i^{\frac{1}{\theta}} a_{ij}^{\frac{1}{\theta}} c_{ij}^{-\frac{1}{\theta}}] \\
& + \sum_{\ell \in \{f,c,r,k\}} \sum_{i,j \in \mathbb{I}} \omega_i \vartheta^\ell_{ij} x^\ell_{ij} [(1{+}\mathrm{t}_{ij}) \tau_{ij} \mathrm{p}_j - \mathbb{P}_i (x^\ell_i)^{\frac{1}{\theta}} a_{ij}^{\frac{1}{\theta}} (x^\ell_{ij})^{-\frac{1}{\theta}}]
\end{aligned}
$$

Let us go over the optimality conditions of the planner. Note, that – the problem being statics/stationary – the planner does not distort the consumption/saving decision of the household, which implies $\phi_i^c = 0$ – as that can be seen by optimizing over the household marginal value of wealth λ_i^h.

The optimality conditions writes:

- Consumption: c_{ij}

$$\omega_i \mathcal{P}_i u'(c_i) c_i^{\frac{1}{\theta}} a_{ij}^{\frac{1}{\theta}} c_{ij}^{-\frac{1}{\theta}} - \omega_i \mathcal{P}_i \lambda_i \tau_{ij} \mathrm{p}_j - \omega_j \mathcal{P}_i \mu_j \tau_{ij} \mathrm{p}_j + \omega_i \mathcal{P}_i c_{ij} \eta_{ij} \frac{1}{\theta} \frac{\tau_{ij}(1+\mathrm{t}_{ij})\mathrm{p}_j}{c_{ij}}(1-s_{ij}) = 0$$

$$c_{ij} = a_{ij} c_i \Big((\tau_{ij}\mathrm{p}_j)[1 + \frac{\omega_j}{\omega_i}\frac{\mu_j}{\lambda_i} - \frac{\eta_{ij}}{\lambda_i}\frac{1}{\theta}(1+\mathrm{t}_{ij}^b)(1-s_{ij})] \Big)^{-\theta} \underbrace{(\frac{u'(c_i)}{\lambda_i})^{\theta}}_{=\mathbb{P}_i}$$

$$u'(c_i) = \lambda_i \Big(\sum_j a_{ij} (\tau_{ij}\mathrm{p}_j)^{1-\theta} \underbrace{[1 + \frac{\omega_j}{\omega_i}\frac{\mu_j}{\lambda_i} - \frac{\eta_{ij}}{\lambda_i}\frac{1}{\theta}(1+\mathrm{t}_{ij})(1-s_{ij})]^{1-\theta}}_{=1+\mathrm{t}_{ij}} \Big)^{\frac{1}{1-\theta}} = \lambda_i \mathbb{P}_i$$

We see that the the consumption choice c_{ij} is distorted due to (i) the fact that demand for good j change the market clearing of country j, hence with shadow value μ_j, and (ii) the FOC is distorted with value η_{ij}.
If η_{ij} is positive, planner would like to relax the FOC $\mathrm{p}_j(1+\mathrm{t}_{ij}) - u'(c_{ij})$ implying it would like to increase the price.
To give intuition for the good demand distortion, let us give an expression for η_{ij}:

$$\eta_{ij}\frac{1}{\theta}\tau_{ij}(1+\mathrm{t}_{ij})\mathrm{p}_j(1-s_{ij}) = u'(c_i) c_i^{\frac{1}{\theta}} a_{ij}^{\frac{1}{\theta}} c_{ij}^{-\frac{1}{\theta}} - \lambda_i \tau_{ij}\mathrm{p}_j - \frac{\omega_j}{\omega_i}\mathcal{P}_i \mu_j \tau_{ij}\mathrm{p}_j$$

$$\Rightarrow \qquad \eta_{ij} = \frac{\theta}{(1-s_{ij})}\Big(\frac{u'(c_i)}{\mathbb{P}_i}\frac{1}{\lambda_i} - \frac{1+\frac{\omega_j}{\omega_i}\frac{\mu_j}{\lambda_i}}{1+\mathrm{t}_{ij}^b} \Big)$$

The distortion is positive $\eta_{ij} > 0$ for redistributive reasons, related to the budget of i and the market clearing of j. If $u'(c_i)/\mathbb{P}_i > \lambda_i$ and $\mathrm{t}_{ij}^b < \frac{\omega_j}{\omega_i}\frac{\mu_j}{\lambda_i}$, then the planner would like to distort the FOC by increasing the bilateral cost $(1+\mathrm{t}_{ij}^b)\tau_{ij}\mathrm{p}_j$.

If tariffs are set optimally, we have $\eta_{ij} = 0$, – from the above equation – we obtain that $\lambda_i + \frac{\omega_j}{\omega_i}\mu_j = \frac{u'(c_i)}{\mathbb{P}_i}(1+\mathrm{t}_{ij})$ and hence

$$1+\mathrm{t}_{ij}^b = 1 + \frac{\omega_j}{\omega_i}\frac{\mu_j}{\lambda_i}$$

for a hypothetical optimal tariffs on consumption imports. By consequence, we would also obtain naturally that $\frac{u'(c_i)}{\mathbb{P}_i} = \lambda_i$. However, for arbitrary policies t_{ij}^b, the FOC of the household is distorted and $\eta_{ij} \neq 0$.

- Price p_i

$$\omega_i \mathcal{P}_i \lambda_i \overline{\mathcal{D}}^y(\mathcal{E}) F(\ell_i, k_i, e_i) - \sum_k \omega_k \mathcal{P}_k \lambda_k \tau_{ki} c_{ki} - \sum_\ell \sum_k \omega_k \mathcal{P}_k \lambda_k \tau_{ki} x^\ell_{ki}$$
$$- \sum_{i \in \mathbb{I}} \omega_i\, \mathcal{P}_i \big[\upsilon^f_i MPe^f_i + \upsilon^c_i MPe^c_i + \upsilon^r_i MPe^r_i \big] + \sum_{\ell \in \{f,c,r\}} \sum_k \omega_k \theta^\ell_{kt} \Big(\frac{\tau_{ki}(1+\mathrm{t}_{ki})\mathrm{p}_i}{\mathbb{P}_k} \frac{\partial \mathbb{P}_k}{\partial \mathrm{p}_i} \Big)$$
$$+ \sum_k \omega_k \mathcal{P}_k c_{ki} \eta_{ki} \big[\tau_{ki}(1+\mathrm{t}_{ki}) - \frac{\tau_{ki}(1+\mathrm{t}_{ki})\mathrm{p}_i}{\mathbb{P}_k} \frac{\partial \mathbb{P}_k}{\partial \mathrm{p}_i} \big]$$
$$+ \sum_{\ell \in \{f,c,r,k\}} \sum_k \omega_k \mathcal{P}_k x^\ell_{ki} \vartheta^\ell_{ki} \tau_{ki}(1+\mathrm{t}_{ki})[1 - s_{ki}] = 0$$

$$\omega_i \mathcal{P}_i \lambda_i \overline{\mathcal{D}}^y(\mathcal{E}) F(k_i, e_i) - \sum_k \omega_k \mathcal{P}_k \lambda_k \big[\tau_{ki} c_{ki} + \sum_\ell \tau_{ki} x^\ell_{ki} \big] - \sum_{i \in \mathbb{I}} \omega_i\, \mathcal{P}_i \big[\upsilon^f_i MPe^f_i + \upsilon^c_i MPe^c_i + \upsilon^r_i MPe^r_i + \upsilon^k_i MPe^k_i \big]$$
$$+ \sum_{\ell \in \{f,c,r\}} \sum_k \omega_k \theta^\ell_{kt} \big(\tau_{ki}(1+\mathrm{t}_{ki}) s_{ki} \big) + \sum_k \omega_k \mathcal{P}_k \tau_{ki}(1+\mathrm{t}_{ki})[1 - s_{ki}] \Big(c_{ki}\eta_{ki} + \sum_{\ell \in \{f,c,r,k\}} x^\ell_{ki} \vartheta^\ell_{ki} \Big) = 0$$

This balances out all the redistributive effects: through λ_i on supply from i and on k's demand λ_k, the distortionary effects on energy choice υ^ℓ and energy production θ^ℓ_i, and the distortion on the bilateral import good choice η_{ki} for consumption and ϑ^ℓ_{ki} for inputs in energy inputs.

- Energy inputs: x^ℓ_{ij}, for $\ell = \{f, c, r, k\}$

$$\omega_i [\lambda_i + \mu^\ell_i] q^\ell_i g'(x^\ell_i)(x^\ell_i)^{\frac{1}{\theta}} a_{ij}^{\frac{1}{\theta}} (x^\ell_{ij})^{-\frac{1}{\theta}} - \omega_i \theta^\ell_i q^\ell_i g''(x^\ell_i)(x^\ell_i)^{\frac{1}{\theta}} a_{ij}^{\frac{1}{\theta}} (x^\ell_{ij})^{-\frac{1}{\theta}}$$
$$- \omega_i \lambda_i \tau_{ij} \mathrm{p}_j - \omega_j \mu_j \tau_{ij} \mathrm{p}_j - \omega_i \vartheta_{ij} \frac{1}{\theta} \tau_{ij}(1+\mathrm{t}_{ij}) \mathrm{p}_j (s_{ij} - 1) = 0$$
$$\{ \mu^\ell_i - \theta^\ell_i \frac{g''(x^\ell_i)}{g'(x^\ell_i)} \} q^\ell_i g'(x^\ell_i)(x^\ell_i)^{\frac{1}{\theta}} a_{ij}^{\frac{1}{\theta}} (x^\ell_{ij})^{-\frac{1}{\theta}} = \tau_{ij} \mathrm{p}_j \big[\mathrm{t}^b_{ij} \lambda_i + \frac{\omega_j}{\omega_i} \mu_j - \vartheta_{ij} \frac{1}{\theta} (1+\mathrm{t}^b_{ij})(1 - s_{ij}) \big]$$
$$\{ \mu^\ell_i - \theta^\ell_i \frac{g''(x^\ell_i)}{g'(x^\ell_i)} \} = \frac{1}{1+\mathrm{t}^b_{ij}} \big[\mathrm{t}^b_{ij} \lambda_i + \frac{\omega_j}{\omega_i} \mu_j - \vartheta_{ij} \frac{1}{\theta} (1+\mathrm{t}^b_{ij})(1 - s_{ij}) \big]$$

As for the consumption good above, this input choice x^ℓ_{ij} for energy production – which resembles a production networks/supply chain problem – bring additional distortions for each energy price ℓ, i.e. ϑ^ℓ_{ij}, which we can reexpress:

$$\vartheta^\ell_{ij} = \frac{\theta}{1 - s_{ij}} \Big[\{ \mu^\ell_i - \theta^\ell_i \frac{g''(x^\ell_i)}{g'(x^\ell_i)} \} - \frac{1 + \frac{\omega_j}{\omega_i} \frac{\mu_j}{\lambda_i}}{1+\mathrm{t}^b_{ij}} \Big]$$

which resemble the expression for η_{ij} for consumption good. This time the distortion for energy inputs ℓ from j are distorted if the shadow value of the market clearing for that energy sources μ^ℓ_i outweighs the distortion from the supply of that energy θ^ℓ_i – weighted by supply elasticity, related to g''/g', in the case tariffs are such that $\mathrm{t}^b_{ij} < \frac{\omega_j}{\omega_i} \frac{\mu_j}{\lambda_i}$.

Again, in the hypothetical case, where tariffs are set optimally, we obtain $\vartheta_{ij}^{\ell} = 0$ and thus:

$$1{+}\mathrm{t}_{ij}^{b} = 1 + \frac{\omega_j}{\omega_i}\frac{\mu_j}{\lambda_i}$$

and therefore: $\mu_i^{\ell} = \theta_i^{\ell}\frac{g''(x_i^{\ell})}{g'(x_i^{\ell})}$. However, in the standard case where tariffs are set suboptimally, we have $\vartheta_{ij}^{\ell} \neq 0$.

- Price q_i^{ℓ}, for coal and renewable $\ell = r, c$

$$\omega_i \mathcal{P}_i \lambda_i \Big[\frac{1}{\mathcal{P}_i} g(x_i^{\ell}) - e_i^{\ell}\Big] + \omega_i(\mathcal{P}_i \upsilon_i^{\ell} - g'(x_i^{\ell})\theta_i^{\ell}) = 0 \qquad \Rightarrow \qquad \mathcal{P}_i \upsilon_i^{\ell} = g'(x_i^{\ell})\theta_i^{\ell}$$

since the market clearing is local at the country level, there are no redistributive effect across countries and the distortion of demand υ_i^{ℓ} equates the distortion of supply θ_i^{ℓ}.

- Price q^f, for oil/gas

$$\sum_{\mathbb{I}} \omega_i \mathcal{P}_i \lambda_i [\frac{1}{\mathcal{P}_i} g(x_i^f) - e_i^f] + \sum_i \omega_i(\mathcal{P}_i \upsilon_i^f - g'(x_i^f)\theta_i^f) = 0$$

At the difference of the FOC for q_i^{ℓ}, for local energy sources, the oil-gas is traded internationally and therefore, changing its price has redistributive effects between countries depending on net-exports $g(x_i^{\ell}) - \mathcal{P}_i e_i^f$, through the covariance between those net-exports and the marginal value of income λ_i:

$$\sum_{\mathbb{I}} \omega_i \mathcal{P}_i \lambda_i [\frac{1}{\mathcal{P}_i} g(x_i^f) - e_i^f] = \mathbb{C}\mathrm{ov}(\omega_i \lambda_i, g(x_i^f) - \mathcal{P}_i e_i^f)$$

- Energy demand e_i^{ℓ}

$$\omega_i \mathcal{P}_i \lambda_i (\mathrm{p}_i MPe_i^{\ell} - q^{\ell}) + \omega_i \mathcal{P}_i \mu_i \mathrm{p}_i MPe_i^{\ell} - \omega_i q_i^{\ell} \mu_i^{\ell} \mathcal{P}_i - \phi^{\varepsilon} \mathcal{P}_i \xi^{\ell}$$
$$- \sum_{\ell'} \omega_i \mathcal{P}_i \mathrm{p}_i \upsilon_i^{\ell'} \partial_{e_i^{\ell}} MPe_i^{\ell'} = 0$$

We see that the energy demand choice by the planner internalize multiple effects that will be key in the formulation of the carbon tax: First it internalizes the climate externality, as summarized by the multiplier ϕ^{ε}. Second, it also accounts for the redistributive effect through the change on the energy market clearing μ_i^{ℓ} for that particular energy source. Third, it also distorts the FOC of the firm in all its energy and input sourcing ℓ', as summarized by the multipliers $\upsilon_i^{\ell'}$, and weighted by the terms $\partial_{e_i^{\ell}} MPe_i^{\ell'}$ which relates to the cross elasticity between energies ℓ and ℓ'. Moreover, it internalizes the effects that energy use has on good production, through multiplier μ_i. All these effects are detailed in more details below.
Let us more specific about each energy sources.

Fossil:

$$\omega_i \lambda_i \xi^f \mathrm{t}^{\varepsilon} + \omega_i \mu_i \mathrm{p}_i MPe_i^f - q^f \mu^f - \phi^{\varepsilon} \xi^f - \sum_{\ell'} \omega_i \mathrm{p}_i \upsilon_i^{\ell'} \partial_{e_i^f} MPe_i^{\ell'} = 0$$

Coal:

$$\omega_i\lambda_i\xi^c\mathrm{t}^\varepsilon + \omega_i\mu_i\mathrm{p}_i MPe_i^c - \omega_i q_i^c\mu_i^c - \phi^\varepsilon\xi^c - \sum_{\ell'}\omega_i\mathrm{p}_i\upsilon_i^{\ell'}\partial_{e_i^c}MPe_i^{\ell'} = 0$$

Renewable / non-carbon

$$\omega_i\mu_i\mathrm{p}_i MPe_i^r - \omega_i q_i^r\mu^f - \sum_{\ell'}\omega_i\mathrm{p}_i\upsilon_i^{\ell'}\partial_{e_i^r}MPe_i^{\ell'} = 0$$

- Carbon tax t_i^ε:

$$\sum_{\mathbb{I}}\omega_i\mathcal{P}_i(\upsilon_i^f\xi^f + \upsilon_i^c\xi^c) = 0$$

  The choice of the optimal carbon tax is a uniform tax that does not impose any additional aggregate distortion on the world economy. As a result, the sum of the country-levels distortions sum to zero: a positive distortion – multiplier $\upsilon_i^f > 0$ – need to be compensated by a negative distortion $\upsilon_{i'}^c < 0$ for another country, or across energy sources.

- Climate damage:

$$\begin{aligned}\phi^\varepsilon &= \sum_{\mathbb{I}}\mathcal{P}_i\omega_i\phi_i^\varepsilon \\ &= -\sum_{\mathbb{I}}\mathcal{P}_i\omega_i\Big[u(c_i)\overline{\mathcal{D}}^{u\prime}(\mathcal{E}) + (\lambda_i+\mu_i)\mathcal{D}_i^{y\prime}(\mathcal{E})\mathrm{p}_i z_i F(e_i,\ell_i,k_i) - \mathcal{D}_i^{y\prime}(\mathcal{E})\mathrm{p}_i\sum_{\ell'}\upsilon_i^{\ell'}MPe_i^{\ell'}\Big]\end{aligned}$$

  The marginal cost of climate change can be summarized by ϕ^ε and it internalizes the direct cost $\mathcal{D}_i^{y\prime}(\mathcal{E})$ and $\mathcal{D}_i^{u\prime}(\mathcal{E})$ of climate change on income – hence the multiplier λ_i – but also the effects of climate on good production μ_i and on the distortion of the energy demand optimality υ_i^ℓ.

Reformulation of the carbon tax

We take the example of the carbon tax on fossil fuels (oil-gas) to provide details the formulation of the tax:

$$\omega_i\lambda_i\xi^f\mathrm{t}^\varepsilon + \omega_i\mu_i\mathrm{p}_i MPe_i^f - q^f\mu^f - \phi^\varepsilon\xi^f - \sum_{\ell'}\omega_i\mathrm{p}_i\upsilon_i^{\ell'}\partial_{e_i^f}MPe_i^{\ell'} = 0$$

$$\underbrace{\sum_i\omega_i\widehat{\mathcal{P}}_i\lambda_i}_{=\overline{\lambda}}\,\xi^f\mathrm{t}^\varepsilon = \xi^f\phi^\varepsilon + q^f\mu^f - (q^f + \xi^f\mathrm{t}^\varepsilon)\sum_i\omega_i\widehat{\mathcal{P}}_i\mu_i + \sum_{\ell'}\sum_i\omega_i\widehat{\mathcal{P}}_i\upsilon_i^{\ell'}\;q_i^{\ell'}\frac{\partial_{e_i^f}MPe_i^{\ell'}}{MPe_i^{\ell'}}$$

Which gives, when aggregating over all countries i and rescaling the multipliers for the good market clearing $\widehat{\mu}_i$, the ones for energy $e_i^{\ell'}$ distortion $\widehat{\upsilon}_i^{\ell'}$, a formula for the carbon tax:

$$\xi^f\mathrm{t}^\varepsilon = \xi^f\underbrace{\frac{\phi^\varepsilon}{\overline{\lambda}}}_{=\text{SCC}} + \underbrace{q^f\frac{\mu^f}{\overline{\lambda}}}_{\substack{E^f\text{ supply}\\ \text{redistribut}^\circ}} - (q^f + \xi^f\mathrm{t}^\varepsilon)\sum_i\underbrace{\widehat{\mu}_i}_{\substack{y_i\text{Trade}\\ \text{redistribut}^\circ}} - \sum_{\ell'}\sum_i\underbrace{\widehat{\upsilon}_i^{\ell'}\;q_i^{\ell'}\frac{\partial_{e_i^f}MPe_i^{\ell'}}{MPe_i^{\ell'}}}_{\substack{e_i^{\ell'}\text{ demand}\\ \text{distort}^\circ}}$$

We now use the functional forms assumptions in our model to simplify this formula further.

Simplifying the formula

Assumptions, for formula in the paper

- Rewriting the social cost of carbon (SCC):

$$SCC = \frac{\phi^\varepsilon}{\overline{\lambda}} = \frac{\mathcal{P}\sum_i \omega_i \widehat{\mathcal{P}}_i \phi_i^\varepsilon}{\sum_i \omega_i \widehat{\mathcal{P}}_i \lambda_i} = \mathcal{P}\sum_i \frac{\omega_i \widehat{\mathcal{P}}_i \lambda_i}{\sum_i \omega_i \widehat{\mathcal{P}}_i \lambda_i}\frac{\phi_i^\varepsilon}{\lambda_i} = \mathcal{P}\sum_i \frac{\omega_i \widehat{\mathcal{P}}_i \lambda_i}{\sum_i \omega_i \widehat{\mathcal{P}}_i \lambda_i} LCC_i$$

$$\phi_i^\varepsilon = -\Big[u(c_i)\overline{\mathcal{D}}^{u\prime}(\mathcal{E}) + (\lambda_i+\mu_i)\mathcal{D}_i^{y\prime}(\mathcal{E})\mathrm{p}_i z_i F(e_i,\ell_i)\dots\Big]$$

$$\text{with} \qquad \lambda_i = \frac{u'(c_i)\overline{\mathcal{D}}^u(\mathcal{E})}{\mathbb{P}_i} \qquad\qquad \text{and CRRA} \qquad u'(c_i) = \frac{c_i^{1-\eta}}{1-\eta}$$

$$\text{damage} \qquad \overline{\mathcal{D}}^u(\mathcal{E}) = \left(\overline{\mathcal{D}}^{\widetilde{u}\prime}(\mathcal{E})\right)^{1-\eta} \qquad\qquad \overline{\mathcal{D}}^{u\prime}(\mathcal{E}) = (1-\eta)(\overline{\mathcal{D}}^{\widetilde{u}}(\mathcal{E}))^{-\eta}\overline{\mathcal{D}}^{\widetilde{u}\prime}(\mathcal{E})$$

$$\frac{u(c_i)\overline{\mathcal{D}}^{u\prime}(\mathcal{E})}{u'(c_i)\overline{\mathcal{D}}^{u}(\mathcal{E})} = \frac{c_i^{1-\eta}}{1-\eta}\frac{1}{\overline{\mathcal{D}}^{\widetilde{u}}(\mathcal{E})^{1-\eta}}(1-\eta)(\overline{\mathcal{D}}^{\widetilde{u}}(\mathcal{E}))^{-\eta}\overline{\mathcal{D}}^{\widetilde{u}\prime}(\mathcal{E}) = \frac{c_i\overline{\mathcal{D}}^{\widetilde{u}\prime}(\mathcal{E})}{\overline{\mathcal{D}}^{\widetilde{u}}(\mathcal{E})}$$

$$\frac{\phi_i^\varepsilon}{\overline{\lambda}} = -\Big[\frac{\lambda_i\mathbb{P}_i c_i}{\overline{\lambda}}\frac{\overline{\mathcal{D}}^{\widetilde{u}\prime}(\mathcal{E})}{\overline{\mathcal{D}}^{\widetilde{u}}(\mathcal{E})} + \frac{\lambda_i+\mu_i}{\overline{\lambda}}\frac{\overline{\mathcal{D}}^{y\prime}(\mathcal{E})}{\overline{\mathcal{D}}^{y}(\mathcal{E})}\mathrm{p}_i\mathcal{D}_i^y(\mathcal{E})z_iF(e_i,\ell_i)\dots\Big]$$

- Nordhaus DICE quadratic damage function and simple climate system: c.f. above.

$$\mathcal{D}^y(T-T^\star) = e^{-\frac{\gamma^y}{2}(T-T_i^\star)^2} \qquad \Rightarrow \qquad \mathcal{D}_i^{y\prime}(T-T^\star) = -\mathcal{D}_i^y(T-T^\star)\gamma^y(T-T_i^\star)$$

$$\dot{\mathcal{S}}_t = \mathcal{E} - \delta_s\mathcal{S}_t$$

$$T_{it} = T_{it_0} + \Delta\chi\mathcal{S}_t$$

$$\frac{\overline{\mathcal{D}}^{y\prime}(\mathcal{E})}{\overline{\mathcal{D}}^{y}(\mathcal{E})} \to_{t\to\infty, T_{it}\to T_i} -\frac{\Delta\chi}{\rho - n + (1-\eta)\bar{g} + \delta_s}\gamma^y(T_i - T_i^\star)$$

$$\frac{\phi_i^\varepsilon}{\overline{\lambda}} = \frac{\Delta\chi(T_i-T_i^\star)}{\rho - n + (1-\eta)\bar{g} + \delta_s}\Big[\frac{\lambda_i}{\overline{\lambda}}\mathbb{P}_i c_i\gamma^c + \frac{\lambda_i+\mu_i}{\overline{\lambda}}\mathrm{p}_i y_i\gamma^y - \gamma^y\big(\frac{\upsilon_i^f}{\overline{\lambda}}(q^f + \xi^f\mathrm{t}^\varepsilon) + \frac{\upsilon_i^c}{\overline{\lambda}}(q_i^c + \xi^c\mathrm{t}^\varepsilon) + \frac{\upsilon_i^r}{\overline{\lambda}}q_i^r + \frac{\upsilon_i^k}{\overline{\lambda}}(\rho+\eta\bar{g})\big)\Big]$$

The Local cost of carbon, for country i if $\omega_i = 1, \omega_j = 0$

$$LCC_i = \frac{\phi_i^\varepsilon}{\lambda_i} = \frac{\Delta\chi(T_i-T_i^\star)}{\rho - n + (1-\eta)\bar{g} + \delta_s}\Big[\mathbb{P}_i c_i\gamma^c + (1+\frac{\mu_i}{\lambda_i})\mathrm{p}_i y_i\gamma^y - \gamma^y\big(\frac{\upsilon_i^f}{\lambda_i}(q^f + \xi^f\mathrm{t}^\varepsilon) + \frac{\upsilon_i^c}{\lambda_i}(q_i^c + \xi^c\mathrm{t}^\varepsilon) + \frac{\upsilon_i^r}{\lambda_i}q_i^r + \frac{\upsilon_i^k}{\lambda_i}(\rho+\eta\bar{g})\big)\Big]$$

Reexpressing the total global social cost of carbon:

$$\begin{aligned}
SCC &= \mathcal{P}\sum_i \omega_i\widehat{\mathcal{P}}_i\frac{\phi_i^\varepsilon}{\overline{\lambda}} = \mathcal{P}\sum_i \widehat{\lambda}_i LCC_i \\
&= \mathcal{P}\sum_i \frac{\Delta\chi(T_i-T_i^\star)}{\rho - n + (1-\eta)\bar{g} + \delta_s}\Big[\widehat{\lambda}_i\mathbb{P}_i c_i\gamma^c + (\widehat{\lambda}_i + \widehat{\mu}_i)\mathrm{p}_i y_i\gamma^y \\
&\qquad - \gamma^y\big(\widehat{\upsilon}_i^f(q^f + \xi^f\mathrm{t}^\varepsilon) + \widehat{\upsilon}_i^c(q_i^c + \xi^c\mathrm{t}^\varepsilon) + \widehat{\upsilon}_i^r q_i^r + \widehat{\upsilon}_i^k(\rho+\eta\bar{g})\big)\Big]
\end{aligned}$$

with the rescaled multipliers for the budget constraint: $\widehat{\lambda}_i = \frac{\omega_i \widehat{\mathcal{P}}_i \lambda_i}{\overline{\lambda}} = \frac{\omega_i \widehat{\mathcal{P}}_i \lambda_i}{\sum_i \omega_i \widehat{\mathcal{P}}_i \lambda_i}$, the multiplier the FOC demand $\widehat{\upsilon}_i = \frac{\omega_i \widehat{\mathcal{P}}_i \upsilon_i}{\overline{\lambda}}$, for the multiplier for market clearing for good: $\widehat{\mu}_i = \frac{\omega_i \widehat{\mathcal{P}}_i \mu_i}{\overline{\lambda}}$ and population share $\widehat{\mathcal{P}}_i = \frac{\mathcal{P}_i}{\mathcal{P}}$

- Isoelastic energy supply curve: $x^f = \frac{\bar{\nu}_i}{1+\nu_i}\Big(\frac{e_i^x}{\mathcal{R}_i}\Big)^{1+\nu_i}\mathcal{R}_i$

$$g_i(x) = \mathcal{R}_i^{\frac{\nu_i}{1+\nu_i}} x^{\frac{1}{1+\nu_i}} \qquad \frac{g''(x^f)}{g'(x^f)} = -\frac{\nu_i}{1+\nu_i}\frac{1}{x_i}$$

As a result, FOC of energy inputs $[x_i^f]$ becomes:

$$\mu^f = \omega_i \theta_i^f \frac{g''(x^f)}{g'(x^f)} = \omega_i \theta_i^f \frac{-\nu_i}{1+\nu_i}\frac{1}{x_i^f} \qquad \omega_i \theta_i^f = -\frac{1+\nu_i}{\nu_i} x_i^f \mu^f$$

Moreover, the FOC of energy price $[q^f]$ becomes:

$$\begin{aligned}
\sum_i \omega_i \theta_i^f g'(x^f) &= -\mu^f \sum_i \frac{1+\nu_i}{\nu_i} x_i^f g'(x_i^f) \\
&= -\mu^f \sum_i \frac{1+\nu_i}{\nu_i}\frac{\bar{\nu}_i}{1+\nu_i}\Big(\frac{e_i^x}{\mathcal{R}_i}\Big)^{1+\nu_i}\mathcal{R}_i\Big(\frac{e_i^x}{\mathcal{R}_i}\Big)^{-\nu_i}\bar{\nu}_i^{-1} \\
\sum_i \omega_i \theta_i^f g'(x^f) &= -\mu^f \sum_i \frac{e_i^x}{\nu_i} = -\mu^f / \Big(\frac{1}{E^f}\underbrace{\big(\sum_i \frac{e_i^x}{E^f}\frac{1}{\nu_i}\big)^{-1}}_{=\bar{\nu}}\Big)
\end{aligned}$$

As a result, we have, with the aggregate elasticity $\bar{\nu}$

$$\begin{aligned}
\mu^f &= \frac{\bar{\nu}}{E^f}\sum_i \omega_i \mathcal{P}_i \lambda_i \big(e_i^f - \frac{e_i^x}{\mathcal{P}_i}\big) \qquad \bar{\nu} = \big(\sum_i \tfrac{e_i^x}{E^f}\nu_i^{-1}\big)^{-1} \\
\frac{\mu^f}{\overline{\lambda}} &= \mathcal{P}\frac{\bar{\nu}}{E^f}\sum_i \frac{\omega_i \widehat{\mathcal{P}}_i \lambda_i}{\overline{\lambda}}\big(e_i^f - \frac{e_i^x}{\mathcal{P}_i}\big)
\end{aligned}$$

- Nested CES framework:

$$\text{Energy} \quad e_i = \Big(\sum_\ell (\omega^\ell)^{\frac{1}{\sigma_e}}(e_i^\ell)^{\frac{\sigma_e-1}{\sigma_e}}\Big)^{\frac{\sigma_e}{\sigma_e-1}} \qquad \text{Output} \quad y_i = \Big((1-\varepsilon)^{\frac{1}{\sigma}}(e_i)^{\frac{\sigma-1}{\sigma}} + \varepsilon^{\frac{1}{\sigma}}(k_i^\alpha \ell_i^{1-\alpha})^{\frac{\sigma-1}{\sigma}}\Big)^{\frac{\sigma}{\sigma-1}}$$

FOC for fossil energy demand:

$$\begin{aligned}
\bar{\upsilon}_i^f &= [\upsilon_i^f \partial_{e^f} MPe_i^f + \upsilon_i^c \partial_{e^f} MPe_i^c + \upsilon_i^r \partial_{e^f} MPe_i^r + \upsilon_i^k \partial_{e^f} MPk_i] \\
&= \frac{1}{e_i^f}\Big[-\upsilon_i^f(q^f+\xi^f \mathrm{t}^\varepsilon)\big[\frac{1-s^f}{\sigma^e} + s^f\frac{1-s^e}{\sigma^y}\big] + \upsilon_i^c(q_i^c+\xi^f \mathrm{t}^\varepsilon)s_i^f\big[\frac{1}{\sigma^e} - \frac{1-s^e}{\sigma^y}\big] + \upsilon_i^r q_i^r s_i^f\big[\frac{1}{\sigma^e} - \frac{1-s^e}{\sigma^y}\big] + \upsilon_i^k(r^\star+\bar{\delta})\frac{s_i^{e^r/y}}{\sigma^y}\Big]
\end{aligned}$$

and when normalizing by $\overline{\lambda}$

$$\widehat{\upsilon}_i^f = \frac{1}{e_i^f}\Big[-\widehat{\upsilon}_i^f(q^f+\xi^f \mathrm{t}^\varepsilon)\big[\frac{1-s^f}{\sigma^e} + s^f\frac{1-s^e}{\sigma^y}\big] + \widehat{\upsilon}_i^c(q_i^c+\xi^f \mathrm{t}^\varepsilon)s_i^f\big[\frac{1}{\sigma^e} - \frac{1-s^e}{\sigma^y}\big] + \widehat{\upsilon}_i^r q_i^r s_i^f\big[\frac{1}{\sigma^e} - \frac{1-s^e}{\sigma^y}\big] + \widehat{\upsilon}_i^k(r^\star+\bar{\delta})\frac{s_i^{e^r/y}}{\sigma^y}\Big]$$

If the production function only contains fossil (oil-gas) in energy $\omega^f = 1$, then we obtain:

$$\bar{\upsilon}_i^f = -\frac{q^f + \xi^f \mathrm{t}^\varepsilon}{e_i^f} \upsilon_i^f \Big[\frac{1-s^e}{\sigma^y}\Big] \qquad\qquad s_i^e = \frac{q^e e_i}{\mathrm{p}_i y_i}$$

Proposition: Using these assumptions, we can reexpress the carbon tax:

$$\xi^f \mathrm{t}^\varepsilon = \xi^f \mathcal{P} \sum_i \widehat{\lambda}_i LCC_i + q^f \mathcal{P} \frac{\bar{\nu}}{E^f} \sum_i \widehat{\lambda}_i (e_i^f - \frac{e_i^x}{\mathcal{P}_i}) - (q^f + \xi^f \mathrm{t}^\varepsilon) \sum_i \widehat{\mu}_i - \sum_i \widehat{\upsilon}_i^f$$

$$\xi^f \mathrm{t}^\varepsilon = \xi^f \underbrace{\mathcal{P}\mathbb{E}_i[LCC_i] + \mathcal{P}\mathbb{C}\mathrm{ov}_i(\widehat{\lambda}_i, LCC_i)}_{=\text{Social Cost of Carbon}} + \underbrace{q^f \mathcal{P} \frac{\bar{\nu}}{E^f} \mathbb{C}\mathrm{ov}_i(\widehat{\lambda}_i, e_i^f - \frac{e_i^x}{\mathcal{P}_i})}_{\substack{E^f \text{ supply} \\ \text{redistribut}^\circ}} - (q^f + \xi^f \mathrm{t}^\varepsilon) \underbrace{\mathbb{E}_i[\widehat{\mu}_i]}_{\substack{y_i \text{Trade} \\ \text{redistribut}^\circ}} - \underbrace{\mathbb{E}_i[\widehat{\upsilon}_i^f]}_{\substack{e_i^{\ell'} \text{ demand} \\ \text{distort}^\circ}}$$

With the demand distortion of fossil fuels $\widehat{\upsilon}_i^f$

$$\widehat{\upsilon}_i^f = \frac{1}{e_i^f} \Big[-\widehat{\upsilon}_i^f (q^f + \xi^f \mathrm{t}^\varepsilon) \Big[\frac{1-s^f}{\sigma^e} + s^f \frac{1-s^e}{\sigma^y}\Big] + \widehat{\upsilon}_i^c (q_i^c + \xi^f \mathrm{t}^\varepsilon) s_i^f \Big[\frac{1}{\sigma^e} - \frac{1-s^e}{\sigma^y}\Big]$$
$$+ \widehat{\upsilon}_i^r q_i^r s_i^f \Big[\frac{1}{\sigma^e} - \frac{1-s^e}{\sigma^y}\Big] + \widehat{\upsilon}_i^k (r^\star + \bar{\delta}) \frac{s_i^{e^r/y}}{\sigma^y} \Big]$$
$$\widehat{\upsilon}_i^f = -\frac{q^f + \xi^f \mathrm{t}^\varepsilon}{e_i^f} \widehat{\upsilon}_i^f \Big[\frac{1-s_i^e}{\sigma^y}\Big] \qquad \text{if } \; s^f = 1, s^r = s^c = 0$$

and the social cost of carbon SCC as:

$$SCC = \mathcal{P} \sum_i \widehat{\lambda}_i LCC_i = \mathcal{P}\mathbb{E}_i[LCC_i] + \mathcal{P}\mathbb{C}\mathrm{ov}_i(\widehat{\lambda}_i, LCC_i)$$
$$= \mathcal{P} \sum_i \frac{\Delta\chi(T_i - T_i^\star)}{\rho - n + (1-\eta)\bar{g} + \delta_s} \Big[\widehat{\lambda}_i \mathbb{P}_i c_i \gamma^c + (\widehat{\lambda}_i + \widehat{\mu}_i) \mathrm{p}_i y_i \gamma^y$$
$$- \gamma^y \big(\widehat{\upsilon}_i^f (q^f + \xi^f \mathrm{t}^\varepsilon) + \widehat{\upsilon}_i^c (q_i^c + \xi^c \mathrm{t}^\varepsilon) + \widehat{\upsilon}_i^r q_i^r + \widehat{\upsilon}_i^k (\rho + \eta \bar{g})\big)\Big]$$

with the rescaled multipliers for the budget constraint: $\widehat{\lambda}_i = \frac{\omega_i \widehat{\mathcal{P}}_i \lambda_i}{\bar{\lambda}} = \frac{\omega_i \widehat{\mathcal{P}}_i \lambda_i}{\sum_i \omega_i \widehat{\mathcal{P}}_i \lambda_i}$, the multiplier the FOC demand $\widehat{\upsilon}_i = \frac{\omega_i \widehat{\mathcal{P}}_i \upsilon_i}{\bar{\lambda}}$, for the multiplier for market clearing for good: $\widehat{\mu}_i = \frac{\omega_i \widehat{\mathcal{P}}_i \mu_i}{\bar{\lambda}}$

Simplifying the multiplier for the FOC for energy demand

$$\xi^f \mathrm{t}^\varepsilon = \xi^f \mathcal{P} \Big(\mathbb{E}_j[LCC_j] + \mathbb{C}\mathrm{ov}_j(\widehat{\lambda}_j, LCC_j)\Big) + \mathcal{P} \frac{q^f \bar{\nu}}{E} \mathbb{C}\mathrm{ov}_j(\widehat{\lambda}_j, e_j^f - \frac{e_j^x}{\mathcal{P}_j})$$
$$- (q^f + \xi^f \mathrm{t}^\varepsilon) \mathbb{E}_j[\widehat{\mu}_j] - (q^f + \xi^f \mathrm{t}^\varepsilon) \mathbb{C}\mathrm{ov}(\widehat{\upsilon}_i^f, \frac{1-s^e}{\sigma e_i^f})$$

where the last equality comes from the fact that $\mathbb{E}_i[\widehat{\upsilon}_i^f] = \sum_i \widehat{\upsilon}_i^f = 0$ by the assumption that there is no aggregate distortion – only individual distortion – when the uniform carbon tax is set at the world level and $\mathbb{E}_i[e_i^f - \frac{e_i^x}{\mathcal{P}_i}] = \sum_i e_i^f - \frac{e_i^x}{\mathcal{P}_i} = 0$ by market clearing on the fossil energy market.

To investigate the demand distortion $\widehat{\upsilon}_i^f$ further, we can see that, with the planner's FOC for energy demand, the individual distortion becomes:

$$\omega_i \mathcal{P}_i \upsilon_i^f = \frac{\mathcal{P}_i}{q^f + \xi^f \mathrm{t}^\varepsilon} \frac{\sigma^y e_i^f}{1 - s_i^e} \Big[\xi^f \phi^S + q^f \mu^f - \omega_i \mu_i (q^f + \xi^f \mathrm{t}^\varepsilon) - \omega_i \lambda_i \xi^f \mathrm{t}^\varepsilon \Big]$$

The distortion is higher if the welfare motives in terms of cost of climate change ϕ^S, supply redistribution μ^f and trade effect $-\mu_i$ etc. outweighs the welfare cost of the carbon tax $\lambda_i \mathrm{t}^\varepsilon$. As mentioned earlier, the averate/aggregate distortion is null, so – in the case where fossil (oil-gas) is the only energy – we obtain:

$$\sum_i \omega_i \mathcal{P}_i \upsilon_i^f = 0$$

$$\frac{1}{q^f + \xi^f \mathrm{t}^\varepsilon} (\xi^f \phi^S + q^f \mu^f) \sum_i \mathcal{P}_i \frac{\sigma^y e_i^f}{1 - s_i^e} - \frac{1}{q^f + \xi^f \mathrm{t}^\varepsilon} \xi^f \mathrm{t}^\varepsilon \sum_i \omega_i \mathcal{P}_i \lambda_i \frac{\sigma^y e_i^f}{1 - s_i^e} - \frac{q^f + \xi^f \mathrm{t}^\varepsilon}{q^f + \xi^f \mathrm{t}^\varepsilon} \sum_i \omega_i \mathcal{P}_i \mu_i \frac{\sigma^y e_i^f}{1 - s_i^e} = 0$$

$$\xi^f \mathrm{t}^\varepsilon \sum_i \omega_i \mathcal{P}_i \lambda_i \frac{\sigma^y e_i^f}{1 - s_i^e} = (\xi^f \phi^S + q^f \mu^f) \sum_i \mathcal{P}_i \frac{\sigma^y e_i^f}{1 - s_i^e} - (q^f + \xi^f \mathrm{t}^\varepsilon) \sum_i \omega_i \mathcal{P}_i \mu_i \frac{\sigma^y e_i^f}{1 - s_i^e}$$

Divide both side by $\overline{\lambda} = \sum_i \omega_i \widehat{\mathcal{P}}_i \lambda_i$ and by $E^{s,\sigma} = \sum_i \widehat{\mathcal{P}}_i \frac{\sigma^y e_i^f}{1 - s_i^e}$ and $\widehat{e}_i^{s,\sigma} = \frac{\frac{\sigma^y e_i^f}{1 - s_i^e}}{\sum_i \widehat{\mathcal{P}}_i \frac{\sigma^y e_i^f}{1 - s_i^e}}$, it becomes;

$$\xi^f \mathrm{t}^\varepsilon \sum_i \frac{\omega_i \widehat{\mathcal{P}}_i \lambda_i}{\overline{\lambda}} \frac{1}{E^{s,\sigma}} \frac{\sigma^y e_i^f}{1 - s_i^e} = (\xi^f \frac{\phi^S}{\overline{\lambda}} + q^f \frac{\mu^f}{\overline{\lambda}}) \frac{1}{E^{s,\sigma}} \sum_i \widehat{\mathcal{P}}_i \frac{\sigma^y e_i^f}{1 - s_i^e} - (q^f + \xi^f \mathrm{t}^\varepsilon) \sum_i \frac{\omega_i \widehat{\mathcal{P}}_i \mu_i}{\overline{\lambda}} \frac{1}{E^{s,\sigma}} \frac{\sigma^y e_i^f}{1 - s_i^e}$$

$$\xi^f \mathrm{t}^\varepsilon = \frac{1}{1 + \mathbb{C}\mathrm{ov}_j(\widehat{\lambda}_j, \widehat{e}_i^{s,\sigma})} \Big[\xi^f \mathcal{P} SCC + \mathcal{P} \frac{q^f \bar{\nu}}{E} \mathbb{C}\mathrm{ov}_j(\widehat{\lambda}_j, e_j^f - \tfrac{e_j^x}{\mathcal{P}_j}) - (q^f + \xi^f \mathrm{t}^\varepsilon) \mathbb{C}\mathrm{ov}_j(\widehat{\mu}_j, \widehat{e}_i^{s,\sigma}) \Big]$$

It implies that the carbon tax is dampened if the Planner puts larger social weights on countries that use a lot of energy e_i^f, with a higher elasticity σ^y, and as a larger energy share in production s_i^e, resulting in $\mathbb{C}\mathrm{ov}_j(\widehat{\lambda}_j, \widehat{e}_i^{s,\sigma}) > 0$. If this covariance is negative, then the carbon tax is amplified, and the planner optimally chooses a higher carbon tax.

C.3 Unilaleral policy

In this allocation, the Ramsey planner now maximizes country i's welfare $\mathcal{P}_i \mathcal{U}_i$, choosing the allocation in country i. It chooses $\boldsymbol{x}_i = \{c_{ij}, x_{ij}^\ell, e_i^\ell\}$, i.e. the traded good for consumption c_{ij}, for energy inputs for the production of in fossil x_{ij}^f, coal x_{ij}^c, non-carbon x_{ij}^r or capital x_{ij}^k, and the energy demand, in fossil e_i^f, coal e_i^c and non-carbon e_i^r, and the prices $\boldsymbol{p}_i = \{\mathrm{p}_i, q^f, q_i^c, q_i^r\}$. For the policy instruments, they choose the country i carbon tax t_i^ε as well as the set of trade tariffs $\{\mathrm{t}_{ij}^b\}_j$ against country j. Moreover, we consider the Nash equilibrium, and the planner i take as given the policies of the other countries j, i.e. $\boldsymbol{x}_j = \{c_{jk}, x_{jk}^\ell, e_j^\ell, \mathrm{t}_j\}_j, \forall j \neq i$.

Again, the allocation and prices are constrained to be a competitive equilibrium and the planner chooses controls that respect the individual optimality conditions. However, due to all the general equilibrium and redistributive effects, we need to take a stance of what the planner internalizes. I make the assumption that the planner in country i only internalizes the optimality conditions (FOC) of its own country, but still internalize the market clearing of other countries when it involves the good traded by country i.

As a result, it uses the multiplier for the market clearing for the good traded from j, i.e. $\mu_j^{(i)}$: so $\mu_j^{(i)}$ represents the shadow value of relaxing country j market clearing (i.e. producing one extra unit of variety j), as internalized by planner i. Similarly, oil-gas is traded internationally so $\mu^{f,(i)}$ represents the shadow value of producing one extra unit of oil-gas (in the market, so produced by any country), as internalized by planner i. However, the market clearing of j for coal and renewable are not affected directly, only the market in i, which we keep denoting μ_i^c and μ_i^r.

Moreover, for climate damage ϕ_i^ε is the shadow cost of one additional ton of carbon emissions in the atmosphere as seen by country i. For the other multipliers, we use the same as before: λ_i for the budget constraint, ϕ^c for the consumption decision, θ_i^ℓ for the production quantity (supply) choice of energy firms $\ell = f, c, r$ for fossil, coal and renewable of country i, υ_i^ℓ for the quantity (demand) of energy ℓ chosen by the good firm, η_{ij} for the consumption choice for imports j by the household in i, ϑ_{ij}^ℓ for the import choice for inputs from j for the energy firm j. We are only considering those values for country i as the planner does not affect directly the distortion in country j.

As a result, the controls are $\boldsymbol{x}_i = \{c_{ij}, x_{ij}^\ell, e_i^\ell, \mathrm{p}_i, q^f, q_i^c, q_i^r, \mathrm{t}_i^\varepsilon, \mathrm{t}_{ij}^b\}_j$ and the multipliers are $\boldsymbol{\lambda}_i = \{\lambda_i, \mu_j^{(i)}, \mu_i^c, \mu_i^r, \mu^{f,(i)}, \phi_i^c, \theta_i^\ell, \upsilon_i^\ell, \phi^\varepsilon, \eta_{ij}, \vartheta_{ij}^\ell\}_{\ell,j}$.

The planner i's Lagrangian writes:

$$
\begin{aligned}
\mathcal{L}(\boldsymbol{x}_i, \boldsymbol{\lambda}_i) = \; & \mathcal{P}_i u(c_i)\overline{\mathcal{D}}_i^u(\mathcal{E}) + \mathcal{P}_i \lambda_i \Big(\mathrm{p}_i \overline{\mathcal{D}}^y(\mathcal{E}) f(\ell_i, k_i, e_i^f, e_i^c, e_i^r) + \frac{1}{\mathcal{P}_i}[q^f g^f(x_i^\ell) - \sum_j x_{ij}^\ell \tau_{ij} \mathrm{p}_j (1+\mathrm{t}_{ij}^b)] \\
& + \sum_\ell \{q^\ell g^\ell(x_i^\ell) - \sum_j x_{ij}^\ell \tau_{ij} \mathrm{p}_j (1+\mathrm{t}_{ij}^b)\} - ((q^f+\xi^f \mathrm{t}_i^\varepsilon) e_i^f + (q^c+\xi^c \mathrm{t}_i^\varepsilon) e_i^c + q_i^r e_i^r + (n_i+\bar{g}_i+\delta) k_i + c_i \mathbb{P}_i + \mathrm{t}_i^{ls}) \Big) \\
& + \sum_{\mathbb{I}} \omega_i^{(i)} \mathrm{p}_i \mu_i^{(i)} \Big(\mathcal{P}_i \mathcal{D}_i(T_i) z_i f(e_i, \ell_i) - \sum_{k\in\mathbb{I}} \mathcal{P}_k \tau_{ki} c_{ki} + \sum_{k\in\mathbb{I}} \tau_{ki}(x_{ki}^f + x_{ki}^c + x_{ki}^r)\Big) \\
& + \mu^{f(i)} q^f [\sum_{i\in\mathbb{I}} e_i^x - \mathcal{P}_i e_i^f] + \mu_i^c q_i^c (\bar{e}_i^c - \mathcal{P}_i e_i^c) + \mu_i^r q_i^r (\bar{e}_i^r - \mathcal{P}_i e_i^r) \\
& + \mathcal{P}_i \phi_i^\varepsilon [\mathcal{S} - \sum_{i\in\mathbb{I}} \mathcal{P}_i (\xi^f e_i^f + \xi^c e_i^c)] + \mathcal{P}_i \phi_i^c (\mathbb{P}_i \lambda_i^h - u'(c_i) \overline{\mathcal{D}}^u(\mathcal{E})) + \sum_{\ell\in\{f,c,r\}} \theta_i^\ell (\mathbb{P}_i - q_i^\ell g'(x_i^\ell)) \\
& + \mathcal{P}_i \Big(\upsilon_i^f [(q^f+\xi^f \mathrm{t}^\varepsilon) - \mathrm{p}_i MP e_i^f] + \upsilon_i^c [(q_i^c+\xi^c \mathrm{t}^\varepsilon) - \mathrm{p}_i MP e_i^c] + \upsilon_i^r [q_i^r - \mathrm{p}_i MP e_i^r] + \upsilon_i^k [\underbrace{\rho+\eta\bar{g}_i}_{r_i+\bar{g}+n} - \mathrm{p}_i MP k_i - \delta]\Big) \\
& + \sum_{j\in\mathbb{I}} \mathcal{P}_i \eta_{ij} c_{ij} [(1+\mathrm{t}_{ij}) \tau_{ij} \mathrm{p}_j - \mathbb{P}_i c_i^{\frac{1}{\theta}} a_{ij}^{\frac{1}{\theta}} c_{ij}^{-\frac{1}{\theta}}] + \sum_{\ell\in\{f,c,r\}} \sum_{j\in\mathbb{I}} \vartheta_{ij}^\ell x_{ij}^\ell [(1+\mathrm{t}_{ij}) \tau_{ij} \mathrm{p}_j - \mathbb{P}_i (x_i^\ell)^{\frac{1}{\theta}} a_{ij}^{\frac{1}{\theta}} (x_{ij}^\ell)^{-\frac{1}{\theta}}]
\end{aligned}
$$

where $\omega_j^{(i)}$ are the weights planner i puts on market clearing j, and $\omega_i = 1$.

The First Order Conditions of the control over home variables $c_{ij}, x_{ij}^\ell, e_i^\ell, \mathrm{p}_i$ write:

- Consumption c_{ij}

$$\mathcal{P}_i u'(c_i) c_i^{\frac{1}{\theta}} a_{ij}^{\frac{1}{\theta}} c_{ij}^{-\frac{1}{\theta}} - \mathcal{P}_i \lambda_i \tau_{ij} \mathrm{p}_j - \omega_j^{(i)} \mathcal{P}_i \mu_j^{(i)} \tau_{ij} \mathrm{p}_j + \omega_i \mathcal{P}_i c_{ij} \eta_{ij} \frac{1}{\theta} \frac{\tau_{ij}(1+\mathrm{t}_{ij})\mathrm{p}_j}{c_{ij}} (s_{ij} - 1)$$

$$u'(c_i) = \lambda_i \Big(\sum_j a_{ij} (\tau_{ij} \mathrm{p}_j)^{1-\theta} \underbrace{[1 + \omega_j^{(i)} \frac{\mu_j^{(i)}}{\lambda_i} - \frac{\eta_{ij}}{\lambda_i} \frac{1}{\theta} (1+\mathrm{t}_{ij}^b)(1-s_{ij})]}_{=1+\mathrm{t}_{ij}^b}{}^{1-\theta} \Big)^{\frac{1}{1-\theta}}$$

As before, if t_{ij}^b are set optimally, we obtain:

$$1 + \mathrm{t}_{ij}^b = 1 + \omega_j^{(i)} \frac{\mu_j^{(i)}}{\lambda_i}$$

The optimal tariff on j internalizes the change in demand through the market clearing of good j and uses this shadow value to set the tariffs: if $\mu_j^{(i)} > 0$, planner i would like to relax the constraint for j and thus reduce demand, hence setting a positive tariff t_{ij}^b.

Moreover, in this context, the distortionary effect on the ij trade, denoted by η_{ij}, is zero. Indeed, now, since the planner can manipulate that decisions freely, they choose to avoid causing a distortion $\eta_{ij} = 0$.

- Energy inputs: x_{ij}^ℓ

$$q_i^\ell g'(x_i^\ell)(x_i^\ell)^{\frac{1}{\theta}} a_{ij}^{\frac{1}{\theta}} (x_{ij}^\ell)^{-\frac{1}{\theta}} - \theta_i^\ell q_i^\ell g''(x_i^\ell)(x_i^\ell)^{\frac{1}{\theta}} a_{ij}^{\frac{1}{\theta}} (x_{ij}^\ell)^{-\frac{1}{\theta}} - \lambda_i \tau_{ij} \mathrm{p}_j - \omega_j^{(i)} \mu_j \tau_{ij} \mathrm{p}_j - \vartheta_{ij} \frac{1}{\theta} \tau_{ij} (1+\mathrm{t}_{ij}) \mathrm{p}_j (s_{ij} - 1) =$$

$$\{\mu_i^\ell - \theta_i^\ell \frac{g''(x_i^\ell)}{g'(x_i^\ell)}\} q_i^\ell g'(x_i^\ell)(x_i^\ell)^{\frac{1}{\theta}} a_{ij}^{\frac{1}{\theta}} (x_{ij}^\ell)^{-\frac{1}{\theta}} = \tau_{ij} \mathrm{p}_j [\mathrm{t}_{ij}^b \lambda_i + \omega_j^{(i)} \mu_j^{(i)} - \vartheta_{ij} \frac{1}{\theta} (1+\mathrm{t}_{ij}^b)(1-s_{ij})]$$

$$\{\mu_i^\ell - \theta_i^\ell \frac{g''(x_i^\ell)}{g'(x_i^\ell)}\} = \frac{1}{1+\mathrm{t}_{ij}^b} \big[- \mathrm{t}_{ij}^b \lambda_i + \omega_j^{(i)} \mu_j^{(i)} - \vartheta_{ij} \frac{1}{\theta} (1+\mathrm{t}_{ij}^b)(1-s_{ij})]$$

Choosing tariffs optimally yield again:

$$\mathrm{t}_{ij}^b = \omega_j^{(i)} \frac{\mu_j^{(i)}}{\lambda_i}$$
$$\vartheta_{ij} = 0 \qquad \mu_i^\ell = \theta_i^\ell \frac{g''(x_i^\ell)}{g'(x_i^\ell)}$$

The optimal tariff is the same as above, and there is no distortion on the FOC for good inputs for energy productions, i.e. $\vartheta_{ij} = 0$.
Moreover, the shadow value of the market clearing for energy i equals the distortion from the supply of that energy θ_i^ℓ – weighted by the supply elasticity, related to g''/g'.
For oil-gas, since the market is global and the supply curve is strictly convex, we get:

$$\mu^{f(i)} = \theta_i^f \frac{g''(x_i^f)}{g'(x_i^f)}$$

For coal and renewable energy this implies that $\mu_i^c = \mu_i^r = 0$ and there is no distortion of the market clearing for these local, again because the planner can completely control the demand and supply for those energy sources.

- Price q_i^ℓ, first with non-oil/gas, $\ell = r$ or $\ell = c$, the local market implies:

$$\mathcal{P}_i\lambda_i[\frac{1}{\mathcal{P}_i}g(x_i^\ell) - e_i^\ell] + (\mathcal{P}_i\upsilon_i^\ell - g'(x_i^\ell)\theta_{it}^\ell) = 0 \qquad \Rightarrow \qquad \mathcal{P}_i\upsilon_i^\ell = g'(x_i^\ell)\theta_{it}^\ell$$

- Price q^f, for oil/gas

$$\lambda_i[g(x_i^f) - \mathcal{P}_i e_i^f] + (\mathcal{P}_i\upsilon_i^f - g'(x_i^f)\theta_i^f) = 0$$

This implies that the total net distortion, between demand $\mathcal{P}_i\upsilon_i^f$ and supply θ_i^f equals the net-import $\mathcal{P}_i e_i^f - g(x_i^f)$.

- Tariffs t_{ij}^b

$$\sum_{\ell\in\{f,c,r\}}\theta_i^\ell(\frac{\tau_{ij}(1+\mathrm{t}_{ij})\mathrm{p}_j}{\mathbb{P}_i}\frac{\partial\mathbb{P}_i}{\partial(1+\mathrm{t}_{ij}^b)}) + \mathcal{P}_i\eta_{ij}(\tau_{ij}\mathrm{p}_j - \frac{\tau_{ij}(1+\mathrm{t}_{ij})\mathrm{p}_j}{\mathbb{P}_i}\frac{\partial\mathbb{P}_i}{\partial(1+\mathrm{t}_{ij}^b)}) + \sum_{\ell\in\{f,c,r,k\}}\vartheta_{ij}^\ell x_{ij}^\ell(\tau_{ij}\mathrm{p}_j - \frac{\tau_{ij}(1+\mathrm{t}_{ij})\mathrm{p}_j}{\mathbb{P}_i}\frac{\partial\mathbb{P}_i}{\partial(1+\mathrm{t}_{ij}^b)}) = 0$$

$$\sum_{\ell\in\{f,c,r\}}\theta_i^\ell s_{ij} + \eta_{ij}c_{ij}\mathcal{P}_i(1-s_{ij}) + \sum_{\ell\in\{f,c,r,k\}}\vartheta_{ij}^\ell x_{ij}^\ell(1-s_{ij}) = 0$$

As we saw that the tariff is set optimally, we have no distortion for each good sourcing: consumption $\eta_{ij} = 0$ and energy inputs $\vartheta_{ij}^\ell = 0, \forall\ell$.
This implies that for all the goods sourced, we obtain:

$$\sum_{\ell\in\{f,c,r\}}\theta_i^\ell = 0$$

the shadow values of the supply distortions are summing to zero at the country level: since all energy source from the same input bundle at price $\mathbb{P}_i$, some supply are distorted positively while some are distorted negatively.

- Price p_i – the planner can only control the price p_i from local good i

$$\mathcal{P}_i\lambda_i\overline{\mathcal{D}}^y(\mathcal{E})f(k_i,e_i) - \mathcal{P}_i\lambda_i[\tau_{ii}c_{ii} + \sum_\ell\tau_{ii}x_{ii}^\ell] - \mathcal{P}_i[\upsilon_i^f MPe_i^f + \upsilon_i^c MPe_i^c + \upsilon_i^r MPe_i^r + \upsilon_i^k MPe_i^k]$$
$$+ \sum_{\ell\in\{f,c,r,k\}}\theta_i^\ell(\tau_{ii}(1+\mathrm{t}_{ii})s_{ii}) + \mathcal{P}_i\tau_{ii}(1+\mathrm{t}_{ii})[1-s_{ii}]\Big(c_{ii}\eta_{ii} + \sum_{\ell\in\{f,c,r,k\}}x_{ii}^\ell\vartheta_{ii}^\ell\Big) = 0$$

$$\Rightarrow \qquad \lambda_i[\overline{\mathcal{D}}^y(\mathcal{E})f(\ell_i,k_i,e_i) - \tau_{ii}(c_{ii} + \sum_\ell x_{ii}^\ell)] - [\upsilon_i^f MPe_i^f + \upsilon_i^c MPe_i^c + \upsilon_i^r MPe_i^r + \upsilon_i^k MPe_i^k] = 0$$

since we have no distortion good sourcings: $\eta_{ij} = \vartheta_{ij}^\ell = 0, \forall\ell$, and use $\sum_{\ell\in\{f,c,r\}}\theta_i^\ell = 0$, we have that the planner would like to balance out different effects.
This condition is akin to a terms-of-trade manipulation: the planner would like to increase the price p_i, as it increases its purchasing power $\lambda_i\mathrm{p}_i y_i$ via income, and allow to demand more energy υ_i^ℓ, but at the same time balance out the cost for its own household and energy inputs $\lambda_i(c_{ii} + \sum_\ell x_{ii}^\ell)$.

- Climate damage per capita ϕ_i^{ε} from changing emissions $\mathcal{E}$:

$$\phi_i^{\varepsilon} = \Big[u(c_i)\overline{\mathcal{D}}_i^{u\prime}(\mathcal{E}) + (\lambda_i + \mu_i^{(i)})\mathcal{D}_i^{y\prime}(\mathcal{E})\mathrm{p}_i z_i F(\ell_i, k_i, e_i) - \mathcal{D}_i^{y\prime}(\mathcal{E})\mathrm{p}_i\{\upsilon_i^f MPe_i^f + \upsilon_i^c MPe_i^c + \upsilon_i^r MPe_i^r + \upsilon_i^k MPe_i^k\}$$
$$+ \sum_j \omega_j^{(i)}\mu_j^{(i)}\mathcal{D}_j^{y\prime}(\mathcal{E})\mathrm{p}_j z_j F(\ell_j, k_j, e_j)\Big]$$

This represents the local social cost of carbon in welfare units. Climate change has different impact on country i: it affects utility $u(\cdot)$, affect production and thus budget $\lambda_i y_i$. It also distorts the input demand υ_i^{ℓ} through its impact on firm productivity.

Moreover, it also impacts production of good i through market clearing $\mu_i^{(i)}$. Interestingly, climate also impacts the goods production from countries j, which affects indirectly the country i through imports i: as a result, the planner i does indirectly account for the impact of climate on other countries j through international trade – because it cares of its own imported consumption c_{ij} through value $\mu_j^{(i)}$

This channel is novel when computing the social cost of carbon and I plan to investigate further how local social of carbon can be correlated across countries through international trade, an idea also discussed in Dingel et al. (2019).

- Carbon tax $\mathrm{t}_i^{\varepsilon}$:

$$\mathcal{P}_i(\upsilon_i^f \xi^f + \upsilon_i^c \xi^c) = 0$$

given that the planner has a single tax instruments for both energy inputs – fossil (oil/gas) and coal – they would like to avoid create distortion at the country level, and hence the value of these two distortions should offset each others. In practice, the distortion for oil is positive while the one for coal is negative: if the planner had two instruments it would set a higher tax on oil and a lower one of coal, to attenuate the distortionary effects.

- Energy demand e_i^{ℓ}

$$\mathcal{P}_i\lambda_i(\mathrm{p}_i MPe_i^{\ell} - q^f) + \mathcal{P}_i\mu_i^{(i)}\mathrm{p}_i MPe_i^{\ell} - q_i^{\ell}\mu_i^{\ell}\mathcal{P}_i - (\mathcal{P}_i\phi_i^{\varepsilon})\mathcal{P}_i\xi^{\ell} - \sum_{\ell'}\mathcal{P}_i\mathrm{p}_i\upsilon_i^{\ell'}\partial_{e_i^{\ell}} MPe_i^{\ell'} = 0$$

Oil-gas:

$$\lambda_i\xi^f\mathrm{t}_i^{\varepsilon} + \mu_i^{(i)}\mathrm{p}_i MPe_i^f - q^f\mu_i^{f(i)} - (\mathcal{P}_i\phi_i^{\varepsilon})\xi^f - \sum_{\ell'}\mathrm{p}_i\upsilon_i^{\ell'}\partial_{e_i^f} MPe_i^{\ell'} = 0$$

Coal

$$\lambda_i\xi^c\mathrm{t}_i^{\varepsilon} + \mu_i^{(i)}\mathrm{p}_i MPe_i^c - q_i^c\mu_i^c - (\mathcal{P}_i\phi_i^{\varepsilon})\xi^c - \sum_{\ell'}\mathcal{P}_i\mathrm{p}_i\upsilon_i^{\ell'}\partial_{e_i^c} MPe_i^{\ell'} = 0$$

Renewable:

$$\mu_i^{(i)}\mathrm{p}_i MPe_i^r - q_i^r\mu_i^r - \sum_{\ell'}\mathcal{P}_i\mathrm{p}_i\upsilon_i^{\ell'}\partial_{e_i^r} MPe_i^{\ell'} = 0$$

Using these last conditions, we can express the optimal carbon tax.

Unilateral carbon tax

From the optimality condition for energy we can express the carbon tax in function of different motives. Let us consider the example of oil and gas:

$$\lambda_i \xi^f \mathrm{t}_i^\varepsilon = -\mu_i^{(i)} \mathrm{p}_i MPe_i^f + q^f \mu_i^{f(i)} + (\mathcal{P}_i \phi_i^\varepsilon)\xi^f + \sum_{\ell'} \mathrm{p}_i \upsilon_i^{\ell'} \partial_{e_i^f} MPe_i^{\ell'}$$

$$\xi^f \mathrm{t}_i^\varepsilon = -\frac{\mu_i^{(i)}}{\lambda_i} q^f + \frac{\mu_i^{f(i)}}{\lambda_i} q^f + (\frac{\mathcal{P}_i \phi_i^\varepsilon}{\lambda_i})\xi^f$$
$$+ \big[\mathrm{p}_i \frac{\upsilon_i^f}{\lambda_i} \partial_{e_i^f} MPe_i^f + \frac{\upsilon_i^c}{\lambda_i} \partial_{e_i^f} MPe_i^c + \frac{\upsilon_i^r}{\lambda_i} \partial_{e_i^f} MPe_i^r + \frac{\upsilon_i^k}{\lambda_i} \partial_{e_i^f} MPe_i^k\big]$$

We see that the carbon balance four different effects – as in the global Second-Best but this time at the country level:

First, a climate externality Pigouvian motive $\mathcal{P}_i \phi_i^\varepsilon / \lambda_i$: the tax needs to be larger to account for the local cost of carbon defined as :

$$LCC_i = \mathcal{P}_i \frac{\phi^\varepsilon}{\lambda_i}$$

Moreover, this scale with population: if the country is larger $\mathcal{P}_i$, the tax internalize more the damage effect that energy consumption has on its' population.

Second, a distortionary effect: it would like to offset the distortion effect it has across the different energy inputs, and this weights by cross-elasticities, which relates to the terms $\partial_{e_i^f} MPe_i^\ell$.

Third, a redistributive term linked to the energy market clearing $\mu^{f(i)}$, if the planner i would like to relax the market for oil, or reduce oil demand, i.e. a positive $\mu^{f(i)}$, it would set a higher tax to lower its own demand

Fourth, a terms-of-trade manipulation effect, linked to the market clearing of it's own good $\mu_i^{(i)}$. This term is usually positive: the planner would like to increase its own production and to that purpose it could lower the carbon tax, or even subsidize carbon (!) to manipulate terms-of-trade.

Finally, all these terms are weighted by the country own marginal utility of consumption $\lambda_i = u'(c_i)/\mathbb{P}_i$. They are amplified if the country is higher income and dampened for lower income/consumption countries.

To provide even more intuitions, let us consider that the production only use labor and fossil (oil/gas) energy and use our isoelastic supply function for oil and gas.

In that case, the carbon tax is solely a tax on oil: there is no demand distortion effect and $\upsilon_i^f = 0$ since the tax can completely offset distortion. The second distortionary effect drops out.

Unfortunately, terms-of-trade effect $\mu_j^{(i)}$ and $\mu_i^{(i)}$ can not be expressed in closed-form easily as they depend on the general equilibrium effects and international trade in the Armington model, usually not tractable

Moreover, we have the isoelastic supply curve $e_i^x = g(x_i^f)$ that implies the term:

$$\frac{g''(x_i^f)}{g'(x_i^f)} = -\frac{\nu_i}{1+\nu_i}\frac{1}{x_i}$$

From the optimality condition for energy inputs goods x_i^f we had:

$$\mu^{f(i)} = \theta_i^f \frac{g''(x_i^f)}{g'(x_i^f)} = -\frac{\nu_i}{1+\nu_i}\frac{1}{x_i^f}\theta_i^f$$

where ν_i is the oil-gas inverse elasticity for country i, and θ_i^f is the supply distortion.
From the optimality condition for q^f we have, with simplification thanks to isoelastic supply:

$$\begin{aligned}
\theta_i^f &= \frac{1}{g'(x_i^f)}\lambda_i(e_i^x - \mathcal{P}_i e_i^f -) \\
\mu^{f(i)} &= -\frac{\nu_i}{1+\nu_i}\frac{1}{x_i^f}\theta_i^f = -\frac{\nu_i}{1+\nu_i}\frac{1}{x_i^f}\frac{1}{g'(x_i^f)}\lambda_i(e_i^x - \mathcal{P}_i e_i^f) = -\frac{\nu_i}{1+\nu_i}\frac{\bar{\nu}_i}{e_i^x}\frac{1+\nu_i}{\bar{\nu}_i}\lambda_i(e_i^x - \mathcal{P}_i e_i^f) \\
\mu^{f(i)} &= \frac{\nu_i}{e_i^x}\lambda_i(\mathcal{P}_i e_i^f - e_i^x)
\end{aligned}$$

As a result, this redistribution terms is positive for net-importers and negative for exporters. This is intuitive: energy dependent countries would like to reduce their dependence to imports, by taxing the energy (and redistributing lump-sum those revenues) they hope to reduce the equilibrium price q^f to benefit for better terms-of-trade. This is the same logic as the global social planner of Appendix C.2, as a smaller scale, and weighted by the country i production e_i^x and inverse elasticity ν_i: a more inelastic supply, with large ν_i would amplify this effect: as usual in Ramsey taxation, it is more relevant to tax inelastic supply goods.

As a result the unilateral carbon tax can write:

$$\xi^f \mathrm{t}_i^\varepsilon = -q^f \frac{\mu_i^{(i)}}{\lambda_i} + q^f \nu_i \frac{\mathcal{P}_i e_i^f - e_i^x}{e_i^x} + \xi^f LCC_i$$

As a result, we see that the optimal carbon tax can become a *subsidy* if the terms-of-trade manipulation motive $\mu_i^{(i)}$ is large enough, if the energy-supply redistribution term is negative, for example for oil-gas exporters, and if the local cost of carbon LCC_i is small.

D Welfare decomposition

D.1 Summary

This welfare decomposition is described in thorough detail in the companion paper Bourany and Rosenthal-Kay (2025), and is inspired Kleinman et al. (2020), where we follow the same steps, using a richer model, with trade a la Armington, energy in production and carbon and trade policy instruments.

There we compute the change in welfare, linearizing the model around the competitive equilibrium where $\mathrm{t}^\varepsilon = \bar{\mathrm{t}}^\varepsilon = 0$ and $\mathrm{t}^b_{ij} = \bar{\mathrm{t}}^b_{ij} = 0$, where policies are identical to the "status-quo". I start from a climate agreement $\mathcal{J}$ of J countries. Those countries are indifferent between being in the club or not, since the policy $(\mathrm{t}^\varepsilon_i, \mathrm{t}^b_{ij}) = (0,0)$ does not affect the equilibrium. I then consider a log-linear perturbation where those policy instruments are increased by a small amount, $d\mathrm{t}^\varepsilon_i$ and $d\mathrm{t}^b_{ij}$ respectively for the club members $i \in \mathcal{J}$. I

denoting $d\ln z_i = \frac{dz_i}{z_i}$ for any variable z_i. The carbon tax and tariffs are increased by $d\mathrm{t}^\varepsilon_i$ and $d\mathrm{t}^b_{ij}$ respectively for the club members $i \in \mathcal{J}$.

$$\mathbf{J} d\mathbf{t}^\varepsilon = \big\{ \mathbb{1}_{\{i\in\mathcal{J}\}}\, d\mathrm{t}^\varepsilon_i \big\}_i \qquad\qquad \bar{\mathbf{J}} \odot d\mathbf{t}^b = \big\{ \mathbb{1}_{\{i\in\mathcal{J}, j\notin\mathcal{J}\}} d\mathrm{t}^b_{ij} \big\}_{ij}$$

with $\mathrm{J} = \mathrm{J}_i = \mathbb{1}_{\{i\in\mathcal{J}\}}$, and $\mathbf{J} \equiv \mathrm{J}_{ki} = \mathbb{1}_{\{i\in\mathcal{J}, j\notin\mathcal{J}\}}$.

The welfare decomposition of individual country i, defined as $\mathcal{U}_i = u(\{c_{ij}\}_j)$ the indirect utility is computed as the consumption-equivalent welfare change:

$$\frac{d\mathcal{U}_i}{u'(c_i)c_i} = \eta^c_i d\ln \mathrm{p}_i + \Big[-\eta^c_i \bar{\gamma}_i \tfrac{1}{\bar{\nu}} - \eta^c_i s^e_i s^f_i + \eta^\pi_i (1 + \tfrac{1}{\bar{\nu}}) \Big] d\ln q^f - \Big[\eta^c_i s^e_i (s^c_i + s^r_i) + \eta^\pi_i \tfrac{1}{\nu} + 1 \Big] d\ln \mathbb{P}_i$$

where $\eta^c_i = \frac{y_i \mathrm{p}_i}{x_i}$, with $x_i = c_i \mathbb{P}_i$ is the ratio of final good output in comparison to consumption – which can also come from energy rent. The counterpart is $\eta^\pi_i = \frac{\pi^f_i}{x_i}$. The energy share $s^e_i = \frac{e_i q^e_i}{y_i \mathrm{p}_i}$ and the share of oil-gas/coal/renewable $s^\ell_i = \frac{e^\ell_i q^\ell_i}{e_i q^e_i}$ governs the impact of energy prices. The aggregate supply elasticity $\bar{\nu} = \big(\sum_i \lambda^x_i \nu_i^{-1} \big)^{-1}$ represents the oil-gas supply curve, and the climate damage $\bar{\gamma}_i = \gamma(T_i - T^\star_i) T_i\, s^{E/S}$ is represented in a static fashion – with $\mathcal{E}$ the emission of that period and $s^{E/S} = \mathcal{E}/\mathcal{S}$ with $\mathcal{S}$ the carbon concentration in the atmosphere.

We observe that most of the impacts arise through aggregate quantity of emissions and fossil fuels consumption, which then affect world prices q^f. For conciseness, I express all the General Equilibrium effects on fossil quantities as a function of price q^f: $d\ln q^f \approx \bar{\nu} d\ln E^f + \dots$

The countries affected the most by a change in equilibrium quantity of fossil fuels consumed E^f, price q^f, and thus by carbon taxation, are the countries with high sensitivity to $d\ln q^f$. A reduction in fossil demand benefits the countries that have large damages from climate changes $\bar{\gamma}_i$, as well as large energy share from fossil s^f_i. This latter effect dampens the cost of taxation: if a larger coalition lower energy demand, it benefits other countries through a reduction in fossil price. This is called the "energy price leakage effect" in the literature. However, this decrease in price hurt fossil fuel producers as it dries out their energy rents, as summarized by $\eta^\pi_i(1+\frac{1}{\bar{\nu}})$. Moreover, there are additional equilibrium effects through trade and good prices p_i and $\mathbb{P}_i$ as we see below.

To see the direct effect of carbon taxation – at the intensive margin – and the extensive margin effect of the size of the club J_i, I simplify the model further to obtain an analytical formula for the fossil price. In the following, I assume that the energy mix is concentrated on oil and gas $s_i^f = 1, s_i^c = s_i^r = 0$. The details of the derivation are provided in Appendix D:

$$d \ln q^f = - \frac{\bar{\nu}}{1+\bar{\gamma}+\mathbb{C}\mathrm{ov}_i(\widetilde{\lambda}_i^f, \bar{\gamma}_i) + \bar{\nu}\overline{\lambda}^{\sigma,f}} \sum_i \widetilde{\lambda}_i^f \mathrm{J}_i d\mathrm{t}_i^{\varepsilon} + \sum_i \beta_i d \ln \mathrm{p}_i$$

with the market share $\lambda_i^f = \frac{\mathcal{P}_i e_i^f}{E^f}$ for fossil, and weighted by elasticity $\widetilde{\lambda}_i^f = \lambda_i^f \frac{\sigma^y}{1-s_i^e}$ and its average $\overline{\lambda}^{\sigma,f} = \sum_i \widetilde{\lambda}_i^f \frac{\sigma^y}{1-s_i^e}$.

As we see, the higher the carbon tax $d\mathrm{t}_i^{\varepsilon}$ – at the intensive margin – or the size of the club J_i – at the extensive margin – the more it reduces the demand for fossil-fuel energy, and hence lower the fossil-fuel price q^f, by a factor $\bar{\nu}$ – the aggregate energy supply elasticity. Moreover, the energy curve q^f is affected by climate change: more emissions imply larger damages $\bar{\gamma} = \sum_i \bar{\gamma}_i$, which in turn reduces energy demand and hence emissions. The price impact of taxation is higher – analogous to the slope of the demand curve – as we see in the denominator of the first term. Moreover, the covariance term indicates that if the large energy consumers are also the most affected by climate change– with a larger share of the market λ_i^f and high elasticity σ in the term $\widetilde{\lambda}_i^f = \lambda_i^f \frac{\sigma^y}{1-s_i^e}$ associated with larger cost $\bar{\gamma}_i$, this effect is stronger and the demand curve is even steeper and more inelastic.

Moreover, carbon taxation $\mathrm{t}_j^{\varepsilon}$ and tariffs t_{ij}^b have large general equilibrium effects, through leakage and reallocation, and affect y_i and p_i. I can compute these changes in prices in equilibrium:

$$\begin{aligned} d \ln \mathrm{p} = \mathbf{A}^{-1} \Big[&- (\mathbf{I}-\mathbf{T}\odot v^y)\alpha^{y,qf} + \mathbf{T}(v^{e^x} \odot \tfrac{1}{\nu} + v^{e^f} \tfrac{\sigma^y}{1-s^e} + v^{ne}) - \Big((\mathbf{I}-\mathbf{T}\odot v^y)\alpha^{y,z} - \tfrac{\sigma^y}{1-s^e} \Big) \bar{\gamma} \tfrac{1}{\bar{\nu}} \Big] d \ln q^f \\ &+ \Big[- (\mathbf{I}-\mathbf{T}\odot v^y)\alpha^{y,qf} + \mathbf{T}(v^{e^f} \odot \tfrac{\sigma^y}{1-s^e}) \Big] \odot \mathrm{J} d \ln \mathrm{t}^{\varepsilon} + \theta (\mathbf{TS} \odot \mathbf{J} \odot d\ln \mathrm{t}^b - \mathbf{T}(\mathbb{1}+\mathbf{S}') \odot (\mathbf{J} \odot d\ln \mathrm{t}^b)') \end{aligned}$$

with parameters: $\mathbf{S}$ for the trade share matrix, $\mathbf{T}$ income flow matrix, θ, Armington CES. Moreover, the general equilibrium (and leakage) effects are summarized in a complicated matrix $\mathbf{A}$ that summarizes the fact that the price p_i also affects energy demand, oil-gas extraction, energy trade balance and output. Further description can be found in the companion paper Bourany and Rosenthal-Kay (2025).

E Climate agreement – Solution method

In this section, we detail the general formulation for the climate agreement design, which join together a policy choice and a combinatorial discrete choice problem.

E.1 Inner problem and Outer problem

The world welfare maximization requires to separate the maxima between the choice of countries and the choice of policy instruments $(\mathbb{J}, \mathbf{t})$. We can split the choice in two ways:

$$\begin{aligned}\max_{\mathbb{J},\mathbf{t}} \sum_i \omega_i \mathcal{U}_i(\mathbb{J}, \mathbf{t}) &= \max_{\mathbb{J}} \max_{\mathbf{t}(\mathbb{J})} \sum_i \omega_i \mathcal{U}_i(\mathbb{J}, \mathbf{t}) \\ &= \max_{\mathbf{t}} \max_{\mathbb{J}(\mathbf{t})} \sum_i \omega_i \mathcal{U}_i(\mathbb{J}, \mathbf{t})\end{aligned}$$

where $\mathcal{U}_i(\mathbb{J}, \mathbf{t})$ the indirect utility can also be written with

$$\mathcal{U}_i(\mathbb{J}, \mathbf{t}) = \mathcal{U}_i(\mathbb{J}, \mathrm{t}^f, \mathrm{t}^b) = \mathcal{U}_i(I_i, \{I_j\}_{j \neq i}, \mathrm{t}^f, \mathrm{t}^b)$$

with indicators $I_i = \mathbb{1}\{i \in \mathbb{J}\}$.

Depending on how we split the joint problem, it leads to different treatments. In addition, I also add K additional constraints $g_k(\mathbb{J}, \mathbf{t}) \geq 0$ that could arise in case of policy constraints and imperfect instruments: for example if a preexisting trade agreement constrains the value of tariffs, or if a political/security agreement may not be feasible because of geopolitical motives.

First representation: policy for each coalition

In this first, naive, approach, I solve for the optimal policy (carbon tax and tariffs) for every possible coalition:

$$\begin{aligned}\max_{\mathbb{J} \in \mathcal{P}(\mathbb{I}),\, \mathbf{t} \in \mathbb{T}} \mathcal{W}(\mathbb{J}, \mathbf{t}) &= \max_{\mathbb{J}} \max_{\mathbf{t} \in \mathbb{T}} \sum_i \omega_i \mathcal{U}_i(\mathbb{J}, \mathbf{t}) && (P_1) \\ s.t. \qquad \mathcal{U}_i(\mathbb{J}, \mathbf{t}) &\geq \mathcal{U}_i(\mathbb{J} \backslash \{i\}, \mathbf{t}) \qquad \forall\, i \in \mathbb{J} && [\lambda_i] \\ g_k(\mathbb{J}, \mathbf{t}) &\geq 0 \qquad [\mu_k]\end{aligned}$$

Define the marginal gain of $j \in \mathbb{J}$ for utility of agent i

$$\Delta_j \mathcal{U}_i(\mathbb{J}, \mathbf{t}) = \mathcal{U}_i(\mathbb{J} \cup \{j\}, \mathbf{t}) - \mathcal{U}_i(\mathbb{J} \backslash \{j\}, \mathbf{t})$$

We define the Lagrangian function for every possible coalition $\mathbb{J}$

$$\mathcal{L}(\mathbf{t}, \boldsymbol{\lambda}, \boldsymbol{\mu} | \mathbb{J}) = \mathcal{W}(\mathbb{J}, \mathbf{t}) + \sum_{i \in \mathbb{J}} \lambda_i \Delta_i \mathcal{U}_i(\mathbb{J}, \mathbf{t}) + \sum_k \mu_k g_k(\mathbb{J}, \mathbf{t})$$

Rewriting and using standard optimization arguments for the Lagrangian, and then maximizing over $\mathbb{J}$:

$$\max_{\mathbb{J}} \max_{\mathbf{t}} \min_{\boldsymbol{\lambda},\boldsymbol{\mu}} \mathcal{L}(\mathbb{J}, \mathbf{t}, \boldsymbol{\lambda}, \boldsymbol{\mu}) \leq \max_{\mathbb{J}} \min_{\boldsymbol{\lambda},\boldsymbol{\mu}} \max_{\mathbf{t}} \mathcal{L}(\mathbb{J}, \mathbf{t}, \lambda)$$

If $\mathcal{W}$ is concave in $\mathbf{t}$, for every $\mathbb{J}$, where the optimal $\mathbf{t}$ and λ are implicit functions of $\mathbb{J}$

$$\begin{aligned}(P_1(\mathbb{J})) = \max_{\mathbf{t}} \min_{\boldsymbol{\lambda},\boldsymbol{\mu}} \mathcal{L}(\mathbf{t}, \boldsymbol{\lambda}|\mathbb{J}) &= \min_{\boldsymbol{\lambda},\boldsymbol{\mu}} \max_{\mathbf{t}} \mathcal{L}(\mathbf{t}, \boldsymbol{\lambda}, \boldsymbol{\mu}|\mathbb{J}) \\ &= \min_{\boldsymbol{\lambda},\boldsymbol{\mu}} \max_{\mathbf{t}} \sum_{i\in\mathbb{I}} \omega_i \mathcal{U}_i(\mathbb{J}, \mathbf{t}) + \sum_{i\in\mathbb{J}} \lambda_i \Delta_i \mathcal{U}_i(\mathbb{J}, \mathbf{t}) + \sum_k \mu_k g_k(\mathbb{J}, \mathbf{t})\end{aligned}$$

Step 1: First, if the constraints functions $\mathbf{t} \to \Delta_i \mathcal{U}_i(\mathbb{J}, \mathbf{t})$ and $\mathbf{t} \to g_k(\mathbb{J}, \mathbf{t})$ maps into $\mathbb{R}^2_+$ (for $\mathbf{t} \in \mathbb{R}^2$), then we can solve the first-order conditions, with the gradient $D_{\mathbf{t}}$:

$$\begin{aligned}\mathbf{t}^\star(\mathbb{J}) \qquad s.t. \qquad & \sum_{i\in\mathbb{I}} \omega_i D_{\mathbf{t}} \mathcal{U}_i(\mathbb{J}, \mathbf{t}^\star) + \sum_{i\in\mathbb{J}} \lambda_i D_{\mathbf{t}} \Delta_i \mathcal{U}_i(\mathbb{J}, \mathbf{t}^\star) + \sum_k \mu_k D_{\mathbf{t}} g_k(\mathbb{J}, \mathbf{t}) = 0 \\ & \sum_{i\in\mathbb{J}} \lambda_i \Delta_i \mathcal{U}_i(\mathbb{J}, \mathbf{t}^\star) = 0 \\ & \sum_k \mu_k g_k(\mathbb{J}, \mathbf{t}^\star) = 0\end{aligned}$$

where the two last equations are the complementary slackness

Step 2: Given $\mathbf{t}^\star(\mathbb{J})$, the dual problem allows to recover the multipliers:

$$\min_{\boldsymbol{\lambda},\boldsymbol{\mu}} \sum_{i\in\mathbb{I}} \omega_i \mathcal{U}_i(\mathbb{J}, \mathbf{t}^\star) + \sum_{i\in\mathbb{J}} \lambda_i \Delta_i \mathcal{U}_i(\mathbb{J}, \mathbf{t}^\star) + \sum_k \mu_k g_k(\mathbb{J}, \mathbf{t}^\star)$$

Step 3: Given $\mathbf{t}^\star(\mathbb{J}), \boldsymbol{\lambda}(\mathbb{J}), \boldsymbol{\mu}(\mathbb{J})$, we have to solve the combinatorial discrete choice problem:

$$\max_{\mathbb{J}} \sum_{i\in\mathbb{I}} \omega_i \mathcal{U}_i(\mathbb{J}, \mathbf{t}^\star(\mathbb{J}))$$

This approach suffers from major problems: First, it can be costly to solve iteratively the primal-dual problem for $\{\mathbf{t}^\star(\mathbb{J}), \boldsymbol{\lambda}(\mathbb{J}), \boldsymbol{\mu}(\mathbb{J})\}$ for every combination of countries $\mathbb{J}$. Second, in a subgame perfect equilibrium, the problem becomes nested: the policy and multipliers $\{\mathbf{t}^\star(\mathbb{J}), \boldsymbol{\lambda}(\mathbb{J}), \boldsymbol{\mu}(\mathbb{J})\}$ depends on the allocation in every deviation $\mathcal{U}_j(\mathbb{J}\backslash\{j\})$, $\forall j \in \mathbb{J}$, and hence on the policy and multipliers $\{\mathbf{t}^\star(\mathbb{J}\backslash\{j\}), \boldsymbol{\lambda}(\mathbb{J}\backslash\{j\}), \boldsymbol{\mu}(\mathbb{J}\backslash\{j\})\}$ which also depends on the allocation of every deviation $\mathbb{J}\backslash\{j, k\}$, $\forall j, k \in \mathbb{J}$, etc. This becomes a joint problem where the allocation of each coalition $\mathbb{J}$ is linked to the allocation of all the other sub coalitions, and it requires solving them jointly: a problem that becomes untractable for any realistic quantitative model.

Second representation: coalition for each policy

In this second approach, I fix the policy $\mathbf{t}$ (carbon tax and tariffs) and solve the model for every possible *stable* coalitions $\mathbb{J}$:

$$\begin{aligned}
&\max_{\mathbb{J}\in\mathcal{P}(\mathbb{I}),\,\mathbf{t}\in\mathbb{T}} \mathcal{W}(\mathbb{J},\mathbf{t}) = \max_{\mathbf{t}\in\mathbb{T}} \max_{\mathbb{J}} \sum_i \omega_i \mathcal{U}_i(\mathbb{J},\mathbf{t}) && (P_2)\\
s.t. \quad & \mathcal{U}_i(\mathbb{J},\mathbf{t}) \geq \mathcal{U}_i(\mathbb{J}\backslash\{i\},\mathbf{t}) \qquad \forall\, i\in\mathbb{J} \qquad [\lambda_i]\\
& g_k(\mathbb{J},\mathbf{t}) \geq 0 \qquad [\mu_k]
\end{aligned}$$

For every policy $\mathbf{t}$ we can solve for the optimal coalition $\mathbb{J}^\star(\mathbf{t})$, by solving the combinatorial discrete choice problem:

$$\max_{\mathbb{J}\in\mathcal{P}(\mathbb{I})} \mathcal{L}(\mathbb{J}|\mathbf{t}) = \max_{\mathbb{J}} \sum_i \omega_i \mathcal{U}_i(\mathbb{J},\mathbf{t}) + \sum_{i\in\mathbb{J}} \lambda_i(\mathbf{t})\Delta_i\mathcal{U}_i(\mathbb{J},\mathbf{t}) + \sum_k \mu_k(\mathbf{t}) g_k(\mathbb{J},\mathbf{t})$$

Note that this $\mathcal{L}(\mathbb{J}|\mathbf{t})$ is defined with a slight abuse of notation given that this problem is discrete: it is well-defined only for $\lambda_i(\mathbf{t}) = +\infty$ if $\Delta_i\mathcal{U}_i(\mathbb{J},\mathbf{t}) < 0$ and $\lambda_i(\mathbf{t}) = 0$ otherwise when the stability constraint is satisfied. Similarly $\mu_i(\mathbf{t}) = +\infty$ if $g_k(\mathbb{J},\mathbf{t}) < 0$ and $\mu_i(\mathbf{t}) \geq 0$, otherwise.

This yields the optimal $\mathbb{J}^\star(\mathbf{t})$ that maximizes welfare subject to participation constraints and policy constraints, and this for every policy $\mathbf{t}$. The procedure to solve this combinatorial discrete choice problem is detailed in Appendix E.2.

Finally, after we solve the model and the optimal coalition $\mathbb{J}^\star(\mathbf{t})$ for every policy $\mathbf{t}$, the maximization of welfare implies to solve:

$$\max_{\mathbf{t}} \sum_{i\in\mathbb{I}} \omega_i \mathcal{U}_i(\mathbb{J}^\star(\mathbf{t}),\mathbf{t})$$

Note that the constraints due to participation and policy constraints disappear, as they are already embedded implicitly in $\mathbb{J}^\star(\mathbf{t})$.

This problem becomes noncontinuous, non-differentiable, and a priori non-convex. We can not use the Kuhn-Tucker theorem or any standard results of calculus of variation. For that reason, I solve it using a grid search over $\mathbf{t}$. This is computationally expensive, but it is feasible in finite time and with reasonable accuracy, provided that the instruments we are searching over are in small dimensions $\mathbf{t} = \{\mathrm{t}^\varepsilon, \mathrm{t}^b\} \in \mathbb{R}^2$.

In extensions, we consider additional instruments: transfers, with a simple rule indexed by a parameter α, as explained in Section 7.1, and fossil-fuels-specific tariffs, using a simple rule indexed by β, as explained in Section 7.3. As a result, the maximization amounts to searching over a grid $\mathbf{t} = \{\mathrm{t}^\varepsilon, \mathrm{t}^b, \alpha, \beta\} \in \mathbb{R}^4$. This is feasible in several hours for a relatively coarse grid for α, β but allows us to solve this problem both for optimal policies and optimal coalition, which was previously intractable.

E.2 Combinatorial discrete choice problem

Second method: Squeezing procedure for CDCP with Participation Constraints Second, since full enumeration is costly, I provide an alternative algorithm inspired by methods used in the international trade literature to solve combinatorial discrete choice problems. The additional difficulty that needs to be considered is the presence of participation constraints. In this section, we only consider unilateral deviations. The idea behind this method is greatly inspired by Arkolakis, Eckert and Shi (2023) and Farrokhi and Lashkaripour (2024).

The idea is to build iteratively sets that are lower bound $\underline{\mathcal{J}}$ and upper bound $\overline{\mathcal{J}}$ sets for the optimal coalition $\mathbb{J}$: $\underline{\mathcal{J}}$ is a subset which *includes* all the countries that we know to be part of the optimal set $\mathbb{J}$ and $\overline{\mathcal{J}}$ is a superset, such that it excludes the countries that we know are not part of the optimal set. The set $\overline{\mathcal{J}}\backslash\underline{\mathcal{J}}$ is the set of potential countries. The natural starting point is $\underline{\mathcal{J}} = \varnothing$, $\overline{\mathcal{J}} = \mathbb{I}$.

The squeezing step in standard CDCP is a mapping from $\mathcal{J}$ to members that bring a positive marginal value to the objective $\mathcal{W}(\mathbb{J}) := \mathcal{W}(\mathbb{J}, \mathbf{t})$. The modification needed in our setting with participation constraints is that the country also needs have marginal *individual* value $\mathcal{U}_i(\mathcal{J}) = \mathcal{U}_i(\mathcal{J}, \mathbf{t})$ to be part of the coalition:

$$\Phi(\mathcal{J}, \mathbf{t}) \equiv \{j \in \mathbb{I} \,|\, \Delta_j \mathcal{W}(\mathcal{J}, \mathbf{t}) > 0 \;\&\; \Delta_j \mathcal{U}_j(\mathcal{J}, \mathbf{t}) > 0\} \tag{21}$$

where the marginal values for global welfare and individual welfare are

$$\Delta_j \mathcal{W}(\mathcal{J}, \mathbf{t}) \equiv \mathcal{W}(\mathcal{J} \cup \{j\}, \mathbf{t}) - \mathcal{W}(\mathcal{J}\backslash\{j\}, \mathbf{t}) = \sum_{i \in \mathbb{I}} \mathcal{P}_i \omega_i \big(\mathcal{U}_i(\mathcal{J} \cup \{j\}, \mathbf{t}) - \mathcal{U}_i(\mathcal{J}\backslash\{j\}, \mathbf{t})\big)$$
$$\Delta_j \mathcal{U}_i(\mathcal{J}, \mathbf{t}) \equiv \mathcal{U}_i(\mathcal{J} \cup \{j\}, \mathbf{t}) - \mathcal{U}_i(\mathcal{J}\backslash\{j\}, \mathbf{t})$$

The iterative procedure builds the lower bound $\underline{\mathcal{J}}$ and upper bound $\overline{\mathcal{J}}$ by successive application of the squeezing step.

$$\underline{\mathcal{J}}^{(k+1)} = \Phi(\underline{\mathcal{J}}^{(k)}, \mathbf{t}) \qquad\qquad \overline{\mathcal{J}}^{(k+1)} = \Phi(\overline{\mathcal{J}}^{(k)}, \mathbf{t}) \tag{22}$$

Under some conditions – complementarity, as defined next section – this sequential procedure yields two sets $\underline{\mathcal{J}}$ and $\overline{\mathcal{J}}$ such that $\underline{\mathcal{J}} \subseteq \mathbb{J} \subseteq \overline{\mathcal{J}}$. In some cases $\underline{\mathcal{J}} = \overline{\mathcal{J}} = \mathbb{J}$ which gives the optimal coalition. If not, with $\overline{\mathcal{J}}\backslash\underline{\mathcal{J}} = \mathcal{J}^{pot}$, we find the optimal coalition by searching exhaustively over all coalitions in:

$$\mathcal{J}^{rem} = \{\mathcal{J} \,|\, \mathcal{J} = \underline{\mathcal{J}} \cup \hat{\mathcal{J}}, \text{ with } \hat{\mathcal{J}} \in \mathcal{P}(\mathcal{J}^{pot})\}$$

Applicability of the squeezing procedure

From the combinatorial discrete choice literature, Arkolakis, Eckert and Shi (2023), we know that the squeezing procedure applies in cases where the model exhibit "complementarity" or single-crossing differences in choices.

Indeed, we say that the objective $\mathcal{W}(\mathcal{J})$ obeys the property of *single crossing differences in choice* (SCD-C) from below if:

$$\Delta_j \mathcal{W}(\mathcal{J}) \geq 0 \quad \Rightarrow \quad \Delta_j \mathcal{W}(\mathcal{J}') \geq 0 \qquad \text{for } \mathcal{J} \subset \mathcal{J}' \quad \& \quad j \in \mathbb{I}$$

A simple sufficient condition for SCD-C, from below to be respected is that the marginal value of the objective is monotone in the set $\mathcal{J}$, also called "complementarity":

$$\Delta_j \mathcal{W}(\mathcal{J}) \leq \Delta_j \mathcal{W}(\mathcal{J}') \qquad \text{for } \mathcal{J} \subseteq \mathcal{J}' \qquad \& \quad j \in \mathbb{I}$$

Theorem (Arkolakis, Eckert and Shi (2023)) The SCD-C from below is *sufficient* for the application of squeezing algorithm to yield $\underline{\mathcal{J}} \subseteq \mathbb{J} \subseteq \overline{\mathcal{J}}$ in standard CDCPs.

In this setting, considering participation constraints requires to adjust the welfare objective, from $\mathcal{W}(\mathbb{J})$ to $\widetilde{\mathcal{W}}(\mathbb{J})$ as in eq. (17). In this context, the single crossing differences in choice with participation constraints (SCD-C, PC) take an intricate form, which we detail below:

$$\left\{\begin{array}{l} \Delta_i \mathcal{U}_i(\mathcal{J} \cup \{j\}) \geq 0 \\ \& \left[\begin{array}{l} \Big(\Delta_j \mathcal{W}(\mathcal{J} \cup \{j\}) \geq 0 \quad \& \quad \Delta_i \mathcal{U}_i(\mathcal{J}) \geq 0 \Big) \\ \text{or} \quad \Delta_i \mathcal{U}_i(\mathcal{J}) < 0 \end{array}\right. \end{array}\right. \quad \Rightarrow \quad \left\{\begin{array}{l} \Delta_i \mathcal{U}_i(\mathcal{J}' \cup \{j\}) \geq 0 \\ \& \left[\begin{array}{l} \Big(\Delta_j \mathcal{W}(\mathcal{J}' \cup \{j\}) \geq 0 \quad \& \quad \Delta_i \mathcal{U}_i(\mathcal{J}') \geq 0 \Big) \\ \text{or} \quad \Delta_i \mathcal{U}_i(\mathcal{J}') < 0 \end{array}\right. \end{array}\right.$$
$$\forall\, \mathcal{J} \subseteq \mathcal{J}' \qquad \forall\, j \in \mathbb{I} \qquad \text{(SCD-C, PC)}$$

An intuition for this condition is the following: IF the coalition $\mathcal{J}$ makes (i) allocation outcomes better for welfare with $\{j\}$, if both $\mathcal{J}$ and $\mathcal{J} \cup \{j\}$ are stable, or (ii) the coalition $\mathcal{J} \cup \{j\}$ is stable if $\mathcal{J}$ is unstable, THEN one of these conditions should also be respected for larger coalitions $\mathcal{J}' \supseteq \mathcal{J}$.

The following condition is sufficient for (SCD-C, PC) and provides intuitions on the trade-offs at play in the construction of the optimal coalition:

$$\left\{\begin{array}{rl} \Delta_j \mathcal{W}(\mathcal{J}, \mathbf{t}) & \leq \Delta_j \mathcal{W}(\mathcal{J}', \mathbf{t}) \\ 0 \leq \Delta_i \mathcal{U}_i(\mathcal{J} \cup \{j\}, \mathbf{t}) & \leq \Delta_i \mathcal{U}_i(\mathcal{J}' \cup \{j\}, \mathbf{t}) \qquad \forall\, i \in \mathcal{J} \cup \{j\} \;\&\; i \in \mathcal{J}' \cup \{j\} \\ 0 \leq \Delta_i \mathcal{U}_i(\mathcal{J}, \mathbf{t}) & \leq \Delta_i \mathcal{U}_i(\mathcal{J}', \mathbf{t}) \qquad \forall\, i \in \mathcal{J} \;\&\; i \in \mathcal{J}' \end{array}\right. \tag{23}$$

or

$$\left\{\begin{array}{l} 0 \leq \Delta_i \mathcal{U}_i(\mathcal{J} \cup \{j\}, \mathbf{t}) \leq \Delta_i \mathcal{U}_i(\mathcal{J}' \cup \{j\}, \mathbf{t}) \qquad \forall\, i \in \mathcal{J} \cup \{j\} \;\&\; i \in \mathcal{J}' \cup \{j\} \\ \exists i \in \mathcal{J} \Delta_i \mathcal{U}_i(\mathcal{J}, \mathbf{t}) < 0 \quad \& \quad \exists i \in \mathcal{J}' \quad \Delta_i \mathcal{U}_i(\mathcal{J}', \mathbf{t}) < 0 \end{array}\right. \tag{24}$$
$$\text{for } \forall\, \mathcal{J} \subseteq \mathcal{J}' \subseteq \mathbb{I} \qquad \& \quad j \in \mathbb{I}$$

This sufficient condition states either one of these cases is verified: (1) in the first case of eq. (23), the marginal welfare $\Delta_j \mathcal{W}$ is monotone in $\mathcal{J}$: the welfare gain of adding country j grows with the size of the coalition $\mathcal{J}$. Moreover, the participation constraint of each member i is still respected when we include country j, and this monotonically in the coalition, from $\mathcal{J}$ to $\mathcal{J}'$, and the coalition is also stable without j. (2) In the second case of eq. (24), we do not require any condition of global welfare, but the participation constraint of each member i is respected when including country j, while it is violated when j is not present in $\mathcal{J}$ and $\mathcal{J}'$. Either one of these two conditions needs to be respected for every pairs of sets $\mathcal{J} \subset \mathcal{J}'$ and every country j.

This condition, as well as its weaker counterpart above (SCD-C-PC), are sufficient conditions for SCD-C from below for $\widetilde{\mathcal{W}}$. It shows that the requirements for coalition building are much stronger as they need to verify if adding marginal members still satisfies the participation constraints of all the incumbent members. In this context, the modified squeezing steps account for such constraint and thus:

Theorem The SCD-C-PC from below is *sufficient* for the application of modified squeezing algorithm, i.e. successive application of eq. (21), starting from $\{\varnothing, \mathbb{I}\}$ and eq. (22), to yield bounding sets $\underline{\mathcal{J}} \subseteq \mathbb{J} \subseteq \overline{\mathcal{J}}$ in CDCPs with participation constraints.

One of the advantages of this setting is that, for a small number of countries $\#\mathbb{I} \approx 10$, we can evaluate numerically if the sufficient conditions mentioned above are satisfied. The disadvantage of the model displayed above is that the large amount of heterogeneity, general equilibrium effects through energy markets, and international trade, prevent the simple evaluation of those sufficient conditions analytically.