SolACE - Solar Geoengineering in an Analytic Climate Economy

Felix D. Meier[1] and Christian P. Traeger[2]

August 2022

Abstract

Solar geoengineering is an affordable measure to counteract the global temperature increase. We derive a simple policy rule for sulfur-based geoengineering in a state of the art integrated assessment model of climate change. We show how geoengineering affects optimal carbon taxation, deriving the different components of the Pigovian tax. We show how the globally optimal rational for geoengineering and carbon taxation changes in a dynamic Markov game across regions. A quantitative simulation suggests a non-cooperative equilibrium where China as a single mover reduces temperatures to the Paris accord's most stringent 1.5°C target by 2100. It "free-drives" on the margin but all regions apart from Russia free-ride in absolute terms. The simulated temperature increase peaks around 2165 exceeding the 2°C target.

JEL Codes: C72, D62, H41, Q54

Keywords: solar geoengineering, climate change, dynamic games, social cost of carbon, optimal carbon tax, free-riding, free-driving

[1]Kiel Institute for the World Economy (felix.meier@ifw-kiel.de); German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig. [2]Department of Economics, University of Oslo (christian.traeger@econ.uio.no); ifo Institute, Munich; Frisch Centre, Oslo. We are grateful for feedback from Bard Harstad, Daniel Heyen, Terry Iverson, David Kelly, Christoph Kleinschmitt, Juan Moreno-Cruz, Billy Pizer, Ulrich Platt, Martin Quaas, Francesco Ricci, Wilfried Rickels, Till Requate, Daniele Visioni, Johannes Quaas, and the audiences at the RFF workshop 2022, WCNRM 2022, VfS 2021, Montpellier 2021, EUROFRAME 2021, University of Miami 2020, SURED 2020, ASSA 2020, EAERE 2019, and Tou 2019. We gratefully acknowledge funding from the German Research Foundation (DFG) under grant agreement number RI 1833/4-1 (CDRecon), and the European Union's Horizon 2020 research and innovation program under grant agreement number 869357 (OceanNETs).

"The climatic changes that may be produced by the increased CO_2 content could be deleterious from the point of view of human beings. The possibilities of deliberately bringing about countervailing climatic changes therefore need to be thoroughly explored. A change in the radiation balance in the opposite direction to that which might result from the increase of atmospheric CO_2 could be produced by raising the albedo, or reflectivity, of the earth. Such a change in albedo could be brought about, for example by spreading very small reflecting particles ... Rough estimates indicate that ... costs ... do not seem excessive." (President's Science Advisory Committee 1965, page 127)

1 Introduction

Worldwide greenhouse gas emissions are still on the rise (Tollefson 2021). Future warming will substantially reduce future output apart from destroying ecosystems and driving species to extinction (Hoegh-Guldberg and Bruno 2010, Burke et al. 2015, Urban 2015, Howard and Sterner 2017). Scientific models and observations from large volcanic eruptions suggest that the deliberate injection of sulfur aerosols into the stratosphere can cool our planet, reflecting sunlight back into space (Crutzen 2006). Solar geoengineering (SG) adds a policy instrument that can be illustrated by applying a potentially damaging sunscreen to planet Earth. It can be deployed in a globally coordinated socially optimal effort or by strategically acting regions. So far, the weak international governance of SG makes a globally coordinated deployment unlikely and National Academies of Sciences (2021) and Aldy et al. (2021) call for more realistic models to investigate strategic interactions among potential SG actors.

Our paper develops a deeper understanding of socially optimal SG and develops a novel quantitative model of SG in a world of strategically interacting regions. We derive analytic formulas for optimal sulfur deployment (the sunscreen) in an integrated assessment model of climate change. Simple formulas explain a social planner's and an individual region's optimal sulfur deployment as functions of atmospheric carbon dioxide concentrations, climatic parameters, SG efficiency and damages, and other cost and damage parameters. Geoengineering our climate is a policy instrument that addresses the symptom (warming) rather than the cause of climate change (greenhouse gases). Counteracting the warming can reduce the incentive to cut back on greenhouse gas emissions. The social cost of carbon (SCC) characterizes the incentives to cut emissions. We develop analytic formulas for the SCC under SG in both the social planner and the strategic setting. These formulas explain the drivers of the emission reduction. In the regional Markov game, our formulas show that and why mitigation incentives can also increase under SG as a result of heterogeneous damages and strategic interactions.

Current emissions are far from the trajectories pledged in international agreements (UNEP 2021, CAT 2021) and from those suggested by global social planner models (e.g. Hänsel et al. 2020). Thus, the social planner model is but a benchmark.

Our regional model integrates SG into an updated version of Nordhaus's (2010) RICE model enriched with elements of the Analytic Climate Economy (Traeger 2022). We formulate a dynamic Markov game where two regions have the potential to undertake SG (as well as countermeasures) and all players' respond in terms of their other choice variables, which include CO_2 emissions. The setting incorporates four relevant externalities. First, emissions exceed the (globally) socially optimal level and the regional carbon prices are too low. Second and third, sulfur travels across regions causing both cooling (a positive[1] externality) and damages (a negative externality). Both of these externalities tie directly to the sulfur spillover. The fourth externality is a direct heat transfer across regions. By cooling one region, we cool other regions even in the absence of sulfur spillovers.[2] The operational costs of SG are relatively low when compared to the costs of climate change (Klepper and Rickels 2012, McClellan et al. 2012, Moriyama et al. 2017, Smith and Wagner 2018). We find that a single strategic actor can help the world to meet the Paris Accord's 1.5°C target by 2100 out of mere self-interest (though temperatures in 2100 are still increasing). We discuss the roles of marginal versus overall free-riding and free-driving. The latter terminology was introduced by Weitzman (2015) as a proxy for one region imposing an externality on many. For a free-rider, the SCC is minimal when the marginal externality from sulfur's temperature and damage impact is maximal.[3] In our quantitative simulation, this marginal externality is negative for most regions, whereas the overall externality from a strategically acting "free-driver" is positive for all regions except for Russia.

Literature. Already Schelling (1996) discusses the potential strategic implications of SG. Since his work, many authors found that optimally deployed SG can reduce climate damages (Nordhaus and Boyer 2000, Moreno-Cruz et al. 2012, Bahn et al. 2015). Nordhaus and Boyer (2000) and Rickels et al. (2020) discuss the regional heterogeneity of the SG impact in integrated assessment models, singling out winners and losers. Harding et al. (2020) find that SG can lower income inequality between countries. Barrett (2008), Millard-Ball (2012), and Urpelainen (2012) discuss strategic deployment of SG in stylized (static) games. Ricke et al. (2013) develop a two stage game where countries can form SG coalitions.

[1] In difference to some of the earlier literature, our regions do not have exogenously set target temperatures. Differences in preferred temperatures arise endogenously as a result of differences in climate change and SG damages. In the absence of damages and operational costs both regions would generally like to reduce temperatures.

[2] This fourth externality implies some interesting but second order modifications of the equilibrium strategies and the SCC. Given direct heat exchange does not change the qualitative findings, we relegate its impact on the analytic formulas to Appendix A.2 and only incorporate it in the quantitative Section 4 of the main body of the paper.

[3] Maximal in the sense that we vary the domain of the active players characteristics to generate the highest possible externality (usually a positive externality). More generally, we show that the marginal externality increases concavely in the carbon concentration and falls convexly in the sulfur level (Corollary 3).

Harding and Moreno-Cruz (2016), Heutel et al. (2016) and Flegal et al. (2019) provide summaries of this literature.

The literature distinguishes different effects of SG on climate policy. Keith (2000), Robock (2008), Morrow (2014) and Quaas et al. (2017) discuss whether research into SG could reduce the political will for traditional mitigation efforts and push society onto a "slippery slope". Goeschl et al. (2013) point out that SG can increase emission incentives as an optimal response by partly substituting away from mitigation. Moreno-Cruz (2015) and Moreno-Cruz and Smulders (2017) show that the impact of SG on mitigation depends on the similarity between countries. For similar countries, the option of SG leads to lower mitigation levels. However, when damages differ across countries, mitigation levels can increase. Our framework underpins these findings in a quantitative dynamic framework where regions play a Markov game; our analytic formulas for the SCC explain the underlying competing trade-offs.

We follow Parker et al. (2018) and Heyen et al. (2019) in permitting for countermeasures (CM), which limit the ability to free-drive. Heyen et al. (2019) make this point in a linear-quadratic static game with two players and discuss (marginal) free-driving. Manoussi and Xepapadeas (2017) develop the only other dynamic SG game we are aware of, extended to a setting of uncertainty and robust control in Manoussi et al. (2018). In this model, two regions control a joint temperature balancing quadratic costs of geoengineering and global warming against linear-quadratic emission benefits. As in Moreno-Cruz (2015), their numeric steady state solutions find that the asymmetry of countries crucially influences emission and temperature levels.[4] We incorporate these inquiries into a full-blown quantitative integrated assessment model with general (and calibrated) regional production functions relying on fossil fuels and (two different) climate zones that follow state of the art temperature dynamics and SG cooling estimates, and we derive analytic solutions for both sulfur deployment and the SCC.

SG introduces several new uncertainties to IAMs, including uncertainties governing SG damages and the radiative forcing response to stratospheric sulfur injections (Heutel et al. 2018). Goes et al. (2011) analyze the robustness of SG strategies under a wide set of scenarios. Heutel et al. (2016) investigate the effectiveness of SG in dealing with tipping points. Emmerling and Tavoni (2018a) analyze the impact of an uncertain future implementation of SG on present emissions in a two period model of climate change and Kelly et al. (2021) show that sulfur-based SG slows down the learning of the climate sensitivity in a stochastic extension of the DICE model.

Our contribution connects the SG literature above to the recently emerging liter-

[4]Even in the authors' asymmetric equilbrium, they find that both regions contribute to geoengineering. In our setting, such an equilibrium ("climate match") is only one of three types of equilibria and only occurs when regions are sufficiently similar.

ature on analytic integrated assessment models of climate change that derive closed-form solutions for the optimal carbon tax to derive a better understanding of its drivers (Golosov et al. 2014, Gerlagh and Liski 2018, Traeger 2022). Our study is among the first to analyze the strategic interaction of regions within a dynamic integrated assessment model.[5]

Starting with a global social planner, the next section integrates sulfur-based SG into the analytic climate economy (ACE) by Traeger (2022) and derives the optimal sulfur deployment strategy and its impact on the optimal carbon tax. Section 3 splits the global model into two non-cooperatively acting regions and a passive rest of the world. We derive the optimal regional sulfur deployment strategies, characterize a set of Markov perfect equilibria, and characterize the corresponding carbon taxes. Section 4 presents a quantitative simulation of the regional model. Appendix A discusses extensions of the model.

2 Global model

This section introduces SG into the analytic climate economy model ACE (Traeger 2022). First, we summarize a slightly simplified version of the ACE model (for details see Appendix B). Then, we introduce SG and calibrate the forcing effect of sulfur to scientific data. Finally, we discuss the optimal cooling strategy of the social planner and the difference that SG makes for the optimal carbon tax.

2.1 Economic production, climate, and damages

Final output is a function of capital, labor, fossil energy, renewable energy, and the technology levels in different sectors. We write gross world output as

$$Y_t = \mathcal{F}(\boldsymbol{A}_t, \boldsymbol{K}_t, \boldsymbol{N}_t, \boldsymbol{E}_t) \tag{1}$$

where the vector $\boldsymbol{A}_t$ characterizes the exogenously evolving technology levels, the vectors $\boldsymbol{K}_t$ and $\boldsymbol{N}_t$ optimally distributed capital and labor across sectors, and $\boldsymbol{E}_t$ a vector of energy inputs. Our only assumption on the production function is homogeneity of degree $\kappa \in (0, 1)$ in capital, a setting that includes the Cobb-Douglas final production with a CES energy sector of Golosov et al. (2014) as well as the DICE setting of Nordhaus and Sztorc (2013).

[5]Hassler and Krusell (2012) develop a dynamic stochastic general-equilibrium model with multiple regions that can be solved in closed-form, suggesting that only taxes on oil producers can mitigate climate change and taxes on oil consumers have no effect. Hambel et al. (2018) introduce international trade in an analytic climate-economy and show that the regional SCC increases in trade volume.

A subset of the energy inputs $E_{1,t}, ..., E_{j,t}$ are fossil fuels and cause CO_2 emissions. We measures these energy inputs in terms of their CO_2 content so that total emissions are $E_t^{tot} = \sum_{i=1}^{j} E_{i,t} + E_t^{\text{exo}}$, where E_t^{exo} denotes exogenous emissions including those from land use change and forestry. The model includes (endogenous) Hotelling rents for scarce fossil fuels. CO_2 emissions accumulate in the atmosphere and other carbon sinks, following a classical carbon cycle model.[6] It is convenient to measure the resulting atmospheric CO_2 concentration $M_{1,t}$ relative to the preindustrial concentration as $m_t = \frac{M_{1,t}}{M_{\text{pre}}}$. Atmospheric CO_2 causes a greenhouse effect that increases atmospheric temperatures, which we model using Traeger's (2022) non-linear atmosphere-ocean temperature dynamic system.[7] The resulting global atmospheric temperature $T_{1,t}$ measures the increase over 1900 in degree Celsius.

Temperature increase, carbon concentration, and sulfur S_t cause (net) damages $D(T_{1,t}, S_t, m_t)$ that we measure as a fraction of output. These damages are composed of three contributions

$$D(T_{1,t}, S_t, m_t) = 1 - \exp\left[-D_T(T_{1,t}) - D_G(S_t) - D_m(m_t)\right]. \tag{2}$$

Each damage contribution specifies a particular damage contribution per unit of output. We take a convex temperature-based damage function $D_T(T_{1,t})$ from Traeger (2022)[8] and assume SG damages of the form

$$D_G(S_t) = d\, S_t, \tag{3}$$

making d the semi-elasticity of damages from stratospheric sulfur injections (the percentage loss of output resulting from an additional ton of sulfur injections). Damages from SG include changes in the precipitation patterns, a reduction in the upper ozone layer, acid precipitation and sulfur deposition (Crutzen 2006, Heckendorn et al. 2009, Keith and MacMartin 2015). The parameter d includes operational costs. The net costs of an increase of atmospheric carbon above preindustrial levels $(m_t - 1)$ are

$$D_m(m_t) = a\,(m_t - 1), \tag{4}$$

[6] Our quantitative model employs the DICE 2013 carbon cycle, which is slightly better calibrated than the DICE 2016 carbon cycle (Dietz et al. 2021). Swapping the classical carbon cycle against an impulse response model of, e.g., Joos et al. (2013) is straight-forward and changes none of the theoretical results presented here, nor does it make a significant quantitative difference, see Traeger (2022) for details.

[7] Traeger's (2022) nonlinear temperature model is calibrated to the MAGICC 6.0 model and eliminates the exaggerated warming delay of, e.g., DICE's linear atmosphere-ocean diffusion model. It delivers a close match to middle-of the road scientific models and, e.g., Dietz et al. (2021) best match to the model comparison study CMIP5.

[8] It is parametrized by the production semi-elasticity to an exponential temperature increase ξ_0, see Appendix B.1

where a is the semi-elasticity of production with respect to changes in the carbon dioxide concentration. Costs include ocean-acidification and benefits include the fertilizer effect that increases plant production and crop yields.

Our theoretic analysis assumes full depreciation of capital over the course of a decade, the model's time step, and the aggregate capital stock evolves as

$$K_{t+1} = Y_t \left[1 - D_t \left(T_{1,t}, S_t, m_t\right)\right] - C_t. \tag{5}$$

The quantitative application in Section 4 introduces capital persistence following Traeger (2022).

2.2 Geoengineering

Radiative Forcing. Volcanic eruptions taught us that small sulfur particles (aerosols) injected into the stratosphere reflect sunlight back into space cooling our planet. The temperature on our planet results from a balance of incoming and outgoing radiation, sunlight coming in and infrared radiation leaving. Greenhouse gases trap some of the outgoing radiation and SG reduces the incoming radiation. The net effect is summarized by the resulting (anthropogenic) radiative forcing F_t, which is measured in Watts per square meter (W/m^2). We can think of this radiative forcing as the additional anthropogenic heating. More heating eventually results in higher temperatures. The medium-run equilibrium temperature is approximately proportional to the radiative forcing (see Appendix B.4 for details). Physics teaches us that radiative forcing increases logarithmically in atmospheric CO_2 and falls as a consequence of SG measures $G_t(S_t)$ that inject sulfur aerosols S_t into the stratosphere

$$F_t^{exact} = \frac{\eta}{\log 2} \log(m_t) - G_t(S_t) = \frac{\eta}{\log 2} \log\left(m_t \exp\left(-\frac{\log 2}{\eta} G_t(S_t)\right)\right).$$

It is common to express radiative forcing in CO_2 equivalents, which corresponds to the argument of the logarithm on the right side of the equation. In terms of CO_2 equivalents, the forcing from sulfur and CO_2 are no longer independent.

Cooling. At high injection rates, sulfur particles lump together decreasing their cooling efficiency. As a result, scientists expect an asymptotic limit for the cooling we can achieve using stratospheric aerosol injections (a maximal possible cooling, Lawrence et al. (2018)). The uncertainty governing the forcing efficiency is high and the instantaneous radiative forcing effect of sulfur injections varies strongly across different climate models (Niemeier and Timmreck 2015, Niemeier and Schmidt 2017, Kleinschmitt et al. 2018, Lawrence et al. 2018). Table 1 presents Kleinschmitt et al.'s (2018) modeling results on the effective radiative forcing exerted by sulfur

injections.[9]

Table 1: Effective radiative forcing effect from sulfur injections.

Kleinschmitt et al. (2018)				
2 TgS	5 TgS	10 TgS	20 TgS	50 TgS
-1.11	-1.64	-2.91	-4.34	-5.63

Units. The negative 5.6 W/m^2 (Watts per square meter) in the table is approximately double the cooling power of what we produced so far in terms of anthropogenic warming (IPCC 2021). 1 TgS/yr (Tera grams sulfur per year) are 1 Million tons of sulfur annually deployed into the stratosphere. Each TgS/yr corresponds to approximately 25 Boeing 747 loads deployed daily for a year. We note that a Boeing cannot make it into the stratosphere, but it can fuel fighter jets that deploy the sulfur in the necessary altitude. Other options to deploy the sulfur include stratospheric balloons (Dykema et al. 2014). The annual injection rates discussed are rather low compared to already existing anthropogenic and natural (pollution) flows of about 136 TgS/yr (Kravitz et al. 2009).

Calibration. We calibrate our model to the recent study by Kleinschmitt et al. (2018). For this purpose, we develop a new functional form with several degrees of freedom that will permit an analytic solution of the dynamic programming problem and calibrates well to the exact radiative forcing equation. This approximate radiative forcing is a function of atmospheric carbon m_t and the annual sulfur injections S_t

$$F_t(m_t, S_t) = \frac{\eta}{\log(2)} \log \left(\underbrace{f_0 + f_1\, m_t + \left(f_2 - f_3 \left(\frac{m_t}{S_t} \right)^n \right) S_t}_{\equiv\, F_t^{CO2}} \right). \tag{6}$$

The expression F_t^{CO2} characterizes the joint radiative forcing of CO_2 and sulfur in CO_2 equivalents. In the absence of SG, only the second term in F_t^{CO2} would contribute. The inner bracket relying on the parameters f_2 and f_3 reduces the forcing in response to sulfur injections S_t. The main contribution derives from the term $f_3 \left(\frac{m_t}{S_t} \right)^n$, multiplying the stratospheric sulfur. Sulfur forcing is more efficient relative to CO_2 the larger the atmospheric carbon concentration and the lower the

[9]The effective radiative forcing, which also includes rapid adjustments such as changes in atmospheric temperature, is larger than the instantaneous radiative forcing effect (Boucher et al. 2017). The publication only cites the instantaneous radiative forcing impact of sulfur and we obtained the effective radiative forcing effect in Table 1 from the authors in personal correspondence. The literature also proposed alternative aerosols like alumina and diamond particles (Weisenstein et al. 2015, Dykema et al. 2016), calcite or limestone (Keith et al. 2016). Given the lack of a natural experiment with such aerosols, our knowledge about the resulting forcing effect is even more limited.

sulfur concentration. For high levels of sulfur, particles lump together reducing their cooling efficiency. The higher the CO_2 concentration, the lower the warming implied by the marginal ton of CO_2 and the higher the relative forcing reduction of sulfur, which we measure in CO_2 equivalents. We summarize both of these nonlinearities in the joint term whose level effect is captured by f_3 and whose nonlinearity is captured by $n > 0$.

We fit the function to Kleinschmitt et al.'s (2018) data from Table 1.[10] Our fit combines Kleinschmitt et al.'s (2018) forcing data for sulfur injections with the well-known forcing from atmospheric carbon dioxide over the interval $m_t \in [1.5, 3]$, i.e., up to a tripling of preindustrial carbon dioxide concentrations. Our fit minimizes the squared differences of 80 data points. Figure 9 in Appendix B.2 graphs the resulting radiative forcing function (6) and Figure 10 demonstrates the quality of the fit. We list the resulting parameters in Table 2. We assume that total radiative forcing

Table 2: Estimated forcing parameters

f_0	f_1	f_2	f_3	n
0.254	1.16	0.014	0.46	0.69

remains positive (above preindustrial levels) and impose the following constraints based on our empirical fit

Assumption 1. *The (fit-)parameters f_i, $i \in \{0, ..., 3\}$, are positive and $0 < n < 1$. Radiative forcing remains above the preindustrial level, $F_t > 0$, and sulfur injections are between 2 TgS $\leq S_t \leq$ 50 TgS.*

Operational costs and damages of geoengineering. Table 3 shows recent cost estimates of stratospheric sulfur injections based on newly designed airplanes. These studies estimate the costs for either a given reduction in radiative forcing (W/m^2) or a given quantity of sulfur injected into the stratosphere. Stars in Table 3 denote the values in the study. The values in Table 3 suggest average operational

Table 3: Annual operational costs of stratospheric sulfur injections

Authors	**Estimate**					
Klepper and Rickels (2012)	billion USD	2-18	for	-1* W/m^2	$\approx$ 2	TgS
Moriyama et al. (2017)	billion USD	10	for	-2* W/m^2	$\approx$ 7	TgS
McClellan et al. (2012)	billion USD	1-3	for		1*	TgS
"	billion USD	2-8	for		5*	TgS
Smith and Wagner (2018)	billion USD	1.5	for		1*	TgS

[10]The new functional form can also be calibrated well to data from other studies (see Appendix B.2).

costs around $2.3 billion per TgS or a cost contribution to d measured as fraction of global output of 0.0017% per TgS.[11]

Damages and potential worries accompanying sulfur-based SG include changes in precipitation, potential impact on the ozone layer, changes in the temperature profile (day-night and high-vs-low lattitude), acid rain, and resulting impact on crop yields and ecosystems. Quantitative assessments of these impacts suffer from insufficient observation. Some authors, for example Moreno-Cruz and Keith (2013), therefore analyze optimal policy as a function of the damage parameter. Others make explicit assumptions, acknowledging a limited or hardly existing empirical basis. We show several of those estimates in Table 4, where we translate the corresponding estimates into our damage parameter d specifying the fraction of global output lost per TgS. For example, Emmerling and Tavoni's (2018b) estimate translates to damages of approximately $135 billion per TgS or $135 per kg of sulfur and implies a damage semi-elasticity of approximately $d = 0.1\%$ per TgS. We discuss details of these estimates and their translation into our model in Appendix B.3.1. We note that the operational costs cited in Table 3 are substantially smaller than the damages estimates. For the global social planner, the semi-elasticity d incorporates both costs and damages.

Table 4: Damages from SG

Authors	**Best guess taken**	d (per TgS)
Emmerling and Tavoni (2018b)	Consumption loss of 3% compensating each 3.5W/m^2 of forcing	$\approx 0.1\%$
Goes et al. (2011)	GDP loss between 0 and 5% per forcing equivalent to a doubling CO_2 forcing	0-0.17%
Heutel et al. (2018)	GDP loss of 3% for resetting forcing to the preindustrial level	$\approx 0.21\%$

2.3 Global planner solution

In the present section, a global social planner maximizes the infinite stream of utility from consumption

$$\max_{C_t, E_t, S_t} \sum_{t=0}^{\infty} \beta^t \log(C_t) \tag{7}$$

[11] At an annual world output of 135 trillion USD (purchasing power parity 2019), a deployment cost of 1 billion USD translates into a fractional output cost of $d = 7.4 \times 10^{-4}$ %. In compiling the average, we give equal weight to authors (putting only half the weight on each of McClellan et al.'s (2012) estimates).

subject to equations (1)-(6), Assumption 1, and some additional model details summarized in Appendix B. The parameter β denotes the utility discount factor (pure time preference). Appendix C solves the intertemporal optimization problem and derives the globally optimal level of sulfur deployment.

Proposition 1. *The optimal level of sulfur deployment is*

$$S_t^* = \left(\frac{(1-n)\,\gamma\, f_3}{d + \gamma\, f_2} \right)^{\frac{1}{n}} m_t \tag{8}$$

with climate change impact $\gamma = \beta\, \xi_0\, \tilde{\sigma}$*, where* $\tilde{\sigma}$ *is a climate system specific multiplier defined in Appendix B.4.*[12]

The optimal deployment of sulfur increases linearly in the atmospheric carbon concentration. We refer to the proportionality factor

$$z \equiv \left(\frac{(1-n)\,\gamma\, f_3}{d + \gamma\, f_2} \right)^{\frac{1}{n}} \tag{9}$$

as the *SG propensity*. It reappears in the strategic setting and characterizes the drivers and moderators of the cooling effort (given atmospheric carbon concentration). This cooling propensity *increases* in the discount factor β, the temperature damage coefficient ξ_0, and the sulfur efficiency f_3. Sulfur deployment *decreases* in SG damages d and the non-linear efficiency loss of sulfur cooling n.[13]

Using the fit parameters from Table 2 and the parameter values from the baseline calibration of the ACE model (see Appendix C.2) we find the optimal sulfur deployment level

$$S_t^* = \left(\frac{1.65}{16\% + 10^3 d} \right)^{1.45} m_t, \tag{10}$$

as a function of the damage (semi-)elasticity of sulfur d, and the atmospheric carbon concentration m_t (expressed relative to preindustrial levels). The calibrated formula shows that the optimal sulfur deployment is extremely sensitive to damages from SG. The left graph in Figure 1 shows the optimal level of annual sulfur deployment as a function of the damage parameter d for different carbon concentrations. The

[12] The closed-form expression for $\tilde{\sigma}$ stated in the Appendix can be interpreted as the discounted infinite sum of the current forcing impact on future atmospheric temperatures given atmosphere-ocean interactions, see Traeger (2022).

[13] $\frac{d}{dn} \log \left(\frac{(1-n)\,\gamma\, f_3}{d+\gamma\, f_2} \right)^{\frac{1}{n}} = -\frac{\log\left(\frac{(1-n)\,\gamma\, f_3}{d+\gamma\, f_2} \right)}{n^2} - \frac{1}{n(1-n)} < 0$ since Assumption 1 requires $S_t \geq 2$ for $m_t \in [1.5, 3]$, and thus $z^n = \frac{(1-n)\,\gamma\, f_3}{d+\gamma\, f_2} \geq 1.22$.

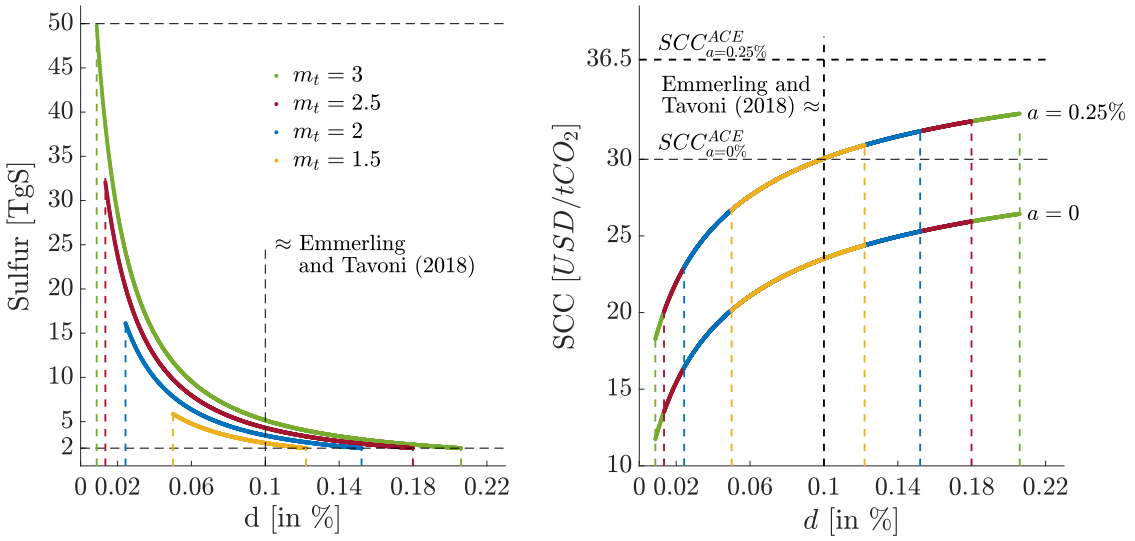

Figure 1: Optimal level of sulfur (left) and the global SCC (right) as a function of SG damages (d in %) for the case without net acidification damages, $a = 0$, and a case with damages of $a = 0.25\%$. The colored dashed lines give the boundaries of our model calibration's validity for the corresponding CO_2 levels (implied by the sulfur deployment constraint and the positive forcing constraint in Assumption 1).

curves are restricted to the invervals satisfying Assumption 1. Inserting S_t^* into equation (6) yields the optimal level of radiative forcing as a function of the damage parameter d and the atmospheric carbon stock (see Figure 12 in Appendix B.2). The higher the damages, the higher the tolerated forcing and, thus, warming levels. Yet, *the tolerated warming is less sensitive to damages than the sulfur injections because a reduction in sulfur levels increases the cooling efficiency per ton.*

We now study the optimal carbon tax in the presence of SG. The SCC reflects the long-term damage from releasing a marginal ton of CO_2 into the atmosphere. Proposition 1 shows that an increase of atmospheric carbon increases the level of SG. As a result, SG partially mitigates the damages from releasing a ton of CO_2.

Proposition 2. *The SCC in money-measured consumption equivalents is*

$$SCC_t = \frac{Y_t^{net}}{M_{pre}} \left[\underbrace{a}_{\substack{ocean \\ acid}} + \underbrace{f_1 \gamma}_{greenhouse} \underbrace{- \left(\left(\frac{f_3}{z^n} - f_2 \right) \gamma - d \right) z}_{geoengineering} \right] \tilde{\phi}. \tag{11}$$

As defined above, the SG propensity is $z = \left(\frac{(1-n)\,\gamma\, f_3}{d + \gamma\, f_2} \right)^{\frac{1}{n}}$ *and* $\gamma = \beta\, \xi_0\, \tilde{\sigma}$ *characterizes the climate change impact. The term* $\tilde{\Phi}$ *is a carbon cycle specific multiplier defined*

in Appendix B.4.[14]

The fraction $\frac{Y_t^{net}}{M_{pre}}$ sets the scale and units of the SCC. The square brackets characterize net damages and the term $\tilde{\phi}$ amplifies the SCC as a result of the long life-time of atmospheric CO_2 (carbon cycle). Earlier analytic integrated assessment models like ACE only contain a term corresponding to our $f_1\gamma$ reflecting the cost resulting from a temperature increase in the absence of climate engineering. First, formula (11) adds direct net damages from atmospheric CO_2 caused by ocean-acidification net of the land-based fertilization effect, represented by the term a. Second, it introduces the term in round brackets, which reduces the SCC as a result of SG (the bracket is always positive). The reduction of the SCC increases in the overall effectiveness of SG f_3. The reduction of the SCC increases as well in the overall SG propensity z, but at a falling rate; the SG propensity in the denominator of $\frac{f_3}{z^n}$ reflects that the effectiveness of sulfur-based cooling decreases in the level of SG. Finally, the damages from SG d augment the SCC, and more so the higher the sulfur deployment per unit of carbon (SG propensity z).

The SCC's composition in Proposition (2) explains how the SCC responds to the SG propensity. Yet, the SG propensity is itself a function of damages and the sulfur's cooling effectiveness. Breaking the SG propensity up into its various contributions, we find

$$\left(\left(\frac{f_3}{z^n}-f_2\right)\gamma-d\right)z=n\frac{(1-n)^{\frac{1-n}{n}}\ (\gamma\, f_3)^{\frac{1}{n}}}{(d+\gamma\, f_2)^{\frac{1-n}{n}}}. \tag{12}$$

The SCC's qualitative dependence on damages d and cooling effectiveness f_3 remains as above. We observe that *the SCC-reduction is less responsive to SG damages than is the SG propensity* z (comparing equations 9 and 12 the power $\frac{1-n}{n} < \frac{1}{n}$ as $n < 1$). The following intuition governs this damage dependence of the SCC. Lower damages increases the SG propensity. As a result, each unit of carbon emissions is accompanied by more cooling, reducing the resulting damage and the SCC. However, more cooling also releases more sulfur into the atmosphere, partially offsetting the cooling benefit, which explains why the SCC is less responsive to changes in damages than the SG propensity itself.

Quantifying the SCC, we follow ACE using a timestep of 10 years and the parameter specification summarized in Table 10 in Appendix C.2,[15] and we use our

[14]The closed-form expression for $\tilde{\Phi}$ stated in the Appendix can be interpreted as a discounted sum over that part of the marginal ton of CO_2 released that remains in (or returns to) the atmosphere over the course of time, see Traeger (2022).

[15]We re-calibrated ACE's temperature dynamics to use two rather than three temperature layers. This common simplification hardly affects the model's ability to replicate the temperature dynamics of scientific climate models and substantially eases the presentation of the regional model with and without heat transfers in Appendix B. The calibration follows the same method as in Traeger

radiative forcing estimates from Table 2. The right panel in Figure 1 shows the SCC in 2019-USD as a function of the SG damages d for a given ocean acidification net damage parameter $a = 0$ and for $a = 0.25\%$[16].

3 Non-cooperative regions

After discussing the benchmark, we now split the world into non-cooperatively interacting regions. We focus on two regions A and B that consider actively engaging in geoengineering. We also permit these regions to engage in countermeasures (CM), an option we can exclude by imposing prohibitively high costs of CM. The rest of the world interacts with regions A and B only through CO_2 emissions and we discuss those regions in more detail in our quantitative Section 4 and in Appendix A.1. The regions play a dynamic Markov game and we identify a natural set of strategies that result in one of three qualitatively different subgame perfect Nash equilibria, conditional on regional characteristics, sulfur efficiency, and SG damages.

3.1 Regional economies and climate dynamics

This section explains the required changes when splitting the world in several economic regions and two different climate zones.

Regional economies and emissions. The regional economies follow equations (1-5), where functional forms and parameters are idiosyncratic to the regions. CO_2 mixes globally and the CO_2 dynamics are described in Appendix B.4. Total CO_2 emissions now result from region A, region B, and the rest of the world.

Sulfur spillovers. Sulfur injected into the stratosphere spreads widely. It travels mostly within latitude and towards the poles (Lawrence et al. 2018). It is impossible to contain stratospheric SG to a region.[17] If region A engages in SG with regional target level $S_t^A > 0$, we denote by $\alpha_A S_t^A$ the spillovers from region A into region B. Accordingly, $\alpha_B S_t^B$ denote the spillovers from region B to region A in case region B engages in SG.[18]

(2022).

[16] Colt and Knapp (2016) estimate the loss of ecosystem services from ocean acidification between as 0.09% and 0.28% of current global GDP (flow damages). Brander et al. (2012) estimate the net present value of economic damages from the loss of (only) coral reefs between 0.14% and 0.18% of current GDP. We base our damage estimate on Colt and Knapp's (2016) lower bound of 0.1% of global GDP. This estimate translates into a damage parameter of $a = \frac{0.1\%}{(1.4-1)} = 0.25\%$.

[17] Recent studies suggest that it might be possible to optimize the geographic distribution of the cooling by varying the altitude, latitude and season of injections (Kravitz et al. 2017, MacMartin et al. 2017, Dai et al. 2018, Jones et al. 2018, Visioni et al. 2019). We stick with an exogenous parametrization of spillovers across regions.

[18] On average, the level targeted by region A, S_t^A, remains in climate zone A's stratosphere and the level $\alpha_A S_t^A$ remains in region B's stratosphere. In this notation, the forcing of a unit of S_t in

The spillover parameters α is low if one region is located on the northern and the other region on the southern hemisphere (e.g. the US and Brazil). It will be close to unity if both regions are located on the same hemisphere and at similar latitude (e.g. Europe and North America). It will be asymmetric if one region lies North of the other on the same hemisphere (e.g. Canada would get almost perfect spillover from the US, but the US much less spillover from Canada). We require $0 \leq \alpha_A, \alpha_B < 1$. In line with Assumption 1, we assume 2 TgS $\leq S_t^A + \alpha_B\, S_t^B \leq 50$ TgS, and 2 TgS $\leq S_t^B + \alpha_A\, S_t^A \leq 50$ TgS.

Regional climate dynamics. We split the globe into two climate zones. For simplicity, we assume that the rest of the world is part of region B's climate zone. Each climate zone experiences region-specific radiative forcing, $F_t^A(m_t, S_t^A + \alpha_B\, S_t^B)$ and $F_t^B(m_t, S_t^B + \alpha_A\, S_t^A)$ respectively. See Appendix B.4 for details.

Regional damages and countermeasures. SG creates damages and operational costs. For region A, d_{AA} summarizes the marginal costs of the region's own action, and d_{BA} the marginal costs imposed on region A by region B. Damages including operational cots as a fraction of output in region A are

$$D^A(T_{1,t}^A, S_t, m_t) = 1 - \exp\left[D_T^A(T_{1,t}^A) - (d_{AA} S_t^A + d_{BA}\, \alpha_B\, S_t^B) - a^A\,(m_t - 1)\right]. \quad (13)$$

We represent region A's engagement in CM by $S_t^A < 0$. CM can represent deployment of a chemical counter-geoengineering agent that partially offsets the radiative forcing effect of the sulfur emitted by the other region (see e.g. Parker et al. 2018). Our preferred interpretation, however, is that region A exerts political pressure on region B to reduce the sulfur spillovers. Giving in to costly political pressure, region B changes vertical injection profile, geographical distribution, and quantity of the sulfur injections to reduce the impact on region A. From region B's perspective, this response leads to higher costs for similar cooling. The costs usually leading to a forcing reduction S_t^B now lead to a forcing reduction of only $S_t^B + \alpha_A S_t^A < S_t^B$ in region B. By construction, the reduction in forcing is even stronger in region A where forcing is $\alpha_B S_t^B + S_t^A$.[19]

the regional model corresponds to that of the globally calibrated model only in the case of high spillovers (α's close to unity). For more moderate spillovers the regional efficiency will be lower, reflecting that indeed regionally deployed sulfur will be (globally) less efficient because it will not be injected at the right locations (equator). The cost coefficients do not translate one to one from the global into the regional model but have to be recalibrated.

[19] In the political pressure interpretation, the α-parameters for cooling and CM are not tied to the same physical diffusion properties of stratospheric sulfur. Thus, one could introduce a second set of α parameters for CM. Both sets of α-parameters would likely lie below but not far below unity. Thus, we simplify the representation using a single set of parameters.

Denoting operational costs by ϵ, we define region A's total costs by

$$d_{AA}(S_t^A) = \begin{cases} d_{AA}^g + \epsilon_A^g & \text{for} \quad S_t^A > 0 \qquad \text{(A geoengineers)} \\ d_{AA}^c - \epsilon_A^c & \text{for} \quad S_t^A < 0 \qquad \text{(A counters)} \\ 0 & \text{for} \quad S_t^A = 0 \end{cases} \tag{14}$$

where d_{AA}^g is the damage from sulfur-based cooling and $\epsilon_A^g > 0$ is the cost of injecting the sulfur into the stratosphere. The parameter d_{AA}^c characterizes the damage *reduction* (noting that $S_t^A < 0$) from employing CM, and $\epsilon_A^c \geq 0$ is the cost of CM. For the damages imposed by region B onto region A we distinguish whether region B engages in sulfur-based cooling or CM

$$d_{BA}(S_t^B) = \begin{cases} d_{BA}^g & \text{for} \quad \alpha_B\, S_t^B > 0 \qquad \text{(B geoengineers)} \\ d_{BA}^c & \text{for} \quad \alpha_B\, S_t^B < 0 \qquad \text{(B counters)} \\ 0 & \text{for} \quad \alpha_B\, S_t^B = 0 \end{cases} \tag{15}$$

where d_{BA}^c characterizes the damage "offset" in region A as a result of region B's CM impeding the desired deployment. The parameters d_{BA}^c will generally be strictly lower than d_{AA}^g; CM can at most partially offset the damages. The parameter can also be negative in case the CM imply a more damaging sulfur deployment profile. We restrict the damage parameters in the active regions as follows.

Assumption 2. *The damage relief from CM is smaller than the damage caused by SG:* $d_{ij}^c \leq d_{il}^g$ *for* $i, j, l \in \{A, B\}$.

We note that Assumption 2 combines necessary assumptions to guarantee a unique optimal response ($d_{AA}^c \leq d_{AA}^g$ and $d_{BB}^c \leq d_{BB}^g$), and assumptions that we only impose because they seem economically sensible. By symmetry we obtain the same damage definitions for region B (see Appendix B.3.2).

For ease of presentation, the following subsections turn off the direct heat transfer between the regions.

Assumption 3. *There is no direct heat exchange between climate zones.*

See Appendix B.4 for a formal statement of this assumption. As a result of the assumption, the regional climates interact only through the spillover of the cooling agents. The assumption simplifies the functional expressions without changing the qualitative results. Our quantitative Section 4 drops Assumption 3 and Appendix A.2 presents the corresponding generalization of the theoretical results.

3.2 Deployment strategies

The present section discusses a natural set of regional strategies resulting in a subgame perfect Nash equilibrium. Appendix D solves the corresponding dynamic Markov game. In every period, regions control their sulfur (and CO_2) emissions optimally, anticipating the future reaction of the other region to its own actions; such a solution is sometimes referred to as a feedback equilibrium (as opposed to an open-loop equilibrium requiring commitment devices or implying continuous revision of plans). Our regions condition sulfur deployment on carbon concentration in direct analogy to the unique optimal strategy of our global social planner. As in this earlier setting, a linear deployment strategy also supports a subgame perfect Nash equilibrium in the dynamic game.[20] We consider such a linear response of SG to carbon concentrations a reasonable assumption governing possible actions in the real world.

Proposition 3 (Strategies)**.** *The following strategies characterize a Markov perfect Nash equilibrium of the dynamic game. If (i) $S_t^B = 0$ region A chooses $S_t^A(m_t) = z_A^g m_t$ and if (ii) $S_t^B \neq 0$ region A chooses*

$$S_t^A(m_t) = \frac{m_t}{1-\alpha_A\,\alpha_B}\left(z_A^g - \alpha_B\, z_B\right) \quad \textit{if} \quad z_A^g > \alpha_B z_B \quad \textit{(SG)} \tag{16}$$

$$S_t^A(m_t) = \frac{m_t}{1-\alpha_A\,\alpha_B}\left(z_A^c - \alpha_B\, z_B^g\right) \quad \textit{if} \quad z_A^c < \alpha_B z_B^g \quad \textit{(CM)}$$

$$S_t^A = 0 \quad \textit{otherwise} \qquad \textit{(inactive)}$$

with SG propensity and CM reluctance

$$z_A^g = \left(\frac{(1-n)\,f_3\,\gamma_A}{f_2\,\gamma_A + (d_{AA}^g + \epsilon_A^g)}\right)^{\frac{1}{n}}, \quad z_A^c = \left(\frac{(1-n)\,f_3\,\gamma_A}{f_2\,\gamma_A + (d_{AA}^c - \epsilon_A^c)}\right)^{\frac{1}{n}}$$

and climate change impact $\gamma_A = \beta^A\,\xi_0^A\,\tilde{\sigma}^A$. In equation (16) $z_B \in \{z_B^g, z_B^c\}$ depending on whether $S_t^B \gtrless 0$. By Assumption 2, $z_A^g < z_A^c$ and the cases above are mutually exclusive. Swapping region indices characterizes region B's strategies.

[20] Also the static linear-quadratic model by Heyen et al. (2019) and the linear quadratic dynamic game by Manoussi and Xepapadeas (2017) give rise to linear deployment strategies (our solution does not rely on a linear-quadratic model). Despite permitting for some asymmetries, Manoussi and Xepapadeas (2017) only find a "climate match" solution where both regions are cooling jointly. Heyen et al.'s (2019) stylized static game also identifies the three different types of equilibria that we classify in Section 3.3, yet, in their setting the availability of CM eliminates free-driving whereas in our case both types of equilbria can co-exist over the heterogeneity domain of the regions (see e.g. Figure 2).

One active region. In the case where one of the regions remains inactive (say $S_t^B = 0$), the other region's optimal cooling effort is structurally equivalent to the social planner setting in Proposition 2; sulfur deployment increases proportional to the atmospheric carbon concentration and to the SG propensity z_A^g. As in the social planner setting, this cooling propensity *increases* in the climate change impact γ_A and the sulfur efficiency f_3. Sulfur deployment *decreases* in SG damages and the non-linear efficiency loss n of sulfur cooling. In contrast to the social planner, the active region only accounts for its own climate impact γ_A and for its own damages d_{AA}^g and costs ϵ_A^g from SG (damages and costs were combined into a single term d in the social planner's problem).

Both regions cooling. In the case where both regions engage in cooling, the strategic interaction introduces two novel motives for sulfur deployment. First, a region acknowledges the other region's contribution and reduces its own sulfur deployment accordingly (free-riding). In equation (16), this direct response subtracts the spillover-weighted SG propensity $\alpha_B z_B^g$ of the other region from region A's original SG propensity z_A^g. Second, each region *anticipates* that the other region will respond to its own response. Because of the linear response functions, this higher order reaction leads to the multiplier $\frac{1}{1-\alpha_A\alpha_B}$; it scales up region A's action as a result of its anticipation that region B reduces its sulfur deployment (free-rides) in response to A's action. This higher order response counteracts the initial free-riding incentive; it is always smaller than the initial response and there is only a partial crowding out.[21] *In summary, with both regions cooling, each region's sulfur deployment (i) decreases strategically as a result of joint action (or free-riding), but this decrease is (ii) partially offset by a region's anticipation of the other region's cooling reduction (or free-riding).*

Countermeasures and climate-clash. In the case that regional interests clash, one region, say region B, is cooling. Region B's "excessive" SG causes spillovers that drive region A to engage in CM ($S_t^A < 0$). The propensity z_A^c characterizes the *reluctance* to engage in the CM. A higher reluctance z_A^c delays the onset of CM and, once region A counters, reduces the strength of the CM.[22] The CM engagement increases in the other region's (spillover-weighted) SG propensity. The anticipation

[21] The finding is less obvious at second thought. The initial free-riding incentive grows in the other region's SG propensity, whereas a region's anticipation of the other regions free-riding is (analogously) driven by its own SG propensity. In principle, a region's anticipation effect could therefore dominate the initial free-riding effect as the following calculation shows: $\frac{z_A^g - \alpha_B z_B^g}{1-\alpha_A \alpha_B} m_t > z_A^g m_t \Leftrightarrow z_A^g - \alpha_B z_B^g > z_A^g - \alpha_A\alpha_B z_A^g \Leftrightarrow z_B^g < \alpha_A z_A^g$. However, if region A's spillover weighted SG propensity indeed dominates region B's own SG propensity ($\alpha_A z_A^g > z_B^g$), the two regions would be in a different equilibrium where only region A is cooling and region B free-rides (or engages in CM) as we will establish in Proposition 4.

[22] The sign in front of the deployment costs ϵ_A switches because reducing S_t^A now imposes deployment costs of the CM (rather than reducing sulfur deployment costs). *This discrete jump from reducing sulfur deployment costs to incurring CM costs ensures a non-trivial (z_B^g-)interval where A simply remains inactive.*

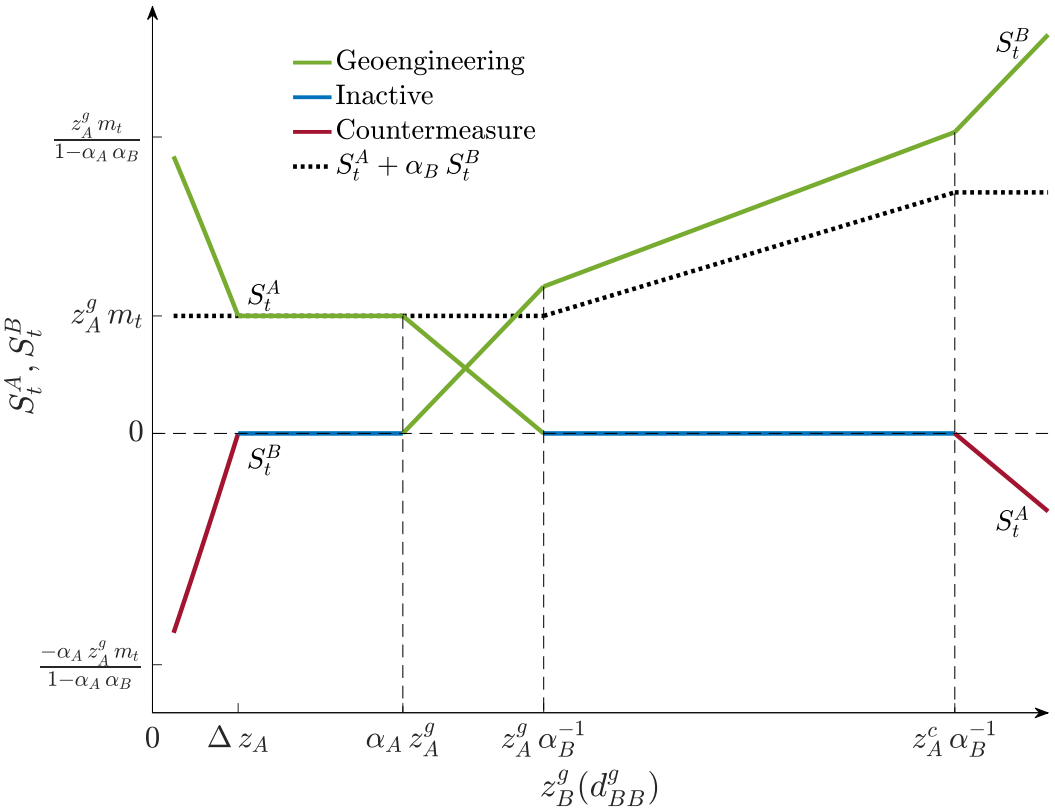

Figure 2: Regional strategies as a function of B's SG propensity z_B^g by varying d_{BB}^g. Each region either deploys sulfur (green), a CM (red), or stays inactive (blue). The dotted black line depicts the target level of sulfur in A ($S_t^A + \alpha_B\, S_t^B$). Note that $\Delta\, z_A \equiv \alpha_A\, z_A^g - (z_B^c - z_B^g)$.

of each other's responses imply once again the multiplier $\frac{1}{1-\alpha_A\alpha_B}$. Here, *the anticipation of each other's response intensifies a clash. Each region tries to push the climate harder into the opposite direction. Despite the friction, the actions bring both countries closer to their targets.*[23]

Figure 2 shows the deployment of sulfur or the engagement in CM for regions A and B as a function of region B's SG propensity z_B^g by varying B's SG damages d_{BB}^g.[24] The graphs keeps region A's parametrization fixed. In the middle of the graph, both regions deploy sulfur. Here, regional cooling propensities are higher

[23] In the case of a physical CM, the incomplete spillovers ($\alpha_A\alpha_B < 1$) imply that simultaneous warming and cooling in the two regions brings both regions closer to their desired temperatures. In the political CM interpretation, a region is pressured into reducing spillovers as well as possible.

[24] We assume $d_{BB}^c = 0.5\, d_{BB}^g$. To get a unique parametrization in terms of propensities, the graph shows z_B^g even if the chosen strategy is z_B^c.

than the spillovers received from the other region ($\alpha_A\, z_A^g < z_B^g < z_A^g\, \alpha_B^{-1}$). Moving to the left, region B's SG propensity falls. Eventually, the spillovers from region A cause sufficient cooling and region B ceases its SG activities ($z_B^g < \alpha_A\, z_A^g$). Moving further left, region B's desire for SG turns into a tolerance for SG. Eventually this tolerance will be exhausted and region B starts to fight back with CM ($z_B^g < \Delta\, z_A \equiv \alpha_A\, z_A^g - (z_B^c - z_B^g) \Leftrightarrow z_B^c < \alpha_A\, z_A^g$). Analogously, the right side of the graph shows how region A becomes inactive if region B's SG propensity is sufficiently high and will eventually resort to CM if region B's SG becomes excessive.

The dashed black line shows the resulting sulfur level in region A. As long as A engages in SG, it keeps sulfur at its target level ($S_t^A + \alpha_B\, S_t^B = z_A^g m_t$). Once region A turns inactive, the sulfur starts exceeding the target level and will eventually be stabilized at a higher level using CM. These levels are determined by the CO_2 concentration and the region's SG propensity and CM reluctance, respectively. Dropping Assumption 3 on the absence of direct heat exchange eliminates the constancy of the target levels and the piece-wise linearity of the graph, but preserves the qualitative features (see Figure 5 in Appendix A.2).[25]

3.3 Characterization of equilibria

We now identify the parameter ranges that give rise to the different types of equilibria and show that they are mutually exclusive and cover the full parameter domain.

Proposition 4 (Equilibria)**.** *The strategies in Proposition 3 give rise to 5 qualitatively different Markov perfect Nash equilibria. They are mutually exclusive and classified based on fundamentals as follows:*

Climate clash	$S_t^A > 0, S_t^B < 0:$	$\alpha_A^{-1} < h$
Free-driver/rider	$S_t^A > 0, S_t^B = 0:$	$h \leq \alpha_A^{-1} \leq H$
Climate match	$S_t^A > 0, S_t^B > 0:$	$\alpha_B < H < \alpha_A^{-1}$
Free-driver/rider	$S_t^A = 0, S_t^B > 0:$	$H \leq \alpha_B \leq \hat{H}$
Climate clash	$S_t^A < 0, S_t^B > 0:$	$\hat{H} < \alpha_B$

[25] Heyen et al. (2019) depict the equilibrium of their static game over a similar "asymmetry domain". In their case, the asymmetry between the two regions is the desired target temperature. Their linear-quadratic setting gives rise to strategies that are piecewise linear and concave over this temperature target asymmetry. In our IAM, countries cannot simply set the temperature, but we note that the above strategies are highly non-linear (and neither concave nor convex) over the relevant parameters characterizing a country's implied "temperature preferences" like climate or SG damages and depend explicitly on climate dynamics, cooling efficiency, and spillover characteristics.

with the ratios of SG propensities (or CM reluctance)

$$h = \frac{z_A^g}{z_B^c} \quad , \quad H = \frac{z_A^g}{z_B^g} \quad , \textit{ and } \quad \hat{H} = \frac{z_A^c}{z_B^g}.$$

We note that $h \leq H \leq \hat{H}$ *and that* $\alpha_B \leq \alpha_A^{-1}$.

The spillover parameters and the ratios of the SG propensities (or CM reluctance) characterize the type of equilibrium that arises.

Symmetric regions. Two fully symmetric regions target the same SG levels, implying a SG propensity ratio $H = 1$ and a climate match equilibrium where both regions contribute to cooling the planet. In the absence of operational costs, this non-cooperative equilibrium would be efficient in cooling the planet, given carbon stocks and emissions.[26]

Asymmetric regions. Departing from the case of symmetry, we assume that region A's SG propensity z_A^g falls, e.g., because perceived damages from SG are larger or because its climate impacts are lower than in region B. As a result, the SG propensity ratio H falls and so is region A's contribution to cooling the planet. Once $H \leq \alpha_B \Leftrightarrow z_A^g \leq \alpha_B z_B^g$ ceases activity; the spillovers from region B provide sufficient cooling. At least initially, region A becomes a free-rider. As long as A unambiguously benefits from region B's actions, we would resist calling this unilateral cooling a "free-driver" equililibrium. However, eventually the SG externality turns negative for region A and the "free-driver" label might be suitable. Once, spillovers from region B dominate region A's reluctance to engage in CM, the regions enter a climate clash ($\alpha_B < \hat{H} \Leftrightarrow \alpha_B z_B^g > z_A^c$).

3.4 Emissions

By reducing global temperatures, SG can reduce the incentive for greenhouse gas mitigation. As discussed in the introduction, a measure of a region's incentive to reduce emissions is the SCC.[27] Traeger (2021) discusses the relation between the SCC and emissions for a variety of production structures including DICE and Golosov et al.'s (2014) model. In general, emissions for region A are

$$E_{i,t}^A = \frac{\sigma_{Y,E_i}^A Y_{A,t}^{net}}{HOT_{i,t}^A + \beta SCC^A} \tag{17}$$

[26] In the absence of operational costs, each region merely balances costs and benefits of SG at the optimal level. As a result of the symmetry, the regions agree on the optimal level. In the case of operational costs, the target level is too low because a region does not incorporate the other regions' spillover benefits. Of course, the CO_2 externality remains unresolved.

[27] Our quantitative Section 4 incorporates that some regions seem to incorporate more than only their own damages in setting prices on carbon emissions. In the present discussion, we can think of it as a slightly increased climate damage factor γ.

where $\sigma^A_{Y,E^A_i} = \frac{\partial \mathcal{F}^A(\cdot)}{\partial E^A_{i,t}} \frac{E^A_{i,t}}{Y^A_t}$ denotes the production elasticity of fossil resource i and $HOT^A_{i,t}$ denotes the scarcity value of the resource (Hotelling rent). In general, both σ^A_{Y,E_i} and $HOT^A_{i,t}$ are endogenous and reflect how strengthening climate measures restructures the economy and changes resource scarcity. Traeger (2021) states conditions under which an increase in the SCC always reduces current emissions (standard case). He also derives a closed-form solution of emissions under the DICE/RICE production structure which we employ in the next section for our simulation.

3.5 Regional social cost of carbon

SG reduces future warming and future damages *per ton of carbon emitted today*. The present section analyzes the corresponding determinants of the SCC and, thus, the incentive to mitigate. In our strategic setting, the SCC depends on the type of equilibrium and other regions' actions.

Proposition 5 (Regional SCC)**.** *(i) If $S^B_t = 0$, the SCC in region A is*

$$SCC^A_t(z^g_A, d_{AA}) \equiv \frac{Y^{net}_{A,t}}{M_{pre}} \Bigg[\underbrace{a^A}_{\substack{ocean \\ acid}} + \underbrace{f_1\,\gamma_A}_{\substack{green- \\ house}} \underbrace{- \left(\left(\frac{f_3}{(z^g_A)^n} - f_2 \right) \gamma_A - (d_{AA}) \right) z^g_A}_{geoengineering} \Bigg] \tilde{\phi}^A$$

where $d_{AA} = d^g_{AA} + \epsilon^g_A$. (ii) If region A itself is inactive ($S^A_t = 0$),

$$SCC^A_t = SCC^A_t(\alpha_B\, z^g_B, d^g_{BA}).$$

(iii) If both regions are active ($S^B_t \neq 0$, $S^A_t \neq 0$),

$$SCC^A_t = SCC^A_t(z_A, d_{AA}) - \frac{Y^{net}_{A,t}}{M_{pre}} \Bigg[\underbrace{- \frac{\alpha_B (z_B - \alpha_A\, z_A)(d_{AA} - d_{BA})}{1 - \alpha_A\, \alpha_B}}_{spillover} \Bigg] \tilde{\phi}^A$$

with $z_A \in \{z^g_A, z^c_A\}$, $z_B \in \{z^g_B, z^c_B\}$, $d_{AA} \in \{d^g_{AA} + \epsilon^g_A, d^c_{AA} - \epsilon^c_A\}$, and $d_{BA} \in \{d^g_{BA}, d^c_{BA}\}$ depending on whether the corresponding region engages in SG (g) or CM (c). Swapping regional indices characterizes region B's SCC.

Before discussing the SCC-formulas in detail, we summarize some consequences in the following corollary.

Corollary 1. *(i) The availability of SG reduces the SCC of a unilaterally acting region. (ii) In all other types of equilibria, the impact of the availability of SG on the SCC is ambiguous; it can increase, decrease or leave the SCC unchanged depending on the heterogeneity of damages, climate impacts, and spillovers. (iii) Therefore, SG can increase or decrease global emissions in all types of equilibria.*

The regional SCC in the absence of SG is given by the first two terms in square brackets (and coincides with setting the SG propensities to zero).

Unilaterally acting region. If region B is inactive ($S_t^B = 0$), region A's SCC has the same structure as in the global model. In contrast to the social planner, A only accounts for its own climate impact γ_A and for its own damages d_{AA}^g and costs ϵ_A^g. As in the global social planner case, *the availability of SG always reduces the active region's SCC.*

Inactive region. If region A is inactive ($S_t^A = 0$), the structure of its SCC resembles that of the unilaterally acting region. However, the SCC is no longer driven by its own SG propensity (z_A) or damages (d_{AA}); rather it is driven by the spillovers from region B, $\alpha_B\, z_B^g$, and the corresponding damages d_{BA}^g. As a result, the inactive region's SCC can increase or decrease depending on the regional differences in climate impact and SG damages.

Both regions active: the spillover term. If both regions are active an additional term enters region A's SCC, which we label the

$$\text{spillover term:} \quad -\frac{\alpha_B(z_B - \alpha_A\, z_A)(d_{AA} - d_{BA})}{1 - \alpha_A\, \alpha_B} = -\alpha_B \frac{S_t^B(m_t)}{m_t}(d_{AA} - d_{BA}). \quad (18)$$

Climate-match. The spillover term characterizes how region B's action changes region A's SCC w.r.t. unilateral action. At the heart of the spillover term are the *excess costs $d_{AA} - d_{BA}$ from region A's own as compared to region B's SG.* They capture the difference between the marginal damages and deployment costs per unit of SG that region A incurs from its own versus B's action. If damages are independent of where SG is deployed, then excess costs equal the positive deployment costs ϵ_A^g, and region A benefits from B's support in cooling the planet. This benefit from B's engagement grows with the spillover weighted sulfur emissions of region B, $\alpha_B \frac{S_t^B(m_t)}{m_t}$, here per unit of atmospheric carbon.

To understand the implications of joint cooling for a region's SCC, we have to understand how cooling support affects the costs resulting from the marginal ton of CO_2 emissions. This ton increases the atmospheric carbon concentration and, thereby, triggers additional and enduring SG in both regions. The spillover term represents the benefits from the other region's cooling support, which reduces the costs of (partially) offsetting the damages from releasing a unit of CO_2. *For sufficiently similar regions, the shared cooling effort therefore reduces the SCC as compared to the case of unilateral action.* However, if the spillover damages are substantially larger than a region's self-imposed damages, e.g., as a result of the sulfur distribution and resulting acid rain patterns, then excess costs are negative and *the spillover term can also increase a region's SCC.*

Climate-clash. In the case where *region A is cooling* and region B engages in CM, region B's deployment term $\frac{S_t^B(m_t)}{m_t}$ turns negative and region A's excess costs

in the spillover term are positive.[28] As a result, the spillover term increases the SCC for the cooling country. In the case where *region A engages in CM* and region B deploys the sulfur, region B's deployment term $\frac{S_t^B(m_t)}{m_t}$ is positive and region A's excess costs turn negative.[29] Again, the spillover term increases the SCC. Thus, *in a climate clash, the spillover term always increases the SCC in both regions. Here, regions always counteract each other. The interference grows stronger for higher CO_2 concentrations, thereby increasing the cost of emitting another ton of carbon.* Depending on whether the positive spillover term dominates the cooling term, the availability of SG can increase or decrease the SCC.

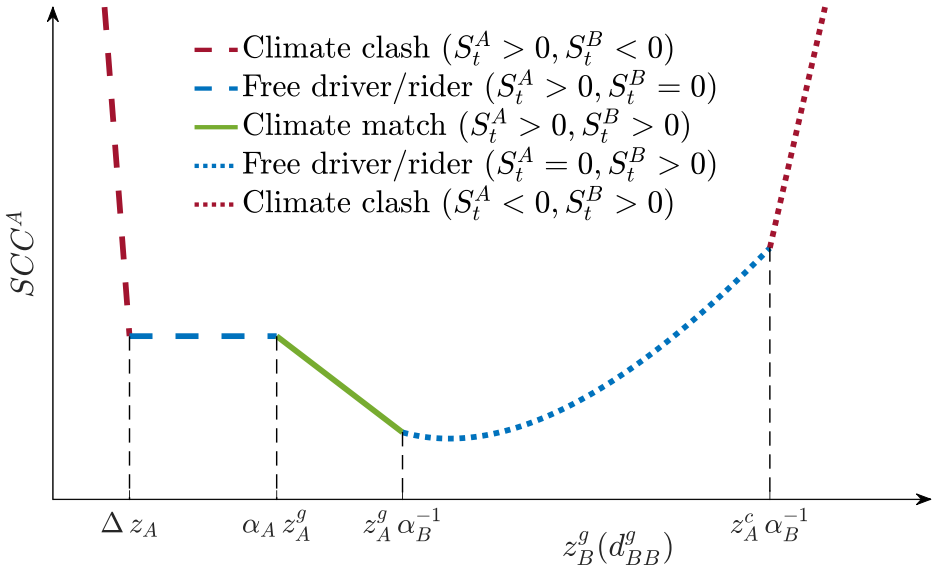

Figure 3: Region A's SCC as a function of region B's SG propensity (by varying d_{BB}^g). Note that $\Delta\, z_A = \alpha_A\, z_A^g - (z_B^c - z_B^g)$.

Figure 3 illustrates region A's SCC across the different equilibrium domains. As in Figure 2, we increase region B's SG propensity along the horizontal axis by varying its SG damages d_{BB}^g.[30] On the left, region B's cooling propensity (and CM reluctance) is small; region A is cooling and region B engages in CM. Here, the clash between the regions grows with each unit of carbon emissions, which raises the social costs of emitting carbon. As region B's damages fall, the conflict and the spillover term decline; region A's SCC falls (red dashes). Once region B's

[28] Region B's CM can at most offset damages in region A ($d_{AA} - d_{BA} = d_{AA}^g + \epsilon_A^g - d_{BA}^c > 0$ by Assumption 2.

[29] Now A's CM can at most offset the damages, $d_{AA} - d_{BA} = d_{AA}^c - \epsilon_A^c - d_{BA}^g < 0$ by Assumption 2.

[30] Region A's SG propensity and CM aversion are fixed and we set $d_{BB}^c = 0.5\, d_{BB}^g$; in addition, we set $d_{AA}^g = d_{BA}^g$.

damages fall sufficiently, region B stops countering and turns inactive. Region A acts unilaterally without interference and its SCC remains constant (blue dashes). As B's SG damages fall further, its SG propensity eventually exceeds the spillovers from region A. Then, region B joins the cooling effort and region A's SCC decreases; the impact of an additional unit of carbon emitted by region A will increasingly be offset by region B's cooling (saving A the effort of sulfur deployment). Eventually, region A becomes inactive (blue dots). From this point onward, only region B cools.

At the beginning of the blue dotted segment, the two regions' SG propensities are still close and region B's cooling causes an unambiguously positive externality for region A; region A is a free-rider in the classical sense. Appendix A.3 show that the marginal sulfur externality is increasing and concave in the carbon concentration m_t and decreasing and convex in the sulfur level S_t. As region B's SG propensity grows, region A's SCC continues to decline until it takes on a minimum at $\alpha_B \, z_B^g = \left(\frac{f_3(1-n)\gamma_A}{f_2\gamma_A + d_{AA}^g}\right)^{\frac{1}{n}}$. Here, the *marginal* externality is maximal. As we move further to the right, the SCC increases again. The marginal externality starts to fall, but the *overall* externality continues to increase up to the point where the marginal externality turns negative. Once the marginal externality turns negative, cooling gains an aspect of free-driving. We can label it "marginal free-driving"; here the marginal externality is negative but the overall externality is still positive ("too much of the good thing"). We simultaneously have a free-driver that overdoes the cooling and a free-rider who still benefits overall ("free-driding"?). Eventually also the overall externality can turn negative. As long as region A abstains from CM, we can label region B a "pure free-driver". However, it depends on the cost-effectiveness of the CM whether this situation arises. If CM are sufficiently cost-effective, region A can engage in CM even before the overall externality turns negative. Then, there is "no free driving" (no lunch punch intended). In the clash, the SCC increases even more steeply (red dotted segment).

4 Quantitative results in the regional setting

This section simulates the strategic model. It incorporates direct heat exchange (dropping Assumption 3) and capital persistence as in Traeger (2022). We choose the US and China as the potentially active SG regions A and B. The rest of the world is divided into 10 regions who can only react by adapting their CO_2 emissions. Our calibration without SG is based on the RICE 2010 model (Nordhaus 2010), which we update based on the current Penn World Tables (Feenstra et al. 2015) and the Global Carbon Project (2021). We increase RICE's climate damages by 50% based on the finding that RICE and DICE damages are lower than more recent damage estimates (Burke et al. 2015, Howard and Sterner 2017, Nordhaus and Moffat 2017, Newell et al. 2021, Traeger 2022). This calibration indicates that the EU, Japan,

and "Other High Income" (OHI) set a higher carbon price than suggested by the Markov game. We assume that these regions have a higher effective damage function, e.g., because they also account for damages in other regions. For these regions, we calibrate a damage multiplier or "care factor" based on the relation between current regional CO_2 price (World Bank 2021) and the model-predicted SCC without SG (and without "care factor"). We discuss details of the calibration and state these adjustment factors in Appendix B.5.

Motivated by Emmerling and Tavoni (2018b), we assume that all countries apart from China have SG damages of 0.1% of GDP per TgS. We assume that China's perceived SG damages are lower and accounted at only 0.05% of GDP per TgS. This assumption will illustrate how a country that is less worried about side-effects from SG can single-handedly regulate the global climate. Figure 4 shows the results of our simulation for a world with an SG option (solid lines) compared to a world without SG (dotted lines).

Geoengineering. China becomes the (only) active region and starts with sulfur injections of about 6 TgS in 2022 that rise to about 10 TgS by the end of the century (upper left panel in Figure 4). This increase follows the increase of the atmospheric carbon concentration according to $S_t^{China}(m_t) = 3.77 \cdot m_t$. The fact that players anticipate the direct heat exchange across regions increases China's geoengineering propensity by 12% (we evaluate Proposition 7 in Appendix A.2 rather than the slightly simplified Proposition 3).

SCC and emissions. The availability of SG leads to a decrease in the region-specific SCC in all regions. As a result, CO_2 emissions increase. The increase in emissions is moderate for the next 50 years. It grows stronger towards the end of the century when the model predicts a curbing of emissions in China and the developing world absent SG. We observe that China's SCC catches up with the European SCC because of its stronger economic growth, which implies that China worries over a (faster) increasing share of global output, even if restricting attention to its own damages.

Table 5 splits up the SCC into its contributions deriving from ocean acidification, the greenhouse effect, and the reduction resulting from from China's SG deployment. Ocean acidification contributes a non-negligible share for all regions.[31] India sees the strongest reduction in mitigation incentives as a result of geoengineering; its SCC falls by over 30%. China, the active player, experiences a reduction of over 25%, and the US and the EU see a reduction around 20%. Russia, who suffers less under climate change and loses out similarly on the geoengineering damages sees the lowest reduction in the SCC; yet, even Russia's SCC is falling under geoengineering. Our numeric simulation uses the extended formula that includes explicit heat exchange

[31] Running the model without ocean acidification damages results in qualitatively similar results. The SCC is slightly lower and temperatures almost approach the 2°C target by the end of the century.

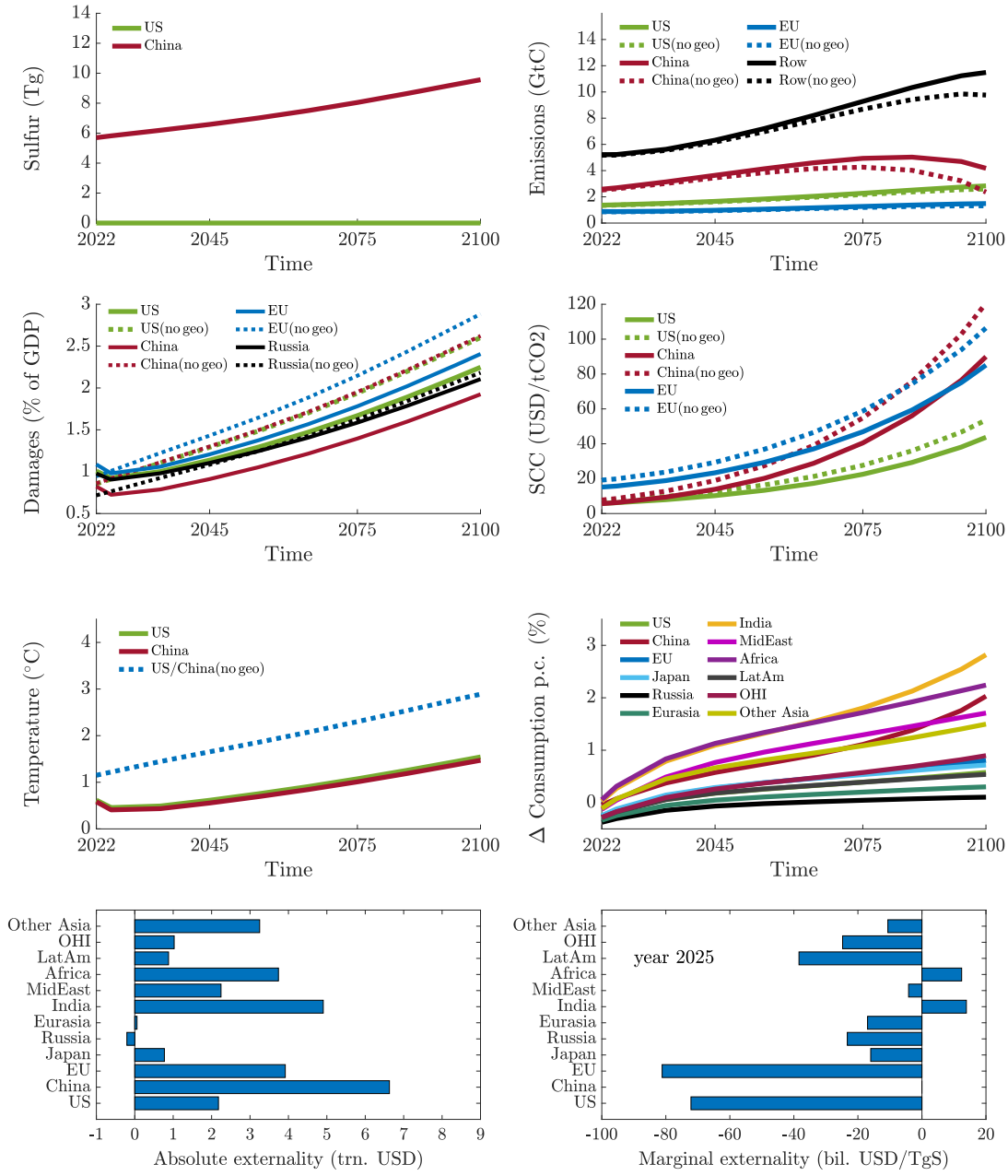

Figure 4: Calibration results with SG (solid) and without SG (dashed) for sulfur deployment, emissions, total damages (SG plus climate), SCC, temperature increase, and the effect of SG on per capita consumption. The bottom left panel shows the absolute externality of sulfur, and the bottom right panel shows the marginal externality of sulfur in the year 2025.

across regions (see AppendixA.2). The last column of Table 5 shows that Proposition 5's simpler formula incurs a minor quantitative error; the column states the net change of the regional SCCs as a result of direct heat exchange.[32]

Table 5: Contributions to the regional Social Cost of Carbon.

Region	SCC in world without geoengineering (Ocean acid + Greenhouse = 100%)	Geo SCC reduction	Δ SCC due to heat exchange
USA	15% + 85%	− 19%	+ 6%
China	15% + 85%	− 27%	+ 7%
EU	14% + 86%	− 21%	+ 7%
Russia	18% + 82%	− 12%	+ 6%
India	9% + 91%	− 31%	+ 8%

Temperature and damages. Our simulations shows that the strong cooling effect of sulfur implies that SG much more than compensates for the increase in emissions, also in the long-run. Temperatures fall initially before increasing again alongside the global carbon concentration, but at a slightly lower rate. Temperature levels under SG remain well below those in a world without SG. Without SG, our calibration implies a temperature increase of 3°C by 2100.[33] This increase drops to 1.5°C in both climate zones as a result of China's SG. However, temperatures keep increasing until they peak in 2165 at 2.2°C. Given strong sulfur spill-overs and heat exchange, the temperature response of the two climate zones is similar. Figure 4 also compares total damages with and without SG. Given the temperature reduction, damages in most regions fall in response to China's SG. However, total damages in Russia are higher under SG until mid century; here, the SG damages cannot make up for the limited cooling benefits.

Consumption impact, marginal externality, and free-riding/driving. Figure 4 shows that the effect of SG on per capita consumption differs substantially across regions and across time. The strongest increase in per capita consumption by 2100 occurs in regions with high climate change damages such as India (+2.8%) and Africa (+2.2%). Regions with low climate change damages, such as Russia, experience a decrease in per capita consumption until mid of the century. The results suggest that the poorer regions tend to benefit more from SG, which might

[32]The model's actual temperature dynamics are the same in both simulations. In the simplified formula, the decision maker does not incorporate the heat exchange when calculating the SCC. Direct heat exchange increases both the greenhouse gas induced warming and the SG-based reduction. Table 5 reports the net difference, where the greenhouse gas induced warming dominates. For the USA (China) SG reduces the SCC by 16% (25%) instead of 19% (27%) when omitting direct heat exchange in the SCC formula.

[33]This temperature increase is slightly higher than the projected increase of 2.7°C from the latest UN Emissions Gap Report (UNEP 2021). The increase of 2.7°C by 2100 is UNEP's (2021) best-estimate if all unconditional 2030 pledges are fully implemented.

not surprise given that these regions also have the most to lose from climate change. This result complements distributional studies of SG undertaken in non-strategic frameworks (e.g. Nordhaus and Boyer 2000, Rickels et al. 2020, Harding et al. 2020). Moreover, the graph shows that – despite higher SG damages – almost all regions are free-riding on China's cooling. Even Russia will eventually benefit from the SG; however, today's net present value is negative for Russia in our simulation (see bottom left panel in Figure 4).

The bottom right panel in Figure 4 presents the marginal externality in 2025, see Appendix A.3 for the underlying formula. We note that the marginal externality is a net present value of costs and benefits. The marginal externality is negative for all regions but Africa and India who suffer the most from climate change. The negative externality is largest for the US and the EU only because of their large GDP, in utils the marginal externality is the most negative for Russia followed by Eurasia and Latin America.[34] We conclude that, for the vast majority of regions, China is a marginal free-driver in our simulation; most regions object to the last ton of sulfur injected, but all regions except for Russia benefit from China's sulfur injections overall. The present section is an example of a calibration of a possible future and a proof of concept. We hope that our model will be used to explore a variety of alternative scenarios.

5 Conclusions

Solar geoengineering (SG) is considered a possible remedy for dangerous climate change, possibly complementing current mitigation efforts. The paper introduces SG into a state of the art integrated assessment model of climate change. We derive closed-form solutions that incorporate current scientific knowledge about the cooling response to stratospheric sulfur injections in a global social planner setting and in a dynamic game. We contribute to understanding and quantifying the interaction between SG, greenhouse gas mitigation, and temperature increase resulting from optimal and strategic actions.

The global model shows that the social planner's optimal sulfur deployment is very sensitive to potential damages from SG, which are mostly unknown. SG lowers the optimal carbon tax; a change that is slightly less sensitive to SG damages than the sulfur injections themselves (as a result of decreasing cooling efficiency). The reduction of the optimal carbon tax is larger for a high cooling efficiency and high climate damages. SG could cut the SCC into half if damages turn out negligible. However, under current damage guesstimates SG would reduce the globally optimal SCC by only 18%.

[34] Appendix A.3 shows the sulfur levels that would be optimal for each region and graphs the marginal externality for 2025 as a function of the sulfur level.

Our strategic model assumes that regional deployment follows a linear Markov strategy where sulfur deployment is conditioned on the atmospheric carbon concentration. Such a linear deployment strategy is the unique optimal course of action in the case of a global planner. We characterize the regional strategies and identify three qualitatively distinct types of equilibria; unilateral action where only one region is active, a climate match where two regions cool the world, and a climate clash where one region cools and another region engages in countermeasures (CM). We show that these equilibria are mutually exclusive and depend on the regional damage characteristics. We study how changes in SG damages, climate impact, and the effectiveness of CM affect the equilibrium domains.

If the active regions are symmetric, they both contribute to global cooling. As the regions become heterogeneous, regions with lower climate damages or higher SG damages will increasingly free-ride on the other region's actions. Given this ability to free-ride, the region's SCC falls and its greenhouse gas emissions increase. Eventually, the free-riding region will stop to contribute entirely and we enter a unilateral cooling equilibrium. There exists a small domain that can be unambiguously labeled a free-rider equilibrium; the inactive region benefits from every unit of SG. As the regions' benefit-cost ratios of SG become more asymmetric, the marginal sulfur injections will eventually harm the free-riding region even if it still benefits overall. Relating to the Weitzmannian "free-driving" metaphor, the active region is free-driving on the margin by "overdoing" SG, but the other region(s) is/are still free-riding overall. Once the overall externality turns negative, the inactive country is unambiguously worse off than without SG; a "pure free-driver equilibrium". If CM exist, such a pure free-driver equilibrium might fall victim to a climate clash. Even in a climate clash, regions can still be better off than without SG. In addition, the climate clash increases the regional SCCs which reduces the CO_2 externality.

By assumption, the rest of the world is either unable or credibly unwilling to engineer the planet's climate. Such regions are always able to free-ride on the cooling, which is beneficial if they are sufficiently similar to the region undertaking SG. In such a scenario, the incentives to mitigate CO_2 drop almost everywhere piling up an increasing stock of atmospheric CO_2. In principle, such a mitigation loss can increase not only the carbon concentration but also the global long-run temperature. However, it is reasonable to assume that the incentives to engage in climate engineering are largest for regions that benefit the most. If these regions have a sufficiently high benefit-cost ratio of SG as compared to the rest of the world, the SCC in the rest of the world can increase. As a result, CO_2 emissions fall under (and despite of) SG.

Our quantitative regional integrated assessment finds that, most likely, the availability of SG increases emissions but reduces temperatures. It also suggests that most regions benefit from an individual country's strategic SG, turning the potential free-driver into the delivery man of a public good. In our simulation, China free-drives on the margin and all other regions apart from Russia experience an

overall positive externality even though we assumed higher SG damages for those regions. Only Africa and India also experience a positive marginal externality and, thus, benefit from every unit of sulfur injected into the stratosphere. We suggested the label "free-driving equilibrium of the first kind" for this situation where a country or region is a free-driver on the margin but still delivers a public good overall with some distributional repercussions.

Direct heat exchange between the regions was omitted in the formulas of the main text but part of the quantitative simulation. We show in the Appendix that direct heat exchange between regions increases a region's incentive to cool the planet because the region benefits not only from its own direct cooling, but also from the spillover cooling in the neighboring regions. Under unilateral action, heat transfer therefore increases a region's cooling incentive. However, if both regions are active, they anticipate the other's response, resulting in a free-riding incentive that reduces or even inverts the original incentive for additional action. These incentives also affect the regional SCCs.

Like any model of a complex real world phenomenon, our framework has many short-comings including a limited freedom in the choice of functional forms. On the upside, the functional forms deliver a good fit of climate dynamics and sulfur-based cooling and, apart from a restriction to log-utility, our regional integrated assessment is state of the art. Another short-coming is that our framework is deterministic, whereas large uncertainties surround both climate change and SG (Emmerling and Tavoni 2018a, Heutel et al. 2018, Kelly et al. 2021). The present framework also sets the stage for an extension incorporating some of these uncertainties following the methods presented in Traeger (2018), which would make it the first quantitative stochastic integrated assessment model with strategic interactions, an interesting challenge for future work.

References

Aldy, Joseph E.; Felgenhauer, Tyler; Pizer, William A.; Tavoni, Massimo; Belaia, Mariia; Borsuk, Mark E.; Ghosh, Arunabha; Heutel, Garth; Heyen, Daniel; Horton, Joshua; Keith, David; Merk, Christine; Moreno-Cruz, Juan; Reynolds, Jesse L.; Ricke, Katharine; Rickels, Wilfried; Shayegh, Soheil; Smith, Wake; Tilmes, Simone; Wagner, Gernot, and Wiener, Jonathan B. Social science research to inform solar geoengineering. Science, 374(6569):815–818, 2021.

Bahn, Olivier; Chesney, Marc; Gheyssens, Jonathan; Knutti, Reto, and Pana, Anca Claudia. Is there room for geoengineering in the optimal climate policy mix? Environmental Science & Policy, 48:67–76, 2015.

Barrett, Scott. The incredible economics of geoengineering. Environmental and resource economics, 39(1):45–54, 2008.

Boucher, Olivier; Kleinschmitt, Christoph, and Myhre, Gunnar. Quasi-Additivity of the Radiative Effects of Marine Cloud Brightening and Stratospheric Sulfate Aerosol Injection. Geophysical Research Letters, 44(21):158–11, 2017.

Brander, Luke M; Rehdanz, Katrin; Tol, Richard SJ, and Van Beukering, Pieter JH. The economic impact of ocean acidification on coral reefs. Climate Change Economics, 3(01):1250002, 2012.

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

CAT, Climate Action Tracker. Cat emissions gap, 2021. URL `https://climateactiontracker.org/global/cat-emissions-gaps/`. Accessed: November 2021.

Colt, Stephen G and Knapp, Gunnar P. Economic effects of an ocean acidification catastrophe. American Economic Review, 106(5):615–19, 2016.

Crutzen, Paul J. Albedo enhancement by stratospheric sulfur injections: A contribution to resolve a policy dilemma? Climatic Change, 77(3-4):211–219, 2006.

Dai, Z.; Weisenstein, D. K., and Keith, D. W. Tailoring Meridional and Seasonal Radiative Forcing by Sulfate Aerosol Solar Geoengineering. Geophysical Research Letters, 45(2):1030–1039, 2018.

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

Dykema, J. A.; Keith, D. W., and Keutsch, F. N. Improved aerosol radiative properties as a foundation for solar geoengineering risk assessment. Geophysical Research Letters, 43(14):7758–7766, 2016.

Dykema, John A; Keith, David W; Anderson, James G, and Weisenstein, Debra. Stratospheric controlled perturbation experiment: a small-scale experiment to improve understanding of the risks of solar geoengineering. Philosophical

Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 372(2031):20140059, 2014.

Emmerling, Johannes and Tavoni, Massimo. Climate Engineering and Abatement: A 'flat' Relationship Under Uncertainty. Environmental and Resource Economics, 69(2):395–415, 2018a.

Emmerling, Johannes and Tavoni, Massimo. Exploration of the interactions between mitigation and solar radiation management in cooperative and non-cooperative international governance settings. Global Environmental Change, 53:244–251, 2018b.

Feenstra, Robert C; Inklaar, Robert, and Timmer, Marcel P. The next generation of the penn world table. American Economic Review, 105(10):3150–3182, 2015. URL `www.ggdc.net/pwt`. Accessed: April 2021.

Flegal, Jane A; Hubert, Anna-Maria; Morrow, David R, and Moreno-Cruz, Juan B. Solar geoengineering: Social science, legal, ethical, and economic frameworks. Annual Review of Environment and Resources, 44:399–423, 2019.

Gerlagh, Reyer and Liski, Matti. Carbon Prices for the Next Hundred Years. Economic Journal, 128(609):728–757, 2018.

Global Carbon Project. Global Carbon Atlas, 2021. URL `http://www.globalcarbonatlas.org/en/content/welcome-carbon-atlas`. Accessed: April 2021.

Goes, Marlos; Tuana, Nancy, and Keller, Klaus. The economics (or lack thereof) of aerosol geoengineering. Climatic Change, 109(3-4):719–744, 2011.

Goeschl, Timo; Heyen, Daniel, and Moreno-Cruz, Juan. The intergenerational transfer of solar radiation management capabilities and atmospheric carbon stocks. Environmental and resource economics, 56(1):85–104, 2013.

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

Hambel, Christoph; Kraft, Holger, and Schwartz, Eduardo S. The carbon abatement game. Technical report, National Bureau of Economic Research, 2018.

Hänsel, Martin C; Drupp, Moritz A; Johansson, Daniel JA; Nesje, Frikk; Azar, Christian; Freeman, Mark C; Groom, Ben, and Sterner, Thomas. Climate economics support for the un climate targets. Nature Climate Change, 10(8):781–789, 2020.

Harding, Anthony and Moreno-Cruz, Juan B. Solar geoengineering economics: From incredible to inevitable and half-way back. Earth's Future, 4(12):569–577, 2016.

Harding, Anthony R; Ricke, Katharine; Heyen, Daniel; MacMartin, Douglas G, and Moreno-Cruz, Juan. Climate econometric models indicate solar geoengineering would reduce inter-country income inequality. Nature communications, 11(1): 1–9, 2020.

Hassler, John and Krusell, Per. Economics and climate change: integrated assessment in a multi-region world. Journal of the European Economic Association, 10 (5):974–1000, 2012.

Heckendorn, P.; Weisenstein, D.; Fueglistaler, S.; Luo, B. P.; Rozanov, E.; Schraner, M.; Thomason, L. W., and Peter, T. The impact of geoengineering aerosols on stratospheric temperature and ozone. Environmental Research Letters, 4(4), 2009.

Heutel, Garth; Moreno-Cruz, Juan, and Shayegh, Soheil. Climate tipping points and solar geoengineering. Journal of Economic Behavior & Organization, 132: 19–45, 2016.

Heutel, Garth; Moreno-Cruz, Juan, and Shayegh, Soheil. Solar geoengineering, uncertainty, and the price of carbon. Journal of Environmental Economics and Management, 87:24–41, 2018.

Heyen, Daniel; Horton, Joshua, and Moreno-Cruz, Juan. Strategic Implications of Counter-Geoengineering: Clash or Cooperation? Journal of Environmental Economics and Management, 95:153–177, 2019.

Hoegh-Guldberg, Ove and Bruno, John F. The impact of climate change on the world's marine ecosystems. Science, 328(5985):1523–1528, 2010.

Howard, Peter H and Sterner, Thomas. Few and not so far between: a meta-analysis of climate damage estimates. Environmental and Resource Economics, 68(1):197–225, 2017.

IPCC. Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, 2021.

Jones, Anthony C.; Hawcroft, Matthew K.; Haywood, James M.; Jones, Andy; Guo, Xiaoran, and Moore, John C. Regional Climate Impacts of Stabilizing Global Warming at 1.5 K Using Solar Geoengineering. Earth's Future, 6(2):230–251, 2018.

Joos, Fortunat; Roth, Raphael; Fuglestvedt, Jan S; Peters, Glen P, et al. Carbon dioxide and climate impulse response functions for the computation of greenhouse gas metrics: a multi-model analysis. Atmospheric Chemistry and Physics, 13(5): 2793–2825, 2013.

Keith, David W. Geoengineering the climate: History and prospect. Annual review of energy and the environment, 25(1):245–284, 2000.

Keith, David W. and MacMartin, Douglas G. A temporary, moderate and responsive scenario for solar geoengineering. Nature Climate Change, 5(3):201–206, 2015.

Keith, David W.; Weisenstein, Debra K.; Dykema, John A., and Keutsch, Frank N. Stratospheric solar geoengineering without ozone loss. Proceedings of the National Academy of Sciences, 113(52):14910–14914, 2016.

Kelly, David L; Heutel, Garth; Moreno-Cruz, Juan B, and Shayegh, Soheil. Solar

geoengineering, learning, and experimentation. Technical report, National Bureau of Economic Research, 2021.

Kleinschmitt, Christoph; Boucher, Olivier, and Platt, Ulrich. Sensitivity of the radiative forcing by stratospheric sulfur geoengineering to the amount and strategy of the SO2injection studied with the LMDZ-S3A model. Atmospheric Chemistry and Physics, 18(4):2769–2786, 2018.

Klepper, Gernot and Rickels, Wilfried. The Real Economics of Climate Engineering. Economics Research International, 2012:1–20, 2012.

Kravitz, Ben; Robock, Alan; Oman, Luke; Stenchikov, Georgiy, and Marquardt, Allison B. Sulfuric acid deposition from stratospheric geoengineering with sulfate aerosols. Journal of Geophysical Research Atmospheres, 114(14):1–7, 2009.

Kravitz, Ben; MacMartin, Douglas G.; Mills, Michael J.; Richter, Jadwiga H.; Tilmes, Simone; Lamarque, Jean-Francois; Tribbia, Joseph J., and Vitt, Francis. First simulations of designing stratospheric sulfate aerosol geoengineering to meet multiple simultaneous climate objectives. Journal of Geophysical Research: Atmospheres, 2017.

Lawrence, Mark G; Schäfer, Stefan; Muri, Helene; Scott, Vivian; Oschlies, Andreas; Vaughan, Naomi E; Boucher, Olivier; Schmidt, Hauke; Haywood, Jim, and Scheffran, Jürgen. Evaluating climate geoengineering proposals in the context of the Paris Agreement temperature goals. Nature Communications, 9(1):3734, 2018.

MacMartin, Douglas G.; Kravitz, Ben; Tilmes, Simone; Richter, Jadwiga H.; Mills, Michael J.; Lamarque, Jean-Francois; Tribbia, Joseph J., and Vitt, Francis. The climate response to stratospheric aerosol geoengineering can be tailored using multiple injection locations. Journal of Geophysical Research: Atmospheres, pages 574–590, 2017.

Manoussi, Vassiliki and Xepapadeas, Anastasios. Cooperation and competition in climate change policies: mitigation and climate engineering when countries are asymmetric. Environmental and Resource Economics, 66(4):605–627, 2017.

Manoussi, Vassiliki; Xepapadeas, Anastasios, and Emmerling, Johannes. Climate engineering under deep uncertainty. Journal of Economic Dynamics and Control, 94:207–224, 2018.

McClellan, Justin; Keith, David W., and Apt, Jay. Cost analysis of stratospheric albedo modification delivery systems. Environmental Research Letters, 7(3): 034019, 2012.

Millard-Ball, Adam. The tuvalu syndrome. Climatic Change, 110(3):1047–1066, 2012.

Moreno-Cruz, Juan B. Mitigation and the geoengineering threat. Resource and Energy Economics, 41:248–263, 2015.

Moreno-Cruz, Juan B. and Keith, David W. Climate policy under uncertainty: A case for solar geoengineering. Climatic Change, 121(3):431–444, 2013.

Moreno-Cruz, Juan B and Smulders, Sjak. Revisiting the economics of climate change: the role of geoengineering. Research in Economics, 71(2):212–224, 2017.

Moreno-Cruz, Juan B; Ricke, Katharine L, and Keith, David W. A simple model to account for regional inequalities in the effectiveness of solar radiation management. Climatic Change, 110(3):649–668, 2012.

Moriyama, Ryo; Sugiyama, Masahiro; Kurosawa, Atsushi; Masuda, Kooiti; Tsuzuki, Kazuhiro, and Ishimoto, Yuki. The cost of stratospheric climate engineering revisited. Mitigation and Adaptation Strategies for Global Change, 22(8):1207–1228, 2017.

Morrow, David R. Ethical aspects of the mitigation obstruction argument against climate engineering research. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 372(2031):20140062, 2014.

National Academies of Sciences, Engineering and Medicine. Reflecting Sunlight: Recommendations for Solar Geoengineering Research and Research Governance. Washington, DC: The National Academies Press, 2021.

Newell, Richard G; Prest, Brian C, and Sexton, Steven E. The GDP-temperature relationship: implications for climate change damages. Journal of Environmental Economics and Management, 108:102445, 2021.

Niemeier, U and Timmreck, C. What is the limit of climate engineering by stratospheric injection of SO2? Atmospheric Chemistry and Physics, 15(16):9129–9141, 2015.

Niemeier, Ulrike and Schmidt, Hauke. Changing transport processes in the stratosphere by radiative heating of sulfate aerosols. Atmospheric Chemistry and Physics, 17(24):14871–14886, 2017.

Nordhaus, William D. Economic aspects of global warming in a post-copenhagen environment. Proceedings of the National Academy of Sciences, 107(26):11721–11726, 2010.

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

Nordhaus, William D and Moffat, Andrew. A survey of global impacts of climate change: Replication, survey methods, and a statistical analysis. Working Paper 23646, National Bureau of Economic Research, 2017. URL `http://www.nber.org/papers/w23646`.

Nordhaus, William D and Sztorc, Paul. DICE2013R: Introduction and user's manual, 2013. URL `http://www.econ.yale.edu/~nordhaus/homepage/homepage/documents/DICE_Manual_100413r1.pdf`.

Parker, A.; Horton, J.B., and Keith, D. W. Stopping Solar Geoengineering Through Technical Means: A Preliminary Assessment of Counter-Geoengineering. Earth's Future, 6:1058–1065, 2018.

President's Science Advisory Committee. Restoring the quality of our environment, 1965. Appendix Y4 in The White House Report of the Environmental Pollution Panel.

Quaas, Martin F; Quaas, Johannes; Rickels, Wilfried, and Boucher, Olivier. Are there reasons against open-ended research into solar radiation management? a model of intergenerational decision-making under uncertainty. Journal of Environmental Economics and Management, 84:1–17, 2017.

Ricke, Katharine L; Moreno-Cruz, Juan B, and Caldeira, Ken. Strategic incentives for climate geoengineering coalitions to exclude broad participation. Environmental Research Letters, 8(1):014021, 2013.

Rickels, Wilfried; Quaas, Martin F; Ricke, Katharine; Quaas, Johannes; Moreno-Cruz, Juan, and Smulders, Sjak. Who turns the global thermostat and by how much? Energy Economics, 91:104852, 2020.

Robock, Alan. 20 Reasons Why Geoengineering May Be a Bad Idea. Bulletin of the Atomic Scientists, 64(2):14–18, 2008.

Schelling, Thomas C. The economic diplomacy of geoengineering. Climatic Change, 33(3):303–307, 1996.

Smith, Wake and Wagner, Gernot. Stratospheric aerosol injection tactics and costs in the first 15 years of deployment. Environmental Research Letters, 13(12): 124001, 2018.

Tollefson, Jeff. Carbon emissions rapidly rebounded following COVID pandemic dip. Nature News, 2021. URL `https://www.nature.com/articles/d41586-021-03036-x`.

Traeger, Christian P. ACE - Analytic Climate Economy (with Temperature and Uncertainty), 2018. URL `https://ssrn.com/abstract=3307622`.

Traeger, Christian P. IAMs and CO_2 emissions – An analytic discussion. EAERE Working Paper, 2021.

Traeger, Christian P. ACE – Analytic Climate Economy. American Economic Journal: Economic Policy, 2022. accepted.

UNEP, United Nations Environment Programme. Emissions Gap Report 2021: The Heat Is On – A World of Climate Promises Not Yet Delivered. Nairobi, 2021.

Urban, Mark C. Accelerating extinction risk from climate change. Science, 348 (6234):571–573, 2015.

Urpelainen, Johannes. Geoengineering and global warming: a strategic perspective. International Environmental Agreements: Politics, Law and Economics, 12(4): 375–389, 2012.

Visioni, Daniele; MacMartin, Douglas G.; Kravitz, Ben; Tilmes, Simone; Mills, Michael J.; Richter, Jadwiga H., and Boudreau, Matthew P. Seasonal Injec-

tion Strategies for Stratospheric Aerosol Geoengineering. Geophysical Research Letters, 46(13):7790–7799, 2019.

Weisenstein, D. K.; Keith, D. W., and Dykema, J. A. Solar geoengineering using solid aerosol in the stratosphere. Atmospheric Chemistry and Physics, 15(20): 11835–11859, 2015.

Weitzman, Martin L. A Voting Architecture for the Governance of Free-Driver Externalities, with Application to Geoengineering. Scandinavian Journal of Economics, 117(4):1049–1068, 2015.

World Bank. Carbon pricing dashboard, 2021. URL `https://carbonpricingdashboard.worldbank.org/map_data`. Accessed: June 2021.

Appendices

Appendix A discusses further results governing the SCC formulas in the rest of the world (the regions not able to undertake geoengineering, Appendix A.1), the extended analytics with explicit heat transfer across regions (Appendix A.2), and the marginal externality imposed by geoengineering on other regions (Appendix A.3). Appendix B discusses the details of the models starting with production, climate damages and (potential) resource scarcity (Appendix B.1), raditative forcing (Appendix B.2), geoengineering damages (Appendix B.3), climate dynamics (Appendix B.4), and some more details on the regional calibration (Appendix B.5).

Appendix C solves the Bellman equation of the global model (Appendix C.1) and proves the first two propositions (Appendices C.2 & C.3). Appendix D solves the regional model, first setting up the Bellman equation and the Markov strategies for the active regions (Appendix D.1), then verifying the solution to the Bellman equation (Appendix D.2), calculating the shadow values of the state variable (Appendix D.3), and finally taking care of the rest of the world (Appendix D.4). Appendix E summarizes the proofs for the regional model building on the solutions derived in Appendix D.

A Further results

Appendix A.1 discusses the SCC formulas in the rest of the world. Appendix A.2 presents the formulas for optimal sulfur deployment and the SCC under direct heat transfer. It discusses how direct heat exchange alters the strategies, equilibria, deployment, and SCC. Finally, Appendix A.3 further discusses the marginal externality both analytically and for our quantitative simulation.

A.1 Rest of the world

We discussed the strategic SG and mitigation decisions in technologically advanced and politically powerful regions that are willing and able to engage in climate engineering. The large part of the world will not be able or willing to perform such planetary alterations. Yet, even a single region's activity will change the mitigation incentives around the globe. This section discusses how the availability of SG (and possibly CM) for some region changes the mitigation incentives around the world.

The "rest of the world" does not engage in climate engineering and affects regions' A and B only indirectly through its emissions and the resulting changes of atmospheric carbon concentrations affecting SG levels. The rest of the world has an aggregate economy similar to that of regions A and B with idiosyncratic production function, damages, and resulting capital and emission dynamics. We denote the sul-

fur related SG damages (or damage reductions) caused to the rest of the world by region $i \in \{A, B\}$ through its deployment of sulfur (or the CM) by $d_{iW} \in \{d^g_{iW}, d^c_{iW}\}$. If region i remains inactive, the respective damage parameter is zero. The resulting overall damages incurred by the rest of the world, as a fraction of its output, are

$$D^W_t(\tau^B_{1,t}, S_t, m_t) = 1 - \exp\left[\xi^W_0 \left(1 - \tau^B_{1,t}\right) - \left(d_{BW} S^B_t + d_{AW}\, \alpha_A\, S^A_t\right) - a^W (m_t - 1)\right]. \tag{19}$$

As we assumed that the rest of the world is part of climate zone B, we also assume that sulfur levels are the same as in region B, including both local deployment S^B_t and spillovers $\alpha_A\, S^A_t$ from region A. We now characterize the impact of SG on the mitigation incentives in the rest of world.

Proposition 6. *If region $i \in \{A, B\}$ acts unilaterally, the SCC in the rest of the world is given by*

$$SCC^W(z) = \frac{Y^{net}_{W,t}}{M_{pre}} \left[a^W + f_1\, \gamma_W - \left(\left(\frac{f_3}{z^n} - f_2 \right) \gamma_W - d^g_{iW} \right) z \right] \tilde{\phi}^W$$

where $z = \alpha_A\, z^g_A$ for $i = A$ (other climate zone active) and $z = z^g_B$ for $i = B$ (same climate zone active). If both regions are active ($S^A_t \neq 0$ and $S^B_t \neq 0$), the SCC in the rest of the world is

$$SCC^W = \frac{Y^{net}_{W,t}}{M_{pre}} \left[a^W + f_1\, \gamma_W - \left(\left(\frac{f_3}{z^n_B} - f_2 \right) \gamma_W - d_{BW} \right) z_B - \frac{\alpha_A (z_A - \alpha_B\, z_B)(d_{BW} - d_{AW})}{1 - \alpha_A\, \alpha_B} \right] \tilde{\phi}^W$$

with $z_A \in \{z^c_A, z^g_A\}$, $z_B \in \{z^c_B, z^g_B\}$, $d_{AW} \in \{d^c_{AW}, d^g_{AW}\}$, and $d_{BW} \in \{d^c_{BW}, d^g_{BW}\}$.

Proof. See Appendix E.5. □

The structure of the SCC in the rest of the world is similar to that of the active regions discussed in Proposition 5. As in the previous section, the first two contributions characterize the SCC in the absence of SG, and a similar corollary follows.

Corollary 2. *The availability of SG in regions A and B can increase, decrease or leave the SCC in the rest of the world unchanged relative to the regional world without SG.*

One region acts unilaterally. If only one region is active, the SCC in the rest of the world is the exact analogue of that of the inactive region in Proposition 5. The one difference lies not in the structure of the SCC, but in the fact that the rest of the world will not match the cooling no matter how beneficial, and it will not be able to engage in CM no matter how damaging SG is. As a result, if climate engineering is beneficial, the rest of the world can free-ride much more than a region with SG potential (it has the credible "strategy" not to engage). But it can also be hit by SG without any ability to respond. As we point out in section 3.2, avoiding the climate clash does not have to be a good thing; the clash allows both regions to get close to their desired climate targets. Without the ability to clash, the rest of the world is forced to inhabit an unfavorable environment.

Without loss of generality we assume that region A is the active region. If the rest of the world's benefit-cost ratio of (passively incurred) SG exceeds that of the active region $\frac{f_3\,\gamma_W}{f_2\,\gamma_W+d^g_{AW}} > \alpha_A^n \frac{(1-n)\,f_3\,\gamma_A}{f_2\,\gamma_A+(d^g_{AA}+\epsilon^g_A)}$ $\left[\,=(\alpha_A z^g_A)^n\right]$ then the rest of world benefits from SG; its marginal cost of emitting carbon is reduced by the cooling term $\left(\left(\frac{f_3}{(\alpha_A\,z^g_A)^n}-f_2\right)\gamma_W - d^g_{AW}\right)\alpha_A\,z^g_A > 0$. Unsurprisingly, this reduction grows with climate impact γ_W in the rest of the world and falls with the incurred SG damages d^g_{AW}. We note that the benefit-cost ratio of SG for the active region has to account for the deployment costs ϵ^g_A and the decreasing effectiveness of SG[35] both of which are absent for the passive region. Thus, *if the regions are somewhat symmetric, the rest of the world will tend to benefit (in the short run) from SG and, as a consequence, increase its CO_2 emissions (worsening the long-run situation). We point out that even a rest of the world that has no interest in ever engaging in climate engineering can increase the World's dependency on SG if only some country is willing to engage in SG. The increase in emissions and therefore the dependency becomes worse the higher the assumed effectiveness of the cooling, the higher the climate damages, and the lower the SG damages to other regions.*

However, it is reasonable to assume that the incentive to do climate engineering are largest for those countries that benefit the most. If the benefit-cost ratios and, thus, SG propensities of the active region is much higher than in the rest of the world, the above inequalities flip and the SCC in the rest of the world increases, making CO_2 emissions even more pricey than in the absence of SG. Then the dependency argument turns around, at least on the planetary scale. Overall emissions decrease to avoid more SG damages from the unilaterally cooling region, and as global emissions decrease, also the cooling country will voluntarily reduce the SG levels.

Climate match. With both regions active, the SCC in the rest of the world resembles that of the active regions (see Proposition 5). However, the rest of the world does not participate in the costly cooling. If its marginal damages from SG

[35]Sulfur deployment scales with the CO_2 concentrations. The active decision maker takes the decreasing effectiveness of sulfur into account and, therefore, has a higher opportunity cost of sulfur deployment.

are independent of the origin, $d^g_{BW} = d^g_{AW}$, then the spillover term is zero. In this case, the situation is identical to the scenario of unilateral action. The rest of the world does not care whether one or both regions engage in SG. Moreover, its SCC reduction (or increase) is proportional to the SG propensity of region B, which shares the climate zone and sets the "local" climate target. If both active regions have similar SG propensities and the marginal damages from SG in region A are lower than those from SG in region B, then the spillover term will reduce the rest of the world's SCC relative to that based on region B's unilateral action.

Climate clash. If region A cools and region B engages in CM, then $\alpha_A(z^g_A - \alpha_B\, z^c_B) > 0$. Provided that CM cannot perfectly offset the damages from SG, $d^c_{BW} < d^g_{AW}$, the spillover term increases the SCC. In the opposite scenario where region B cools and region A engages in CM, we have $\alpha_A(z^c_A - \alpha_B\, z^g_B) < 0$ and $d^g_{BW} \geq d^c_{AW}$. Again, the spill-over term is positive and increases the SCC. *Thus, in a climate clash the spillover term is always positive, and increases the SCC in the rest of the world.* If region B cools and region A is countering it, then the positive spillover term increases the SCC compared to case where region B acts unilaterally. If SG damages from region B's cooling are high and the effectiveness of CM is low, the SCC in the rest of the world can be higher than in a world without SG, *helping to turn the slippery of SG uphill increasing global mitigation.*[36]

A.2 Heat transfer

This section lifts Assumption 3 of the baseline model, introducing direct heat exchange between the regions; temperature change in one region now directly affects the temperature in the other region. As a result, we find adjustments to the SG targets as well as the strategic interactions. Propositions 7 states that under according modifications of the SG propensity (and CM reluctance), Propositions 3 and 4 characterizing the subgame-perfect Nash equilibrium remain valid. Proposition 8 shows how heat transfers alter the SCC.

In our baseline model, interaction between regions was reduced to the spill-over of the cooling sulfur deployment (or CM). Now, cooling one region also directly alters the temperature in the other region, even without sulfur spillover, a natural consequence of heat exchange across the globe.

Proposition 7. *With heat transfers, Propositions 3 and 4 remain valid under the following modifications of the definitions of the SG propensities z^g_A and CM reluctance z^c_A:*

[36] The climate clash equilibrium where $S^A_t < 0$ and $S^B_t > 0$ requires $z^c_A - \alpha_B\, z^g_B < 0$ and $z^g_B - \alpha_A\, z^c_A > 0$. Thus, the availability of SG increases the SCC if $d^g_{BW} > \frac{(1-\alpha_A\,\alpha_B)\left(\frac{f_3}{(z^g_B)^n} - f_2\right)\gamma_W\, z^g_B - \alpha_A(z^c_A - \alpha_B\, z^g_B)d^c_{AW}}{z^g_B - \alpha_A\, z^c_A} > 0$.

If (i) $S_t^B = 0$*, then*

$$z_A^g = \left(\frac{a_A + b_A \, \alpha_A^{-n}}{\delta_A^g} \right)^{\frac{1}{n}},$$

and if (ii) $S_t^B \neq 0$*, then*

$$z_A^g = \left(\frac{a_A - b_A \frac{b_B}{a_B}}{\delta_A^g - b_A \frac{\delta_B}{a_B}} \right)^{\frac{1}{n}}, \quad z_A^c = \left(\frac{a_A - b_A \frac{b_B}{a_B}}{\delta_A^c - b_A \frac{\delta_B}{a_B}} \right)^{\frac{1}{n}},$$

where

$$\begin{aligned} a_A = \quad & (1-n) f_3 \, \gamma_A \quad , \qquad \delta_A^g = (d_{AA}^g + \epsilon_A^g) + \gamma_A \, f_2 + \alpha_A \, \gamma_A^{heat} \, f_2, \\ b_A = \alpha_A & (1-n) f_3 \, \gamma_A^{heat}, \qquad \delta_A^c = (d_{AA}^c - \epsilon_A^c) + \gamma_A \, f_2 + \alpha_A \, \gamma_A^{heat} \, f_2, \end{aligned}$$

and direct climate change impact $\gamma_A \equiv \beta^A \, \xi_0^A \, \tilde{\sigma}_{11}^A \, \sigma_{forc}$*, as well as heat transfer driven climate change impact* $\gamma_A^{heat} \equiv \beta^A \, \xi_0^A \, \tilde{\sigma}_{12}^A \, \sigma_{forc}$*. Swapping region indices characterizes region B's strategies.*

Proof. See Appendix E.1. □

The new term γ_A^{heat} reflects the climate change impact from heat exchange with the other region. Its component $\tilde{\sigma}_{12}^A = \left[(\mathbf{1} - \beta^A \, \boldsymbol{\sigma}^{\boldsymbol{A}})^{-1} \right]_{12}$ characterizes the discounted long-term heat flux from region B to region A resulting from a present heating (forcing) change in region B. This heat exchange affects both, heat increase as a result of greenhouse gas emissions and heat reduction as a result of SG.

We note that the SG propensities and CM reluctance without heat transfers ($\gamma_A^{heat} = 0$) in Propositions 3 and 4 are simply

$$\text{w/o heat transfer:} \quad z_A^g = \left(\frac{a_A}{\delta_A^g} \right)^{\frac{1}{n}}, \quad z_A^c = \left(\frac{a_A}{\delta_A^c} \right)^{\frac{1}{n}},$$

where a_A captures sulfur's direct cooling efficiency f_3 in region A, and the damage impact of global warming γ_A, both of which increase the region's SG propensity. The component δ_A^g captures the damages from and costs of SG, which reduces the SG propensity (as well as a negligible efficiency loss in forcing captured by the close-to-zero f_2, which here gains an additional term capturing heat spillover from region B caused by region A's cooling). Comparing a model with heat transfer to one without heat transfers implies a re-calibration of the heat transfer matrix $\boldsymbol{\sigma}$. In order to describe the same equilibrium temperature response, the matrix elements $[(\mathbf{1} - \boldsymbol{\sigma})^{-1}]_{ij}$ for $i, j \in \{1, 2\}$ have to coincide. For perfectly patient decision

makers, this condition implies that the parameters $\tilde{\sigma}_{ij}$ and, thus, γ_A, γ_B are directly comparable across models. For the impatient decision maker, the climate impact parameters γ_A, γ_B would slightly differ, a minor difference that we ignore in the subsequent discussion.

One active region. If only region A is active, the SG propensity increases by the term $b_A \alpha_A^{-n}$, where b_A *is the heat transfer benefit.* It characterizes the climate–impact–γ_A–weighted cooling benefits in region A that result from its cooling of region B through sulfur spillovers α_A at efficiency f_3. These spillovers then feed back into region A through the heat transfer captured in $\tilde{\sigma}_{12}^A$. The SG propensity adjustment weighs this heat transfer benefit by α_A^{-n}. A higher sulfur spill-over has two implications. First, it increases the marginal benefit from deploying sulfur through the heat transfer benefit (part of b_A). But second, it also implies that the level of sulfur in region B is already high, and the effectiveness of deploying yet another ton of sulfur to cool region B is lower. The net effect of the spillover α_A on the additional SG propensity as a result of heat transfer is positive ($\alpha_A^{-n}\alpha_A = \alpha_A^{1-n}$, which increases in α_A), but it is lower than one might expect when merely considering the heat transfer benefits. It might be more intuitive to call this (negative) heat transfer a "cooling transfer". *In summary, under unilateral action, the cooling transfer gives the region an additional incentive for cooling as it benefits from the global forcing impact of its sulfur deployment.*

Both regions cooling. With two active regions, strategic effects set in. We focus on the additional strategic effects implied by heat transfer. Other strategic implications discussed in Proposition 3 apply alongside. We explain the adjustment of the SG propensity in case (ii) in two steps. First, we assume that region B does not experience any heat transfer benefit ($b_B = 0$). Then, region B's SG propensity remains $z_B^g = \left(\frac{a_B}{\delta_B}\right)^{\frac{1}{n}}$ as in the case without heat transfer. Region A increases its SG propensity according to the contribution $b_A \frac{\delta_B}{a_B}$; its SG increases in response to a higher heat transfer benefit b_A, but less so if region B already has a high SG propensity, which reduces the efficiency of additional cooling.[37] Second, let us turn back on region B's heat transfer benefit b_B. Now also region B benefits from the cooling in region A. Similarly to region A in the first case, it will ramp up sulfur deployment. Anticipating this response of region B, region A lowers its own target as captured by the term $b_A \frac{b_B}{a_B}$. This strategic "free-riding" response increases in the heat transfer benefit (connectedness) of both regions.[38] A short calculation shows that region A's cooling increases as a result of heat (or cooling) transfer, if its own benefit-cost ratio

[37] The term $\frac{\delta_B}{a_B} = \left(\frac{1}{z_B^g}\right)^n$ falls in region B's SG propensity z_B^g.

[38] This additional "free-riding" as a result of heat transfer falls in a_B, which characterizes the climate impact weighted cooling efficiency within region B. A higher a_B relative to b_B makes region B relatively less responsive to switching on the heat transfer. Then, region A lowers its SG propensity less than in the case where it anticipates a stronger response of region B.

before heat transfer $\frac{a_A}{\delta_A^g}$ is larger than the spill-over benefit over region B's SG damages $\frac{b_A}{\delta_B^g}$.[39] *In summary, the heat transfer increases the incentive for cooling. With both regions cooling, efficiency loss and free-riding incentive counter this additional cooling incentive. A region's cooling incentive increases overall under heat transfer whenever its benefit-cost ratio without heat transfers exceeds the spillover benefits to other region's costs ratio (otherwise the free-riding incentive will dominate).*

Countermeasures. Now, let us consider the climate clash scenario where region B cools, and region A deploys a CM. Following the two step interpretation, first, suppose region B does not experience any heat transfer ($\Rightarrow b_B = 0$). Then region B's SG propensity remains $z_B^g = \left(\frac{a_B}{\delta_B}\right)^{\frac{1}{n}}$. Region A's reluctance to engage in CM increases (=less CM) with heat transfer as the denominator in z_A^c decreases by the term $b_A \frac{\delta_B}{a_B}$; region A gets more cooling for less damages and results less contrarian to region B's cooling. We note that the reluctance to CM is more sensitive to heat transfers than the SG propensity in the case above where both regions are cooling since $\delta_A^c > \delta_A^g$. Second, turning on heat transfer for region B ($b_B > 0$) leads to an increase in region B's SG propensity (as in the case above). Region A anticipates the reaction of region B and decreases its target sulfur level (as the numerator in z_A^c decreases by the term $b_A \frac{b_B}{a_B}$). The strategic component renders the regions more contrarian under heat transfers. Overall, the effect of heat transfer on deployment levels depends on the cost-benefit ratios of SG and CM, and the heat benefits in both regions. If heat transfer increases region B's propensity to cool and decreases region A's CM reluctance, both regions increase their deployment levels (the climate clash gets worse). The opposite occurs if heat transfer decreases region B's propensity and increase A's reluctance. In this case, both regions lower their deployment levels, and the climate clash turns less extreme. If heat transfer increases region B's cooling propensity, and increases region A's reluctance the net effect on deployment levels is ambiguous. We note that a decrease in the SG propensity of region B and a decrease in the CM reluctance of region A cannot occur in combination.[40] *In summary, in the*

[39] Region A's propensity increases if $\frac{a_A - b_A \frac{b_B}{a_B}}{\delta_A^g - b_A \frac{\delta_B}{a_B}} > \frac{a_A}{\delta_A^g} \Leftrightarrow \frac{a_A}{\delta_A^g} > \frac{b_B}{\delta_B^g}$, and decreases if $\frac{a_A}{\delta_A^g} \leq \frac{b_B}{\delta_B^g}$. Analogously, region B's propensity increases if $\frac{a_B}{\delta_B^g} > \frac{b_A}{\delta_A^g}$, and decreases if $\frac{a_B}{\delta_B^g} \leq \frac{b_A}{\delta_A^g}$. Assuming $\tilde{\sigma}_{11}^i > \tilde{\sigma}_{12}^i$ implies $\gamma_i > \gamma_i^{heat}$, and thus $a_i > b_i$ for $i \in \{A, B\}$. As a result $\frac{a_A}{\delta_A^g} \leq \frac{b_B}{\delta_B^g}$ can never occur in combination with $\frac{a_B}{\delta_B^g} \leq \frac{b_A}{\delta_A^g}$ since the former implies $\frac{a_A}{\delta_A^g} < \frac{a_B}{\delta_B^g}$ and the latter implies $\frac{a_B}{\delta_B^g} < \frac{a_A}{\delta_A^g}$. Thus, a decrease in both regions' propensities cannot occur.

[40] Region B's SG propensity increases if $\frac{a_B - b_B \frac{b_A}{a_A}}{\delta_B^g - b_B \frac{\delta_A^c}{a_A}} > \frac{a_B}{\delta_B^g} \Leftrightarrow \frac{a_B}{\delta_B^g} > \frac{b_A}{\delta_A^c}$, and decreases if $\frac{a_B}{\delta_B^g} \leq \frac{b_A}{\delta_A^c}$. Analogously, region A's CM reluctance increases if $\frac{a_A - b_A \frac{b_B}{a_B}}{\delta_A^c - b_A \frac{\delta_B^g}{a_B}} > \frac{a_A}{\delta_A^c} \Leftrightarrow \frac{a_A}{\delta_A^c} > \frac{b_B}{\delta_B^g}$, and decreases if $\frac{a_A}{\delta_A^c} \leq \frac{b_B}{\delta_B^g}$. Assuming $\tilde{\sigma}_{11}^i > \tilde{\sigma}_{12}^i$ implies $\gamma_i > \gamma_i^{heat}$, and thus $a_i > b_i$ for $i \in \{A, B\}$. As a result $\frac{a_B}{\delta_B^g} \leq \frac{b_A}{\delta_A^c}$ and $\frac{a_A}{\delta_A^c} \leq \frac{b_B}{\delta_B^g}$ cannot occur in combination since the former implies $\frac{a_B}{\delta_B^g} < \frac{a_A}{\delta_A^c}$ and the

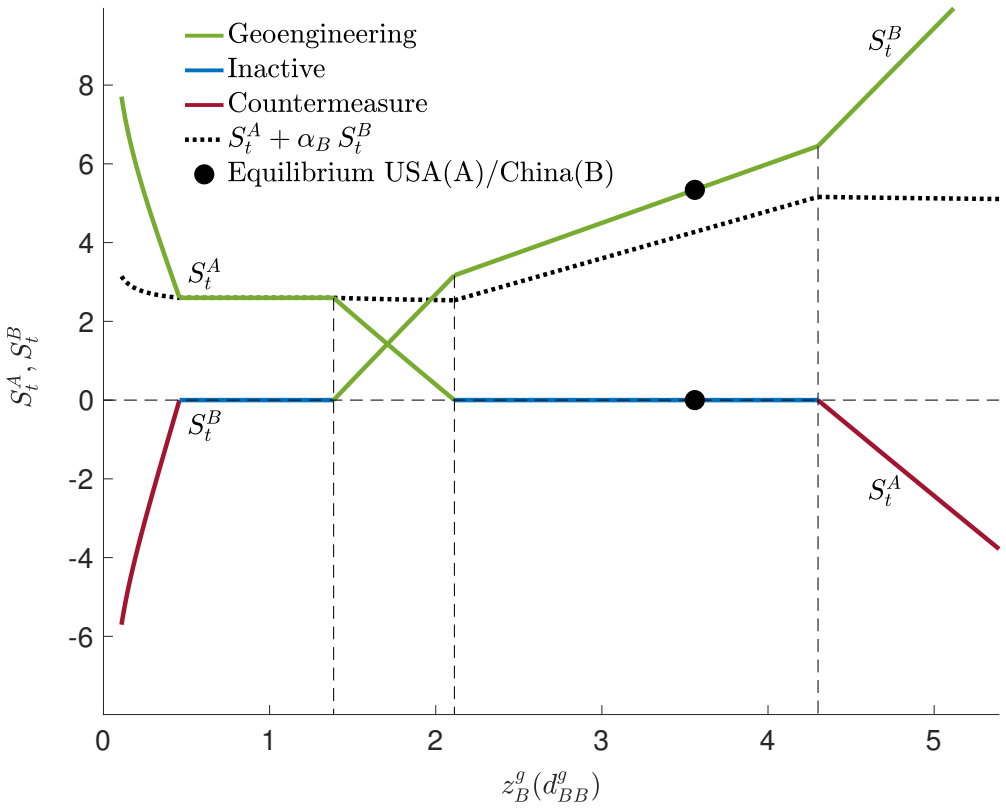

Figure 5: Regions' strategies with heat flows as a function of region B's SG propensity z_B^g by varying d_{BB}^g (at $m_t = 1.5$). Each region either deploys sulfur (shown in green), a CM (shown in red), or stays inactive (shown in blue). The dotted black line depicts the target level of sulfur in region A ($S_t^A + \alpha_B\, S_t^B$). Note that $\Delta\, z_A = \alpha_A\, z_A^g - (z_B^c - z_B^g)$.

clash scenario, the non-strategic implications of heat transfer reduce CM. However, they also increase the cooling region's SG propensity. Thus, the strategic incentives boost the clash. Overall, heat transfer can both turn the climate clash worse or improve the situation, depending on the regions' cost-benefit ratios for SG and CM.

We now derive the regional SCC characterizing the mitigation implications of heat exchange. We focus on the new contributions resulting from heat exchange and abbreviate the original SCC formula of our baseline model as $SCC^A_{w/o}(\cdot)$, indicating as arguments the engineering propensities on which it depends. These propensities change as discussed above, whereas the structure of $SCC^A_{w/o}(\cdot)$ remains identical to that observed in Proposition 5.

latter implies $\frac{\alpha_B}{\delta_B^g} > \frac{\alpha_A}{\delta_A^c}$. Thus, a decrease in region B's SG propensity cannot occur in combination with a decrease in region A's CM reluctance.

Proposition 8. *Under unilateral action, the SCC in region A is*

$$SCC^A = SCC^A_{w/o}(z^g_A) + \frac{Y^{net}_{A,t}}{M_{pre}}\left[f_1\,\gamma^{heat}_A - \left(\frac{f_3}{(\alpha_A z^g_A)^n} - f_2\right)\gamma^{heat}_A\,\alpha_A z^g_A\right]\tilde{\phi}^A \quad \textit{for } S^B_t = 0,$$

$$SCC^A = SCC^A_{w/o}(z^g_B) + \frac{Y^{net}_{A,t}}{M_{pre}}\left[f_1\,\gamma^{heat}_A - \left(\frac{f_3}{(z^g_B)^n} - f_2\right)\gamma^{heat}_A\,z^g_B\right]\tilde{\phi}^A \quad \textit{for } S^A_t = 0.$$

If both regions are active ($S^B_t \neq 0$ and $S^A_t \neq 0$) the SCC is

$$SCC^A = SCC^A_{w/o}(z_A, z_B) + \frac{Y^{net}_{A,t}}{M_{pre}}\left[f_1\,\gamma^{heat}_A - \left(\frac{f_3}{z^n_B} - f_2\right)\gamma^{heat}_A\,z_B\right]\tilde{\phi}^A,$$

with $z_A \in \{z^c_A, z^g_A\}$, and $z_B \in \{z^c_B, z^g_B\}$, depending on whether the corresponding region engages in SG (g) or CM (c). Swapping region indices characterizes region B's SCC. We note that Corollary 1 from the base model still holds.

Proof. See Appendix E.3. □

All cases still contain the original terms abbreviated $SCC^A_{w/o}(\cdot)$. Their contributions depends on the SG propensities and CM reluctancies, which now change because of the strategical responses to heat transfer that we discussed above in Proposition 7. In addition, heat transfer introduces two new terms. The first in the square brackets reflects the heat flow across regions related to greenhouse gas emissions. It is positive and increases in γ^{heat}_A characterizing the heat transfer based climate change impact. This contribution also arises in a regional model with heat flows but without SG. Other damage terms are absent because heat transfer does not affect ocean acidification or sulfur-based damages. The second term in the square brackets is negative and reflects the cooling from SG under heat transfer.

A.3 Marginal externalities (formula)

We obtain the marginal sulfur externalities studying the welfare impact of the marginal ton of sulfur injected for a given region. For this purpose, we take the derivative of the (r.h.s. of the) corresponding Bellman equation w.r.t. sulfur. E.g., we obtain the marginal impact of B's sulfur injections on A's welfare by deriving

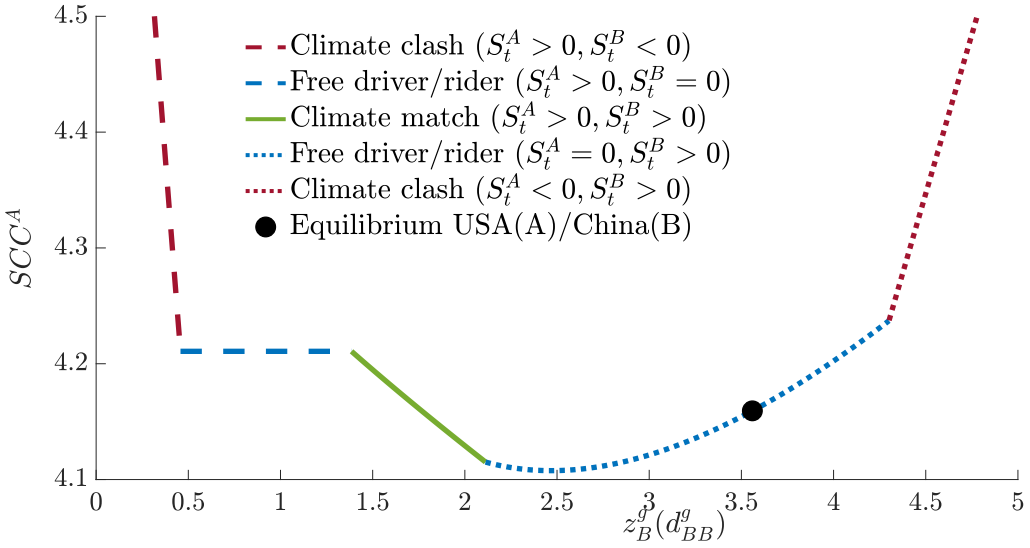

Figure 6: Region A's SCC with heat flows as a function of region B's SG propensity (by varying d_{BB}^g). Note that $\Delta\, z_A = \alpha_A\, z_A^g - (z_B^c - z_B^g)$.

equation (60) with respect to S_t^B delivering[41]

$$\begin{aligned}\frac{\partial B_{nc}^A(m_t, S_t^A)}{\partial S_t^B} = {} & \alpha_B(n-1)\,\beta^A\,\varphi_{\tau 1}^{AA}\,\sigma_{\text{forc}}\,f_3 m_t^n (S_t^A + \alpha_B\, S_t^B)^{-n} \\ & + (n-1)\beta^A\,\varphi_{\tau 1}^{BA}\,\sigma_{\text{forc}}\;f_3 m_t^n (S_t^B + \alpha_A\, S_t^A)^{-n} \\ & + \left[\beta^A\,\varphi_{\tau 1}^{AA}\,\sigma_{\text{forc}}\,f_2\,\alpha_B + \beta^A\,\varphi_{\tau 1}^{BA}\,\sigma_{\text{forc}}\,f_2 - (1+\beta^A\,\varphi_k^A)\alpha_B d_{BA}\right],\end{aligned}$$

which characterizes the marginal externality in utils per TgS injected. The analogous formula applies for the marginal externality in other regions, replacing regional indices and climate zone indices correspondingly. We obtain the marginal externality in consumption equivalents dividing the externality in utils by the region's marginal utility.

Corollary 3. *The marginal sulfur externality is decreasing and convex in the sulfur level S_t and increasing and concave in the relative atmospheric carbon concentration m_t.*

Figure 7 shows the marginal sulfur externality in 2025 for all regions as a function of the sulfur level for our calibration where China acts unilaterally. We note that China's sulfur deployment in 2025 is 5.5 TgS. Figure 8 graphs the optimal regional

[41] To evaluate the marginal externality we have to derive w.r.t. S_t^B, even if S_t^B is itself a strategic function in the dynamic game.

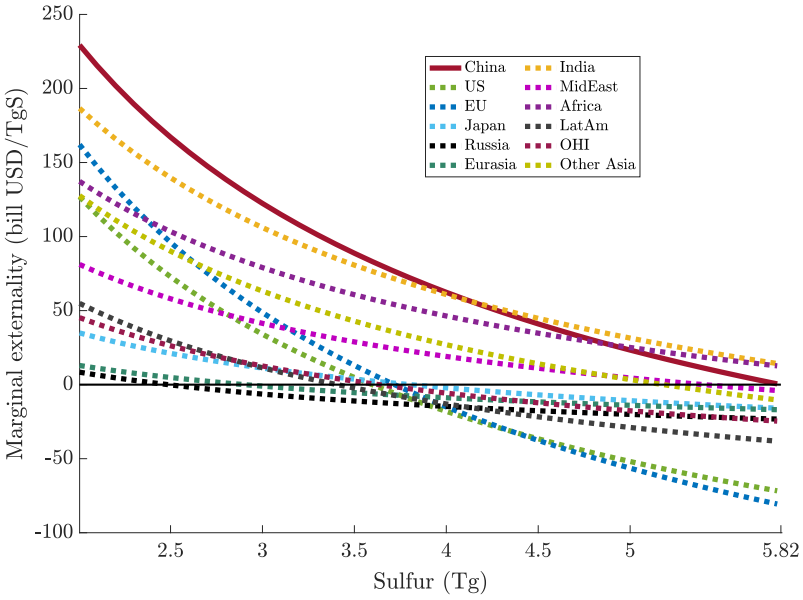

Figure 7: The graph shows the marginal externality for all regions in the year 2025 as a function of the sulfur level.

sulfur levels over time. It shows that relative to China only Africa and India prefer a higher sulfur deployment. All other regions favour a lower sulfur level.

B Model details

B.1 Production, climate damages, resource scarcity.

Global gross output is a function of vectors of dimension I_j with $j \in \{A, N, K, E\}$ and the production function is stated in equation (1). Homogeneity of degree κ in capital is defined as

$$\mathcal{F}(\boldsymbol{A}_t,\, \lambda\, \boldsymbol{K}_t,\, \boldsymbol{N}_t,\, \boldsymbol{E}_t) = \lambda^{\kappa} \mathcal{F}(\boldsymbol{A}_t,\, \boldsymbol{K}_t,\, \boldsymbol{N}_t,\, \boldsymbol{E}_t) \quad \forall \quad \lambda \in \mathbb{R}_+. \tag{20}$$

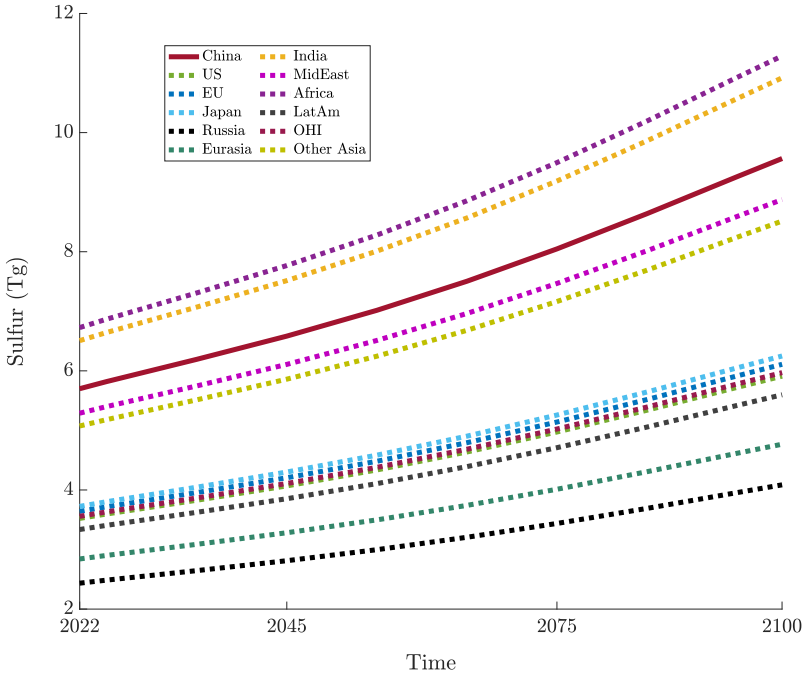

Figure 8: The graph shows the optimal sulfur level for every region.

Population size is normalized to unity $\sum_{i=1}^{I_N} N_{i,t} = 1$.[42] Temperature-based damages are of the form

$$D_T(T_{1,t}) = \xi_0 \exp(\xi_1 T_{1,t}) - \xi_0, \tag{21}$$

see Traeger (2022) for a detailed discussion of the functional form and calibrations to various estimates in the literature. In the regional model, all vectors, the production function, temperature, and the damage parameters carry regional indices.

The first I^d energy inputs $E_1, ..., E_{I^d}$ causing CO_2 emission are collected in the subvector $\boldsymbol{E}_t^d$. The vector $\boldsymbol{R}_t \in \mathbb{R}_+^{I^d}$ characterize fossil fuel resource stocks. The dynamics of the resource stock are

$$\boldsymbol{R}_{t+1} = \boldsymbol{R}_t - \boldsymbol{E}_t^d \tag{22}$$

with initial stock size $\boldsymbol{R}_0 \in \mathbb{R}_+^{I^d}$ given. Renewable energies are indexed by I^{d+1} to

[42] We do not use population weighting in the objective function. The DICE model gives more weight to larger future population, in which case we could not normalize the population to unity. See Traeger (2022) for details.

I_E. To avoid boundary value complications we assume that scarce resources are essential.

B.2 Radiative forcing

Approximation to Kleinschmitt et al. (2018). Figure 9 shows radiative forcing resulting from a given annual flow of sulfur injections at a given atmospheric carbon dioxide concentration in a 3D graph.

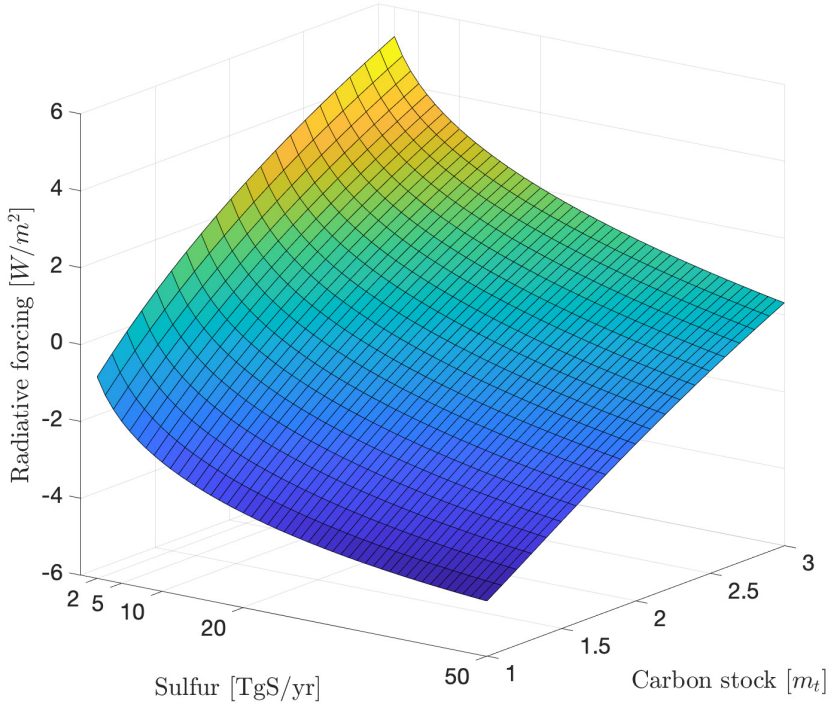

Figure 9: Radiative forcing as a function of the relative atmospheric carbon concentration and sulfur injections, calibrated to data from Kleinschmitt et al. (2018).

Figure 10 illustrates the goodness of our fit, slicing the 3D graph in the two dimensions and adding the data points from Kleinschmitt et al. (2018). The left graph shows radiative forcing as a function of sulfur for different atmospheric carbon concentrations. The right graph shows radiative forcing as a function of the atmospheric carbon concentration for different sulfur injection rates.

Approximation to Niemeier and Schmidt (2017). We also calibrate our radiative forcing equation F_t to data from Niemeier and Schmidt (2017) (see Table 6) over the relative atmospheric carbon interval $m_t \in [1.5, 3]$. The estimated forcing parameters are given in Table 7 and are in line with Assumption 1.

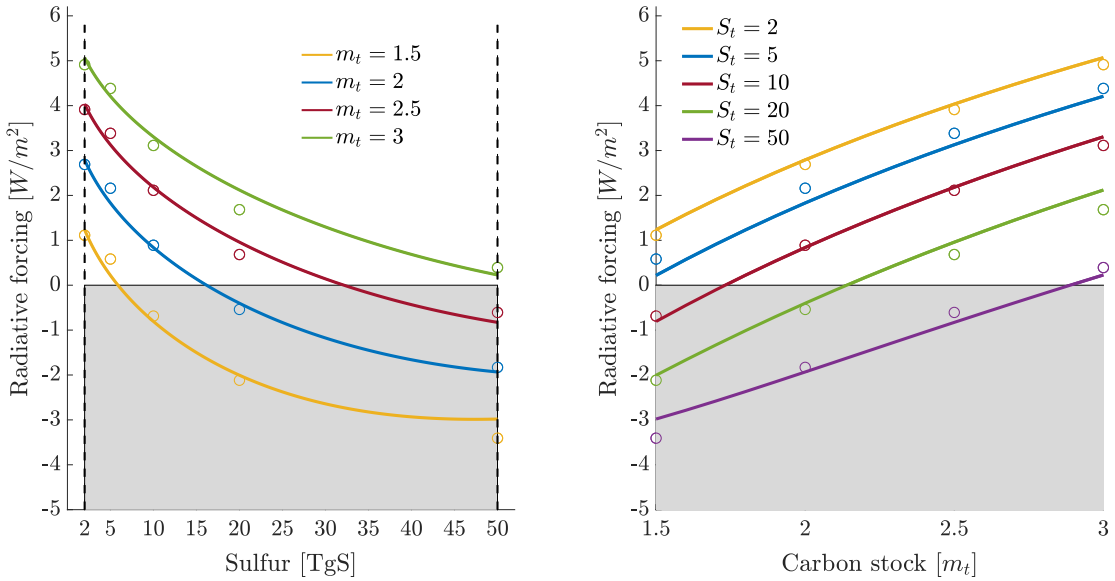

Figure 10: Approximation of radiative forcing based on data (shown by circles) from Kleinschmitt et al. (2018) for sulfur injections between 2 and 50 TgS and positive radiative forcing levels as stated in Assumption 1.

Table 6: Effective radiative forcing effect from sulfur injections (Niemeier and Schmidt 2017).

4 TgS	6 TgS	8 TgS	10 TgS	30 TgS	40 TgS	50 TgS
-0.34	-1.30	-1.54	-1.78	-4.04	-4.76	-5.18

The left graph in Figure 11 shows radiative forcing as a function of sulfur for different atmospheric carbon concentrations. The right graph shows radiative forcing as a function of the atmospheric carbon concentration for different sulfur injection rates.

Optimal radiative forcing. Inserting optimal sulfur deployment S_t^* as given in Proposition 1 into equation (6) yields the optimal level of radiative forcing as a function of the damage parameter d and the atmospheric carbon stock m_t,

$$F_t^* = \frac{\eta}{\log(2)} \log \left[f_0 + f_1 \, m_t + \left(f_2 - f_3 \left(\frac{m_t}{S_t^*} \right)^n \right) S_t^* \right]. \tag{23}$$

Table 7: Estimated forcing parameters.

f_0	f_1	f_2	f_3	n
0.4	2.9	0.004	2.08	0.9

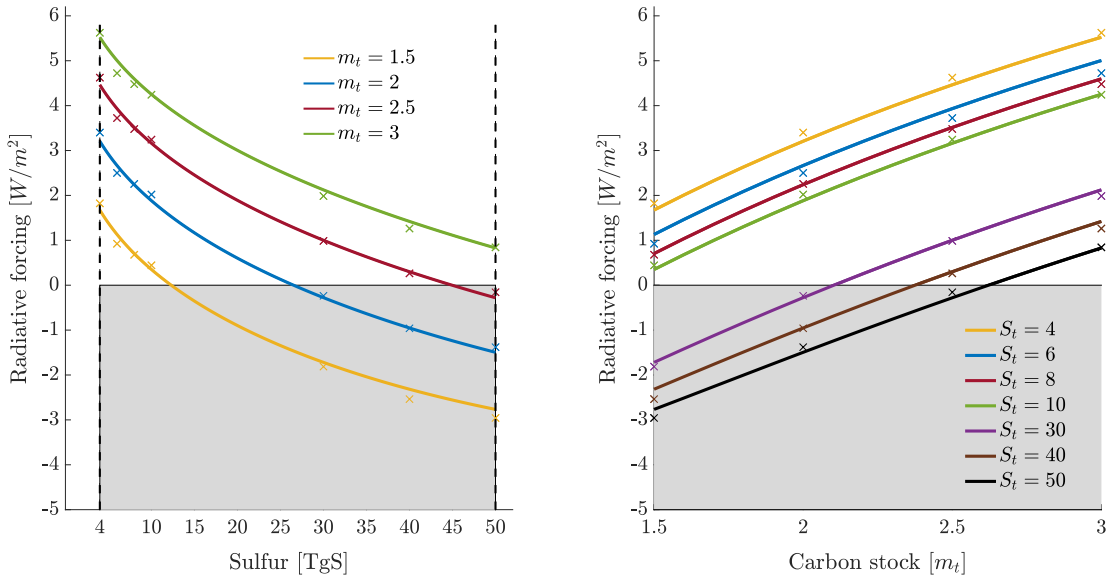

Figure 11: Approximation of radiative forcing based on data (shown by cross markers) from Niemeier and Schmidt (2017) for sulfur injections between 4 and 50 TgS and positive radiative forcing levels as stated in Assumption 1.

We show optimal radiative forcing in Figure 12.

B.3 Geoengineering damages

B.3.1 Damage estimates

Here we disucss the damage estimates in Table 4 and our translation into the damage semi-elasticity d. Heutel et al. (2018) assume a cost of 3% of world output for resetting radiative forcing to its preindustrial level, independent of the prevailing forcing level. In general, higher CO_2 concentrations also increase the costs of cooling the planet back to preindustrial levels. We interpret their costs as the average of neutralizing the forcing of carbon concentrations $m \in \{1.5, 2, 2.5, 3\}$ and find an approximate cost guesstimate of $d_H \approx 0.21\%$ per TgS. [43] Emmerling and Tavoni's (2018b) guesstimate is for a forcing reduction of $3.5\text{W}/\text{m}^2$, independent of the prevailing forcing level. Our model captures a decreasing efficiency of sulfur

[43] To translate the value from Heutel et al. (2018) into our model, we denote by S_m^{pre} the sulfur levels required to neutralize the antropogenic forcing of carbon concentrations of $m \in \{1.5, 2, 2.5, 3\}$. This range of carbon concentrations corresponds approximately to the concentrations along the simulated paths in Heutel et al. (2018). Then a damage of 3 percent of output implies $d_H = \frac{1}{4} \sum_{m=1.5}^{3} \frac{0.03}{S_m^{pre}} \approx 0.21\%$ per TgS.

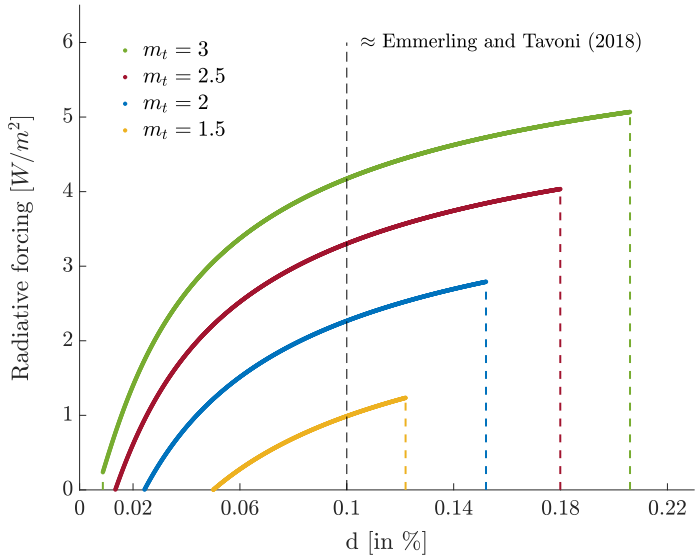

Figure 12: Optimal radiative forcing as a function of the SG damage parameter d for different relative atmospheric carbon stocks m_t.

deployment, and a 3.5W/m^2 reduction at low levels of cooling requires lower sulfur injections than the same forcing reduction at an already high level of cooling. Using an intermediate value, we convert Emmerling and Tavoni's (2018b) guesstimate into a damage of 0.1% per TgS.[44] Goes et al. (2011) state damages for counteracting a doubling CO_2, which corresponds to a forcing of 3.5W/m^2, translating their suggested damage range of 0-5% into a range of 0-0.17% per TgS for d.

B.3.2 Regional geoengineering damages.

We define damages for region B symmetric to damages for region A, which we defined in the main part of the paper. Thus, damages for region B as a fraction of output are given by

$$D_t^B(\tau_{1,t}^B, S_t, m_t) = 1 - \exp\left[\xi_0^B \left(1 - \tau_{1,t}^B\right) - \left(d_{BB} S_t^B + d_{AB}\, \alpha_A\, S_t^A\right) - a^B \left(m_t - 1\right)\right], \tag{24}$$

[44] Figure 10 in Appendix B.2 shows how, at low initial cooling levels, a cooling by 3.5W/m^2 from 4W/m^2 to 0.5W/m^2 can be achieved with 25 TgS (along the red curve). In contrast, with a higher CO_2 concentration, reducing forcing from 4W/m^2 to 0.5W/m^2 requires over 40 TgS (along the green curve). Using a value of 30 TgS, we obtain the damage guesstimate of $d_E = 0.1\%$ per TgS.

with

$$d_{BB} = \begin{cases} d^g_{BB} + \epsilon^g_B & \text{for} \quad S^B_t > 0 \\ d^c_{BB} - \epsilon^c_B & \text{for} \quad S^B_t < 0 \\ 0 & \text{for} \quad S^B_t = 0 \end{cases} \tag{25}$$

and

$$d_{AB} = \begin{cases} d^g_{AB} & \text{for} \quad \alpha_A \, S^A_t > 0 \\ d^c_{AB} & \text{for} \quad \alpha_A \, S^A_t < 0 \\ 0 & \text{for} \quad \alpha_A \, S^A_t = 0. \end{cases} \tag{26}$$

B.4 Climate dynamics (global and regional model)

B.4.1 Global climate dynamics

Carbon dioxide. Following DICE, we consider three carbon reservoirs, atmosphere (carbon content M_1), upper ocean (carbon content M_2) and lower ocean (carbon content M_3) which we summarize in the vector $\boldsymbol{M}$. An extension to additional carbon reservoirs is straight-forward. The dynamics of the carbon reservoirs is

$$\boldsymbol{M}_{t+1} = \boldsymbol{\Phi} \boldsymbol{M}_t + \tilde{\boldsymbol{e}}_t, \tag{27}$$

with the carbon cycle's transition matrix $\boldsymbol{\Phi}$. Further we define $\tilde{\boldsymbol{e}}_t = \boldsymbol{e}_1 \, E^{tot}_t$, with total CO_2 emissions $E^{tot}_t = \sum_{i=1}^{I^d} E_{i,t} + E^{\text{exo}}_t$ resulting from industrial fossil fuel burning and other exogenous processes including land use change and forestry. Similar to Traeger (2022), we define $\tilde{\phi} = [(\mathbf{1} - \beta \, \boldsymbol{\Phi})^{-1}]_{11}$. Instead of a simple decay, it captures how much carbon inserted into the atmosphere remains in or returns to the atmosphere over the discounted infinite time horizon.

Temperature dynamics. In the medium to long run a new level of radiative forcing implies the new atmospheric equilibrium temperature $T_{0,t} = \frac{s}{\eta} \, F_t$. Following ACE, we model the evolution of atmospheric temperature $T_{1,t}$ as a generalized mean of last period's atmospheric temperature (persistence), the last period's ocean temperature (currently cooling), and the new equilibrium temperature corresponding to radiative forcing $T_{0,t}$. Similarly ocean temperature $T_{2,t}$ evolves as a generalized mean of own past and atmospheric temperature

$$\begin{aligned} T_{1,t+1} &= \frac{1}{\xi_1} \log \Big((1 - \sigma^{forc} - \sigma_{21}) \exp(\xi_1 \, T_{1,t}) + \sigma^{forc} \exp(\xi_1 \, T_{0,t}) + \sigma_{21} \exp(\xi_1 \, T_{2,t}) \Big) \\ T_{2,t+1} &= \frac{1}{\xi_1} \log \Big((1 - \sigma_{12}) \exp(\xi_1 \, T_{2,t}) + \sigma_{12} \exp(\xi_1 \, T_{1,t}) \Big) \end{aligned} \tag{28a}$$

with $\xi_1 = \frac{\log 2}{s}$. We rewrite these equations in terms of transformed temperatures

$\tau_{it} = \exp(\xi_1 T_{i,t})$ as

$$\begin{pmatrix} \tau_{1,t+1} \\ \tau_{2,t+1} \end{pmatrix} = \underbrace{\begin{pmatrix} 1-\sigma_{\text{forc}}-\sigma_{21} & \sigma_{21} \\ \sigma_{12} & 1-\sigma_{12} \end{pmatrix}}_{\equiv \boldsymbol{\sigma}} \begin{pmatrix} \tau_{1,t} \\ \tau_{2,t} \end{pmatrix} + \begin{pmatrix} \sigma_{\text{forc}} \exp\left(\frac{\log(2)}{\eta} F_t\right) \\ 0 \end{pmatrix}. \tag{28b}$$

Similar to Traeger (2022), we define $\tilde{\sigma} = [(\mathbf{1} - \beta\,\boldsymbol{\sigma})^{-1}]_{1,1}$, characterizing the discounted heat increase over the infinite time horizon resulting from a heat influx into the atmosphere in the present.

B.4.2 Regional climate dynamics

Carbon dioxide. Similar to the global model, we again consider three carbon reservoirs. The dynamics of the carbon reservoirs are given by

$$\begin{pmatrix} M_{1,t+1} \\ M_{2,t+1} \\ M_{3,t+1} \end{pmatrix} = \begin{pmatrix} \phi_{11} & \phi_{21} & \phi_{31} \\ \phi_{12} & \phi_{22} & \phi_{32} \\ \phi_{13} & \phi_{23} & \phi_{33} \end{pmatrix} \begin{pmatrix} M_{1,t} \\ M_{2,t} \\ M_{3,t} \end{pmatrix} + \begin{pmatrix} \sum_{i=1}^{I^d} E_{A,i,t} + \sum_{i=1}^{I^d} E_{B,i,t} + \sum_{i=1}^{I^d} E_{W,i,t} + E_t^{\text{exo}} \\ 0 \\ 0 \end{pmatrix} \tag{29}$$

or equivalently

$$\boldsymbol{M}_{t+1} = \boldsymbol{\Phi} \boldsymbol{M}_t + \tilde{\boldsymbol{e}}_{\boldsymbol{t}}. \tag{30}$$

Regional temperature dynamics. We characterize two climate zones by temperature levels $T_{1,t}^A$ and $T_{1,t}^B$, containing the two active regions. For simplicity, we assume that the rest of the world is part of region B's climate zone. Analogously to the global case, we let $T_{0,t}^A = \frac{s}{\eta} F_t^A$ and $T_{0,t}^B = \frac{s}{\eta} F_t^B$. The two regional atmospheric temperatures and the ocean temperature evolve as

$$\begin{aligned} T_{1,t+1}^A = \frac{1}{\xi_1} \log \big(&(\sigma_A^A) \exp\left(\xi_1 T_{1,t}^A\right) + \sigma_{\text{forc}} \exp\left(\xi_1 T_{0,t}^A\right) + \sigma_B^A \exp\left(\xi_1 T_{1,t}^B\right) \\ &+ \sigma_O^A \exp\left(\xi_1 T_{2,t}\right) \big) \end{aligned}$$

$$T^B_{1,t+1} = \frac{1}{\xi_1}\log\Big((\sigma^B_B)\,\exp\left(\xi_1\,T^B_{1,t}\right) + \sigma_{\text{forc}}\,\exp\left(\xi_1\,T^B_{0,t}\right) + \sigma^B_A\,\exp\left(\xi_1\,T^A_{1,t}\right) + \sigma^B_O\,\exp\left(\xi_1\,T_{2,t}\right)\Big)$$

$$T_{2,t+1} = \frac{1}{\xi_1}\log\Big((\sigma^O_O)\,\exp\left(\xi_1\,T_{2,t}\right) + \sigma^O_A\exp\left(\xi_1\,T^A_{1,t}\right) + \sigma^O_B\exp\left(\xi_1\,T^B_{1,t}\right)\Big),$$

where $\sigma^A_A = 1-\sigma^A_B-\sigma^A_O-\sigma_{\text{forc}}$, $\sigma^B_B = 1-\sigma^B_A-\sigma^B_O-\sigma_{\text{forc}}$, and $\sigma^O_O = 1-\sigma^O_A-\sigma^O_B$.

Analogously to the global case, we define transformed temperatures $\tau^A_{1,t} = \exp(\xi_1\,T^A_{1,t})$, $\tau^B_{1,t} = \exp(\xi_1\,T^B_{1,t})$, and $\tau_{2,t} = \exp(\xi_1\,T_{2,t})$ with $\xi_1 = \frac{\log 2}{s}$. Then the regional temperature dynamics evolves according to

$$\begin{pmatrix}\tau^A_{1,t+1}\\ \tau^B_{1,t+1}\\ \tau_{2,t+1}\end{pmatrix} = \underbrace{\begin{pmatrix}\sigma^A_A & \sigma^A_B & \sigma^A_O\\ \sigma^B_A & \sigma^B_B & \sigma^B_O\\ \sigma^O_A & \sigma^O_B & \sigma^O_O\end{pmatrix}}_{\equiv\boldsymbol{\sigma}^A}\begin{pmatrix}\tau^A_{1,t}\\ \tau^B_{1,t}\\ \tau_{2,t}\end{pmatrix} + \begin{pmatrix}\sigma_{\text{forc}}\exp\left(\frac{\log 2}{\eta}F_t(S^A_t+\alpha_B\,S^B_t)\right)\\ \sigma_{\text{forc}}\exp\left(\frac{\log 2}{\eta}F_t(S^B_t+\alpha_A\,S^A_t)\right)\\ 0\end{pmatrix}, \quad (31)$$

or equivalently

$$\boldsymbol{\tau}_{t+1} = \boldsymbol{\sigma}^A\boldsymbol{\tau}_t + \tilde{\boldsymbol{F}}_t, \quad \text{with} \quad \tilde{\boldsymbol{F}}_t = \begin{pmatrix}\sigma_{\text{forc}}F^{CO2}_t(m_t, S^A_t+\alpha_B\,S^B_t)\\ \sigma_{\text{forc}}F^{CO2}_t(m_t, S^B_t+\alpha_A\,S^A_t)\\ 0\end{pmatrix}, \quad (32)$$

where $\sigma^A_A = 1-\sigma^A_B-\sigma^A_O-\sigma_{\text{forc}}$, $\sigma^B_B = 1-\sigma^B_A-\sigma^B_O-\sigma_{\text{forc}}$, and $\sigma^O_O = 1-\sigma^O_A-\sigma^O_B$. To preserve symmetry in notation, we define the corresponding matrix $\boldsymbol{\sigma}^B$ by swapping the first and second rows and columns, characterizing the identical dynamics from the perspective of region B. In anticipation of a similar climate impact as in the global case, we define

$$\tilde{\sigma}^A_{ij} = \left[(\mathbf{1}-\beta^A\,\boldsymbol{\sigma}^{\boldsymbol{A}})^{-1}\right]_{ij} \text{ and } \tilde{\sigma}^B_{ij} = \left[(\mathbf{1}-\beta^B\,\boldsymbol{\sigma}^{\boldsymbol{B}})^{-1}\right]_{ij} \text{ for } i,j\in\{1,2\}.$$

The term $\tilde{\sigma}^A_{11}$ characterizes the discounted heat increase in region A over the infinite time horizon resulting from a heat influx into region A's atmosphere in the present, and $\tilde{\sigma}^A_{12}$ characterizes the discounted heat increase in region A over the infinite time horizon resulting from an influx into region B's atmosphere in the present.[45] Equipped with the formal characterization of temperature dynamics, we can now

[45] We obtain this interpretation from expanding the inverse $(\mathbf{1}-\beta^A\,\boldsymbol{\sigma}^{\boldsymbol{A}})^{-1} = \sum_{l=0}^{\infty}\beta^{A^l}\boldsymbol{\sigma}^{\boldsymbol{A}^l}$ using the von Neumann series. The entry i-j of the l^{th} power of the transition matrix σ^A, i.e. $\left[\boldsymbol{\sigma}^{\boldsymbol{A}^l}\right]_{ij}$, characterizes how much of the heat flow entering temperature layer j in the present still remains in or returns to layer i after l periods.

state Assumption 3 formally.

Assumption 3. *The heat flow coefficients $\sigma_B^A, \sigma_A^B, \sigma_A^O$, and σ_B^O are zero.*

B.5 Regional calibration

We calibrate the strategic model with region A being the US and region B being China. These are the potentially active SG-players. We split the rest of the world into 10 regions. The calibration is based on the (most recent official) RICE 2010 model by Nordhaus (2010). We update production and growth using the latest Penn World Tables (Feenstra et al. 2015) and we update emissions using the Global Carbon Project (2021). The carbon cycle is based on DICE 2013 (Nordhaus and Sztorc 2013), which – in combination with our updated temperature dynamics – delivers an SCC close to what we would get using Joos et al.'s (2013) recently promoted box model, see Traeger (2022). We calibrate the climate damage coefficients to RICE's regional climate damages and increase the damage coefficient by 50 percent reflecting the increase in global damages estimates since the appearance of RICE 2010, see Traeger (2022) for an extended discussion. Note that the numerical simulations also include direct heat exchange between regions as discussed in Appendix A.2. We solve the infinite horizon game and simulate from 2015 forward.

CO_2 emissions and abatement. Our calibration uses the analytic solution for DICE's emissions derived by Traeger (2021)

$$E_t = \sigma_t \, (1 - \mu_t) \, Y_t^{\text{gross}}, \tag{33}$$

with gross output

$$Y_t^{\text{gross}} = A_t \, K_t^{\kappa} \, N_t^{1-\kappa}, \tag{34}$$

and abatement rate

$$\mu_t = \left(\frac{\Gamma_t}{p_t^{back} [1 - D_t(T_{1,t})]} \right)^{\frac{1}{\theta_2 - 1}}, \tag{35}$$

where p_t^{back} denotes the backstop price in period t, θ_2 measures the convexity of abatement costs, and σ_t denotes the carbon intensity. We adopt the regional values for the backstop price and the declines rates for the carbon intensities from RICE 2010 (Nordhaus 2010). Initial carbon intensities are updated based on current GDP and emissions data (Feenstra et al. 2015, Global Carbon Project 2021).

Capital growth. To obtain analytic solutions, most approaches to analytic IAMs assume full capital depreciation (over a 10 year time step). However, this approximation does not perform particularly well. To account for the incomplete depreciation of capital over a decade, we use ACE's extended capital accumulation

formula (Traeger 2022),

$$K_{t+1} = (Y_t^{net} - C_t) \left[\frac{1 + g_t^K}{\delta^K + g_t^K} \right], \tag{36}$$

incorporating the correction factor for capital persistence where δ^K is the capital depreciation and g_t^K is an exogenous approximation of the growth rate of capital. We estimate δ^K and g_t^K using a 10 year trend in the regional aggregates of the Penn World Tables (Feenstra et al. 2015).

Investment rates and capital shares. We use regional investment rates from the Penn World Tables (Feenstra et al. 2015) and we use the same rate of pure time preference as in the global model for all regions.[46] We calculate the regional capital shares based on this data. We bound capital shares between 0.3 and 0.4, adjusting investment rates accordingly where needed.

Growth of Total Factor Productivity (TFP). We estimate the growth rate of world GDP based on PWT output data over the last 10 years (Feenstra et al. 2015). Using RICE's population growth (Nordhaus 2010) and our calculated capital shares delivers the world TFP growth rate, denoted by $g_{A,0}$. We assume that the growth rate is composed of a fixed long-term trend, g_A^{long}, and a part that changes over time, $g_{A,t}^{\Delta}$,

$$g_{A,t} = g_A^{long} + g_{A,t}^{\Delta}.$$

We assume that long-run TFP growth is half of the current TFP growth,

$$g_A^{long} = 0.5\, g_{A,0},$$

and that the time-changing part of the trend falls slowly over time

$$g_{A,t}^{\Delta} = g_{A,0}^{\Delta} \exp(-\delta_A \cdot t),$$

where we set the rate of decline of TFP growth to 2 percent per year.[47]

We calculate regional initial TFP growth rates analogous to the global initial TFP growth rate using the 10 year trends in the PWT. We then assume that regional TFP growth of region i converges to the global growth rate over time at a speed parameterized by the regional weights ω_i, i.e., regional growth rates are

$$g_{A,t+1}^i = \omega^i\, g_{A,t}^i + (1 - \omega^i)\, g_{A,t}.$$

[46] We want to avoid a built-in permanent difference between the inhabitants of different regions.

[47] Our assumptions are a "compromise" between the usual macroeconomic assumptions of (higher) mean-reverting growth and the assumptions of DICE and RICE that the growth rate eventually falls to zero.

Care factors. We calculate the current regional SCCs using the World Bank's (2021) carbon pricing dashboard. The EU, Japan, and other high income countries (OHI) have a higher SCC than our non-cooperative equilibrium (absent SG) would suggest. In our interpretation, these regions incorporate damages beyond those incurred to their own region and we raise their concern about climate change in our model accordingly using what we coined "care factors" (CF_i). These care factors multiply a region's damages. We calculate these care factors as the ratio of a region's implemented SCC and its non-cooperative SCC predicted by the model without SG. We calculate the care factor of region i by summing over a region's carbon pricing initiatives $j \in J \subset \mathbb{N} \cup \{0\}$. Let initiative j's carbon price be p_j^{carbon} and cover emissions $E_{i,j}^{cov}$ in region i. The care factors are

$$\text{CF}_i = \frac{\sum_{j \in J} E_{i,j}^{cov} \cdot p_j^{carbon}}{E_i \cdot \text{SCC}_i^{no\,geo}},$$

where E_i are the total emissions of region i and $\text{SCC}_i^{no\,geo}$ is the region's SCC predicted by the model absent SG. We adopt care factor for those regions where $\text{CF}_i > 1$, i.e., for those regions that appear to care for more than their own regional damages.[48] The full set of parameter values for our calibration is shown in Table 9.

Table 8: Regional care factors (regions with $\text{CF}_i > 1$).

Region	EU	Japan	OHI
CF_i	2.21	1.34	5.39

[48]For all other regions we set the care factors equal to one, i.e., we do not further adjust the SCC which then reflects merely their own regional damages.

Table 9: Economic and climate parameters.

	Parameter	Value	Source
$Y_{0,i}$	initial GDPs	regional	PWT
$K_{0,i}$	initial capital stocks	regional	PWT
$E_{0,i}$	initial emissions	regional	GCP
$N_{t,i}$	population levels	regional	RICE2010
$A_{0,i}$	initial TFPs	regional	Calculated
$\sigma_{0,i}$	initial carbon intensities	regional	Calculated
δ_i^σ	decline rates carbon intensities	regional	RICE2010
p_i^{back}	backstop prices	regional	RICE2010
κ_i	capital shares	regional	PWT ($0.3 \leq \kappa_i \leq 0.4$)
β_i	discount factors	0.986^{10}	ACE
g_i^K	capital growth rates	regional	PWT
δ_i^K	capital depreciation rates	regional	PWT
g_i^A	TFP growth rates	regional	RICE/PWT
$\xi_{0,i}$	climate damages	regional	Calibrated to RICE
a_i	OA damages	0.25%	Colt and Knapp (2016)
d_{AA}^g	SG damages A	0.1%	Literature proxy
d_{BB}^g	SG damages B	0.05%	$0.5 \cdot$ Literature proxy
ϵ_A^g	SG costs to A	$6.25 \cdot 0.0017\%$	GDP weighted lit. proxy
ϵ_B^g	SG costs to B	$6.25 \cdot 0.0017\%$	GDP weighted lit. proxy
d_{BA}^g	SG damages B to A	0.1%	Literature proxy
d_{AB}^g	SG damages A to B	0.05%	$0.5 \cdot$ Literature proxy
d_{ARow}^g	SG damages A to Row	0.1%	Literature proxy
d_{BRow}^g	SG damages B to Row	0.1%	Literature proxy
d_{AA}^c	CM effectiveness A	0.05%	Assumption
d_{BB}^c	CM effectiveness B	0.025%	Assumption
d_{AB}^c	CM effectiveness A to B	0.025%	Assumption
d_{BA}^c	CM effectiveness B to A	0.05%	Assumption
d_{ARow}^c	CM effectiveness A to Row	0.05%	Assumption
d_{BRow}^c	CM effectiveness B to Row	0.05%	Assumption
ϵ_A^c	CM costs A	$6.25 \cdot 0.0017\%$	Assumption
ϵ_B^c	CM costs B	$6.25 \cdot 0.0017\%$	Assumption
CF_i	regional care factors	regional	Calculated

Parameter	Value	Source
f_0	0.254	Global Solace
f_1	1.16	Global Solace
f_2	0.014	Global Solace
f_3	1.16	Global Solace
n	0.69	Global Solace
α_A	0.9	Assumption
α_B	0.9	Assumption
σ_{forc}	0.5198	Regional Temp cal
s	3	Regional Temp cal
η	3.8	Regional Temp cal
$M_{1,\text{pre}}$	600	ACE
$M_{1,0}$	862.86	Updated DICE2013
$M_{2,0}$	1541.11	Updated DICE2013
$M_{3,0}$	10010.44	Updated DICE2013
$\tau_{1,0}^A$	1.2567	Regionl Temp cal
$\tau_{1,0}^B$	1.2567	Regionl Temp cal
$\tau_{2,0}$	1.1855	Regionl Temp cal
σ_A^A	0.11445	Regionl Temp cal
σ_B^B	0.11445	Regionl Temp cal
σ_B^A	0.05225	Regionl Temp cal
σ_A^B	0.05225	Regionl Temp cal
σ_O^A	0.3135	Regionl Temp cal
σ_O^B	0.3135	Regionl Temp cal
σ_O^O	0.9771	Regionl Temp cal
σ_A^O	0.01145	Regionl Temp cal
σ_B^O	0.01145	Regionl Temp cal
ϕ_{11}	0.824000	DICE2013
ϕ_{21}	0.076657	DICE2013
ϕ_{12}	0.176000	DICE2013
ϕ_{22}	0.918342	DICE2013
ϕ_{23}	0.005000	DICE2013
ϕ_{32}	0.000675	DICE2013
ϕ_{33}	0.999325	DICE2013

C Global model

C.1 Solving the Bellman equation

Definitions. We note that aggregate capital $K_t = \sum_{i=1}^{I_K} K_{i,t}$ and the share of capital in industry i is $\mathcal{K}_{i,t} = \frac{K_{i,t}}{K_t}$.

We define the consumption rate as

$$x_t = \frac{C_t}{Y_t \left[1 - D_t\left(T_{1,t}, G_t(S_t), m_t\right)\right]}. \tag{37}$$

Homogeneity of the production function implies

$$Y_t = \mathcal{F}(\boldsymbol{A}_t,\, \boldsymbol{K}_t,\, \boldsymbol{N}_t,\, \boldsymbol{E}_t) = K_t^{\kappa} \mathcal{F}(\boldsymbol{A}_t,\, \boldsymbol{\mathcal{K}}_t,\, \boldsymbol{N}_t,\, \boldsymbol{E}_t), \tag{38}$$

such that

$$\log C_t = \log x_t + \kappa \log K_t + \log \mathcal{F}(\boldsymbol{A}_t,\, \boldsymbol{\mathcal{K}}_t,\, \boldsymbol{N}_t,\, \boldsymbol{E}_t) + \xi_0\, (1 - \tau_{1,t}) - d\, S_t - a(m_t - 1). \tag{39}$$

We transform the optimization problem into its dynamic programming form (Bellman equation)

$$V(k_t, \boldsymbol{\tau}_t, \boldsymbol{M}_t, \boldsymbol{R}_t, t) = \max_{x_t, \boldsymbol{N}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{E}_t, S_t} \Bigg\{ \log x_t + \kappa\, k_t + \log \mathcal{F}(\boldsymbol{A}_t,\, \boldsymbol{\mathcal{K}}_t,\, \boldsymbol{N}_t,\, \boldsymbol{E}_t) + \xi_0\, (1 - \tau_{1,t}) - d\, S_t - a(m_t - 1) + \beta\, V(k_{t+1}, \boldsymbol{\tau}_{t+1}, \boldsymbol{M}_{t+1}, \boldsymbol{R}_{t+1}, t+1) \Bigg\} \tag{40}$$

where $k_t = \log K_t$ with the equation of motion

$$k_{t+1} = \kappa\, k_t + \log \mathcal{F}(\boldsymbol{A}_t,\, \boldsymbol{\mathcal{K}}_t,\, \boldsymbol{N}_t,\, \boldsymbol{E}_t) + \log(1 - x_t) + \xi_0\, (1 - \tau_{1,t}) - d\, S_t - a(m_t - 1). \tag{41}$$

We take a linear affine guess for the value function

$$V(k_t, \boldsymbol{\tau}_t, \boldsymbol{M}_t, \boldsymbol{R}_t, t) = \varphi_k\, k_t + \boldsymbol{\varphi}_{\tau}^T\, \boldsymbol{\tau}_t + \boldsymbol{\varphi}_M^T\, \boldsymbol{M}_t + \boldsymbol{\varphi}_{R,t}^T\, \boldsymbol{R}_t + \varphi_t. \tag{42}$$

We show that the system is linear in states and that the affine value function, (42), solves the system. Inserting the trial solution and the next period's states

(equations (22), (27), (28b) and (41)) into the Bellman equation delivers

$$
\begin{aligned}
&\varphi_k\, k_t + \boldsymbol{\varphi}_\tau^T\, \boldsymbol{\tau}_t + \boldsymbol{\varphi}_M^T\, \boldsymbol{M}_t + \boldsymbol{\varphi}_{R,t}^T\, \boldsymbol{R}_t + \varphi_t \\
&= \max_{x_t, \boldsymbol{N}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{E}_t, S_t} \Bigg\{ \log x_t + \kappa\, k_t + \log \mathcal{F}(\boldsymbol{A}_t,\, \boldsymbol{\mathcal{K}}_t,\, \boldsymbol{N}_t,\, \boldsymbol{E}_t) + \xi_0\, (1 - \tau_{1,t}) - d\, S_t - a(m_t - 1) \\
&+ \lambda_t^{\mathcal{K}} (1 - \sum_{i=1}^{I_K} \mathcal{K}_{i,t}) + \lambda_t^N (1 - \sum_{i=1}^{I_N} N_{i,t}) + \beta\, \varphi_k \left(\kappa\, k_t + \log \mathcal{F}(\boldsymbol{A}_t,\, \boldsymbol{\mathcal{K}}_t,\, \boldsymbol{N}_t,\, \boldsymbol{E}_t) + \log(1 - x_t) \right. \\
&\qquad \left. + \xi_0\, (1 - \tau_{1,t}) - d\, S_t - a(m_t - 1) \right) + \beta\, \boldsymbol{\varphi}_\tau^T \left(\boldsymbol{\sigma} \boldsymbol{\tau}_t + \tilde{\boldsymbol{F}}_t \right) + \beta\, \boldsymbol{\varphi}_M^T\, (\boldsymbol{\Phi} \boldsymbol{M}_t + \tilde{\boldsymbol{e}}_{\boldsymbol{t}}) \\
&\qquad\qquad + \beta\, \boldsymbol{\varphi}_{R,t+1}^T\, \left(\boldsymbol{R}_t - \boldsymbol{E}_t^d \right) + \beta\, \varphi_{t+1} \Bigg\}. \quad (43)
\end{aligned}
$$

First order conditions. Maximizing the right hand side over x_t yields

$$
\frac{1}{x_t} - \beta\, \varphi_k \frac{1}{1 - x_t} = 0 \quad \Longrightarrow \quad x_t = \frac{1}{1 + \beta\, \varphi_k}. \quad (44)
$$

Maximizing the right hand side over $\mathcal{K}_{i,t}$ yields

$$
(1 + \beta\, \varphi_k) \frac{\frac{\partial \mathcal{F}(\boldsymbol{A}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{N}_t, \boldsymbol{E}_t)}{\partial \mathcal{K}_{i,t}}}{\mathcal{F}(\boldsymbol{A}_t,\, \boldsymbol{\mathcal{K}}_t,\, \boldsymbol{N}_t,\, \boldsymbol{E}_t)} = \lambda_t^{\mathcal{K}}
$$

which is equivalent to

$$
\mathcal{K}_{i,t} = \frac{\sigma_{Y,\mathcal{K}_i}(\boldsymbol{A}_t,\, \boldsymbol{\mathcal{K}}_t,\, \boldsymbol{N}_t,\, \boldsymbol{E}_t)}{\sum_{i=1}^{I_K} \sigma_{Y,\mathcal{K}_i}(\boldsymbol{A}_t,\, \boldsymbol{\mathcal{K}}_t,\, \boldsymbol{N}_t,\, \boldsymbol{E}_t)} \quad (45)
$$

with

$$
\sigma_{Y,\mathcal{K}_i}(\boldsymbol{A}_t,\, \boldsymbol{\mathcal{K}}_t,\, \boldsymbol{N}_t,\, \boldsymbol{E}_t) \equiv \frac{\partial \mathcal{F}(\boldsymbol{A}_t,\, \boldsymbol{\mathcal{K}}_t,\, \boldsymbol{N}_t,\, \boldsymbol{E}_t)}{\partial \mathcal{K}_{i,t}} \frac{\mathcal{K}_{i,t}}{\mathcal{F}(\boldsymbol{A}_t,\, \boldsymbol{\mathcal{K}}_t,\, \boldsymbol{N}_t,\, \boldsymbol{E}_t)}.
$$

Similarly, the first order conditions for the labor input is

$$
(1 + \beta\, \varphi_k) \frac{\frac{\partial \mathcal{F}(\boldsymbol{A}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{N}_t, \boldsymbol{E}_t)}{\partial N_{i,t}}}{\mathcal{F}(\boldsymbol{A}_t,\, \boldsymbol{\mathcal{K}}_t,\, \boldsymbol{N}_t,\, \boldsymbol{E}_t)} = \lambda_t^N
$$

and hence

$$N_{i,t} = \frac{\sigma_{Y,N_i}(\boldsymbol{A}_t,\, \boldsymbol{\mathcal{K}}_t,\, \boldsymbol{N}_t,\, \boldsymbol{E}_t)}{\sum_{i=1}^{I_N} \sigma_{Y,N_i}(\boldsymbol{A}_t,\, \boldsymbol{\mathcal{K}}_t,\, \boldsymbol{N}_t,\, \boldsymbol{E}_t)} \tag{46}$$

with

$$\sigma_{Y,N_i}(\boldsymbol{A}_t,\, \boldsymbol{\mathcal{K}}_t,\, \boldsymbol{N}_t,\, \boldsymbol{E}_t) \equiv \frac{\partial \mathcal{F}(\boldsymbol{A}_t,\, \boldsymbol{\mathcal{K}}_t,\, \boldsymbol{N}_t,\, \boldsymbol{E}_t)}{\partial N_{i,t}} \frac{N_{i,t}}{\mathcal{F}(\boldsymbol{A}_t,\, \boldsymbol{\mathcal{K}}_t,\, \boldsymbol{N}_t,\, \boldsymbol{E}_t)}$$

The first order condition for the optimal input of fossil fuels is given by

$$(1+\beta\,\varphi_k)\frac{\frac{\partial \mathcal{F}(\boldsymbol{A}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{N}_t, \boldsymbol{E}_t)}{\partial E_{i,t}}}{\mathcal{F}(\boldsymbol{A}_t,\, \boldsymbol{\mathcal{K}}_t,\, \boldsymbol{N}_t,\, \boldsymbol{E}_t)} = \beta(\varphi_{R,i,t+1} - \varphi_{M1})$$

which is equivalent to

$$E_{i,t} = \frac{(1+\beta\,\varphi_k)\sigma_{Y,E_i}(\boldsymbol{A}_t,\, \boldsymbol{\mathcal{K}}_t,\, \boldsymbol{N}_t,\, \boldsymbol{E}_t)}{\beta(\varphi_{R,i,t+1} - \varphi_{M1})} \tag{47}$$

with

$$\sigma_{Y,E_i}(\boldsymbol{A}_t,\, \boldsymbol{\mathcal{K}}_t,\, \boldsymbol{N}_t,\, \boldsymbol{E}_t) \equiv \frac{\partial \mathcal{F}(\boldsymbol{A}_t,\, \boldsymbol{\mathcal{K}}_t,\, \boldsymbol{N}_t,\, \boldsymbol{E}_t)}{\partial E_{i,t}} \frac{E_{i,t}}{\mathcal{F}(\boldsymbol{A}_t,\, \boldsymbol{\mathcal{K}}_t,\, \boldsymbol{N}_t,\, \boldsymbol{E}_t)}.$$

So far, our results are equivalent to those of the ACE model. Next, we spell out the part of the Bellman equation that depends on sulfur

$$B_t = \beta\,\varphi_{\tau 1}\,\sigma_{\text{forc}}(f_2\,S_t - f_3\,m_t^n\,S_t^{1-n}) - (1+\beta\,\varphi_k)\,d\,S_t$$

and find the first order condition for optimal sulfur deployment

$$\beta\,\varphi_{\tau 1}\,\sigma_{\text{forc}}(f_2 + (n-1)f_3\,m_t^n\,S_t^{-n}) - (1+\beta\,\varphi_k)\,d = 0.$$

Solving the first order condition for S_t gives the optimal level of sulfur deployment

$$S_t = \underbrace{\left(\frac{\beta\,\varphi_{\tau 1}\,\sigma_{\text{forc}}\,(n-1)f_3}{(1+\beta\,\varphi_k)\,d - \beta\,\varphi_{\tau 1}\,\sigma_{\text{forc}}\,f_2}\right)^{\frac{1}{n}}}_{\equiv z} m_t. \tag{48}$$

Solving the system of first order conditions gives us $\boldsymbol{N}_t^*(\boldsymbol{A}_t, \varphi_k, \boldsymbol{\varphi}_M, \boldsymbol{\varphi}_{R,t+1})$,

$\boldsymbol{\mathcal{K}}_t^*(\boldsymbol{A}_t, \varphi_k, \boldsymbol{\varphi}_M, \boldsymbol{\varphi}_{R,t+1})$, and $\boldsymbol{E}_t^*(\boldsymbol{A}_t, \varphi_k, \boldsymbol{\varphi}_M, \boldsymbol{\varphi}_{R,t+1})$ which are independent of the states and $S_t^*(\varphi_k, \varphi_{\tau 1}, M_{1,t})$ which depends on the atmospheric carbon stock. In the following we show that given these optimal controls the maximized Bellman equation is linear in all states.

Shadow values. Inserting the optimal control rules into the maximized Bellman equation gives us

$$\begin{aligned}
&\varphi_k\, k_t + \boldsymbol{\varphi}_\tau^T\, \boldsymbol{\tau}_t + \boldsymbol{\varphi}_M^T\, \boldsymbol{M}_t + \boldsymbol{\varphi}_{R,t}^T\, \boldsymbol{R}_t + \varphi_t \\
&\quad = \log x_t^* + \kappa\, k_t + \log \mathcal{F}(\boldsymbol{A}_t,\, \boldsymbol{\mathcal{K}}_t^*,\, \boldsymbol{N}_t^*,\, \boldsymbol{E}_t^*) + \xi_0\, (1 - \tau_{1,t}) - d\, S_t^* - a(m_t - 1) \\
&+ \beta\, \varphi_k \left(\kappa\, k_t + \log \mathcal{F}(\boldsymbol{A}_t,\, \boldsymbol{\mathcal{K}}_t^*,\, \boldsymbol{N}_t^*,\, \boldsymbol{E}_t^*) + \log(1 - x_t^*) + \xi_0\, (1 - \tau_{1,t}) - d\, S_t^* - a(m_t - 1)\right) \\
&\quad + \beta\, \boldsymbol{\varphi}_\tau^T \left(\boldsymbol{\sigma}\boldsymbol{\tau}_t + \tilde{\boldsymbol{F}}_t\right) + \beta\, \boldsymbol{\varphi}_M^T\, (\boldsymbol{\Phi}\boldsymbol{M}_t + \tilde{\boldsymbol{e}}_{\boldsymbol{t}}) + \beta\, \boldsymbol{\varphi}_{R,t+1}^T \left(\boldsymbol{R}_t - \boldsymbol{E}_t^{d*}\right) + \beta\, \varphi_{t+1} \quad (49)
\end{aligned}$$

Arranging terms with respect to their states and using the propensity definition z yields

$$\begin{aligned}
&\varphi_k\, k_t + \boldsymbol{\varphi}_\tau^T\, \boldsymbol{\tau}_t + \boldsymbol{\varphi}_M^T\, \boldsymbol{M}_t + \boldsymbol{\varphi}_{R,t}^T\, \boldsymbol{R}_t + \varphi_t = \Big[(1 + \beta\, \varphi_k)\kappa\Big] k_t + \Big[\beta\, \boldsymbol{\varphi}_\tau^T\, \boldsymbol{\sigma} - (1 + \beta\, \varphi_k)\xi_0\, \boldsymbol{e}_1^T\Big] \boldsymbol{\tau}_t \\
&+ \Big[\beta\, \boldsymbol{\Phi} \boldsymbol{\varphi}_M^T + \Big((\beta\, \varphi_{\tau 1}\, \sigma_{\text{forc}}) \left(f_1 + f_2\, z - f_3\, z^{1-n}\right) - (1 + \beta\, \varphi_k)(a + d\, z)\Big) M_{pre}^{-1}\, \boldsymbol{e}_1^T\Big] \boldsymbol{M}_t \\
&+ \Big[\beta\, \boldsymbol{\varphi}_{R,t+1}^T\Big] R_t + \log\, x_t^* + \beta\, \varphi_k\, \log(1 - x_t^*) + (1 + \beta\, \varphi_k) \log \mathcal{F}(\boldsymbol{A}_t,\, \boldsymbol{\mathcal{K}}_t^*,\, \boldsymbol{N}_t^*,\, \boldsymbol{E}_t^*) \\
&\quad + (1 + \beta\, \varphi_k)(\xi_0 + a) + \beta\, \varphi_{\tau 1}\, \sigma_{\text{forc}}\, f_0 + \beta\, \boldsymbol{\varphi}_M^T\, \tilde{\boldsymbol{e}}_t - \beta\, \boldsymbol{\varphi}_{R,t+1}^T\, \boldsymbol{E}_t^{d*} + \beta\, \varphi_{t+1}. \quad (50)
\end{aligned}$$

Hence, the system is linear in all states. Deriving both sides of the equation with respect to capital, k_t, yields

$$\varphi_k = (1 + \beta\, \varphi_k)\kappa \quad \Leftrightarrow \quad \varphi_k = \frac{\kappa}{1 - \beta\, \kappa} \quad (51)$$

Inserting φ_k into equation (44) yield the optimal consumption rate $x_t^* = 1 - \beta\, \kappa$.

Coefficient matching with respect to transformed temperatures delivers

$$\boldsymbol{\varphi}_\tau^T = -\xi_0\, (1 + \beta\, \varphi_k)\, \boldsymbol{e}_1^T\, (\mathbf{1} - \beta\, \boldsymbol{\sigma})^{-1}$$

where we denote the entries of the inverted matrix as follows

$$\begin{pmatrix} \tilde{\sigma}_{11} & \tilde{\sigma}_{12} \\ \tilde{\sigma}_{21} & \tilde{\sigma}_{22} \end{pmatrix} \equiv (\mathbf{1} - \beta\, \boldsymbol{\sigma})^{-1}$$

yielding

$$\tilde{\sigma}_{11} = \left[(\mathbf{1} - \beta\,\boldsymbol{\sigma})^{-1}\right]_{1,1} \tag{52}$$

where $[\cdot]_{1,1}$ denotes the first element of the inverted matrix in square brackets.

Hence,

$$\varphi_{\tau 1} = -\xi_0\,(1 + \beta\,\varphi_k)\tilde{\sigma}_{11}. \tag{53}$$

Coefficient matching with respect to the atmospheric carbon stock leads to

$$\varphi_M^T = \Big((\beta\,\varphi_{\tau 1}\,\sigma_{\text{forc}})\left(f_1 + f_2\,z - f_3\,z^{1-n}\right) - (1 + \beta\,\varphi_k)(a + d\,z)\Big) M_{pre}^{-1}\,\boldsymbol{e}_1^T\,(\mathbf{1} - \beta\,\boldsymbol{\Phi})^{-1}. \tag{54}$$

We define

$$\tilde{\phi}_{ij} = \left[(\mathbf{1} - \beta\,\boldsymbol{\Phi})^{-1}\right]_{ij} \text{ for } i, j \in \{1, 2, 3\},$$

yielding

$$\tilde{\phi}_{11} = \left[(\mathbf{1} - \beta\,\boldsymbol{\Phi})^{-1}\right]_{1,1} \tag{55}$$

where $[\cdot]_{1,1}$ denotes the first element of the inverted matrix in square brackets. Note that for ease of representation we drop the subscript of the term $\tilde{\phi}_{11}$ and instead use $\tilde{\phi}$.

Coefficient matching with respect to the resource stock yields

$$\boldsymbol{\varphi}_{R,t}^T = \beta\,\boldsymbol{\varphi}_{R,t+1}^T \quad \Leftrightarrow \quad \boldsymbol{\varphi}_{R,t} = \beta^{-t}\boldsymbol{\varphi}_{R,0} \quad \text{(Hotelling's rule)}.$$

The initial resource values $\boldsymbol{\varphi}_{R,0}^T$ depend on the set up of the economy, including assumptions about production and the energy sector. Given the coefficients and the optimal rate of consumption equation (50) turns to the following condition:

$$\begin{aligned}\varphi_t - \beta\,\varphi_{t+1} = \log\,x_t^* + \beta\,\varphi_k\,\log(1 - x_t^*) + (1 + \beta\,\varphi_k)\log\mathcal{F}(\boldsymbol{A}_t,\,\boldsymbol{\mathcal{K}}_t^*,\,\boldsymbol{N}_t^*,\,\boldsymbol{E}_t^*) \\ + (1 + \beta\,\varphi_k)(\xi_0 + a) + \beta\,\varphi_{\tau 1}\,\sigma_{\text{forc}}\,f_0 + \beta\,\boldsymbol{\varphi}_M^T\,\tilde{\boldsymbol{e}}_t - \beta\,\boldsymbol{\varphi}_{R,t+1}^T\,\boldsymbol{E}_t^{d*}\end{aligned} \tag{56}$$

This condition will be satisfied by picking the sequence $\varphi_0, \varphi_1, \varphi_2, ...$ The additional condition $\lim_{t\to\infty}\beta^t V(\cdot) = 0 \Rightarrow \lim_{t\to\infty}\beta^t\varphi_t = 0$ pins down this initial value φ_0.

C.2 Proof of Proposition 1 (optimal level of sulfur)

In Appendix C.1 we have shown that the optimal level of sulfur is given by

$$S_t^* = \left(\frac{(n-1)\beta\,\varphi_{\tau 1}\,\sigma_{\text{forc}}\,f_3}{(1+\beta\,\varphi_k)\,d - \beta\,\varphi_{\tau 1}\,\sigma_{\text{forc}}\,f_2} \right)^{\frac{1}{n}} m_t. \tag{57}$$

The endogenous shadow value of capital $\varphi_k > 0$ is positive (see (51)), while the endogenous shadow value of (transformed) temperature is negative $\varphi_{\tau 1} < 0$ (a bad, see (53)). Therefore, both numerator and denominator are positive. The optimal level of sulfur deployment increases in the absolute value of the shadow price of atmospheric temperature. Note that for ease of representation we use the following definition in the main part of the paper, $\tilde{\sigma} = \tilde{\sigma}_{11}\,\sigma_{\text{forc}}$.

Inserting (51) and (53) for the shadow values $\varphi_{\tau 1}$ and φ_k into our expression for sulfur deployment (57) and using the definition $\gamma = \beta\,\xi_0\,\tilde{\sigma}$ delivers

$$S_t^* = \left(\frac{(1-n)\,\gamma\,f_3}{d + \gamma\,f_2} \right)^{\frac{1}{n}} m_t.$$

The parameter $\tilde{\sigma}_{11}$ is defined as in (52).

Table 10: Parameter values from ACE re-calibrated for 2 temperature layers

Y_t^{net}	M_{pre}	β	ξ_0	$\tilde{\sigma}$	$\tilde{\phi}$
135×10^{13}	$3.667 \times 600 \times 10^9$	0.986^{10}	0.021	0.63	4.26

Inserting the fit parameters from Table 2, and using the parameter values from the baseline calibration of the ACE model (re-calibrated for 2 temperature layers) from Table 10 leads to

$$S_t^* = \left(\frac{1.65}{16\% + 10^3 d} \right)^{1.45} m_t.$$

C.3 Proof of Proposition 2 (SCC)

Inserting equation (53) for the shadow value $\varphi_{\tau 1}$ and $\gamma = \beta\,\xi_0\,\tilde{\sigma}$ into equation (54) for the shadow value of the atmospheric carbon stock delivers

$$\varphi_{M1} = -(1+\beta\,\varphi_k)\Big(\gamma\left(f_1 + f_2\,z - f_3\,z^{1-n}\right) + a + d\,z\Big) M_{pre}^{-1}\,\tilde{\phi}.$$

Inserting equation (51) for the shadow value φ_k leads to

$$\varphi_{M1} = -\frac{1}{1-\beta\,\kappa}\Big(\gamma\left(f_1 + f_2\,z - f_3\,z^{1-n}\right) + a + d\,z\Big)M_{pre}^{-1}\,\tilde{\phi}.$$

The SCC is the negative of the shadow value of atmospheric carbon stock expressed in money-measured consumption units.

$$\begin{aligned} SCC &= -(1-\beta\,\kappa)Y_t^{net}\,\varphi_{M1} \\ &= \frac{Y_t^{net}}{M_{pre}}\left[a + f_1\,\gamma - \left(\frac{f_3}{z^n} - f_2\right)\gamma z + dz\right]\tilde{\phi}. \end{aligned}$$

Using parameter values from the ACE model (see Table 10) yields $\gamma = 0.0108$. Thus, the SCC in (USD-2019-) money-measured consumption equivalents is

$$SCC = 613\left[a + 2\% - 6.65\left(\frac{1.6}{(4\% + 10^3 d)^{0.13}}\right)^{1.15}\right]4.3.$$

D Regional model

D.1 Bellman equation and Markov strategies

We show the existence of linear Markov strategies $S_t^i(m_t) = s_t^i m_t$ for $i \in \{A, B\}$ forming a subgame perfect Nash equilibrium of the dynamic game.

Bellman equation. In the following we show for region A that the system is linear in states and that the affine value function

$$V(k_t, \boldsymbol{\tau}_t, \boldsymbol{M}_t, \boldsymbol{R}_t, t) = \varphi_k^A\,k_t + \boldsymbol{\varphi}_{\tau A}^T\,\boldsymbol{\tau}_t + \boldsymbol{\varphi}_{MA}^T\,\boldsymbol{M}_t + \boldsymbol{\varphi}_{R,t}^T\,\boldsymbol{R}_t + \varphi_t, \tag{58}$$

with

$$\boldsymbol{\varphi}_{\tau A}^T = \begin{pmatrix} \varphi_{\tau 1}^{AA} \\ \varphi_{\tau 1}^{BA} \\ \varphi_{\tau 2}^{A} \end{pmatrix} \qquad \text{and} \qquad \boldsymbol{\varphi}_{MA}^T = \begin{pmatrix} \varphi_{M1}^{A} \\ \varphi_{M2}^{A} \\ \varphi_{M3}^{A} \end{pmatrix},$$

solves the system. The proof for region B is analogous. We suppress regional indices when there is no ambiguity to ease notation. We denote region A's shadow price for temperature $\tau_{1,t}^A$ in region A by $\varphi_{\tau 1}^{AA}$, and region A's shadow price for temperature $\tau_{1,t}^B$ in region B by $\varphi_{\tau 1}^{BA}$.

Inserting the trial solution and the equations of motion for the next period's

states into the Bellman equation delivers

$$\begin{aligned}
&\varphi_k^A\, k_t + \boldsymbol{\varphi}_{\tau A}^T\, \boldsymbol{\tau}_t + \boldsymbol{\varphi}_{MA}^T\, \boldsymbol{M}_t + \boldsymbol{\varphi}_{R,t}^T\, \boldsymbol{R}_t + \varphi_t \\
&= \max_{x_t, \boldsymbol{N}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{E}_t, S_t^A} \Bigg\{ \log x_t + \kappa\, k_t + \log \mathcal{F}(\boldsymbol{A}_t,\, \boldsymbol{\mathcal{K}}_t,\, \boldsymbol{N}_t,\, \boldsymbol{E}_t) + \xi_0^A\, \left(1 - \tau_{1,t}^A\right) - d_{AA} S_t^A \\
&+ d_{BA} \alpha_B\, S_t^B(m_t) - a^A\, (m_t - 1) + \lambda_t^{\mathcal{K}} (1 - \sum_{i=1}^{I_K} \mathcal{K}_{i,t}) + \lambda_t^N (1 - \sum_{i=1}^{I_N} N_{i,t}) + \beta^A\, \varphi_k^A \Big(\kappa\, k_t \\
&+ \log \mathcal{F}(\boldsymbol{A}_t,\, \boldsymbol{\mathcal{K}}_t,\, \boldsymbol{N}_t,\, \boldsymbol{E}_t) + \log(1 - x_t) + \xi_0^A\, \left(1 - \tau_{1,t}^A\right) - d_{AA} S_t^A + d_{BA} \alpha_B\, S_t^B(m_t) - a^A\, (m_t - 1) \Big) \\
&+ \beta^A\, \boldsymbol{\varphi}_{\tau A}^T \Big(\boldsymbol{\sigma}^A \boldsymbol{\tau}_t + \tilde{\boldsymbol{F}}_t(S_t^A, S_t^B(m_t)) \Big) + \beta^A\, \boldsymbol{\varphi}_{MA}^T\, (\boldsymbol{\Phi} \boldsymbol{M}_t + \tilde{\boldsymbol{e}}_{\boldsymbol{t}}) + \beta^A\, \boldsymbol{\varphi}_{R,t+1}^T\, \left(\boldsymbol{R}_t - \boldsymbol{E}_t^d \right) \\
&\qquad + \beta^A\, \varphi_{t+1} \Bigg\}. \quad (59)
\end{aligned}$$

First order conditions (apart from SG). The first order conditions for $x_t, \boldsymbol{N}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{E}_t$ are structurally the same as in the global model.

Optimal response functions: collecting terms. The optimal SG deployment has to be compatible with the assumed strategies of both regions. Region A takes region B's strategy as given while maximizing its welfare over its own sulfur deployment (or, for $S_t^A < 0$ CM). The part of the r.h.s. Bellman equation depending on sulfur is

$$\begin{aligned}
B_{nc}^A(m_t, S_t^A) \equiv\; & \beta^A\, \varphi_{\tau 1}^{AA}\, \sigma_{\text{forc}} F_t^{CO2}(m_t, S_t^A + \alpha_B\, S_t^B(m_t)) \\
& + \beta^A\, \varphi_{\tau A}^{BA}\, \sigma_{\text{forc}} F_t^{CO2}(m_t, \alpha_A S_t^A + S_t^B(m_t)) \\
& - (1 + \beta^A\, \varphi_k^A)[d_{AA} S_t^A + d_{BA}\, \alpha_B\, S_t^B(m_t)] \\
=\; & \underbrace{\left[\beta^A\, \varphi_{\tau 1}^{AA}\, \sigma_{\text{forc}}\, f_2 + \beta^A\, \varphi_{\tau 1}^{BA}\, \sigma_{\text{forc}}\, f_2\, \alpha_A - (1 + \beta^A\, \varphi_k^A) d_{AA}\right]}_{\equiv\, -\tilde{\delta}_A} S_t^A \\
& + \left[\beta^A\, \varphi_{\tau 1}^{AA}\, \sigma_{\text{forc}}\, f_2\, \alpha_B + \beta^A\, \varphi_{\tau 1}^{BA}\, \sigma_{\text{forc}}\, f_2 - (1 + \beta^A\, \varphi_k^A) \alpha_B d_{BA}\right] S_t^B(m_t) \\
& - \beta^A\, \varphi_{\tau 1}^{AA}\, \sigma_{\text{forc}}\, f_3\, m_t^n (S_t^A + \alpha_B\, S_t^B(m_t))^{1-n} \\
& - \beta^A\, \varphi_{\tau 1}^{BA}\, \sigma_{\text{forc}}\, f_3\, m_t^n (S_t^B(m_t) + \alpha_A\, S_t^A)^{1-n}. \quad (60)
\end{aligned}$$

We note that only the term δ_A depends on the damage term d_{AA}, which discretely switches sign and magnitude as the country changes action between CM, no action, and SG at $S_t^A = 0$. All other terms are continuous. Given d_{AA} multiplies S_t^A, also the term $\delta_A S_t^A$ remains continuous. The shadow prices of a temperature increase $\varphi_{\tau 1}^i < 0$, $i \in \{AA, BA\}$ are negative, therefore, $\delta_A > 0$.

First order condition. Deriving (60) with respect to S_t^A delivers the equation

$$\frac{\partial B_{nc}^A(m_t, S_t^A)}{\partial S_t^A} = \underbrace{(n-1)\,\beta^A\,\varphi_{\tau 1}^{AA}\,\sigma_{\text{forc}}\,f_3}_{\equiv \tilde{a}_A} m_t^n (S_t^A + \alpha_B\, S_t^B(m_t))^{-n} + \underbrace{(n-1)\beta^A\,\varphi_{\tau 1}^{BA}\,\sigma_{\text{forc}}\;f_3\alpha_A}_{\equiv \tilde{b}_A} m_t^n (S_t^B(m_t) + \alpha_A\, S_t^A)^{-n} - \tilde{\delta}_A.$$

We note that $a_A, b_A > 0$ because $n < 1$ and the shadow prices of a temperature increase are negative. Moreover recall that region A takes region B's strategy $S_t^B(m_t) = s_t^B m_t$ as given. Defining $s^A \equiv \frac{S_t^A}{m_t}$ we rewrite the derivative as

$$\begin{aligned}\frac{\partial B_{nc}^A(m_t, S_t^A)}{\partial S_t^A} &= \tilde{a}_A m_t^n (S_t^A + \alpha_B s_t^B m_t)^{-n} + \tilde{b}_A m_t^n (m_t s_t^B + \alpha_A S_t^A)^{-n} - \tilde{\delta}_A \\ &= \tilde{a}_A (s^A + \alpha_B s_t^B)^{-n} + \tilde{b}_A (s_t^B + \alpha_A s^A)^{-n} - \tilde{\delta}_A. \end{aligned} \tag{61}$$

Strict concavity. The second order derivative in S_t^A is strictly negative so that the function $B_{nc}^A(m_t, S_t^A)$ is strictly concave at all points of continuity. We still have to check the discontinuity at $S_t^A = 0$ ($\Leftrightarrow s^A = 0$). The left and right limits of the objective function's slope at $s^A = 0$ are

$$\lim_{s^a \to -0} \frac{\partial B_{nc}^A(m_t, S_t^A)}{\partial S_t^A} = (\tilde{a}_A \alpha_B^{-n} + \tilde{b}_A) s_t^{B^{-n}} - \tilde{\delta}_A^c \text{ and}$$

$$\lim_{s^a \to +0} \frac{\partial B_{nc}^A(m_t, S_t^A)}{\partial S_t^A} = (\tilde{a}_A \alpha_B^{-n} + \tilde{b}_A) s_t^{B^{-n}} - \tilde{\delta}_A^g$$

where $\tilde{\delta}_A \in \{\tilde{\delta}_A^g, \tilde{\delta}_A^c\}$ was defined in equation (60) and depends on damages d_{AA}. The superindex on $\tilde{\delta}_A^c$ refers to the case of CM where $S_t^A < 0$ and $d_{AA} \equiv d_{AA}^c - \epsilon_A^c$ and $\tilde{\delta}_A^g$ refers to the case of (sulfur-based) SG where $S_t^A > 0$ and $d_{AA} \equiv d_{AA}^g + \epsilon_A^g$. Because $d_{AA}^c \leq d_{AA}^g$ by Assumption 2 and because operational cost of SG are positive, we have $\tilde{\delta}_A^c < \tilde{\delta}_A^g$ and

$$\lim_{s^a \to -0} \frac{\partial B_{nc}^A(m_t, S_t^A)}{\partial S_t^A} > \lim_{s^a \to +0} \frac{\partial B_{nc}^A(m_t, S_t^A)}{\partial S_t^A}.$$

Therefore, the function $B_{nc}^A(m_t, S_t^A)$ has a concave kink at $S_t^A = 0$.

Three qualitatively distinct response functions $\boldsymbol{S_t^A = m_t\, s_t^A}$. We have the following cases for region A's optimal sulfur deployment:

1. if $\lim_{s^a \to -0} \frac{\partial B_{nc}^A(m_t, S_t^A)}{\partial S_t^A} = (\tilde{a}_A \alpha_B^{-n} + \tilde{b}_A) s_t^{B^{-n}} - \tilde{\delta}_A^c < 0$ then $s^A, S_t^A < 0$,

the interior optimum lies to the left of the kink, and the region engages in CM.

$$2. \quad \text{if } \lim_{s^a \to -0} \frac{\partial B^A_{nc}(m_t, S^A_t)}{\partial S^A_t} > 0 > \lim_{s^a \to +0} \frac{\partial B^A_{nc}(m_t, S^A_t)}{\partial S^A_t}, \text{ then } S^A_t = s^A = 0,$$

$B^A_{nc}(m_t, S^A_t)$ is maximal at the kink, and the region remains inactive.

$$3. \quad \text{if } \lim_{s^a \to +0} \frac{\partial B^A_{nc}(m_t, S^A_t)}{\partial S^A_t} = (\tilde{a}_A \alpha_B^{-n} + \tilde{b}_A) s_t^{B-n} - \tilde{\delta}^g_A > 0 \text{ then } s^A, S^A_t > 0,$$

an interior optimum exists to the right of the kink, and the region engages in SG.

Thus, region A's strategy is $S^A_t = s^A m_t$, consistent with our assumption that both regions follow a climate engineering strategy proportional to the CO_2 concentration m_t. We obtain the same result for region B by exchanging region labels.

Equilibrium strategies. We now solve for the proportionality constants s^i, $i \in \{A, B\}$, characterized by the optimality conditions above, such that the regions' strategies are mutually best responses.

(i) Let $S^A_t \neq 0$ and $S^B_t \neq 0$. Then we have shown that the optimal responses follow from the interior solution to the first order condition of equation (61)

$$\underbrace{(n-1)\,\beta^A\,\varphi^{AA}_{\tau 1}\,\sigma_{\text{forc}}\,f_3}_{\equiv \tilde{a}_A > 0}(s^A + \alpha_B\, s^B)^{-n} + \underbrace{(n-1)\beta^A\,\varphi^{BA}_{\tau 1}\,\sigma_{\text{forc}}\;f_3\,\alpha_A}_{\equiv \tilde{b}_A > 0}(s^B + \alpha_A\, s^A)^{-n} - \tilde{\delta}_A = 0 \tag{62}$$

$$\underbrace{(n-1)\,\beta^B\,\varphi^{BB}_{\tau 1}\,\sigma_{\text{forc}}\,f_3}_{\equiv \tilde{a}_B > 0}(s^B + \alpha_A\, s^A)^{-n} + \underbrace{(n-1)\beta^B\,\varphi^{AB}_{\tau 1}\,\sigma_{\text{forc}}\;f_3\,\alpha_B}_{\equiv \tilde{b}_B > 0}(s^A + \alpha_B\, s^B)^{-n} - \tilde{\delta}_B = 0 \tag{63}$$

Similarly as for region A, we denoted the shadow price of region B for temperature $\tau^B_{1,t}$ by $\varphi^{BB}_{\tau 1}$, and the shadow price of region B for temperature $\tau^A_{1,t}$ by $\varphi^{AB}_{\tau 1}$. Rearranging (63) gives

$$(s^A + \alpha_B\, s^B)^{-n} = \frac{\tilde{\delta}_B - \tilde{a}_B (s^B + \alpha_A\, s^A)^{-n}}{\tilde{b}_B}$$

Using this result in (62) yields

$$(s^B + \alpha_A\, s^A)^{-n} = \frac{\tilde{\delta}_A\,\tilde{b}_B - \tilde{a}_A\,\tilde{\delta}_B}{\tilde{b}_A\,\tilde{b}_B - \tilde{a}_A\,\tilde{a}_B}.$$

From this we get

$$s^B = \left(\frac{\tilde{\delta}_A \, \tilde{b}_B - \tilde{a}_A \, \tilde{\delta}_B}{\tilde{b}_A \, \tilde{b}_B - \tilde{a}_A \, \tilde{a}_B} \right)^{-\frac{1}{n}} - \alpha_A \, s^A \tag{64}$$

and

$$s^A = \left(\frac{\tilde{\delta}_B \, \tilde{b}_A - \tilde{a}_B \, \tilde{\delta}_A}{\tilde{b}_B \, \tilde{b}_A - \tilde{a}_B \, \tilde{a}_A} \right)^{-\frac{1}{n}} - \alpha_B \, s^B \tag{65}$$

and hence

$$s^B = \frac{1}{1 - \alpha_A \, \alpha_B} \left[\underbrace{\left(\frac{\tilde{\delta}_A \, \tilde{b}_B - \tilde{a}_A \, \tilde{\delta}_B}{\tilde{b}_A \, \tilde{b}_B - \tilde{a}_A \, \tilde{a}_B} \right)^{-\frac{1}{n}}}_{\equiv z_B} - \alpha_A \underbrace{\left(\frac{\tilde{\delta}_B \, \tilde{b}_A - \tilde{a}_B \, \tilde{\delta}_A}{\tilde{b}_B \, \tilde{b}_A - \tilde{a}_B \, \tilde{a}_A} \right)^{-\frac{1}{n}}}_{\equiv z_A} \right] \tag{66}$$

and analogously

$$s^A = \frac{1}{1 - \alpha_A \, \alpha_B} \left[\underbrace{\left(\frac{\tilde{\delta}_B \, \tilde{b}_A - \tilde{a}_B \, \tilde{\delta}_A}{\tilde{b}_B \, \tilde{b}_A - \tilde{a}_B \, \tilde{a}_A} \right)^{-\frac{1}{n}}}_{\equiv z_A} - \alpha_B \underbrace{\left(\frac{\tilde{\delta}_A \, \tilde{b}_B - \tilde{a}_A \, \tilde{\delta}_B}{\tilde{b}_A \, \tilde{b}_B - \tilde{a}_A \, \tilde{a}_B} \right)^{-\frac{1}{n}}}_{\equiv z_B} \right]. \tag{67}$$

We further define

$$z_A^g = \left(\frac{\tilde{\delta}_B \, \tilde{b}_A - \tilde{a}_B \, \tilde{\delta}_A^g}{\tilde{b}_B \, \tilde{b}_A - \tilde{a}_B \, \tilde{a}_A} \right)^{-\frac{1}{n}} \qquad \text{for} \qquad S_t^A > 0$$

$$z_A^c = \left(\frac{\tilde{\delta}_B \, \tilde{b}_A - \tilde{a}_B \, \tilde{\delta}_A^c}{\tilde{b}_B \, \tilde{b}_A - \tilde{a}_B \, \tilde{a}_A} \right)^{-\frac{1}{n}} \qquad \text{for} \qquad S_t^A < 0$$

This gives us

$$S_t^A = \frac{m_t}{1 - \alpha_A \, \alpha_B} \left(z_A^g - \alpha_B \, z_B \right) \qquad \text{for} \qquad S_t^A > 0 \tag{68}$$

and

$$S_t^A = \frac{m_t}{1 - \alpha_A \, \alpha_B} \left(z_A^c - \alpha_B \, z_B \right) \qquad \text{for} \qquad S_t^A < 0 \tag{69}$$

(ii) In the second case where $S_t^A > 0$ and $S_t^B = 0$, the first order condition for

region A simplifies to

$$\tilde{a}_A \, (s^A)^{-n} + \tilde{b}_A (\alpha_A \, s^A)^{-n} - \tilde{\delta}_A^g = 0 \tag{70}$$

which is equivalent to

$$s^A = \left(\frac{\tilde{\delta}_A^g}{\tilde{b}_A \, \alpha_A^{-n} + \tilde{a}_A} \right)^{-\frac{1}{n}} \tag{71}$$

and gives us

$$S_t^A = z_A^g \, m_t \tag{72}$$

with

$$z_A^g = \left(\frac{\tilde{\delta}_A^g}{\tilde{b}_A \, \alpha_A^{-n} + \tilde{a}_A} \right)^{-\frac{1}{n}}.$$

(iii) The last possible case is $S_t^A = 0$ and $S_t^B > 0$, which is symmetric to case (i).

Summary of strategies. In conclusion, the following reaction functions characterize a Nash equilibrium of the dynamic game: If $S_t^B = 0$, region A chooses $S_t^A = z_A^g m_t$ and if $S_t^B \neq 0$ region A chooses

$$S_t^A = \frac{m_t}{1 - \alpha_A \, \alpha_B} \left(z_A^g - \alpha_B \, z_B \right) \quad \text{for} \quad S_t^A > 0$$

$$S_t^A = \frac{m_t}{1 - \alpha_A \, \alpha_B} \left(z_A^c - \alpha_B \, z_B \right) \quad \text{for} \quad S_t^A < 0$$

$$S_t^A = 0 \quad \text{otherwise.}$$

Swapping country indices characterizes region B's strategies.

D.2 Verifying solution to the Bellman equation.

Inserting the optimal control rules $\boldsymbol{N}_t^*(\boldsymbol{A}_t, \varphi_k^A, \boldsymbol{\varphi}_{MA}, \boldsymbol{\varphi}_{R,t+1})$, $\boldsymbol{\mathcal{K}}_t^*(\boldsymbol{A}_t, \varphi_k^A, \boldsymbol{\varphi}_{MA}, \boldsymbol{\varphi}_{R,t+1})$, $\boldsymbol{E}_t^*(\boldsymbol{A}_t, \varphi_k^A, \boldsymbol{\varphi}_{MA}, \boldsymbol{\varphi}_{R,t+1})$, and $S_t^{A^*}(m_t)$, which are analogous to the global solutions characterized in equations (44) to (48), into the maximized Bellman equation (59)

gives us

$$\begin{aligned}
&\varphi_k^A\, k_t + \boldsymbol{\varphi}_{\tau A}^T\, \boldsymbol{\tau}_t + \boldsymbol{\varphi}_{MA}^T\, \boldsymbol{M}_t + \boldsymbol{\varphi}_{R,t}^T\, \boldsymbol{R}_t + \varphi_t \\
&= \log x_t^* + \kappa\, k_t + \log \mathcal{F}(\boldsymbol{A}_t,\, \boldsymbol{\mathcal{K}}_t^*,\, \boldsymbol{N}_t^*,\, \boldsymbol{E}_t^*) + \xi_0^A\, \left(1 - \tau_{1,t}^A\right) - d_{AA} S^{A^*}{}_t(m_t) + d_{BA}\, \alpha_B\, S_t^B(m_t) \\
&- a^A\, (m_t - 1) + \beta^A\, \varphi_k^A \Big(\kappa\, k_t + \log \mathcal{F}(\boldsymbol{A}_t,\, \boldsymbol{\mathcal{K}}_t^*,\, \boldsymbol{N}_t^*,\, \boldsymbol{E}_t^*) + \log(1 - x_t^*) + \xi_0^A\, \left(1 - \tau_{1,t}^A\right) \\
&- d_{AA} S_t^{A^*}(m_t) + d_{BA}\, \alpha_B\, S_t^B(m_t) - a^A\, (m_t - 1) \Big) + \beta^A\, \boldsymbol{\varphi}_{\tau A}^T \left(\boldsymbol{\sigma}^A \boldsymbol{\tau}_t + \tilde{\boldsymbol{F}}_t(S_t^{A^*}(m_t), S_t^B(m_t)) \right) \\
&+ \beta^A\, \boldsymbol{\varphi}_{MA}^T\, (\boldsymbol{\Phi} \boldsymbol{M}_t + \tilde{\boldsymbol{e}}_{\boldsymbol{t}}) + \beta^A\, \boldsymbol{\varphi}_{R,t+1}^T\, \left(\boldsymbol{R}_t - \boldsymbol{E}_t^{d^*} \right) + \beta^A\, \varphi_{t+1} \quad (73)
\end{aligned}$$

where $\tilde{\boldsymbol{F}}_t$ is the forcing vector defined in equation (32) making its sulfur dependencies explicit. Arranging terms by states for the different Nash equilibria yields

(i): $S_t^A \neq 0, S_t^B \neq 0$

$$\begin{aligned}
&\varphi_k^A\, k_t + \boldsymbol{\varphi}_{\tau A}^T\, \boldsymbol{\tau}_t + \boldsymbol{\varphi}_{MA}^T\, \boldsymbol{M}_t + \boldsymbol{\varphi}_{R,t}^T\, \boldsymbol{R}_t + \varphi_t = \Big[(1 + \beta^A\, \varphi_k^A) \kappa \Big] k_t \\
&+ \Big[\beta^A\, \boldsymbol{\varphi}_{\tau A}^T\, \boldsymbol{\sigma}^A - (1 + \beta^A\, \varphi_k^A) \xi_0^A\, \boldsymbol{e}_1^T \Big] \boldsymbol{\tau}_t + \Big[\beta^A\, \boldsymbol{\varphi}_{MA}^T\, \boldsymbol{\Phi} + \Big(\beta^A\, \varphi_{\tau 1}^{AA}\, \sigma_{\text{forc}} (f_1 + f_2\, z_A - f_3\, z_A^{1-n}) \\
&+ \beta^A\, \varphi_{\tau 1}^{BA}\, \sigma_{\text{forc}}\, (f_1 + f_2\, z_B - f_3\, z_B^{1-n}) - (1 + \beta^A\, \varphi_k^A) \Big(d_{AA}\, \frac{1}{1 - \alpha_A\, \alpha_B} (z_A - \alpha_B\, z_B) \\
&+ d_{BA}\, \frac{\alpha_B}{1 - \alpha_A\, \alpha_B} (z_B - \alpha_A\, z_A) + a^A \Big) \Big) M_{pre}^{-1}\, \boldsymbol{e}_1^T \Big] \boldsymbol{M}_t + \Big[\beta^A\, \boldsymbol{\varphi}_{R,t+1}^T \Big] R_t + \log\, x_t^* \\
&+ \beta^A\, \varphi_k^A\, \log(1 - x_t^*) + (1 + \beta^A\, \varphi_k^A) \log \mathcal{F}(\boldsymbol{A}_t,\, \boldsymbol{\mathcal{K}}_t^*,\, \boldsymbol{N}_t^*,\, \boldsymbol{E}_t^*) + (1 + \beta^A\, \varphi_k^A)(\xi_0^A + a^A) \\
&+ (\beta^A\, \varphi_{\tau 1}^{AA}\, \sigma_{\text{forc}} + \beta^A\, \varphi_{\tau 1}^{BA}\, \sigma_{\text{forc}})\, f_0 + \beta^A\, \boldsymbol{\varphi}_{MA}^T\, \tilde{\boldsymbol{e}}_{\boldsymbol{t}} - \beta^A\, \boldsymbol{\varphi}_{R,t+1}^T\, \boldsymbol{E}_t^{d^*} + \beta^A\, \varphi_{t+1}, \quad (74)
\end{aligned}$$

(ii): $S_t^A > 0, S_t^B = 0$

$$\begin{aligned}
&\varphi_k^A\, k_t + \boldsymbol{\varphi}_{\tau A}^T\, \boldsymbol{\tau}_t + \boldsymbol{\varphi}_{MA}^T\, \boldsymbol{M}_t + \boldsymbol{\varphi}_{R,t}^T\, \boldsymbol{R}_t + \varphi_t = \Big[(1 + \beta^A\, \varphi_k^A) \kappa \Big] k_t \\
&+ \Big[\beta^A\, \boldsymbol{\varphi}_{\tau A}^T\, \boldsymbol{\sigma}^A - (1 + \beta^A\, \varphi_k^A) \xi_0^A\, \boldsymbol{e}_1^T \Big] \boldsymbol{\tau}_t + \Big[\beta^A\, \boldsymbol{\varphi}_{MA}^T\, \boldsymbol{\Phi} + \Big(\beta^A\, \varphi_{\tau 1}^{AA}\, \sigma_{\text{forc}} (f_1 + f_2\, z_A^g \\
&- f_3\, (z_A^g)^{1-n}) + \beta^A\, \varphi_{\tau 1}^{BA}\, \sigma_{\text{forc}} (f_1 + f_2\, \alpha_A\, z_A^g - f_3\, (\alpha_A\, z_A^g)^{1-n}) \\
&- (1 + \beta^A\, \varphi_k^A)(d_{AA}\, z_A^g + a^A) \Big) M_{pre}^{-1}\, \boldsymbol{e}_1^T \Big] \boldsymbol{M}_t + \Big[\beta^A\, \boldsymbol{\varphi}_{R,t+1}^T \Big] R_t + \log\, x_t^* \\
&+ \beta^A\, \varphi_k^A\, \log(1 - x_t^*) + (1 + \beta^A\, \varphi_k^A) \log \mathcal{F}(\boldsymbol{A}_t,\, \boldsymbol{\mathcal{K}}_t^*,\, \boldsymbol{N}_t^*,\, \boldsymbol{E}_t^*) + (1 + \beta^A\, \varphi_k^A)(\xi_0^A + a^A) \\
&+ (\beta^A\, \varphi_{\tau 1}^{AA}\, \sigma_{\text{forc}} + \beta^A\, \varphi_{\tau 1}^{BA}\, \sigma_{\text{forc}})\, f_0 + \beta^A\, \boldsymbol{\varphi}_{MA}^T\, \tilde{\boldsymbol{e}}_{\boldsymbol{t}} - \beta^A\, \boldsymbol{\varphi}_{R,t+1}^T\, \boldsymbol{E}_t^{d^*} + \beta^A\, \varphi_{t+1}, \quad (75)
\end{aligned}$$

(iii): $S_t^A = 0, S_t^B > 0$

$$\varphi_k^A\, k_t + \boldsymbol{\varphi}_{\tau A}^T\, \boldsymbol{\tau}_t + \boldsymbol{\varphi}_{MA}^T\, \boldsymbol{M}_t + \boldsymbol{\varphi}_{R,t}^T\, \boldsymbol{R}_t + \varphi_t = \Big[(1+\beta^A\,\varphi_k^A)\kappa\Big] k_t$$
$$+\Big[\beta^A\, \boldsymbol{\varphi}_{\tau A}^T\, \boldsymbol{\sigma}^A - (1+\beta^A\,\varphi_k^A)\xi_0^A\, \boldsymbol{e}_1^T\Big]\boldsymbol{\tau}_t + \Big[\beta^A\, \boldsymbol{\varphi}_{MA}^T\, \boldsymbol{\Phi} + \Big(\beta^A\, \varphi_{\tau 1}^{AA}\, \sigma_{\text{forc}}(f_1 + f_2\, \alpha_B\, z_B^g$$
$$-f_3\,(\alpha_B\, z_B^g)^{1-n}) + \beta^A\, \varphi_{\tau 1}^{BA}\, \sigma_{\text{forc}}(f_1 + f_2\, z_B^g - f_3\,(z_B^g)^{1-n})$$
$$-(1+\beta^A\,\varphi_k^A)(d_{BA}\,\alpha_B\, z_B^g + a^A)\Big) M_{pre}^{-1}\, \boldsymbol{e}_1^T\Big]\boldsymbol{M}_t + \Big[\beta^A\, \boldsymbol{\varphi}_{R,t+1}^T\Big] R_t + \log\, x_t^*$$
$$+\beta^A\,\varphi_k^A\, \log(1-x_t^*) + (1+\beta^A\,\varphi_k^A)\log \mathcal{F}(\boldsymbol{A}_t,\, \boldsymbol{\mathcal{K}}_t^*,\, \boldsymbol{N}_t^*,\, \boldsymbol{E}_t^*) + (1+\beta^A\,\varphi_k^A)(\xi_0^A + a^A)$$
$$+(\beta^A\,\varphi_{\tau 1}^{AA}\,\sigma_{\text{forc}} + \beta^A\,\varphi_{\tau 1}^{BA}\,\sigma_{\text{forc}})\, f_0 + \beta^A\, \boldsymbol{\varphi}_{MA}^T\, \tilde{\boldsymbol{e}}_{\boldsymbol{t}} - \beta^A\, \boldsymbol{\varphi}_{R,t+1}^T\, \boldsymbol{E}_t^{d*} + \beta^A\,\varphi_{t+1}. \tag{76}$$

Hence, for all Nash equilibria the controlled dynamics remains linear in states.

D.3 Shadow values of the states.

Coefficient matching with respect to capital, k_t, yields

$$\varphi_k^A = (1+\beta^A\,\varphi_k^A)\kappa \quad \Leftrightarrow \quad \varphi_k^A = \frac{\kappa}{1-\beta^A\,\kappa}.$$

Structurally similar to the global model we get the consumption rate $x_t^* = 1-\beta^A\,\kappa$.

Coefficient matching with respect to transformed temperatures delivers

$$\boldsymbol{\varphi}_{\tau A}^T = -\xi_0^A\,(1+\beta^A\,\varphi_k^A)\,\boldsymbol{e}_1^T\,[\mathbf{1} - \beta^A\,\boldsymbol{\sigma}^A]^{-1}. \tag{77}$$

We define

$$\tilde{\sigma}_{ij}^A = \big[(\mathbf{1} - \beta^A\,\boldsymbol{\sigma}^{\boldsymbol{A}})^{-1}\big]_{ij} \text{ and } \tilde{\sigma}_{ij}^B = \big[(\mathbf{1} - \beta^B\,\boldsymbol{\sigma}^{\boldsymbol{B}})^{-1}\big]_{ij} \text{ for } i,j \in \{1,2\}.$$

Note that to preserve symmetry in notation, we defined the corresponding matrix $\boldsymbol{\sigma}^B$ by swapping the first and second rows and columns of the matrix $\boldsymbol{\sigma}^{\boldsymbol{A}}$.

Region A's shadow values of atmospheric temperature in regions A and B are therefore

$$\varphi_{\tau 1}^{AA} = -\xi_0^A\,(1+\beta^A\,\varphi_k^A)\tilde{\sigma}_{11}^A \tag{78}$$

$$\varphi_{\tau 1}^{BA} = -\xi_0^A\,(1+\beta^A\,\varphi_k^A)\tilde{\sigma}_{12}^A. \tag{79}$$

The temperature shadow values for region B follow by switching region indices. We

now define $\gamma_A \equiv \beta^A \xi_0^A \tilde{\sigma}_{11}^A \sigma_{\text{forc}}$, and $\gamma_A^{heat} \equiv \beta^A \xi_0^A \tilde{\sigma}_{12}^A \sigma_{\text{forc}}$. This gives us $\tilde{\delta}_A$, $\tilde{a}_A$, and $\tilde{b}_A$ (which we defined in equations (60) to (63)) as a function of the γ's.

$$-\tilde{\delta}_A = (1+\beta^A \varphi_k^A)(-\gamma_A f_2 - \alpha_A \gamma_A^{heat} f_2 - d_{AA})$$

$$\tilde{a}_A = (1+\beta^A \varphi_k^A)(1-n)\gamma_A f_3$$
$$\tilde{b}_A = (1+\beta^A \varphi_k^A)(1-n)\alpha_A \gamma_A^{heat} f_3$$

By switching region indices we gain $\tilde{\delta}_B$, $\tilde{a}_B$, and $\tilde{b}_B$. Thus, z_A and z_B are also functions of the γ's.

$$z_A(\gamma_A, \gamma_A^{heat}, \gamma_B, \gamma_B^{heat}) = \left(\frac{\delta_B(\gamma_B, \gamma_B^{heat})\, b_A(\gamma_A^{heat}) - a_B(\gamma_B)\, \delta_A(\gamma_A, \gamma_A^{heat})}{b_B(\gamma_B^{heat})\, b_A(\gamma_A^{heat}) - a_B(\gamma_B)\, a_A(\gamma_A)}\right)^{-\frac{1}{n}} \quad (80)$$

$$z_B(\gamma_B, \gamma_B^{heat}, \gamma_A, \gamma_A^{heat}) = \left(\frac{\delta_A(\gamma_A, \gamma_A^{heat})\, b_B(\gamma_B^{heat}) - a_A(\gamma_A)\, \delta_B(\gamma_B, \gamma_B^{heat})}{b_A(\gamma_A^{heat})\, b_B(\gamma_B^{heat}) - a_A(\gamma_A)\, a_B(\gamma_B)}\right)^{-\frac{1}{n}} \quad (81)$$

where the term $(1+\beta^A \varphi_k^A)$ cancels out and thus

$$\delta_A = \gamma_A f_2 + \alpha_A \gamma_A^{heat} f_2 + d_{AA}$$

$$a_A = (1-n)\gamma_A f_3$$
$$b_A = (1-n)\alpha_A \gamma_A^{heat} f_3.$$

Coefficient matching with respect to the carbon stocks, and using the γ's, yields (i): $S_t^A \neq 0$ and $S_t^B \neq 0$

$$\boldsymbol{\varphi}_{MA}^T = (1+\beta^A \varphi_k^A)\left(-f_2\,(\gamma_A\, z_A + \gamma_A^{heat}\, z_B) - \frac{d_{AA}\,(z_A - \alpha_B\, z_B) + \alpha_B\, d_{BA}(z_B - \alpha_A\, z_A)}{1-\alpha_A\,\alpha_B}\right.$$
$$\left. + f_3\,(\gamma_A\, z_A^{1-n} + \gamma_A^{heat}\, z_B^{1-n}) - a^A - f_1\,(\gamma_A + \gamma_A^{heat})\right) M_{pre}^{-1}\, \boldsymbol{e}_1^T\, (\mathbf{1} - \beta^A\, \boldsymbol{\Phi})^{-1}, \quad (82)$$

(ii): $S_t^A > 0$ and $S_t^B = 0$

$$\boldsymbol{\varphi}_{MA}^T = (1+\beta^A\,\varphi_k^A)\Bigg(-f_2\,(\gamma_A\,z_A^g+\gamma_A^{heat}\,\alpha_A\,z_A^g)-d_{AA}\,z_A^g+f_3\Big(\gamma_A\,(z_A^g)^{1-n} + \gamma_A^{heat}\,(\alpha_A\,z_A^g)^{1-n}\Big)-a^A-f_1\,(\gamma_A+\gamma_A^{heat})\Bigg)M_{pre}^{-1}\,\boldsymbol{e}_1^T\,(\mathbf{1}-\beta^A\,\boldsymbol{\Phi})^{-1}, \quad (83)$$

(iii): $S_t^A = 0$ and $S_t^B > 0$

$$\boldsymbol{\varphi}_{MA}^T = (1+\beta^A\,\varphi_k^A)\Bigg(-f_2\,(\gamma_A\,\alpha_B\,z_B^g+\gamma_A^{heat}\,z_B^g)-\,d_{BA}\,\alpha_B\,z_B^g+f_3\Big(\gamma_A\,(\alpha_B\,z_B^g)^{1-n} + \gamma_A^{heat}\,(z_B^g)^{1-n}\Big)-a^A-f_1\,(\gamma_A+\gamma_A^{heat})\Bigg)M_{pre}^{-1}\,\boldsymbol{e}_1^T\,(\mathbf{1}-\beta^A\,\boldsymbol{\Phi})^{-1}. \quad (84)$$

We define

$$\tilde{\phi}_{ij}^A = \big[(\mathbf{1}-\beta^A\,\boldsymbol{\Phi})^{-1}\big]_{ij} \text{ for } i,j\in\{1,2,3\},$$

yielding

$$\tilde{\phi}_{11}^A = \big[(\mathbf{1}-\beta^A\,\boldsymbol{\Phi})^{-1}\big]_{1,1}. \quad (85)$$

Similar to the global model, for ease of representation we omit the subscript of the term $\tilde{\phi}_{11}^A$ and instead use $\tilde{\phi}^A$. Switching region indices delivers the symmetric result for region B.

Coefficient matching with respect to the resource stock leads to

$$\boldsymbol{\varphi}_{R,t}^T = \beta\,\boldsymbol{\varphi}_{R,t+1}^T \quad \Leftrightarrow \quad \boldsymbol{\varphi}_{R,t} = \beta^{-t}\boldsymbol{\varphi}_{R,0} \quad \text{(Hotelling's rule)}.$$

The initial resource values $\boldsymbol{\varphi}_{R,0}^T$ depend on the set up of the economy, including assumptions about production and the energy sector. Given the coefficients and the optimal rate of consumption equation (74),(75), and (76) turn to the following condition:

$$\varphi_t-\beta^A\,\varphi_{t+1} = \log\,x_t^*+\beta^A\,\varphi_k^A\,\log(1-x_t^*)+(1+\beta^A\,\varphi_k^A)\log\mathcal{F}(\boldsymbol{A}_t,\,\boldsymbol{\mathcal{K}}_t^*,\,\boldsymbol{N}_t^*,\,\boldsymbol{E}_t^*) + (1+\beta^A\,\varphi_k^A)(\xi_0^A+a^A)+(\beta^A\,\varphi_{\tau 1}^{AA}\,\sigma_{\text{forc}}+\beta^A\,\varphi_{\tau 1}^{BA}\,\sigma_{\text{forc}})\,f_0+\beta\,\boldsymbol{\varphi}_{MA}^T\,\tilde{\boldsymbol{e}}_{\boldsymbol{t}}-\beta^A\,\boldsymbol{\varphi}_{R,t+1}^T\,\boldsymbol{E}_t^{d*} \quad (86)$$

This condition will be satisfied by picking the sequence $\varphi_0,\varphi_1,\varphi_2,....$ The additional condition $\lim_{t\to\infty}(\beta^A)^t V(\cdot)=0 \Rightarrow \lim_{t\to\infty}(\beta^A)^t\varphi_t = 0$ pins down this initial value

φ_0.

D.4 Rest of the world

In the following we show for the rest of the world that the system is linear in states and that the affine value function

$$V(k_t, \boldsymbol{\tau}_t, \boldsymbol{M}_t, \boldsymbol{R}_t, t) = \varphi_k^W \, k_t + \boldsymbol{\varphi}_{\tau W}^T \, \boldsymbol{\tau}_t + \boldsymbol{\varphi}_{MW}^T \, \boldsymbol{M}_t + \boldsymbol{\varphi}_{R,t}^T \, \boldsymbol{R}_t + \varphi_t, \qquad (87)$$

where

$$\boldsymbol{\varphi}_{\tau W}^T = \begin{pmatrix} \varphi_{\tau 1}^{AW} \\ \varphi_{\tau 1}^{BW} \\ \varphi_{\tau 2}^{W} \end{pmatrix} \qquad \text{and} \qquad \boldsymbol{\varphi}_{MW}^T = \begin{pmatrix} \varphi_{M1}^{W} \\ \varphi_{M2}^{W} \\ \varphi_{M3}^{W} \end{pmatrix},$$

solves the system. We suppress regional indices when there is no ambiguity to ease notation. Inserting the trial solution and the next periods states into the Bellman equation delivers

$$\begin{aligned}
&\varphi_k^W \, k_t + \boldsymbol{\varphi}_{\tau W}^T \, \boldsymbol{\tau}_t + \boldsymbol{\varphi}_{MW}^T \, \boldsymbol{M}_t + \boldsymbol{\varphi}_{R,t}^T \, \boldsymbol{R}_t + \varphi_t \\
&= \max_{x_t, \boldsymbol{N}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{E}_t} \Bigg\{ \log x_t + \kappa \, k_t + \log \mathcal{F}(\boldsymbol{A}_t, \, \boldsymbol{\mathcal{K}}_t, \, \boldsymbol{N}_t, \, \boldsymbol{E}_t) + \xi_0^W \, \left(1 - \tau_{1,t}^B\right) \\
&- (d_{BW} S_t^B + d_{AW} \, \alpha_A \, S_t^A) - a^W \, (m_t - 1) + \lambda_t^{\mathcal{K}} (1 - \sum_{i=1}^{I_K} \mathcal{K}_{i,t}) + \lambda_t^N (1 - \sum_{i=1}^{I_N} N_{i,t}) \\
&+ \beta^W \, \varphi_k^W \, \big(\kappa \, k_t + \log \mathcal{F}(\boldsymbol{A}_t, \, \boldsymbol{\mathcal{K}}_t, \, \boldsymbol{N}_t, \, \boldsymbol{E}_t) + \log(1 - x_t) + \xi_0^W \, \left(1 - \tau_{1,t}^B\right) \\
&- (d_{BW} S_t^B + d_{AW} \, \alpha_A \, S_t^A) - a^W \, (m_t - 1) \big) + \beta^W \, \boldsymbol{\varphi}_{\tau W}^T \left(\boldsymbol{\sigma}^A \boldsymbol{\tau}_t + \tilde{\boldsymbol{F}}_t \right) \\
&+ \beta^W \, \boldsymbol{\varphi}_{MW}^T \, (\boldsymbol{\Phi} \boldsymbol{M}_t + \tilde{\boldsymbol{e}}_{\boldsymbol{t}}) + \beta^W \, \boldsymbol{\varphi}_{R,t+1}^T \, \left(\boldsymbol{R}_t - \boldsymbol{E}_t^d \right) + \beta^W \, \varphi_{t+1} \Bigg\}.
\end{aligned}$$

First order conditions. The first order conditions for $x_t, \boldsymbol{N}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{E}_t$ are structurally the same as in the global model.

Verifying solution to the Bellman equation.

Inserting the optimal control rules $\boldsymbol{N}_t^*(\boldsymbol{A}_t, \varphi_k^W, \boldsymbol{\varphi}_{MW}, \boldsymbol{\varphi}_{R,t+1})$, $\boldsymbol{\mathcal{K}}_t^*(\boldsymbol{A}_t, \varphi_k^W, \boldsymbol{\varphi}_{MW}, \boldsymbol{\varphi}_{R,t+1})$,

and $\boldsymbol{E}_t^*(\boldsymbol{A}_t, \varphi_k^W, \boldsymbol{\varphi}_{MW}, \boldsymbol{\varphi}_{R,t+1})$ into the maximized Bellman equation gives us

$$\begin{aligned}
&\varphi_k^W\, k_t + \boldsymbol{\varphi}_{\tau W}^T\, \boldsymbol{\tau}_t + \boldsymbol{\varphi}_{MW}^T\, \boldsymbol{M}_t + \boldsymbol{\varphi}_{R,t}^T\, \boldsymbol{R}_t + \varphi_t = \log x_t^* + \kappa\, k_t + \log F(\boldsymbol{A}_t,\, \boldsymbol{\mathcal{K}}_t^*,\, \boldsymbol{N}_t^*,\, \boldsymbol{E}_t^*) \\
&+\xi_0^W\, \left(1 - \tau_{1,t}^B\right) - (d_{BW} S_t^B + d_{AW}\, \alpha_A\, S_t^A) - a^W\, (m_t - 1) + \beta^W\, \varphi_k^W\, \left(\kappa\, k_t + \log F(\boldsymbol{A}_t,\, \boldsymbol{\mathcal{K}}_t^*,\, \boldsymbol{N}_t^*,\, \boldsymbol{E}_t^*)\right. \\
&\left.+\log(1 - x_t^*) + \xi_0^W\, \left(1 - \tau_{1,t}^B\right) - (d_{BW} S_t^B + d_{AW}\, \alpha_A\, S_t^A) - a^W\, (m_t - 1)\right) + \beta^W\, \boldsymbol{\varphi}_{\tau W}^T \left(\boldsymbol{\sigma}^A \boldsymbol{\tau}_t + \tilde{\boldsymbol{F}}_t\right) \\
&\qquad + \beta^W\, \boldsymbol{\varphi}_{MW}^T\, (\boldsymbol{\Phi} \boldsymbol{M}_t + \tilde{\boldsymbol{e}}_{\boldsymbol{t}}) + \beta^W\, \boldsymbol{\varphi}_{R,t+1}^T \left(\boldsymbol{R}_t - \boldsymbol{E}_t^{d*}\right) + \beta^W\, \varphi_{t+1} \quad (88)
\end{aligned}$$

Arranging terms by states for the different Nash equilibria yields
(i): $S_t^A \neq 0, S_t^B \neq 0$

$$\begin{aligned}
&\varphi_k^W\, k_t + \boldsymbol{\varphi}_{\tau W}^T\, \boldsymbol{\tau}_t + \boldsymbol{\varphi}_{MW}^T\, \boldsymbol{M}_t + \boldsymbol{\varphi}_{R,t}^T\, \boldsymbol{R}_t + \varphi_t = \left[(1 + \beta^W\, \varphi_k^W)\kappa\right] k_t \\
&+\left[\beta^W\, \boldsymbol{\varphi}_{\tau W}^T\, \boldsymbol{\sigma}^A - (1 + \beta^W\, \varphi_k^W)\xi_0^W\, \boldsymbol{e}_1^T\right] \boldsymbol{\tau}_t + \left[\beta^W\, \boldsymbol{\varphi}_{MW}^T\, \boldsymbol{\Phi}\right. \\
&+\Big(\beta^W\, \varphi_{\tau 1}^{AW}\, \sigma_{\text{forc}}(f_1 + f_2\, z_A - f_3\, z_A^{1-n}) + \beta^W\, \varphi_{\tau 1}^{BW}\, \sigma_{\text{forc}}(f_1 + f_2\, z_B - f_3\, z_B^{1-n}) \\
&-(1 + \beta^W\, \varphi_k^W) \left(d_{AW}\, \frac{\alpha_A}{1 - \alpha_A\, \alpha_B}(z_A - \alpha_B\, z_B) + d_{BW}\, \frac{1}{1 - \alpha_A\, \alpha_B}(z_B - \alpha_A\, z_A)\right. \\
&\left.\left.+a^W\right)\Big) M_{pre}^{-1}\, \boldsymbol{e}_1^T\right] \boldsymbol{M}_t + \left[\beta^W\, \boldsymbol{\varphi}_{R,t+1}^T\right] R_t + \log\, x_t^* + \beta^W\, \varphi_k^W\, \log(1 - x_t^*) \\
&+(1 + \beta^W\, \varphi_k^W) \log F(\boldsymbol{A}_t,\, \boldsymbol{\mathcal{K}}_t^*,\, \boldsymbol{N}_t^*,\, \boldsymbol{E}_t^*) + (1 + \beta^W\, \varphi_k^W)(\xi_0^W + a^W) \\
&+f_0\, \beta^W (\varphi_{\tau 1}^{AW}\, \sigma_{\text{forc}} + \varphi_{\tau 1}^{BW}\, \sigma_{\text{forc}}) + \beta^W\, \boldsymbol{\varphi}_{MW}^T\, \tilde{\boldsymbol{e}}_{\boldsymbol{t}} - \beta^W\, \boldsymbol{\varphi}_{R,t+1}^T\, \boldsymbol{E}_t^{d*} + \beta^W\, \varphi_{t+1}.
\end{aligned}$$

with $z_A \in \{z_A^c, z_A^g\}$, $z_B \in \{z_B^c, z_B^g\}$, $d_{AA} \in \{d_{AA}^c, d_{AA}^g\}$, and $d_{BB} \in \{d_{BB}^c, d_{BB}^g\}$ depending on whether region A and B engage in CM or SG.

(ii): $S_t^A > 0, S_t^B = 0$

$$\begin{aligned}
&\varphi_k^W\, k_t + \boldsymbol{\varphi}_{\tau W}^T\, \boldsymbol{\tau}_t + \boldsymbol{\varphi}_{MW}^T\, \boldsymbol{M}_t + \boldsymbol{\varphi}_{R,t}^T\, \boldsymbol{R}_t + \varphi_t = \left[(1 + \beta^W\, \varphi_k^W)\kappa\right] k_t \\
&+\left[\beta^W\, \boldsymbol{\varphi}_{\tau W}^T\, \boldsymbol{\sigma}^A - (1 + \beta^W\, \varphi_k^W)\xi_0^W\, \boldsymbol{e}_1^T\right] \boldsymbol{\tau}_t + \left[\beta^W\, \boldsymbol{\varphi}_{MW}^T\, \boldsymbol{\Phi}\right. \\
&+\Big(\beta^W \sigma_{\text{forc}}\, \varphi_{\tau 1}^{AW} (f_1 + f_2\, z_A^g - f_3\, (z_A^g)^{1-n}) + \beta^W\, \sigma_{\text{forc}}\, \varphi_{\tau 1}^{BW} (f_1 + f_2\, \alpha_A\, z_A^g - f_3\, (\alpha_A\, z_A^g)^{1-n}) \\
&\left.-(1 + \beta^W\, \varphi_k^W)\, (d_{AW}\, \alpha_A\, z_A^g + a^W)\Big) M_{pre}^{-1}\, \boldsymbol{e}_1^T\right] \boldsymbol{M}_t + \left[\beta^W\, \boldsymbol{\varphi}_{R,t+1}^T\right] R_t + \log\, x_t^* \\
&+\beta^W\, \varphi_k^W\, \log(1 - x_t^*) + (1 + \beta^W\, \varphi_k^W) \log F(\boldsymbol{A}_t,\, \boldsymbol{\mathcal{K}}_t^*,\, \boldsymbol{N}_t^*,\, \boldsymbol{E}_t^*) + (1 + \beta^W\, \varphi_k^W)(\xi_0^W + a^W) \\
&+f_0\, \beta^W (\varphi_{\tau 1}^{AW}\, \sigma_{\text{forc}} + \varphi_{\tau 1}^{BW}\, \sigma_{\text{forc}}) + \beta^W\, \boldsymbol{\varphi}_{MW}^T\, \tilde{\boldsymbol{e}}_{\boldsymbol{t}} - \beta^W\, \boldsymbol{\varphi}_{R,t+1}^T\, \boldsymbol{E}_t^{d*} + \beta^W\, \varphi_{t+1}.
\end{aligned}$$

(iii): $S_t^A = 0, S_t^B > 0$

$$\varphi_k^W\, k_t + \boldsymbol{\varphi}_{\tau W}^T\, \boldsymbol{\tau}_t + \boldsymbol{\varphi}_{MW}^T\, \boldsymbol{M}_t + \boldsymbol{\varphi}_{R,t}^T\, \boldsymbol{R}_t + \varphi_t = \Big[(1+\beta^W\,\varphi_k^W)\kappa\Big] k_t$$
$$+\Big[\beta^W\,\boldsymbol{\varphi}_{\tau W}^T\,\boldsymbol{\sigma}^A - (1+\beta^W\,\varphi_k^W)\xi_0^W\,\boldsymbol{e}_1^T\Big]\boldsymbol{\tau}_t + \Big[\beta^W\,\boldsymbol{\varphi}_{MW}^T\,\boldsymbol{\Phi}$$
$$+\Big(\beta^W\,\sigma_{\text{forc}}\,\varphi_{\tau 1}^{AW}(f_1 + f_2\,z_B^g - f_3\,(\alpha_B\,z_B^g)^{1-n}) + \beta^W\,\sigma_{\text{forc}}\,\varphi_{\tau 1}^{BW}(f_1 + f_2\,\alpha_A\,z_B^g - f_3\,(z_B^g)^{1-n})$$
$$-(1+\beta^W\,\varphi_k^W)\,(d_{BW}\,z_B^g + a^W)\Big) M_{pre}^{-1}\,\boldsymbol{e}_1^T\Big]\boldsymbol{M}_t + \Big[\beta^W\,\boldsymbol{\varphi}_{R,t+1}^T\Big] R_t + \log\, x_t^*$$
$$+\beta^W\,\varphi_k^W\,\log(1-x_t^*) + (1+\beta^W\,\varphi_k^W)\log F(\boldsymbol{A}_t,\,\boldsymbol{\mathcal{K}}_t^*,\,\boldsymbol{N}_t^*,\,\boldsymbol{E}_t^*) + (1+\beta^W\,\varphi_k^W)(\xi_0^W + a^W)$$
$$+f_0\,\beta^W(\varphi_{\tau 1}^{AW}\,\sigma_{\text{forc}} + \varphi_{\tau 1}^{BW}\,\sigma_{\text{forc}}) + \beta^W\,\boldsymbol{\varphi}_{MW}^T\,\tilde{\boldsymbol{e}}_{\boldsymbol{t}} - \beta^W\,\boldsymbol{\varphi}_{R,t+1}^T\,\boldsymbol{E}_t^{d*} + \beta^W\,\varphi_{t+1}.$$

Hence, for all Nash equilibria the system is linear in states.

Shadow values of the states. Coefficient matching with respect to capital, k_t, yields

$$\varphi_k^W = (1+\beta^W\,\varphi_k^W)\kappa \quad \Leftrightarrow \quad \varphi_k^W = \frac{\kappa}{1-\beta^W\,\kappa}$$

Structurally similar to the global model we get the consumption rate $x_t^* = 1-\beta^W\,\kappa$.

Coefficient matching with respect to transformed temperatures delivers

$$\boldsymbol{\varphi}_{\tau W}^T = -\xi_0^W\,(1+\beta^W\,\varphi_k^W)\,\boldsymbol{e}_1^T\,(\mathbf{1}-\beta^W\,\boldsymbol{\sigma}^A)^{-1}. \tag{89}$$

Since the rest of the world is part of climate zone B, we define

$$\tilde{\sigma}_{ij}^W = \big[(\mathbf{1}-\beta^W\,\boldsymbol{\sigma}^{\boldsymbol{B}})^{-1}\big]_{ij} \text{ for } i,j \in \{1,2\}.$$

Note that to preserve symmetry in notation, we defined the corresponding matrix $\boldsymbol{\sigma}^B$ by swapping the first and second rows and columns of the matrix $\boldsymbol{\sigma}^{\boldsymbol{A}}$.

The rest of the world's shadow values of atmospheric temperature in regions A and B are therefore

$$\varphi_{\tau 1}^{AW} = -\xi_0^W\,(1+\beta^W\,\varphi_k^W)\tilde{\sigma}_{12}^W \tag{90}$$

$$\varphi_{\tau 1}^{BW} = -\xi_0^W\,(1+\beta^W\,\varphi_k^W)\tilde{\sigma}_{11}^W \tag{91}$$

We define $\gamma_W^{heat} \equiv \beta^W\,\xi_0^W\,\tilde{\sigma}_{12}^W\,\sigma_{\text{forc}}$ and $\gamma_W \equiv \beta^W\,\xi_0^W\,\tilde{\sigma}_{11}^W\,\sigma_{\text{forc}}$.

Coefficient matching with respect to carbon stocks and using the γ's yields

(i): $S_t^A \neq 0, S_t^B \neq 0$

$$\varphi_{M1}^W = (1+\beta^W \varphi_k^W)\Bigg(- f_2(\gamma_W^{heat} z_A + \gamma_W z_B) - d_{AW}\frac{\alpha_A}{1-\alpha_A \alpha_B}(z_A - \alpha_B z_B) - d_{BW}\frac{1}{1-\alpha_A \alpha_B}(z_B - \alpha_A z_A) + f_3\left(\gamma_W^{heat} z_A^{1-n} + \gamma_W z_B^{1-n}\right) - a^W - f_1(\gamma_W^{heat} + \gamma_W)\Bigg) M_{pre}^{-1} \boldsymbol{e}_1^T (\mathbf{1} - \beta^W \boldsymbol{\Phi})^{-1} \quad (92)$$

(ii): $S_t^A > 0, S_t^B = 0$

$$\varphi_{M1}^W = (1+\beta^W \varphi_k^W)\Bigg(- f_2(\gamma_W^{heat} z_A^g + \gamma_W \alpha_A z_A^g) - d_{AW}\alpha_A z_A^g + f_3\Big(\gamma_W^{heat}(z_A^g)^{1-n} + \gamma_W (\alpha_A z_A^g)^{1-n}\Big) - a^W - f_1(\gamma_W^{heat} + \gamma_W)\Bigg) M_{pre}^{-1} \boldsymbol{e}_1^T (\mathbf{1} - \beta^W \boldsymbol{\Phi})^{-1} \quad (93)$$

(iii): $S_t^A = 0, S_t^B > 0$

$$\varphi_{M1}^W = (1+\beta^W \varphi_k^W)\Bigg(- f_2(\gamma_W^{heat} \alpha_B z_B^g + \gamma_W z_B^g) - d_{BW} z_B^g + f_3\Big(\gamma_W^{heat}(\alpha_B z_B^g)^{1-n} + \gamma_W (z_B^g)^{1-n}\Big) - a^W - f_1(\gamma_W^{heat} + \gamma_W)\Bigg) M_{pre}^{-1} \boldsymbol{e}_1^T (\mathbf{1} - \beta^W \boldsymbol{\Phi})^{-1}. \quad (94)$$

We define

$$\tilde{\phi}_{ij}^W = \left[(\mathbf{1} - \beta^W \boldsymbol{\Phi})^{-1}\right]_{ij} \text{ for } i,j \in \{1,2,3\},$$

yielding

$$\tilde{\phi}_{11}^W = \left[(\mathbf{1} - \beta^W \boldsymbol{\Phi})^{-1}\right]_{1,1}. \quad (95)$$

Similar to the global model, for ease of representation we omit the subscript of the term $\tilde{\phi}_{11}^W$ and instead use $\tilde{\phi}^W$.

From coefficient matching with respect to the resource stock we have

$$\boldsymbol{\varphi}_{R,t}^T = \beta^W \boldsymbol{\varphi}_{R,t+1}^T \quad \Leftrightarrow \quad \boldsymbol{\varphi}_{R,t} = (\beta^W)^{-t} \boldsymbol{\varphi}_{R,0} \quad \text{(Hotelling's rule)}.$$

The initial resource values $\boldsymbol{\varphi}_{R,0}^T$ depend on the set up of the economy, including assumptions about production and the energy sector. Given the coefficients and the optimal rate of consumption equation (74),(75), and (76) turn to the following condition:

$$\varphi_t - \beta^W \varphi_{t+1} = \log x_t^* + \beta^W \varphi_k^W \log(1-x_t^*) + (1+\beta^W \varphi_k^W) \log F(\boldsymbol{A}_t, \boldsymbol{\mathcal{K}}_t^*, \boldsymbol{N}_t^*, \boldsymbol{E}_t^*) + (1+\beta^W \varphi_k^W)(\xi_0^W + a^W) + \beta^W (\varphi_{\tau 1}^{AW} \sigma_{\text{forc}} + \varphi_{\tau 1}^{BW} \sigma_{\text{forc}}) f_0 + \beta^W \boldsymbol{\varphi}_{MW}^T \tilde{\boldsymbol{e}}_{\boldsymbol{t}} - \beta^W \boldsymbol{\varphi}_{R,t+1}^T \boldsymbol{E}_t^{d*} \tag{96}$$

This condition will be satisfied by picking the sequence $\varphi_0, \varphi_1, \varphi_2, ...$ The additional condition $\lim_{t\to\infty}(\beta^W)^t V(\cdot) = 0 \Rightarrow \lim_{t\to\infty}(\beta^W)^t \varphi_t = 0$ pins down this initial value φ_0.

E Proofs for the regional model

E.1 Proof of Proposition 3 & 7 (strategies)

This proof makes use of section D.1, where we derive the Markov strategies, section D.2, where we show that the strategies are consistent with the assumed linear form of trial solution for the Bellman equation, and section D.3, where we derive the solutions for the shadow values.

General model (Proposition 7). In section D.1 we show that the following reaction functions characterize a Nash equilibrium of the dynamic game: If (i) $S_t^B = 0$, then

$$z_A^g = \left(\frac{\tilde{a}_A + \tilde{b}_A \alpha_A^{-n}}{\tilde{\delta}_A^g} \right)^{\frac{1}{n}},$$

and if (ii) $S_t^B \neq 0$, then

$$z_A^g = \left(\frac{\tilde{a}_A - \tilde{b}_A \frac{\tilde{b}_B}{\tilde{a}_B}}{\tilde{\delta}_A^g - \tilde{b}_A \frac{\tilde{\delta}_B}{\tilde{a}_B}} \right)^{\frac{1}{n}}, \quad z_A^c = \left(\frac{\tilde{a}_A - \tilde{b}_A \frac{\tilde{b}_B}{\tilde{a}_B}}{\tilde{\delta}_A^c - \tilde{b}_A \frac{\tilde{\delta}_B}{\tilde{a}_B}} \right)^{\frac{1}{n}},$$

where

$$-\tilde{\delta}_A^g = \beta^A \varphi_{\tau 1}^{AA} \sigma_{\text{forc}} f_2 + \beta^A \varphi_{\tau 1}^{BA} \sigma_{\text{forc}} f_2 \alpha_A - (1+\beta^A \varphi_k^A) d_{AA}^g$$
$$-\tilde{\delta}_A^c = \beta^A \varphi_{\tau 1}^{AA} \sigma_{\text{forc}} f_2 + \beta^A \varphi_{\tau 1}^{BA} \sigma_{\text{forc}} f_2 \alpha_A - (1+\beta^A \varphi_k^A) d_{AA}^c$$
$$\tilde{a}_A = (n-1)\beta^A \varphi_{\tau 1}^{AA} \sigma_{\text{forc}} f_3$$
$$\tilde{b}_A = (n-1)\beta^A \varphi_{\tau 1}^{BA} \sigma_{\text{forc}} f_3 \alpha_A.$$

Swapping country indices characterizes region B's strategies.

In section D.3 we show that region A's shadow values of atmospheric temperature in regions A and B are

$$\varphi_{\tau 1}^{AA} = -\xi_0^A \,(1 + \beta^A \,\varphi_k^A)\tilde{\sigma}_{11}^A \tag{97}$$

$$\varphi_{\tau 1}^{BA} = -\xi_0^A \,(1 + \beta^A \,\varphi_k^A)\tilde{\sigma}_{12}^A. \tag{98}$$

The temperature shadow values for region B follow by switching region indices.

Defining $\gamma_A \equiv \beta^A \,\xi_0^A \,\tilde{\sigma}_{11}^A \,\sigma_{\text{forc}}$, and $\gamma_A^{heat} \equiv \beta^A \,\xi_0^A \,\tilde{\sigma}_{12}^A \,\sigma_{\text{forc}}$, gives us δ_A, a_A, and b_A as a function of the γ's.

$$\begin{aligned}
a_A &\equiv \frac{\tilde{a}_A}{(1 + \beta^A \,\varphi_k^A)} = (1 - n)\gamma_A \,f_3 \\
b_A &\equiv \frac{\tilde{b}_A}{(1 + \beta^A \,\varphi_k^A)} = (1 - n)\alpha_A \,\gamma_A^{heat} \,f_3 \\
\delta_A^g &\equiv \frac{\tilde{\delta}_A^g}{(1 + \beta^A \,\varphi_k^A)} = (d_{AA}^g + \epsilon_A^g) + \gamma_A \,f_2 + \alpha_A \,\gamma_A^{heat} \,f_2 \\
\delta_A^c &\equiv \frac{\tilde{\delta}_A^c}{(1 + \beta^A \,\varphi_k^A)} = (d_{AA}^c - \epsilon_A^c) + \gamma_A \,f_2 + \alpha_A \,\gamma_A^{heat} \,f_2.
\end{aligned}$$

The solutions for δ_B, a_B, and b_B follow by switching region indices. Thus,

$$z_A^g = \left(\frac{a_A + b_A \,\alpha_A^{-n}}{\delta_A^g}\right)^{\frac{1}{n}},$$

and if (ii) $S_t^B \neq 0$, then

$$z_A^g = \left(\frac{a_A - \tilde{b}_A \frac{b_B}{a_B}}{\delta_A^g - b_A \frac{\delta_B}{a_B}}\right)^{\frac{1}{n}}, \quad z_A^c = \left(\frac{a_A - b_A \frac{b_B}{a_B}}{\delta_A^c - b_A \frac{\delta_B}{a_B}}\right)^{\frac{1}{n}}, \tag{99}$$

Base model (Proposition 3). No heat flows, i.e. setting $\sigma_B^A = \sigma_A^B = \sigma_A^O = \sigma_B^O = 0$, simplifies the shadow values of temperature such that $\varphi_{\tau 1}^{AA} = -\xi_0^A \,(1 + \beta^A \,\varphi_k^A)\,\tilde{\sigma}_{11}^A$ and $\varphi_{\tau 1}^{BA} = 0$ since $\tilde{\sigma}_{12}^A = 0$. For the base model, we use a slightly simpler notation and define $\varphi_{\tau 1}^A \equiv \varphi_{\tau 1}^{AA}$ and $\varphi_{\tau 1}^B \equiv \varphi_{\tau 1}^{BB}$.

No heat flows imply that $b_A = b_B = 0$, and therefore equations (99) turn to

$$z_A^g = \left(\frac{\delta_A^g}{a_A}\right)^{-\frac{1}{n}} = \left(\frac{(n-1)\beta^A \,\varphi_{\tau 1}^A \,\sigma_{\text{forc}} \,f_3}{(1 + \beta^A \,\varphi_k)d_{AA}^g - \beta^A \varphi_{\tau 1}^A \,\sigma_{\text{forc}} \,f_2}\right)^{\frac{1}{n}}$$

$$z_A^c = \left(\frac{\delta_A^c}{a_A}\right)^{-\frac{1}{n}} = \left(\frac{(n-1)\beta^A\,\varphi_{\tau 1}^A\,\sigma_{\text{forc}}\,f_3}{(1+\beta^A\,\varphi_k)d_{AA}^c - \beta^A\varphi_{\tau 1}^A\,\sigma_{\text{forc}}\,f_2}\right)^{\frac{1}{n}}$$

and

$$z_B = \left(\frac{\delta_B}{a_B}\right)^{-\frac{1}{n}} = \left(\frac{(n-1)\beta^B\,\varphi_{\tau 1}^B\,\sigma_{\text{forc}}\,f_3}{(1+\beta^B\,\varphi_k)d_{BB} - \beta^B\varphi_{\tau 1}^B\,\sigma_{\text{forc}}\,f_2}\right)^{\frac{1}{n}}$$

with $z_B \in \{z_B^c, z_B^g\}$. Note that in this simplification $z_A^{go} = \left(\frac{\delta_A^g}{a_A}\right)^{-\frac{1}{n}} = z_A^g$ since $b_A = 0$. Switching region indices gives us the analogous result for region B. Similar to the global model and to ease representation we define $\tilde{\sigma}^A = \tilde{\sigma}_{11}^A\,\sigma_{\text{forc}}$ and $\tilde{\sigma}^B = \tilde{\sigma}_{11}^B\,\sigma_{\text{forc}}$. Using $\gamma_A = \beta^A\,\xi_0^A\,\tilde{\sigma}^A$, and $\gamma_B = \beta^B\,\xi_0^B\,\tilde{\sigma}^B$, and noting that $\gamma_A^{heat} = 0$ and $\gamma_B^{heat} = 0$, leads to

$$z_A^g = \left(\frac{(1-n)\,f_3\,\gamma_A}{f_2\,\gamma_A + (d_{AA}^g + \epsilon_A^g)}\right)^{\frac{1}{n}} \quad \text{and} \quad z_A^c = \left(\frac{(1-n)\,f_3\,\gamma_A}{f_2\,\gamma_A + (d_{AA}^c - \epsilon_A^c)}\right)^{\frac{1}{n}}. \tag{100}$$

Switching region indices shows that for region B

$$z_B^g = \left(\frac{(1-n)\,f_3\,\gamma_B}{f_2\,\gamma_B + (d_{BB}^g + \epsilon_B^g)}\right)^{\frac{1}{n}} \quad \text{and} \quad z_B^c = \left(\frac{(1-n)\,f_3\,\gamma_B}{f_2\,\gamma_B + (d_{BB}^c - \epsilon_B^c)}\right)^{\frac{1}{n}}. \tag{101}$$

E.2 Proof of Proposition 4 (Nash equilibria)

Using the reaction functions for region A and B from section D.1 we derive the Nash equilibria. We have excluded the case that both countries would engage in CM by assumption.

i.a) In the case where both regions are cooling ($S_t^A > 0, S_t^B > 0$) we obtain

$$S_t^{A^*}(m_t) = \frac{m_t}{1-\alpha_A\,\alpha_B}\left(z_A^g - \alpha_B\,z_B^g\right) > 0 \quad \Rightarrow \quad z_A^g > \alpha_B\,z_B^g \tag{102}$$

and

$$S_t^{B^*}(m_t) = \frac{m_t}{1-\alpha_A\,\alpha_B}\left(z_B^g - \alpha_A z_A^g\right) > 0 \quad \Rightarrow \quad z_B^g > \alpha_A z_A^g. \tag{103}$$

Together, the two equations imply

$$\alpha_B < \underbrace{\frac{z_A^g}{z_B^g}}_{\equiv H} < \alpha_A^{-1} \tag{104}$$

Note that this condition defines a non-empty range of parameter values unless $\alpha_A^n =$

$\alpha_B^n = 1$. Thus, condition (104) states the range of parameter values α_B, α_A, d_{AA}^g, d_{BB}^g, ϵ_A^g, ϵ_B^g, and f_2 for which there exists an equilibrium in which both regions are cooling the world given γ_A, γ_A^{heat}, γ_B and γ_B^{heat}.

i.b) In the case where region A is cooling ($S_t^A > 0$) and region B is warming ($S_t^B < 0$) we obtain

$$S_t^{A^*}(m_t) = \frac{m_t}{1 - \alpha_A \, \alpha_B} \left(z_A^g - \alpha_B z_B^c \right) > 0 \quad \Leftrightarrow \quad z_A^g > \alpha_B z_B^c \tag{105}$$

$$S_t^{B^*}(m_t) = \frac{m_t}{1 - \alpha_A \, \alpha_B} \left(z_B^c - \alpha_A z_A^g \right) < 0 \quad \Leftrightarrow \quad z_B^c < \alpha_A z_A^g. \tag{106}$$

Therefore, the parameter range in which this case defines the Nash equilibrium is characterized by

$$\frac{z_A^g}{z_B^c} > \max\{\alpha_B, \alpha_A^{-1}\} \quad \Leftrightarrow \quad h \equiv \frac{z_A^g}{z_B^c} > \alpha_A^{-1} = \max\{\alpha_B, \alpha_A^{-1}\} \tag{107}$$

i.c) The case where region A is warming ($S_t^A < 0$) and region B is cooling ($S_t^B > 0$) follows by symmetry (switching the region indices)

$$\frac{z_B^g}{z_A^c} > \alpha_B^{-1} = \max\{\alpha_A, \alpha_B^{-1}\} \quad \Rightarrow \quad \hat{H} \equiv \frac{z_A^c}{z_B^g} < \alpha_B. \tag{108}$$

ii) In the case where region A is cooling ($S_t^A > 0$) and region B is not acting ($S_t^B = 0$) it has to be optimal for region B to neither engage in cooling, nor in CM. Given region A is taking the same actions as in scenarios i and ii, region B's reaction function can neither satisfy equation (103) nor (106). Therefore, it must be that $H \geq \alpha_A^{-1}$ and $h \leq \alpha_A^{-1}$. In addition, region A's reaction function becomes

$$S_t^{A^*}(m_t) = z_A^g \, m_t > 0, \tag{109}$$

which will always be satisfied. The reaction function of region B is obviously $S_t^{B^*} = 0$.

iii) Finally, the symmetric reasoning for region B cooling ($S_t^B > 0$) and region A not acting ($S_t^A = 0$) delivers $\frac{1}{H} \geq \alpha_B^{-1}$ and $\frac{1}{\hat{H}} \leq \alpha_B^{-1}$ or

$$\frac{1}{\hat{H}} \leq \alpha_B^{-1} \leq \frac{1}{H} \quad \Leftrightarrow \quad \hat{H} \geq \alpha_B \geq H. \tag{110}$$

Here the reaction functions are

$$S_t^{B^*}(m_t) = z_B^g \, m_t > 0 \tag{111}$$

and locally $S_t^{A^*} = 0$. These 5 cases are mutually exclusive and cover the full parameter domain.

E.3 Proof of Proposition 5 & 8 (regional SCC)

The proof makes use of the solutions for the shadow values from section D.3.

General model (Proposition 8). Inserting φ_k^A into (82), (83), and (84) delivers:

(i): $S_t^A \neq 0$ and $S_t^B \neq 0$

$$\varphi_{M1}^A = \frac{1}{1-\beta^A \kappa}\Bigg(-f_2\,(\gamma_A\, z_A + \gamma_A^{heat}\, z_B) - \frac{d_{AA}\,(z_A - \alpha_B\, z_B) + \alpha_B\, d_{BA}(z_B - \alpha_A\, z_A)}{1-\alpha_A\,\alpha_B} + f_3\,(\gamma_A\, z_A^{1-n} + \gamma_A^{heat}\, z_B^{1-n}) - a^A - f_1\,(\gamma_A + \gamma_A^{heat})\Bigg) M_{pre}^{-1}\,\tilde{\phi}^A$$

(ii): $S_t^A > 0$ and $S_t^B = 0$

$$\varphi_{M1}^A = \frac{1}{1-\beta^A \kappa}\Bigg(-f_2\,(\gamma_A\, z_A^g + \gamma_A^{heat}\,\alpha_A\, z_A^g) - d_{AA}\, z_A^g + f_3\Big(\gamma_A\,(z_A^g)^{1-n} + \gamma_A^{heat}\,(\alpha_A\, z_A^g)^{1-n}\Big) - a^A - f_1\,(\gamma_A + \gamma_A^{heat})\Bigg) M_{pre}^{-1}\,\tilde{\phi}^A$$

(iii): $S_t^A = 0$ and $S_t^B > 0$

$$\varphi_{M1}^A = \frac{1}{1-\beta^A \kappa}\Bigg(-f_2\,(\gamma_A\,\alpha_B\, z_B^g + \gamma_A^{heat}\, z_B^g) - d_{BA}\,\alpha_B\, z_B^g + f_3\Big(\gamma_A\,(\alpha_B\, z_B^g)^{1-n} + \gamma_A^{heat}\,(z_B^g)^{1-n}\Big) - a^A - f_1\,(\gamma_A + \gamma_A^{heat})\Bigg) M_{pre}^{-1}\,\tilde{\phi}^A$$

The regional SCC is the negative of the regional shadow value of atmospheric carbon expressed in money-measured consumption units. Thus,

(i): $S_t^A \neq 0$ and $S_t^B \neq 0$

$$SCC^A = -(1-\beta^A\,\kappa)Y_{A,t}^{net}\,\varphi_{M1}^A$$
$$= \frac{Y_{A,t}^{net}}{M_{pre}}\Bigg[a^A + f_1\,(\gamma_A + \gamma_A^{heat}) - \Bigg(\Bigg(\frac{f_3}{z_A^n} - f_2\Bigg)\gamma_A - d_{AA}\Bigg)z_A - \Bigg(\frac{f_3}{z_B^n} - f_2\Bigg)\gamma_A^{heat}\,z_B$$
$$-\frac{\alpha_B(z_B - \alpha_A\,z_A)(d_{AA} - d_{BA})}{1-\alpha_A\,\alpha_B}\Bigg]\,\tilde{\phi}^A, \tag{112}$$

where we extended the term by $\frac{-d_{AA}+\alpha_A\,\alpha_B\,d_{AA}}{1-\alpha_A\,\alpha_B}z_A + d_{AA}\,z_A\ (=0)$.

(ii): $S_t^A > 0$ and $S_t^B = 0$

$$SCC^A = -(1-\beta^A\,\kappa)Y_{A,t}^{net}\,\varphi_{M1}^A$$
$$= \frac{Y_{A,t}^{net}}{M_{pre}}\Bigg[a^A + f_1\,(\gamma_A + \gamma_A^{heat}) - \Bigg(\Bigg(\frac{f_3}{(z_A^g)^n} - f_2\Bigg)\gamma_A - d_{AA}\Bigg)z_A^g$$
$$-\Bigg(\frac{f_3}{(\alpha_A\,z_A^g)^n} - f_2\Bigg)\gamma_A^{heat}\,\alpha_A\,z_A^g\Bigg]\,\tilde{\phi}^A, \tag{113}$$

(iii): $S_t^A = 0$ and $S_t^B > 0$

$$SCC^A = -(1-\beta^A\,\kappa)Y_{A,t}^{net}\,\varphi_{M1}^A$$
$$= \frac{Y_{A,t}^{net}}{M_{pre}}\Bigg[a^A + f_1\,(\gamma_A + \gamma_A^{heat}) - \Bigg(\Bigg(\frac{f_3}{(\alpha_B\,z_B^g)^n} - f_2\Bigg)\gamma_A - d_{BA}\Bigg)\alpha_B\,z_B^g$$
$$-\Bigg(\frac{f_3}{(z_B^g)^n} - f_2\Bigg)\gamma_A^{heat}\,z_B^g\Bigg]\,\tilde{\phi}^A. \tag{114}$$

Summarizing all terms in the SCC that do not depend on γ_A^{heat} in the term $SCC_{w/o}^A$ leads to Proposition 8.

Base model (Proposition 5). In the special case, where σ_B^A, σ_A^B, σ_A^O, σ_B^O are equal to zero and thus also $\gamma_A^{heat} = 0$, the general equations (112), (113), and (114) for the regional SCC turn to
(i): $S_t^A \neq 0$ and $S_t^B \neq 0$

$$SCC^A = \frac{Y_{A,t}^{net}}{M_{pre}}\Bigg[a^A + f_1\,\gamma_A - \Bigg(\Bigg(\frac{f_3}{z_A^n} - f_2\Bigg)\gamma_A - d_{AA}\Bigg)z_A - \frac{\alpha_B(z_B - \alpha_A\,z_A)(d_{AA} - d_{BA})}{1-\alpha_A\,\alpha_B}\Bigg]\,\tilde{\phi}^A.$$

with $z_A \in \{z_A^c, z_A^g\}$, $z_B \in \{z_B^c, z_B^g\}$, $d_{AA} \in \{d_{AA}^c + \epsilon^c, d_{AA}^g + \epsilon^g\}$, and $d_{BB} \in \{d_{BB}^c + \epsilon^c, d_{BB}^g + \epsilon^g\}$.

(ii): $S_t^A > 0$ and $S_t^B = 0$

$$SCC^A = \frac{Y_{A,t}^{net}}{M_{pre}} \left[a^A + f_1 \, \gamma_A - \left(\left(\frac{f_3}{(z_A^g)^n} - f_2 \right) \gamma_A - d_{AA} \right) z_A^g \right] \tilde{\phi}^A.$$

(iii): $S_t^A = 0$ and $S_t^B > 0$

$$SCC^A = \frac{Y_{A,t}^{net}}{M_{pre}} \left[a^A + f_1 \, \gamma_A - \left(\left(\frac{f_3}{(\alpha_B \, z_B^g)^n} - f_2 \right) \gamma_A - d_{BA} \right) \alpha_B \, z_B^g \right] \tilde{\phi}^A.$$

E.4 Proof of Corollary 1

Region B inactive ($S_t^B = 0$). Analogously to equation (12) from the global model, the cooling contribution

$$\left(\left(\frac{f_3}{(z_A^g)^n} - f_2 \right) \gamma_A - (d_{AA}^g + \epsilon_A^g) \right) z_A^g = n \frac{(1-n)^{\frac{1-n}{n}} \, (\gamma_A \, f_3)^{\frac{1}{n}}}{((d_{AA}^g + \epsilon_A^g) + \gamma_A \, f_2)^{\frac{1-n}{n}}} > 0$$

always reduces the SCC.

Region A inactive ($S_t^A = 0$). The effect of SG on the SCC of region A (Δ_{SCC^A}) is determined by the sign of the cooling term

$$\left(\frac{f_3}{(\alpha_B \, z_B^g)^n} - f_2 \right) \gamma_A - d_{BA}^g \Big) \lessgtr 0 \qquad \Leftrightarrow \qquad \Delta_{SCC^A} \lessgtr 0, \tag{115}$$

Rearranging (115) leads to

$$\left(\frac{f_3 \, \gamma_A}{f_2 \, \gamma_A + d_{BA}^g} \right)^{\frac{1}{n}} \lessgtr \alpha_B \, z_B^g \qquad \Leftrightarrow \qquad \Delta_{SCC^A} \lessgtr 0,$$

Thus, the cooling term decrease the SCC, if $\alpha_B z_B^g < \left(\frac{f_3 \, \gamma_A}{f_2 \, \gamma_A + d_{BA}^g} \right)^{\frac{1}{n}}$. The unilateral action equilibrium with $S_t^A = 0$ requires $z_A^g \leq \alpha_B z_B^g \leq z_A^c$. If $d_{BA}^g < d_{AA}^g + \epsilon_A^g$ then there exists a non-empty z_B^g-interval where $z_A^g \leq \alpha_B z_B^g < \left(\frac{f_3 \, \gamma_A}{f_2 \, \gamma_A + d_{BA}^g} \right)^{\frac{1}{n}}$.

We note that region A will move to the climate clash equilibrium before its SCC dominates that of a world without SG if spillover damages are very high and CM are very cheap and effective, $d_{BA}^g < d_{AA}^c - \epsilon_A^c$.

Both regions active: If both regions are active, the effect of SG on the SCC^A

depends on two terms:

$$\text{cooling term: } \left(\left(\frac{f_3}{(z_A)^n} - f_2\right)\gamma_A - d_{AA}\right) z_A = n\frac{(1-n)^{\frac{1-n}{n}}\,(\gamma_A\, f_3)^{\frac{1}{n}}}{(d_{AA} + \gamma_A\, f_2)^{\frac{1-n}{n}}} > 0$$

$$\text{spillover term: } -\frac{\alpha_B(z_B - \alpha_A\, z_A)(d_{AA} - d_{BA})}{1 - \alpha_A\,\alpha_B} = -\alpha_B \frac{S_t^B(m_t)}{m_t}(d_{AA} - d_{BA})$$

We note that the cooling term always decreases the SCC^A, whereas the sign of the spillover term can vary across the different equilibria.

(i) $\underline{S_t^A > 0 \text{ and } S_t^B > 0}$: In the case where both regions cool, the spillover term can be positive or negative. If the spillover damages d_{BA} are substantially larger than region A's self-imposed damages and costs captured by d_{AA}, the positive spillover term can dominate the cooling term. In this case, the availability of SG increases the SCC^A. If, however, the spillover damages are lower than region A's self-imposed damages and costs, the spillover term will be negative and the availability of SG decreases the SCC^A. Overall, the availability of SG can increase, decrease or leave the SCC^A unchanged.

(ii) $\underline{S_t^A > 0 \text{ and } S_t^B < 0}$: In the climate clash case where region A cools, the spillover term is always positive as $d_{AA} - d_{BA} = d_{AA}^g + \epsilon_A^g - d_{BA}^c > 0$ by Assumption 2. Depending on whether the positive spillover term dominates the cooling term, the availability of SG can increase, decrease or leave the SCC^A unchanged.

(iii) $\underline{S_t^A < 0 \text{ and } S_t^B > 0}$: In the climate clash where region A engages in CM, the spillover term is again positive as $d_{AA} - d_{BA} = d_{AA}^c - \epsilon_A^c - d_{BA}^g < 0$ by Assumption 2). Depending on whether the positive spillover term dominates the cooling term, the availability of SG can increase, decrease or leave the SCC^A unchanged.

E.5 Proof of Proposition 6 (SCC in rest of the world)

The proof makes use of the solutions for the shadow values from section D.4.

General model. Inserting φ_k^W into (92), (93), and (94) delivers
(i): $S_t^A \neq 0, S_t^B \neq 0$

$$\varphi_{M1}^W = \frac{1}{1-\beta^W\,\kappa}\Bigg(- f_2(\gamma_W^{heat}\, z_A + \gamma_W\, z_B) - d_{AW}\,\frac{\alpha_A}{1-\alpha_A\,\alpha_B}(z_A - \alpha_B\, z_B)$$

$$-d_{BW}\,\frac{1}{1-\alpha_A\,\alpha_B}(z_B - \alpha_A\, z_A) + f_3\left(\gamma_W^{heat}\, z_A^{1-n} + \gamma_W\, z_B^{1-n}\right) - a^W - f_1(\gamma_W^{heat} + \gamma_W)\Bigg) M_{pre}^{-1}\,\tilde{\phi}^W$$

(ii): $S_t^A > 0, S_t^B = 0$

$$\varphi_{M1}^W = \frac{1}{1-\beta^W\,\kappa}\Bigg(- f_2(\gamma_W^{heat}\,z_A^g + \gamma_W\,\alpha_A\,z_A^g) - d_{AW}\,\alpha_A\,z_A^g + f_3\Big(\gamma_W^{heat}\,(z_A^g)^{1-n} + \gamma_W\,(\alpha_A\,z_A^g)^{1-n}\Big) - a^W - f_1(\gamma_W^{heat} + \gamma_W)\Bigg) M_{pre}^{-1}\,\tilde{\phi}^W$$

(iii): $S_t^A = 0, S_t^B > 0$

$$\varphi_{M1}^W = \frac{1}{1-\beta^W\,\kappa}\Bigg(- f_2(\gamma_W^{heat}\,\alpha_B z_B^g + \gamma_W\ z_B^g) - d_{BW}\,z_B^g + f_3\Big(\gamma_W^{heat}\,(\alpha_B\,z_B^g)^{1-n} + \gamma_W\,(z_B^g)^{1-n}\Big) - a^W - f_1(\gamma_W^{heat} + \gamma_W)\Bigg) M_{pre}^{-1}\,\tilde{\phi}^W$$

The regional SCC is the negative of the regional shadow value of atmospheric carbon expressed in money-measured consumption units. Thus,
(i): $S_t^A \neq 0$ and $S_t^B \neq 0$

$$\begin{aligned} SCC^W &= -(1-\beta^W\,\kappa) Y_{W,t}^{net}\,\varphi_{M1}^W \\ &= \frac{Y_{W,t}^{net}}{M_{pre}}\Bigg[a^W + f_1(\gamma_W^{heat} + \gamma_W) - \left(\frac{f_3}{z_A^n} - f_2\right)\gamma_W^{heat}\,z_A \\ &\quad - \left(\left(\frac{f_3}{z_B^n} - f_2\right)\gamma_W - d_{BW}\right) z_B - \frac{\alpha_A(z_A - \alpha_B\,z_B)(d_{BW} - d_{AW})}{1-\alpha_A\,\alpha_B}\Bigg]\,\tilde{\phi}^W, \end{aligned} \tag{116}$$

where we extended the term by $\frac{-d_{BW}+\alpha_A\,\alpha_B\,d_{BW}}{1-\alpha_A\,\alpha_B} z_B + d_{BW}\,z_B\ (=0)$.
(ii): $S_t^A > 0$ and $S_t^B = 0$

$$\begin{aligned} SCC^W &= -(1-\beta^W\,\kappa) Y_{W,t}^{net}\,\varphi_{M1}^W \\ &= \frac{Y_{W,t}^{net}}{M_{pre}}\Bigg[a^W + f_1\,(\gamma_W^{heat} + \gamma_W) - \left(\left(\frac{f_3}{(z_A^g)^n} - f_2\right)\gamma_W^{heat}\right) z_A^g \\ &\quad - \left(\left(\frac{f_3}{(\alpha_A\,z_A^g)^n} - f_2\right)\gamma_W - d_{AW}\right)\alpha_A\,z_A^g\Bigg]\,\tilde{\phi}^W, \end{aligned} \tag{117}$$

(iii): $S_t^A = 0$ and $S_t^B > 0$

$$\begin{aligned} SCC^W &= -(1-\beta^W\,\kappa)Y_{W,t}^{net}\,\varphi_{M1}^W \\ &= \frac{Y_{W,t}^{net}}{M_{pre}}\Bigg[a^W + f_1\,(\gamma_W^{heat}+\gamma_W) - \Bigg(\Bigg(\frac{f_3}{(\alpha_B\,z_B^g)^n} - f_2\Bigg)\gamma_W^{heat}\Bigg)\alpha_B\,z_B^g \\ &\quad -\Bigg(\Bigg(\frac{f_3}{(z_B^g)^n} - f_2\Bigg)\gamma_W - d_{BW}\Bigg)z_B^g\Bigg]\tilde{\phi}^W. \end{aligned} \tag{118}$$

Base model (Proposition 6). In the special case, where σ_B^A, σ_A^B, σ_A^O, σ_B^O are equal to zero, $\gamma_W^{heat} = 0$. The general equations (116), (117), and (118) for the SCC in the rest of the world turn to
(i): $S_t^A \neq 0$ and $S_t^B \neq 0$

$$\begin{aligned} SCC^W = \frac{Y_{W,t}^{net}}{M_{pre}}\Bigg[a^W + f_1\,\gamma_W - \Bigg(\Bigg(\frac{f_3}{z_B^n} - f_2\Bigg)\gamma_W - d_{BW}\Bigg)z_B \\ -\frac{\alpha_A(z_A - \alpha_B\,z_B)(d_{BW} - d_{AW})}{1-\alpha_A\,\alpha_B}\Bigg]\tilde{\phi}^W, \end{aligned}$$

with $z_A \in \{z_A^c, z_A^g\}$, $z_B \in \{z_B^c, z_B^g\}$, $d_{AA} \in \{d_{AA}^c, d_{AA}^g\}$, and $d_{BB} \in \{d_{BB}^c, d_{BB}^g\}$.
(ii): $S_t^A > 0$ and $S_t^B = 0$

$$SCC^W = \frac{Y_{W,t}^{net}}{M_{pre}}\Bigg[a^W + f_1\,\gamma_W - \Bigg(\Bigg(\frac{f_3}{(\alpha_A\,z_A^g)^n} - f_2\Bigg)\gamma_W - d_{AW}\Bigg)\alpha_A\,z_A^g\Bigg]\tilde{\phi}^W,$$

(iii): $S_t^A = 0$ and $S_t^B > 0$

$$SCC^W = \frac{Y_{W,t}^{net}}{M_{pre}}\Bigg[a^W + f_1\,\gamma_W - \Bigg(\Bigg(\frac{f_3}{(z_B^g)^n} - f_2\Bigg)\gamma_W - d_{BW}\Bigg)z_B^g\Bigg]\tilde{\phi}^W.$$