Solar Geoengineering in a Regional Analytic Climate Economy

Felix Meier and Christian Traeger

December 31, 2019

Abstract

The paper analyzes geoengineering and strategic interactions in an integrated assessment model (IAM) of climate change. For this purpose, we (i) derive a new class of solutions to analytic IAMs that allows us to (ii) solve an integrated assessment model with sulfur-based geoengineering and damages in closed form, and to (iii) model realistic strategic interactions between regions. Temperatures respond to carbon dioxide (standard carbon cycle), sulfur injections into the stratosphere (fitted to scientific data), and a potential counter-engineering agent that can offset some of the sulfur-based cooling. Damages arise from the increase in temperatures, the chemical agent(s) employed for geoengineering, and the modulation of the radiative energy balance through geoengineering. Our dynamic game involves two active players that are either partially or fully affected by the other region's geoengeneering measures and have the ability to contribute, remain inactive, or offset some of the other region's cooling measures. We shine new light on the "free-driver" problem popularized by Weitzman (2015), the climate-clash equilibria suggested by Heyen et al. (2019), a somewhat extreme sensitivity of geoengineering measures to potential damages, and the colloquial "slippery slope" argument showing how the active regions and the rest of the world respond to (some region's) availability of geoengineering measures. We discuss these findings using analytic solutions for the social cost of carbon (globally or regionally optimal carbon tax).

1 Introduction

Worldwide greenhouse gas emissions are still on the rise (Tollefson 2018) leading to potentially severe consequences for the world and its economy. Future warming will substantially reduce future output and may reduce global economic growth rates (Carleton and Hsiang 2016) apart from destroying ecosystems and driving species to extinction. In light of these developments, engineering a cooler climate remains a hot topic. The two most studied solar geoengineering techniques are stratospheric aerosol injections and marine cloud brightening (Boucher et al. 2017). We focus on the first measure, which has already been studied in natural experiments:

a series of volcanic eruptions, including Pinatubo in 1991, released large amounts of sulfur into the atmosphere and cooled the planet by scattering sunlight back to space (Crutzen 2006).

We build a full-fledged integrated assessment model with state of the art climate dynamics and an economy that transforms non-renewable and renewable resources into energy, final goods, and emissions. Introducing a new class of analytic solutions, we are able to integrate the temperature response to stratospheric sulfur as well as damages from geoengineering into our model, and to calibrate the cooling efficiency of sulfur well to recent scientific work of Kleinschmitt et al. (2018). We separate climate change damages into damages from rising temperatures and damages from increasing atmospheric carbon concentrations. Geoengineering implies a third type of damages: side-effects of the employed chemicals and changes in the radiative spectrum.

Our first results are closed-form solutions for the optimal sulfur-deployment strategy and the optimal carbon tax in a fully cooperating world with geoengineering. We explain the components that reduce earlier formulas for the optimal carbon tax because of our ability to cool the planet. The resulting reduction in the incentive to abate CO_2 is a delicate balance between the effectiveness of sulfur-based cooling and potential damages from geoengineering. Yet, the economics of solar geoengineering combines two externality problems with a regional conflict over targeting the right level of a global public good (or bad). We study the non-cooperative solutions in a dynamic IAM with two active players and a passive "rest of the world". Our study is among the first to analyze the strategic interaction of regions within an integrated assessment model of climate change.

We derive free-driving equilibria where one region determines the global temperatures, climate match equilibria where both active regions engage in cooling activities, and equilibria where one regions cools and the other region tries to offset the geoengineering agent (climate clash). In our model with heterogeneous regions and the rest of the world these equilibria are mutually exclusive and we explain the fundamental determinants giving rise to each of these situations. We obtain closed-form solutions for the regional geoengineering (Markov-) strategies and for the regionally optimal carbon taxes. We show how each region's geoengineering effort depends on the characteristics of both players. Also the mitigation incentives expressed through the SCC depend on the characteristics of both regions. It is straight-forward that there are scenarios where the availability of geoengineering reduces the SCC and, thereby, the mitigation efforts. The rest of the world can be both better off or worse off as a result depending on the precise spill-overs and its susceptibility to climate and geoengineering damages. Yet, it can also happen that the potential of geoengineering in other regions increases the mitigation efforts as a result of geoengineering damages. This scenario not only occurs in the climate clash equilibria, but also in cases where both active regions are cooling the planet. In fact, the availability of geoengineering can increase the mitigation effort in both active regions and the rest of the world.

The cooling effect of sulfur does not come at zero cost. Solar geoengineering is not able to

stabilize global precipitation and temperature simultaneously resulting in heterogeneous preferences regarding the optimal level (Ricke et al. 2010). It also leads to various side effects such as a reduction in the upper ozone layer (Heckendorn et al. 2009) or acid precipitation and deposition (Crutzen 2006). However, annual injection rates discussed are rather low compared to already existing anthropogenic and natural inputs of about 136 TgS/yr (Kravitz et al. 2009). Our model introduces differential damages between sulfur-based cooling and counterengineering as well as costs for the injection of these chemical agents into the stratosphere. We show that such a careful cost distinction has important implications for the equilibrium strategies of different regions. Stratospheric agents travel from the tropics to the poles and spread quickly across longitude. As a consequence, it is impossible to do regional climate management using stratospheric engineering without major and partly perfect spillovers to the other regions. Other less cooling. Our model analyzes carefully how limitations in and opportunities of avoiding spill-overs affects the game.

Heutel et al. (2016) argue that solar geoengineering should be part of a welfare maximizing policy for two reasons. It allows to control the mean surface temperature at (potentially) lower costs than mitigation and reduces the risk of reaching climate tipping points. For their numerical simulations, the authors assume damages from geoengineering of 3 percent of gross output for a reduction in radiative forcing to pre-industrial levels. The authors consider this a conservative guess (biased against solar geoengineering) since damages are still largely unknown. In a recent study, Proctor et al. (2018) estimate agricultural damages from solar geoengineering using data from volcanic eruptions. They find two competing effects. The cooling from solar geoengineering has an increasing effect on crop yields. Scattering light on the one hand decreases total available sunlight which affects photosynthesis negatively but on the other hand influences photosynthesis positively since the fraction of diffuse light increases. Their results suggest that in the mid-twenty-first century the positive effect of cooling on crop yields is completely offset by a negative insolation effect. Heutel et al. (2018) look at how uncertainty with respect to the damages of geoengineering alters optimal policy. They find that introducing uncertainty lowers the level of geoengineering and increases the level of abatement to compensate for the reduction. Numeric assessment of the uncertainty in regional (including strategic) models is severely challenged by the curse of dimensionality in dynamic programming. Our analytic model will be able to integrate uncertainty into the high-dimensions space of a regional integrated assessment model, even though we currently have to leave it for future work (maybe by the time of the conference).

Solar geoengineering has been titled "symptomatic approach" since it does not address the consequences from increasing atmospheric carbon concentrations (Klepper and Rickels 2014). Therefore, it can never fully substitute for mitigation. Increasing atmospheric carbon concentrations lead to damages from ocean acidification but also create benefits due to an increase in the rate of photosynthesis in plants (fertilization effect) (Moreno-Cruz and Smulders 2017). Another reason for limited substitutability is explored in Emmerling and Tavoni (2018a). The

paper analyzes uncertainty governing the effectiveness of solar geoengineering. The authors find a decreasing concave relationship between today's mitigation effort and the probability of success of geoengineering.

Moreno-Cruz (2015) studies how solar geoengineering impacts the free riding effect on mitigation. The option of geoengineering is explored in a cost-minimizing set up with two sequential stages. In the first stage each country chooses its optimal level of mitigation. In the second stage countries select the optimal level of geoengineering. The paper shows that the impact of geoengineering depends on the similarity between countries. Assuming similar countries (i.e. similar with respect to damages from geoengineering and climate change) the option of geoengineering leads to lower mitigation levels in both countries. When countries differ with respect to damages, mitigation in both countries can increase.

Recently, another risk of solar geoengineering has gathered attention. The so called "free driver" incentive (Weitzman 2015). Because geoengineering affects the global temperature, regions have conflicting target levels of the global temperature level. Operational costs for stratospheric sulfur injections are presumably low (Smith and Wagner 2018, McClellan et al. 2012) in the sense that many countries would be capable to deploy geoengineering unilaterally. This has raised the concern that a country or a club of countries might implement solar geoengineering at high levels to achieve its optimal climate at the expense of others (Pasztor et al. 2017). Emmerling and Tavoni (2018b) quantify the free-riding effect in an integrated assessment model and show that without cooperation geoengineering is used at inefficiently high levels. The option of counter-geoengineering has been discussed as a possible solution to this problem. Heyen et al. (2019) find that the free-driver outcome becomes unstable once counter-geoengineering is available. However, the option of counter-geoengineering might instead lead to a "climate clash" when no moratorium treaty (countries abstain from climate engineering) and no cooperative deployment is realized. Their outcome strongly depends on the degree of asymmetry in temperature preferences. There are different proposals for counter-geoengineering. One idea is to countervail the cooling effect from sulfur aerosols by the additional release of greenhouse gases such as sulphur hexafluoride, chlorofluorocarbons or hydrochlorofluorocarbons. Another idea is to inject a base into the stratosphere that decreases or even neutralizes the cooling effect of the aerosols (Parker et al. 2018). We will take the strategic interactions governing climate engineering from highly stylized and often static models into a full-fletched dynamic integrated assessment model of climate change.

2 Global Model

This section introduces geoengineering into the analytic climate economy model ACE (Traeger 2018). First, we summarize a slightly simplified version of the ACE model. Then, we introduce geoengineering 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 geoengineering makes

for the optimal carbon tax.

2.1 Economy

Production. Gross output is a function of vectors of dimension I_j with $j \in \{A, N, K, E\}$. The technology levels $\boldsymbol{A}_t$ are exogenous. Capital is optimally distributed of the different sectors, resulting in the capital levels summarized by the vector $\boldsymbol{K}_t$. Labor distribution is $\boldsymbol{N}_t$ and energy inputs by $\boldsymbol{E}_t$.

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

The production function is homogeneous of degree κ in capital such that

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

We note that aggregate capital is $K_t = \sum_{i=1}^{I_K} K_{i,t}$ and the share of capital in industry i is $\mathcal{K}_{i,t} = \frac{K_{i,t}}{K_t}$. Population size is normalized to unity $\sum_{i=1}^{I_N} N_{i,t} = 1$.

Damages. We denote the atmospheric carbon stock (or concentration) by $M_{1,t}$. It is convenient to measure it relative to the pre-industrial stock level as $m_t = \frac{M_{1,t}}{M_{\text{pre}}}$. Global atmospheric temperature $T_{1,t}$ measures the temperature increase over 1900 levels (in degree Celsius). Atmospheric temperature increase, atmospheric carbon concentration, as well as the level of the cooling agent S_t cause (net) damages $D_t(T_{1,t}, S_t, m_t)$ that we measure as a fraction of output. They are of the form

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

We take the temperature-based damages

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

from the ACE model's calibration to DICE. Our global model assumes

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

making d the semi-elasticity of damages from stratospheric sulfur injections (the percentage loss of output resulting from an additional ton of sulfur injections). The parameter d includes linear operational costs. Our regional version of the model in section 3 further refines the damage parameter into damages from geoengineering, counter-geoengineering, and the costs of

injecting the chemical agents into the stratosphere. The net costs of atmospheric carbon are

$$D_m(m_t) = a\,(m_t - 1),\qquad(4)$$

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

Emissions, resources and capital. The first I^d energy inputs $E_1, ..., E_{I^d}$ are fossil fuels causing CO_2 emission and are collected in subvector $\boldsymbol{E}_t^d$. These energy inputs are measured in terms of their CO_2 content so that total emissions from production are $\sum_{i=1}^{I^d} E_{i,t}$. Fossil fuel resource stocks are collected in vector $\boldsymbol{R}_t \in \mathbb{R}_+^{I^d}$. The dynamics of the resource stock are

$$\boldsymbol{R}_{t+1} = \boldsymbol{R}_t - \boldsymbol{E}_t^d \qquad(5)$$

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

We assume full depreciation of capital over the course of a decade, the model's time step. We choose this simplifying assumption noting that Traeger's (2018) extension for capital persistence would also go through in our setting. The global economy's capital stock evolves as

$$\begin{aligned}
K_{t+1} &= Y_t\,[1 - D_t\,(T_{1,t}, S_t, m_t)] - C_t \qquad(6)\\
&= Y_t\,\exp\left[-\xi_0\,\exp\left(\xi_1\,T_{1,t}\right) + \xi_0 - d\,S_t - a(m_t - 1)\right] - C_t.
\end{aligned}$$

2.2 Climate

Carbon dioxide. We consider two carbon reservoirs, atmosphere (carbon content M_1) and ocean (carbon content M_2). The extension to additional carbon reservoirs is straight-forward. The dynamics of the carbon reservoirs is

$$\begin{pmatrix} M_{1,t+1} \\ M_{2,t+1} \end{pmatrix} = \begin{pmatrix} \phi_{11} & \phi_{21} \\ \phi_{12} & \phi_{22} \end{pmatrix} \begin{pmatrix} M_{1t} \\ M_{2t} \end{pmatrix} + \begin{pmatrix} E_t^{tot} \\ 0 \end{pmatrix} \qquad(7)$$

with transition matrix Φ and total CO_2 emissions $E_t^{tot} = \sum_{i=1}^{I^d} E_{i,t} + E_t^{\text{exo}}$ resulting from industrial fossil fuel burning and other exogenous processes including lang use change and forestry.

Greenhouse effect and cooling. Greenhouse gases in the atmosphere as well as geo-engineering affect out planet's temperature. The net heating with respect to preindustrial times is summarized by the resulting radiative forcing F_t (measured in W/m^2). It increases (logarithmically) in the increase of atmospheric CO_2 relative to preindustrial levels m_t and it falls as a consequence of geoengineering measures $G_t(S_t)$ that inject sulfur aerosols S_t into the

stratosphere

$$F_t = \frac{\eta}{\log 2} \log(m_t) - G_t(S_t)$$

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

$$= \frac{\eta}{\log 2} \log\left(m_t \exp\left(-\frac{\log 2}{\eta} G_t(S_t)\right)\right) \tag{8}$$

We ignore non-CO_2 greenhouse gases, which can easily be added as in the ACE model. The next subsection will fit a data-based approximation to equation (8) that will replace the generic formulation above.

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 temperatures $T_{2,t}$ evolves as a generalized mean of own past and atmospheric temperature

$$T_{1,t+1} = \frac{1}{\xi_1} \log\left((1 - \sigma_{01} - \sigma_{21}) \exp\left(\xi_1 T_{1,t}\right) + \sigma_{01} \exp\left(\xi_1 T_{0,t}\right) + \sigma_{21} \exp\left(\xi_1 T_{2,t}\right)\right) \tag{9a}$$

$$T_{2,t+1} = \frac{1}{\xi_2} \log\left((1 - \sigma_{12}) \exp\left(\xi_2 T_{2,t}\right) + \sigma_{12} \exp\left(\xi_2 T_{1,t}\right)\right) \tag{9b}$$

with $\xi_1 = \xi_2 = \frac{\log 2}{s}$. Defining $\sigma_{01} = \sigma_{\text{forc}}$, we rewrite these equations in terms of transformed temperatures $\tau_{it} = \exp(\xi_i T_{i,t})$:

$$\begin{pmatrix} \tau_{1,t+1} \\ \tau_{2,t+1} \end{pmatrix} = \begin{pmatrix} 1 - \sigma_{01} - \sigma_{21} & \sigma_{21} \\ \sigma_{12} & 1 - \sigma_{12} \end{pmatrix} \begin{pmatrix} \tau_{1,t} \\ \tau_{2,t} \end{pmatrix} + \begin{pmatrix} \sigma_{\text{forc}} \exp(\frac{\log(2)}{\eta} F_t) \\ 0 \end{pmatrix} \tag{10}$$

2.3 Geoengineering

Best on the experience of many volcanic eruptions scientists have learned that the injection of small sulfur particles (aerosols) into the atmosphere reflects sunlight back into space cooling our planet. Yet, at high injection rates, sulfur particles lump together which decreases their cooling efficiency. As a result, scientists expect that the cooling from stratospheric aerosol injections has an asymptotic limit (Lawrence et al. 2018). The uncertainty governing the forcing efficiency of sulfur injections is very high (Kleinschmitt et al. 2018). Niemeier and Timmreck (2015) find an exponential relationship between sulfur injections and radiative forcing. Their model reaches a negative instantaneous radiative forcing effect of -6 W/m^2 with an injection rate of 45 TgS/yr. Niemeier and Schmidt (2017) find that Niemeier and Timmreck (2015) overestimate the forcing efficiency of sulfur and conclude that it would require injection rates of 70 TgS/yr to reach a cooling of -6 W/m^2. Kleinschmitt et al. (2018) find that it might be

impossible to achieve more cooling than a negative instantaneous radiative forcing of -2 W/m^2. In a recent literature review, Lawrence et al. (2018) conclude that the cooling potential from sulfur injections ranges between -2 and -5 W/m^2. The effective radiative forcing effect, which also includes rapid adjustments such as changes in atmospheric temperature, is larger than the instantaneous radiative forcing effect (Boucher et al. 2017). The cooling potential does not only depend on the injection rate but also on the location of the injections. Recent studies suggest that it might be possible to optimize the geographic distribution of the cooling by varying the altitude and latitude of injections (Jones et al. 2018, MacMartin et al. 2017, Kravitz et al. 2017). In addition, the literature has 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 is even more limited.

We calibrate our model to the recent study by Kleinschmitt et al. (2018).[1] For this purpose,

Table 1: Radiative forcing effect (Kleinschmitt et al. 2018)

2 TgS	5 TgS	10 TgS	20 TgS	50 TgS
-1.11	-1.64	-2.91	-4.34	-5.63

we develop a new functional form with several degrees of freedom that will permit an analytic solution of the dynamic programming problem. We approximate radiative forcing (equation 8) for the use of stratospheric sulfur injections by

$$F_t = \frac{\eta}{\log(2)} \log \left(\underbrace{f_0 + f_1 \, m_t + \left(f_2 - f_3 \left(\frac{m_t}{S_t} \right)^n \right) S_t}_{\equiv F_t^c} \right). \tag{11}$$

The expression F_t^c characterizes the effect of sulfur on radiative forcing in CO_2 equivalents. In the absence of geoengineering, the term m_t would capture the CO_2 forcing. The round inner bracket reduces the forcing in response to sulfur injections. The main contribution derives from the term $f_3 \left(\frac{m_t}{S_t} \right)^n$. Sulfur forcing is more efficient relative to CO_2 the larger the atmospheric carbon concentration and the lower the 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 on effective radiative forcing from sulfur injections (Table 1).[2] Our fit combines Kleinschmitt et al.'s (2018) forcing data for sulfur

---

[1]Note that the publication only cites the direct radiative forcing impact of sulfur. We obtained the total radiative forcing in Table 1 from the authors in personal correspondence.

[2]The new functional form can also be calibrated to data from other studies that deal with the radiative forcing effect from stratospheric sulfur injection (see Appendix A).

injections with and 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 emissions. Our fit minimizes the squared differences for those 80 data points. We list the resulting parameters in Table 2.

Table 2: Estimated forcing parameters

f_0	f_1	f_2	f_3	n
0.228	1.83	0.00403	1.12	0.869

In Appendix A, we show radiative forcing resulting from a given annual flow of sulfur injections at a given atmospheric carbon dioxide concentration in a 3D graph. Figure 1 illustrates the goodness of our fit, slicing the 3D graph in the two dimensions and adding the data points from Kleinschmitt et al. (2018). We assume that total radiative forcing remains positive (above preindustrial levels). Based on our empirical fit of the radiative forcing equation we take the following assumption.

Assumption 1. *The (fit-)parameters f_i, $i \in \{0, ..., 3\}$, are positive and $0 < n < 1$.*
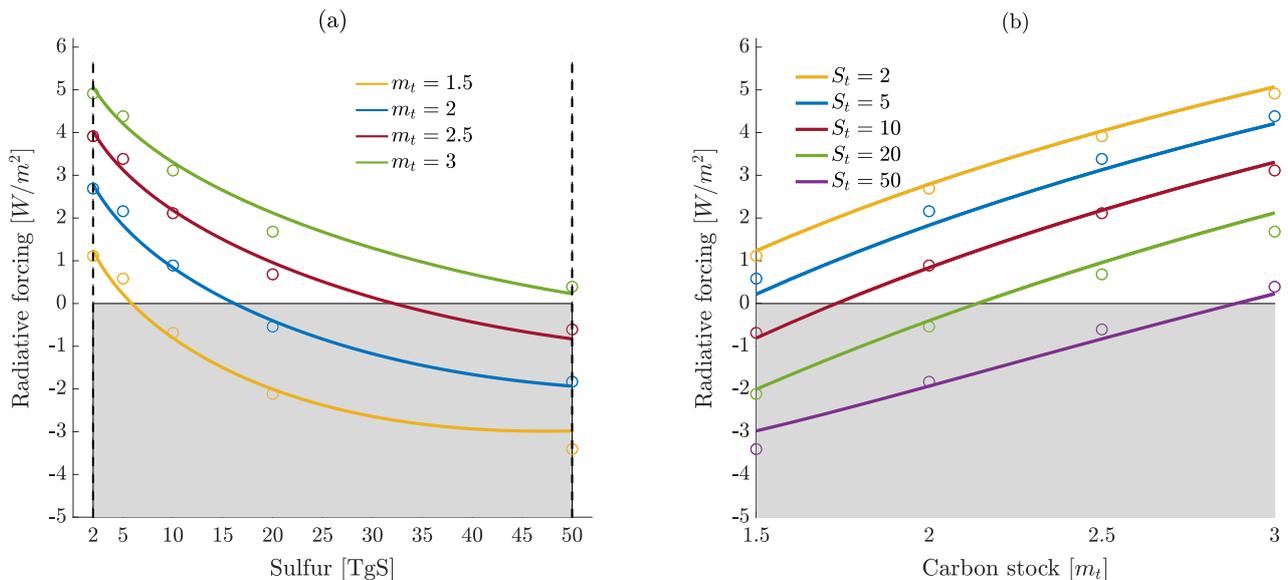
Figure 1: Approximation to model data from Kleinschmitt et al. (2018). The forcing effect from sulfur for different carbon stocks is shown in (a). The forcing effect from the carbon stock for different injection rates is shown in (b).

Operational costs and damages of geoengineering Several studies have investigated the operational costs of solar geoengineering. Table 3 shows recent cost estimates on stratospheric sulfur injections by newly designed airplanes. Estimates are given for a reduction in radiative forcing (W/m^2) or the quantity of sulfur injected into the stratosphere (Mt). Stars in Table 3 denote the original values from the study.

Current assessments on operational costs are by far more reliable than assessments on economic impacts from geoengineering (e.g. from acid rain). Insufficient information does not

Table 3: Operational costs of stratospheric sulfur injections

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

allow to quantify the damages from solar geoengineering. Some authors, for example Moreno-Cruz and Keith (2013), have analyzed optimal policy as a function of the damage parameter. Others have made explicit assumptions, but acknowledge that there is only limited or non-existent empirical bases. We show several of those estimates in Table 4.

Table 4: Damages from solar geoengineering

Authors	Best guess
Emmerling and Tavoni (2018b)	Consumption loss of 3% compensating each every 3.5 W/m^2 of forcing
Goes et al. (2011)	GDP loss between 0 and 5% per forcing equivalent to a doubling CO_2 forcing
Heutel et al. (2016; 2018)	GDP loss of 3% for setting forcing back to the pre-industrial level

2.4 Global Planner Solution

In the present section, a social planner maximizes the infinite stream of consumption flows

$$\max_{C_t, E_t, S_t} \sum_{t=0}^{\infty} \beta^t \log(C_t) \tag{12}$$

subject to equations (1), (5), (6), (7), and (10). The parameter β denotes the utility discount factor (pure time preference). In the regional model, each region follows the analogous objective for their own region's welfare (no trade). In Appendix B we solve the inter-temporal optimization problem and derive the global optimal level of sulfur deployment.

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

$$S_t^* = \left(\frac{(1-n)\,\gamma\,f_3}{d + \gamma\,f_2} \right)^{\frac{1}{n}} m_t \tag{13}$$

with geoengineering propensity $z = \left(\frac{(1-n)\,\gamma\,f_3}{d+\gamma\,f_2} \right)^{\frac{1}{n}}$, climate change impact $\gamma = \beta\,\xi_0\,\tilde{\sigma}_{11}\,\sigma_{forc}$, and temperature dynamics contribution $\tilde{\sigma}_{11} = \left(1 - \beta(1 - \sigma_{01} - \sigma_{21}) - \frac{\beta^2\,\sigma_{21}\,\sigma_{12}}{1-\beta(1-\sigma_{12})} \right)^{-1}$.

Proof. See Appendix B.2. $\qquad\square$
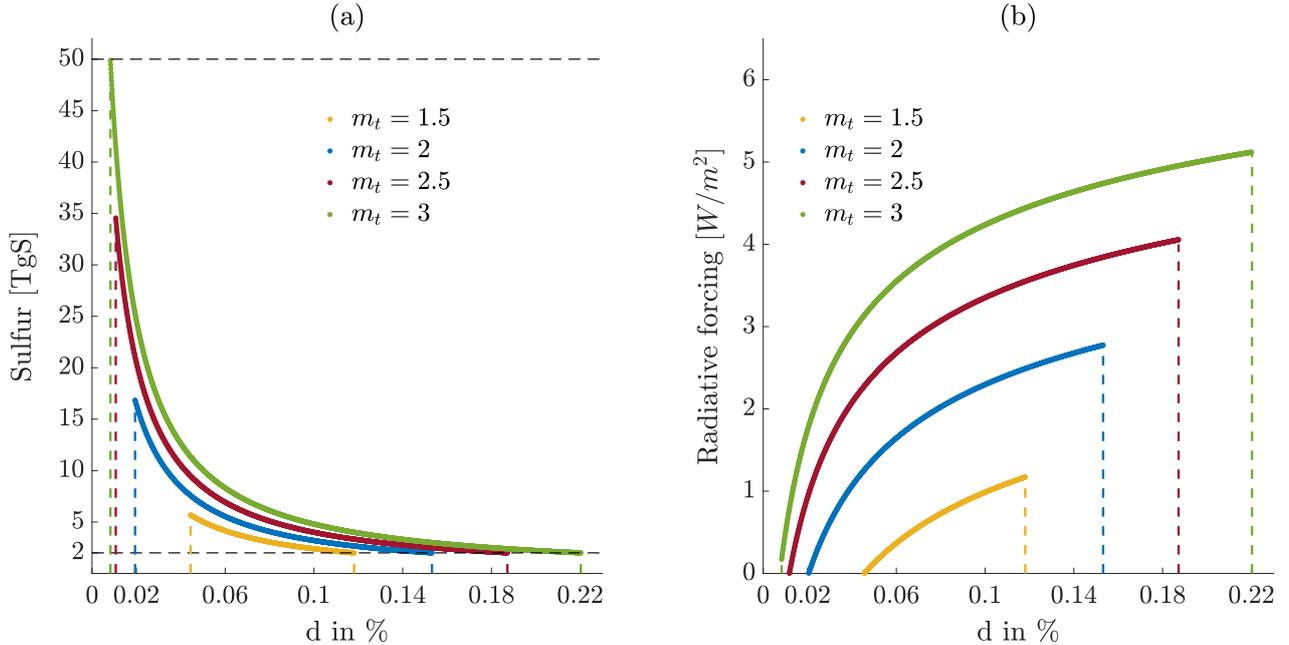
Figure 2: (a) Optimal level of sulfur and (b) optimal level of radiative forcing as a function of the damage parameter d (in %).

The optimal level of sulfur deployment increases linearly in the atmospheric carbon content. We call the proportionality factor the geoengineering propensity. It characterizes the drivers and moderators of the cooling effort given atmospheric CO_2. 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 geoengineering damages (d) and the non-linear efficiency loss of sulfur cooling (n).

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

$$S_t^* = \left(\frac{1.6}{4\% + 10^3 d} \right)^{1.15} m_t, \tag{14}$$

as a function of the damage (semi-)elasticity of sulfur, and the atmospheric carbon concentration (relative to pre-industrial levels). The calibrated formula shows that the optimal sulfur deployment is extremely sensitive to damages from geoengineering. Figure 2(a) shows the optimal level of sulfur as a function of the damage parameter d for different carbon stocks. Inserting S_t^* into equation (11) yields the optimal level of radiative forcing as a function of the damage parameter d and the atmospheric carbon stock, which we show in Figure 2(b). The higher the damages, the higher the tolerated forcing and, thus, warming levels. Given the efficiency loss of sulfur-based cooling, the tolerated hearing does not increase as sharply in the damages level as the sulfur injections are falling. We restrict our model to a "well-calibrated" region defined by the interval $[\underline{d}(m_t), \bar{d}(m_t)]$: the boundaries of the interval depend on the carbon stock $m_t \in [1.5, 3]$ and they result from (i) our assumption that radiative forcing remains posi-

tive, and (ii) that we approximate radiative forcing for sulfur levels between 2 and 50 TgS. For $d < \underline{d}(m_t)$, it is either the case that $S_t > 50$ TgS or $F_t < 0$. For $d > \bar{d}(m_t)$ it must be that $S_t < 2$.

We now study how the social cost of carbon is affected by the availability of solar geoengineering.

Proposition 2. *The Social Cost of Carbon in money-measured consumption equivalents is*

$$SCC = \frac{Y_t^{net}}{M_{pre}} \left[a + \gamma \, f_1 - \frac{n}{1-n} \left(d + \gamma \, f_2 \right) z \right] \tilde{\phi} \tag{15}$$

with the carbon dynamics contribution $\tilde{\phi} = \left(1 - \beta \, \phi_{11} - \frac{\beta^2 \, \phi_{12} \, \phi_{21}}{1 - \beta \, \phi_{22}} \right)^{-1}$ and, as above, the geoengineering propensity $z = \left(\frac{(1-n)\,\gamma\,f_3}{d + \gamma\,f_2} \right)^{\frac{1}{n}}$ and the climate change impact $\gamma = \beta \, \xi_0 \, \tilde{\sigma}_{11} \, \sigma_{forc}$.

Proof. See Appendix B.3. □

The fraction $\frac{Y_t^{net}}{M_{pre}}$ sets the scale and units of the SCC. The square brackets characterizes 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 the term γ in the square bracket reflecting the cost resulting from temperature increase in the absence of climate engineering. Formula (15) adds the term a reflecting the direct net damages from atmospheric CO_2 caused by ocean-acidification net of the land-based fertilization effect. The final term reduces the SCC as a result of the available geoengineering. This reduction is proportional to the cooling propensity z, and is amplified by the increase in sulfur-based cooling efficiency for high levels of CO_2 measured by n as well as the geoengineering damages. The combined expression might be more easily interpreted in the form

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

noting that $n < 1$. This transformation shows more clearly that the (negative) adjustment of the optimal carbon tax increases in sulfur-based cooling efficiency f_3 and falls with geoengineering damages d. However, the SCC adjustment falls much slower in the damages than the geoengineering propensity itself.

Following ACE, we use a time step of 10 years and use the parameter specification summarized in Table 5. Together with our radiative forcing estimates from Table 2 we obtain the

Table 5: Parameter values from ACE (Traeger 2018)

Y_t^{net}	M_{pre}	β	ξ_0	σ_{forc}	$\tilde{\sigma}_{11}$	$\tilde{\phi}$
$135 \, 10^{13}$	$3.667\,600\,10^9$	0.986^{10}	0.021	0.54	1.1	4.3

optimal carbon tax in (USD-2018-) money-measured consumption equivalents as a function of

the damage parameters a and d (see Appendix B.3)

$$SCC = 613 \left[a + 2\% - 6.65 \left(\frac{1.6}{(4\% + 10^3 d)^{0.13}} \right)^{1.15} \right] 4.3 \qquad (16)$$

Figure 3 graphs this SCC as a function of the geoengineering damages d for a given damage parameter $a = 0$ (no direct damages from an increase in the atmospheric carbon stock) and $a = 0.1\%$.
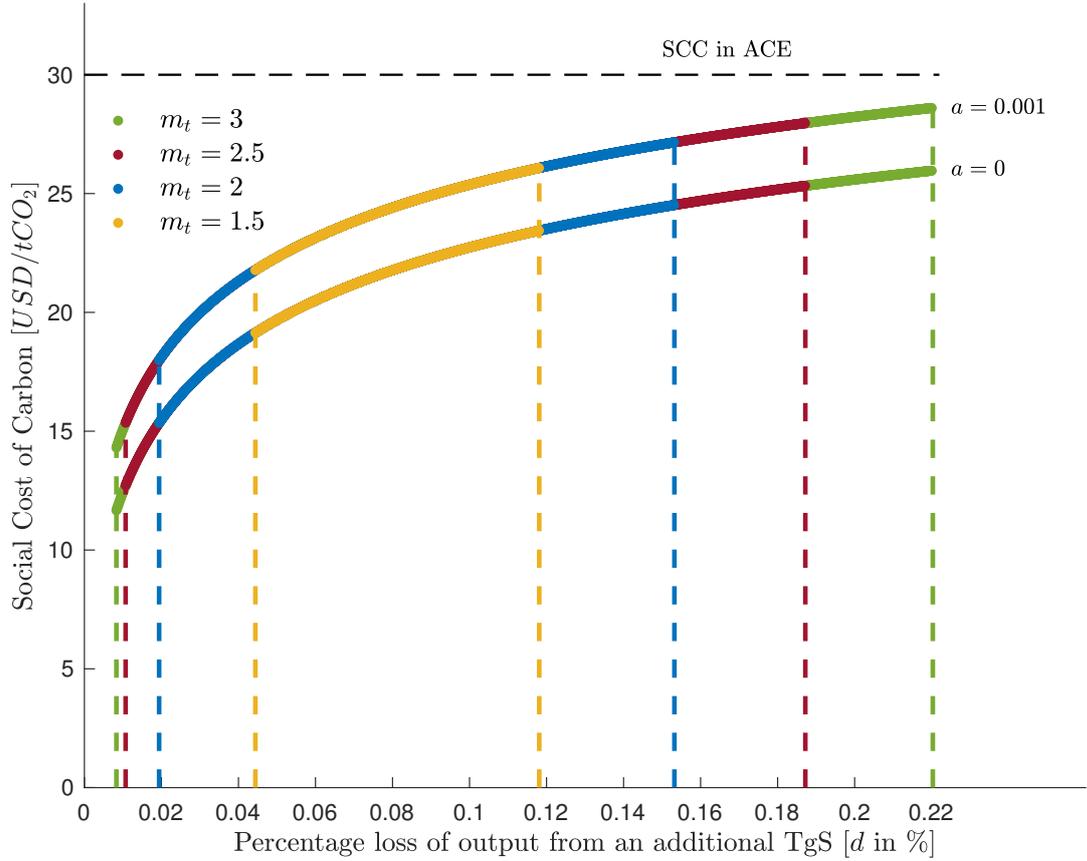
Figure 3: Social cost of carbon as a function of the damage parameter d

3 Non-Cooperative Regional Model

We now split the world into regions that act non-cooperatively. We focus on two active regions A and B that consider engaging in climate engineering, either sulphur-based cooling or counter-geoengineering. The rest of the world only affects the decisions of regions A and B through their contributions to the global carbon stock.

3.1 Regional Model

This section explains the changes required to split up the world in several economic and climate regions. It also introduces more detailed damages specifications and the option to engage in counter-geoengineering.

Regional economies. The regional economies follow equations (1-6), where functional forms and parameters are idiosyncratic to the regions.

Regional climate. CO_2 mixes globally and the CO_2 dynamics are still described by equation (7). However, the total CO_2 emissions are now resulting from region A, region B, and the rest of the world W:

$$E_t^{tot} = \sum_{i=1}^{I^d} E_{A,i,t} + \sum_{i=1}^{I^d} E_{B,i,t} + \sum_{i=1}^{I^d} E_{W,i,t} + E_t^{\text{exo}} \tag{17}$$

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 temperature zone. Analogously to the global case, we let $T_{0,t}^A = \frac{s}{\eta} F_t^A$ and $T_{0,t}^B = \frac{s}{\eta} F_t^B$ and the two regional atmospheric temperatures and the ocean temperature evolve as

$$T_{1,t+1}^A = \frac{1}{\xi_1} \log\left((1 - \sigma^A) \exp\left(\xi_1 T_{1,t}^A\right) + \sigma_{\text{forc}} \exp\left(\xi_1 T_{0,t}^A\right) + \sigma_B^A \exp\left(\xi_1 T_{1,t}^B\right) + \sigma_O^A \exp\left(\xi_1 T_{2,t}\right)\right) \tag{18a}$$

$$T_{1,t+1}^B = \frac{1}{\xi_1} \log\left((1 - \sigma^B) \exp\left(\xi_1 T_{1,t}^B\right) + \sigma_{\text{forc}} \exp\left(\xi_1 T_{0,t}^B\right) + \sigma_A^B \exp\left(\xi_1 T_{1,t}^A\right) + \sigma_O^B \exp\left(\xi_1 T_{2,t}\right)\right) \tag{18b}$$

$$T_{2,t+1} = \frac{1}{\xi_2} \log\left((1 - \sigma^O) \exp\left(\xi_2 T_{2,t}\right) + \sigma_A^O \exp\left(\xi_2 T_{1,t}^A\right) + \sigma_B^O \exp\left(\xi_2 T_{1,t}^B\right)\right), \tag{18c}$$

where $\sigma^A = \sigma_B^A + \sigma_O^A$, $\sigma^B = \sigma_A^B + \sigma_O^B$, and $\sigma^O = \sigma_A^O + \sigma_B^O$. In terms of transformed temperatures the system of equations simplifies to the matrix equation

$$\begin{pmatrix} \tau_{1,t+1}^A \\ \tau_{1,t+1}^B \\ \tau_{2,t+1} \end{pmatrix} = \begin{pmatrix} 1 - \sigma^A & \sigma_B^A & \sigma_O^A \\ \sigma_A^B & 1 - \sigma^B & \sigma_O^B \\ \sigma_A^O & \sigma_B^O & 1 - \sigma^O \end{pmatrix} \begin{pmatrix} \tau_{1,t}^A \\ \tau_{1,t}^B \\ \tau_{2,t} \end{pmatrix} + \begin{pmatrix} \sigma_{\text{forc}} \exp(\frac{\log 2}{\eta} F_t^A) \\ \sigma_{\text{forc}} \exp(\frac{\log 2}{\eta} F_t^B) \\ 0 \end{pmatrix} \tag{19}$$

Regional forcing. Radiative forcing in region A is a function of global atmospheric carbon and the engineering undertaken in the two regions. Let $\tilde{S}_t^B = \alpha_B S_t^B$ and $\tilde{S}_t^A = \alpha_A S_t^A$ denote the spill-over of the direct cooling or the counter-engineering from one region to the other. The parameter α_A (α_B) determines the share of region A's (B's) injection of the cooling or counter-geoengineering agent that travels to region B (A). The magnitude of the α parameter depends in particular on the geographic location of the two regions. For example, the α parameters will be relatively 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 perfect spill-over from the US, but the US much less spill-over from Canada). Then we have the regional forcing levels

$$F_t^A = \frac{\eta}{\log(2)} \log \left[f_0 + f_1 m_t + \left(f_2 - f_3 \left(\frac{m_t}{(S_t^A + \tilde{S}_t^B)} \right)^n \right) (S_t^A + \tilde{S}_t^B) \right] \qquad (20)$$

for region A and for region B

$$F_t^B = \frac{\eta}{\log(2)} \log \left[f_0 + f_1 m_t + \left(f_2 - f_3 \left(\frac{m_t}{(S_t^B + \tilde{S}_t^A)} \right)^n \right) (S_t^B + \tilde{S}_t^A) \right]. \qquad (21)$$

We assume that counter-geoengineering ($S_t < 0$) can only be used to offset (part of) the cooling imposed by the other region and, therefore, $S_t^A + \tilde{S}_t^B > 0$ and $S_t^B + \tilde{S}_t^A > 0$.

Regional damages. Geoengineering creates damages and operational costs which, for region A, we summarize in d_{AA} for the marginal costs of the region's own action, and d_{BA} for the marginal damages imposed on region A by region B. Damages as a fraction of output in region A are given by

$$D_t^A(\tau_{1,t}^A, S_t, m_t) = 1 - \exp \left[\xi_0^A \left(1 - \tau_{1,t}^A \right) - (d_{AA} S_t^A + d_{BA} \tilde{S}_t^B) - a^A (m_t - 1) \right]. \qquad (22)$$

Region A's self-imposed marginal costs d_{AA} depend on whether it is cooling or conducting counter-geoengineering

$$d_{AA} = \begin{cases} d_{AA}^g + \epsilon_A^g & \text{for} \quad S_t^A > 0 \\ d_{AA}^c - \epsilon_A^c & \text{for} \quad S_t^A < 0 \\ 0 & \text{for} \quad S_t^A = 0 \end{cases} \qquad (23)$$

where d_{AA}^g is the damage from sulfur-based cooling and ϵ_A^g 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 counter-engineering, and ϵ_A^c is the cost of counter-geoengineering. For the damages imposed by region B onto region A we distinguish whether region B engages in sulfur-based cooling or counter-engineering

$$d_{BA} = \begin{cases} d_{BA}^g & \text{for} \quad \tilde{S}_t^B > 0 \\ d_{BA}^c & \text{for} \quad \tilde{S}_t^B < 0 \\ 0 & \text{for} \quad \tilde{S}_t^B = 0 \end{cases} \qquad (24)$$

where d_{BA}^c is again a partial offsetting of the damages from region A's climate engineering in case region B is countering it. This damage reduction from counter-engineering can at most

offset the original damages: $d_{AA}^c \leq d_{BA}^g$ and $d_{BA}^c \leq d_{AA}^g$. While the direct radiative forcing damage will be offset, the damages caused by the chemical agent sulfur will probably not be offset, only partially offset, or maybe even enhanced by the counter-geoengineering agent. Thus d_{AA}^c will generally be strictly lower than d_{AA}^g in the real world (and it could potentially even be negative).

Analogously, for region B we define damages as a fraction of output

$$D_t^B(\tau_{1,t}^B, S_t, m_t) = 1 - \exp\left[\xi_0^B\left(1 - \tau_{1,t}^B\right) - (d_{BB}S_t^B + d_{AB}\tilde{S}_t^A) - a^B(m_t - 1)\right],\qquad (25)$$

with

$$d_{BB} = \begin{cases} d_{BB}^g + \epsilon_B^g & \text{for} \quad S_t^B > 0 \\ d_{BB}^c - \epsilon_B^c & \text{for} \quad S_t^B < 0 \\ 0 & \text{for} \quad S_t^B = 0 \end{cases} \qquad (26)$$

and

$$d_{AB} = \begin{cases} d_{AB}^g & \text{for} \quad \tilde{S}_t^A > 0 \\ d_{AB}^c & \text{for} \quad \tilde{S}_t^A < 0 \\ 0 & \text{for} \quad \tilde{S}_t^A = 0. \end{cases} \qquad (27)$$

We make the following restriction on the damage parameters in the active regions.

Assumption 2. *Within a region, the damage relieve from counter-geoengineering is smaller than the damage caused by geoengineering: $d_{AA}^c < d_{AA}^g$ and $d_{BB}^c < d_{BB}^g$.*

3.2   Results of the Base Model

To simplify the exposition of the main results we turn off the direct heat transfer between the regions.

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

As a result, the regional climates interact only through the spill-over of the cooling and, potentially, counter-geoengineering agents. The assumption simplifies the functional expressions in the results without changing the qualitative results (as we show later/in the Appendix).

The section first identifies a set of equilibrium strategies, second characterizes the domain of the resulting subgame perfect Nash-equilibria and, third, derives a formula for the resulting non-cooperative SCC levels in the different regions. We spell out and solve the two simultaneous and interacting dynamic programming problems of regions A and B in Appendix C. We thereby solve the dynamic Markov game which results in a subgame perfect equilibirum. It is sometimes referred to as a feedback solution of the game (as opposed to an open loop equilibrium). The regions foresee how, in the future, the other region responds to its action.

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

$$S_t^A = \frac{m_t}{1 - \alpha_A \, \alpha_B} \left(z_A^g - \alpha_B \, z_B \right) \quad for \quad S_t^A > 0 \qquad (28)$$

$$S_t^A = \frac{m_t}{1 - \alpha_A \, \alpha_B} \left(z_A^c - \alpha_B \, z_B \right) \quad for \quad S_t^A < 0 \qquad (29)$$

$$S_t^A = 0 \quad otherwise.$$

with geoengineering propensity and counter-geoengineering aversion

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

and climate change impact $\gamma_A = \beta^A \, \xi_0^A \, \tilde{\sigma}_A^{-1} \, \sigma_{forc}$. Swapping country indices characterizes region B's strategies.

Proof. See Appendix C.1.1. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

The game solves in linear strategies. In the case where one of the regions remains inactive, the other region's optimal cooling effort is structurally the same as that of the social planner in Proposition 2. In difference 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 geoengineering (damages and costs were combined into a single term d in the social planner's problem).

In the case where both regions engage in cooling, their optimal strategies incorporate the other region's efforts. Region B responds to the share α_A of A's cooling that spill over to its territory. Region A responds to region B's share of cooling α_B that spill over to its territory and, thereby, also responds to B's response to its own action. Given that each region responds proportionally to the other region with factors α_i, $i \in \{A, B\}$, the feedback gives rise to the multiplier $\frac{1}{\alpha_A \alpha_B}$ – somewhat like a Keynesian multiplier in macroeconomics. Yet, despite the cooling reinforcement, the fact that both regions are cooling also reduces region A's incentive to cool, see equation (28) with $z_B = z_B^g$. Therefore, region A's base propensity to cool, which by itself would be z_A^g, is reduced by the cooling $\alpha_B z_B^g$ that arrives in equilibrium from the other region. If countries and spill-overs are fully symmetric, region A's optimal sulfur deployment will always be lower than when it acts alone.

In the case where the regions' interests are clashing, one region, say region B, is cooling while region A is deploying the counter-geoengineering agent (equation 29 with $z_B = z_B^g$). Now z_A^c no longer captures region A's engineering propensity (z_A^g), but its reluctance to engage in counter-geoengineering ($S_t^A < 0$). This reluctance increases with the effectiveness of geoengineering because a more effective geoengineering reduces temperatures with less damages; it also increases in the cost of deploying the counter-geoengineering agent ϵ_A^c. The reluctance to undertake counter-geoengineering decreases with the damage reduction d_{AA}^c that it can achieve,

the main incentive to undertake counter-geoengineering. As before, the linear response functions give rise to the multiplier $\frac{1}{\alpha_A \alpha_B}$. Yet, here, region A's incentive to undertake counter-geoengineering ($S_t^A < 0$) is driven entirely by region B's geoengineering spillover $\alpha_B z_B$. Only if this geoengineering spill-over dominates region A's reluctance to undertake counter-engineering region A will deploy the counter-engineering agent ($S_t^A < 0$).

The cooling undertaken by region A decreases in the other regions cooling (e.g., as region B's perceived damages from geoengineering fall or those from temperature increase) and it increase in the other region's counter-geoengineering. There is a parameter range between the two situations where a region engages in cooling and where it is deploying the counter-geoengineering agent. In this parameter range, the region remains inactive and the other region sets the global temperature distribution alone. This region results from the cost-benefit-differential between no longer facing enough benefits to bear the costs of geoengineering and suffering enough from geoengineering to bear the costs of engaging into counter-geoengineering to at least partially offset the costs. Note that, if a region's reluctance to undertake counter-geoengineering becomes very high, the other region's effort approaches the geoengineering level it would undertake if the region is inactive.

We now identify the parameter ranges that give rise to the different equilibria and show that they are mutually exclusive.

Proposition 4. *There are 5 qualitatively different Nash-equilibria. They are mutually exclusive and classified based on fundamentals as follows:*

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

where

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

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

Proof. See Appendix C.1.2. □

Figure 4 shows how the domain of Nash-equilibria is affected by a) an increase in the cost of geoengineering (i.e either an increase in the damages or the operational costs) and b) an increase in the cost of counter-geoengineering (i.e. a decrease in the effectiveness or an increase in operational costs). Dashed lines show the domain before the change occurs.

An increase in the cost of geoengineering for region A decreases the propensity to do geo-engineering, i.e. z_A^g declines. The domain in which region B engages in geoengineering expands to the right. The domain where region B is taking no action shift to the right. Thus, the domain where region B engages in counter-geoengineering decreases. The domain in which region A is taking no action increases by expanding downwards.

An increase in the cost of counter-geoengineering in region A increases the aversion to do counter-geoengineering, i.e. z_A^c increases. This expands the domain in which region A taking no action upwards and thus decreases the domain where region A engages in counter-geoengineering. Note that due to the assumption that α_A, $\alpha_B \in (0, 1)$ only a subarea in each graph is possible.
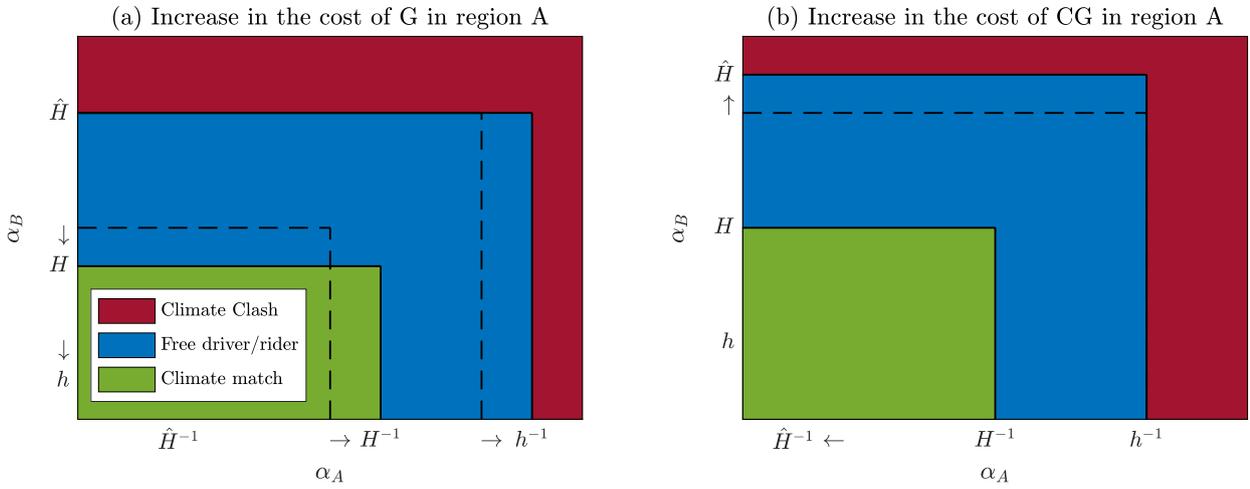
Figure 4: Effect of a change in the cost of geoengineering (G) and counter-geoengineering (CG) on the domain of Nash-equilibria.

Figure 5 shows a numeric example of Proposition 4. We set up two completely symmetric regions (shown in green), and then demonstrate how different regional damage parameters can shift the equilibrium away from a climate match to a free rider/driver or even to a climate clash equilibrium.

We now derive the regional social cost of carbon, which exhibits a different structure for each type of equilibrium.

Proposition 5. *If $S_t^B = 0$, the social cost of carbon in region A is given by*

$$SCC^A = \frac{Y_{A,t}^{net}}{M_{pre}} \left[a^A + f_1 \gamma_A - \frac{n}{1-n} z_A^g (f_2 \gamma_A + (d_{AA}^g + \epsilon^g)) \right] \tilde{\phi}_A.$$

If both regions are active ($S_t^B \neq 0$ and $S_t^A \neq 0$) the social cost of carbon is

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

with $z_A \in \{z_A^g, z_A^c\}$, $z_B \in \{z_B^g, z_B^c\}$, $d_{AA} \in \{d_{AA}^g + \epsilon^g, d_{AA}^c - \epsilon^c\}$, and $d_{BA} \in \{d_{BA}^g, d_{BA}^c\}$ depending on whether the corresponding region engages in geoengineering (g) or counter-geoengineering (c). If region A itself is inactive ($S_t^A = 0$) its social cost of carbon is

$$SCC^A = \frac{Y_{A,t}^{net}}{M_{pre}}\left[a^A + f_1\,\gamma_A + \alpha_B z_B^g d_{BA}^g - \gamma_A\Big(f_3(\alpha_B z_B^g)^{1-n} - \alpha_B z_B^g f_2\Big)\right]\tilde{\phi}_A.$$

Swapping region indices characterizes region B's social cost of carbon.

Proof. See Appendix C.1.3. □

If region B is inactive ($S_t^B = 0$), then the region A's SCC has the same structure as in the global model. In difference to the social planner, region A only accounts for its own climate impact γ_A and for its own damages d_{AA}^g and costs ϵ_A^g. In the other cases where region B is active ($S_t^B \neq 0$), an additional term enters region A's SCC. It results from the spillovers from region B to region A. If both regions are active, this spillover effect is given by $-\frac{\alpha_B(z_B - \alpha_A z_A)(d_{AA} - d_{BA})}{1 - \alpha_A \alpha_B}$.

In the case where region B is cooling, the term $\alpha_B(z_B - \alpha_A z_A)$ is positive. In a climate clash $d_{AA} - d_{BA} < 0$ since counter-geoengineering by region A can at most offset the damages from geoengineering and, therefore, the spillover effect increases the SCC. Similarly, if region A is cooling but region B engages in counter-geoengineering the spillover effect increases the SCC: the term $\alpha_B(z_B - \alpha_A z_A)$ is negative because region B engages in counter-geoengineering and $d_{AA} - d_{BA} > 0$ because region B's counter-geoengineering agent can at most offset the damages in region A.

Thus, in a climate clash, the spillover effect always increases the SCC in both regions. The easy intuition is that the other region will always interfere with what the region would like to do if it was acting alone.

In the climate match, the sign of the spillover effect depends on the marginal damage of geoengineering in both regions. If the marginal damage from the spillover cooling is larger than the marginal damage from the cooling done by region A itself, $d_{BA}^g > d_{AA}^g$, the spillover effect is positive. In the opposite case, where $d_{BA}^g < d_{AA}^g$, the spillover effect is negative. In the most intuitive scenario where the sulfur traveling over from the other region causes the same damages as the sulfur deployed locally we have $d_{AA} - d_{BA}) = \epsilon_A^g > 0$ and the spillover term from the other cooling region indeed reduces the local SCC.

If regions A is inactive itself ($S_t^A = 0$), the typical geoengineering term $\frac{n}{1-n}\,z_A(f_2\,\gamma_A + d_{AA})$ disappears from the equation: it relies on the active balancing of the local temperature in

accordance to damage, impact, and cooling parameters. Instead, we find two spillover terms. The first term $\alpha_B\, z_B^g\, d_{BA}^g$ is positive and captures the spillover of damages from region B's geoengineering onto region A. The second term $-\gamma_A\left(f_3(\alpha_B\, z_B^g)^{1-n} - \alpha_B\, z_B^g\, f_2\right)$ is negative and accounts for the reduction in climate change impacts due to the spillover of cooling from region B to region A. If the effect from the cooling spillover is larger (smaller) than the effect from the damage spillover, the overall effect on region A's social cost of carbon is negative (positive).
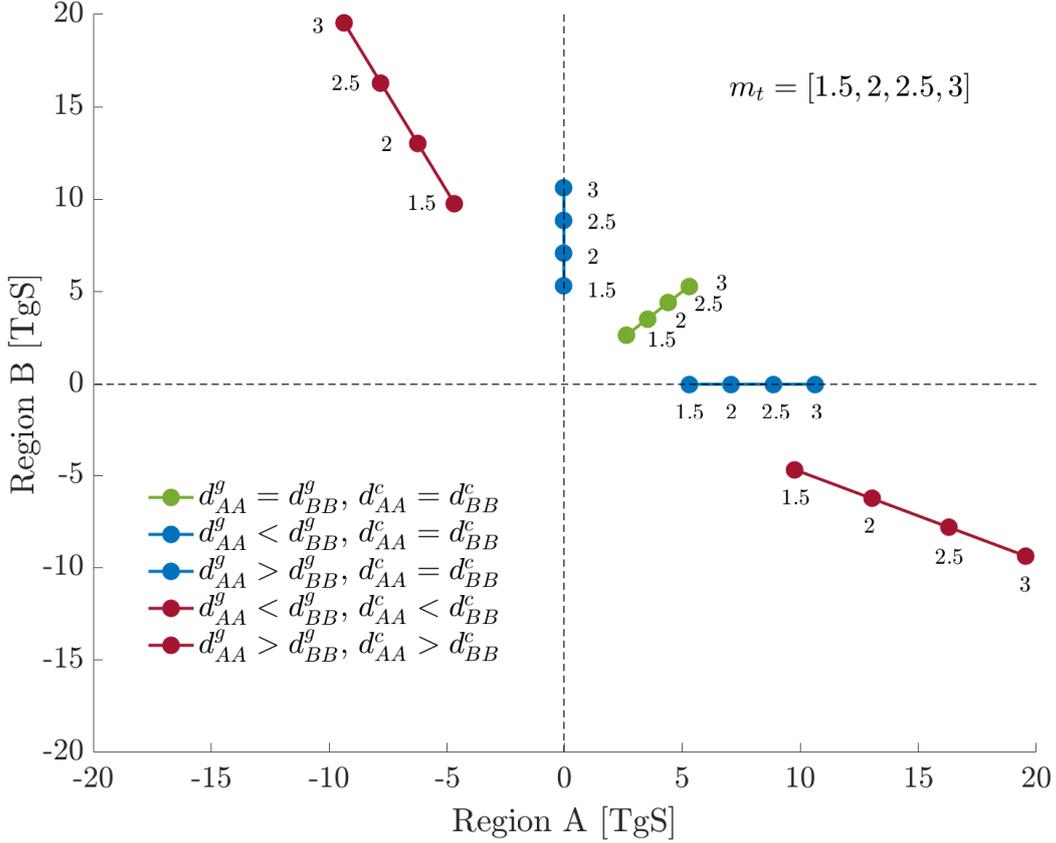
Figure 5: Nash-equilibria for different regional damage parameters of two otherwise symmetric regions.

3.3 Rest of the world

We have assumed that the rest of the world does not engage in geoengineering measures. Therefore, it affects regions A and B only through its emissions that change the global stock of carbon and, therefore, the regions' geoengineering levels. We now analyze how the rest of world responds to the prospect that other regions engage in geoengineering and potential counter-geoengineering activities.

The rest of the world has an aggregate economy similar to that of regions A and B with idiosyncratic functional forms and parameters, and we denote the sulfur related damages (or damage reductions) that occur in the rest of the world from region A's sulfur deployment (or

counter-geoengineering agent) by $d_{AW} \in \{d_{AW}^g, d_{AW}^c\}$ and the damage (or damage reduction) that occurs from region B's sulfur use (or counter-geoengineering agent) by $d_{BW} \in \{d_{BW}^g, d_{BW}^c\}$. If a region chooses to be inactive the respective damage parameter is equal to zero. The damages in the rest of the world as a fraction of output are

$$D_t^W(\tau_{1,t}^W, S_t, m_t) = 1 - \exp\left[\xi_0^W\left(1 - \tau_{1,t}^W\right) - (d_{BW}S_t^B + d_{AW}\tilde{S}_t^A) - a^W(m_t - 1)\right] \qquad (30)$$

As we assumed that the rest of the world is part of climate zone B, the share of sulfur that travels from region A to the rest of the world is given by $\tilde{S}_t^A$ and we simply assume that the sulfur level, also in terms of damaging fallout, is the same as in region B (S_t^B). We ask whether the availability of geoengineering and/or counter-geoengineering increases or reduces the incentive to abate in the rest of the world.

Proposition 6. *If both regions are active ($S_t^A \neq 0$ and $S_t^B \neq 0$), the social cost of carbon in the rest of the world is*

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

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 $d_{BB} \in \{d_{BB}^c, d_{BB}^g\}$. If region A is active ($S_t^A > 0$) and region B is inactive ($S_t^B = 0$), the social cost of carbon is given by

$$SCC^W = \frac{Y_{W,t}^{net}}{M_{pre}}\left[a^W + f_1\,\gamma_W + \alpha_A\,z_A^g\,d_{AW} - \gamma_W\left(f_3\,(\alpha_A\,z_A^g)^{1-n} - \alpha_A\,z_A^g\,f_2\right)\right]\tilde{\phi}_W.$$

If region A is inactive ($S_t^A = 0$) and region B is active ($S_t^B > 0$), the social cost of carbon is given by

$$SCC^W = \frac{Y_{W,t}^{net}}{M_{pre}}\left[a^W + f_1\,\gamma_W + z_B^g\,d_{BW} - \gamma_W\left(f_3\,(z_B^g)^{1-n} - f_2\,z_B^g\right)\right]\tilde{\phi}_W.$$

The first two contributions from ocean acidification and temperature increase remain the same as in the earlier expressions for the SCC. If both regions A and B are active, the third and fourth term represents the damages that spill over from each region: $\frac{(z_A - \alpha_B\,z_B)}{1 - \alpha_A\,\alpha_B}$ is the level of sulfur deployment by region A per additional ton of CO_2 released into the atmosphere (see Proposition 3), which is multiplied with the damages per ton of sulfur reaching the rest of the world. Since the rest of the world is located in the same climate zone as region B, the α parameter only reduces the sulfur carry-over from region A. The last term in the SCC is negative and accounts for the reduction in climate change impacts due to the cooling spillover. Whenever region B is actively managing the radiative forcing in the climate zone B, the rest of the world's cooling is controlled only by region B and its cooling propensity.

If region B remains inactive, there are no more damages from region B's geoengineering

and the cooling spillover is now determined by region A's geoengineering propensity (and the effectiveness of cooling f_3). If region A remains inactive, there is no more damage carry-over from region A, and region B geoengineering propensity is once again responsible for the additional cooling in response to an additional ton of CO_2 released in the rest of the world.

In conclusion, the incentive to abate can both increase or decrease in the rest of the world as a result of regions that engage in geoengineering. If the damages from geoengineering are low but the temperature impact is high, the rest of the world will reduce its mitigation effort in response to other regions willingness to engage in geoengineering. That is a formalization and extension the wide-spread "slippery slope argument". The fact that we consider under-taking geoengineering in the future makes us emit more. And worse: it is enough that some region considers undertaking geoengineering and the rest of the world is likely to reduce their mitigation efforts, increasing the pressure on the geoengineering region to increase their level of cooling and sulfur injections.

However, we show that the opposite can happen as well. If the rest of the world is relatively more afraid of the damages from geoengineering and/or less affected by temperature increase, then the availability of geoengineering in a region suffering more from climate change can increase the mitigation effort in the rest of the world. Currently, many countries have a large part of the population that is quite afraid of geoengineering. Indeed, we find that optimal sulfur deployment is extremely sensitivity to the a damage level that is currently unknown. Thus, it can be perfectly rational to push for even more mitigation in the face of other region's potential use of geoengineering.

4   Conclusions

We have integrated solar geoengineering into a state of the art integrated assessment model. For this purpose, we found a new solution class to closed-form dynamic models which allows us to incorporate the current scientific knowledge about the radiative forcing response to stratospheric sulfur injections. In the global model, the optimal level of sulfur deployment is very sensitive to the potential damages from geoengineering. The globally optimal carbon tax decreases due to the availability of geoengineering. The size of this reduction increases in the sulfur-based cooling efficiency and falls with geoengineering damages.

We solve the dynamic game between heterogenous regions. Our analytic solution explains the strategic interactions between two active regions that have the ability to engage in geo-engineering or counter-geoengineering, and a passive rest of the world that merely responds by changing CO_2 emissions. The regional model gives rise to several Markov-perfect Nash equi-libria: a climate match where both regions cool the world, a free driver outcome where one region cools and the other region stays inactive, and a climate clash where one region cools and the other region offsets a part of this cooling. These equilibria are mutually exclusive and we characterize their occurrence based on the region's characteristics.

We characterize the regionally optimal carbon taxes for the different players (regions) and show how they depend on the strategies and characteristics of the different regions. As is to be expected, the availability of geoengineering can reduce the incentive to reduce CO_2 emissions and, thus, the regionally optimal carbon tax. However, we show that the availability of geoengineering in some region can also increase the optimal carbon tax and, thus, mitigation incentive in the rest of the world, as a result of differences in perceived damages.

Our paper is one of the few regional integrated assessment models that can handel strategic interactions. To the best of our knowledge, it is the only full-fledged analytic model where regions interact "non-trivially" in that strategies are not dominant. Moreover, the framework permits the introduction of uncertainty, which is ubiquitous in the economics of climate engineering. We are currently working on completing the numeric calibration to deliver quantitative results also for the game.

References

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.

Carleton, T. A. and Hsiang, S. M. Social and economic impacts of climate. Science, 353(6304): 9837–1, 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.

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.

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.

Goes, Marlos; Tuana, Nancy, and Keller, Klaus. The economics (or lack thereof) of aerosol geoengineering. Climatic Change, 109(3-4):719–744, 2011. ISSN 01650009. doi: 10.1007/s10584-010-9961-z.

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 and 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.

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.

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.

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.

Klepper, Gernot and Rickels, Wilfried. Climate engineering: Economic considerations and research challenges. Review of Environmental Economics and Policy, 8(2):270–289, 2014.

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.

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

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.

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.

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.

Pasztor, Janos; Scharf, Cynthia, and Schmidt, Kai-Uwe. How to govern geoengineering? Science, 357(6348):231–231, 2017.

Proctor, Jonathan; Hsiang, Solomon; Burney, Jennifer; Burke, Marshall, and Schlenker, Wolfram. Estimating global agricultural effects of geoengineering using volcanic eruptions. Nature, 560(7719):480–483, 2018.

Ricke, Katharine L.; Morgan, M. Granger, and Allen, Myles R. Regional climate response to solar-radiation management. Nature Geoscience, 3(8):537–541, 2010.

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. Global industrial carbon emissions to reach all-time high in 2018. Nature News, 2018. URL https://www.nature.com/articles/d41586-018-07666-6#references.

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

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.

Appendices

A   Radiative forcing approximation
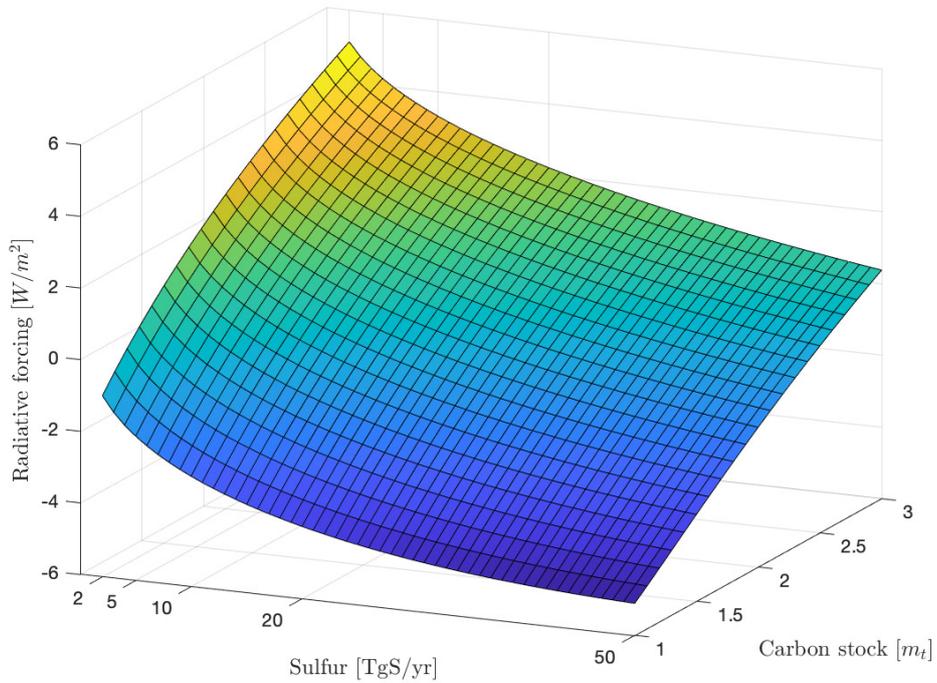
Figure 6: Forcing effect from carbon stock and injection rate
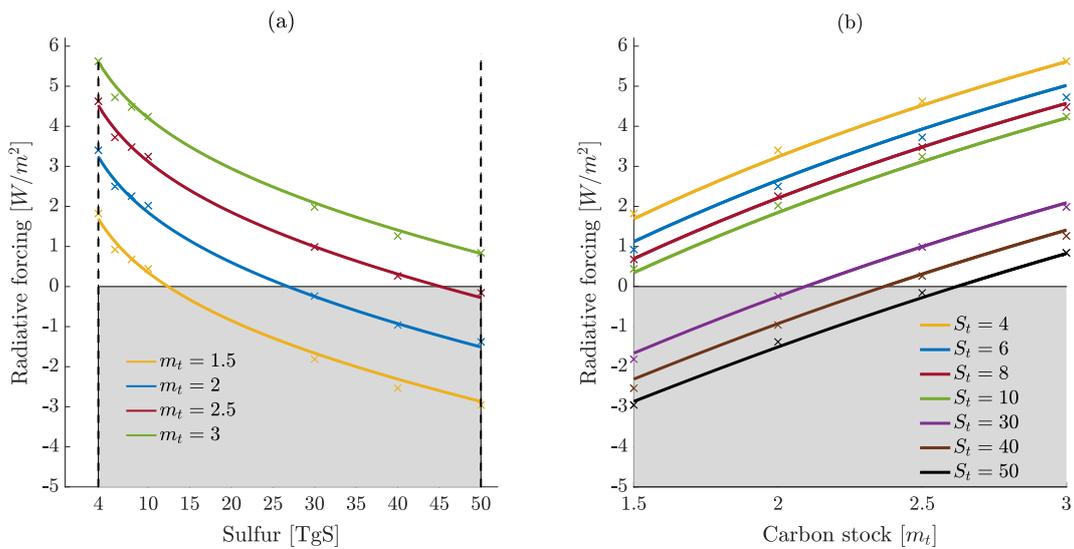
Figure 7: Approximation to model data from Niemeier and Timmreck (2015). The forcing effect from sulfur for different carbon stocks is shown in (a). The forcing effect from the carbon stock for different injection rates is shown in (b).

B Global model

B.1 Proof of sufficiency

The consumption rate can be written as

$$x_t = \frac{C_t}{Y_t \left[1 - D_t\left(T_{1,t}, G_t(S_t), m_t\right)\right]}. \tag{31}$$

Homogeneity of the production function implies

$$Y_t = F(\boldsymbol{A}_t, \boldsymbol{K}_t, \boldsymbol{N}_t, \boldsymbol{E}_t) = K_t^\kappa F(\boldsymbol{A}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{N}_t, \boldsymbol{E}_t), \tag{32}$$

such that

$$\log C_t = \log x_t + \kappa \log K_t + \log F(\boldsymbol{A}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{N}_t, \boldsymbol{E}_t) + \xi_0 \left(1 - \tau_{1,t}\right) - d\, S_t - a(m_t - 1). \tag{33}$$

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

$$V(k_t, \boldsymbol{\tau}_t, \boldsymbol{M}_t, \boldsymbol{R}_t, t) = \max_{x_t, \boldsymbol{N}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{E}_t, S_t} \left\{ \log x_t + \kappa\, k_t + \log F(\boldsymbol{A}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{N}_t, \boldsymbol{E}_t) \right.$$
$$\left. + \xi_0 \left(1 - \tau_{1,t}\right) - d\, S_t - a(m_t - 1) + \beta\, V\left(k_{t+1}, \boldsymbol{\tau}_{t+1}, \boldsymbol{M}_{t+1}, \boldsymbol{R}_{t+1}, t+1\right) \right\} \tag{34}$$

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

$$k_{t+1} = \kappa\, k_t + \log F(\boldsymbol{A}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{N}_t, \boldsymbol{E}_t) + \log(1 - x_t) + \xi_0 \left(1 - \tau_{1,t}\right) - d\, S_t - a(m_t - 1). \tag{35}$$

The linear affine guess for the value function

$$V(k_t, \boldsymbol{\tau}_t, \boldsymbol{M}_t, \boldsymbol{R}_t, t) = \varphi_k\, k_t + \boldsymbol{\varphi}_\tau^T\, \boldsymbol{\tau}_t + \boldsymbol{\varphi}_M^T\, \boldsymbol{M}_t + \boldsymbol{\varphi}_{R,t}^T\, \boldsymbol{R}_t + \varphi_t \tag{36}$$

turns the Bellman equation into the form

$$
\varphi_k \, k_t + \boldsymbol{\varphi}_\tau^T \, \boldsymbol{\tau}_t + \boldsymbol{\varphi}_M^T \, \boldsymbol{M}_t + \boldsymbol{\varphi}_{R,t}^T \, \boldsymbol{R}_t + \varphi_t
$$

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

$$
+ \lambda_t^{\mathcal{K}} (1 - \sum_{i=1}^{I_K} \mathcal{K}_{i,t}) + \lambda_t^N (1 - \sum_{i=1}^{I_N} N_{i,t}) + \beta \, \varphi_k \left(\kappa \, k_t + \log F(\boldsymbol{A}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{N}_t, \boldsymbol{E}_t) + \log(1 - x_t) + \xi_0 \, (1 - \tau_{1,t}) \right.
$$

$$
\left. - d \, S_t - a(m_t - 1) \right) + \beta \, \boldsymbol{\varphi}_\tau^T \left(\boldsymbol{\sigma} \boldsymbol{\tau}_t + \tilde{\boldsymbol{F}}_t \right) + \beta \, \boldsymbol{\varphi}_M^T \left(\boldsymbol{\Phi} \boldsymbol{M}_t + \tilde{\boldsymbol{e}} \right) + \beta \, \boldsymbol{\varphi}_{R,t+1}^T \left(\boldsymbol{R}_t - \boldsymbol{E}_t^d \right) + \beta \, \varphi_{t+1} \Bigg\}.
$$

$$(37)$$

We show that the system (2 layer carbon and 2 layer temperature system) is linear in states and that the affine value function, (36), solves the system. Inserting the trial solution and the next periods states (equations (5), (7), (10) and (35)) into the Bellman equation delivers

$$
\varphi_k \, k_t + \boldsymbol{\varphi}_\tau^T \, \boldsymbol{\tau}_t + \boldsymbol{\varphi}_M^T \, \boldsymbol{M}_t + \boldsymbol{\varphi}_{R,t}^T \, \boldsymbol{R}_t + \varphi_t
$$

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

$$
+ \lambda_t^{\mathcal{K}} (1 - \sum_{i=1}^{I_K} \mathcal{K}_{i,t}) + \lambda_t^N (1 - \sum_{i=1}^{I_N} N_{i,t}) + \beta \, \varphi_k \left(\kappa \, k_t + \log F(\boldsymbol{A}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{N}_t, \boldsymbol{E}_t) + \log(1 - x_t) \right.
$$

$$
\left. + \xi_0 \, (1 - \tau_{1,t}) - d \, S_t - a(m_t - 1) \right) + \beta \, \boldsymbol{\varphi}_\tau^T \left(\boldsymbol{\sigma} \boldsymbol{\tau}_t + \tilde{\boldsymbol{F}}_t \right) + \beta \, \boldsymbol{\varphi}_M^T \left(\boldsymbol{\Phi} \boldsymbol{M}_t + \tilde{\boldsymbol{e}}_t \right) + \beta \, \boldsymbol{\varphi}_{R,t+1}^T \left(\boldsymbol{R}_t - \boldsymbol{E}_t^d \right) + \beta \, \varphi_{t+1} \Bigg\}.
$$

$$(38)$$

We now look at the first order conditions. Maximizing the right hand side over x_t yields

$$
\frac{1}{x_t} - \beta \, \varphi_k \frac{1}{1 - x_t} = 0 \quad \Longrightarrow \quad x_t = \frac{1}{1 + \beta \, \varphi_k}. \tag{39}
$$

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

$$
(1 + \beta \, \varphi_k) \frac{\frac{\partial F(\boldsymbol{A}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{N}_t, \boldsymbol{E}_t)}{\partial \mathcal{K}_{i,t}}}{F(\boldsymbol{A}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{N}_t, \boldsymbol{E}_t)} = \lambda_t^{\mathcal{K}} \tag{40}
$$

which is equivalent to

$$
\mathcal{K}_{i,t} = \frac{\sigma_{Y,\mathcal{K}_i}(\boldsymbol{A}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{N}_t, \boldsymbol{E}_t)}{\sum_{i=1}^{I_K} \sigma_{Y,\mathcal{K}_i}(\boldsymbol{A}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{N}_t, \boldsymbol{E}_t)} \tag{41}
$$

with

$$\sigma_{Y, \mathcal{K}_i}(\boldsymbol{A}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{N}_t, \boldsymbol{E}_t) \equiv \frac{\partial F(\boldsymbol{A}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{N}_t, \boldsymbol{E}_t)}{\partial \mathcal{K}_{i,t}} \frac{\mathcal{K}_{i,t}}{F(\boldsymbol{A}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{N}_t, \boldsymbol{E}_t)}. \tag{42}$$

Similarly, the first order conditions for the labor input is

$$(1 + \beta \, \varphi_k) \frac{\frac{\partial F(\boldsymbol{A}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{N}_t, \boldsymbol{E}_t)}{\partial N_{i,t}}}{F(\boldsymbol{A}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{N}_t, \boldsymbol{E}_t)} = \lambda_t^N \tag{43}$$

and hence

$$N_{i,t} = \frac{\sigma_{Y, N_i}(\boldsymbol{A}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{N}_t, \boldsymbol{E}_t)}{\sum\limits_{i=1}^{I_N} \sigma_{Y, N_i}(\boldsymbol{A}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{N}_t, \boldsymbol{E}_t)} \tag{44}$$

with

$$\sigma_{Y, N_i}(\boldsymbol{A}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{N}_t, \boldsymbol{E}_t) \equiv \frac{\partial F(\boldsymbol{A}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{N}_t, \boldsymbol{E}_t)}{\partial N_{i,t}} \frac{N_{i,t}}{F(\boldsymbol{A}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{N}_t, \boldsymbol{E}_t)} \tag{45}$$

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

$$(1 + \beta \, \varphi_k) \frac{\frac{\partial F(\boldsymbol{A}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{N}_t, \boldsymbol{E}_t)}{\partial E_{i,t}}}{F(\boldsymbol{A}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{N}_t, \boldsymbol{E}_t)} = \beta(\varphi_{R,i,t+1} - \varphi_{M1}) \tag{46}$$

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

$$\sigma_{Y, E_i}(\boldsymbol{A}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{N}_t, \boldsymbol{E}_t) \equiv \frac{\partial F(\boldsymbol{A}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{N}_t, \boldsymbol{E}_t)}{\partial E_{i,t}} \frac{E_{i,t}}{F(\boldsymbol{A}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{N}_t, \boldsymbol{E}_t)}. \tag{48}$$

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

$$B_t = (\beta \, \varphi_{\tau 1} \, \sigma_{\text{forc}} \, f_2 - (1 + \beta \, \varphi_k) \, d) \, S_t - \beta \, \varphi_{\tau 1} \, \sigma_{\text{forc}} \, f_3 \, m_t^n \, S_t^{1-n} \tag{49}$$

and find the first order condition for optimal sulfur deployment

$$(n - 1) \, \beta \, \varphi_{\tau 1} \, \sigma_{\text{forc}} \, f_3 \, m_t^n \, S_t^{-n} + \beta \, \varphi_{\tau 1} \, \sigma_{\text{forc}} \, f_2 - (1 + \beta \, \varphi_k) \, d = 0. \tag{50}$$

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

$$S_t = \underbrace{\left(\frac{(n-1)\beta\, \varphi_{\tau 1}\, \sigma_{\text{forc}}\, f_3}{(1 + \beta\, \varphi_k)\, d - \beta\, \varphi_{\tau 1}\, \sigma_{\text{forc}}\, f_2} \right)^{\frac{1}{n}}}_{\equiv z} m_t. \tag{51}$$

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.

Inserting the optimal control rules into the maximized Bellman equation gives us

$$\begin{aligned}
\varphi_k\, k_t &+ \boldsymbol{\varphi}_\tau^T\, \boldsymbol{\tau}_t + \boldsymbol{\varphi}_M^T\, \boldsymbol{M}_t + \boldsymbol{\varphi}_{R,t}^T\, \boldsymbol{R}_t + \varphi_t \\
&= \log x_t^* + \kappa\, k_t + \log F(\boldsymbol{A}_t, \boldsymbol{\mathcal{K}}_t^*, \boldsymbol{N}_t^*, \boldsymbol{E}_t^*) + \xi_0\, (1 - \tau_{1,t}) - d\, S_t^* - a(m_t - 1) \\
&+ \beta\, \varphi_k \left(\kappa\, k_t + \log F(\boldsymbol{A}_t, \boldsymbol{\mathcal{K}}_t^*, \boldsymbol{N}_t^*, \boldsymbol{E}_t^*) + \log(1 - x_t^*) + \xi_0\, (1 - \tau_{1,t}) - d\, S_t^* - a(m_t - 1) \right) \\
&+ \beta\, \boldsymbol{\varphi}_\tau^T \left(\boldsymbol{\sigma} \boldsymbol{\tau}_t + \tilde{\boldsymbol{F}}_t \right) + \beta\, \boldsymbol{\varphi}_M^T \left(\boldsymbol{\Phi} \boldsymbol{M}_t + \tilde{\boldsymbol{e}}_t \right) + \beta\, \boldsymbol{\varphi}_{R,t+1}^T \left(\boldsymbol{R}_t - \boldsymbol{E}_t^{d*} \right) + \beta\, \varphi_{t+1}
\end{aligned} \tag{52}$$

Arranging terms with respect to their states yields

$$\begin{aligned}
\varphi_k\, k_t + \boldsymbol{\varphi}_\tau^T\, \boldsymbol{\tau}_t &+ \varphi_{M1}\, M_{1,t} + \varphi_{M2}\, M_{2,t} + \boldsymbol{\varphi}_{R,t}^T\, \boldsymbol{R}_t + \varphi_t = \left[(1 + \beta\, \varphi_k)\kappa \right] k_t + \left[\beta\, \boldsymbol{\varphi}_\tau^T\, \boldsymbol{\sigma} - (1 + \beta\, \varphi_k)\xi_0\, \boldsymbol{e}_1^T \right] \boldsymbol{\tau}_t \\
&+ \left[\left(-(1 + \beta\, \varphi_k) d\, z + f_2\, \beta\, \varphi_{\tau 1}\, \sigma_{\text{forc}}\, z - f_3\, \beta\, \varphi_{\tau 1}\, \sigma_{\text{forc}}\, z^{1-n} - (1 + \beta\, \varphi_k)a + f_1\, \beta\, \varphi_{\tau 1}\, \sigma_{\text{forc}} \right) M_{pre}^{-1} \right. \\
&+ \left. \beta\, (\varphi_{M1}\, \phi_{11} + \varphi_{M2}\, \phi_{12}) \right] M_{1,t} + \left[\beta(\varphi_{M1}\, \phi_{21} + \varphi_{M2}\, \phi_{22}) \right] M_{2,t} + \left[\beta\, \boldsymbol{\varphi}_{R,t+1}^T \right] \boldsymbol{R}_t + \log x_t^* + \beta\, \varphi_k\, \log(1 - x_t^*) \\
&+ (1 + \beta\, \varphi_k) \log F(\boldsymbol{A}_t, \boldsymbol{\mathcal{K}}_t^*, \boldsymbol{N}_t^*, \boldsymbol{E}_t^*) + (1 + \beta\, \varphi_k)(\xi_0 + a) + \beta\, \varphi_{\tau 1}\, \sigma_{\text{forc}}\, f_0 + \beta\, \varphi_{M1}(\sum_{i=1}^{I^d} E_{i,t}^* + E_t^{\text{exo}}) \\
&- \beta\, \boldsymbol{\varphi}_{R,t+1}^T\, \boldsymbol{E}_t^{d*} + \beta\, \varphi_{t+1}.
\end{aligned} \tag{53}$$

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

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

Inserting φ_k into equation (39) yield the optimal consumption rate $x_t^* = 1 - \beta\, \kappa$.

Coefficient matching with respect to transformed temperatures delivers

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

with

$$[\boldsymbol{1} - \beta\, \boldsymbol{\sigma}]^{-1} = \begin{pmatrix} \tilde{\sigma}_{11} & \tilde{\sigma}_{12} \\ \tilde{\sigma}_{21} & \tilde{\sigma}_{22} \end{pmatrix},$$

and hence

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

The first element of the inverted matrix $\tilde{\sigma}_{11} = \left(1 - \beta(1 - \sigma_{01} - \sigma_{21}) - \frac{\beta^2 \, \sigma_{21} \, \sigma_{12}}{1 - \beta(1 - \sigma_{12})}\right)^{-1}$. Coefficient matching with respect to the atmospheric carbon stock leads to

$$\varphi_{M1} = \left(-(1 + \beta \, \varphi_k)d \, z + f_2 \, \beta \, \varphi_{\tau 1} \, \sigma_{\text{forc}} \, z - f_3 \, \beta \, \varphi_{\tau 1} \, \sigma_{\text{forc}} \, z^{1-n} - (1 + \beta \, \varphi_k)a + f_1 \, \beta \, \varphi_{\tau 1} \, \sigma_{\text{forc}}\right) M_{pre}^{-1}$$
$$+ \beta \, (\varphi_{M1} \, \phi_{11} + \varphi_{M2} \, \phi_{12}).$$

Coefficient matching with respect to the carbon stock in the ocean leads to

$$\varphi_{M2} = \beta(\varphi_{M1} \, \phi_{21} + \varphi_{M2} \, \phi_{22}) \quad \Leftrightarrow \quad \varphi_{M2} = \frac{\beta \, \varphi_{M1} \, \phi_{21}}{1 - \beta \, \phi_{22}}.$$

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 (53) turns to the following condition:

$$\varphi_t - \beta \, \varphi_{t+1} = \log x_t^* + \beta \, \varphi_k \, \log(1 - x_t^*) + (1 + \beta \, \varphi_k) \log F(\boldsymbol{A}_t, \, \boldsymbol{\mathcal{K}}_t^*, \, \boldsymbol{N}_t^*, \, \boldsymbol{E}_t^*) + (1 + \beta \, \varphi_k)(\xi_0 + a)$$
$$+ \beta \, \varphi_{\tau 1} \, \sigma_{\text{forc}} \, f_0 + \beta \, \varphi_{M1}(\sum_{i=1}^{I^d} E_{i,t}^* + E_t^{\text{exo}}) - \beta \, \boldsymbol{\varphi}_{R,t+1}^T \, \boldsymbol{E}_t^{d*} \quad (54)$$

This condition will be satisfied by picking the sequence $\varphi_0, \varphi_1, \varphi_2, \ldots$ 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.

B.2 Optimal Level of Sulfur

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

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

The endogenous shadow value of capital $\varphi_k > 0$ is positive, while the endogenous shadow value of (transformed) temperature is negative $\varphi_{\tau 1} < 0$ (a bad). 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.

Inserting $\varphi_{\tau 1} = -\xi_0 \left(1 + \beta\, \varphi_k\right) \tilde{\sigma}_{11}$, and defining $\gamma = \beta\, \xi_0\, \tilde{\sigma}_{11}\, \sigma_{\text{forc}}$ delivers

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

Inserting the fit parameters from Table 2, and using the parameter values from the baseline calibration of the ACE model from Table 5 leads to

$$S_t^* = \left(\frac{0.0016}{d + 0.00004} \right)^{1.15} m_t.$$

B.3   Social cost of carbon

Using $\gamma = \beta\, \xi_0\, \tilde{\sigma}_{11}\, \sigma_{\text{forc}}$ in the shadow price of the atmospheric carbon concentration yields

$$\varphi_{M1} = (1 + \beta\, \varphi_k)\left(-d\, z - f_2\, \gamma\, z + f_3\, \gamma\, z^{1-n} - a - f_1\, \gamma \right) M_{pre}^{-1} + \beta\left(\varphi_{M1}\, \phi_{11} + \varphi_{M2}\, \phi_{12} \right).$$

Inserting φ_{M2} and solving for φ_{M1} delivers

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

where we abbreviated $\tilde{\phi} = \left(1 - \beta\, \phi_{11} - \frac{\beta^2\, \phi_{12}\, \phi_{21}}{1 - \beta\, \phi_{22}}\right)^{-1}$. Inserting φ_k leads to

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

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

$$SCC = -(1 - \beta\, \kappa) Y_t^{net}\, \varphi_{M1}$$

$$= \frac{Y_t^{net}}{M_{pre}} \left[d\, z + f_2\, \gamma\, z - f_3\, \gamma\, z^{1-n} + a + f_1\, \gamma \right] \tilde{\phi}$$

Which is equivalent to

$$SCC = \frac{Y_t^{net}}{M_{pre}} \left[a + \gamma\, f_1 - \frac{n}{1-n}(d + \gamma\, f_2)z \right] \tilde{\phi}$$

Using parameter values from the ACE model (see Table 5) yields $\gamma = 0.0108$. Thus, the SCC turns to

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

C Regional model

C.1 Active regions

In terms of transformed temperatures in the regional model we have

$$
\underbrace{\begin{pmatrix} \tau_{1,t+1}^A \\ \tau_{1,t+1}^B \\ \tau_{2,t+1} \end{pmatrix}}_{\equiv\,\boldsymbol{\tau}_{t+1}} = \underbrace{\begin{pmatrix} 1-\sigma^A & \sigma_B^A & \sigma_O^A \\ \sigma_A^B & 1-\sigma^B & \sigma_O^B \\ \sigma_A^O & \sigma_B^O & 1-\sigma^O \end{pmatrix}}_{\equiv\,\boldsymbol{\sigma}} \underbrace{\begin{pmatrix} \tau_{1,t}^A \\ \tau_{1,t}^B \\ \tau_{2,t} \end{pmatrix}}_{\equiv\,\boldsymbol{\tau}_t} + \underbrace{\begin{pmatrix} \sigma_{\text{forc}}\exp(\frac{\log 2}{\eta}F_t^A) \\ \sigma_{\text{forc}}\exp(\frac{\log 2}{\eta}F_t^B) \\ 0 \end{pmatrix}}_{\equiv\,\tilde{\boldsymbol{F}}_t(S_t^A,S_t^B)}
\tag{55}
$$

where $\sigma^A = \sigma_B^A + \sigma_O^A$, $\sigma^B = \sigma_A^B + \sigma_O^B$, and $\sigma^O = \sigma_A^O + \sigma_B^O$, or equivalently

$$
\boldsymbol{\tau}_{t+1} = \boldsymbol{\sigma}\boldsymbol{\tau}_t + \tilde{\boldsymbol{F}}_t(S_t^A, S_t^B).
\tag{56}
$$

The dynamics of the carbon reservoirs are given by

$$
\underbrace{\begin{pmatrix} M_{1,t+1} \\ M_{2,t+1} \end{pmatrix}}_{\equiv\,\boldsymbol{M}_{t+1}} = \underbrace{\begin{pmatrix} \phi_{11} & \phi_{21} \\ \phi_{12} & \phi_{22} \end{pmatrix}}_{\equiv\,\boldsymbol{\Phi}} \underbrace{\begin{pmatrix} M_{1,t} \\ M_{2,t} \end{pmatrix}}_{\equiv\,\boldsymbol{M}_t} + \underbrace{\begin{pmatrix} \sum_{i=1}^{I^d} E_{A,i,t} + \sum_{i=1}^{I^d} E_{B,i,t} + \sum_{i=1}^{I^d} E_{W,i,t} + E_t^{\text{exo}} \\ 0 \end{pmatrix}}_{\equiv\,\tilde{\boldsymbol{e}}_t}
\tag{57}
$$

or equivalently

$$
\boldsymbol{M}_{t+1} = \boldsymbol{\Phi}\boldsymbol{M}_t + \tilde{\boldsymbol{e}}_t.
\tag{58}
$$

C.1.1 Markov strategies

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

Bellman equation. In the following we show for region A that the system (3 temperatures and 2 carbon reservoirs) is linear in states and that the affine value function

$$
V(k_t, \boldsymbol{\tau}_t, \boldsymbol{M}_t, \boldsymbol{R}_t, t) = \varphi_k^A k_t + \boldsymbol{\varphi}_{\tau A}^T \boldsymbol{\tau}_t + \boldsymbol{\varphi}_{MA}^T \boldsymbol{M}_t + \boldsymbol{\varphi}_{R,t}^T \boldsymbol{R}_t + \varphi_t,
\tag{59}
$$

with

$$
\boldsymbol{\varphi}_{\tau A}^T = \begin{pmatrix} \varphi_{\tau1}^{AA} \\ \varphi_{\tau1}^{BA} \\ \varphi_{\tau2}^A \end{pmatrix} \quad \text{and} \quad \boldsymbol{\varphi}_{MA}^T = \begin{pmatrix} \varphi_{M1}^A \\ \varphi_{M2}^A \end{pmatrix},
$$

solves the system. The proof for region B follows analogously. We only use region indices when they are needed. We denote the shadow price of temperature $\tau_{1,t}^A$ in region A by $\varphi_{\tau1}^{AA}$, and the shadow price of temperature $\tau_{1,t}^B$ in region A by $\varphi_{\tau1}^{BA}$.

Inserting the trial solution and the next periods states into the Bellman equation delivers

$$
\varphi_k^A \, k_t + \boldsymbol{\varphi}_{\tau A}^T \, \boldsymbol{\tau}_t + \boldsymbol{\varphi}_{MA}^T \, \boldsymbol{M}_t + \boldsymbol{\varphi}_{R,t}^T \, \boldsymbol{R}_t + \varphi_t
$$

$$
= \max_{x_t, \boldsymbol{N}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{E}_t, S_t^A} \Bigg\{ \log x_t + \kappa \, k_t + \log F(\boldsymbol{A}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{N}_t, \boldsymbol{E}_t) + \xi_0^A \left(1 - \tau_{1,t}^A\right) - [d_{AA} S_t^A(m_t) + d_{BA} \tilde{S}_t^B(m_t)]
$$

$$
- a^A \, (m_t - 1) + \lambda_t^{\mathcal{K}} (1 - \sum_{i=1}^{I_K} \mathcal{K}_{i,t}) + \lambda_t^N (1 - \sum_{i=1}^{I_N} N_{i,t}) + \beta^A \, \varphi_k^A \left(\kappa \, k_t + \log F(\boldsymbol{A}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{N}_t, \boldsymbol{E}_t) + \log(1 - x_t) \right.
$$

$$
+ \xi_0^A \left(1 - \tau_{1,t}^A\right) - [d_{AA} S_t^A + d_{BA} \tilde{S}_t^B] - a^A \, (m_t - 1) \Big) + \beta^A \, \boldsymbol{\varphi}_{\tau A}^T \left(\boldsymbol{\sigma \tau}_t + \tilde{\boldsymbol{F}}_t(S_t^A(m_t), S_t^B(m_t)) \right)
$$

$$
+ \beta^A \, \boldsymbol{\varphi}_{MA}^T \left(\boldsymbol{\Phi M}_t + \tilde{\boldsymbol{e}}_t\right) + \beta^A \, \boldsymbol{\varphi}_{R,t+1}^T \left(\boldsymbol{R}_t - \boldsymbol{E}_t^d\right) + \beta^A \, \varphi_{t+1} \Bigg\}. \quad (60)
$$

First order conditions (apart from geoengineering). Maximizing the right hand side over x_t yields

$$
\frac{1}{x_t} - \beta^A \, \varphi_k^A \frac{1}{1 - x_t} = 0 \quad \Longrightarrow \quad x_t = \frac{1}{1 + \beta^A \, \varphi_k^A}. \quad (61)
$$

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

$$
(1 + \beta^A \, \varphi_k^A) \frac{\frac{\partial F(\boldsymbol{A}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{N}_t, \boldsymbol{E}_t)}{\partial \mathcal{K}_{i,t}}}{F(\boldsymbol{A}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{N}_t, \boldsymbol{E}_t)} = \lambda_t^{\mathcal{K}} \quad (62)
$$

which is equivalent to

$$
\mathcal{K}_{i,t} = \frac{\sigma_{Y,\mathcal{K}_i}(\boldsymbol{A}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{N}_t, \boldsymbol{E}_t)}{\sum_{i=1}^{I_K} \sigma_{Y,\mathcal{K}_i}(\boldsymbol{A}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{N}_t, \boldsymbol{E}_t)} \quad (63)
$$

with

$$
\sigma_{Y,\mathcal{K}_i}(\boldsymbol{A}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{N}_t, \boldsymbol{E}_t) \equiv \frac{\partial F(\boldsymbol{A}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{N}_t, \boldsymbol{E}_t)}{\partial \mathcal{K}_{i,t}} \frac{\mathcal{K}_{i,t}}{F(\boldsymbol{A}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{N}_t, \boldsymbol{E}_t)}. \quad (64)
$$

Similarly, the first order conditions for the labor input is

$$
(1 + \beta^A \, \varphi_k^A) \frac{\frac{\partial F(\boldsymbol{A}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{N}_t, \boldsymbol{E}_t)}{\partial N_{i,t}}}{F(\boldsymbol{A}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{N}_t, \boldsymbol{E}_t)} = \lambda_t^N \quad (65)
$$

and hence

$$
N_{i,t} = \frac{\sigma_{Y,N_i}(\boldsymbol{A}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{N}_t, \boldsymbol{E}_t)}{\sum_{i=1}^{I_N} \sigma_{Y,N_i}(\boldsymbol{A}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{N}_t, \boldsymbol{E}_t)} \quad (66)
$$

with

$$\sigma_{Y,N_i}(\boldsymbol{A}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{N}_t, \boldsymbol{E}_t) \equiv \frac{\partial F(\boldsymbol{A}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{N}_t, \boldsymbol{E}_t)}{\partial N_{i,t}} \frac{N_{i,t}}{F(\boldsymbol{A}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{N}_t, \boldsymbol{E}_t)} \tag{67}$$

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

$$(1 + \beta^A \varphi_k^A)\frac{\frac{\partial F(\boldsymbol{A}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{N}_t, \boldsymbol{E}_t)}{\partial E_{i,t}}}{F(\boldsymbol{A}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{N}_t, \boldsymbol{E}_t)} = \beta^A(\varphi_{R,i,t+1} - \varphi_{M1}^A) \tag{68}$$

which is equivalent to

$$E_{i,t} = \frac{(1 + \beta^A \varphi_k^A)\sigma_{Y,E_i}(\boldsymbol{A}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{N}_t, \boldsymbol{E}_t)}{\beta(\varphi_{R,i,t+1} - \varphi_{M1}^A)} \tag{69}$$

with

$$\sigma_{Y,E_i}(\boldsymbol{A}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{N}_t, \boldsymbol{E}_t) \equiv \frac{\partial F(\boldsymbol{A}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{N}_t, \boldsymbol{E}_t)}{\partial E_{i,t}} \frac{E_{i,t}}{F(\boldsymbol{A}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{N}_t, \boldsymbol{E}_t)}. \tag{70}$$

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

$$B_{nc}^A(m_t, S_t^A) \equiv \beta^A \varphi_{\tau1}^{AA} \sigma_{\text{forc}} f_2\left(S_t^A + \tilde{S}_t^B(m_t)\right) + \beta^A \varphi_{\tau1}^{BA} \sigma_{\text{forc}} f_2\left(S_t^B(m_t) + \tilde{S}_t^A\right) \tag{71}$$

$$- (1 + \beta^A \varphi_k^A)[d_{AA}S_t^A + d_{BA}\tilde{S}_t^B] - \beta^A \varphi_{\tau1}^{AA} \sigma_{\text{forc}} f_3 m_t^n(S_t^A + \tilde{S}_t^B(m_t))^{1-n}$$

$$- \beta^A \varphi_{\tau1}^{BA} \sigma_{\text{forc}} f_3 m_t^n(S_t^B(m_t) + \tilde{S}_t^A)^{1-n}.$$

We reorder these terms to

$$B_{nc}^A(m_t, S_t^A) = \underbrace{\left[\beta^A \varphi_{\tau1}^{AA} \sigma_{\text{forc}} f_2 + \beta^A \varphi_{\tau1}^{BA} \sigma_{\text{forc}} f_2 \alpha_A - (1 + \beta^A \varphi_k^A)d_{AA}\right]}_{\equiv -\delta_A} S_t^A$$

$$+ \left[\beta^A \varphi_{\tau1}^{AA} \sigma_{\text{forc}} f_2 \alpha_B + \beta^A \varphi_{\tau1}^{BA} \sigma_{\text{forc}} f_2 - (1 + \beta^A \varphi_k^A)\alpha_B d_{BA}\right] S_t^B$$

$$- \beta^A \varphi_{\tau1}^{AA} \sigma_{\text{forc}} f_3 m_t^n(S_t^A + \tilde{S}_t^B(m_t))^{1-n} - \beta^A \varphi_{\tau1}^{BA} \sigma_{\text{forc}} f_3 m_t^n(S_t^B(m_t) + \tilde{S}_t^A)^{1-n}.$$

We note that only the term δ_A depends on the damage term d_{AA}, which switches magnitude as the country changes action between counter-geoengineering, no action, and geoengineering 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 $\varphi_{\tau1}^i < 0$, $i \in \{AA, BA\}$ are negative, therefore, $\delta_A > 0$.

Deriving (71) with respect to S_t^A delivers the equation

$$\frac{\partial B_{nc}^A(m_t, S_t^A)}{\partial S_t^A} = \underbrace{(n-1)\,\beta^A\,\varphi_{\tau 1}^{AA}\,\sigma_{\text{forc}}\,f_3}_{\equiv\, a_A}\, m_t^n(S_t^A + \tilde{S}_t^B(m_t))^{-n} +$$

$$\underbrace{(n-1)\beta^A\,\varphi_{\tau 1}^{BA}\,\sigma_{\text{forc}}\,f_3\alpha_A}_{\equiv\, b_A}\, m_t^n(S_t^B(m_t) + \tilde{S}_t^A)^{-n} - \delta_A.$$

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 and $\tilde{S}_t^i(m_t) = \alpha_i S_t^i(m_t)$. Defining $s^A \equiv \frac{S_t^A}{m_t}$ we rewrite the derivative as

$$\frac{\partial B_{nc}^A(m_t, S_t^A)}{\partial S_t^A} = a_A m_t^n(S_t^A + \alpha_B s_t^B m_t)^{-n} + b_A m_t^n(m_t s_t^B + \alpha_A S_t^A)^{-n} - \delta_A$$

$$= a_A(s^A + \alpha_B s_t^B)^{-n} + b_A(s_t^B + \alpha_A s^A)^{-n} - \delta_A. \tag{72}$$

The second order derivative 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

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

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

where the superindex on δ_A^c refers to the case of counter-geoengineering where $S_t^A < 0$ and $d_{AA}^c \equiv d_{AA}^c - \epsilon_A^c$ and δ_A^g refers to the case of (sulfur-based) geoengineering where $S_t^A > 0$ and $d_{AA}^g \equiv d_{AA}^g + \epsilon_A^g$. Because $d_{AA}^c < d_{AA}^g$
we have $\delta_A^c < \delta_A^g$ 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 $S_t^A = 0$. As a result we have the following cases for country A's optimal sulfur deployment:

$$\text{if } \lim_{s^a \to -0} \frac{\partial B_{nc}^A(m_t, S_t^A)}{\partial S_t^A} = (a_A \alpha_B^{-n} + b_A)s_t^{B-n} - \delta_A^c < 0 \text{ then } s^A, S_t^A < 0$$

(the interior optimum lies to the left of the kink)

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

($B_{nc}^A(m_t, S_t^A)$ is maximal at the kink)

$$\text{if } \lim_{s^a \to +0} \frac{\partial B_{nc}^A(m_t, S_t^A)}{\partial S_t^A} = (a_A \alpha_B^{-n} + b_A)s_t^{B-n} - \delta_A^g > 0 \text{ then } s^A, S_t^A > 0$$

(an interior optimum exists to the right of the kink).

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

We have shown that there exist proportionality constants s^i, $i \in \{A, B\}$, characterized by the optimality conditions above, such that the mutual optimal best responses are indeed linear functions of the emission stock. Now we solve for the values of s^A and s^B.

(i) Let $S_t^A \neq 0$ and $S_t^B \neq 0$. Then we have shown that the optimal responses follow from the interior solution to the first order condition of equation (72)

$$\underbrace{(n-1)\,\beta^A\,\varphi_{\tau 1}^{AA}\,\sigma_{\text{forc}}\,f_3}_{\equiv\, a_A \,>\, 0}\,(s^A + \alpha_B\, s^B)^{-n} + \underbrace{(n-1)\beta^A\,\varphi_{\tau 1}^{BA}\,\sigma_{\text{forc}}\,f_3\,\alpha_A}_{\equiv\, b_A \,>\, 0}(s^B + \alpha_A\, s^A)^{-n} - \delta_A = 0$$

$$(73)$$

$$\underbrace{(n-1)\,\beta^B\,\varphi_{\tau 1}^{BB}\,\sigma_{\text{forc}}\,f_3}_{\equiv\, a_B \,>\, 0}\,(s^B + \alpha_A\, s^A)^{-n} + \underbrace{(n-1)\beta^B\,\varphi_{\tau 1}^{AB}\,\sigma_{\text{forc}}\,f_3\,\alpha_B}_{\equiv\, b_B \,>\, 0}(s^A + \alpha_B\, s^B)^{-n} - \delta_B = 0$$

$$(74)$$

Similarly as for region A, we denoted the shadow price of temperature $\tau_{1,t}^B$ in region B by $\varphi_{\tau 1}^{BB}$, the shadow price of temperature $\tau_{1,t}^A$ in region B by $\varphi_{\tau 1}^{AB}$. Rearranging (74) gives

$$(s^A + \alpha_B\, s^B)^{-n} = \frac{\delta_B - a_B(s^B + \alpha_A\, s^A)^{-n}}{b_B}$$

Using this result in (73) yields

$$(s^B + \alpha_A\, s^A)^{-n} = \frac{\delta_A\, b_B - a_A\, \delta_B}{b_A\, b_B - a_A\, a_B}$$

. From this we get

$$s^B = \left(\frac{\delta_A\, b_B - a_A\, \delta_B}{b_A\, b_B - a_A\, a_B}\right)^{-\frac{1}{n}} - \alpha_A\, s^A \tag{75}$$

and

$$s^A = \left(\frac{\delta_B\, b_A - a_B\, \delta_A}{b_B\, b_A - a_B\, a_A}\right)^{-\frac{1}{n}} - \alpha_B\, s^B \tag{76}$$

and hence

$$s^B = \frac{1}{1 - \alpha_A \alpha_B} \left[\underbrace{\left(\frac{\delta_A \, b_B - a_A \, \delta_B}{b_A \, b_B - a_A \, a_B}\right)^{-\frac{1}{n}}}_{\equiv z_B} - \alpha_A \underbrace{\left(\frac{\delta_B \, b_A - a_B \, \delta_A}{b_B \, b_A - a_B \, a_A}\right)^{-\frac{1}{n}}}_{\equiv z_A} \right] \tag{77}$$

and

$$s^A = \frac{1}{1 - \alpha_A \alpha_B} \left[\underbrace{\left(\frac{\delta_B \, b_A - a_B \, \delta_A}{b_B \, b_A - a_B \, a_A}\right)^{-\frac{1}{n}}}_{\equiv z_A} - \alpha_B \underbrace{\left(\frac{\delta_A \, b_B - a_A \, \delta_B}{b_A \, b_B - a_A \, a_B}\right)^{-\frac{1}{n}}}_{\equiv z_B} \right] \tag{78}$$

We further define

$$z_A^g = \left(\frac{\delta_B \, b_A - a_B \, \delta_A^g}{b_B \, b_A - a_B \, a_A}\right)^{-\frac{1}{n}} \qquad \text{for} \qquad S_t^A > 0$$

$$z_A^c = \left(\frac{\delta_B \, b_A - a_B \, \delta_A^c}{b_B \, b_A - a_B \, a_A}\right)^{-\frac{1}{n}} \qquad \text{for} \qquad S_t^A < 0$$

This gives us

$$S_t^A = \frac{m_t}{1 - \alpha_A \alpha_B} \left(z_A^g - \alpha_B \, z_B \right) \qquad \text{for} \qquad S_t^A > 0 \tag{79}$$

and

$$S_t^A = \frac{m_t}{1 - \alpha_A \alpha_B} \left(z_A^c - \alpha_B \, z_B \right) \qquad \text{for} \qquad S_t^A < 0 \tag{80}$$

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

$$a_A \, (s^A)^{-n} + b_A (\alpha_A \, s^A)^{-n} - \delta_A^g = 0 \tag{81}$$

which is equivalent to

$$s^A = \left(\frac{\delta_A^g}{b_A \, \alpha_A^{-n} + a_A} \right)^{-\frac{1}{n}} \equiv z_A^{go} \tag{82}$$

Hence, for $S_t^B = 0$ we have $S_t^A = z_A^{go} \, m_t$. Obviously, the third and last possible case (iii) is that $S_t^A = 0$ and $S_t^B > 0$.

The following reaction functions characterize a Nash equilibrium of the dynamic game: If

$S_t^B = 0$, region A chooses $S_t^A = z_A^{go} m_t$ 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.

C.1.2 Nash Equilibria

Using the reaction functions for region A and B we derive the Nash equilibria. We have excluded the case that both countries would engage in counter-engineering by assumption.

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

$$S_t^{A*} = \frac{m_t}{1 - \alpha_A \, \alpha_B} \left(z_A^g - \alpha_B \, z_B^g \right) > 0 \quad \Rightarrow \quad z_A^g > \alpha_B \, z_B^g \tag{83}$$

and

$$S_t^{B*} = \frac{m_t}{1 - \alpha_A \, \alpha_B} \left(z_B^g - \alpha_A z_A^g \right) > 0 \quad \Rightarrow \quad z_B^g > \alpha_A z_A^g. \tag{84}$$

Together, the two equations imply

$$\alpha_B < \underbrace{\frac{z_A^g}{z_B^g}}_{\equiv H} < \alpha_A^{-1} \tag{85}$$

Note that this condition defines a non-empty range of parameter values unless $\alpha_A^n = \alpha_B^n = 1$. Thus, condition (85) states the range of parameter values $\alpha_B, \alpha_A, d_{AA}^g, d_{BB}^g, \epsilon_A^g, \epsilon_B^g, f_2$ for which there exists an equilibrium in which both regions are cooling the world given $\gamma_{AA}, \gamma_{BA}, \gamma_{BB}$ and γ_{AB}.

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

$$S_t^{A*} = \frac{m_t}{1 - \alpha_A \, \alpha_B} \left(z_A^g - \alpha_B z_B^c \right) > 0 \quad \Leftrightarrow \quad z_A^g > \alpha_B z_B^c \tag{86}$$

$$S_t^{B*} = \frac{m_t}{1 - \alpha_A \, \alpha_B} \left(z_B^c - \alpha_A z_A^g \right) < 0 \quad \Leftrightarrow \quad z_B^c < \alpha_A z_A^g. \tag{87}$$

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

by

$$\frac{z_A^g}{z_B^c} > \max\{\alpha_B, \alpha_A^{-1}\} \quad \Leftrightarrow \quad h \equiv \frac{z_A^g}{z_B^c} > \alpha_A^{-1} = \max\{\alpha_B, \alpha_A^{-1}\} \tag{88}$$

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)

$$\frac{z_B^g}{z_A^c} > \alpha_B^{-1} = \max\{\alpha_A, \alpha_B^{-1}\} \quad \Rightarrow \quad \hat{H} \equiv \frac{z_A^c}{z_B^g} < \alpha_B. \tag{89}$$

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 counter-engineering. Given region A is taking the same actions as in scenarios i and ii, region B's reaction function can neither satisfy equation (84) nor (87). Therefore, it must be that $H \geq \alpha_A^{-1}$ and $h \leq \alpha_A^{-1}$. In addition, region A's reaction function becomes

$$S_t^{A*} = z_A^{go} m_t > 0, \tag{90}$$

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

$$\frac{1}{\hat{H}} \leq \alpha_B^{-1} \leq \frac{1}{H} \quad \Leftrightarrow \quad \hat{H} \geq \alpha_B \geq H. \tag{91}$$

Here the reaction functions are

$$S_t^{B*} = z_B^{go} m_t > 0 \tag{92}$$

and locally $S_t^{A*} = 0$. These 5 cases are mutually exclusive.

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 $S_t^{A*}(\varphi_k^A, \varphi_k^B, \boldsymbol{\varphi}_{\tau A}^T, M_{1,t})$ into the maximized Bellman equation gives us

$$
\begin{aligned}
&\varphi_k^A \, k_t + \boldsymbol{\varphi}_{\tau A}^T \, \boldsymbol{\tau}_t + \boldsymbol{\varphi}_{MA}^T \, \boldsymbol{M}_t + \boldsymbol{\varphi}_{R,t}^T \, \boldsymbol{R}_t + \varphi_t \\
&= \log x_t^* + \kappa \, k_t + \log F(\boldsymbol{A}_t, \boldsymbol{\mathcal{K}}_t^*, \boldsymbol{N}_t^*, \boldsymbol{E}_t^*) + \xi_0^A \left(1 - \tau_{1,t}^A\right) - [d_{AA}S_t^{A*} + d_{BA}\tilde{S}_t^B] - a^A (m_t - 1) \\
&+ \beta^A \, \varphi_k^A \left(\kappa \, k_t + \log F(\boldsymbol{A}_t, \boldsymbol{\mathcal{K}}_t^*, \boldsymbol{N}_t^*, \boldsymbol{E}_t^*) + \log(1 - x_t^*) + \xi_0^A \left(1 - \tau_{1,t}^A\right) - [d_{AA}S_t^{A*} + d_{BA}\tilde{S}_t^B] - a^A (m_t - 1)\right) \\
&+ \beta^A \, \boldsymbol{\varphi}_{\tau A}^T \left(\boldsymbol{\sigma}\boldsymbol{\tau}_t + \tilde{\boldsymbol{F}}_t(S_t^{A*}, S_t^B)\right) + \beta^A \, \boldsymbol{\varphi}_{MA}^T \left(\boldsymbol{\Phi}\boldsymbol{M}_t + \tilde{\boldsymbol{e}}_t\right) + \beta^A \, \boldsymbol{\varphi}_{R,t+1}^T \left(\boldsymbol{R}_t - \boldsymbol{E}_t^{d*}\right) + \beta^A \, \varphi_{t+1}
\end{aligned}
\tag{93}
$$

Arranging terms with respect to their states for all Nash equilibria yields

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

$$\varphi_k^A\, k_t + \boldsymbol{\varphi}_{\tau A}^T\, \boldsymbol{\tau}_t + \varphi_{M1}^A\, M_{1,t} + \varphi_{M2}^A\, M_{2,t} + \boldsymbol{\varphi}_{R,t}^T\, \boldsymbol{R}_t + \varphi_t = \Big[(1+\beta^A\,\varphi_k^A)\kappa\Big]k_t + \beta^A\,\boldsymbol{\varphi}_{\tau A}^T\,\boldsymbol{\sigma}\boldsymbol{\tau}_t$$

$$-(1+\beta^A\,\varphi_k^A)\xi_0^A\, \boldsymbol{e}_1^T\, \boldsymbol{\tau}_t + \Bigg[\Bigg(f_2(\beta^A\,\varphi_{\tau 1}^{AA}\,\sigma_{\text{forc}}\, z_A + \beta^A\,\varphi_{\tau 1}^{BA}\,\sigma_{\text{forc}}\, z_B) - (1+\beta^A\,\varphi_k^A)\Big(d_{AA}\,\frac{1}{1-\alpha_A\,\alpha_B}(z_A - \alpha_B\, z_B)$$

$$+ d_{BA}\,\frac{\alpha_B}{1-\alpha_A\,\alpha_B}(z_B - \alpha_A\, z_A)\Big) - f_3(\beta^A\,\varphi_{\tau 1}^{AA}\,\sigma_{\text{forc}}\, z_A^{1-n} + \beta^A\,\varphi_{\tau 1}^{BA}\,\sigma_{\text{forc}}\, z_B^{1-n}) - a^A(1+\beta^A\,\varphi_k^A)$$

$$+ (\beta^A\,\varphi_{\tau 1}^{AA}\,\sigma_{\text{forc}} + \beta^A\,\varphi_{\tau 1}^{BA}\,\sigma_{\text{forc}})\, f_1\Bigg)M_{pre}^{-1} + \beta^A\,(\varphi_{M1}^A\,\phi_{11} + \varphi_{M2}^A\,\phi_{12})\Bigg]M_{1,t} + \Big[\beta^A(\varphi_{M1}^A\,\phi_{21} + \varphi_{M2}^A\,\phi_{22})\Big]M_{2,t}$$

$$+ \Big[\beta^A\,\boldsymbol{\varphi}_{R,t+1}^T\Big]R_t + \log x_t^* + \beta^A\,\varphi_k^A\,\log(1-x_t^*) + (1+\beta^A\,\varphi_k^A)\log F(\boldsymbol{A}_t, \boldsymbol{\mathcal{K}}_t^*, \boldsymbol{N}_t^*, \boldsymbol{E}_t^*) + (1+\beta^A\,\varphi_k^A)(\xi_0^A + a^A)$$

$$+ (\beta^A\,\varphi_{\tau 1}^{AA}\,\sigma_{\text{forc}} + \beta^A\,\varphi_{\tau 1}^{BA}\,\sigma_{\text{forc}})\, f_0 + \beta^A\,\varphi_{M1}^A(\sum_{i=1}^{I^d} E_{i,t}^* + E_t^{\text{exo}}) - \beta^A\,\boldsymbol{\varphi}_{R,t+1}^T\, \boldsymbol{E}_t^{d*} + \beta^A\,\varphi_{t+1}, \quad (94)$$

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

$$\varphi_k^A\, k_t + \boldsymbol{\varphi}_{\tau A}^T\, \boldsymbol{\tau}_t + \varphi_{M1}^A\, M_{1,t} + \varphi_{M2}^A\, M_{2,t} + \boldsymbol{\varphi}_{R,t}^T\, \boldsymbol{R}_t + \varphi_t = \Big[(1+\beta^A\,\varphi_k^A)\kappa\Big]k_t + \beta^A\,\boldsymbol{\varphi}_{\tau A}^T\,\boldsymbol{\sigma}\boldsymbol{\tau}_t$$

$$- (1+\beta^A\,\varphi_k^A)\xi_0^A\, \boldsymbol{e}_1^T\, \boldsymbol{\tau}_t + \Bigg[\Bigg(f_2\,(\beta^A\,\varphi_{\tau 1}^{AA}\,\sigma_{\text{forc}}\, z_A^{go} + \beta^A\,\varphi_{\tau 1}^{BA}\,\sigma_{\text{forc}}\,\alpha_A\, z_A^{go}) - (1+\beta^A\,\varphi_k^A)\, d_{AA}\, z_A^{go}$$

$$- f_3\Big(\beta^A\,\varphi_{\tau 1}^{AA}\,\sigma_{\text{forc}}\,(z_A^{go})^{1-n} + \beta^A\,\varphi_{\tau 1}^{BA}\,\sigma_{\text{forc}}\,(\alpha_A\, z_A^{go})^{1-n}\Big) - (1+\beta^A\,\varphi_k^A)\, a^A + f_1\,(\beta^A\,\varphi_{\tau 1}^{AA}\,\sigma_{\text{forc}}$$

$$+ \beta^A\,\varphi_{\tau 1}^{BA}\,\sigma_{\text{forc}})\Bigg)M_{pre}^{-1} + \beta^A\,(\varphi_{M1}^A\,\phi_{11} + \varphi_{M2}^A\,\phi_{12})\Bigg]M_{1,t} + \Big[\beta^A(\varphi_{M1}^A\,\phi_{21} + \varphi_{M2}^A\,\phi_{22})\Big]M_{2,t} + \Big[\beta^A\,\boldsymbol{\varphi}_{R,t+1}^T\Big]R_t$$

$$+ \log x_t^* + \beta^A\,\varphi_k^A\,\log(1-x_t^*) + (1+\beta^A\,\varphi_k^A)\log F(\boldsymbol{A}_t, \boldsymbol{\mathcal{K}}_t^*, \boldsymbol{N}_t^*, \boldsymbol{E}_t^*) + (1+\beta^A\,\varphi_k^A)(\xi_0^A + a^A)$$

$$+ (\beta^A\,\varphi_{\tau 1}^{AA}\,\sigma_{\text{forc}} + \beta^A\,\varphi_{\tau 1}^{BA}\,\sigma_{\text{forc}})\, f_0 + \beta^A\,\varphi_{M1}^A(\sum_{i=1}^{I^d} E_{i,t}^* + E_t^{\text{exo}}) - \beta^A\,\boldsymbol{\varphi}_{R,t+1}^T\, \boldsymbol{E}_t^{d*} + \beta^A\,\varphi_{t+1}, \quad (95)$$

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

$$\varphi_k^A \, k_t + \boldsymbol{\varphi}_{\tau A}^T \, \boldsymbol{\tau}_t + \varphi_{M1}^A \, M_{1,t} + \varphi_{M2}^A \, M_{2,t} + \boldsymbol{\varphi}_{R,t}^T \, \boldsymbol{R}_t + \varphi_t = \left[(1 + \beta^A \, \varphi_k^A)\kappa\right] k_t + \beta^A \, \boldsymbol{\varphi}_{\tau A}^T \, \boldsymbol{\sigma}\boldsymbol{\tau}_t$$

$$- (1 + \beta^A \, \varphi_k^A)\xi_0^A \, \boldsymbol{e}_1^T \, \boldsymbol{\tau}_t + \Bigg[\bigg(\Big(f_2 \,(\beta^A \, \varphi_{\tau 1}^{AA} \, \sigma_{\text{forc}} \, \alpha_B \, z_B^{go} + \beta^A \, \varphi_{\tau 1}^{BA} \, \sigma_{\text{forc}} \, z_B^{go}) - (1 + \beta^A \, \varphi_k^A) \, d_{BA} \, \alpha_B \, z_B^{go}$$

$$- f_3\Big(\beta^A \, \varphi_{\tau 1}^{AA} \, \sigma_{\text{forc}} \, (\alpha_B \, z_B^{go})^{1-n} + \beta^A \, \varphi_{\tau 1}^{BA} \, \sigma_{\text{forc}} \, (z_B^{go})^{1-n}\Big) - (1 + \beta^A \, \varphi_k^A) \, a^A + f_1 \,(\beta^A \, \varphi_{\tau 1}^{AA} \, \sigma_{\text{forc}}$$

$$+ \beta^A \, \varphi_{\tau 1}^{BA} \, \sigma_{\text{forc}}\,)\bigg) M_{pre}^{-1} + \beta^A \,(\varphi_{M1}^A \, \phi_{11} + \varphi_{M2}^A \, \phi_{12})\Bigg] M_{1,t} + \Big[\beta^A(\varphi_{M1}^A \, \phi_{21} + \varphi_{M2}^A \, \phi_{22})\Big] M_{2,t} + \Big[\beta^A \, \boldsymbol{\varphi}_{R,t+1}^T\Big] R_t$$

$$+ \log x_t^* + \beta^A \, \varphi_k^A \log(1 - x_t^*) + (1 + \beta^A \, \varphi_k^A) \log F(\boldsymbol{A}_t, \boldsymbol{\mathcal{K}}_t^*, \boldsymbol{N}_t^*, \boldsymbol{E}_t^*) + (1 + \beta^A \, \varphi_k^A)(\xi_0^A + a^A)$$

$$+ (\beta^A \, \varphi_{\tau 1}^{AA} \, \sigma_{\text{forc}} + \beta^A \, \varphi_{\tau 1}^{BA} \, \sigma_{\text{forc}}) \, f_0 + \beta^A \, \varphi_{M1}^A \Big(\sum_{i=1}^{I^d} E_{i,t}^* + E_t^{\text{exo}}\Big) - \beta^A \, \boldsymbol{\varphi}_{R,t+1}^T \, \boldsymbol{E}_t^{d*} + \beta^A \, \varphi_{t+1}. \quad (96)$$

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

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

Inserting φ_k into equation (61) yields the optimal consumption rate $x_t^* = 1 - \beta^A \, \kappa$.

Coefficient matching with respect to transformed temperatures delivers

$$\boldsymbol{\varphi}_{\tau A}^T = -\xi_0^A \,(1 + \beta^A \, \varphi_k^A) \, \boldsymbol{e}_1^T \,[\boldsymbol{1} - \beta^A \, \boldsymbol{\sigma}]^{-1}. \quad (97)$$

with

$$[\boldsymbol{1} - \beta^A \, \boldsymbol{\sigma}]^{-1} = \begin{pmatrix} \tilde{\sigma}_{11}^A & \tilde{\sigma}_{12}^A & \tilde{\sigma}_{13}^A \\ \tilde{\sigma}_{21}^A & \tilde{\sigma}_{22}^A & \tilde{\sigma}_{23}^A \\ \tilde{\sigma}_{31}^A & \tilde{\sigma}_{32}^A & \tilde{\sigma}_{33}^A \end{pmatrix}$$

and hence

$$\varphi_{\tau 1}^{AA} = -\xi_0^A \,(1 + \beta^A \, \varphi_k^A)\tilde{\sigma}_{11}^A, \quad (98)$$

$$\varphi_{\tau 1}^{BA} = -\xi_0^A \,(1 + \beta^A \, \varphi_k^A)\tilde{\sigma}_{21}^A, \quad (99)$$

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

- -

Simplification: If we set the heat flow coefficients σ_B^A, σ_A^B, σ_A^O, σ_B^O equal to zero, then $\varphi_{\tau 1}^{AA} = -\xi_0^A \,(1 + \beta^A \, \varphi_k^A) \, \tilde{\sigma}_{11}^A$ and $\varphi_{\tau 1}^{BA} = 0$ since $\tilde{\sigma}_{21}^A = 0$. This implies that $b_A = b_B = 0$, and therefore

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

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

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 $z_B \in \{z_B^c, z_B^g\}$. We used the simpler notation $\varphi_{\tau 1}^{AA} = \varphi_{\tau 1}^A = -\xi_0^A\left(1+\beta^A\,\varphi_k^A\right)\tilde{\sigma}_{11}^A$ and $\varphi_{\tau 1}^{BB} = \varphi_{\tau 1}^B = -\xi_0^B\left(1+\beta^B\,\varphi_k^B\right)\tilde{\sigma}_{11}^B$. Note that with 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 analoug result for region B.

We can further see that the cooling undertaken by the cooling region will always go up if the other region engages in counter-engineering. The increase in cooling is less strong if the other region's reluctance to engage in counter-engineering is higher. For this statement to be true, the following equation must hold:

$$\frac{m_t}{1-\alpha_A\,\alpha_B}\left(z_B^g - \alpha_A\,z_A^c\right) > m_t\,z_B^g$$

Rearranging yields $\alpha_B\,z_B^c > z_A^c$ which is always satisfied for $S_t^A < 0$.

– –

We now define $\gamma_{AA} \equiv \beta^A\,\xi_0^A\,\tilde{\sigma}_{11}^A\,\sigma_{\text{forc}}$, and $\gamma_{BA} \equiv \beta^A\,\xi_0^A\,\tilde{\sigma}_{21}^A\,\sigma_{\text{forc}}$. This gives us δ_A, δ_B, a_A, a_B, b_A, and b_B as a function of the γ's.

$$-\delta_A = (1+\beta^A\,\varphi_k^A)(\gamma_{AA}\,f_2 + \alpha_A\,\gamma_{BA}\,f_2 - d_{AA})$$

$$a_A = (1+\beta^A\,\varphi_k^A)(n-1)\gamma_{AA}\,f_3$$

$$b_A = (1+\beta^A\,\varphi_k^A)(n-1)\alpha_A\,\gamma_{BA}\,f_3$$

Thus, z_A and z_B are also functions of the γ's.

$$z_A(\gamma_{AA},\gamma_{BA},\gamma_{BB},\gamma_{AB}) = \left(\frac{\delta_B(\gamma_{BB},\gamma_{AB})\,b_A(\gamma_{BA}) - a_B(\gamma_{BB})\,\delta_A(\gamma_{AA},\gamma_{BA})}{b_B(\gamma_{AB})\,b_A(\gamma_{BA}) - a_B(\gamma_{BB})\,a_A(\gamma_{AA})}\right)^{-\frac{1}{n}}$$

$$z_B(\gamma_{BB},\gamma_{AB},\gamma_{AA},\gamma_{BA}) = \left(\frac{\delta_A(\gamma_{AA},\gamma_{BA})\,b_B(\gamma_{AB}) - a_A(\gamma_{AA})\,\delta_B(\gamma_{BB},\gamma_{AB})}{b_A(\gamma_{BA})\,b_B(\gamma_{AB}) - a_A(\gamma_{AA})\,a_B(\gamma_{BB})}\right)^{-\frac{1}{n}}$$

– –

Simplification: Setting σ_B^A, σ_A^B, σ_A^O, σ_B^O equal to zero leads to

$$z_A^g = \left(\frac{(1-n)\,f_3\,\gamma_A}{f_2\,\gamma_A + (d_{AA}^g + \epsilon_A^g)}\right)^{\frac{1}{n}}, \quad z_A^c = \left(\frac{(1-n)\,f_3\,\gamma_A}{f_2\,\gamma_A + (d_{AA}^c - \epsilon_A^c)}\right)^{\frac{1}{n}} \quad \text{where} \quad \gamma_A = \beta^A\,\xi_0^A\,\tilde{\sigma}_A^{-1}\,\sigma_{\text{forc}}.$$

Switching region indices leads shows that for region B

$$z_B^g = \left(\frac{(1-n)\,f_3\,\gamma_B}{f_2\,\gamma_B + (d_{BB}^g + \epsilon_B^g)}\right)^{\frac{1}{n}}, \quad z_B^c = \left(\frac{(1-n)\,f_3\,\gamma_B}{f_2\,\gamma_B + (d_{BB}^c - \epsilon_B^c)}\right)^{\frac{1}{n}} \quad \text{where} \quad \gamma_B = \beta^B\,\xi_0^B\,\tilde{\sigma}_B^{-1}\,\sigma_{\text{forc}}.$$

– –

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

$$\varphi_{M1}^A = (1+\beta^A\,\varphi_k^A)\left(-f_2\,(\gamma_{AA}\,z_A + \gamma_{BA}\,z_B) - \left(d_{AA}\,\frac{1}{1-\alpha_A\,\alpha_B}(z_A - \alpha_B\,z_B) + d_{BA}\,\frac{\alpha_B}{1-\alpha_A\,\alpha_B}(z_B - \alpha_A\,z_A)\right)\right.$$

$$\left. + f_3\,(\gamma_{AA}\,z_A^{1-n} + \gamma_{BA}\,z_B^{1-n}) - a^A - f_1\,(\gamma_{AA} + \gamma_{BA})\right) M_{pre}^{-1} + \beta^A\,(\varphi_{M1}^A\,\phi_{11} + \varphi_{M2}^A\,\phi_{12}) \quad (100)$$

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

$$\varphi_{M1}^A = (1+\beta^A\,\varphi_k^A)\left(-f_2\,(\gamma_{AA}\,z_A^{go} + \gamma_{BA}\,\alpha_A\,z_A^{go}) - d_{AA}\,z_A^{go} + f_3\left(\gamma_{AA}\,(z_A^{go})^{1-n} + \gamma_{BA}\,(\alpha_A\,z_A^{go})^{1-n}\right) - a^A \right.$$

$$\left. - f_1\,(\gamma_{AA} + \gamma_{BA})\right) M_{pre}^{-1} + \beta^A\,(\varphi_{M1}^A\,\phi_{11} + \varphi_{M2}^A\,\phi_{12}) \quad (101)$$

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

$$\varphi_{M1}^A = (1+\beta^A\,\varphi_k^A)\left(-f_2\,(\gamma_{AA}\,\alpha_B\,z_B^{go} + \gamma_{BA}\,z_B^{go}) - d_{BA}\,\alpha_B\,z_B^{go} + f_3\left(\gamma_{AA}\,(\alpha_B\,z_B^{go})^{1-n} + \gamma_{BA}\,(z_B^{go})^{1-n}\right) - a^A \right.$$

$$\left. - f_1\,(\gamma_{AA} + \gamma_{BA})\right) M_{pre}^{-1} + \beta^A\,(\varphi_{M1}^A\,\phi_{11} + \varphi_{M2}^A\,\phi_{12}) \quad (102)$$

Coefficient matching with respect to the carbon stock in the ocean gets us

$$\varphi_{M2}^A = \beta^A(\varphi_{M1}^A\,\phi_{21} + \varphi_{M2}^A\,\phi_{22}) \quad \Leftrightarrow \quad \varphi_{M2}^A = \frac{\beta^A\,\varphi_{M1}^A\,\phi_{21}}{1 - \beta^A\,\phi_{22}},$$

and for the resource stock we have

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

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 (94),(95), and (96) turn to the following condition:

$$\varphi_t - \beta^A \varphi_{t+1} = \log x_t^* + \beta^A \varphi_k^A \log(1-x_t^*) + (1+\beta^A \varphi_k^A) \log F(\boldsymbol{A}_t, \boldsymbol{\mathcal{K}}_t^*, \boldsymbol{N}_t^*, \boldsymbol{E}_t^*) + (1+\beta^A \varphi_k^A)(\xi_0^A + a^A)$$

$$+ (\beta^A \varphi_{\tau 1}^{AA} \sigma_{\text{forc}} + \beta^A \varphi_{\tau 1}^{BA} \sigma_{\text{forc}}) f_0 + \beta \varphi_{M1}^A (\sum_{i=1}^{I^d} E_{i,t}^* + E_t^{\text{exo}}) - \beta^A \boldsymbol{\varphi}_{R,t+1}^T \boldsymbol{E}_t^{d*} \quad (103)$$

This condition will be satisfied by picking the sequence $\varphi_0, \varphi_1, \varphi_2, \dots$ 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.

C.1.3 Regional social cost of carbon

Inserting φ_{M2}^A, and φ_k^A into (100), (101), and (102) delivers:
(i): $S_t^A \neq 0$ and $S_t^B \neq 0$

$$\varphi_{M1}^A = \frac{1}{1-\beta^A \kappa} \left(-f_2 (\gamma_{AA} z_A + \gamma_{BA} z_B) - \left(d_{AA} \frac{1}{1-\alpha_A \alpha_B} (z_A - \alpha_B z_B) + d_{BA} \frac{\alpha_B}{1-\alpha_A \alpha_B} (z_B - \alpha_A z_A) \right) \right.$$

$$\left. + f_3 (\gamma_{AA} z_A^{1-n} + \gamma_{BA} z_B^{1-n}) - a^A - f_1 (\gamma_{AA} + \gamma_{BA}) \right) M_{pre}^{-1} \underbrace{\left(1 - \beta^A \phi_{11} - \frac{(\beta^A)^2 \phi_{12} \phi_{21}}{1-\beta^A \phi_{22}} \right)^{-1}}_{\equiv \tilde{\phi}_A}$$

$$(104)$$

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

$$\varphi_{M1}^A = \frac{1}{1-\beta^A \kappa} \left(-f_2 (\gamma_{AA} z_A^{go} + \gamma_{BA} \alpha_A z_A^{go}) - d_{AA} z_A^{go} + f_3 \left(\gamma_{AA} (z_A^{go})^{1-n} + \gamma_{BA} (\alpha_A z_A^{go})^{1-n} \right) - a^A \right.$$

$$\left. - f_1 (\gamma_{AA} + \gamma_{BA}) \right) M_{pre}^{-1} \underbrace{\left(1 - \beta^A \phi_{11} - \frac{(\beta^A)^2 \phi_{12} \phi_{21}}{1-\beta^A \phi_{22}} \right)^{-1}}_{\equiv \tilde{\phi}_A} \quad (105)$$

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

$$\varphi_{M1}^A = \frac{1}{1-\beta^A \kappa} \left(-f_2 (\gamma_{AA} \alpha_B z_B^{go} + \gamma_{BA} z_B^{go}) - d_{BA} \alpha_B z_B^{go} + f_3 \left(\gamma_{AA} (\alpha_B z_B^{go})^{1-n} + \gamma_{BA} (z_B^{go})^{1-n} \right) - a^A \right.$$

$$\left. - f_1 (\gamma_{AA} + \gamma_{BA}) \right) M_{pre}^{-1} \underbrace{\left(1 - \beta^A \phi_{11} - \frac{(\beta^A)^2 \phi_{12} \phi_{21}}{1-\beta^A \phi_{22}} \right)^{-1}}_{\equiv \tilde{\phi}_A} \quad (106)$$

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 $S_t^B \neq 0$

$$SCC^A = -(1 - \beta^A \kappa)Y_{A,t}^{net} \varphi_{M1}^A$$

$$= \frac{Y_{A,t}^{net}}{M_{pre}} \Bigg(f_2\left(\gamma_{AA}\, z_A + \gamma_{BA}\, z_B\right) + \Big(d_{AA}\, \frac{1}{1 - \alpha_A \alpha_B}(z_A - \alpha_B z_B) + d_{BA}\, \frac{\alpha_B}{1 - \alpha_A \alpha_B}(z_B - \alpha_A z_A)\Big)$$

$$-f_3\left(\gamma_{AA}\, z_A^{1-n} + \gamma_{BA}\, z_B^{1-n}\right) + a^A + f_1\left(\gamma_{AA} + \gamma_{BA}\right) \Bigg) \tilde{\phi}_A,$$

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

$$SCC^A = -(1 - \beta^A \kappa)Y_{A,t}^{net} \varphi_{M1}^A$$

$$= \frac{Y_{A,t}^{net}}{M_{pre}} \Bigg(f_2\left(\gamma_{AA}\, z_A^{go} + \gamma_{BA}\, \alpha_A\, z_A^{go}\right) + d_{AA}\, z_A^{go} - f_3\Big(\gamma_{AA}\, (z_A^{go})^{1-n} + \gamma_{BA}\, (\alpha_A\, z_A^{go})^{1-n}\Big) + a^A$$

$$+f_1\left(\gamma_{AA} + \gamma_{BA}\right) \Bigg) \tilde{\phi}_A,$$

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

$$SCC^A = -(1 - \beta^A \kappa)Y_{A,t}^{net} \varphi_{M1}^A$$

$$= \frac{Y_{A,t}^{net}}{M_{pre}} \Bigg(f_2\left(\gamma_{AA}\, \alpha_B\, z_B^{go} + \gamma_{BA}\, z_B^{go}\right) + d_{BA}\, \alpha_B\, z_B^{go} - f_3\Big(\gamma_{AA}\, (\alpha_B\, z_B^{go})^{1-n} + \gamma_{BA}\, (z_B^{go})^{1-n}\Big) + a^A$$

$$+f_1\left(\gamma_{AA} + \gamma_{BA}\right) \Bigg) \tilde{\phi}_A.$$

$- -$

Simplification: Setting σ_B^A, σ_A^B, σ_A^O, σ_B^O equal to zero, the regional social cost of carbon becomes

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

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

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

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

$$SCC^A = \frac{Y_{A,t}^{net}}{M_{pre}} \Bigg(a^A + f_1\, \gamma_A - \frac{n}{1-n}\, z_A^g(f_2\, \gamma_A + d_{AA}) \Bigg) \tilde{\phi}_A,$$

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

$$SCC^A = \frac{Y_{A,t}^{net}}{M_{pre}} \left[a^A + f_1\, \gamma_A + \alpha_B\, z_B^g\, d_{BA}^g - \gamma_A \Big(f_3 (\alpha_B\, z_B^g)^{1-n} - \alpha_B\, z_B^g\, f_2 \Big) \right] \tilde{\phi}_A,$$

- -

C.2 Rest of the world

In terms of transformed temperatures in the regional model we have

$$
\underbrace{\begin{pmatrix} \tau_{1,t+1}^A \\ \tau_{1,t+1}^B \\ \tau_{2,t+1} \end{pmatrix}}_{\equiv\, \boldsymbol{\tau}_{t+1}}
=
\underbrace{\begin{pmatrix} 1-\sigma^A & \sigma_B^A & \sigma_O^A \\ \sigma_A^B & 1-\sigma^B & \sigma_O^B \\ \sigma_A^O & \sigma_B^O & 1-\sigma^O \end{pmatrix}}_{\equiv\, \boldsymbol{\sigma}}
\underbrace{\begin{pmatrix} \tau_{1,t}^A \\ \tau_{1,t}^B \\ \tau_{2,t} \end{pmatrix}}_{\equiv\, \boldsymbol{\tau}_t}
+
\underbrace{\begin{pmatrix} \sigma_{\text{forc}} \exp(\frac{\log 2}{\eta} F_t^A) \\ \sigma_{\text{forc}} \exp(\frac{\log 2}{\eta} F_t^B) \\ 0 \end{pmatrix}}_{\equiv\, \tilde{\boldsymbol{F}}_t(S_t^A, S_t^B)}
\tag{107}
$$

where $\sigma^A = \sigma_B^A + \sigma_O^A$, $\sigma^B = \sigma_A^B + \sigma_O^B$, and $\sigma^O = \sigma_A^O + \sigma_B^O$, or equivalently

$$\boldsymbol{\tau}_{t+1} = \boldsymbol{\sigma} \boldsymbol{\tau}_t + \tilde{\boldsymbol{F}}_t(S_t^A, S_t^B). \tag{108}$$

The dynamics of the carbon reservoirs are given by

$$
\underbrace{\begin{pmatrix} M_{1,t+1} \\ M_{2,t+1} \end{pmatrix}}_{\equiv\, \boldsymbol{M}_{t+1}}
=
\underbrace{\begin{pmatrix} \phi_{11} & \phi_{21} \\ \phi_{12} & \phi_{22} \end{pmatrix}}_{\equiv\, \boldsymbol{\Phi}}
\underbrace{\begin{pmatrix} M_{1,t} \\ M_{2,t} \end{pmatrix}}_{\equiv\, \boldsymbol{M}_t}
+
\underbrace{\begin{pmatrix} \sum_{i=1}^{I^d} E_{A,i,t} + \sum_{i=1}^{I^d} E_{B,i,t} + \sum_{i=1}^{I^d} E_{W,i,t} + E_t^{\text{exo}} \\ 0 \end{pmatrix}}_{\equiv\, \tilde{\boldsymbol{e}}_t}
\tag{109}
$$

or equivalently

$$\boldsymbol{M}_{t+1} = \boldsymbol{\Phi} \boldsymbol{M}_t + \tilde{\boldsymbol{e}}_t. \tag{110}$$

In the following we show for the rest of the world that the system (4 temperatures and 2 carbon reservoirs) is linear in states and that the affine value function

$$V(k_t, \boldsymbol{\tau}_t, \boldsymbol{M}_t, \boldsymbol{R}_t, t) = \varphi_k^W\, k_t + \boldsymbol{\varphi}_{\tau W}^T\, \boldsymbol{\tau}_t + \boldsymbol{\varphi}_{MW}^T\, \boldsymbol{M}_t + \boldsymbol{\varphi}_{R,t}^T\, \boldsymbol{R}_t + \varphi_t, \tag{111}$$

where

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

solves the system. We only use region indices when they are needed. Inserting the trial solution and the next periods states into the Bellman equation delivers

$$\varphi_k^W k_t + \boldsymbol{\varphi}_{\tau W}^T \boldsymbol{\tau}_t + \boldsymbol{\varphi}_{MW}^T \boldsymbol{M}_t + \boldsymbol{\varphi}_{R,t}^T \boldsymbol{R}_t + \varphi_t$$

$$= \max_{x_t, \boldsymbol{N}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{E}_t} \left\{ \log x_t + \kappa\, k_t + \log F(\boldsymbol{A}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{N}_t, \boldsymbol{E}_t) + \xi_0^W \left(1 - \tau_{1,t}^W\right) - (d_{BW} S_t^B + d_{AW} \tilde{S}_t^A) \right.$$

$$-a^W (m_t - 1) + \lambda_t^{\mathcal{K}} (1 - \sum_{i=1}^{I_K} \mathcal{K}_{i,t}) + \lambda_t^N (1 - \sum_{i=1}^{I_N} N_{i,t}) + \beta^W \varphi_k^W \left(\kappa\, k_t + \log F(\boldsymbol{A}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{N}_t, \boldsymbol{E}_t) + \log(1 - x_t) \right.$$

$$+ \xi_0^W \left(1 - \tau_{1,t}^W\right) - (d_{BW} S_t^B + d_{AW} \tilde{S}_t^A) - a^W (m_t - 1) \Big) + \beta^W \boldsymbol{\varphi}_{\tau W}^T \left(\boldsymbol{\sigma} \boldsymbol{\tau}_t + \tilde{\boldsymbol{F}}_t (S_t^A, S_t^B) \right)$$

$$\left. + \beta^W \boldsymbol{\varphi}_{MW}^T \left(\boldsymbol{\Phi} \boldsymbol{M}_t + \tilde{\boldsymbol{e}}_t \right) + \beta^W \boldsymbol{\varphi}_{R,t+1}^T \left(\boldsymbol{R}_t - \boldsymbol{E}_t^d \right) + \beta^W \varphi_{t+1} \right\}. \quad (112)$$

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

$$\frac{1}{x_t} - \beta^W \varphi_k^W \frac{1}{1 - x_t} = 0 \quad \Longrightarrow \quad x_t = \frac{1}{1 + \beta^W \varphi_k^W}. \quad (113)$$

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

$$(1 + \beta^W \varphi_k^W) \frac{\frac{\partial F(\boldsymbol{A}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{N}_t, \boldsymbol{E}_t)}{\partial \mathcal{K}_{i,t}}}{F(\boldsymbol{A}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{N}_t, \boldsymbol{E}_t)} = \lambda_t^{\mathcal{K}} \quad (114)$$

which is equivalent to

$$\mathcal{K}_{i,t} = \frac{\sigma_{Y,\mathcal{K}_i}(\boldsymbol{A}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{N}_t, \boldsymbol{E}_t)}{\sum_{i=1}^{I_K} \sigma_{Y,\mathcal{K}_i}(\boldsymbol{A}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{N}_t, \boldsymbol{E}_t)} \quad (115)$$

with

$$\sigma_{Y,\mathcal{K}_i}(\boldsymbol{A}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{N}_t, \boldsymbol{E}_t) \equiv \frac{\partial F(\boldsymbol{A}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{N}_t, \boldsymbol{E}_t)}{\partial \mathcal{K}_{i,t}} \frac{\mathcal{K}_{i,t}}{F(\boldsymbol{A}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{N}_t, \boldsymbol{E}_t)}. \quad (116)$$

Similarly, the first order conditions for the labor input is

$$(1 + \beta^W \varphi_k^W) \frac{\frac{\partial F(\boldsymbol{A}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{N}_t, \boldsymbol{E}_t)}{\partial N_{i,t}}}{F(\boldsymbol{A}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{N}_t, \boldsymbol{E}_t)} = \lambda_t^N \quad (117)$$

and hence

$$N_{i,t} = \frac{\sigma_{Y,N_i}(\boldsymbol{A}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{N}_t, \boldsymbol{E}_t)}{\sum_{i=1}^{I_N} \sigma_{Y,N_i}(\boldsymbol{A}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{N}_t, \boldsymbol{E}_t)} \quad (118)$$

with

$$\sigma_{Y,N_i}(\boldsymbol{A}_t,\,\boldsymbol{\mathcal{K}}_t,\,\boldsymbol{N}_t,\,\boldsymbol{E}_t) \equiv \frac{\partial F(\boldsymbol{A}_t,\,\boldsymbol{\mathcal{K}}_t,\,\boldsymbol{N}_t,\,\boldsymbol{E}_t)}{\partial N_{i,t}} \frac{N_{i,t}}{F(\boldsymbol{A}_t,\,\boldsymbol{\mathcal{K}}_t,\,\boldsymbol{N}_t,\,\boldsymbol{E}_t)} \tag{119}$$

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

$$(1+\beta^W\,\varphi_k^W)\frac{\frac{\partial F(\boldsymbol{A}_t,\boldsymbol{\mathcal{K}}_t,\boldsymbol{N}_t,\boldsymbol{E}_t)}{\partial E_{i,t}}}{F(\boldsymbol{A}_t,\,\boldsymbol{\mathcal{K}}_t,\,\boldsymbol{N}_t,\,\boldsymbol{E}_t)} = \beta^W(\varphi_{R,i,t+1}-\varphi_{M1}^W) \tag{120}$$

which is equivalent to

$$E_{i,t} = \frac{(1+\beta^W\,\varphi_k^W)\sigma_{Y,E_i}(\boldsymbol{A}_t,\,\boldsymbol{\mathcal{K}}_t,\,\boldsymbol{N}_t,\,\boldsymbol{E}_t)}{\beta(\varphi_{R,i,t+1}-\varphi_{M1}^W)} \tag{121}$$

with

$$\sigma_{Y,E_i}(\boldsymbol{A}_t,\,\boldsymbol{\mathcal{K}}_t,\,\boldsymbol{N}_t,\,\boldsymbol{E}_t) \equiv \frac{\partial F(\boldsymbol{A}_t,\,\boldsymbol{\mathcal{K}}_t,\,\boldsymbol{N}_t,\,\boldsymbol{E}_t)}{\partial E_{i,t}} \frac{E_{i,t}}{F(\boldsymbol{A}_t,\,\boldsymbol{\mathcal{K}}_t,\,\boldsymbol{N}_t,\,\boldsymbol{E}_t)}. \tag{122}$$

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

$$
\begin{aligned}
&\varphi_k^W\,k_t + \boldsymbol{\varphi}_{\tau W}^T\,\boldsymbol{\tau}_t + \boldsymbol{\varphi}_{MW}^T\,\boldsymbol{M}_t + \boldsymbol{\varphi}_{R,t}^T\,\boldsymbol{R}_t + \varphi_t \\
&= \log x_t^* + \kappa\,k_t + \log F(\boldsymbol{A}_t,\,\boldsymbol{\mathcal{K}}_t^*,\,\boldsymbol{N}_t^*,\,\boldsymbol{E}_t^*) + \xi_0^W\,\left(1-\tau_{1,t}^W\right) - (d_{BW}S_t^B + d_{AW}\tilde{S}_t^A) - a^W\,(m_t-1) \\
&+\beta^W\,\varphi_k^W\,\left(\kappa\,k_t + \log F(\boldsymbol{A}_t,\,\boldsymbol{\mathcal{K}}_t^*,\,\boldsymbol{N}_t^*,\,\boldsymbol{E}_t^*) + \log(1-x_t^*) + \xi_0^W\,\left(1-\tau_{1,t}^W\right) - (d_{BW}S_t^B + d_{AW}\tilde{S}_t^A) - a^W\,(m_t-1)\right) \\
&+\beta^W\,\boldsymbol{\varphi}_{\tau W}^T\,\left(\boldsymbol{\sigma}\boldsymbol{\tau}_t + \tilde{\boldsymbol{F}}_t(S_t^A,S_t^B)\right) + \beta^W\,\boldsymbol{\varphi}_{MW}^T\,\left(\boldsymbol{\Phi}\boldsymbol{M}_t + \tilde{\boldsymbol{e}}_t\right) + \beta^W\,\boldsymbol{\varphi}_{R,t+1}^T\,\left(\boldsymbol{R}_t - \boldsymbol{E}_t^{d*}\right) + \beta^W\,\varphi_{t+1}
\end{aligned}
\tag{123}
$$

Arranging with respect to states for all Nash equilibria yields

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

$$\varphi_k^W \, k_t + \boldsymbol{\varphi}_{\tau W}^T \, \boldsymbol{\tau}_t + \varphi_{M1}^W \, M_{1,t} + \varphi_{M2}^W \, M_{2,t} + \boldsymbol{\varphi}_{R,t}^T \, \boldsymbol{R}_t + \varphi_t = \Big[(1 + \beta^W \, \varphi_k^W)\kappa\Big] k_t + \beta^W \, \boldsymbol{\varphi}_{\tau W}^T \, \boldsymbol{\sigma \tau}_t$$

$$-(1+\beta^W \, \varphi_k^W)\xi_0^W \, \boldsymbol{e}_1^T \, \boldsymbol{\tau}_t + \Bigg[\Bigg(f_2 \, \beta^W (\varphi_{\tau 1}^{AW} \, \sigma_{\text{forc}} \, z_A + \varphi_{\tau 1}^{BW} \, \sigma_{\text{forc}} \, z_B) - (1+\beta^W \, \varphi_k^W) \Big(d_{AW} \, \frac{\alpha_A}{1 - \alpha_A \, \alpha_B}(z_A - \alpha_B \, z_B)$$

$$+ \, d_{BW} \, \frac{1}{1 - \alpha_A \, \alpha_B}(z_B - \alpha_A \, z_A) \Big) - f_3 \, \beta^W (\sigma_{\text{forc}} \, \varphi_{\tau 1}^{AW} \, z_A^{1-n} + \sigma_{\text{forc}} \, \varphi_{\tau 1}^{BW} \, z_B^{1-n}) - a^W (1 + \beta^W \, \varphi_k^W)$$

$$+ f_1 \, \beta^W (\sigma_{\text{forc}} \, \varphi_{\tau 1}^{AW} + \sigma_{\text{forc}} \, \varphi_{\tau 1}^{BW}) \Bigg) M_{pre}^{-1} + \beta^W \, (\varphi_{M1}^W \, \phi_{11} + \varphi_{M2}^W \, \phi_{12}) \Bigg] M_{1,t} + \Big[\beta^W (\varphi_{M1}^W \, \phi_{21} + \varphi_{M2}^W \, \phi_{22}) \Big] M_{2,t}$$

$$+ \Big[\beta^W \, \boldsymbol{\varphi}_{R,t+1}^T \Big] R_t + \log x_t^* + \beta^W \, \varphi_k^W \, \log(1 - x_t^*) + (1 + \beta^W \, \varphi_k^W) \log F(\boldsymbol{A}_t, \, \boldsymbol{\mathcal{K}}_t^*, \, \boldsymbol{N}_t^*, \, \boldsymbol{E}_t^*) + (1 + \beta^W \, \varphi_k^W)(\xi_0^W + a^W)$$

$$+ f_0 \, \beta^W (\varphi_{\tau 1}^{AW} \, \sigma_{\text{forc}} + \varphi_{\tau 1}^{BW} \, \sigma_{\text{forc}}) + \beta^W \, \varphi_{M1}^W (\sum_{i=1}^{I^d} E_{i,t}^* + E_t^{\text{exo}}) - \beta^W \, \boldsymbol{\varphi}_{R,t+1}^T \, \boldsymbol{E}_t^{d*} + \beta^W \, \varphi_{t+1}. \quad (124)$$

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 counter-geoengineering or geoengineering.

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

$$\varphi_k^W \, k_t + \boldsymbol{\varphi}_{\tau W}^T \, \boldsymbol{\tau}_t + \varphi_{M1}^W \, M_{1,t} + \varphi_{M2}^W \, M_{2,t} + \boldsymbol{\varphi}_{R,t}^T \, \boldsymbol{R}_t + \varphi_t = \Big[(1 + \beta^W \, \varphi_k^W)\kappa \Big] k_t + \beta^W \, \boldsymbol{\varphi}_{\tau W}^T \, \boldsymbol{\sigma \tau}_t$$

$$- (1 + \beta^W \, \varphi_k^W)\xi_0^W \, \boldsymbol{e}_1^T \, \boldsymbol{\tau}_t + \Bigg[\Bigg(f_2 \, \beta^W (\sigma_{\text{forc}} \, \varphi_{\tau 1}^{AW} \, z_A^{go} + \sigma_{\text{forc}} \, \varphi_{\tau 1}^{BW} \, \alpha_A \, z_A^{go}) - (1 + \beta^W \, \varphi_k^W) \, d_{AW} \, \alpha_A \, z_A^{go}$$

$$- f_3 \, \beta^W (\sigma_{\text{forc}} \, \varphi_{\tau 1}^{AW} \, (z_A^{go})^{1-n} + \sigma_{\text{forc}}^B \, \varphi_{\tau 1}^{BW} \, (\alpha_A \, z_A^{go})^{1-n}) - a^W (1 + \beta^W \, \varphi_k^W) + f_1 \, \beta^W (\sigma_{\text{forc}}^A \, \varphi_{\tau 1}^{AW} + \sigma_{\text{forc}}^B \, \varphi_{\tau 1}^{BW}) \Bigg) M_{pre}^{-1}$$

$$+ \beta^W \, (\varphi_{M1}^W \, \phi_{11} + \varphi_{M2}^W \, \phi_{12}) \Bigg] M_{1,t} + \Big[\beta^W (\varphi_{M1}^W \, \phi_{21} + \varphi_{M2}^W \, \phi_{22}) \Big] M_{2,t} + \Big[\beta^W \, \boldsymbol{\varphi}_{R,t+1}^T \Big] R_t + \log x_t^* + \beta^W \, \varphi_k^W \, \log(1 - x_t^*)$$

$$+ (1 + \beta^W \, \varphi_k^W) \log F(\boldsymbol{A}_t, \, \boldsymbol{\mathcal{K}}_t^*, \, \boldsymbol{N}_t^*, \, \boldsymbol{E}_t^*) + (1 + \beta^W \, \varphi_k^W)(\xi_0^W + a^W) + f_0 \, \beta^W (\varphi_{\tau 1}^{AW} \, \sigma_{\text{forc}} + \varphi_{\tau 1}^{BW} \, \sigma_{\text{forc}})$$

$$+ \beta^W \, \varphi_{M1}^W (\sum_{i=1}^{I^d} E_{i,t}^* + E_t^{\text{exo}}) - \beta^W \, \boldsymbol{\varphi}_{R,t+1}^T \, \boldsymbol{E}_t^{d*} + \beta^W \, \varphi_{t+1}. \quad (125)$$

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

$$\varphi_k^W \, k_t + \boldsymbol{\varphi}_{\tau W}^T \, \boldsymbol{\tau}_t + \varphi_{M1}^W \, M_{1,t} + \varphi_{M2}^W \, M_{2,t} + \boldsymbol{\varphi}_{R,t}^T \, \boldsymbol{R}_t + \varphi_t = \Big[(1 + \beta^W \, \varphi_k^W)\kappa\Big] k_t + \beta^W \, \boldsymbol{\varphi}_{\tau W}^T \, \boldsymbol{\sigma} \boldsymbol{\tau}_t$$

$$- (1 + \beta^W \, \varphi_k^W)\xi_0^W \, \boldsymbol{e}_1^T \, \boldsymbol{\tau}_t + \Bigg[\bigg(f_2 \, \beta^W (\sigma_{\text{forc}} \, \varphi_{\tau 1}^{AW} \, z_B^{go} + \sigma_{\text{forc}} \, \varphi_{\tau 1}^{BW} \, \alpha_A \, z_B^{go}) - (1 + \beta^W \, \varphi_k^W) \, d_{BW} \, z_B^{go}$$

$$- f_3 \, \beta^W (\sigma_{\text{forc}} \, \varphi_{\tau 1}^{AW} \, (\alpha_B \, z_B^{go})^{1-n} + \sigma_{\text{forc}} \, \varphi_{\tau 1}^{BW} \, (z_B^{go})^{1-n}) - a^W (1 + \beta^W \, \varphi_k^W) + f_1 \, \beta^W (\sigma_{\text{forc}} \, \varphi_{\tau 1}^{AW} + \sigma_{\text{forc}} \, \varphi_{\tau 1}^{BW}) \bigg) M_{pre}^{-1}$$

$$+ \beta^W (\varphi_{M1}^W \, \phi_{11} + \varphi_{M2}^W \, \phi_{12}) \Bigg] M_{1,t} + \Big[\beta^W (\varphi_{M1}^W \, \phi_{21} + \varphi_{M2}^W \, \phi_{22}) \Big] M_{2,t} + \Big[\beta^W \, \boldsymbol{\varphi}_{R,t+1}^T \Big] R_t + \log x_t^* + \beta^W \, \varphi_k^W \, \log(1 - x_t^*)$$

$$+ (1 + \beta^W \, \varphi_k^W) \log F(\boldsymbol{A}_t, \boldsymbol{\mathcal{K}}_t^*, \boldsymbol{N}_t^*, \boldsymbol{E}_t^*) + (1 + \beta^W \, \varphi_k^W)(\xi_0^W + a^W) + f_0 \, \beta^W (\varphi_{\tau 1}^{AW} \, \sigma_{\text{forc}} + \varphi_{\tau 1}^{BW} \, \sigma_{\text{forc}})$$

$$+ \beta^W \, \varphi_{M1}^W (\sum_{i=1}^{I^d} E_{i,t}^* + E_t^{\text{exo}}) - \beta^W \, \boldsymbol{\varphi}_{R,t+1}^T \, \boldsymbol{E}_t^{d*} + \beta^W \, \varphi_{t+1}. \quad (126)$$

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

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

Inserting φ_k^W into equation (113) yields the optimal consumption rate $x_t^* = 1 - \beta^W \, \kappa$.

Coefficient matching with respect to transformed temperatures delivers

$$\boldsymbol{\varphi}_{\tau W}^T = -\xi_0^W \, (1 + \beta^W \, \varphi_k^W) \, \boldsymbol{e}_1^T \, [\boldsymbol{1} - \beta^W \, \boldsymbol{\sigma}]^{-1}. \quad (127)$$

with

$$[\boldsymbol{1} - \beta^W \, \boldsymbol{\sigma}]^{-1} = \begin{pmatrix} \tilde{\sigma}_{11}^W & \tilde{\sigma}_{12}^W & \tilde{\sigma}_{13}^W \\ \tilde{\sigma}_{21}^W & \tilde{\sigma}_{22}^W & \tilde{\sigma}_{23}^W \\ \tilde{\sigma}_{31}^W & \tilde{\sigma}_{32}^W & \tilde{\sigma}_{33}^W \end{pmatrix}$$

and hence

$$\varphi_{\tau 1}^{AW} = -\xi_0^W \, (1 + \beta^W \, \varphi_k^W) \tilde{\sigma}_{21}^W \quad (128)$$

$$\varphi_{\tau 1}^{BW} = -\xi_0^W \, (1 + \beta^W \, \varphi_k^W) \tilde{\sigma}_{31}^W \quad (129)$$

We define $\gamma_{AW} \equiv \beta^W \, \xi_0^W \, \tilde{\sigma}_{21}^W \, \sigma_{\text{forc}}$ and $\gamma_{BW} \equiv \beta^W \, \xi_0^W \, \tilde{\sigma}_{31}^W \, \sigma_{\text{forc}}$.

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

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

$$\varphi_{M1}^W = (1+\beta^W\,\varphi_k^W)\left(-f_2(\gamma_{AW}\,z_A+\gamma_{BW}\,z_B)-d_{AW}\,\frac{\alpha_A}{1-\alpha_A\,\alpha_B}(z_A-\alpha_B\,z_B)-d_{BW}\,\frac{1}{1-\alpha_A\,\alpha_B}(z_B-\alpha_A\,z_A)\right.$$

$$\left.+\,f_3\left(\gamma_{AW}\,z_A^{1-n}+\gamma_{BW}\,z_B^{1-n}\right)-a^W-f_1(\gamma_{AW}+\gamma_{BW})\right)M_{pre}^{-1}+\beta^W\,(\varphi_{M1}^W\,\phi_{11}+\varphi_{M2}^W\,\phi_{12}) \quad (130)$$

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

$$\varphi_{M1}^W = (1+\beta^W\,\varphi_k^W)\left(-f_2(\gamma_{AW}\,z_A^{go}+\gamma_{BW}\,\alpha_A\,z_A^{go})-d_{AW}\,\alpha_A\,z_A^{go}+f_3\left(\gamma_{AW}\,(z_A^{go})^{1-n}+\gamma_{BW}\,(\alpha_A\,z_A^{go})^{1-n}\right)\right.$$

$$\left.-\,a^W-f_1(\gamma_{AW}+\gamma_{BW})\right)M_{pre}^{-1}+\beta^W\,(\varphi_{M1}^W\,\phi_{11}+\varphi_{M2}^W\,\phi_{12}) \quad (131)$$

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

$$\varphi_{M1}^W = (1+\beta^W\,\varphi_k^W)\left(-f_2(\gamma_{AW}\,\alpha_B z_B^{go}+\gamma_{BW}\,z_B^{go})-d_{BW}\,z_B^{go}+f_3\left(\gamma_{AW}\,(\alpha_B\,z_B^{go})^{1-n}+\gamma_{BW}\,(z_B^{go})^{1-n}\right)\right.$$

$$\left.-\,a^W-f_1(\gamma_{AW}+\gamma_{BW})\right)M_{pre}^{-1}+\beta^W\,(\varphi_{M1}^W\,\phi_{11}+\varphi_{M2}^W\,\phi_{12}) \quad (132)$$

From coefficient matching with respect to the carbon stock in the ocean we get

$$\varphi_{M2}^W = \beta^W(\varphi_{M1}^W\,\phi_{21}+\varphi_{M2}^W\,\phi_{22}) \quad \Leftrightarrow \quad \varphi_{M2}^W = \frac{\beta^W\,\varphi_{M1}^W\,\phi_{21}}{1-\beta^W\,\phi_{22}},$$

and for 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 (94),(95), and (96) turn to the following condition:

$$\varphi_t - \beta^W\,\varphi_{t+1} = \log x_t^* + \beta^W\,\varphi_k^W\,\log(1-x_t^*) + (1+\beta^W\,\varphi_k^W)\log F(\boldsymbol{A}_t,\,\boldsymbol{\mathcal{K}}_t^*,\,\boldsymbol{N}_t^*,\,\boldsymbol{E}_t^*)$$

$$+(1+\beta^W\,\varphi_k^W)(\xi_0^W+a^W)+\beta^W\,(\varphi_{\tau 1}^{AW}\,\sigma_{\text{forc}}+\varphi_{\tau 1}^{BW}\,\sigma_{\text{forc}})\,f_0+\beta^W\,\varphi_{M1}^W(\sum_{i=1}^{I^d}E_{i,t}^*+E_t^{\text{exo}})-\beta^W\,\boldsymbol{\varphi}_{R,t+1}^T\,\boldsymbol{E}_t^{d*}$$

$$(133)$$

This condition will be satisfied by picking the sequence $\varphi_0, \varphi_1, \varphi_2, \dots$. 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.

C.2.1 Social cost of carbon in the rest of the world

Inserting φ_{M2}^W, and φ_k^W into (130), (131), and (132) delivers

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

$$
\begin{aligned}
\varphi_{M1}^W = \frac{1}{1-\beta^W \kappa} \Bigg(&-f_2(\gamma_{AW} z_A + \gamma_{BW} z_B) - d_{AW} \frac{\alpha_A}{1-\alpha_A \alpha_B}(z_A - \alpha_B z_B) - d_{BW} \frac{1}{1-\alpha_A \alpha_B}(z_B - \alpha_A z_A) \\
&+ f_3\left(\gamma_{AW} z_A^{1-n} + \gamma_{BW} z_B^{1-n}\right) - a^W - f_1(\gamma_{AW} + \gamma_{BW}) \Bigg) M_{pre}^{-1} \underbrace{\left(1 - \beta^W \phi_{11} - \frac{(\beta^W)^2 \phi_{12} \phi_{21}}{1-\beta^W \phi_{22}}\right)^{-1}}_{\equiv \tilde\phi_W}
\end{aligned}
$$

$$(134)$$

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

$$
\begin{aligned}
\varphi_{M1}^W = \frac{1}{1-\beta^W \kappa} \Bigg(&-f_2(\gamma_{AW} z_A^{go} + \gamma_{BW} \alpha_A z_A^{go}) - d_{AW} \alpha_A z_A^{go} + f_3\left(\gamma_{AW} (z_A^{go})^{1-n} + \gamma_{BW} (\alpha_A z_A^{go})^{1-n}\right) \\
&- a^W - f_1(\gamma_{AW} + \gamma_{BW}) \Bigg) M_{pre}^{-1} \underbrace{\left(1 - \beta^W \phi_{11} - \frac{(\beta^W)^2 \phi_{12} \phi_{21}}{1-\beta^W \phi_{22}}\right)^{-1}}_{\equiv \tilde\phi_W}
\end{aligned}
$$

$$(135)$$

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

$$
\begin{aligned}
\varphi_{M1}^W = \frac{1}{1-\beta^W \kappa} \Bigg(&-f_2(\gamma_{AW} \alpha_B z_B^{go} + \gamma_{BW} z_B^{go}) - d_{BW} z_B^{go} + f_3\left(\gamma_{AW} (\alpha_B z_B^{go})^{1-n} + \gamma_{BW} (z_B^{go})^{1-n}\right) \\
&- a^W - f_1(\gamma_{AW} + \gamma_{BW}) \Bigg) M_{pre}^{-1} \underbrace{\left(1 - \beta^W \phi_{11} - \frac{(\beta^W)^2 \phi_{12} \phi_{21}}{1-\beta^W \phi_{22}}\right)^{-1}}_{\equiv \tilde\phi_W}
\end{aligned}
$$

$$(136)$$

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 $S_t^B \neq 0$

$$SCC^W = -(1 - \beta^W \kappa)Y_{W,t}^{net}\, \varphi_{M1}^W$$

$$= \frac{Y_{W,t}^{net}}{M_{pre}} \left(f_2(\gamma_{AW}\, z_A + \gamma_{BW}\, z_B) + d_{AW}\, \frac{\alpha_A}{1 - \alpha_A \alpha_B}(z_A - \alpha_B\, z_B) + d_{BW}\, \frac{1}{1 - \alpha_A \alpha_B}(z_B - \alpha_A\, z_A) \right.$$

$$\left. - f_3\left(\gamma_{AW}\, z_A^{1-n} + \gamma_{BW}\, z_B^{1-n} \right) + a^W + f_1(\gamma_{AW} + \gamma_{BW}) \right) \tilde{\phi}_W,$$

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

$$SCC^W = -(1 - \beta^W \kappa)Y_{W,t}^{net}\, \varphi_{M1}^W$$

$$= \frac{Y_{W,t}^{net}}{M_{pre}} \left(f_2(\gamma_{AW}\, z_A^{go} + \gamma_{BW}\, \alpha_A\, z_A^{go}) + d_{AW}\, \alpha_A\, z_A^{go} - f_3\left(\gamma_{AW}\, (z_A^{go})^{1-n} + \gamma_{BW}\, (\alpha_A\, z_A^{go})^{1-n} \right) \right.$$

$$\left. + a^W + f_1(\gamma_{AW} + \gamma_{BW}) \right) \tilde{\phi}_W,$$

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

$$SCC^W = -(1 - \beta^W \kappa)Y_{W,t}^{net}\, \varphi_{M1}^W$$

$$= \frac{Y_{W,t}^{net}}{M_{pre}} \left(f_2(\gamma_{AW}\, \alpha_B z_B^{go} + \gamma_{BW}\, z_B^{go}) + d_{BW}\, z_B^{go} - f_3\left(\gamma_{AW}\, (\alpha_B\, z_B^{go})^{1-n} + \gamma_{BW}\, (z_B^{go})^{1-n} \right) \right.$$

$$\left. + a^W + f_1(\gamma_{AW} + \gamma_{BW}) \right) \tilde{\phi}_W.$$

- -

Simplification: We set σ_B^A, σ_A^B, σ_A^O, σ_B^O equal to zero. After defining $\gamma_W = \gamma_{BW}$, the social cost of carbon in the rest of the world becomes

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

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

$$(137)$$

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 $S_t^B = 0$

$$SCC^W = \frac{Y_{W,t}^{net}}{M_{pre}} \left[a^W + f_1 \, \gamma_W + \alpha_A \, z_A^g \, d_{AW} - \gamma_W \left(f_3 \, (\alpha_A \, z_A^g)^{1-n} - \alpha_A \, z_A^g \, f_2 \right) \right] \tilde{\phi}_W, \qquad (138)$$

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

$$SCC^W = \frac{Y_{W,t}^{net}}{M_{pre}} \left[a^W + f_1 \, \gamma_W + z_B^g \, d_{BW} - \gamma_W \left(f_3 \, (z_B^g)^{1-n} - z_B^g \, f_2 \right) \right] \tilde{\phi}_W. \qquad (139)$$