Online Appendix to

The Social Cost of Carbon in a Non-cooperative World

Christoph Hambel[a], Holger Kraft[b,*], Eduardo Schwartz[c]

[a,b] *Goethe University Frankfurt, Faculty of Economics and Business Administration, Theodor-W.-Adorno-Platz 3, 60323 Frankfurt am Main, Germany.*

[c] *UCLA Anderson School of Management, Los Angeles.*

This version: May 4, 2021

Abstract

This online appendix contains additional material such as details of the calibrations and two model extensions.

* Corresponding author. Phone: +49 69 798 33699.

E-mail addresses: christoph.hambel@finance.uni-frankfurt.de (Christoph Hambel), holgerkraft@finance.uni-frankfurt.de (Holger Kraft), eduardo.schwartz@anderson.ucla.edu (Eduardo Schwartz)

	Carbon Cycle	
φ_{12}	rate of carbon diffusion from atmosphere to upper ocean	0.0174
φ_{21}	rate of carbon diffusion from upper ocean to atmosphere	7.6×10^{-4}
φ_{23}	rate of carbon diffusion from upper ocean to lower ocean	0.0005
φ_{32}	rate of carbon diffusion from lower ocean to upper ocean	6.48×10^{-5}
$M^{at,\mathrm{PI}}$	pre-industrial atmospheric CO_2 concentration	588
$M^{uo,\mathrm{PI}}$	pre-industrial upper oceanic CO_2 concentration	1,350
$M^{lo,\mathrm{PI}}$	pre-industrial lower oceanic CO_2 concentration	10,000
	Temperature Dynamics	
ϕ	rate of temperature decrease due to infrared radiation to space	0.0283
ϕ_{12}	rate of heat diffusion from atmosphere to ocean	0.0050
ϕ_{21}	rate of heat diffusion from ocean to atmosphere	0.0024
κ_τ	temperature transition parameter	0.0238
η_τ	radiative forcing parameter	3.8
η_0	linearization parameter	-0.4925
η_1	linearization parameter	2.0544

Table 7: Climate System Calibration. This table summarizes the parameters of the climate system.

D Calibration

This section presents the calibration of the different versions of the model. Additionally, we describe how DICE-2013R can be formulated in a continuous-time setting. For the global model, we drop the subscripts n throughout the section as there is only one country.

D.1 Preferences

Nordhaus and Sztorc (2013) use a CRRA utility function

$$u(\mathscr{C}, L) = \frac{\mathscr{C}^{1-\eta} L^{\eta} - 1}{1-\eta}$$

where the inverse of the elasticity of intertemporal substitution (EIS) is $\eta = 1.45$ and the pure rate of time preferences is fixed at $\delta = 1.5\%$. Our closed-form solution, however, relies on logarithmic preferences, where η is equal to one. Following Nordhaus (2008) and Golosov et al. (2014),[25] we thus increase the pure rate of time preference and fix the time preference rate at $\delta = 2.2\%$. This adjustment leads to evolutions of the relevant key variables in DICE that are very close to the optimal scenario resulting from $\eta = 1.45$ and $\delta = 1.5\%$.

[25] Both authors use a similar adjustment in DICE-2007.

D.2 Climate Model

DICE-2013R evolves in 5-year time steps, whereas our model is formulated in continuous time. To obtain a continuous-time version of the DICE-2013R climate system, we adopt the calibration strategy used by Cai and Lontzek (2018). We discretize our model using one-year time steps and choose the relevant carbon cycle parameters such that the model closely matches the baseline evolution in DICE-2013R. Notice that using finer discretizations does not change the calibration results significantly. We denote the pre-industrial equilibrium concentrations in the three layers by $M^{at,\mathrm{PI}}$, $M^{uo,\mathrm{PI}}$, and $M^{lo,\mathrm{PI}}$, respectively. In equilibrium, the climate system is in a steady state and thus $\varphi_{21}M^{uo,\mathrm{PI}} = \varphi_{12}M^{at,\mathrm{PI}}$ and $\varphi_{32}M^{lo,\mathrm{PI}} = \varphi_{23}M^{uo,\mathrm{PI}}$, see Nordhaus (2008). Therefore, the carbon cycle is fully determined by φ_{12} and φ_{23}. We denote the CO_2 concentrations in the three layers in the DICE-2013R model by $\big(\widehat{M}^{at}(t_j), \widehat{M}^{uo}(t_j), \widehat{M}^{lo}(t_j)\big)$ where $t_j \in \{1, \ldots, 60\}$ denote the time nodes of the DICE model.

We take the rate of emissions in the DICE-2013R baseline scenario and use a cubic-spline interpolation to generate data for the baseline emissions path, denoted by $E(t)$ for each year t. Then, we minimize the relative quadratic deviations of the model-implied concentrations determined by (8) from the DICE-2013R concentrations, i.e., we minimize

$$\min_{\varphi_{12},\varphi_{23}} \sum_{j=1}^{60} \left[\Big(\frac{\widehat{M}^{at}(t_j) - M^{at}(t_j)}{M^{at}(t_j)}\Big)^2 + \Big(\frac{\widehat{M}^{uo}(t_j) - M^{uo}(t_j)}{M^{uo}(t_j)}\Big)^2 + \Big(\frac{\widehat{M}^{lo}(t_j) - M^{lo}(t_j)}{M^{lo}(t_j)}\Big)^2 \right].$$

We obtain $\varphi_{12} = 0.0174$, $\varphi_{23} = 5.0 \times 10^{-4}$ and thus $\varphi_{21} = \varphi_{12}\frac{M^{at,\mathrm{PI}}}{M^{uo,\mathrm{PI}}} = 0.0076$ and $\varphi_{32} = \varphi_{23}\frac{M^{uo,\mathrm{PI}}}{M^{lo,\mathrm{PI}}} = 6.48 \times 10^{-5}$.

Similarly, we calibrate the parameters determining the climate system. As in DICE-2013R, we fix an equilibrium climate sensitivity of ECS $= 3.2$°C and set the radiative forcing parameter to $\eta_\tau = 3.8$. Then, ϕ is determined by $\phi = \frac{\kappa_\tau \eta_\tau}{\mathrm{ECS}}$. Using a similar notation as before, we determine the remaining climate system parameters by minimizing the relative quadratic deviations of the model-implied temperatures determined by (8) from the DICE-2013R temperatures, i.e.,

$$\min_{\phi_{12},\phi_{21},\kappa_\tau} \sum_{j=1}^{60} \left[\Big(\frac{\widehat{T}^{at}(t_j) - T^{at}(t_j)}{T^{at}(t_j)}\Big)^2 + \Big(\frac{\widehat{T}^{o}(t_j) - T^{o}(t_j)}{T^{o}(t_j)}\Big)^2 \right].$$

The minimization yields $\phi_{12} = 0.0050$, $\phi_{21} = 0.0024$ and $\kappa_\tau = 0.0238$. Consequently, we set $\phi = \frac{\kappa_\tau \eta_\tau}{\mathrm{ECS}} = 0.0283$.

D.3 Global Economic Model

Continuous-time DICE-2013R For the continuous-time version of DICE-2013R, we use the deterministic functions $A(t)$ (total factor productivity), $L(t)$ (world population), $\sigma(t)$ (emission intensity), $E^{\text{land}}(t)$ (CO_2 emissions from deforestation and land-use), $F^{\text{ex}}(t)$ (exogenous radiative forcing), and $a(t)$ (abatement cost coefficient) as in Nordhaus and Sztorc (2013). The DICE model assumes that climate change has a level impact captured by the damage function

$$D(T) = \frac{1}{1 + \theta T^2}, \tag{63}$$

which scales down output in response to higher temperatures. Following Nordhaus and Sztorc (2013), we fix the damage parameter at $\theta = 0.00266375$, the depreciation rate at $\delta_k = 0.1$ and the capital share at $\alpha = 0.3$.

AK-Model We calibrate the expected gross growth rate g in (17) such that our analytical AK-model closely matches the evolution of GDP growth in the DICE baseline scenario. We first simulate the output in DICE disregarding abatement ($\mu = 0$) and the impact of climate damage ($D = 0$). This yields data points $\big(t_j, \widehat{Y}(t_j)\big)$, which are used to extract the corresponding future GDP gross growth rates of DICE. Notice that these growth rates depend on the optimal consumption rates. It turns out that these growth rates (before abatement and damages) can be fitted well using the following functional form:

$$g^{\text{DICE}}(t) = g_0 + g_1 \mathrm{e}^{-\delta_g t} \tag{64}$$

where $g_0 = 0.0044$, $g_1 = 0.029$, and $\delta_g = 0.012$, i.e., the growth rate declines at a rate of 1.1% to its long-term steady-state level of 0.44%.

In a second step, we follow Pindyck and Wang (2013) and fix the productivity parameter at $A = 0.113$. To calibrate δ_k and θ, we equate the gross growth rate $g(t, \chi^*, 0)$ in the absence of climate change, which is given in (18), and the DICE gross growth rate (64). To separately identify δ_k and θ, we must make an assumption about the consumption rate χ^*, which is an endogenous variable. The optimal χ^* in DICE is is pretty stable over time and close to 76%.[26]

[26] This is also in line with the literature: Pindyck and Wang (2013) and Golosov et al. (2014) use a value of 75%.

Notice that for $\mu = 0$ the optimal consumption rate is a constant given by

$$\chi_n^* = \frac{\theta A - 1 + \sqrt{(\theta A - 1)^2 + 4\theta\delta}}{2\theta A}.$$

Therefore, assuming that the optimal consumption rate of the society is $\chi^* = 0.76$ yields $\theta = 27.43$.

In a third step, we calibrate the damage parameter such that the damages in the AK-model are close to the damages in DICE. For this purpose, we simulate output net of damages in DICE. This yields data points $\big(t_j, \widehat{Y}^{net}(t_j)\big)$, $j = 1, \ldots, \mathcal{J}$. Then, we minimize the quadratic deviations of the model-implied output determined by (17) from the net output in DICE, i.e.,

$$\min_{\xi} \sum_{j=1}^{\mathcal{J}} \big(\widehat{Y}^{net}(t_j) - Y(t_j)\big)^2.$$

This minimization is sensitive to the time horizon $\mathcal{J}$ and we focus on the first 100 years. The minimization yields $\xi = 6.8644 \times 10^{-4}$.

D.4 Regional Economic Model

The calibration of the regional model follows closely Section D.3. We consider a model with five heterogeneous regions as in the representative concentration pathways (RCPs) provided by the AR5 Scenario Database of IPCC (2014). Table 2 summarizes the definitions of these regions and the calibration of other relevant parameters.

To calibrate the economic model, we use the predictions from RICE-2010 and aggregate national data from RICE into the five regions. This yields data points $\big(t_j, \widehat{Y}_n(t_j)\big)$, which are used to extract the corresponding future GDP gross growth rates and to calibrate the growth rate parameters A_n, δ_n^K, and θ_n such that the model closely matches the evolution of GDP and aggregate consumption in RICE as described in Section D.3. To obtain the regional damage parameters ξ_n, we use the relative estimates provided by Stanton et al. (2012) and scale them such that they are consistent with the global damage parameter estimated in Appendix D.3.

The results are summarized in Table 2. In our calibration, the damage parameter of the MAF region is three times higher than the one of the OECD90region, which is consistent with

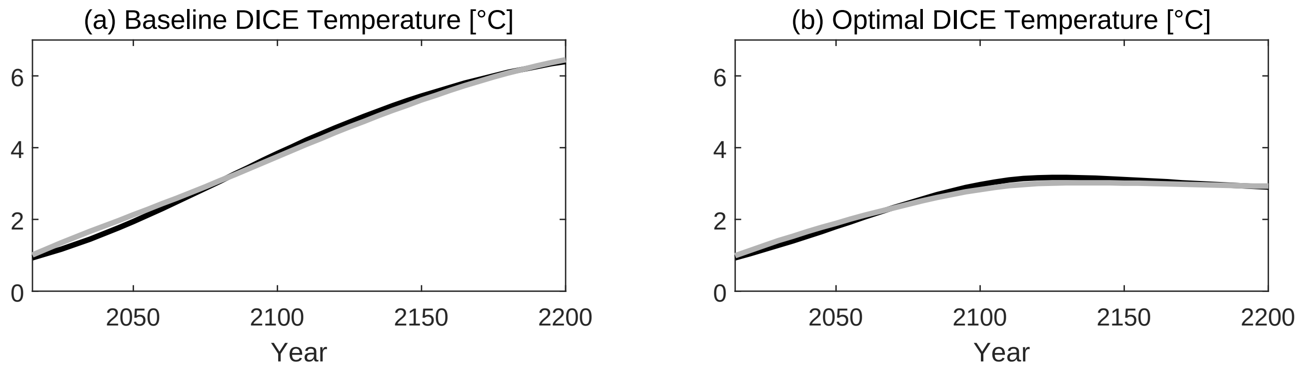

Figure 6: Temperature Evolution. The figure depicts the temperature evolution in the continuous-time version of DICE-2013R. For the baseline scenario, Panel (a) shows the simulated global average temperature using the correct temperature dynamics (black line) and the simulated temperature using the linear approximation (19). Analogously, Panel (b) depicts the correct and approximated temperature evolutions for the optimal scenario. The approximation uses $\eta_0 = -0.4925$ and $\eta_1 = 2.0544$.

other studies such as Dell et al. (2009, 2012) or Burke et al. (2015). Finally, we use RCP 8.5 predictions of regional CO_2 emissions to determine the regional emission intensities.

E Linearization of Radiative Forcing

To obtain a closed-form solution to the model, we linearize radiative forcing in the temperature dynamics (10), i.e., we apply the first-order approximation

$$\eta_\tau \log\left(\frac{M_t^{at}}{M^{\mathrm{PI}}}\right) \approx \eta_0 + \eta_1 \frac{M_t^{at}}{M^{\mathrm{PI}}}. \tag{65}$$

We calibrate the coefficients η_0 and η_1 such that the climate model closely matches the temperature evolution of the DICE-2013R model. This can be achieved by jointly minimizing squared errors between the correct temperature evolution, T^j, and the approximated temperature evolution, T^{j,η_0,η_1}, in the optimal and the baseline scenario.

$$\min_{\eta_0,\eta_1} \sum_{j\in\{\mathrm{opt},\mathrm{base}\}} \sum_{i=1}^{I} [T_i^j - T_i^{j,\eta_0,\eta_1}]^2. \tag{66}$$

The minimization yields $\eta_0 = -0.4925$ and $\eta_1 = 2.0544$. Figure 6 depicts the correct and the approximated temperature evolution until the year 2200 in both scenarios and shows that the approximation performs well. We find that the maximal deviation $\max_{i,j} |T_i^j - T_i^{j,\eta_0,\eta_1}|$ of the correct temperature evolution from its approximation is below 0.15°C in both scenarios indicating a high accuracy of the approximation (65).

F Optimal Carbon Tax and the SCC

If no climate policy is implemented at all, a country's industrial BAU-emissions are given by

$$E_{nt}^{\text{BAU}} = \sigma_n(t) Y_{nt}.$$

The SCC is the marginal damage of one additional ton of carbon. The output causing one ton of carbon emissions $\Delta E_{nt} = 1\,\text{tC}$ is thus given by

$$\Delta Y_{nt} = \frac{1}{\sigma_n(t)}.$$

The marginal costs of avoiding this ton of carbon are equal to

$$\Delta Y_{nt} a_n(t) b_n \mu_{nt}^{b_n - 1} = \mu_{nt}^{b_n - 1} \frac{a_n(t) b_n}{\sigma_n(t)}$$

In the optimum and if climate policies are unlimited available, marginal damages are equal to marginal costs, i.e.,

$$\text{SCC}_{\text{nt}} = \tau_{nt}^* = (\mu_{nt}^*)^{b_n - 1} \frac{a_n(t) b_n}{\sigma_n(t)}.$$

G Extension: Capital Transfers

This appendix briefly studies the implications of capital transfers on our results. In particular, we are interested in the answers to the following questions: Would a country be willing to sponsor abatement in another country if this can be achieved via transfers? And if the answer is positive, what is the effect on the SCC? A way to address these points in our framework is that we allow country n to donate some of their imports from country k so that country k can implement additional abatement.[27] There are two possible scenarios.

1st scenario without a commitment device. In this case, all countries are allowed to optimize consumption, abatement, and transfers simultaneously, i.e., we add transfers as an additional decision variable in (21). One can show that in such a setting optimal abatement stays the same in all countries.[28] Only the financing of the abatement policies changes since some of it might be financed by transfers. This is presumably not a satisfying outcome from the perspective of a giving country.

2nd scenario with a commitment device. Alternatively, one can analyze a setting with a commitment device that binds a receiving country to maintain its optimal abatement expenditures before transfers and to use the transfers to implement additional abatement. To determine the equilibrium in such a scenario, we assume that decisions are made in *three steps*: First, all countries optimize over consumption and abatement without transfers. This leads to the solution of the non-cooperative game presented in Section 4. Second, all countries determine whether it is optimal for them to make transfers to other countries. To do so, we assume that all countries optimize over transfers and again over consumption, but keep abatement from the first step fixed. Third, those countries that do not receive any transfers are allowed to reoptimize both abatement and consumption, whereas receiving countries are bound by the commitment device.

As explained above, we model a transfer from country n to country k in such a way that country n leaves some of its imports of good k in country k, i.e., $\mathcal{C}_k^n$ is reduced by the size of the transfer $\mathcal{T}_k^n$. In the presence of capital transfers, the consumption bundles (16) and the abatement expenditures can thus be rewritten as

$$\widehat{\mathscr{C}}_n = \prod_{k=1}^{N} (\mathcal{C}_k^n - \mathcal{T}_k^n)^{\beta_k^n}, \qquad \widehat{\mathcal{A}}_n = \mathcal{A}_n + \sum_{k=1}^{N} \mathcal{T}_n^k. \tag{67}$$

Let $\mathcal{A}_n^*$ denote the optimal abatement expenditures of the non-cooperative game that are determined in Section 4. In the first scenario, $\mathcal{A}_n$ can be different from the optimal abatement $\mathcal{A}_n^*$ without transfers. In the second scenario, $\mathcal{A}_n$ is identical to $\mathcal{A}_n^*$ for receiving countries due to the commitment device, but it can differ for countries that do not receive any transfers.

[27] This is in line with our earlier assumption that the abatement of a country is financed through its output.

[28] The proof is available upon request.

In any case, we impose the constrained that transfers must be positive, i.e., $\mathcal{T}_n^k \geq 0$ for all $n, k \in \{1, \ldots, N\}$. The following proposition summarizes the effects of capital transfers in the presence of a commitment device.

Proposition G.1 (Capital Transfers and Commitment Device)**.** *The country-specific social cost of carbon is given by*

$$\mathrm{SCC}_{nt} = \frac{Y_{nt}\chi_{nt}^*}{M^{\mathrm{PI}}\delta_n} \sum_{\ell=1}^{N} \beta_\ell^n \xi_\ell \frac{\kappa_\tau \eta_1}{\delta_n + \phi + \frac{\phi_{21}\delta_n}{\delta_n + \phi_{12}}} (\delta_n \mathbf{1}_{3\times 3} - \mathbf{\Phi}^\top)_{1,1}^{-1},$$

i.e., initially the social cost of carbon is the same for the non-cooperative game with and without capital transfers. Let $\varepsilon_\ell^n = \mathcal{T}_\ell^n / \mathcal{C}_\ell^n$. The optimal consumption and transfer decisions of country n satisfy the following non-linear system

$$\begin{aligned}
&p_{M^{at}}^n Y_\ell \sigma_\ell \Big(\frac{a_\ell(t)(\mu_\ell^*)^{b_\ell} + \sum_{k\neq\ell} \chi_\ell^k \varepsilon_\ell^k}{a_\ell(t)}\Big)^{\frac{1-b_\ell}{b_\ell}} \frac{\chi_\ell^n}{b_\ell a_\ell(t)} - \frac{\delta_n \beta_\ell^n}{1-\varepsilon_\ell^n} &&= \lambda_{\ell t}^n, \\
&p_{M^{at}}^n Y_\ell \sigma_\ell \Big(\frac{a_\ell(t)(\mu_\ell^*)^{b_\ell} + \sum_{k\neq\ell} \chi_\ell^k \varepsilon_\ell^k}{a_\ell(t)}\Big)^{\frac{1-b_\ell}{b_\ell}} \frac{\varepsilon_\ell^n}{b_\ell a_\ell(t)} + \frac{\delta_n \beta_\ell^n}{\chi_\ell^n} \\
&\qquad + \beta_n^n \Big[A_n - \theta_n A_n^2 \big(1 - \chi_n - a_n(t)(\mu_{nt}^*)^{b_n}\big)\Big] \omega_n^\ell &&= 0, \\
&\frac{\delta_n \beta_n^n}{\chi_n^n} + \beta_n^n \Big[A_n - \theta_n A_n^2 \big(1 - \chi_n - a_n(t)(\mu_{nt}^*)^{b_n}\big)\Big] &&= 0, \\
&\lambda_{\ell t}^n \varepsilon_\ell^n &&= 0,
\end{aligned}$$

where $\lambda_k^n \geq 0$ is the Kuhn-Tucker multiplier associated with the non-negativity constraint $\varepsilon_k^n \geq 0$. If at least two countries face different marginal abatement costs there is a positive optimal transfer strategy $(\varepsilon_k^n)_{k,n\in\{1,\ldots,N\},k\neq n}$, which increases global abatement compared to the non-cooperative game without transfers.

Proof. To prove the result in the second scenario, we follow the three-step procedure explained in the main text.

1st step. The solution to the optimization problem of the first step is given by Theorem 4.2.

2nd step. To formulate the optimization problem of the second step (over consumption and transfers), we express transfers in relative terms and use the notation $\varepsilon_\ell^n = \mathcal{T}_\ell^n / \mathcal{C}_\ell^n$. By (5) and (67), the abatement policy of country ℓ after transfers is given by

$$\widehat{\mu}_\ell = \Big(\frac{a_\ell(t)(\mu_\ell^*)^{b_\ell} + \sum_{k\neq\ell}\chi_\ell^k\varepsilon_\ell^k}{a_\ell(t)}\Big)^{1/b_\ell}$$

The HJB equation of country n thus reads

$$\begin{aligned}\delta_n J^n = \sup_{\chi_1^n,\dots,\chi_N^n,(\varepsilon_k^n)_{k\neq n}} \Big\{ & \delta_n \sum_{\ell=1}^N \beta_\ell^n \log\big(\chi_\ell^n A_\ell K_\ell(1-\varepsilon_\ell^n)\big) + J_t^n + \sum_{\ell=1}^N J_{K_\ell}^n K_\ell\big[g_\ell(t,\chi_\ell,\mu_\ell^*) - \xi_\ell T^{at}\big] \\ & + J_{M^{at}}^n\Big(-\varphi_{12}M^{at} + \varphi_{21}M^{uo} + \sum_{\ell=1}^N \overline{Y}_\ell\sigma_\ell(1-\widehat{\mu}_\ell) + E_\ell^{\text{land}}\Big) \\ & + J_{M^{uo}}^n(\varphi_{12}M^{at} + (-\varphi_{21}-\varphi_{23})M^{uo} + \varphi_{32}M^{lo}) + J_{M^{lo}}^n(\varphi_{23}M^{uo} - \varphi_{32}M^{lo}) \\ & + \big(\kappa_\tau[\eta_0 + \eta_1 M^{at}/M^{\text{PI}} + F^{\text{ex}}] - (\phi+\phi_{21})T^{at} + \phi_{21}T^o\big)\, J_{T^{at}}^n + \phi_{12}(T^{at}-T^o)J_{T^o}^n \Big\}\end{aligned}$$

To determine the indirect utility function J^n, we now substitute the conjecture (49) into the HJB system and choose $p_{M^{at}}^n, p_{M^{uo}}^n, p_{M^{lo}}^n, p_{T^{at}}^n, p_{T^o}^n$ such that the separation holds true, i.e., we choose the same sensitivities as in Theorem 4.2. The simplified HJB system is thus given by:

$$\begin{aligned}\delta_n p^n(t) = \sup_{\chi_1^n,\dots,\chi_N^n,(\varepsilon_k^n)_{k\neq n}} \Big\{ & \dot{p}^n(t) + \delta_n \sum_{\ell=1}^N \beta_\ell^n \log\big(\chi_\ell^n A_\ell K_\ell(1-\varepsilon_\ell^n)\big) + \sum_{\ell=1}^N \beta_\ell^n K_\ell g_\ell(t,\chi_\ell,\mu_\ell^*) \\ & - p_{M^{at}}^n\Big(\sum_{\ell=1}^N \overline{Y}_\ell\sigma_\ell\Big[1 - \Big(\frac{a_\ell(t)(\mu_\ell^*)^{b_\ell} + \sum_{k\neq\ell}\chi_\ell^k\varepsilon_\ell^k}{a_\ell(t)}\Big)^{1/b_\ell}\Big] + E_\ell^{\text{land}}\Big) - p_{T^{at}}^n\kappa_\tau[\eta_0 + \eta_1 M^{at} + F^{\text{ex}}]\Big\}\end{aligned}$$

The first-order conditions for country n imply the following non-linear system for the optimal consumption and transfer decisions

$$p_{M^{at}}^n \overline{Y}_\ell\sigma_\ell\Big(\frac{a_\ell(t)(\mu_\ell^*)^{b_\ell} + \sum_{k\neq\ell}\chi_\ell^k\varepsilon_\ell^k}{a_\ell(t)}\Big)^{\frac{1-b_\ell}{b_\ell}} \frac{\chi_\ell^n}{b_\ell a_\ell(t)} - \frac{\delta_n\beta_\ell^n}{1-\varepsilon_\ell^n} = \lambda_{\ell t}^n, \tag{68}$$

$$\begin{aligned}p_{M^{at}}^n \overline{Y}_\ell\sigma_\ell\Big(\frac{a_\ell(t)(\mu_\ell^*)^{b_\ell} + \sum_{k\neq\ell}\chi_\ell^k\varepsilon_\ell^k}{a_\ell(t)}\Big)^{\frac{1-b_\ell}{b_\ell}} \frac{\varepsilon_\ell^n}{b_\ell a_\ell(t)} + \frac{\delta_n\beta_\ell^n}{\chi_\ell^n} & \\ + \beta_n^n\Big[A_n - \theta_n A_n^2\big(1-\chi_n - a_n(t)(\mu_{nt}^*)^{b_n}\big)\Big]\omega_n^\ell &= 0,\end{aligned} \tag{69}$$

$$\frac{\delta_n\beta_n^n}{\chi_n^n} + \beta_n^n\Big[A_n - \theta_n A_n^2\big(1-\chi_n - a_n(t)(\mu_{nt}^*)^{b_n}\big)\Big] = 0, \tag{70}$$

$$\lambda^n_{\ell t}\varepsilon^n_\ell = 0 \tag{71}$$

for $\ell \neq n$. Here, $\lambda^n_k \geq 0$ is the Kuhn-Tucker multiplier associated with the non-negativity constraint $\varepsilon^n_k \geq 0$. Due to its non-linearity, the above system cannot be solved in closed-form, but it defines a set of state-independent optimal controls which does not compromise our separation (49).

It is left to show that any solution with $\varepsilon^n_\ell = 0$ for all $\ell, n \in \{1, \ldots, N\}$ cannot be optimal. The first-order conditions then imply that the consumption strategies are given as in (54). To show that such a solution constitutes a saddle point and not an optimum, it is sufficient to show that country's n Hessian matrix H^n is indefinite. Consider without loss of generality $n = 1$. The principal diagonal elements $H^1_{\ell,\ell}$ are given by

$$\begin{aligned}
H^n_{\ell-1,\ell-1} &= (1-b_n)p^n_{M^{at}}\overline{Y}_\ell\sigma_\ell\Big(\frac{a_\ell(t)(\mu^*_\ell)^{b_\ell} + \sum_{k\neq\ell}\chi^k_\ell\varepsilon^k_\ell}{a_\ell(t)}\Big)^{\frac{1-2b_\ell}{b_\ell}}\Big(\frac{\chi^n_\ell}{b_\ell a_\ell(t)}\Big)^2 - \frac{\delta_n\beta^n_\ell}{(1-\varepsilon^n_\ell)^2},\\
H^n_{N,N} &= -\frac{\delta_n\beta^n_n}{(\chi^n_n)^2} + \beta^n_n\theta_n A^2_n,\\
H^n_{N+\ell,N+\ell} &= (1-b_n)p^n_{M^{at}}\overline{Y}_\ell\sigma_\ell\Big(\frac{a_\ell(t)(\mu^*_\ell)^{b_\ell} + \sum_{k\neq\ell}\chi^k_\ell\varepsilon^k_\ell}{a_\ell(t)}\Big)^{\frac{1-2b_\ell}{b_\ell}}\Big(\frac{\varepsilon^n_\ell}{b_\ell a_\ell(t)}\Big)^2 - \frac{\delta_n\beta^n_\ell}{(\chi^n_\ell)^2} + \beta^n_n\theta_n A^2_n\omega^\ell_n
\end{aligned}$$

for $\ell = 2, \ldots, N$. Evaluating the Hessian matrix at the critical point $\varepsilon^n_\ell = 0$ for all $\ell, n \in \{1, \ldots, N\}$ implies

$$H^n_{\ell-1,\ell-1} < 0, \qquad H^n_{N+\ell,N+\ell} > 0, \qquad H^n_{N,N} > 0$$

since $b_n > 1$. Therefore, the Hessian matrix is indefinite, which shows that such a solution cannot be optimal. Hence, in the optimum there must be at least one positive transfer between countries provided that there is an inner solution of (68)-(70). However, it is obvious that an inner solution exists if at least two countries face different marginal abatement costs.

3rd step. Now, assume that country n does not receive any transfers, i.e., $\varepsilon^\ell_n = 0$ for all $\ell \neq n$. Such a country is allowed to reoptimize abatement after transfers. If this country reoptimizes μ_n, then the first-order condition is identical to the case without transfers. Therefore, the optimal abatement policy stays the same. However, if abatement does not change, then consumption also does not change since it has already been optimized in the second step. □

H Extension: Non-proportional Damages

Following Rezai and van der Ploeg (2016) we consider a model with damages from climate change whose elasticity to output differs from one, i.e., we replace (17) by

$$\mathrm{d}K_{nt} = K_{nt} g_n(t, \chi_{nt}, \mu_{nt})\mathrm{d}t - \xi_n T_t^{at} K_{nt}^{\varepsilon_n} K_{n0}^{1-\varepsilon_n}\mathrm{d}t, \tag{72}$$

where $\varepsilon_n \in [0,1]$ is the elasticity of marginal damages with respect to capital. If $\varepsilon_n = 1$, damages are proportional to output and the dynamics collapse to the benchmark specification (17). By contrast, $\varepsilon_n = 0$ corresponds to a situation where marginal damages do not grow in line with economic activity, see Rezai and van der Ploeg (2016) and the references therein. The following corollary states the results. Its proof relies on a similar application of Banach's fixed-point theorem as in Theorem 4.2 and works along its lines. It is available upon request.

Corollary H.1 (Non-proportional Climate Damages)**.** *Under Assumptions 1, 2, 4 of Section 3 and the capital dynamics (72), the country-specific social cost of carbon is*

$$\mathrm{SCC}_{nt} = \frac{Y_{nt}^* \chi_{nt}^*}{M^{\mathrm{PI}} \delta_n} \sum_{\ell=1}^{N} \beta_\ell^n \xi_\ell \left(\frac{K_{\ell 0}}{K_{\ell t}^*}\right)^{1-\varepsilon_n} \frac{\kappa_\tau \eta_1}{\delta_n + \phi + \frac{\phi_{21}\delta_n}{\delta_n + \phi_{12}}} (\delta_n \mathbf{1}_{3\times 3} - \mathbf{\Phi}^\top)^{-1}_{1,1}.$$

For the special case of an autarky, $\beta_n^n = 1$, the country-specific SCC is given by

$$\mathrm{SCC}_{nt} = \frac{(Y_{nt}^*)^{\varepsilon_n} Y_{n0}^{1-\varepsilon_n} \chi_{nt}^*}{M^{\mathrm{PI}} \delta_n} \frac{\xi_n \kappa_\tau \eta_1}{\delta_n + \phi + \frac{\phi_{21}\delta_n}{\delta_n + \phi_{12}}} (\delta_n \mathbf{1}_{3\times 3} - \mathbf{\Phi}^\top)^{-1}_{1,1}.$$

As in Result 1 of Rezai and van der Ploeg (2016) the SCC is then proportional to $(Y_{nt}^*)^{\varepsilon_n} Y_{n0}^{1-\varepsilon_n}$. Their rule is slightly more general with respect to the preference specification, however. In particular, it involves the growth rate of GDP if the elasticity of intertemporal substitution departs from one.