DISCUSSION PAPER SERIES

DP15968

ACE - Analytic Climate Economy

Christian Traeger

CLIMATE CHANGE RPN

CEPR

ISSN 0265-8003

ACE - Analytic Climate Economy

Christian Traeger

Discussion Paper DP15968
Published 26 March 2021
Submitted 25 March 2021

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

ACE - Analytic Climate Economy

Abstract

The paper discusses optimal carbon taxation in an analytic quantitative integrated assessment model (IAM). The model links IAM components and parametric assumptions directly to their policy impacts. The paper discusses the distinct tax impact of carbon versus temperature dynamics and uses the see-through model to illustrate various aspects of IAM calibrations including the differentiation between consumption and investments goods. Novel to analytic IAMs are the explicit temperature dynamics, a general economy, energy sectors including capital, various degrees of substitutability across energy sources, an approximation of capital persistence, and objective functions that include CES preferences and population weighting. ACE opens the door to tractable forward-looking stochastic modeling and dynamic strategic interactions in complex IAMs, explored in accompanying work.

JEL Classification: Q54, H23, H43, E13, D80, D61

Keywords: climate change, Integrated assessment, Social cost of carbon, carbon tax, carbon cycle, Temperature, climate sensitivity, technological progress, capital persistence, population weighting

Christian Traeger - christian.traeger@econ.uio.no
University of Oslo, ifo Institute for Economic Research

ACE – Analytic Climate Economy*

Christian P. Traeger
Department of Economics, University of Oslo & ifo Institute, Munich

March 2021

Abstract: The paper discusses optimal carbon taxation in an analytic quantitative integrated assessment model (IAM). The model links IAM components and parametric assumptions directly to their policy impacts. The paper discusses the distinct tax impact of carbon versus temperature dynamics and uses the see-through model to illustrate various aspects of IAM calibrations including the differentiation between consumption and investments goods. Novel to analytic IAMs are the explicit temperature dynamics, a general economy, energy sectors including capital, various degrees of substitutability across energy sources, an approximation of capital persistence, and objective functions that include CES preferences and population weighting. ACE opens the door to tractable forward-looking stochastic modeling and dynamic strategic interactions in complex IAMs, explored in accompanying work.

JEL Codes: Q54, H23, H43, E13, D80, D61

Keywords: climate change, integrated assessment, social cost of carbon, carbon tax, carbon cycle, temperature, climate sensitivity, technological progress, capital persistence, population weighting, model calibration

*I am grateful for feedback and inspiration from Larry Karp, Terry Iverson, Svenn Jensen, Bard Harstad, Kjetil Storesletten, Armon Rezai, Buzz Brock, Lars Hansen, Rick van der Ploeg, Michael Greenstone, Andre Butz, Fortunat Joos, David Anthoff, Valentina Bossetti, Richard Tol, Christian Gollier, Drew Creal, Nour Meddahi, Ravi Bansal, Till Requate, Andreas Lange, Grischa Perino, Rob Nicholas, Klaus Keller, Nancy Tuana and participants of the PET 2015, AERE 2015, SURED 2016, CESifo 2016, SITE 2017, ASSA 2018, WCERE 2018, ESEM 2018, and of department seminars at Stanford University, University of Amsterdam, University of Miami, University of Oslo in 2014, College de France, University of Toulouse, London School of Economics, Ohio State, University of Cambridge in 2015, University of California, Berkeley, University of Colorado, Iowa State, University of Chicago, CESifo Munich in 2016, University of Gothenborg, University of Hamburg, University of Arizona in 2017, and Paris School of Economics and the University of Oldenburg in 2018. This work was supported by the National Science Foundation through the Network for Sustainable Climate Risk Management (SCRiM) under NSF cooperative agreement GEO-1240507. The present paper is the first out of three papers that emerged out of the original project titled "ACE - Analytic Climate Economy (with Temperature and Uncertainty)". The other two papers focus on Uncertainty and Emissions, respectively.

1 Introduction

Integrated assessment models (IAMs) of climate change analyze the long-term interactions of economic production, greenhouse gas (GHG) emissions, and global warming. The present paper develops an analytically tractable integrated assessment model composed of the typical components making up the quantitative numeric models used in policy advising. The Analytic Climate Economy (ACE) combines a general production system with state of the art climate dynamics. In a see-through framework, the model delivers new insights into the roles of climate dynamics, production characteristics, objective functions, and calibration approaches for optimal climate policy.

The Analytic Climate Economy (ACE) is a close relative to Nordhaus' (1994, 2013 with Sztorc, 2017) widely used Nobel-awarded DICE model. Despite being more general in most dimensions, it solves in closed form. ACE bridges the gap between the numeric IAMs used in policy advising and a quickly growing literature of analytic and semi-analytic approaches sparked by Golosov et al.'s (2014) seminal contribution. ACE puts forth a framework that is transparent, open to analytic introspection, and seriously quantitative at the same time. The present paper focuses on the optimal carbon tax. The Economist (2017) declared the immediately linked social cost of carbon (SCC) to be "the most important number in climate economics". The analysis focuses as much on understanding the structural drivers of this number as on their quantitative implications. A set of accompanying papers build on ACE discussing carbon dioxide (CO_2) emissions (Traeger 2021*a*), climate change uncertainties (Traeger 2021*b*) and analyzing strategic interactions between regions (Meier & Traeger 2021).

An important step in closing the gap between analytic and quantitative policy models is that ACE explicitly introduces temperature dynamics and the non-linear greenhouse effect relating global warming to the emissions that drive our economy. Modeling temperature is at the core of analyzing climate change and damages. I show that temperature and the previously (analytically) modeled carbon dynamics have entirely different impacts on the optimal carbon tax. The high persistence of carbon (mass conservation) increases the optimal tax four to thirty-fold, depending on the calibration of time preference. By contrast, the delay in the temperature dynamics (ocean cooling) cuts the carbon tax by approximately 40% to 20%. Moreover, the distinction between between temperature and carbon turns out crucial for incorporating uncertainty, modeling geoengineering, and understanding strategic interactions as shown in the accompanying work.

ACE substantially generalizes the structure of the economy and the energy sectors w.r.t. to earlier analytic models as well as the DICE model. Energy production relies on clean and fossil-fuel-based sectors that employ labor, capital, and potentially scarce resources. Technological progress is exogenous.[1] I show that the only restriction on production required for the present results are that overall production in the economy is homogenous in capital (of arbitrary degree keeping the problem concave). IAMs usually rely on complex and at times crude calibrations of the energy sectors or, as in the case of DICE, reduced form guesstimates governing an exogenous evolution of carbon intensity and abatement costs. ACE fleshes out those assumptions that matter most for the optimal carbon tax. Moreover, it shows that the oft-criticized absence of empirically important capital in the energy production sectors in Golosov et al.'s (2014) model and the subsequent literature is not as crucial as often assumed. ACE also eases the analytic literature's full depreciation assumption. Extensions of the base model discuss the role of DICE's population weighting for the carbon tax, introduce heterogenous consumption levels (only to show how they do not matter), distinguish between consumption and investment goods (discussing the importance for model calibration), and introduce CES-preferences over a variety of consumption goods.

Literature. Analytic approaches to the integrated assessment of climate change date back to at least Heal's (1984) insightful non-quantitative contribution. Several papers have used the linear quadratic model for a quantitative analytic discussion of climate policy (Hoel & Karp 2002, Newell & Pizer 2003, Karp & Zhang 2006, Karp & Zhang 2012). A disadvantage of these linear quadratic approaches is their highly stylized representation of the economy and the climate system. In particular, these models have no production or energy sector.

Golosov et al. (2014) broke new ground by amending the log-utility and full-depreciation version of Brock & Mirman's (1972) stochastic growth model with an energy sector and an impulse response of production to emissions. A decadal time step is not uncommon in IAMs, rendering the full-depreciation assumption more reasonable than in other macroeconomic contexts. The present paper weakens the full-depreciation assumption to improve the descriptive power of the analytic approach. Golosov et al.'s (2014) impulse response to emissions assumes that economic production (and implicitly temperature) responds immediately to atmospheric CO_2. However, the oceans keep

[1] Accompanying work in progress extends the ACE model to encompass endogenous technological progress. This work also relaxes the full depreciation assumption even further.

cooling us for decades. Gerlagh & Liski (2018*b*) extend the model by introducing the empirically important delay between emission accumulation and damages, and analyze the implications of non-constant rates of time preference.[2] The present paper follows the numeric IAMs used in policy advising and explicitly introduces the non-linearity in the relation between atmospheric CO_2 and temperatures caused by the greenhouse effect that drives climatic change (Nordhaus 2008, Hope 2006, Bosetti et al. 2006, Anthoff & Tol 2014). ACE incorporates alternatively DICE's original carbon cycle or an impulse response model used in the IPCC (2013), recently promoted by Van der Ploeg et al. (2020). It also develops a novel model of ocean-atmosphere temperature dynamics that permits an analytic solution, while closely matching the dynamics of numeric climate change models.

Golosov et al.'s (2014) framework has sparked a growing literature on Analytic Integrated Assessment Models (AIAMs), including applications to a multi-regional setting (Hassler & Krusell 2012, Hassler et al. 2018, Hambel et al. 2021), non-constant discounting (Gerlagh & Liski 2018*b*, Iverson & Karp 2020), intergenerational games (Karp 2017), and regime shifts (Gerlagh & Liski 2018*a*). As pointed out by Karp (2017), these frameworks solves analytically because they can be transformed to a system that is linear in the states' equations of motion, a fact I use for a simpler presentation and solution of the more general ACE model. Importantly, the present ACE model is not linear in temperature, which has crucial implications for the accompanying applications to uncertainty and strategic interactions. The present ACE model assumes logarithmic utility to capture intertemporal substitution. The accompanying extension to uncertainty employs an arbitrary degree of relative Arrow-Pratt risk aversion, disentangling it from the unit elasticity of intertemporal substition (using Epstein-Zin-Weil preferences).

Anderson et al. (2014), Brock & Xepapadeas (2017), Dietz & Venmans (2018), and van der Ploeg (2018) spearhead the use of a simple and yet powerful climate model in economic integrated assessment. This so-called transient climate response to cumulative carbon emissions (TCRE) model builds on the convenient observation that several

[2]Matthews et al. (2009) and subsequent work including the IPCC (2013) suggest that explicit models of carbon and temperature dynamics can be approximated by a direct response of temperatures to *cumulative historic emissions*. It is a frequent misunderstanding that these findings render modeling CO_2 *concentrations* sufficient and ocean cooling negligible (see final paragraph of introduction). The models merely suggest that the impulse response to emissions (along particular steady emission scenarios) are fairly flat, rather than peaked, as assumed in Gerlagh & Liski's (2018*b*) DICE-based calibration of damage delay.

dynamic effects in temperature and carbon dynamics "offset" each other, at least along somewhat stable emission scenarios. The clever model constitutes another great starting point for analytic climate modeling. Yet, it has led to a misunderstanding that temperature delay (as well as non-linearities) are absent or irrelevant. By explicitly pinning down the effect of carbon and temperature dynamics on the SCC, ACE hopefully resolves related misunderstandings. I refer to Bosetti (2021) and Metcalf & Stock (2017) for overviews of numeric IAMs of climate change;[3] The most prominent IAMs for SCC calculations have been the DICE, PAGE (Hope 2006), and FUND (Anthoff & Tol 2014) models, partly because the Interagency Working Group on Social Cost of Carbon (2013) chose these three models to establish an official SCC for the US, likely accompanied by WITCH (Emmerling et al. 2016) and REMIND (Luderer et al. 2021) for scenario, energy sector, and emission analysis. I refer to the accompanying paper Traeger (2021*b*) for a review of the fast growing literature on uncertainty in climate change.

2 The Model

ACE's structure follows (and generalizes) that of most IAMs, see Figure 1. Labor, capital, technology, and energy produce output that is either consumed or invested. "Dirty" energy sectors consume fossil fuels and cause emissions, which accumulate in the atmosphere, cause radiative forcing (greenhouse effect), and increase global temperature(s), thus reducing output. This section introduces the basic model of the economy and the climate system.

2.1 ACE's Economy

Production and energy sectors. Final gross output Y_t is a function of vectors of exogenous technologies $\boldsymbol{A}_t$, the optimally allocated labor and capital distributions $\boldsymbol{N}_t$ and $\boldsymbol{K}_t$, and a flow of potentially scarce resource inputs $\boldsymbol{E}_t$

$$Y_t = F(\boldsymbol{A}_t, \boldsymbol{N}_t, \boldsymbol{K}_t, \boldsymbol{E}_t) \quad \text{with} \tag{1}$$
$$F(\boldsymbol{A}_t, \boldsymbol{N}_t, \gamma\boldsymbol{K}_t, \boldsymbol{E}_t) = \gamma^\kappa F(\boldsymbol{A}_t, \boldsymbol{N}_t, \boldsymbol{K}_t, \boldsymbol{E}_t)\ \forall\gamma \in \mathbb{R}_+.$$

[3]The Integrated Assessment Modeling Consortium's website at https://www.iamconsortium.org/ offers a rich set of resources with a focus on the more complex IAMs and scenario analysis. See Pindyck (2013) for a particularly critical analysis of IAMs and the suggestion to shift attention from formal models to experts.

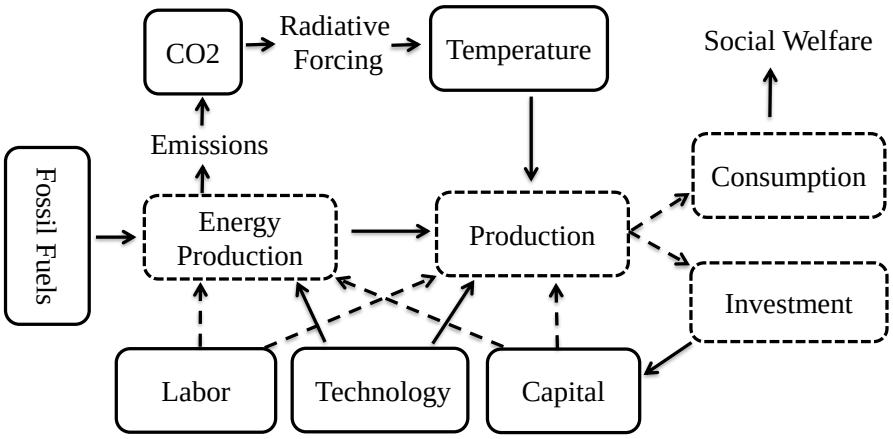

Figure 1: The structure of ACE and most IAMs. Solid boxes characterize the model's state variables, dashed boxes are flows, and dashed arrows mark choice variables.

The production function is homogenous of degree κ in capital and has to be sufficiently well-behaved to deliver well-defined solutions to the optimization problem, which will generally imply $\kappa \leq 1$. The general functional form covers explicit production structures with intermediates and a variety of clean and dirty energy sectors relying on different, possibly time-changing degrees of substitutability. It generalizes special cases in the earlier literature and, in contrast to earlier analytic models, allows the energy sectors to utilize capital. The input vectors are of dimension $I_j \in \mathbb{N}$ with $j \in \{A, N, K, E\}$. Aggregate capital K_t is optimally distributed across sectors such that $\sum_{i=1}^{I_K} K_{i,t} = K_t$ and similarly $\sum_{i=1}^{I_N} N_{i,t} = 1$. I denote by $\mathcal{K}_{i,t} = \frac{K_{i,t}}{K_t}$ the share of capital in sector i. Section 5.2 discusses a concrete example of the present production structure.

Emissions and resources. The first I^d resources $E_1, ..., E_{I^d}$ are fossil fuels and emit CO_2; I collect them in the subvector $\boldsymbol{E}_t^d$ ("dirty"). I measure these fossil fuels in terms of their carbon content and total emissions from production amount to $\sum_{i=1}^{I^d} E_{i,t}$. In addition, land conversion, forestry, and agriculture emit smaller quantities of CO_2. Following the DICE model, I treat these additional anthropogenic emissions as exogenous and denote them by E_t^{exo}.

Renewable energy production relies on the inputs indexed by I^{d+1} to I_E such as water, wind, or sunlight, which I assume to be abundant. By contrast, fossil fuel use reduces the resource stock in the ground $\boldsymbol{R_t} \in \mathbb{R}_+^{I^d}$:

$$\boldsymbol{R}_{t+1} = \boldsymbol{R}_t - \boldsymbol{E}_t^d, \tag{3}$$

with initial stock levels $\boldsymbol{R}_0 \in \mathbb{R}_+^{I^d}$ given. Possible extraction costs are part of the general production function. I take the following assumption to avoid boundary value complications. If a resource is scarce along the optimal path, its use is stretched over the infinite time horizon.[4]

Damages. The next section explains how the carbon emissions increase the *global atmospheric temperature* $T_{1,t}$ measured as the increase over the preindustrial temperature level. This temperature increase causes damages, which destroy a fraction $D_t(T_{1,t})$ of output. *Damages at the preindustrial temperature level are* $D_t(0) = 0$ and Proposition 1 characterizes the class of damage functions $D_t(T_{1,t})$ that permit an analytic solution of the model.

Capital accumulation. The base model assumes that the part of production left after climate damages and consumption is invested

$$I_t = \underbrace{Y_t[1 - D_t(T_{1,t})]}_{\equiv Y_t^{net}} - C_t. \tag{4}$$

Section 5.2's extensions also allow for a dedicated investment composite. A limitation of integrated assessment models building on Golosov et al. (2014) has been the assumption of full depreciation. These models assume that next period's capital is current period's investment, i.e., $K_{t+1} = I_t$. IAMs often run in time steps of several years and Golosov et al. (2014) use full capital depreciation in combination with a 10 years time step. Yet, some applications can require shorter times steps and even over 10 years full depreciation underestimates capital accumulation. I suggest a simple extension that does not fully fix, but at least ameliorate the issues resulting from a full-depreciation assumption. Instead of equating next period's capital stock with current investment, I assume

$$K_{t+1} = I_t \left[\frac{1 + g_{k,t}}{\delta_k + g_{k,t}} \right] , \tag{5}$$

where δ_k is the capital depreciation factor, and $g_{k,t}$ is an exogenous approximation of the growth rate of capital. A full depreciation assumption (i) misses the remaining capital from previous periods and (ii) misses that current investments lead to more capital availability in the future. Equation (5) is a "quick fix" addressing both of these issues. A depreciation factor below unity implies a multiplier of current investments larger

[4] A sufficient but not necessary condition is that the scarce resources are essential in production, i.e., that production is not possible without the input of the scarce resource.

than unity. The additional factor increases both, next period's capital availability and the current investment's payoff in terms of future capital. Of course, the timing is a bit off, compressing the fixes to issues (i) and (ii) into the same period rather than spreading them properly over time. Appendix A.1 shows that equation (5) coincides with the standard equation of motion for capital accumulation

$$K_{t+1} = Y_t[1 - D_t(T_{1,t})] - C_t + (1 - \delta_k)K_t$$

if the exogenous capital growth approximation is correct, $g_{k,t} = \frac{K_{t+1}}{K_t} - 1$, or if $\delta_k = 1$ (full depreciation). To the best of my knowledge, turning the full depreciation model into an approximate model of capital persistence is novel, also to the broader literature. It adjusts ACE's capital dynamics and capital-output ratio to macroeconomic observation and makes the decision maker aware of additional future return to capital investments.[5] Typical applications, including those of ACE, calibrate the underlying Ramsey growth model to observed data, which directly determine capital growth along the balanced growth path. Thus, the correction factor's reliance on the exogenous growth rate is probably a rather mild limitation. Relying on the Penn World Tables (Feenstra et al. 2015), the correction factor for global capital growth is $\left[\frac{1+g_{k,2009-19}}{\delta_k+g_{k,2009-19}}\right] \approx 1.8$ adopting Golosov et al.'s (2014) 10 year time step. For the five year time step adopted in DICE, the correction factor would be 3.1. For, e.g., the U.S. alone the correction factor would be slightly higher (1.9 and 3.6, respectively, see Appendix A.1 for details).

2.2 ACE's Climate System

This section introduces the deterministic baseline specification of ACE's climate system.

Carbon cycle. Carbon released into the atmosphere does not decay, it only cycles through different carbon reservoirs. Let $M_{1,t}$ denote the atmospheric carbon content and let $M_{2,t}, ..., M_{m,t}$, $m \in \mathbb{N}$, denote the carbon content of a finite number of non-atmospheric carbon reservoirs. DICE uses two carbon reservoirs besides the atmosphere: $M_{2,t}$ captures the combined carbon content of the upper ocean and the

[5]The limited depreciation factor has no impact on the optimal carbon policy, given current world output. Yet, it is relevant only for the evolution of the model over time. The relevant implication of the capital accumulation in equation (5) is that the investment rate is independent of the system states. Consequently, climate policy will not operate through the consumption rate. Appendix A.1 shows that the consumption rate is roughly independent of the climate states, also in an annual time-step version of the widespread numeric IAM DICE (using non-logarithmic utility and the standard capital equation of motion), in particular, under optimal climate policy.

biosphere (mostly plants and soil) and $M_{3,t}$ captures the carbon content of the deep ocean. The vector $\boldsymbol{M_t}$ comprises the carbon content of the different reservoirs, and the matrix $\boldsymbol{\Phi}$ captures the transfer coefficients. Then

$$\boldsymbol{M}_{t+1} = \boldsymbol{\Phi}\boldsymbol{M}_t + \boldsymbol{e}_1(\textstyle\sum_{i=1}^{I^d} E_{i,t} + E_t^{exo}) \tag{6}$$

captures the carbon dynamics. The first unit vector $\boldsymbol{e}_1$ channels new emissions from fossil fuel burning $\sum_{i=1}^{I^d} E_{i,t}$ and from land use change, forestry, and agriculture E_t^{exo} into the atmosphere $M_{1,t+1}$. The fact that carbon does not decay, but only moves across reservoirs implies that the columns of the transition matrix Φ sum to unity (*mass conservation of carbon*).

Greenhouse effect. An increase in atmospheric carbon causes a change in our planet's energy balance, which leads to heating.[6] This heating is known as anthropogenic radiative forcing and is concave in atmospheric CO_2:

$$F_t = \eta \frac{\log \frac{M_{1,t}+G_t}{M_{pre}}}{\log 2} \ . \tag{7}$$

The exogenous process G_t captures the contribution from other non-CO_2 GHGs (measured in CO_2 equivalents). Anthropogenic radiative forcing was absent in preindustrial times, when $G_t = 0$ and $M_{1,t}$ was equal to the preindustrial atmospheric CO_2 concentration M_{pre}. The parameter η captures the strength of the greenhouse effect; every time CO_2 concentrations double, the forcing increases by η. Whereas *radiative forcing is immediate*, the planet's *temperature* responds with major *delay*; warming our planet with its oceans is like warming a big pot of soup on a small flame. After decades to centuries, the new equilibrium[7] temperature of the (lower) atmosphere caused by a new level of radiative forcing F^{new} will be $T_{1,eq}^{new} = \frac{s}{\eta}F^{new} = \frac{s}{\log 2}\log\frac{M_{1,eq}+G_{eq}}{M_{pre}}$. The parameter s is the *climate sensitivity*, measuring the medium- to long-term *temperature response to a doubling of preindustrial CO_2 concentrations*. Its best estimate is currently around 3°C, but the true temperature response to a doubling of CO_2 is uncertain.[8]

[6] In equilibrium, our planet radiates the same amount of energy out into space that it receives from the sun. Atmospheric carbon $M_{1,t}$ and other GHGs "trap" some of this outgoing (infrared) radiation, which leads to a warming commonly referred to as the greenhouse effect.

[7] The conventional climate equilibrium incorporates feedback processes that take several centuries, but excludes feedback processes that operate at even longer time scales, e.g., the full response of the ice sheets.

[8] Such uncertainty was first introduced into integrated assessment models by Kelly & Kolstad (1999)

Temperature Dynamics. The next period's atmospheric temperature depends on current atmospheric temperature, current temperature in the upper ocean, and on radiative forcing (heating). I denote the temperatures of a finite number of ocean layers by $T_{i,t}, i \in \{2, ..., l\}, l \in \mathbb{N}$. I abbreviate the atmospheric equilibrium temperature resulting from the radiative forcing level F_t by $T_{0,t} = \frac{s}{\eta} F_t$. Each layer slowly adjusts its own temperature to the temperatures of the surrounding layers. Numeric IAMs usually approximate this temperature adjustment as a linear process, which would prevent an analytic solution of the model. Yet, heat exchange is governed by many nonlinear processes (radiative, convective, evaporative) in addition to linear diffusion. I model the next period's temperature in layer $i \in \{1, ..., l\}$ as a generalized (rather than arithmetic) mean of its current temperature $T_{i,t}$ and the current temperatures in the adjacent layers $T_{i-1,t}$ and $T_{i+1,t}$[9]

$$T_{i,t+1} = \mathfrak{M}_i^{\boldsymbol{\sigma}}(T_{i,t}, T_{i-1,t}, T_{i+1,t}) \text{ for } i \in \{1, ..., l\}, \tag{8}$$

where $T_{0,t} = \frac{s}{\eta} F_t$. The weight matrix $\boldsymbol{\sigma}$ characterizes the (generalized) heat flow between adjacent layers, and $\sigma^{forc} = 1 - \sigma_{1,1} - \sigma_{1,2}$ characterizes the heat influx response to radiative forcing. Proposition 1 in the next section characterizes the class of means (weighting functions f) that permit an analytic solution.

3 Objective, Solution, and Calibration

The present section follows the common approach of the IAM literature of setting up, solving, and calibrating the model in social planner form. Section 5 discusses extensions and alternatives. I characterize the class of damage functions and ocean-atmosphere temperature dynamics that permit an analytic solution (Proposition 1). To make this solution relevant, these functions must permit a reasonable calibration, which I discuss subsequently to the proposition.

---

and has since been discussed in a variety of papers (Jensen & Traeger 2013, van den Bremer & van der Ploeg 2018, Hambel et al. 2018). The accompanying Traeger (2021*b*) integrates this uncertainty into ACE.

[9] A generalized mean is an arithmetic mean enriched by a nonlinear weighting function f. It takes the form $\mathfrak{M}_i(T_{i-1,t}\ ,\ T_{i,t}\ ,\ T_{i+1,t}) = f^{-1}[\sigma_{i,i-1} f(T_{i-1,t}) + \sigma_{i,i} f(T_{i,t}) + \sigma_{i,i+1} f(T_{i+1,t})]$ with weight

3.1 Objective

The social planner's time horizon is infinite and he or she discounts future welfare with a utility discount factor $\beta < 1$. Population growth is exogenous and welfare is logarithmic in aggregate consumption C_t

$$\max \sum_{t=0}^{\infty} \beta^t \log C_t \ . \tag{9}$$

Given logarithmic utility, this assumption is equivalent to maximizing average per capita consumption (see Section 5 for details and extensions). The implied unit elasticity of intertemporal substitution (EIS) is both the mode and the median of Drupp et al.'s (2018) expert survey composed of 200 experts on social discounting. The majority of the macroeconomic literature suggests that this unit EIS is a bit too high (Havránek 2015). The long-run risk literature, which I rely on when extending ACE to uncertainty (Traeger 2021*b*), argues strongly for an even higher EIS (Vissing-Jørgensen & Attanasio 2003, Bansal & Yaron 2004, Chen et al. 2013, Nakamura et al. 2013, Bansal et al. 2014, Kung & Schmid 2015, Collin-Dufresne et al. 2016, Engel 2016, Bansal et al. 2016, Nakamura et al. 2017, Jagannathan & Liu 2019).[10] This extension to uncertainty decouples the EIS from risk aversion and solves the stochastic ACE for a general coefficient of constant relative risk aversion. Section 5.2 extends the model to a variety of consumption goods and CES preferences, which can also explain observed deviations from a unit EIS. Having argued that a unit EIS is *a* reasonable choice, the obvious attraction is that the resulting model permits not only a quantification but also analytic insights.

3.2 General Solution

Equations (1-9) characterize the base ACE model. The policy maker optimizes energy and labor inputs, as well as consumption and investments to maximize discounted

$\sigma_{i,i} = 1 - \sigma_{i,i-1} - \sigma_{i,i+1} > 0$. The weight $\sigma_{i,j}$ characterizes the (generalized) heat-flow coefficient from layer j to layer i. Heat flow between any two non-adjacent layers is zero. Note that the weight $\sigma_{i,i}$ captures the warming persistence (or inertia) in ocean layer i. The weight $\sigma^{forc} \equiv \sigma_{1,0} = 1 - \sigma_{1,1} - \sigma_{1,2}$ determines the heat influx caused by radiative forcing. I define $\sigma_{l,l+1} = 0$: the lowest ocean layer exchanges heat with only the next upper layer. For notational convenience, equation (8) writes a mean of three temperature values also for the deepest layer ($i = l$), with a zero weight on the arbitrary entry T_{l+1}. I collect all weights in the $l \times l$ matrix $\boldsymbol{\sigma}$, which characterizes the heat exchange between the atmosphere and the different ocean layers.

[10] See Traeger (2019) for a suggested explanation of the opposing findings between the macroeconomic and the long-run risk literature.

logarithmic welfare over the infinite time horizon. Appendix A.2 transforms the ACE, making the equations of motion linear in the (transformed) states and finding separable controls. Such models are solved by an affine value function. The transformations also flesh out which changes would maintain (or eliminate) analytic tractability.

Proposition 1 *An affine value function of the form*

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

solves the deterministic ACE if and only if[11] $k_t = \log K_t$*,* $\boldsymbol{\tau}_t$ *is a vector composed of the generalized temperatures* $\tau_{i,t} = \exp(\xi_1 T_{i,t})$*,* $i \in \{1, ..., L\}$*, the damage function takes the form*

$$D(T_{1,t}) = 1 - \exp(-\xi_0 \exp[\xi_1 T_{1,t}] + \xi_0) \tag{10}$$

with $\xi_0 \in \mathbb{R}$ *and the mean in the equation of motion* (8) *for temperature layer* $i \in \{1, ..., l\}$ *takes the form*

$$\begin{aligned}\mathfrak{M}_i^{\boldsymbol{\sigma}}(T_{i,t}, T_{i-1,t}, T_{i+1,t}) = \frac{1}{\xi_1} \log \Big(&(1 - \sigma_{i,i-1} - \sigma_{i,i+1}) \exp[\xi_1 T_{i,t}] \\ &+ \sigma_{i,i-1} \exp[\xi_1 T_{i-1,t}] + \sigma_{i,i+1} \exp[\xi_1 T_{i+1,t}] \Big)\end{aligned} \tag{11}$$

with parameter $\xi_1 = \frac{\log 2}{s} \approx \frac{1}{4}$*. The solutions for the shadow values are summarized in the proof.*

Appendix B provides the proof and subsequent sections discuss the result in detail. The coefficients φ in the value function are the shadow values of the respective state variables. The symbol $^\top$ denotes transposition of the shadow value vectors. For example, $\varphi_{M,1}$ is the shadow value of atmospheric carbon and will play a crucial role in determining the optimal carbon tax. The coefficient vector on the resource stock, $\boldsymbol{\varphi}_{R,t}^\top$, must be time-dependent: the shadow values of scarce exhaustible resources increase over time, following the endogenously derived Hotelling rule (see equation B.8 in the Appendix). The process φ_t captures the value contribution of the exogenous processes, including technological progress.

[11] Affine transformations of the (transformed) state variables are also permitted, which essentially correspond to a change in the measurement scale.

3.3 General Discussion and Calibration

The damage function is of a double-exponential form with a free parameter ξ_0, which scales the severity of damages at a given temperature level. This free parameter ξ_0 is the *semi-elasticity of output with respect to a change of transformed atmospheric temperature* $\tau_{1,t} = \exp(\xi_1 T_{1,t})$, i.e., with respect to the exponential of the change of temperature. ACE's base calibration matches DICE (Nordhaus 2008, Nordhaus & Sztorc 2013, Nordhaus 2017). Figure 2 plots the damage functions of the last three DICE versions (shades of green) together with ACE's base calibration (dashed green). The base calibration is an exact match of the two calibration points 0 and 2.5°C of the 2007 model, which had a slightly higher damage coefficient than the later two generations (to which ACE is closer for higher temperature levels).[12] The calibration delivers the damage semi-elasticity $\xi_0 = 0.022$. Figure 2 also depicts dotted lines presenting a ±50% deviation of this value, which mostly bounds the different DICE damage curves. ACE's base calibration adopts this green dashed damage function, sticking closely to the wide-spread and Nobel-awarded DICE model.

Figure 2 also graphs Howard & Sterner's (2017) preferred damage function resulting from their meta-analysis of earlier studies including DICE. This damage function suggests a good 10% loss of world output at a $3°C$ warming, which is about 4 times the DICE damage. Such a 10% loss surrounding a $3°C$ warming is also supported by Pindyck's (2020) recent survey among economists and climate scientists, which finds a similar 'most likely' GDP loss for the year 2066 following business as usual.[13] Howard & Sterner (2017) do not extrapolate their damage function beyond $6°C$ and I follow DICE's method of replacing quadratic damages of aT^2 by $\frac{1}{1+aT^2}$ to limit damages to 100% of production. While damages can exceed world output in principle, without such curtailing they would also quickly exceed the world capital stock. It is a judg-

[12] Figure 2 plots the damage curves specified in the manuals and papers, which is of the form $D(T) = 1 - \frac{1}{1+aT^2}$ (Nordhaus 2008, Nordhaus & Sztorc 2013, Nordhaus 2017). While consistent with the EXCEL and GAMS codes of the earlier versions of DICE, the GAMS code of the 2013 model and the DICE2016R-091916ap model seem to adopt $D(T) = aT^2$ instead. I decided to adopt the normalized version stated in the texts and papers that avoid damages exceeding world output. While of minor relevance for DICE's damage calibration, it will be more crucial for the high damage calibration discussed below. DICE's damage coefficient is $a = 0.0028$ for 2007, $a = 0.00267$ for 2013, and $a = 00236$ for 2016.

[13] Pindyck asks about the year 2066 (50 years after his survey), assuming business as usual. As seen in Figure 3 for the RCP 8.5 scenario, BAU by 2066 corresponds approximately to a $3°C$ warming, even if "business as usual" leaves room for interpretation. The average response was a most likely loss of 10.8%, where the economists averaged a guesstimate of 8.6%. Pindyck contacted 6833 economists and climate scientists who published on related topics and the results reflect about 600 answers.

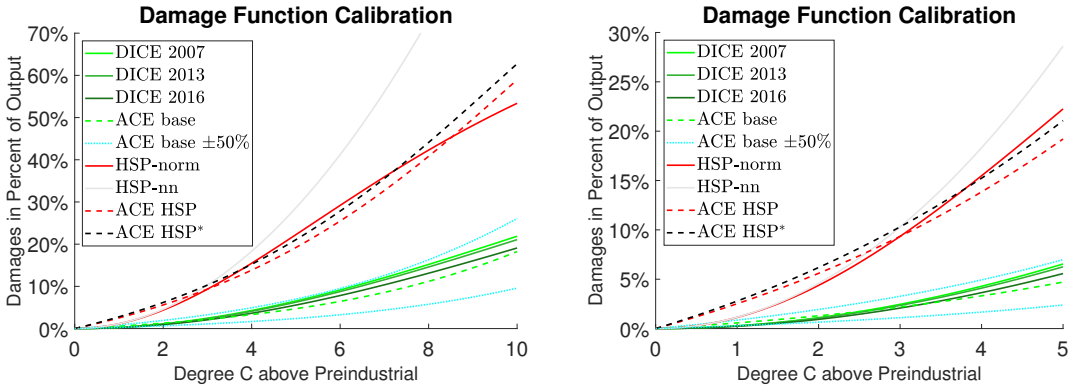

Figure 2: ACE's damage calibration (dashed) to DICE 2013 and to a substantially higher damage function based on Howard & Sterner (2017) and Pindyck (2020) (HSP). HSP-norm extrapolates Howard & Sterner's (2017) damage function using DICE's approach to limit damages to 100% of production, whereas HSP-nn extrapolates without renormalization and damages exceed production at a 9.5C warming. The left graph covers the temperature range of the IPCC scenarios in Figure 3 whereas the right graph zooms in on lower degrees of warming.

ment call and we hopefully never obtain the data for this business as usual range. The red solid curve HSP-norm reflects the result, and the light grey HSP-nn curve reflects the damage curve without re-normalization. The black dashed line calibrates ACE to the original damages at $3°C$, yielding $\xi_0 = 0.011$, and the red dashed line calibrates ACE to the re-normalized damages at $3°C$, yielding $\xi_0 = 0.10$, a difference of 10% in the damage coefficient.

For all calibrations, ACE's damage function delivers somewhat higher damages for a reasonably low temperature change, then slightly lower damage for the medium range, and eventually again higher damages at the high end of business as usual. It is initially less convex than the (normalized) quadratic function, and then more convex.

Temperature dynamics. The generalized mean $\mathfrak{M}_i^{\sigma}$ uses the nonlinear weighting function $\exp[\xi_1 \cdot]$. The calibration of temperature dynamics (equation 11) uses the representative concentration pathways (RCP) of the latest assessment report by the Intergovernmental Panel on Climate Change (IPCC 2013). I use the MAGICC6.0 model by Meinshausen et al. (2011) to simulate the RCP scenarios over a time horizon of 400 years. MAGICC6.0 emulates the results of the large atmosphere-ocean general circulation models (AOGCMs) and is employed in the IPCC's assessment report. DICE was calibrated to a (single) scenario using an earlier version of MAGICC. My calibration of ACE uses two ocean layers (upper and deep), compared to MAGICC's 50 layers and DICE's single ocean layer.

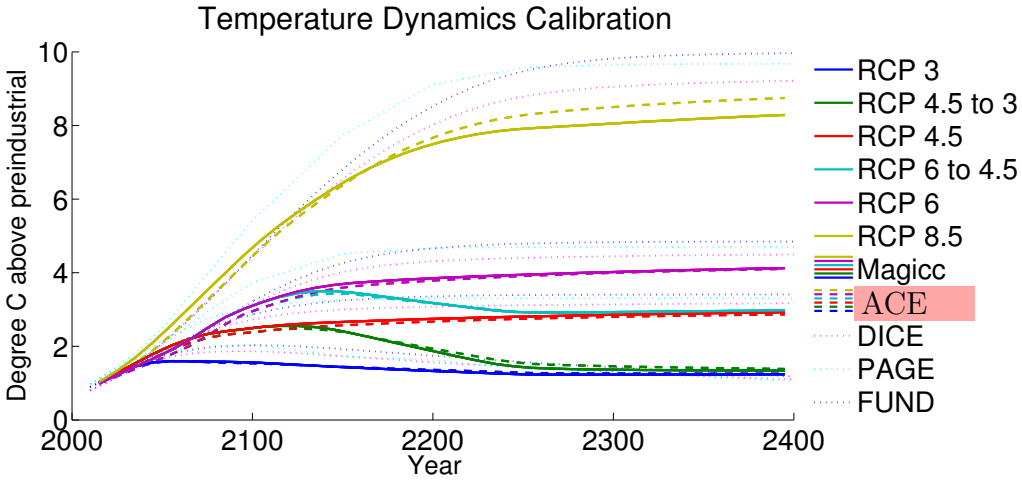

Figure 3: ACE's temperature response compared to MAGICC6.0 using the color-coded radiative forcing scenarios of the latest IPCC assessment report. RCP 3 is the strongest stabilization scenario, and RCP 8.5 is a business-as-usual scenario. The MAGICC model (solid lines) emulates the large AOGCMs and is used in the IPCC's assessment reports. ACE (dashed lines) matches MAGICC's temperature response very well for the "moderate" warming scenarios and reasonably well for RCP 8.5. By courtesy of Calel & Stainforth (2017) the figure also presents the corresponding temperature response of DICE 2013, PAGE 09, and FUND 3.9, the numeric IAMs used for the interagency report determining the official SCC in the US. ACE competes very well in all scenarios.

Figure 3 shows the calibration results. In addition to the original RCP scenarios, I include two scenarios available in MAGICC6.0 that initially follow a higher radiative forcing scenario and then switch over to a lower scenario (RCP 4.5 to 3 and RCP 6 to 4.5). These scenarios would be particularly hard to fit in a model tracing only atmospheric temperature. The ability to fit temperature dynamics across a peak is important for optimal policy analysis. ACE's temperature model does an excellent job in reproducing MAGICC's temperature response for the scenarios up to a radiative forcing of 6W/m^2 as in the RCP 6 scenario. It performs slightly worse for the high "business as usual" scenario RCP 8.5, but still well compared to other IAMs.[14] Transformed to the vector of generalized temperatures $\boldsymbol{\tau}_t$, the temperatures' equations of motion (11) take the linear vector form

$$\boldsymbol{\tau}_{t+1} = \boldsymbol{\sigma}\boldsymbol{\tau}_t + \sigma^{forc}\frac{M_{1,t} + G_t}{M_{pre}}\boldsymbol{e}_1 \ .$$

[14]The fact that all IAMs slightly overestimate the temperature for high carbon concentrations results from the recent findings that climate sensitivity is most likely not constant, but slowly declining in the atmospheric carbon dioxide concentration. An extension of ACE can incorporate a declining climate sensitivity without losing analytic tractability. For ease of exposition, I decided for the simpler climate system.

Further characterization and calibration. The optimal consumption rate is $x^* = 1 - \beta\kappa$. Society consumes less the higher the discounted shadow value of capital ($x_t^* = \frac{1}{1+\beta\varphi_\kappa}$ with $\varphi_k = \frac{\kappa}{1-\beta\kappa}$), resulting in a consumption rate that decreases in the capital share of output κ. The other controls depend on the precise form of the production and energy sectors, but they are not needed to determine the optimal carbon tax (see Appendix B.1 for further details).

I calibrate the remaining parts of ACE as follows (base calibration). A capital share of $\kappa = 0.3$ and the International Monetary Fund's (IMF 2020) investment rate forecast of $1 - x^* = 26\%$ pin down an annual rate of pure time preference of $\rho = 1.4\%$. Present world output Y is 10 times (time step) the IMF's global economic output forecast of $Y_{2020}^{annual} = 130$ trillion USD (purchasing power parity). The base calibration uses the carbon cycle of DICE 2013.

4 Policy Results

The SCC is the money-measured present value welfare loss from adding a ton of CO_2 to the atmosphere. The Pigovian carbon tax is the SCC along the optimal trajectory of the economy. In the present model, the SCC is independent of the future path of the economy. Therefore, this unique SCC is the optimal carbon tax.[15]

4.1 Base Scenario, Analytics, Decomposition

Appendix B solves for the shadow values and derives the optimal CO_2 tax. It is proportional to output Y_t and increases over time at the rate of economic growth as in Golosov et al. (2014). In contrast to earlier models, ACE avoids summing over future periods, emission impulse responses, or reservoirs, yielding a simpler and yet richer description of the dynamic characteristics driving the optimal carbon tax. It is the first formula pinpointing the tax contribution from carbon versus temperature dynamics.

Proposition 2 *(1) Under the assumptions of section 2, the SCC in (USD-2020-)*

[15]The present paper calculations the Pigovian tax in the first-best setting, see e.g. Barrage (2019) for an analysis of how the typical distortions in raising tax revenue modify the Pigovian tax on carbon.

money-measured consumption equivalents is

$$SCC_t = \underbrace{\frac{\beta Y_t^{net}}{M_{pre}} \underbrace{\xi_0}_{2.1\%}}_{11.5\frac{USD}{tCO_2}} \underbrace{\left[(\mathbf{1}-\beta\boldsymbol{\sigma})^{-1}\right]_{1,1}}_{1.1} \underbrace{\sigma^{forc}}_{0.54} \underbrace{\left[(\mathbf{1}-\beta\boldsymbol{\Phi})^{-1}\right]_{1,1}}_{4.3} = 30\frac{USD}{tCO_2} \tag{12}$$

where $[\cdot]_{1,1}$ *denotes the first element of the inverted matrix in square brackets, and the numbers rely on the calibration discussed in section 3.*
(2) A carbon cycle (equation 6) satisfying mass conservation of carbon implies a factor $(1-\beta)^{-1}$ *in the SCC (equation 12), which is approximately proportional to the inverse of the rate of pure time preference* $\frac{1}{\rho}$.

The ratio of world output to preindustrial carbon concentrations M_{pre} sets the units of the carbon tax. The discount factor β reflects a one-period delay between temperature increase and production impact. The damage parameter ξ_0 represents the semi-elasticity of net output with respect to a transformed temperature increase, i.e., to an increase of $\tau_1 = \exp(\xi_1 T_1)$. In the absence of climate dynamics, these terms would imply a carbon tax of $11.50\frac{USD}{tCO_2}$, or 10 cents per gallon at the pump (2 €-cents per liter). The interesting thrust making the carbon tax more serious stems from climate dynamics.

Contribution of carbon dynamics: a major multiplier. Analytic models of climate change previously captured carbon dynamics in a decay or impulse response formulation, whereas numeric IAMs incorporate a carbon cycle that models the physical carbon flows and respects that carbon does not decay. ACE integrates a carbon cycle, and a Neumann series expansion of $\beta\boldsymbol{\Phi}$ interprets the implications for the carbon tax

$$(1-\beta\boldsymbol{\Phi})^{-1} = \textstyle\sum_{i=0}^{\infty} \beta^i \boldsymbol{\Phi}^i \ . \tag{13}$$

The element $[\boldsymbol{\Phi}^i]_{1,1}$ of the transition matrix characterizes how much of the carbon injected into the atmosphere in the present remains in or returns to the atmosphere in period i, after cycling through the different carbon reservoirs. E.g., $[\boldsymbol{\Phi}^2]_{1,1} = \sum_j \Phi_{1,j}\Phi_{j,1}$ characterizes the fraction of carbon leaving the atmosphere for layers $j \in \{1,...,m\}$ in the first time step and arriving back to the atmosphere in the second time step. Thus, the term $[(\mathbf{1}-\beta\boldsymbol{\Phi})^{-1}]_{1,1}$characterizes in closed form the discounted sum of CO_2 persisting in and returning to the atmosphere in all future periods. The discount factor accounts for the delay between the act of emitting CO_2 and the resulting temperature forcing over the course of time. Quantitatively, the persistence of carbon increases the

earlier value by a factor of 4.3 and the resulting carbon tax would be almost $50\frac{USD}{tCO_2}$ or over 40 cents per gallon – ignoring temperature dynamics.

Intuitively, the carbon multiplier $\left[(\mathbf{1}-\beta\mathbf{\Phi})^{-1}\right]_{1,1}$ is a form of Keynesian multiplier. Our emissions diffuse into different "channels", but some of these feed back again into the atmosphere, thereby amplifying the emissions' effectiveness. Similarly, $\left[(\mathbf{1}-\beta\mathbf{\Phi})^{-1}\right]_{1,2}$ characterizes the long-term forcing contribution from CO_2 that is currently in the shallow ocean. The difference between these multipliers can inform back-of-the-envelope calculations for the value of different geoengeneering "solutions" to climate change that channel our CO_2 emissions into the ocean or other non-permanent reservoirs. E.g., channeling emissions into the shallow ocean would relieve the SCC by $\frac{\left[(\mathbf{1}-\beta\mathbf{\Phi})^{-1}\right]_{1,1}-\left[(\mathbf{1}-\beta\mathbf{\Phi})^{-1}\right]_{1,2}}{\left[(\mathbf{1}-\beta\mathbf{\Phi})^{-1}\right]_{1,1}} \approx 67\%$ or $20\frac{USD}{tCO_2}$ in absolute terms (ignoring damages from ocean acidification). This difference in the long-term impact of carbon in the atmosphere versus the ocean and biosphere also play an important role in evaluating the impact of uncertainty governing carbon dynamics (Traeger 2021*b*).

Contribution of temperature dynamics: a net reduction. The terms $\left[(\mathbf{1}-\beta\boldsymbol{\sigma})^{-1}\right]_{1,1}\sigma^{forc}$ capture the atmosphere-ocean temperature dynamics resulting from both delay and persistence. That is, it takes time to warm the atmosphere and oceans, but once they are warm, they conserve some of this warming. Analogously to the case of carbon, the expression $\left[(1-\beta\boldsymbol{\sigma})^{-1}\right]_{1,1}$ characterizes the generalized heat flow that enters, remains, and returns to our atmosphere. Thus, the simple closed-form expression for the carbon tax in equation (12) captures an infinite double sum: an additional ton of carbon emissions today causes radiative forcing in all future periods, and the resulting radiative forcing in any given period increases the temperature in all subsequent periods. The parameter σ^{forc} captures the atmospheric adjustment rate to radiative forcing absent ocean cooling. Ocean-atmosphere temperature dynamics reduces the carbon tax by approximately 40%, resulting in the optimal carbon tax of $30\frac{USD}{tCO_2}$ or 27 cents per gallon (6 €-cents per liter). This carbon tax is somewhat higher than Nordhaus's (2014) optimal carbon tax of $21\frac{USD}{tCO_2}$ for DICE 2013, which has a more sluggish temperature response and uses a much lower estimate of world output.[16]

[16]Golosov et al. (2014) and Gerlagh & Liski (2018*b*) use an emission response model similar to the common carbon cycle models that I adopt here. Their models do not explicitly incorporate radiative forcing, temperature dynamics, and damages as a function of temperature. However, Gerlagh & Liski (2018*b*) introduce a delay between peak emissions and peak damages, motivated by the missing temperature component. This delay multiplier contributes a factor of .45 in their closest scenario ("Nordhaus"), which cuts the tax a little more than ACE's factor of $1.4 \cdot 0.42 \approx .6$, derived from an explicit model of temperature dynamics calibrated to MAGICC6.0.

A conceptual implication for policy making. The SCC in equation (12) is independent of the atmospheric carbon concentration and of the prevailing temperature level. A corresponding independence of past emissions already prevails in Golosov et al. (2014), and contradicts the common perception that slacking on climate policy today will require more mitigation in the future. This result might sound like good news, but what the model really states is as follows. If we delay policy today, we will not compensate in our mitigation effort tomorrow, but will live with the consequences forever. Yet, the result does contain some good news for policy makers and modelers. Setting the optimal carbon tax requires minimal assumptions about future emission trajectories and mitigation technologies. The policy-maker sets an optimal price of carbon, and the economy determines the resulting optimal emission trajectory. The common intuition that the SCC ought to increase in the CO_2 concentration and the prevailing temperature level results from the convexity of damages in temperature. Yet, what common intuition overlooks is that CO_2 traps (absorbs) energy only in a certain range of wavelengths that is increasingly saturated. As a result, warming is logarithmic in the prevailing atmospheric CO_2 concentration (equation 7), and the marginal ton's warming impact is proportional to the inverse of the prevailing concentration.

4.2 Discounting and Time Preference

It is well known that the *consumption discount rate* plays a crucial role in valuing long-run impacts (Nordhaus 2007, Weitzman 2007). Finding (2) in Proposition 2 is different. It states that *the interaction of pure time preference and carbon cycle dynamics* is the main sensitivity when it comes to discounting. The proposition ties this sensitivity directly to the fact that carbon does not decay, but only cycles through the different reservoirs; a fraction of our current emissions remains in or returns to the atmosphere in the long run. In DICE 2013's carbon cycle, about 6% stay in the atmosphere forever.[17] The growth contribution to discounting is less relevant because the damages from climate change grow approximately proportionally to consumption, which offsets the growth-induced reduction in future marginal value. The contribution of temperature dynamics is less sensitive to time preference because heat is not conserved.[18] The

[17]The maximal eigenvalue of the carbon cycle's transition matrix Φ is unity and the corresponding eigenvector governs the long-run distribution (as the transitions corresponding to all other eigenvectors are damped). The first entry of the corresponding eigenvector is 0.06.

[18]Heat is constantly exchanged with outer space. Golosov et al.'s (2014) SCC formula already shows a related sensitivity to pure time preference, arising from modeling emissions as a convex combination

extreme sensitivity weakens as we depart from log utility as van den Bijgaart et al. (2016) and Rezai & van der Ploeg (2016) elaborate with approximate formulas and numeric simulations. As we start reducing the elasticity of intertemporal substitution, we clean up more of our historic sins along the optimal trajectory, and worry a little less about their long-run consequences.

The high sensitivity to pure time preference, coupled with a low sensitivity to the consumption discount rate, has an interesting policy implication. The US Circular A-4 by the Office of Management and Budget prescribes a consumption discount rate of 3%.[19] It does not give any direct guidance regarding its composition, leaving a degree of freedom to the modelers of the Interagency Working Group on the US' federal SCC that I show to have a huge impact on the SCC's magnitude. For this purpose, first, note that the SCC formula holds independently of the economy's growth rate. Second, section 5.2 develops a more sophisticated investment model showing that ACE can easily match observed consumption-investment rates with any of the present choices of time preference.

The base calibration in Section 3.3 implies a rate of pure time preference of $\rho = 1.4\%$. Instead, Drupp et al.'s (2018) recent expert survey suggests $\rho = 0.5\%$ as the median response, changing the SCC and its components to

$$SCC_t = \underbrace{\frac{\beta Y_t^{net}}{M_{pre}}}_{\text{~~11.5~~}12.5\frac{USD}{tCO_2}} \underbrace{\xi_0}_{2.1\%} \underbrace{\left[(\mathbf{1}-\beta\boldsymbol{\sigma})^{-1}\right]_{1,1}}_{\text{~~1.1~~}1.3} \underbrace{\sigma^{forc}}_{0.54} \underbrace{\left[(\mathbf{1}-\beta\boldsymbol{\Phi})^{-1}\right]_{1,1}}_{\text{~~1.3~~}8.4} = \text{~~30~~}\mathbf{72}\ \frac{USD}{tCO_2}.$$

The formula emphasizes that, under certainty, the main determinant of the optimal carbon tax is the interaction of time preference and carbon dynamics, doubling its contribution to a factor of 8. Overall, the SCC increases almost 2.5 times to $72\frac{USD}{tCO_2}$ or 63 cents per gallon at the pump ($\approx$ 14 €-cents per liter).

of decaying and non-decaying carbon. In their published revision, Gerlagh & Liski (2018*b*) explain how the non-decaying carbon box results from a carbon cycle. Finding (2), which was first published in the working paper version of the present paper (Traeger 2015), is related in spirit, but directly factors out the sensitivity factor $(1-\beta)^{-1}$ resulting from mass conservation in the carbon cycle formulation.

[19] More precisely, cost-benefit analysis is undertaken with both a 3% and a 7% discount rate where the 3% reflects consumption discounting, and the 7% pre-tax capital interest. In addition, a third rate between 1-3% can be suggested if future generations are affected. The 3% rate was employed as the base rate under the Obama administration and is likely to return with the Biden administration. For the 3% rate, the Interagency Report on the SCC found a 2020 SCC estimate of $26\frac{2007\text{-}USD}{tCO_2}$ using a cross section of models. They also report SCCs of 7 and $42\frac{2007\text{-}USD}{tCO_2}$ for consumption discount rates of 5 and 2.5%, respectively.

More sophisticated calibrations of time preference than in the present paper, using the long-run-risk model of asset pricing, deliver even lower rates while matching risk-free rate and risk premia substantially better, e.g., Bansal et al. (2012) calibrate a rate of $\rho = 0.11\%$.[20] The Stern (2007) Review argues for the rate of $\rho = 0.1\%$ for normative reasons. Similarly, low rates of pure time preference can arise if observed market equilibria are interpreted as reflecting individual life-cycle choices in the absence of bequest motives (Schneider et al. 2013) and when distinguishing consumption choice from political long-term decision making and voting (Hepburn 2006). A rate of pure time preference of $\rho = 0.1\%$ changes the SCC and its components as follows

$$SCC_t = \underbrace{\frac{\beta Y_t^{net}}{M_{pre}}}_{\cancel{11.5}\,13\,\frac{USD}{tCO_2}} \underbrace{\xi_0}_{2.1\%} \underbrace{\left[(\mathbf{1}-\beta\boldsymbol{\sigma})^{-1}\right]_{1,1}}_{\cancel{1.1}\,1.5} \underbrace{\sigma^{forc}}_{0.54} \underbrace{\left[(\mathbf{1}-\beta\boldsymbol{\Phi})^{-1}\right]_{1,1}}_{\cancel{4.3}\,26} = \cancel{\mathbf{30}}\,\mathbf{276}\ \frac{USD}{tCO_2}.$$

This ninefold increase of the SCC is, once more, mostly driven by the interaction of time preference with the carbon cycle. Temperature dynamics now only reduces the SCC by 20% as also the (moderate) temperature persistence obtains more weight relative to the short-term warming delay. The present evaluations use equation (12) without further recalibration. As a result, these evaluations would correspond to investment rates of 28.5% and 30% rather than the observation-based 26% used in the original model calibration. From a normative point of view, such a difference is warranted (but would suggest that policy should also increase other investments). In response to the Stern (2007) Review, which applies log-utility in combination with $\rho = 0.1\%$, Nordhaus (2007,2014) rejects the resulting high values of the SCC based on this descriptive inconsistency. Yet, the long-run risk model calibrates such a low time preference and explains observed investments and risk premia far better than any IAM. It would explain the high investment rates noted above as a result of the missing investment risks in these IAMs. More generally, Section 5.2 distinguishes consumption and investment sectors, casting into doubt the somewhat simplistic calibration of pure time preferences common to many IAMs. The resulting modification easily reconciles low rates of pure time preference with observed investment rates, also under certainty. The approach results in a slight modification of equation (12), which further increases the SCC to the values shown in Table 1, $75\frac{USD}{tCO_2}$ and $290\frac{USD}{tCO_2}$, respectively.

[20] Traeger (2012*a*) shows how uncertainty-based discounting by an agent whose risk aversion does not coincide with her consumption-smoothing preference (falsely) manifests itself as pure time preference in the standard economic model.

4.3 DICE's Carbon Cycle versus Impulse Response

DICE's climate dynamics has recently been criticized for scientific inaccuracy (Mattauch et al. 2020, Van der Ploeg et al. 2020). ACE's base scenario already avoids the main target of the criticism, DICE's overly sluggish temperature response. But also the carbon cycle responds a little more sluggishly when compared to scientific models (Van der Ploeg et al. 2020). Joos et al. (2013) provides a simple emulation of scientific carbon cycle models. Instead of tracking carbon through different (at times complex) reservoirs, the authors calibrate the atmospheric impulse response for a variety of carbon cycle models, including the best fit to a state of the art ensemble (CMIP 5). The approach distributes a ton of carbon emitted into 4 different "boxes". These boxes no longer correspond to physical reservoirs; instead these boxes are characterized by unique atmospheric life-times. The amount of atmospheric carbon in period t resulting from a unit of today's (t=0) emissions is

$$\Delta M_{1,t} = a_0 + \sum_{i=1}^{3} a_i \exp\left(-\frac{t}{\tau_i}\right),$$

where $\boldsymbol{a} = (a_0, a_1, a_2, a_3)^\top$ characterizes the fraction of carbon going into a particular box. The fraction a_0 stays in the atmosphere forever and the fractions a_i of today's emission unit has an e-folding time τ_i.[21] The IPCC (2013) uses this quantitative model to calculate global warming potentials.

This box model is formally equivalent to the usual carbon cycle model with the following modifications. First, carbon is released not only into the atmosphere, but into all boxes. In addition, some of the carbon simply vanishes immediately ($\sum_{i=0}^{3} a_0 < 1$). Formally, the weight vector $\boldsymbol{a}$ replaces the unit vector $\boldsymbol{e}_1$ channeling emissions (only) into the atmosphere in equation (6). As a result, the SCC will become the a_i-weighted sum of the different boxes' shadow values. Second, all boxes correspond to atmospheric carbon and, thus, contribute to radiative forcing changing equation (7) to

$$F_t = \eta \frac{\log \frac{\sum_{i=0}^{3} M_{i,t} + G_t}{M_{pre}}}{\log 2}. \tag{14}$$

Third, the transition matrix is diagonal as carbon no longer moves across boxes.

[21] The e-folding time represents the time at which only a fraction corresponding to one over the Euler's number remains in the atmosphere. The corresponding e-folding times (half-lifes) are $\tau_1 =$ 394.4 (273), $\tau_2 =$ 36.54 (25), and $\tau_3 =$ 4.304 (3) years. Note the notational difference between e-folding time τ (only used in the present subsection) and generalized temperature τ.

As a result, the inversion of the matrix in the SCC formula becomes trivial and $[(\mathbf{1}-\beta\mathbf{\Phi})^{-1}]_{i,i} = \frac{1}{1-\gamma_i\beta}$, where the coefficients $\gamma_i = \exp\big(-\frac{step}{\tau_i}\big)$ are the (diagonal) entries of the transition matrix; the first diagonal entry corresponds to $\tau_0 = \infty$ and $\gamma_0 = 1$, and $step = 10$ represents the time step. I emphasize that all boxes reflect atmospheric carbon and no longer have a direct physical counterpart; they approximate the different response times of various carbon sinks.

The resulting SCC formula becomes (see Appendix B.3)

$$SCC_t = \frac{\beta Y_t^{net}}{M_{pre}} \left[(\mathbf{1}-\beta\boldsymbol{\sigma})^{-1}\right]_{1,1} \sigma^{forc} \left(\frac{a_0}{1-\beta} + \sum_{i=1}^{3} \frac{a_i}{1-\gamma_i\beta}\right) \tag{15}$$

The only change w.r.t. the SCC formula in equation (12) is the last term, replacing $[(\mathbf{1}-\beta\mathbf{\Phi})^{-1}]_{1,1}$. Interestingly, the base scenario's SCC remains the same $30\frac{USD}{tCO_2}$ as in the base scenario using DICE 2013's carbon cycle. The carbon cycle multiplier changes from 4.31 to 4.29, and using 4 digit precision one would see a drop of the SCC by 14 cent. The exact match in the base scenario is somewhat of a coincidence. Joos et al.'s (2013) calibration suggests that $a_0 = 21\%$ of today's carbon emissions remain in the atmosphere forever, contrasting sharply with DICE 2013's 6%. As a result, the two models should exhibit different sensitivities to time preference. For a rate of pure time preference of 0.5%, the box-model's carbon multiplier increases from 8.3 to 8.9 causing a 4 USD increase of the optimal carbon tax to $77\frac{USD}{tCO_2}$. For a rate of 0.1%, the carbon multiplier increases from 26 to 30 causing a 26 USD increase of the optimal carbon tax to $312\frac{USD}{tCO_2}$. In contrast, high rates of pure time preference reduce the impulse-response-based SCC relative to the DICE 2013 carbon cycle, see e.g. scenarios 8 versus 9 in Table 1, a reduction of 6% for a rate of pure time preference of 4.3%. As is nicely observed in Van der Ploeg et al.'s (2020) simulations, DICE's "carbon-sluggishness" reduces the short-term impact, increases the medium-run impact, and underestimates the long-run impact. Low rates of pure time preference place more weight on the long-run implications and DICE slightly underestimates the SCC relative to Joos et al. (2013). High rates of time preference place more weight on the medium-term impacts and DICE yields a slightly higher SCC.[22]

The advantage of following standard IAMs in using an actual carbon cycle model is that equations (12) and (13) have a direct interpretation in terms of physical carbon

[22]I have used the DICE 2013 carbon cycle because it best matches the scientific carbon cycle models. DICE 2016 was re-calibrated to better match the long-run dynamics, but it comes at the expense of substantially overestimating the medium-run response that receives most of the weight in IAMs.

flows. The model also delivers the social cost of carbon in other sinks, like the biosphere or the oceans. These numbers can be of interest for geoengineering measures and for explicitly modeling damages such as ocean acidification (as in, e.g., the FUND model), but DICE's carbon cycle might be a bit simple to overly emphasize those values. The advantage of the impulse response model is that it permits pinpointing the individual SCC contributions of its different boxes, corresponding to the differential "life-times" of atmospheric carbon. In the base scenario, the 21% of an emitted ton of CO_2 that stay up in the atmosphere forever contribute close to 40% of the SCC, in the $\rho = 0.5\%$ scenario they contribute 50%, and in the $\rho = 0.1\%$ scenario they contribute over 70% of the total SCC. Golosov et al. (2014) adopt a formally equivalent impulse response model to describe carbon dynamics. Given their model omits temperature dynamics, it would be closest to ACE's scenario of "no temperature delay".

4.4 Damage Levels and Capital Productivity

Apart from temperature dynamics and updates in world output, ACE's base calibration leans heavily on choices resembling the DICE model. The present section discusses variations of the DICE based damage calibration and capital share. Regarding the first, DICE's damage function has repeatedly been criticized for being on the low side (yet other wide-spread IAMs like FUND suggest lower damages). Section 3.3 already introduces an alternative and much higher damage calibration based on the meta-analysis by Howard & Sterner (2017) and the survey by Pindyck (2020). It (i) remains a challenge to estimate climate impacts from weather volatility, the small intertemporal changes we have observed so far, and cross-country differences; (ii) many in the field are looking forward to the results of the Climate Impact Lab and other initiatives generating more detailed global impact estimates. First results for the global cost of increased mortality (Carleton et al. 2019) suggest that such projects likely result in values closer to the second, higher damage calibration. And (iii), while current work attempts to incorporate adaptation by current means, it will be even harder to incorporate actual future adaptation based on future technologies. That said, I simply present and compare results for the DICE calibrated damage function (*base*) as well as the calibration based on Howard & Sterner (2017) and Pindyck (2020) labeled *HSP*. The accompanying paper Traeger (2021*b*) explicitly introduces damage and damage-adaptation uncertainty.

HSP damages increase the damage semi-elasticity to an exponential temperature

increase from $\xi_0 = 0.022$ to $\xi_0 = 0.10$. Formula (12) shows that such a 4.4 times increase of the damage semi-elasticity translates straight into a 4.4 times increase of the SCC (across all scenarios). Table 1 states the results. E.g., the SCC of the base calibration ($\rho = 1.4\%$) increases to $134\frac{USD}{tCO_2}$. The SCC increases to $331\frac{USD}{tCO_2}$ for expert discounting ($\rho = 0.5\%$), and to a painstaking $1280\frac{USD}{tCO_2}$ for the low discounting scenario ($\rho = 0.1\%$) – or even to $1450\frac{USD}{tCO_2}$ in combination with the impulse response model by Joos et al. (2013). The alternative calibration of damages to Howard & Sterner (2017), labeled HSP* in Figure 2, increases all of these values by approximately 10%.

ACE's base calibration uses the common capital share of $\kappa = 0.3$, which is also part of the DICE model. However, more recent estimates suggest estimates of κ closer to 0.4 (Inklaar & Timmer 2013, Neiman 2014, Jones 2016). The parameter κ does not explicitly show in the SCC formula (12). Yet, it determines the base model's investment rate and implicitly shows up at various points. First, I used it to calibrate the rate of pure time preference by matching the model's investment rate $1 - x^* = \beta\kappa = 26\%$ to observation. Here, using $\kappa = 0.4$ substantially increases of the rate of pure time preferences from 1.4% to 4.3%. As a result, the optimal carbon tax in the (otherwise) base scenario drops to $11\frac{USD}{tCO_2}$; the main reduction stems from the carbon multiplier's decrease from 4.3 to 2.2. Second, the shadow value of atmospheric carbon in utils is proportional to $\frac{1}{1-\kappa\beta}$ and, third, its conversion to USD makes the SCC proportional to consumption, which is world output times the consumption rate $x^* = 1 - \kappa\beta$. As a result, the two latter appearances of κ cancel in the SCC formula (12). Thus, overall $\kappa = 0.4$ implies a 65% reduction of the SCC to $11\frac{USD}{tCO_2}$ in the base model, using the "base calibration method".

The reasoning of the previous paragraph takes for granted that the world's investment rate is determined by $\kappa\beta$. This simple formula is a result of assuming perfect substitutability between consumption and investment goods (and simplified capital dynamics). Section 5.2 introduces a variety of goods and distinguishes between consumption and investment goods. As a result, the consumption and investment rates are no longer governed by this simplistic formula and can change over time and space for a variety of reasons (as observed in the data). The resulting model has the degrees of freedom to separately calibrate capital productivity κ, time preference β, and the consumption rate x. Then, increasing κ no longer implies a reduction of time preference. Moreover, increasing κ increases the SCC because the higher productivity increases the cost of climate damages and the resulting foregone investments. Instead

of a reduction, the more sophisticated model yields a slight increase of the optimal carbon tax, yielding $34\frac{USD}{tCO_2}$ in response to an increase in capital share. Table 1 shows how the increase in capital affects the SCC in the scenarios using lower discount rates (scenarios 20 & 32) and higher damages (scenarios 10, 25, & 35).

5 Extensions

The first part of this section introduces population weighting within and across generations. The second part presents an example of the production economy and introduces heterogenous consumption and investment goods.

5.1 Population Weighting and Heterogenous Consumption Levels

So far, I assumed that a social planner or representative agent maximizes the log-welfare of consumption. In contrast, the DICE model uses population-weighted utility from average consumption, introducing differential weighting of generational consumption over time. In addition, climate change economics is increasingly concerned with distributional differences within a generation. The present section extends the baseline model to analyze the implications of (i) differing welfare (or Negishi) weights across subgroups of the population at a given point in time and (ii) time varying welfare weights across time as implied by DICE's population weighting. The analysis is restricted to exogenously evolving population weights.

I split the population into $P \in \mathbb{N}$ groups with idiosyncratic group sizes $p_{i,t}$, $i \in P$ and corresponding average consumption levels $c_{i,t}$. The social planner places weight α_i on each of these groups and the objective becomes

$$\max_{(c_{i,t})_{i\in P;t\in\mathbb{N}_0}} \sum_{t=0}^{\infty} \beta^t \left(\sum_{i\in P} \alpha_{i,t} \log c_{i,t} \right) \qquad \text{subject to } \sum_{i\in P} p_{i,t} c_{i,t} = C_t. \tag{16}$$

The group-specific weights $\alpha_{i,t}$ can but do not have to depend on (or coincide with) the group sizes $p_{i,t}$. Both $\alpha_{i,t}$ and $p_{i,t}$ evolve exogenously. In a descriptive world, the weights $\alpha_{i,t}$ are interpreted as Negishi-weights. A low observed equilibrium consumption of subgroup i translates into a low Negishi-weight $\alpha_{i,t}$ in the social planner representation.[23]

[23]Negishi-weights are straight-forward in the static context. Their calculation and interpretation

r.p.t.p.		Scenario	Parameter changes	Carb Mult	w/o TD USD	SCC in USD	cent per Gallon	Euro-cent per liter
1.4%	1	**base**		4.3	50	**30**	27	6
	2	Joos carbon (carb)	Φ	4.3	50	30	27	6
	3	**HSP-damages** (HSP)	ξ_0	4.3	219	**134**	118	27
x 4.3%	4	**productivity simple**	κ, β	2.2	19	**11**	9	2
	5	**productivity 5.2**	κ	4.3	56	**34**	30	7
	6	Population growth (pg)	g		63	39	34	8
x 2.3%	7	Pop growth recalib (pg-re)	g, β		40	24	21	5
x 4.3%	8	**HSP & prod simple**	ξ_0, κ, β	2.2	85	**48**	42	10
x 4.3%	9	carb & HSP & prod simple	$\Phi, \xi_0, \kappa, \beta$	2.1	78	45	39	9
	10	**HSP & prod 5.2**	ξ_0, κ	4.3	248	**152**	335	78
x 4.3%	11	prod simple & pg	κ, β, g		23	13	11	3
x 5.2%	12	prod simple & pg-re	κ, β, g		18	10	9	2
	13	prod 5.2 & pg	κ, g		72	44	39	9
x 4.3%	14	HSP & prod simple & pg	ξ_0, κ, β, g		102	57	50	12
x 5.2%	15	HSP & prod simple & pg-re	ξ_0, κ, β, g		81	45	39	9
	16	HSP & prod 5.2 & pg	ξ_0, κ, g		319	197	173	40
0.5%	17	**expert discounting**		8.4	109	**75**	66	15
	18	Joos carbon (carb)	Φ	8.9	115	79	70	16
	19	**HSP-damages** (HSP)	ξ_0	8.4	480	**331**	291	68
	20	**productivity 5.2** (p)	κ	8.4	125	**86**	76	18
	21	Population growth (pg)	g		146	100	88	21
	22	carb & HSP	Φ, ξ_0	8.9	509	351	308	72
	23	carb & p 5.2	Φ, κ	8.9	133	92	80	19
	24	carb & pg	Φ, g		156	107	94	22
	25	**carb & HSP & p5.2**	Φ, ξ_0, κ	8.9	587	**405**	356	83
	26	carb & HSP & pg	Φ, ξ_0, g		688	473	416	97
	27	carb & p 5.2 & pg	Φ, κ, g		180	124	109	25
	28	carb & HSP & p 5.2 & pg	Φ, ξ_0, κ, g		795	547	480	112
0.1%	29	low discounting		26	361	290	255	59
	30	**Joos carbon** (carb)	Φ	30	409	**328**	288	67
	31	HSP-damages (HSP)	ξ_0	26	1600	1280	1130	262
	32	productivity 5.2 (p5.2)	κ	26	421	338	296	69
	33	Population growth (pg)	g		500	400	351	82
	34	**carb & HSP**	Φ, ξ_0	30	1810	**1450**	1270	296
	35	**carb & prod 5.2**	Φ, κ	30	476	**382**	325	78
	36	carb & pg	Φ, g		567	454	398	93
	37	**carb & HSP & p** 5.2	Φ, ξ_0, κ	30	2100	**1690**	1480	345
	38	carb & HSP & pg	Φ, ξ_0, g		2510	2000	1760	410
	39	carb & p 5.2 & pg	Φ, ξ_0, g		660	528	464	108
	40	carb & HSP & p 5.2 & pg	Φ, ξ_0, κ, g		2920	2330	2050	477

Table 1: Quantitative results. "r.p.t.p."=rate of pure time preference. "Carb Mult"= Carbon-based multiplier in SCC (non-stationary for population growth). "w/o TD"= without temperature delay (cutting temperature related terms in SCC). "Joos carbon" uses the impulse response calibration of Joos et al. (2013) and "HSP-damages" uses Section 3.3's calibration of damages to Howard & Sterner (2017) and Pindyck (2020). For population growth see Section 5.1. The "productivity simple" and "population growth recalibrated" (pg-re) scenarios ("x" in the first column) recalibrate time preference. In contrast, "production 5.2" uses the extended model of Section 5.2 abandoning perfect substitutability between consumption and investment goods and not requiring recalibration.

DICE equates population size with the labor force, $N_{i,t} = \sum_{i \in P} p_{i,t}$, and uses population weighted average utility delivering the objective $\max_{(c_t)_{t \in \mathbb{N}_0}} \sum_{t=0}^{\infty} \beta^t N_t \log \frac{C_t}{N_t}$. The model extension has two alternative interpretations. Either, a social planner treats different groups of the population differently for normative reasons that lie outside of this model. Or, adopting a descriptive perspective, a set of exogenous Negishi weights capture that consumption is distributed heterogenously in the population.

First, I focus on distributional differences within a generation. For this purpose, I assume that I can normalize the welfare weights to unity, i.e., $\sum_{i \in P} \alpha_{i,t} = 1$. The weights and consumption distribution in the population can change over time, but each period obtains the same total weight. Second, I will relax this assumption and assume that the total weight grows at a constant rate, e.g., along with population size. Third, I treat the case of general non-stationary population growth leaving details to Appendix C.

Proposition 3 *Let the welfare objective in equation* (16) *replace ACE's original objective in equation* (9).
i) Let $\sum_{i \in P} \alpha_{i,t} = 1$ *for all* t. *The optimal carbon tax of Proposition 2 discussed in Section 4 remains unchanged.*
ii) Let $\alpha_t \equiv \sum_{i \in P} \alpha_{i,t}$ *grow at a constant rate with growth factor* $g = \frac{\alpha_{t+1}}{\alpha_t}$ *for all* t. *The optimal carbon tax of Proposition 2 changes to the form*

$$SCC_t^{det} = \frac{\beta g Y_t^{net}}{M_{pre}} \left[(\mathbf{1} - \beta g \boldsymbol{\sigma})^{-1}\right]_{1,1} \sigma^{forc} \left[(\mathbf{1} - \beta g \boldsymbol{\Phi})^{-1}\right]_{1,1}.$$

iii) For non-constant growth of α_t *the shadow value of atmospheric carbon and the SCC have to be calculated recursively using equations (C.5-C.7) stated in Appendix C.*

The first part of the proposition finds that *exogenous* differences in the consumption distribution have no impact on the SCC (or the optimal carbon tax). This finding fleshes out that it is not merely the existence of inequality that alters the optimal SCC. Instead, distribution becomes relevant because of *endogenous* interactions of inequality and taxation or climate change and inequality. This result, of course, hinges as well on using log-utility in line with the median expert of Drupp et al.'s (2018) survey. Under log-utility, income and substitution effects cancel. Differences in consumption will matter if income and substitution effects start to differ across agents as a result of differences in income.

is more intricate in the dynamic context, an issue I am side-stepping given my assumption of an exogenous evolution of these weights.

The second part of Proposition 3 shows that a change of the total weight over time, i.e., a change of the intergenerational weight affects the SCC. The intuition is straightforward. An increasing weight on future generations favors future generations the same way as an increase in the discount *factor* (reduction of the utility discount rate). Relating to DICE, population-weighting matters. The result is independent of whether the social planner uses absolute consumption, average consumption, or some other within-generation distribution and weighting. Placing more weight on more populous future generations reduces the effective discount rate and, thereby, increases the SCC. Increasing the welfare weight on more populous generations might be reasonable from a normative perspective (assuming exogenous population dynamics). DICE is usually interpreted as a descriptive model. As such, a tentative explanation of the weight could be that the current generation will be more altruistic towards future generations when they have more offsprings.[24] For a quantitative evaluation of population weighting, I use the United Nation's population growth scenario (see Appendix C) with a stable population after 2100. Here, DICE's population weighting increases the SCC from 30 to $39\frac{USD}{tCO_2}$. If we stick to the original calibration approach, however, we would have to recalibrate the time preference to match the observed investment rate. Then, the rate of pure time preference increases from 1.4% to 2.3% and the SCC falls to a mere $24\frac{USD}{tCO_2}$. If we consider the population weighting a normative argument, or if we follow Section 5.2's more sophisticated investment model that frees time preference and investment rate from their simplistic link, then we might not want to recalibrate time preference. Using accordingly the expert's median of $\rho = 0.5\%$ or the low discount rate of $\rho = 0.1\%$, UN population growth increases the SCC by 25 and 110 USD, respectively (see Table 1).

Finally, I note that the SCC uses the shadow value of aggregate consumption to translate the shadow value of CO_2 into consumption equivalents (and USD). The SCC is an exchange ratio between consumption and CO_2. This exchange ratio differs across the different subgroups. From the perspective of the poor, the value of CO_2 reductions is not worth as many USD.[25] Yet, CO_2 is a global public bad and heterogenous

[24]Where causality goes from more offsprings to more altruistic. See Schneider et al. (2013) for a discussion of the difficulties with a purely descriptive interpretation of intertemporal weighting in integrated assessment models using infinitely lived agents as opposed to explicitly overlapping generations, given limited observed altruism.

[25]The SCC is the *social* price of carbon. As such, it aggregates the global welfare loss of all individuals. Yet, the same aggregate welfare damage in utils translates into different consumption equivalents depending on the consumption levels of different subgroups.

pricing would always be inefficient. By using exogenous welfare (or Negishi) weights, the present model assumes a redistribution that leaves the consumption ratios across different subgroups of the economy untouched. Thus, it assumes that distributional implications of a carbon tax can be mitigated by other channels. Models doing away with the strict assumptions of the present section are a center piece of current research. The present section itself is helpful to (i) flesh out those aspects of a heterogenous population that leave the SCC untouched and (ii) explain the role and implications of population weighting in the DICE model.

5.2 CES-Preferences for Consumption, Investment Goods, & Production Example

This section introduces a variety of final goods. The representative agent will be equipped with CES preferences over final consumption goods. Given the multiplicity of goods, I also have to specify investment explicitly. I retain equation (5), but it now takes a composite of investment goods. The consumption-investment rate will no longer be time-constant and it no longer has a simple closed-form solution. In particular, it is no longer pinned down merely by capital productivity κ and discount rate β.

Let the sets $I_c \in \{1, ..., \bar{I}_c\}$ and $I_I = \{\underline{I}_I, ..., I\}$ denote the sets of final consumption goods and investment components, respectively, where $\bar{I}_c, \underline{I}_I, I \in \mathbb{N}$. These sets can coincide, overlap, or be disjoint. The representative agent consumes the share $x_{l,t}$ of good $c_{l,t}$. For goods that are only used in consumption, $l \in I_c \backslash I_I$, it is $x_{l,t} = 1$. For goods that are used in both the consumption and the investment process, $l \in I_c \cap I_I$, the consumption share $x_{l,t}$ is endogenously chosen and the remaining share $1 - x_{l,t}$ enters investment. For goods that are only used in the production-investment process, $l \in I_I \backslash I_c$, it is $x_{l,t} = 0$.

As in the base model, a representative consumer maximizes the discounted sum of the log of a consumption aggregate C_t (objective 9). But now, the consumption aggregate is a CES-aggregator over a variety of different final goods

$$C_t = \Big(\sum_{l \in I_C} a_{l,t} (x_{l,t} c_{l,t})^{s_t} \Big)^{\frac{1}{s_t}} \tag{17}$$

with good-specific weights $a_{l,t}$ and substitutability index $s_t \leq 1$ for all t and $l \in I_C$.

The final goods follow a production process of the form

$$c_{l,t} = A_{l,t} K_{l,t}^{\alpha} N_{l,t}^{1-\alpha-\nu} d_{l,t}^{\nu} \left[1 - D_t(T_{1,t})\right] \quad \text{for } l = 1, ..., I \tag{18}$$

with good-specific technology, capital, and labor as well as an energy intermediate $d_{l,t}$. I introduce the energy intermediate to represent the different substitutabilities across energy sources in different sectors, in particular, in transport versus common production processes. As compared to, e.g., Golosov et al. (2014) such a model reflects that oil and renewable energy are, at least currently, less substitutable in the transport sector than are coal and renewables in the production sector. In each sector, the energy intermediate

$$d_{l,t} = \Big(\sum_{i \in \Theta_l} e_{i,t}^{\tilde{s}_{l,t}} \Big)^{\frac{1}{\tilde{s}_{l,t}}} \tag{19}$$

is a CES-combination of different energy sources, where $\Theta_l \subset \{1, ..., I_E\}$ is the (index set of the) subset of primary energy sources used for producing good l. E.g., for $l = transport$, the substitutability $\tilde{s}_{transport,t}$ between oil-based-energy ($i = oil$) and renewable energy ($i = renewable$) is currently limited, but it can increase over time ($\tilde{s}_{transport,t}$ is period-specific). To convert a fossil fuel or renewable resource into energy we need capital and labor and I assume

$$e_{i,t} = g_{i,t}(A_{I+i,t}, K_{I+i,t}, N_{I+i,t}, E_{i,t})$$
$$\text{satisfying } \; g_{i,t}(A_{I+i,t}, \gamma K_{I+i,t}, N_{I+i,t}, E_{i,t}) = \gamma^{\tilde{\alpha}} g_{i,t}(A_{I+i,t}, K_{I+i,t}, N_{I+i,t}, E_{i,t}),$$

for all $i \in \{1, ..., I_E\}$. Some fossil fuels $E_{i,t}$ will be subject to the intertemporal scarcity constraint (2). I use a general function g to include Cobb-Douglas as well as formulations with a "bliss point", i.e., a finite emission level beyond which more coal or more oil no longer increase production – a feature satisfied in DICE. I refer to an accompanying paper for a more detailed discussion of these features and their implications for emissions (Traeger 2021*a*). Here, I analyze the implications on the optimal carbon tax. Finally, the investment composite

$$I_t = \Big(\sum_{l \in I_I} b_{l,t} c_{l,t}^{\zeta_t} \Big)^{\frac{1}{\zeta_t}} \tag{20}$$

with investment weights $b_{l,t}$ and substitutability index $\zeta_t \leq 1$ for all t and $l \in I_I$ replaces equation (4) of the base model. This economy features $I_A = I_K = I_N = I + I_E$ different sectors.

Proposition 4 *The preference and production extensions of this section replace equations* (1) *and* (4) *of the base ACE by equations (17-20). The SCC, expressed in terms of aggregate consumption equivalents, becomes*

$$SCC_t = \frac{C_t}{1-\beta\kappa}\frac{\beta}{M_{pre}}\xi_0 \left[(\mathbf{1}-\beta\boldsymbol{\sigma})^{-1}\right]_{1,1} \sigma^{forc} \left[(\mathbf{1}-\beta\boldsymbol{\Phi})^{-1}\right]_{1,1} \tag{21}$$

where $\kappa \equiv \alpha + \tilde{\alpha}\nu$.

Assuming coinciding parameter values, the value of the SCC in equation (21) coincides with that of Proposition 2. Also equation (12) of the base model implicitly contains the expression $\frac{C_t}{1-\beta\kappa}$. However, the base model assumes a single consumption-investment good. This assumption results in the simple and constant consumption rate $1-\beta\kappa$, implying that $\frac{C_t}{1-\beta\kappa} = \frac{(1-\beta\kappa)Y_t^{net}}{1-\beta\kappa} = Y_t^{net}$. In general, changes in production processes and consumption preferences shift the fraction of production that is being consumed. The present model accommodates such shifts. As a result, the SCC is no longer proportional to world output, but it remains proportional to world consumption. In general, the factor $\frac{C_t}{1-\beta\kappa}$ can be smaller or larger than world output.

The important implications of Proposition 4 govern calibration. In contrast to simplistic one-commodity models, the present specification no longer ties the consumption rate straight to pure time preference and capital productivity κ. Figure 5 in Appendix C illustrates how a degree of freedom that is alternatively generated by preferences, calibration of the different technology levels, or composition of the investment good changes the consumption rate much more than the changes of the rate of pure time preference ρ or capital productivity κ in the scenarios discussed in the present model. Thus, the present section shows that there is no contradiction between any of the discussed scenarios and observed macroeconomic consumption rates, even under certainty. In particular, the result backs up the ceteris paribus variations of the rate of pure time preference and capital productivity κ in Table 1.

6 Conclusions

ACE is a full-blown IAM with analytic solution fleshing out the qualitative and quantitative policy contributions of different IAM components. The present paper discusses the implications of temperature versus carbon dynamics, damage estimates, capital productivity, population weighting, consumption versus investment composites, and of

calibration approaches for the optimal carbon tax or SCC. Table 1 summarizes the quantitative results for a variety of calibrations that are reasonably founded in the literature.

Using a single model, a simple "base calibration" using IMF consumption and output data can generate optimal carbon tax recommendations spanning almost all carbon prices currently observed from the low $10\frac{USD}{tCO_2}$ in some cap and trade markets to the high $200\frac{USD}{tCO_2}$ energy taxes imposed on selective sectors in some European countries. The low tax values rely crucially on calibrating a high pure time preference by matching observed consumption rates under population-weighting (as in DICE) or an increased capital productivity (suggested in recent studies). A more sophisticated specification of the economy distinguishes between consumption and investment goods. As a result, commonly invoked simple relationships between pure time preference and macroeconomic consumption no longer hold; observed investment rates can be explained for almost arbitrary rates of time preference based on differences between consumption and investment goods, preferences, and/or production technologies. Recalibrating time preference based on an expert median results in a tax range of 75 to $550\frac{USD}{tCO_2}$. Taking out DICE's population weighting, which I consider slightly less convincing than the other model components, the high value falls to $400\frac{USD}{tCO_2}$. Finally, the paper presents the results under a 0.1% rate of pure time preference. The Stern (2007) Review suggests this value on normative grounds and studies of long-run risk match asset prices well using similarly low time preference, as do studies reinterpreting observed equilibria as those of finitely lived agents without a bequest motive. In this scenario, the SCC ranges from 300 to $2300\frac{USD}{tCO_2}$, or cutting out population weighting to $1700\frac{USD}{tCO_2}$.

A simple and yet under-appreciated message of the analytic insights and Table 1 is that pure time preference remains one of – or possibly *the* – most crucial determinant governing how much society should do about climate change. If we calibrate the rate of pure time preference to a high 4.2% and trust in DICE damages we can continue with a carbon price of $10\frac{USD}{tCO_2}$. If we combine this calibration with high damages estimates instead, we would ramp things up to a price closer to $50\frac{USD}{tCO_2}$ as currently attempted in some countries. If we believe in the cited expert median, or in normative arguments that the future should not be discounted merely for its futurity, then we should ramp up climate action by an additional order of magnitude. Despite the highly relevant progress in modeling climate change and estimating damages and substitutabilities,

we cannot avoid a more basic truth. If we intrinsically care for the future, we should undergo much more serious efforts to limit climate change. If we don't, "we're good" with current attempts. A common argument that low rates of pure time preference would be inconsistent with current investment patterns no longer holds under ACE's disentanglement between consumption and investment goods, as well as for a variety of other reasons cited.

ACE is the first IAM pinning down and quantifying the different multipliers resulting from the different aspects of climate dynamics. Whereas the carbon cycle multiplier of the SCC substantially increases the SCC, temperature dynamics reduces the SCC because of the warming delay. The crucial sensitivity of the SCC to time preference is rooted in its interaction with the carbon cycle. Carbon does not simply decay but only moves across different reservoirs. As a result of this persistence, the carbon multiplier ranges between 2 and 30 depending on the calibration of time preference. Population-weighting further increases this sensitivity. ACE explains how changes in capital productivity affect the SCC in a variety of ways, but most importantly through model calibration. DICE 2013's much critiqued carbon cycle only implies minor differences for the SCC when compared to a prominent science calibration. While DICE's carbon cycle is too sluggish, the inaccuracies mostly balance for the overall SCC, slightly underestimating the SCC for low rates of time preference and overestimating it for high rates.

In most aspects, ACE is far more general than any pre-existing IAM. Its main limiting assumption is a unit elasticity of intertemporal substitution. I emphasize that the unit elasticity only reflects deterministic trade-offs; I show in accompanying work that ACE solves for arbitrary coefficients of (disentangled) relative Arrow-Pratt risk aversion under uncertainty. The present extension to CES preferences slightly weakens the assumption of a unit elasticity. A unit elasticity of intertemporal substitution coincides with the median and the mode of Drupp et al.'s (2018) expert survey and was used in the **?** Review of Climate Change. The macroeconomic literature tends to find lower elasticities, whereas the long-run risk literature suggests higher elasticities. If sticking to a single value, the unit elasticity is likely a reasonable candidate. It implies that income and substitution effects cancel. As a result, I show that also exogenously changing distributional inequalities do not affect the optimal tax. Addressing the interplay of income and substitution effects under deviations from log-utility or as a result of endogenous distributional dynamics linked to climate change and taxation is

an interesting and important research agenda not addressed in the present paper.

The present paper develops the basic ACE model. An accompanying paper uses ACE to "break the curse of dimensionality" by analytically solving a complex stochastic IAM with multiple uncertainties, Epstein-Zin-Weil preferences, and an exact analytic solution to the infinite horizon stochastic fix-point problem. Another application of ACE solves dynamic strategic interactions in a full-blown regional IAM. These applications break new ground by overcoming serious numerical challenges. More generally, ACE provides analytic insights combined with their quantitative implications that constructively accompany the numeric literature on IAMs designed for policy advising.

References

Anderson, E., Brock, W., Hansen, L. P. & Sanstad, A. H. (2014), 'Robust analytical and computational explorations of coupled economic-climate models with carbon-climate response', *RDCEP Working Paper 13-05* .

Anthoff, D. & Tol, R. S. (2014), The climate framework for uncertainty, negotiation, and distribution (fund), technical description, version 3.9, Technical report, http://www.fund-model.org/.

Bansal, R., Kiku, D., Shaliastovich, I. & Yaron, A. (2014), 'Volatility, the macroeconomy, and asset prices', *The Journal of Finance* **69**(6), 2471–2511.

Bansal, R., Kiku, D. & Yaron, A. (2010), 'Long-run risks, the macro-economy and asset prices', *American Economic Review: Papers & Proceedings* **100**, 542–546.

Bansal, R., Kiku, D. & Yaron, A. (2012), 'An empirical evaluation of the long-run risks model for asset prices', *Critical Finance Review* **1**, 183–221.

Bansal, R., Kiku, D. & Yaron, A. (2016), 'Risks for the long run: Estimation with time aggregation', *Journal of Monetary Economics* **82**(C), 52–69.

Bansal, R. & Yaron, A. (2004), 'Risks for the long run: A potential resolution of asset pricing puzzles', *The Journal of Finance* **59**(4), 1481–509.

Barrage, L. (2019), 'Optimal Dynamic Carbon Taxes in a ClimateEconomy Model with Distortionary Fiscal Policy', *The Review of Economic Studies* **87**(1), 1–39.
URL: *https://doi.org/10.1093/restud/rdz055*

Bosetti, V. (2021), 'Integrated assessment models for climate change'.
URL: *https://oxfordre.com/economics/view/10.1093/acrefore/9780190625979.001.0001/acrefore-9780190625979-e-572*

Bosetti, V., Carraro, C., Galeotti, M., Massetti, E. & Tavoni, M. (2006), 'Witch: A world induced technical change hybrid model', *The Energy Journal* **27**(2), 13–38.

Brock, W. A. & Mirman, L. J. (1972), 'Optimal economic growth and uncertainty: The discounted case', *Journal of Economic Theory* **4**, 479–513.

Brock, W. & Xepapadeas, A. (2017), 'Climate change policy under polar amplification', *European Economic Review* **99**, 93 – 112. Combating Climate Change. Lessons from Macroeconomics, Political Economy and Public Finance.

Calel, R. & Stainforth, D. A. (2017), 'On the physics of three integrated assessment models', *Bulletin of the American Meteorological Society* **98**, 1199–1216.

Carleton, T., Delgado, M., Greenstone, M., Houser, T., Hsiang, S., Hultgren, A., Jina, A., Kopp, R., McCusker, K., Nath, I., Rising, J., Rode, A., Seo, H. K., Simcock, J., Viaene, A., Yuan, J. & Zhang, A. (2019), Valuing the global mortality consequences of climate change accounting for adaptation costs and benefits, Technical Report 2018-51, Becker Friedman Institute.

Chen, X., Favilukis, J. & Ludvigson, S. C. (2013), 'An estimation of economic models with recursive preferences', *Quantitative Economics* **4**, 39–83.

Collin-Dufresne, P., Johannes, M. & Lochstoer, L. A. (2016), 'Parameter Learning in General Equilibrium: The Asset Pricing Implications', *American Economic Review* **106**(3), 664–698.

Dietz, S. & Venmans, F. (2018), 'Cumulative carbon emissions and economic policy: in search of general principles', *Grantham Working Paper 283* .

Drupp, M. A., Freeman, M. C., Groom, B. & Nesje, F. (2018), 'Discounting disentangled', *American Economic Journal: Economic Policy* **10**(4), 109–134.

Economist (2017), 'The EPA is rewriting the most important number in climate economics', *The Economist, Nov 17 th* .

Emmerling, J., Drouet, L., Reis, L. A., Bevione, M., Berger, L., Bosetti, V., Carrara, S., Cian, E. D., D'Aertrycke, G. D. M., Longden, T., Malpede, M., Marangoni, G., Sferra, F., Tavoni, M., Witajewski-Baltvilks, J. & Havlik, P. (2016), 'The witch 2016 model documentation and implementation of the shared socioeconomic pathways', *FEEM Note di Lavoro* (42.2016).

Engel, C. (2016), 'Exchange rates, interest rates, and the risk premium', *American Economic Review* **106**(2), 436–74.

Feenstra, R. C., Inklaar, R. & Timmer, M. P. (2015), 'The next generation of the penn world table', *American Economic Review* **105**(10), 3150–3182. PWT 10.0, downloaded 3/05/2021 at http://www.ggdc.net/pwt/.

Gerlagh, R. & Liski, M. (2018*a*), 'Carbon prices for the next hundred years', *The Economic Journal* **128**(609), 728–757.

Gerlagh, R. & Liski, M. (2018*b*), 'Consistent climate policies', *Journal of the European Economic Association* **16**(1), 1–44.

Golosov, M., Hassler, J., Krusell, P. & Tsyvinski, A. (2014), 'Optimal taxes on fossil fuel in general equilibrium', *Econometrica* **82**(1), 41–88.

Hambel, C., Kraft, H. & Schwartz, E. S. (2018), 'Optimal carbon abatement in a stochastic equilibrium model with climate change', *SSRN working paper* **2578064**.

Hambel, C., Kraft, H. & Schwartz, E. S. (2021), 'The carbon abatement game in a stochastic equilibrium model with climate change', *European Economic Review* **132**, 1036–42.

Hardy, G., Littlewood, J. & Polya, G. (1964), *Inequalities*, 2 edn, Cambridge University Press. first puplished 1934.

Hassler, J. & Krusell, P. (2012), 'Economics and climate change: Integrated assessment in a multi-region world', *Journal of the European Economic Association* **10**(5), 974–1000.

Hassler, J., Krusell, P., Olovsson, C. & Reiter, M. (2018), 'Integrated assessment in a multi-region world with multiple energy sources and endogenous technical change', *Working Paper* .

Havránek, T. (2015), 'Measuring Intertemporal Substitution: The Importance Of Method Choices And Selective Reporting', *Journal of the European Economic Association* **13**(6), 1180–1204.

Heal, G. (1984), Interactions between economy and climate: A framework for policy design under uncertainty, *in* V. K. Smith & A. D. White, eds, 'Advances in Applied Microeconomics', Vol. 3, JAI Press, Greenwich, CT, pp. 151–168.

Hepburn, C. (2006), Discounting climate change damages: Working notes for the Stern review, Working note.

Hoel, M. & Karp, L. (2002), 'Taxes versus quotas for a stock pollutant', *Resource and Energy Economics* **24**, 367–384.

Hope, C. (2006), 'The marginal impact of CO2 from PAGE2002: An integrated assessment model incorporating the IPCC's five reasons for concern', *The Integrated Assessment Journal* **6**(1), 19–56.

Howard, P. H. & Sterner, T. (2017), 'Few and not so far between: A meta-analysis of climate damage estimates', *Environmental and Resource Economics* **68**, 197–225.

IMF (2020), 'World Economic Outlook Database. International Monetary Fund. October 2020'.

Inklaar, R. & Timmer, M. P. (2013), Capital, labor and TFP in PWT8.0, Groningen growth and development centre, University of Groningen.

Interagency Working Group on Social Cost of Carbon, U. S. G. (2013), 'Technical support document: Technical update of the social cost of carbon for regulatory impact analysis under executive order 12866', *Department of Energy* .

IPCC (2013), *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Iverson, T. & Karp, L. (2020), 'Carbon taxes and climate commitment with non-constant time preference', *The Review of economic studies* .

Jagannathan, R. & Liu, B. (2019), 'Dividend Dynamics, Learning, and Expected Stock Index Returns', *Journal of Finance* **74**(1), 401–448.

Jensen, S. & Traeger, C. (2013), 'Optimally climate sensitive policy: A comprehensive evaluation of uncertainty & learning', *Department of Agricultural and Resource Economics, UC Berkeley* .

Jones, C. I. (2016), Chapter 1 - The Facts of Economic Growth, Vol. 2 of *Handbook of Macroeconomics*, Elsevier, pp. 3–69.

Joos, F., Roth, R. & Weaver, A. J. (2013), 'Carbon dioxide and climate impulse response functions for the computation of greenhouse gas metrics: a multi-model analysis', *Atmos. Chem. Phys.* **13**, 2793–2825.

Karp, L. (2017), 'Provision of a public good with altruistic overlapping generations and many triibes', *The Economic Journal* **127**(607), 2641–2664.

Karp, L. & Zhang, J. (2006), 'Regulation with anticipated learning about environmental damages', *Journal of Environmental Economics and Management* **51**, 259–279.

Karp, L. & Zhang, J. (2012), 'Taxes versus quantities for a stock pollutant with endogenous abatement costs and asymmetric information', *Economic Theory* **49**, 371–409.

Kelly, D. L. & Kolstad, C. D. (1999), 'Bayesian learning, growth, and pollution', *Journal of Economic Dynamics and Control* **23**, 491–518.

Kung, H. & Schmid, L. (2015), 'Innovation, Growth, and Asset Prices', *Journal of Finance* **70**(3), 1001–1037.

Luderer, G., Leimbach, M., Bauer, N., Kriegler, E., Baumstark, L., Bertram, C., Giannousakis, A., Hilaire, J., Klein, D., Levesque, A., Mouratiadou, I., Pehl, M., Pietzcker, R., Piontek, F., Roming, N., Schultes, A., Schwanitz, V. J. & Strefler, J. (2021), Description of the remind model (version 1.6), Technical report, SSRN Working Paper 2697070.

Mattauch, L., Matthews, H. D., Millar, R., Rezai, A., Solomon, S. & Venmas, F. (2020), 'Steering the climate system - using inertia to lower the cost of policy', *American Economic Review* **114**, 1231–1237.

Matthews, H. D., Gillett, N. P., Stott, P. A. & Zickfeld, K. (2009), 'The proportionality of global warming to cumulative carbon emissions', *Nature* **459**(11), 829–833.

Meier, F. & Traeger, C. (2021), Solace – solar geoengineering in an analytic climate economy, Technical report, Mimeo.

Meinshausen, M., Raper, S. & Wigley, T. (2011), 'Emulating coupled atmosphere-ocean and carbon cycle models with a simpler model, magicc6 - part 1: Model description and calibration', *Atmospheric Chemistry and Physics* **11**, 1417–1456.

Metcalf, G. E. & Stock, J. H. (2017), 'Integrated assessment models and the social cost of carbon: A review and assessment of u.s. experience', *Review of Environmental Economics and Policy* **11**(1), 80–99.

Nakamura, E., Sergeyev, D. & Steinsson, J. (2017), 'Growth-rate and uncertainty shocks in consumption: Cross-country evidence', *American Economic Journal: Macroeconomics* **9**(1), 1–39.

Nakamura, E., Steinsson, J., Barro, R. & Ursua, J. (2013), 'Crises and recoveries in an empirical model of consumption disasters', *American Economic Journal: Macroeconomics* **5**(3), 35–74.

Neiman, K. (2014), 'The global decline of the labor share', *The Quarterly Journal of Economics* **129**(1), 61–103.

Newell, R. G. & Pizer, W. A. (2003), 'Regulating stock externalities under uncertainty', *Journal of Environmental Economics and Management* **45**, 416–432.

Nordhaus, W. (2008), *A Question of Balance: Economic Modeling of Global Warming*, Yale University Press, New Haven. Online preprint: A Question of Balance: Weighing the Options on Global Warming Policies.

Nordhaus, W. (2014), 'Estimates of the social cost of carbon: Concepts and results from the dice-2013r model and alternative approaches', *Journal of the Association of Environmental and Resource Economists* **1**(1), 273–312.

Nordhaus, W. D. (1994), *Managing the Global Commons: The Economics of the Greenhouse Effect*, MIT Press, Cambridge.

Nordhaus, W. D. (2007), 'A review of the Stern review on the economics of climate change', *Journal of Economic Literature* **45**(3), 686–702.

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

Nordhaus, W. & Sztorc, P. (2013), *DICE2013R:Introduction and User's Manual*, dicemodel.net.

Pindyck, R. S. (2013), 'Climate change policy: What do the models tell us?', *Journal of Economic Literature* **51**(3), 860–872.

Pindyck, R. S. (2020), 'The social cost of carbon revisited', *Journal of Environmental Economics and Management* **94(C)**, 140–160.

Rezai, A. & van der Ploeg, F. (2016), 'Intergenerational inequality aversion, growth, and the role of damages: Occam's rule for the global carbon tax', *Journal of the Association of Environmental and Resource Economists* **3**(2), 493–522.

Schneider, M., Traeger, C. P. & Winkler, R. (2013), 'Trading off generations: Infinitely lived agent versus olg', *European Economic Review* **56**, 1621–1644.

Stern, N., ed. (2007), *The Economics of Climate Change: The Stern Review*, Cambridge University Press, Cambridge.

Stokey, N. L. & Lucas, R. E. (1989), *Recursive Economic Dynamics*, Harvard University Press, Cambridge. With Edward C. Prescott.

Traeger, C. (2012*a*), 'Once upon a time preference - how rationality and risk aversion change the rationale for discounting', *CESifo Working Paper* **3793**.

Traeger, C. (2015), 'Analytic integrated assessment and uncertainty', *CESifo working paper no. 5464 & SSRN 2667972* .

Traeger, C. P. (2012*b*), 'A 4-stated dice: Quantitatively addressing uncertainty effects in climate change', *Environmental and Resource Economics* **59**, 1–37.

Traeger, C. P. (2019), 'Capturing intrinsic risk aversion', *SSRN Working Paper 3462905* .

Traeger, C. P. (2021*a*), IAMs and CO_2 emissions – an analytic discussion, Technical report, Mimeo.

Traeger, C. P. (2021*b*), Uncertainty in the analytic climate economy, Technical report, Mimeo.

van den Bijgaart, I., Gerlagh, R. & Liski, M. (2016), 'A simple formula for the social cost of carbon', *Journal of Environmental Economics and Management* **77**, 75–94.

van den Bremer & van der Ploeg (2018), 'The risk-adjusted carbon price', *OxCarre Research Paper* (203).

van der Ploeg, F. (2018), 'The safe carbon budget', *Climatic change* **147**(1), 47–59.

Van der Ploeg, R., Dietz, S., Rezai, A. & Venmans, F. (2020), Are economists getting climate dynamics right and does it matter?, Economics Series Working Papers 900, University of Oxford, Department of Economics.

Vissing-Jørgensen, A. & Attanasio, O. P. (2003), 'Stock-market participation, intertemporal substitution, and risk-aversion', *The American Economic Review* **93**(2), 383–391.

Weitzman, M. (2007), 'A review of the Stern review on the economics of climate change', *Journal of Economic Literature* **45**(3), 703–724.

Appendix

A Capital Depreciation & Model Transformation

A.1 Capital Depreciation

This section derives the capital equation of motion (5), quantifies the correction factor, and discusses the model's implication that the consumption *rate* (but not level) is unaffected by climate states. The usual capital accumulation equation, enriched by climate damages, is

$$K_{t+1} = Y_t[1 - D_t(T_{1,t})] - C_t + (1 - \delta_k)K_t \ .$$

Defining the consumption rate $x_t = \frac{C_t}{Y_t[1-D_t(T_{1,t})]}$ and recognizing that $Y_t[1 - D_t(T_{1,t})] - C_t = K_{t+1} - (1 - \delta_k)K_t$ implies[26]

$$K_{t+1} = Y_t[1 - D_t(T_{1,t})](1 - x_t) \left[1 + \frac{1 - \delta_k}{\frac{K_{t+1}}{K_t} - (1 - \delta_k)}\right] \ .$$

Defining the capital growth rate $g_{k,t} = \frac{K_{t+1}}{K_t} - 1$, I obtain the equation of motion for capital (5) stated in the main text.

I evaluate the correction factor based on the Penn World Tables (Feenstra et al. 2015). For 2019, currently the latest year of the Penn Word Tables 10.0, the global depreciation parameter is $\delta_k^{annual} = 0.0439$, so $\delta_k \approx 0.44$. Averaging capital growth over the past 10 years (using $2017USD$ values) delivers $g_k^{annual} = 0.02949$ or $g_k \approx 0.29$. The resulting correction factor is $\left[\frac{1+g_{k,t}}{\delta_k+g_{k,t}}\right] \approx 1.8$. (for a 5 year time step the correction factor would be 3.1, and for an annual time step 14). For the US, $\delta_k^{annual} = 0.046$ and $g_k^{annual} = 0.013$, resulting in a decadal correction factor of 1.9 (or a correction factor of 3.6 for a five year time step).

Treating the growth and depreciation correction in squared brackets as exogenous remains an approximation. The extension shows that the model is robust against the immediate criticism of not being able to represent the correct capital evolution and capital output ratio, and against the agent's neglecting of capital value beyond immediate next period usage. Yet, the crucial implications of the assumptions underlying equation (5) is that the investment rate is independent of the climate states. It is the

[26] The step uses $K_{t+1} = Y_t[1 - D_t(T_{1,t})] - Y_t[1 - D_t(T_{1,t})]x_t + (Y_t[1 - D_t(T_{1,t})] - C_t) \times \frac{(1-\delta_k)K_t}{Y_t[1-D_t(T_{1,t})]-C_t}$.

price to pay for an analytic solution. The remainder of this section shows that this price seems small.

Figure 4 tests ACE's result (and implicit assumption) that the optimal consumption rate is independent of the climate states. The figure depicts the optimal consumption rate generated by a recursive DICE implementation with an annual time step and, thus, an annual capital decay structure of the usual form (Traeger 2012*b*).[27] It also abandons the assumption of logarithmic utility, further stacking the cards against ACE's assumptions. The first two graphs in the figure depict the control rules for DICE-2013's $\eta = 1.45$ (inverse of the intertemporal elasticity of substitution). These two graphs state the optimal consumption rate for the years 2025 and 2205. The third graph in the figure depicts the optimal consumption rate for the lower value $\eta = 0.66$ calibrated by the long-run risk literature (Vissing-Jørgensen & Attanasio 2003, Bansal & Yaron 2004, Bansal et al. 2010, Chen et al. 2013, Bansal et al. 2012, Bansal et al. 2014, Collin-Dufresne et al. 2016, Nakamura et al. 2017)

The qualitative behavior is the same for all graphs in Figure 4. Overall, the figure shows that the optimal consumption rate is largely independent of the climate states (if the vertical axis started at zero the variation of the control rule would be invisible). At current temperature levels, the optimal consumption rate does not depend on the CO_2 concentrations. This result is in accordance with the theoretical result under ACE's assumption set. However, the optimal consumption rate increases slightly for higher temperatures. It increases by less than a percentage point from no warming to a 3C warming at low CO_2 concentrations. The increase is lower at higher CO_2 concentrations.

The graphs confirm that also in DICE, and in a model with regular annual capital decay structure and not exactly log-utility, the investment rate is not used as a primary measure of climate change policy. The rate does not respond to the CO_2 concentration, which is a measure of expected warming. Only once the temperature dependent damages set in, the consumption rate slightly increases and the investment rate goes down. Instead of reflecting climate policy, this (minor) climate dependence of the consumption rate reflects a response to the damages incurred: these damages reduce the cake to be split into investment and consumption, then, a slightly higher fraction goes

[27] The recursive implementation based on the Bellman equation solves for the optimal control rule as a function of the states derlivering the full control surface depicted here. This recursive implementation has a slightly simplified climate change model compared to the original DICE model, but matches the Maggic6.0 model, used also as the DICE benchmark, similarly well.

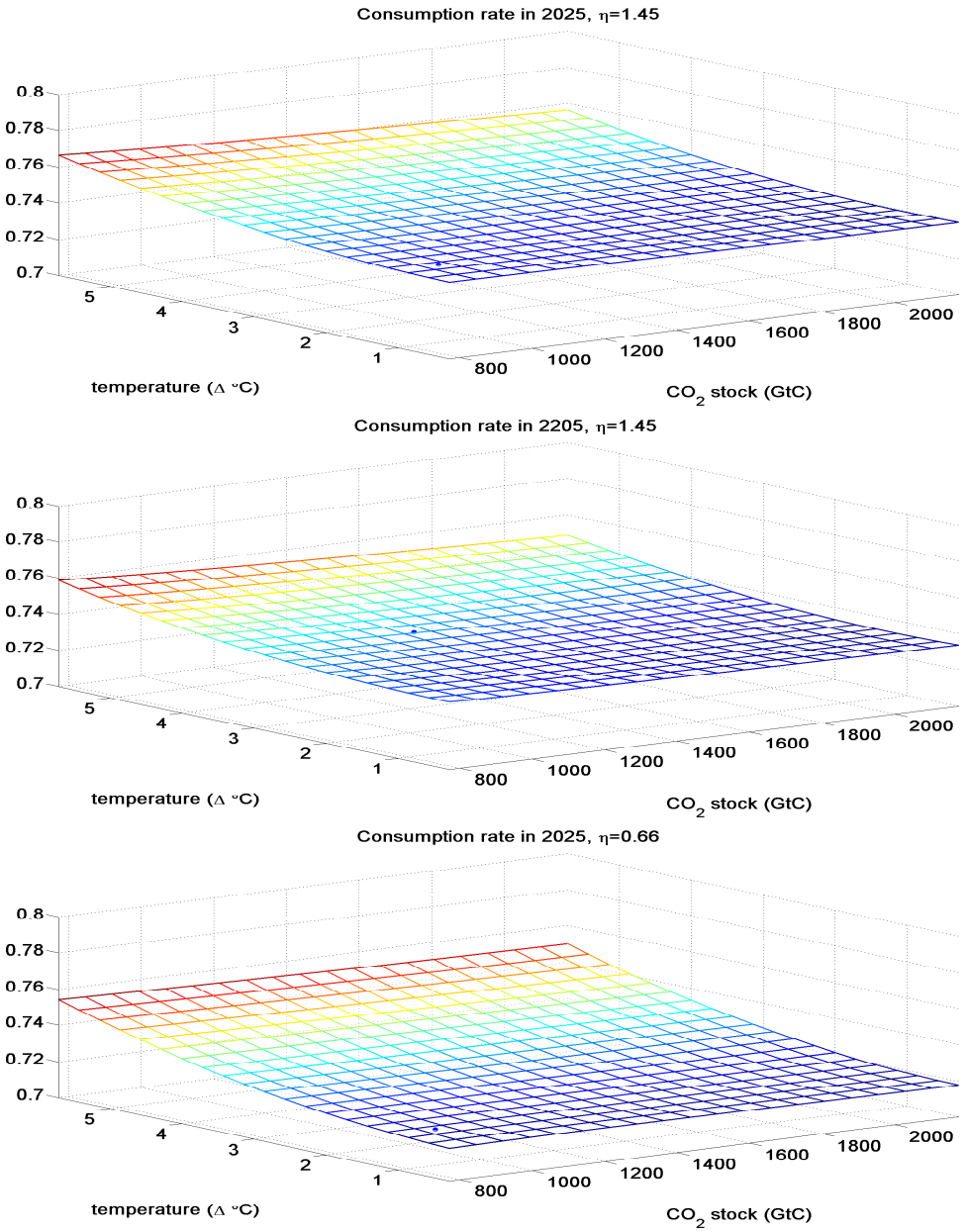

Figure 4: The graphs analyze the climate (in-)dependence of the optimal consumption rate x^* in the wide-spread DICE model, relying on the control rules of the recursive implementation by Traeger (2012*b*). The first two graphs assume the DICE-2013 value $\eta = 1.45$, the third graph follows the long-run risk literature with $\eta = \frac{2}{3}$. The blue dot in each graph indicates the expected optimal control and prevailing temperature-CO_2 combination along the optimal policy path in the given year.

to consumption. This response is lower when CO_2 concentrations are high: then the social planner expects high temperatures and damages also in the future and is more

hesitant to reduce investment.

A.2 Transformation to linearity in states

For notational convenience, I introduce the normalized vector $\boldsymbol{\mathcal{K}}_t \equiv \frac{\boldsymbol{K}_t}{K_t}$ characterizing the distribution of capital over sectors whose components satisfy $\sum_{i=1}^{I_K} \mathcal{K}_{i,t} = 1$. To obtain the equivalent linear-in-state-system, I replace aggregate capital $K_t = \sum_{i=1}^{I_K} K_{i,t}$ by logarithmic capital $k_t \equiv \log K_t$. I replace temperature levels in the atmosphere and the different ocean layers by the transformed exponential temperature states $\tau_{i,t} \equiv \exp(\xi_i T_{i,t})$, $i \in \{1, ..., L\}$. I collect these transformed temperature states in the vector $\boldsymbol{\tau}_t \in \mathbb{R}^L$. Finally, I use the consumption rate $x_t = \frac{C_t}{Y_t[1-D_t(T_{1,t})]}$, rather than absolute consumption, as the consumption-investment control. Only the rate will be separable from the system's states. Homogeneity of the production function implies that

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

Then, welfare as a function of the consumption rate is

$$\begin{aligned} u(x_t) &\equiv \log C_t = \log x_t + \log Y_t + \log[1 - D_t(T_{1,t})] \\ &= \log x_t + \kappa \log K_t + \log F(\boldsymbol{A}_t, \boldsymbol{N}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{E}_t) - \xi_0 \tau_{1,t} + \xi_0. \end{aligned}$$

The Bellman equation in terms of the transformed state variables is

$$\begin{aligned} V(k_t, \boldsymbol{\tau}_t, \boldsymbol{M}_t, \boldsymbol{R}_t, t) = \max_{x_t, \boldsymbol{N_t}, \boldsymbol{\mathcal{K}}_t, \boldsymbol{E_t}} & \log x_t + \kappa k_t + \log F(\boldsymbol{A}_t, \boldsymbol{N}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{E}_t) \\ & -\xi_0 \tau_{1,t} + \xi_0 + \beta V(k_{t+1}, \boldsymbol{\tau}_{t+1}, \boldsymbol{M}_{t+1}, \boldsymbol{R}_{t+1}, t+1) \ , \end{aligned} \tag{A.1}$$

and is subject to the following linear equations of motion and constraints. The equations of motion for the effective capital stock and the carbon cycle are

$$\begin{aligned} k_{t+1} \quad &= \kappa k_t + \log F(\boldsymbol{A}_t, \boldsymbol{N}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{E}_t) - \xi_0 \tau_{1,t} + \xi_0 + \log(1 - x_t) \\ &\quad + \log[1 + g_{k,t}] - \log[\delta_k + g_{k,t}] \end{aligned} \tag{A.2}$$

$$\boldsymbol{M_{t+1}} = \boldsymbol{\Phi} \boldsymbol{M}_t + \left(\sum_{i=1}^{I^d} E_{i,t} + E_t^{exo} \right) \boldsymbol{e}_1 \ . \tag{A.3}$$

Using equation (11), I transform the temperature's equation of motion (8) for layer layer $i \in \{1, ..., l\}$ to

$$T_{i,t+1} = \frac{1}{\xi_1} \log \Big((1 - \sigma_{i,i-1} - \sigma_{i,i+1}) \exp[\xi_1 T_{i,t}] + \sigma_{i,i-1} \exp[\xi_1 T_{i-1,t}] + \sigma_{i,i+1} \exp[\xi_1 T_{i+1,t}] \Big) .$$

Using the definitions $\sigma_{ii} = 1 - \sigma_{i,i-1} - \sigma_{i,i+1}$ and $\tau_{i,t} = \exp(\xi_1 T_{i,t})$ I find

$$\begin{aligned} \exp(\xi_1 T_{i,t+1}) &= \sigma_{i,i-1} \exp[\xi_1 T_{i,t}] + \sigma_{i,i-1} \exp[\xi_1 T_{i-1,t}] + \sigma_{i,i+1} \exp[\xi_1 T_{i+1,t}] \\ \Rightarrow \tau_{i,t+1} \quad &= \sigma_{i,i} \tau_{i,t} + \sigma_{i,i-1} \tau_{i-1,t} + \sigma_{i,i+1} \tau_{i+1,t}, i \in \{2, ..., l\}, \end{aligned}$$

still using $\sigma_{l,l+1} = 0$ for notational convenience (see footnote 9). Noting that

$$\exp[\xi_1 T_{0,t}] = \exp\left[\xi_1 \frac{s}{\eta} F_t\right] = \exp\left[\xi_1 \frac{s}{\log 2} \log \frac{M_{1,t} + G_t}{M_{pre}}\right] = \frac{M_{1,t} + G_t}{M_{pre}},$$

the equation for atmospheric temperature ($i = 1$) becomes

$$\tau_{1,t+1} = \sigma_{1,1} \tau_{1,t} + \sigma_{1,0} \frac{M_{1,t} + G_t}{M_{pre}} + \sigma_{1,2} \tau_{2,t} .$$

Note that the linearity in $M_{1,t}$ requires $\xi_1 = \frac{\log 2}{s}$ as stated in the proposition. Then, using the definition $\sigma^{forc} = \sigma_{1,0}$, the individual equations of motion for generalized temperature can be collected into the vector equation

$$\boldsymbol{\tau}_{t+1} = \boldsymbol{\sigma} \boldsymbol{\tau}_t + \sigma^{forc} \frac{M_{1,t} + G_t}{M_{pre}} \boldsymbol{e}_1 . \tag{A.4}$$

Finally, the equation of motion for the resource stock is

$$\boldsymbol{R}_{t+1} = \boldsymbol{R}_t - \boldsymbol{E}_t^d. \tag{A.5}$$

The underlying constraints within periods are

$$\textstyle\sum_{i=0}^{I} N_{i,t} = 1,\ N_{i,t} \geq 0,\ \sum_{i=1}^{I_K} \mathcal{K}_{i,t} = 1,\ \mathcal{K}_{i,t} \geq 0,\ \boldsymbol{R}_t \geq 0,$$

and the initial states are given. The present paper assumes that the optimal labor and capital allocation across sectors has an interior solution and that the scarce resources are stretched over the infinite time horizon along the optimal path, avoiding boundary value complications.

A.3 Illustrating a Two Layer Carbon Cycle

In the simple and insightful case of two carbon reservoirs the carbon cycle's transition matrix is $\mathbf{\Phi} = \left(\begin{smallmatrix} 1-\delta_{\text{Atm}\to\text{Ocean}} & \delta_{\text{Ocean}\to\text{Atm}} \\ \delta_{\text{Atm}\to\text{Ocean}} & 1-\delta_{\text{Ocean}\to\text{Atm}} \end{smallmatrix}\right)$, where e.g. $\delta_{\text{Atm}\to\text{Ocean}}$ characterizes the fraction of carbon in the atmosphere transitioning into the ocean in a given time step. The conservation of carbon implies that both columns add to unity: carbon that does not leave a layer ($\delta_{\cdot\to\cdot}$) stays ($1-\delta_{\cdot\to\cdot}$). The shadow value becomes

$$\varphi_{M,1} = \beta\varphi_{\tau,1}\sigma^{forc}M_{pre}^{-1}(1-\beta)^{-1}\left[1+\beta\frac{\delta_{\text{Atm}\to\text{Ocean}}}{1-\beta(1-\delta_{\text{Ocean}\to\text{Atm}})}\right]^{-1}.$$

The shadow value becomes less negative if more carbon flows from the atmosphere into the ocean (higher $\delta_{\text{Atm}\to\text{Ocean}}$). It becomes more negative for a higher persistence of carbon in the ocean (higher $1-\delta_{\text{Ocean}\to\text{Atm}}$). These impacts on the SCC are straight forward: the carbon in the ocean is the "good carbon" that does not contribute to the greenhouse effect. In round brackets, the root $(1-\beta)^{-1}$ noted in Proposition 2.2 makes the expression so sensitive to a low rate of pure time preference.

A common approximation of atmospheric carbon dynamics is the equation of motion of the early DICE 1994 model. Here, carbon in excess of preindustrial levels decays as in $M_{1,t+1} = M_{pre} + (1-\delta_{decay})(M_{1,t} - M_{pre})$. The shadow value formula becomes

$$\varphi_{M,1} = \beta\varphi_{\tau,1}\sigma^{forc}M_{pre}^{-1}\big(1-\beta(1-\delta_{decay})\big)^{-1},$$

which misses the long-run carbon impact and the SCC's sensitivity to pure time preference.

A.4 Illustrating a Two Layer Atmosphere-Ocean Temperature System

The two layer example of atmosphere-ocean temperature dynamics has the transition matrix $\boldsymbol{\sigma} = \left(\begin{smallmatrix} 1-\sigma_1^{up}-\sigma_1^{down} & \sigma_1^{down} \\ \sigma_2^{up} & 1-\sigma_2^{up} \end{smallmatrix}\right)$. The corresponding term of the SCC (equation 12) takes the form

$$\left[(1-\beta\boldsymbol{\sigma})^{-1}\right]_{11} = \left(1-\beta\underbrace{(1-\sigma_1^{down}-\sigma_1^{up})}_{\text{persistence in atmosphere}} - \frac{\beta^2\sigma_1^{down}\sigma_1^{up}}{1-\beta\underbrace{(1-\sigma_2^{up})}_{\text{pers. in ocean}}}\right)^{-1}.$$

Persistence of the warming in the atmosphere or in the oceans increases the shadow cost. Persistence of warming in the oceans increases the SCC proportional to the

terms σ_1^{down} routing the warming into the oceans and σ_1^{up} routing the warming back from the oceans into the atmosphere. The discount factor β accompanies each of these transition coefficients as each of them causes a one period delay. Taking the limit of $\beta \to 1$ confirms that (an analogue to) Proposition 2(2) does not hold for the temperature system

$$\lim_{\beta \to 1} \varphi_{\tau,1} = -\xi_0(1+\varphi_k)\left[(1-\boldsymbol{\sigma})^{-1}\right]_{11} = -\frac{\xi_0(1+\varphi_k)}{\sigma_1^{up}} \neq \infty. \tag{A.6}$$

As the discount rate approaches zero, the transient temperature dynamics characterized by σ_1^{down} and σ_2^{up} becomes irrelevant for evaluation, and only the weight σ_1^{up} reducing the warming persistence below unity contributes.[28]

Extending on the "missing time preference sensitivity" in the general case, note that temperature is an intensive variable: it does not scale proportional to mass or volume (as is the case for the extensive variable carbon). The columns of the matrix $\boldsymbol{\sigma}$ do not sum to unity. As a consequence of the mean structure in equation (8), however, the rows in the ocean layers' transition matrix sum to unity. The first row determining next period's atmospheric temperature sums to a value smaller than unity: it "misses" the weight that the mean places on radiative forcing. The "missing weight" is a consequence of the permanent energy exchange with outer space. Radiative forcing characterizes a flow equilibrium of incoming and outgoing radiation.

B Proofs and Calculation for "Base ACE" (Sections 3 & 4)

B.1 Proof of Proposition 1

1) Sufficiency: I show that the affine value function

$$V(k_t, \boldsymbol{\tau}_t, \boldsymbol{M}_t, \boldsymbol{R}_t, t) = \varphi_k k_t + \boldsymbol{\varphi}_M^\top \boldsymbol{M}_t + \boldsymbol{\varphi}_\tau^\top \boldsymbol{\tau}_t + \boldsymbol{\varphi}_{R,t}^\top \boldsymbol{R}_t + \varphi_t \tag{B.1}$$

[28] I note that the carbon cycle lacks the reduction in persistence deriving from the forcing weight σ_1^{up}. With this observation, equation (A.6) gives another illustration of the impact of mass conservation in the case of carbon: "$\sigma_1^{up} \to 0 \Rightarrow \varphi_{\tau,1} \to \infty$". Note that in the SCC formula σ_1^{up} cancels, as it simultaneously increases the impact of a carbon change on temperature. This exact cancellation (in the limit $\beta \to 1$) is a consequence of the weights σ_1^{up} on forcing and $1-\sigma_1^{up}$ on atmospheric temperature summing to unity. The result that a warming pulse has a transitional impact and an emission pulse has a permanent impact on the system is independent of the fact that these weights sum to unity.

solves the linear-in-state system corresponding to the equations of sections 2.1 and 2.2 with the function form assumptions presented in Proposition 1. Appendix A.2 transformed these assumptions into the linear-in-state-system summarized by equations (A.1-A.5), which I take as point of departure. Note that the coefficient on the resource stock has to be time-dependent: the shadow value of the exhaustible resource increases (endogenously) over time following the Hotelling rule.

The controls in the equations of motion (A.2)-(A.5) are additively separated from the states. Therefore, maximizing the right hand side of the resulting Bellman equation delivers optimal control rules that are independent of the state variables. These controls are functions of the shadow values, but independent of the states. Solving the Bellman equation then amounts to a set of coefficient matching conditions determining the shadow values.

In detail, inserting the value function's trial solution (equation B.1) and the next period states (equations A.2-A.5) into the (deterministic) Bellmann equation (A.1) delivers

$$\begin{aligned}
\varphi_k k_t + \boldsymbol{\varphi}_M^\top \boldsymbol{M}_t + \boldsymbol{\varphi}_\tau^\top \boldsymbol{\tau}_t + \boldsymbol{\varphi}_{R,t}^\top \boldsymbol{R}_t + \varphi_t = \max_{x_t, \boldsymbol{N_t}, \boldsymbol{\mathcal{K}}_t, \boldsymbol{E_t}} \; & \log x_t + \beta\varphi_k \log(1-x_t) \\
& + (1+\beta\varphi_k)\kappa k_t + (1+\beta\varphi_k)\log F(\boldsymbol{A}_t, \boldsymbol{N}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{E}_t) \\
& - (1+\beta\varphi_k)\xi_0\tau_{1,t} + (1+\beta\varphi_k)\xi_0 + \lambda_t^N \big(1 - \textstyle\sum_{i=1}^{I_N} N_{i,t}\big) \\
& + \beta\varphi_k(\log[1+g_{k,t}] - \log[\delta_k + g_{k,t}]) + \lambda_t^{\mathcal{K}} \big(1 - \textstyle\sum_{i=1}^{I_K} \mathcal{K}_{i,t}\big) \\
& + \beta\boldsymbol{\varphi}_{R,t+1}^\top \left(\boldsymbol{R}_t - \boldsymbol{E}_t^d\right) + \beta\varphi_{t+1} \\
& + \beta\boldsymbol{\varphi}_M^\top \left(\boldsymbol{\Phi}\boldsymbol{M}_t + \big(\textstyle\sum_{i=1}^{I^d} E_{i,t} + E_t^{exo}\big)\boldsymbol{e}_1\right) \\
& + \beta\boldsymbol{\varphi}_\tau^\top \left(\boldsymbol{\sigma}\boldsymbol{\tau}_t + \sigma^{forc}\frac{M_{1,t}+G_t}{M_{pre}}\boldsymbol{e}_1\right)
\end{aligned} \tag{B.2}$$

Maximizing the right hand side of the Bellman equation over the consumption rate yields

$$\frac{1}{x} - \beta\varphi_k \frac{1}{1-x} = 0 \quad \Rightarrow \quad x^* = \frac{1}{1+\beta\varphi_k} \; . \tag{B.3}$$

The optimal labor, capital, and resource inputs depend on the precise assumptions governing production and energy sector, i.e., the specification of $F(\boldsymbol{A}_t, \boldsymbol{N}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{E}_t)$. For a well-defined energy system, I obtain unique solutions for these optimal inputs as functions of the technology levels, shadow values, and current states. In detail, the

first order conditions for the capital shares deliver

$$
\begin{aligned}
&(1+\beta\varphi_k)\frac{\frac{\partial F(\boldsymbol{A}_t,\boldsymbol{N}_t,\boldsymbol{\mathcal{K}}_t,\boldsymbol{E}_t)}{\partial \mathcal{K}_{i,t}}}{F(\boldsymbol{A}_t,\boldsymbol{N}_t,\boldsymbol{\mathcal{K}}_t,\boldsymbol{E}_t)} = \lambda_t^{\mathcal{K}} \\
\Leftrightarrow\ & \mathcal{K}_{i,t} = \frac{1}{\lambda_t^{\mathcal{K}}}(1+\beta\varphi_k)\sigma_{Y,\mathcal{K}_i}(\boldsymbol{A}_t,\boldsymbol{N}_t,\boldsymbol{\mathcal{K}}_t,\boldsymbol{E}_t) \\
\Rightarrow\ & \lambda_t^{\mathcal{K}} = \sum_{i=1}^{I_K}(1+\beta\varphi_k)\sigma_{Y,\mathcal{K}_i}(\boldsymbol{A}_t,\boldsymbol{N}_t,\boldsymbol{\mathcal{K}}_t,\boldsymbol{E}_t) \\
\Rightarrow\ & \mathcal{K}_{i,t} = \frac{\sigma_{Y,\mathcal{K}_i}(\boldsymbol{A}_t,\boldsymbol{N}_t,\boldsymbol{\mathcal{K}}_t,\boldsymbol{E}_t)}{\sum_{i=1}^{I_K}\sigma_{Y,\mathcal{K}_i}(\boldsymbol{A}_t,\boldsymbol{N}_t,\boldsymbol{\mathcal{K}}_t,\boldsymbol{E}_t)},
\end{aligned}
$$

which is an explicit equation only in the case of constant elasticities $\sigma_{Y,\mathcal{K}_i}(\boldsymbol{A}_t,\boldsymbol{N}_t,\boldsymbol{\mathcal{K}}_t,\boldsymbol{E}_t) \equiv \frac{\partial F(\boldsymbol{A}_t,\boldsymbol{N}_t,\boldsymbol{\mathcal{K}}_t,\boldsymbol{E}_t)}{\partial \mathcal{K}_{i,t}}\frac{\mathcal{K}_{i,t}}{F(\boldsymbol{A}_t,\boldsymbol{N}_t,\boldsymbol{\mathcal{K}}_t,\boldsymbol{E}_t)}$, and an implicit equation that has to be solved together with the other first order conditions otherwise. Analogously, the first order conditions for the labor input deliver

$$
\begin{aligned}
&(1+\beta\varphi_k)\frac{\frac{\partial F(\boldsymbol{A}_t,\boldsymbol{N}_t,\boldsymbol{\mathcal{K}}_t,\boldsymbol{E}_t)}{\partial N_{i,t}}}{F(\boldsymbol{A}_t,\boldsymbol{N}_t,\boldsymbol{\mathcal{K}}_t,\boldsymbol{E}_t)} = \lambda_t^{N} \\
\Rightarrow\ & N_{i,t} = \frac{\sigma_{Y,N_i}(\boldsymbol{A}_t,\boldsymbol{N}_t,\boldsymbol{\mathcal{K}}_t,\boldsymbol{E}_t)}{\sum_{i=1}^{I_N}\sigma_{Y,N_i}(\boldsymbol{A}_t,\boldsymbol{N}_t,\boldsymbol{\mathcal{K}}_t,\boldsymbol{E}_t)},
\end{aligned}
$$

with elasticities $\sigma_{Y,N_i}(\boldsymbol{A}_t,\boldsymbol{N}_t,\boldsymbol{\mathcal{K}}_t,\boldsymbol{E}_t) \equiv \frac{\partial F(\boldsymbol{A}_t,\boldsymbol{N}_t,\boldsymbol{\mathcal{K}}_t,\boldsymbol{E}_t)}{\partial N_{i,t}}\frac{N_{i,t}}{F(\boldsymbol{A}_t,\boldsymbol{N}_t,\boldsymbol{\mathcal{K}}_t,\boldsymbol{E}_t)}$. The first order conditions for a scarce (fossil) resource input are

$$
\begin{aligned}
&(1+\beta\varphi_k)\frac{\frac{\partial F(\boldsymbol{A}_t,\boldsymbol{N}_t,\boldsymbol{\mathcal{K}}_t,\boldsymbol{E}_t)}{\partial E_{i,t}}}{F(\boldsymbol{A}_t,\boldsymbol{N}_t,\boldsymbol{\mathcal{K}}_t,\boldsymbol{E}_t)} = \beta(\varphi_{R,i,t+1}-\varphi_{M,1}) \\
\Leftrightarrow\ & E_{i,t} = \frac{(1+\beta\varphi_k)\sigma_{Y,E_i}(\boldsymbol{A}_t,\boldsymbol{N}_t,\boldsymbol{\mathcal{K}}_t,\boldsymbol{E}_t)}{\beta(\varphi_{R,i,t+1}-\varphi_{M,1})}
\end{aligned}
$$

with elasticities $\sigma_{Y,E_i}(\boldsymbol{A}_t,\boldsymbol{N}_t,\boldsymbol{\mathcal{K}}_t,\boldsymbol{E}_t) \equiv \frac{\partial F(\boldsymbol{A}_t,\boldsymbol{N}_t,\boldsymbol{\mathcal{K}}_t,\boldsymbol{E}_t)}{\partial E_{i,t}}\frac{E_{i,t}}{F(\boldsymbol{A}_t,\boldsymbol{N}_t,\boldsymbol{\mathcal{K}}_t,\boldsymbol{E}_t)}$. The first order conditions for a non-scarce resource input are analogous but without the shadow cost term $\beta\varphi_{R,i,t+1}$.

Solving the (potentially simultaneous) system of first order conditions I obtain the optimal controls $\boldsymbol{N}_t^*(\boldsymbol{A}_t\,,\varphi_k,\boldsymbol{\varphi}_M,\boldsymbol{\varphi}_{R,t+1})$, $\boldsymbol{\mathcal{K}}_t^*(\boldsymbol{A}_t\,,\varphi_k,\boldsymbol{\varphi}_M,\boldsymbol{\varphi}_{R,t+1})$, and $\boldsymbol{E}_t^*(\boldsymbol{A}_t\,,\varphi_k,\boldsymbol{\varphi}_M,\boldsymbol{\varphi}_{R,t+1})$. I will suppress the detailed dependencies below for notational convenience. Knowing these solutions is crucial for determining the precise output path and energy transition under a given policy regime. However, the SCC and, thus,

the carbon tax depend only on the structure and optimization of the controls but not on their quantification.

Inserting the (general) control rules into the maximized Bellman equation and collecting terms that depend on state variables on the left hand side delivers

$$\begin{aligned}
&\big(\boldsymbol{\varphi}_M^\top - \beta\boldsymbol{\varphi}_M^\top\boldsymbol{\Phi} - \beta\varphi_{\tau,1}\frac{\sigma^{forc}}{M_{pre}}\boldsymbol{e}_1^\top\big)\boldsymbol{M}_t + \big(\boldsymbol{\varphi}_\tau^\top - \beta\boldsymbol{\varphi}_\tau^\top\boldsymbol{\sigma} + (1+\beta\varphi_k)\xi_0\boldsymbol{e}_1^\top\big)\boldsymbol{\tau}_t \\
&\big(\varphi_k - (1+\beta\varphi_k)\kappa\big)k_t + \big(\boldsymbol{\varphi}_{R,t}^\top - \beta\boldsymbol{\varphi}_{R,t+1}^\top\big)\boldsymbol{R}_t \\
&+\varphi_t \;=\; \beta\varphi_{t+1} \\
&\quad +\;\log x_t^*(\varphi_k) + \beta\varphi_k\log(1-x_t^*(\varphi_k)) + (N_t+\beta\varphi_k)\xi_0 \\
&\quad + (1+\beta\varphi_k)\kappa k_t + (1+\beta\varphi_k)\log F(\boldsymbol{A}_t, \boldsymbol{N}_t^*, \boldsymbol{\mathcal{K}}_t^*, \boldsymbol{E}_t^*) \\
&\quad + \beta\varphi_k(\log[1+g_{k,t}] - \log[\delta_k + g_{k,t}]) - \beta\boldsymbol{\varphi}_{R,t+1}^\top\boldsymbol{E}_t^{d*} \\
&\quad +\beta\varphi_{M,1}\left(\textstyle\sum_{i=1}^{I^d} E_{i,t}^* + E_t^{exo}\right) + \beta\varphi_{\tau,1}\frac{\sigma^{forc}}{M^{pre}}G_t.
\end{aligned} \tag{B.4}$$

The equality holds for all levels of the state variables if and only if the coefficients in front of the state variables vanish, and the evolution of φ_t satisfies the state independent part of the equation. Setting the states' coefficients to zero yields

$$\varphi_k - (1+\beta\varphi_k)\kappa = 0 \qquad \Rightarrow \varphi_k = \frac{\kappa}{1-\beta\kappa} \tag{B.5}$$

$$\boldsymbol{\varphi}_M^\top - \beta\boldsymbol{\varphi}_M^\top\boldsymbol{\Phi} - \beta\varphi_{\tau,1}\frac{\sigma^{forc}}{M_{pre}}\boldsymbol{e}_1^\top = 0 \Rightarrow \boldsymbol{\varphi}_M^\top = \frac{\beta\varphi_{\tau,1}\sigma^{forc}}{M_{pre}}\boldsymbol{e}_1^\top(\mathbf{1}-\beta\boldsymbol{\Phi})^{-1} \tag{B.6}$$

$$\boldsymbol{\varphi}_\tau^\top + (1+\beta\varphi_k)\xi_0\boldsymbol{e}_1^\top - \beta\boldsymbol{\varphi}_\tau^\top\boldsymbol{\sigma} = 0 \quad \Rightarrow \boldsymbol{\varphi}_\tau = -\xi_0(1+\beta\varphi_k)\boldsymbol{e}_1^\top(1-\beta\boldsymbol{\sigma})^{-1} \tag{B.7}$$

$$\boldsymbol{\varphi}_{R,t}^\top - \beta\boldsymbol{\varphi}_{R,t+1}^\top = 0 \qquad \Rightarrow \boldsymbol{\varphi}_{R,t} = \beta^{-t}\boldsymbol{\varphi}_{R,0}\ . \tag{B.8}$$

The initial values $\boldsymbol{\varphi}_{R,0}^\top$ of the scarce resources depend on the precise evolution of the economy and, thus, depends on assumptions about production and the energy sector. Using the shadow value of log capital in equation (B.3) results in the optimal consumption rate $x^* = 1-\beta\kappa$. Then equation equation (B.4) turns into the condition

$$\varphi_t - \beta\varphi_{t+1} = B(\cdot) + \beta\varphi_{M,1}\left(\textstyle\sum_{i=1}^{I^d} E_{i,t}^* + E_t^{exo}\right) + \beta\varphi_{\tau,1}\frac{\sigma^{forc}}{M^{pre}}G_t. \tag{B.9}$$

This condition will be satisfied by picking the sequence $\varphi_0, \varphi_1, \varphi_2, ...$ Equation (B.9) does not pin down the initial value φ_0. 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 ensuring that the value function is normalized just as the infinite sum of optimized utility (Stokey & Lucas 1989, chapter

4.1). Yet, optimal policy does not dependent on the sequence $\varphi_0, \varphi_1, \varphi_2, \varphi_3, \ldots$.

2) Necessity: The affine value function solves the system if and only if it is linear-in-state. I have to show that no other transformation of capital or temperature, no other damage function, and no other non-linear mean can achieve the linear-in-state transformation of the equations in sections 2.1 and 2.2. I take as common knowledge that only the log-transformation of capital will solve the system with an affine value function.

To obtain a linear-in-state structure, generalized atmospheric temperature has to be linear in atmospheric carbon. By assumption, temperature evolves as a generalized mean:

$$\mathfrak{M}_i(T_{i-1,t}\ ,\ T_{i,t}\ ,\ T_{i+1,t}) = f^{-1}[\sigma_{i,i-1}f(T_{i-1,t}) + \sigma_{i,i}f(T_{i,t}) + \sigma_{i,i+1}f(T_{i+1,t})]$$

and atmospheric equilibrium temperature for a given forcing is

$$T_{0,t} = \frac{s}{\eta}F_t = \frac{s}{\log 2}\log\frac{M_{1,t}+G_t}{M_{pre}}\ ,$$

which is logarithmic in the atmospheric carbon stock. The equation of motion of atmospheric temperature $T_{1,t}$ is therefore

$$\begin{aligned} &T_{1,t+1} = \mathfrak{M}_1(T_{0,t}\ ,\ T_{1,t}\ ,\ T_{2,t}) = f^{-1}[\sigma_{1,0}f(T_{0,t}) + \sigma_{1,1}f(T_{1,t}) + \sigma_{1,2}f(T_{2,t})] \\ &\Leftrightarrow f(T_{1,t+1}) = \sigma_{1,0}f\left(\frac{s}{\log 2}\log\frac{M_{1,t}+G_t}{M_{pre}}\right) + \sigma_{1,1}f(T_{1,t}) + \sigma_{1,2}f(T_{2,t}). \end{aligned} \tag{B.10}$$

First, equation (B.10) implies that $f(T_{1,t})$ and $f(T_{2,t})$ have to be linear to permit a linear-in-state interaction between generalized atmospheric and upper ocean temperature (atmospheric temperature appears on both left and right side of the equality). Second, equation (B.10) implies that $f\left(\frac{s}{\log 2}\log\frac{M_{1,t}+G_t}{M_{pre}}\right)$ has to be linear in $M_{1,t}$ to permit a linear-in-state interaction between generalized atmospheric temperature and atmospheric carbon. Thus, $f(z) = \exp\left(\frac{\log 2}{s}z\right)$ up to positive affine transformation. Yet, positive affine transformations of f leave the generalized mean unchanged as they simply cancel with the inverse (Hardy et al. 1964). Note that this step fixes both the functional form of f and the parameter $\xi_1 = \log\frac{2}{s}$.[29] Consequently, the generalized

[29]The earlier working paper version uses a slightly generalized version of the generalized mean $\mathfrak{M}_1(\cdot)$ permitting additional degrees of freedom (Traeger 2015). However, additional quality of the fit achieved with these additional weight did not warrant the complications in the presentation.

temperature state delivering a linear-in-state dynamics and a linear contribution to the value function has to be $\tau_{i,t} = \exp(\xi_1 T_{i,t})$ for $i \in \{1, 2\}$. It follows inductively from

$$f(T_{i,t+1}) = \sigma_{i,i-1} f(T_{i-1,t}) + \sigma_{i,i} f(T_{i,t}) + \sigma_{i,i+1} f(T_{i+1,t})$$

for $i = 2, ..., l-1$ that $\tau_{i,t} = \exp(\xi_1 T_{i,t})$ has to hold for all $i \in \{1, ..., l\}$, up to affine transformations with a joint multiplicative constant.

Finally, I show that damages have to be of the form stated in equation (10). Taking the logarithm of the capital's equation of motion (5) delivers

$$\log K_{t+1} = \log Y_t + \log[1 - D_t(T_{1,t})] + \log(1 - x_t) + \log\left[\frac{1 + g_{k,t}}{\delta_k + g_{k,t}}\right] ,$$

where $\log Y_t$ is linear in the state $k_t = \log K_t$. To render the system linear in the states, at any time t, there have to exist two constants $c_1, c_2 \in \mathbb{R}$ such that

$$\log[1 - D_t(T_{1,t})] = c_1 \tau_{1,t} + c_2 = c_1 \exp(\xi_1 T_{1,t}) + c_2$$
$$\Rightarrow D_t(T_{1,t}) = 1 - \exp(c_1 \exp(\xi_1 T_{1,t}) + c_2).$$

Moreover, $c_1 = -c_2 \equiv \xi_0 \in \mathbb{R}$ follows from the requirement that damages are zero at $T_{1,t} = 0$.

B.2 Proof of Proposition 2

Proof of Part (1): The SCC is the negative of the shadow value of atmospheric carbon expressed in money-measured consumption units. Inserting equation (B.5) for the shadow value of log-capital and (B.7) for the shadow value of atmospheric temperature (first entry of the vector) into equation (B.6) characterizing the shadow value of carbon in the different reservoirs delivers

$$\boldsymbol{\varphi}_M^\top = -\xi_0 \left(1 + \beta \frac{\kappa}{1 - \beta\kappa}\right) \left[(1 - \beta\boldsymbol{\sigma})^{-1}\right]_{1,1} \frac{\beta \sigma^{forc}}{M_{pre}} \boldsymbol{e}_1^\top (\mathbf{1} - \beta\boldsymbol{\Phi})^{-1} .$$

The expression characterizes the social cost in terms of welfare units. This marginal welfare cost translates into a consumption change as follows: $du_t = \frac{1}{C_t} dC_t = \frac{1}{x^* Y_t^{net}} dC_t \Rightarrow dC_t = (1 - \beta\kappa) Y_t^{net} du_t$. Thus, observing that $\left(1 + \beta \frac{\kappa}{1-\beta\kappa}\right) = \frac{1}{1-\beta\kappa}$, the SCC in consumption units is

$$SCC = -(1 - \beta\kappa) Y_t^{net} \frac{\varphi_{M,1}}{N_t} = Y_t^{net} \, \xi_0 \left[(1 - \beta\boldsymbol{\sigma})^{-1}\right]_{1,1} \frac{\beta \sigma^{forc}}{M_{pre}} \left[(\mathbf{1} - \beta\boldsymbol{\Phi})^{-1}\right]_{1,1} .$$

Proof of Part (2): Mass conservation of carbon implies that the columns of Φ add to unity. In consequence, the vector with unit entry in all dimensions is a left and, thus, right eigenvector. The corresponding eigenvalue is one and the determinant of $\mathbf{1} - \beta\mathbf{\Phi}$ has the root $1 - \beta$. It follows from Cramer's rule (or as an application of the Cayley-Hamilton theorem) that the entries of the matrix $(\mathbf{1} - \beta\mathbf{\Phi})^{-1}$ are proportional to $(1 - \beta)^{-1}$.

B.3 The impulse response model

The impulse response model in section 4.3 changes equation (7) to equation (14) as discussed in the main text. As a result equation (B.6) turns into

$$\boldsymbol{\varphi}_M^\top = \frac{\beta\varphi_{\tau,1}\sigma^{forc}}{M_{pre}}\boldsymbol{e}^\top(\mathbf{1} - \beta\mathbf{\Phi})^{-1}$$

where $\boldsymbol{e}^\top = (1, 1, 1, 1)^\top$ replaces the first unit vector $\boldsymbol{e}_1^\top$. The matrix $\mathbf{1} - \beta\mathbf{\Phi}$ is diagonal with entries $1 - \beta\gamma_i$, $i \in \{0, 1, 2, 3\}$ rendering the inversion trivial. The shadow cost in utils of a unit of emissions is the a_i-weighted combination of the shadow cost's of the different boxes, i.e., $-\boldsymbol{\varphi}_M^\top\boldsymbol{a}$. The conversion of this shadow cost into consumption equivalents works as above in part (1) of the proof of Proposition 2, delivering equation (15).

C Proofs and Calculations for Extended ACE (Section 5)

C.1 General Remarks on Population Change

Population change is a special case of Proposition 3 that I prove below. The Proposition states the analytic closed-form result obtained under the assumption of a constant population growth rate. The quantitative results in Table 1 rely on non-constant population growth. They use the UN population growth scenario delivering decadal growth factors 1.0967, 1.0761, 1.0583, 1.0428, 1.0303, 1.0205, 1.0127, 1.0061. I assume population to be stationary after 2100, which is in line with the corresponding UN data that almost converges by its end year 2100. I first calculate the shadow value of atmospheric carbon using the constant population growth solution (with zero growth) for

2100 (see proposition). Then, I recursively calculate the present shadow value using the equations (C.5-C.7) derived below towards the end of the proof.

No temperature lag under population change. The scenarios omitting temperature lag under a non-stationary population change adjust equations (C.6) and (C.7) as follows. The absence of temperature dynamics or delay eliminates the term containing the generalized heat transition matrix $\boldsymbol{\sigma}$ in equation (C.6). As a result the temperature's shadow value is directly determined by the damage coefficient and the shadow value of log-capital. It also eliminates the parameter σ^{forc} from equation (C.7). Again, the quantitative solution for the UN growth scenario first calculates the stationary shadow values post 2100 (merely omitting σ^{forc} and $\boldsymbol{e}_1^\top(1-\beta\boldsymbol{\sigma})^{-1}$) and then recursively calculates the present shadow values using equation (C.5) and the modified versions of equations (C.6) and (C.7) discussed above.

C.2 Proof of Proposition 3.

The welfare objective (16) replaces the term $\log C_t$ on the r.h.s. of the (untransformed) Bellman equation by the term

$$\textstyle\sum_{i\in P}\alpha_{i,t}\log c_{i,t}+\lambda_{p,t}\Big(C_t-\sum_{i\in P}p_{i,t}c_{i,t}\Big), \tag{C.1}$$

using the Lagrange multiplier $\lambda_{p,t}$. ACE now maximizes over all $c_{i,t}$, $i\in P$, as well as aggregate consumption C_t. Group i's consumption levels $c_{i,t}$ only appears in the Bellman equation in the terms stated in equation (C.1). Thus, the first order condition for group i's consumption is

$$\frac{\alpha_{i,t}}{c_{i,t}}=\lambda_{p,t}p_{i,t}\quad\Rightarrow\quad c_{i,t}=\frac{\alpha_{i,t}}{p_{i,t}}\lambda_{p,t}^{-1}.$$

Thus, the consumption constraint yields

$$C_t=\sum_{i\in P}p_{i,t}c_{i,t}=\sum_{i\in P}p_{i,t}\frac{\alpha_{i,t}}{p_{i,t}}\lambda_{p,t}^{-1}=\lambda_{p,t}^{-1}\sum_{i\in P}\alpha_{i,t}=\lambda_{p,t}^{-1}\alpha_t\quad\Rightarrow\quad\lambda_{p,t}=C_t^{-1}\alpha_t.$$

The FOC for aggregate consumption replaces the earlier marginal utility C_t^{-1} with the shadow value of the consumption constraint $\lambda_{p,t}=C_t^{-1}\alpha_t$. Thus, in part i), where $\alpha_t=1$ for all t, the Bellman equation remains unaltered and so does the solution for the SCC.

For part ii), the FOC for aggregate consumption picks up another constant and I have to recalculate the FOC for the aggregate consumption rate. After optimizing

the individual consumption levels $c_{i,t}$, the equations above imply that period t's wefare contribution is

$$\begin{aligned}\sum_{i\in P}\alpha_{i,t}\log c_{i,t} &= \sum_{i\in P}\alpha_{i,t}\log\left(\frac{\alpha_{i,t}}{p_{i,t}}\lambda_{p,t}^{-1}\right) = \sum_{i\in P}\alpha_{i,t}\log\left(\frac{\alpha_{i,t}}{\alpha_t p_{i,t}}C_t\right)\\ &= \alpha_t\log C_t + \underbrace{\sum_{i\in P}\alpha_{i,t}\left[\log\frac{\alpha_{i,t}}{\alpha_t} - \log p_{i,t}\right]}_{\equiv\bar{\alpha}_t},\end{aligned}$$

where $\bar{\alpha}_t$ is an exogenous additive constant. The crucial change to the Bellman equation derives from the constant α_t multiplying aggregate consumption. After expressing the objective in terms of the aggregate consumption rate, Bellman equation (A.1) turns into

$$\begin{aligned}V(k_t,\boldsymbol{\tau}_t,\boldsymbol{M}_t,\boldsymbol{R}_t,t) = &\max_{x_t,\boldsymbol{N_t},\boldsymbol{\mathcal{K}}_t,\boldsymbol{E_t}} \alpha_{\mathrm{t}}(\log x_t + \kappa k_t + \log F(\boldsymbol{A}_t,\boldsymbol{N}_t,\boldsymbol{\mathcal{K}}_t,\boldsymbol{E}_t))\\ &+\alpha_{\mathrm{t}}(-\xi_0\tau_{1,t}+\xi_0) + \bar{\alpha}_{\mathrm{t}} + \beta V(k_{t+1},\boldsymbol{\tau}_{t+1},\boldsymbol{M}_{t+1},\boldsymbol{R}_{t+1},t+1)\ . \end{aligned}\tag{C.2}$$

From here, there are two ways forward. One can use the original trial solution for the value function permitting the shadow values to change over time. More elegantly, at least for the case of a constant growth rate, take the trial solution

$$V(k_t,\boldsymbol{\tau}_t,\boldsymbol{M}_t,\boldsymbol{R}_t,t) = \alpha_t\varphi_k k_t + \alpha_t\boldsymbol{\varphi}_M^\top\boldsymbol{M}_t + \alpha_t\boldsymbol{\varphi}_\tau^\top\boldsymbol{\tau}_t + \alpha_t\boldsymbol{\varphi}_{R,t}^\top\boldsymbol{R}_t + \varphi_t.\tag{C.3}$$

Plugging this trial-solution into (C.2) and dividing by α_t delivers

$$\begin{aligned}&\varphi_k k_t + \boldsymbol{\varphi}_M^\top\boldsymbol{M}_t + \boldsymbol{\varphi}_\tau^\top\boldsymbol{\tau}_t + \boldsymbol{\varphi}_{R,t}^\top\boldsymbol{R}_t + \varphi_t\\ &= \max_{x_t,\boldsymbol{N_t},\boldsymbol{\mathcal{K}}_t,\boldsymbol{E_t}} \log x_t + \kappa k_t + \log F(\boldsymbol{A}_t,\boldsymbol{N}_t,\boldsymbol{\mathcal{K}}_t,\boldsymbol{E}_t) + -\xi_0\tau_{1,t} + \xi_0 + \frac{\bar{\alpha}_t}{\alpha_t}\\ &\quad +\beta\frac{\alpha_{t+1}}{\alpha_t}\left(\varphi_k k_{t+1} + \boldsymbol{\varphi}_M^\top\boldsymbol{M}_{t+1} + \boldsymbol{\varphi}_\tau^\top\boldsymbol{\tau}_{t+1} + \boldsymbol{\varphi}_{R,t+1}^\top\boldsymbol{R}_{t+1} + \varphi_{t+1}\right)\ .\end{aligned}\tag{C.4}$$

Denoting the, by assumption, constant growth factor of the intergenerational weight α_t by $g = \frac{\alpha_{t+1}}{\alpha_t}$, equation (C.4) corresponds to the original dynamic programming problem with discount factor βg replacing the original discount factor β. The constant $\frac{\bar{\alpha}_t}{\alpha_t}$ affects the absolute welfare level in utils, however, it has no impact on the shadow values (see derivation of the original shadow values). Thus, we obtain the same shadow value formula for $\varphi_{M,1}$ in utils as before. Yet, our new value function (C.3) multiplies this $\varphi_{M,1}$ with the intergenerational weight α_t. At the same time, we have derived the

shadow value of aggregate consumption as $\lambda_{p,t} = C_t^{-1}\alpha_t$, and the conversion factor from utils to consumption changes to $dC_t = C_t\alpha_t^{-1}dW_t$. The novel α_t^{-1} in the conversion factor cancels the multiplier α_t in front of the shadow value $\varphi_{M,1}$ and, therefore, the only change in the SCC formula remains $\beta \to \beta g$.

In the *general case with non-constant populations growth*, the shadow values in equation (C.4) have to pick up time indices

$$\begin{aligned}
&\varphi_{k,t}k_t + \boldsymbol{\varphi}_{M,t}^\top \boldsymbol{M}_t + \boldsymbol{\varphi}_{\tau,t}^\top \boldsymbol{\tau}_t + \boldsymbol{\varphi}_{R,t}^\top \boldsymbol{R}_t + \varphi_t \\
&\quad = \max_{x_t, \boldsymbol{N_t}, \boldsymbol{\mathcal{K}}_t, \boldsymbol{E_t}} \log x_t + \kappa k_t + \log F(\boldsymbol{A}_t, \boldsymbol{N}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{E}_t) + -\xi_0\tau_{1,t} + \xi_0 + \frac{\bar{\alpha}_t}{\alpha_t} \\
&\qquad +\beta\frac{\alpha_{t+1}}{\alpha_t}\left(\varphi_{k,t+1}k_{t+1} + \boldsymbol{\varphi}_{M,t}^\top \boldsymbol{M}_{t+1} + \boldsymbol{\varphi}_{\tau,t+1}^\top \boldsymbol{\tau}_{t+1} + \boldsymbol{\varphi}_{R,t+1}^\top \boldsymbol{R}_{t+1} + \varphi_{t+1}\right) .
\end{aligned}$$

Letting $g_t = \frac{\alpha_{t+1}}{\alpha_t}$, and plugging in the expressions for the next period states as in equation (B.2) delivers the first order condition for the consumption rate

$$\frac{1}{x_t} = \beta\varphi_{k,t+1}\frac{g_t}{1-x_t} \quad \Rightarrow \quad x_t^* = \frac{1}{1+\beta g_t\varphi_{k,t+1}} .$$

Also the other controls remain independent of the states. Next, I have to insert the controls into the maximized Bellman equation and collect terms that depend on the state variables on the left as in equation (B.4). All that changes w.r.t. the earlier derivation is that we distinguish present from next period shadow values, which are multiplied by the growth factor g_t

$$\begin{aligned}
&\left(\boldsymbol{\varphi}_{M,t}^\top - \beta g_t\boldsymbol{\varphi}_{M,t+1}^\top\boldsymbol{\Phi} - \beta g_t\varphi_{\tau,1,t+1}\frac{\sigma^{forc}}{M_{pre}}\boldsymbol{e}_1^\top\right)\boldsymbol{M}_t \\
&+\left(\boldsymbol{\varphi}_{\tau,t}^\top - \beta g_t\boldsymbol{\varphi}_{\tau,t+1}^\top\boldsymbol{\sigma} + (1+\beta g_t\varphi_{k,t+1})\xi_0\boldsymbol{e}_1^\top\right)\boldsymbol{\tau}_t \\
&+\left(\varphi_{k,t} - (1+\beta g_t\varphi_{k,t+1}\kappa)k_t + \left(\boldsymbol{\varphi}_{R,t}^\top - \beta g_t\boldsymbol{\varphi}_{R,t+1}^\top\right)\right)\boldsymbol{R}_t + \varphi_t = ...
\end{aligned}$$

The resulting recursion equations for the shadow values that eliminate the state coef-

ficients are

$$\varphi_{k,t} - (1 + \beta g_t \varphi_{k,t+1})\kappa = 0 \qquad \Rightarrow \varphi_{k,t} = \kappa + \beta g_t \kappa \varphi_{k,t+1} \tag{C.5}$$

$$\boldsymbol{\varphi}_{\tau,t}^{\top} + (1 + \beta g_t \varphi_{k,t+1})\xi_0 \boldsymbol{e}_1^{\top} - \beta g_t \boldsymbol{\varphi}_{\tau,t+1}^{\top} \boldsymbol{\sigma} = 0$$

$$\Rightarrow \boldsymbol{\varphi}_{\tau,t}^{\top} = \beta g_t \boldsymbol{\varphi}_{\tau,t+1}^{\top} \boldsymbol{\sigma} - (1 + \beta g_t \varphi_{k,t+1})\xi_0 \boldsymbol{e}_1^{\top} \tag{C.6}$$

$$\boldsymbol{\varphi}_{M,t}^{\top} - \beta g_t \boldsymbol{\varphi}_{M,t+1}^{\top} \boldsymbol{\Phi} - \beta g_t \varphi_{\tau,1,t+1} \frac{\sigma^{forc}}{M_{pre}} \boldsymbol{e}_1^{\top} = 0$$

$$\Rightarrow \boldsymbol{\varphi}_{M,t}^{\top} = \frac{\beta g_t \varphi_{\tau,1,t+1} \sigma^{forc}}{M_{pre}} \boldsymbol{e}_1^{\top} + \beta g_t \boldsymbol{\varphi}_{M,t+1}^{\top} \boldsymbol{\Phi}. \tag{C.7}$$

Once populations stabilizes we are in a stationary state where our original solution holds. From that stationary state, equations (C.5-C.7) deliver the recursion to calculate the present shadow values, first solving (C.5), then (C.6), and then (C.7).

As earlier, the SCC converts the shadow value of atmospheric carbon back into consumption units using the relation derived above, $dC_t = C_t \alpha_t^{-1} dW_t$. The shadow value of atmospheric carbon is now $\alpha_t \varphi_{M,1,t}$ because of the differing trial solution. Thus, once again the α_t cancels and the conversion of the shadow value into consumption units works equivalently to the cases without population weighting.

C.3 Proof of Proposition 4.

First, observe that the production processes of the final goods $c_{l,t}$ are all homogenous of degree $\kappa \equiv \alpha + \tilde{\alpha}\nu$ in capital. As a result, both C_t and I_t are also homogenous of degree κ in capital. Moreover, I can pull the production damage factor $[1 - D_t(T_{1,t})]$ from equation (18) through the CES aggregators in equations (17) and (20). Thus,

$$\begin{aligned} C_t &= \left(\sum_{l \in I_c} a_{l,t} \left(x_{l,t} A_{l,t} K_{l,t}^{\alpha} N_{l,t}^{1-\alpha-\nu} \left(\Big(\sum_{i \in \Theta_l} (g_{i,t}(A_{i,t}, K_{i,t}, N_{i,t}, E_{i,t}))^{\tilde{s}_{l,t}} \Big)^{\frac{1}{\tilde{s}_{l,t}}} \right)^{\nu} \right)^{s_t} \right)^{\frac{1}{s_t}} \\ &\quad \times [1 - D_t(T_{1,t})] \\ &= \underbrace{\left(\sum_{l \in I_c} a_{l,t} \left(x_{l,t} A_{l,t} \mathcal{K}_{l,t}^{\alpha} N_{l,t}^{1-\alpha-\nu} \left(\Big(\sum_{i \in \Theta_l} (g_{i,t}(A_{i,t}, \mathcal{K}_{i,t}, N_{i,t}, E_{i,t}))^{\tilde{s}_{l,t}} \Big)^{\frac{1}{\tilde{s}_{l,t}}} \right)^{\nu} \right)^{s_t} \right)^{\frac{1}{s_t}}}_{\equiv \Omega_{C,t}(\boldsymbol{A}_t, \boldsymbol{N}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{E}_t)} \\ &\quad \times [1 - D_t(T_{1,t})] K_t^{\alpha + \tilde{\alpha}\nu} = \Omega_{C,t}(\boldsymbol{x}_t, \boldsymbol{A}_t, \boldsymbol{N}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{E}_t)[1 - D_t(T_{1,t})] K_t^{\alpha + \tilde{\alpha}\nu} \end{aligned}$$

and analogously $I_t = \Omega_{I,t}(\boldsymbol{x}_t, \boldsymbol{A}_t, \boldsymbol{N}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{E}_t)[1 - D_t(T_{1,t})]K_t^{\alpha+\tilde{\alpha}\nu}$. I find

$$\begin{aligned}\log C_t &= \log\left(\Omega_{C,t}(\boldsymbol{x}_t, \boldsymbol{A}_t, \boldsymbol{N}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{E}_t)\right) + \log[1 - D_t(T_{1,t})] + (\alpha + \tilde{\alpha}\nu)\log K_t \\ &= \log\left(\Omega_{C,t}(\boldsymbol{x}_t, \boldsymbol{A}_t, \boldsymbol{N}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{E}_t)\right) - \xi_0\tau_{1,t} + \xi_0 + \kappa k_t,\end{aligned}$$

replacing the terms $\log x_t + \log F(\boldsymbol{A}_t, \boldsymbol{N}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{E}_t)$ in equation (A.1) by the term $\log\left(\Omega_{C,t}(\boldsymbol{x}_t, \boldsymbol{A}_t, \boldsymbol{N}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{E}_t)\right)$. Similarly, the equation of motion for log-capital (A.2) changes into

$$\begin{aligned}k_{t+1} = \kappa k_t &+ \log\left(\Omega_{I,t}(\boldsymbol{x}_t, \boldsymbol{A}_t, \boldsymbol{N}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{E}_t)\right) - \xi_0\tau_{1,t} + \xi_0 + \log[1 + g_{k,t}] \\ &- \log[\delta_k + g_{k,t}]\end{aligned}$$

replacing the terms $\log(1 - x_t) + \log F(\boldsymbol{A}_t, \boldsymbol{N}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{E}_t)$ in equation (A.1) by the term $\log\left(\Omega_{I,t}(\boldsymbol{x}_t, \boldsymbol{A}_t, \boldsymbol{N}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{E}_t)\right)$. The maximization on the r.h.s. Bellman equation (A.1) now entails the labor, energy, and capital distribution within and across consumption and investment production processes. Instead of optimizing w.r.t. a single consumption rate, the maximization is now over all the consumption rates $x_{l,t}$ of those goods that can be used in consumption and in investment, i.e., where $l \in I_C \cap I_I$. Let $\hat{\boldsymbol{x}}_t$ denote the vector of endogenously chosen consumption rates, i.e., the vector containing all $x_{l,t}$ with $l \in I_C \cap I_I$. The r.h.s. of the Bellman equation (B.2) on page 48 now reads

$$\begin{aligned}\max_{\hat{\boldsymbol{x}}_t, \boldsymbol{N_t}, \boldsymbol{\mathcal{K}}_t, \boldsymbol{E_t}} \; &\log\left(\Omega_{C,t}(\boldsymbol{x}_t, \boldsymbol{A}_t, \boldsymbol{N}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{E}_t)\right) + \beta\varphi_k \log\left(\Omega_{I,t}(\boldsymbol{x}_t, \boldsymbol{A}_t, \boldsymbol{N}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{E}_t)\right) \\ &+\lambda_t^N\big(\alpha_t - \textstyle\sum_{i=1}^{I_N} N_{i,t}\big) + \lambda_t^{\mathcal{K}}\big(1 - \textstyle\sum_{i=1}^{I_K} \mathcal{K}_{i,t}\big) - \beta\boldsymbol{\varphi}_{R,t+1}^{\top}\boldsymbol{E}_t^d \\ &+\beta\varphi_{M,1}\big(\textstyle\sum_{i=1}^{I^d} E_{i,t} + E_t^{exo}\big) + ...\end{aligned}$$

where I omit terms that are independent of the controls.

As in the original problem, the solutions to the maximization problems of the r.h.s. Bellman equation are independent of the states because the equation additively separates the terms containing the controls from the terms containing the states. As a result, the FOCs and their solutions change, but the shadow values remain the same as in the original problem (equations B.5-B.8). The change to the SCC formula results from dealing with a consumption variety in the consumption-investment trade-off and the conversion of the shadow price of atmospheric carbon into consumption equivalents.

Expressing the shadow price of atmospheric carbon in equivalents of the consumption aggregate C_t, I still have the conversion factor $du_t = \frac{1}{C_t}dC_t \Rightarrow dC_t = C_t du_t$.

However, there no longer exists a simple constant aggregate consumption rate – formerly $x^* = (1-\beta\kappa)$ – such that $C_t = x^* Y_t^{net} = (1-\beta\kappa) Y_t^{net}$ would hold, the factor of which had canceled the shadow value's term $1+\beta\varphi_k = 1+\beta\frac{\kappa}{1-\beta\kappa} = \frac{1}{1-\beta\kappa}$ in the base model. Thus, instead of Y_t^{net} I have the term $\frac{C_t}{1-\beta\kappa}$. I note that the consumption rate of the model is

$$x_t \equiv \frac{\Omega_{C,t}(\boldsymbol{x}_t, \boldsymbol{A}_t, \boldsymbol{N}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{E}_t)}{\Omega_{C,t}(\boldsymbol{x}_t, \boldsymbol{A}_t, \boldsymbol{N}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{E}_t) + \Omega_{I,t}(\boldsymbol{x}_t, \boldsymbol{A}_t, \boldsymbol{N}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{E}_t)},$$

which can be calibrated in many different ways, generally differs from $(1-\beta\kappa)$, and generally changes over time.

C.4 Illustration of the degree of freedom in calibrating the consumption rate.

A simple example illustrates that distinguishing between consumption and investment goods frees the consumption rate from a simplistic relationship between time preference and capital productivity. This finding does not rely on the complexity of the climate economy model, but merely on the distinction between a consumption aggregate

$$C_t = \left(a_{1,t}\left(A_{1,t}\,\mathcal{K}_{1,t}^{\alpha}\right)^{s_t} + a_{2,t}\left(A_{2,t}\,\mathcal{K}_{2,t}^{\alpha}\right)^{s_t}\right)^{\frac{1}{s_t}}$$

and an investment aggregate

$$I_t = \left(a_{3,t}\left(A_{3,t}\,\mathcal{K}_{3,t}^{\alpha}\right)^{\zeta_t} + a_{4,t}\left(A_{4,t}\,\mathcal{K}_{4,t}^{\alpha}\right)^{\zeta_t}\right)^{\frac{1}{\zeta_t}}.$$

Simplifying the example further to a single free parameter, I assume $s_t = \zeta_t = \frac{1}{2}$, $a_{1,t}A_{1,t}^{s_t} = a_{3,t}A_{3,t}^{s_t} = \frac{1}{4}$, $a_{2,t}A_{2,t}^{s_t} = \frac{3}{4}$, and $a_{4,t}A_{4,t}^{s_t} = A \in [0.5, 1.5]$ so that

$$C_t = \left(\frac{1}{4}\,\mathcal{K}_{1,t}^{\frac{\alpha}{2}} + \frac{3}{4}\,\mathcal{K}_{2,t}^{\frac{\alpha}{2}}\right)^2 \quad \text{and} \quad I_t = \left(\frac{1}{4}\,\mathcal{K}_{3,t}^{\frac{\alpha}{2}} + A\,\mathcal{K}_{4,t}^{\frac{\alpha}{2}}\right)^2.$$

Figure 5 depicts, as a function of A, the consumption rate resulting for different scenarios with a rate of pure time preference $\rho \in \{0.1\%, 1.4\%\}$ and $\kappa \in \{0.3, 0.4\}$. Minor adjustments of A, corresponding to relative adjustments of $a_{1,t}, ..., a_{4,t}$ and/or $A_{1,t}, ...A_{4,t}$ easily match the observed consumption rate of 74% for all parameter variations.

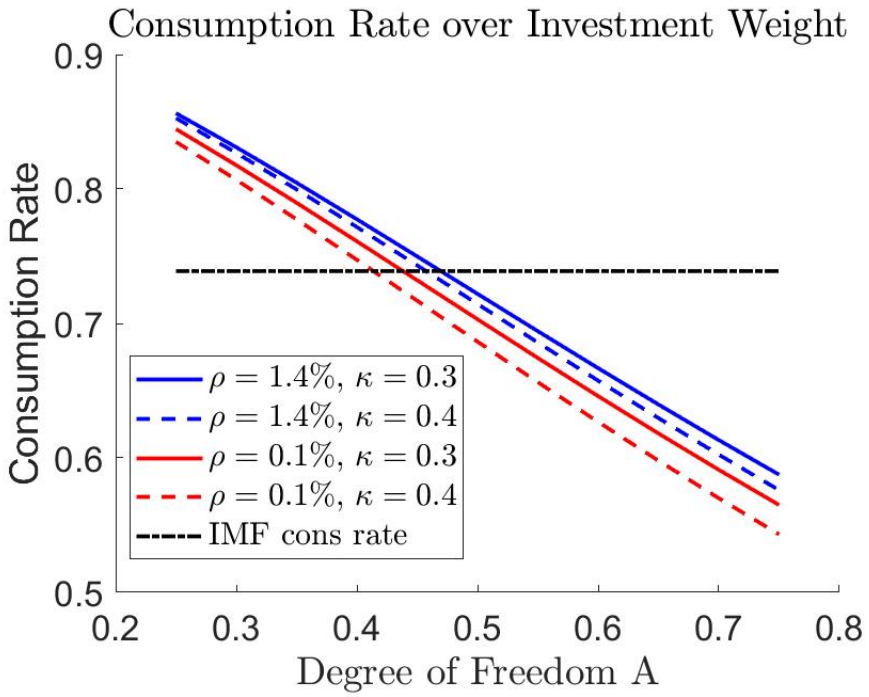

Figure 5: Illustration of optimal consumption rate as a function of A, which represents degrees of freedom in preference specification or technology levels or composition of the investment good.