ACE – Analytic Climate Economy (with Temperature and Uncertainty)[1]

Christian P. Traeger
Department of Economics, University of Oslo
ifo Institute, Munich ; Frisch Centre, Oslo

September 2019

Abstract: The Analytic Climate Economy (ACE) closes a gap between analytic climate change assessments and quantitative numeric integrated assessment models (IAMs) used in policy advising. Its closed-form solution links IAM components and parametric assumptions directly to their policy impacts. Its analytic nature overcomes Bellman's curse of dimensionality for a wide range of stochastic processes. ACE shows that uncertainty flips the main drivers of the carbon tax. Uncertainty also makes IAMs even more sensitive to the discount rate and its composition. Under a recent survey's median estimate for pure time preference, uncertainty almost triples the optimal tax.

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

Keywords: climate change, integrated assessment, uncertainty, risk aversion, recursive utility, social cost of carbon, carbon tax, carbon cycle, climate sensitivity, stochastic volatility, autoregressive gamma

[1]I am grateful for feedback and inspiration from Larry Karp, Terry Iverson, Armon Rezai, Buzz Brock, Lars Hansen, Rick van der Ploeg, Michael Greenstone, 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, Fortunat Joos, 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 project has been ongoing for 5 years and the present paper replaces an early SSRN version of the project (Traeger 2015) as well as several intermediate versions published on conference websites under slight variations of the name.

1 Introduction

Despite observed changes, the true reason to worry about climate change lies in the future. It originates in the wide range of possible climatic responses to any given emission scenario. First, we have a limited understanding of carbon dioxide (CO_2) accumulation in the atmosphere, where it causes the greenhouse effect. Second, a doubling of the atmospheric carbon dioxide can lead to moderate 1.5°C (degree Celsius) warming, intense 4.5°C warming, or even 7 or 8°C warming. Integrated assessment models (IAMs) of climate change analyze interactions of long-run economic growth, greenhouse gas (GHG) emissions, and global warming. Unsurprisingly, they find that the optimal carbon tax varies by orders of magnitude for different scientifically possibly temperature responses. Policy advice remains in the hands of deterministic models that explore and average large samples of deterministic worlds. The federal social cost of carbon (SCC) in the US is based on the average of such deterministic model runs (the technical support document also emphasizes the 95th percentile).

The present analytic climate economy (ACE) is part of an emerging literature that integrates uncertainty into the model, and explores the optimal policy in the face of an uncertain future. Aiming at the general audience, ACE develops analytic closed-form solutions and insights. Quantitatively, ACE competes with (deterministic) numeric models used to derive the US federal SCC. ACE derives structural insights into the determinants of the optimal price on carbon: the roles of uncertainty, risk aversion, discounting, carbon dynamics, temperature response, and their interactions. The solution overcomes the curse of dimensionality, which severely limits the integration of uncertainty into numeric models currently used in policy advising.

To model and calibrate the various climatic uncertainties, ACE is the first quantitative analytic IAM explicitly modeling climate change, i.e., the global temperature response to carbon dioxide emissions, and damages as a function of temperature increase. It is also the first IAM that breaks down the carbon tax into contributions from carbon versus climate dynamics. I show that such disentanglement is crucial as these two climate change drivers have quantitatively and qualitatively different impacts on the optimal carbon tax under both certainty and uncertainty. I find that under certainty, the carbon cycle dynamics cause a major amplification of the carbon tax, whereas warming delay and temperature persistence result in a reduction. In contrast, the uncertainty about carbon flows hardly matters for the optimal carbon tax, whereas temperature uncertainty adds a sizable risk premium.

Guided by the long-run risk literature in asset pricing, ACE disentangles risk aver-

sion from consumption smoothing to calibrate the risk-free discount rate and risk premia separately. Models lacking this feature either discount the future too highly, or disregard the risk premia.[2] Much of the long-run risk literature is concerned with asset *valuation*. In contrast, I focus on climate *policy*, making it crucial to account for the endogeneity of climate risk. Preference estimates of disentangled preferences imply low rates of pure time preference (Bansal et al. 2012), as do normative assessments (Stern 2007). ACE's analytic nature helps in solving the model easily for low rates of pure time preference,[3] adjusting parameters to reflect individual perspectives, and translating philosophical differences or different calibration approaches into their quantitative policy implications.

Analytic approaches to the integrated assessment of climate change date 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). In linear quadratic models *welfare* responds to uncertainty. In the widespread additive noise model, optimal *policy* remains unaffected by risk (weak certainty equivalence). In Hoel & Karp's (2001) multiplicative noise model, the optimal policy responds to uncertainty. 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. Recently, 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.

Golosov et al.'s (2014) elegant analytic IAM (AIAM) uses two climate change characteristics. First, a decadal time step is neither uncommon in IAMs nor particularly problematic given the time scales of the climate change problem. Therefore, the full-depreciation assumption is more reasonable than in other macroeconomic contexts. The present paper further weakens the full-depreciation assumption. Second, planetary "heating" (radiative forcing) is concave in atmospheric carbon and damages are convex in temperature. Consequently, the authors argue for a linear relation between

[2] Alternative explanations include habit formation preferences that result in a closely related modification of the Bellman equation, or the assumption that fat-tailed risk causes observed risk premia. Nakamura et al. (2013) suggest that the second explanation also does substantially better together with the preference disentanglement embraced here.

[3] In deterministic models, these low discount rates require very long and computationally expensive time horizons beyond the common model specifications. In stochastic models, a low discount rate reduces the contraction of the Bellman equation and numeric issues frequently prevent a solution.

past emissions and present damages. Their argument assumes that temperature responds immediately to an atmospheric carbon concentrations. However, the oceans keep 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.[4] The present paper follows the numeric IAMs used in policy advising and explicitly introduces the logarithmic relation between carbon dioxide's radiative forcing and temperature change (Nordhaus 2008, Hope 2006, Bosetti et al. 2006, Anthoff & Tol 2014). I then incorporate a novel model of ocean-atmosphere temperature dynamics that permits an analytic solution. Explicitly modeling temperature is crucial as I show that the uncertainty's policy impact differs substantially between temperature versus carbon.

Golosov et al.'s (2014) framework has sparked a growing AIAM literature, including applications to a multi-regional setting (Hassler & Krusell 2012, Hassler et al. 2018), non-constant discounting (Gerlagh & Liski 2018*b*, Iverson & Karp 2017), intergenerational games (Karp 2017), and regime shifts Gerlagh & Liski (2018*a*). The framework imposes *strong certainty equivalence*: not even welfare responds to uncertainty. I show that this feature arises from simultaneously setting the intertemporal elasticity of substitution and Arrow–Pratt risk aversion to unity. Whereas unity is within the estimated range of the intertemporal elasticity of substitution, Arrow–Pratt risk aversion is ubiquitously estimated higher. I solve ACE for a *flexible* degree of (disentangled) Arrow–Pratt *risk aversion*, accommodating for one of the most prominent criticisms of the model. Constant relative Arrow–Pratt risk aversion implies a decreasing coefficient of absolute risk aversion. This stylized fact is widely believed to hold and contrasts with linear quadratic AIAMs that capture only increasing absolute Arrow–Pratt risk aversion or risk neutrality.

Karp (2017) points out that Golosov et al.'s (2014) model solves analytically because it transforms to a linear-in-state model, a fact I use for a simpler presentation and solution of the more general ACE model. The stochastic extension leaves the class of linear-in-state dynamical systems and merges the original approach with affine stochastic processes developed for asset pricing (Gourieroux & Jasiak 2006, Le et al. 2010).

[4]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 make modeling CO_2 *concentrations* sufficient and ocean cooling negligible. The noted approximation trades off emission decay with warming delay and merely suggests that the impulse response to emissions (along particular steady emission scenarios) are fairly flat rather than peaked as suggested e.g. by Gerlagh & Liski's (2018*b*) DICE-based calibration of damage delay.

Alternatively, Li et al. (2016) and Anderson et al. (2014) leave the world of von Neumann & Morgenstern's (1944) axioms and introduce a preference for robustness to escape the strong certainty equivalence of the Golosov et al. (2014) framework.[5] The present paper breaks with both strong and weak certainty equivalence, while maintaining von Neumann & Morgenstern's (1944) classic axioms for choice under uncertainty, which are often considered desirable for rational or normative choice. In a similar vein, Ha-Duong & Treich (2004) and Traeger (2014*a*) discuss the relevance of disentangling risk aversion from intertemporal substitutability in climate change evaluation using simple two period models. More recently, van den Bremer & van der Ploeg (2018), Hambel et al. (2018*b*), and Hambel et al. (2018*a*) follow ACE's push for AIAMs disentangling risk aversion from intertemporal substitutability. Van den Bremer & van der Ploeg (2018) derive an approximate analytic perturbation solution to a more stylized IAM with many uncertainties, and Hambel et al. (2018*b*) derive exact analytic solutions assuming a simplified emission dynamics. They extend the setting to multiple regions playing a Nash equilibrium in emissions in Hambel et al. (2018*a*).

2 The Model

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

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 $\boldsymbol{E}_t$

$$Y_t = F(\boldsymbol{A}_t, \boldsymbol{N}_t, \boldsymbol{K}_t, \boldsymbol{E}_t) \quad \text{with} \tag{2}$$
$$F(\boldsymbol{A}_t, \boldsymbol{N}_t, \gamma\boldsymbol{K}_t, \boldsymbol{E}_t) = \gamma^{\kappa} F(\boldsymbol{A}_t, \boldsymbol{N}_t, \boldsymbol{K}_t, \boldsymbol{E}_t)\ \forall \gamma \in \mathbb{R}_+.$$

[5] Anderson et al. (2014) deviate from Golosov et al. (2014) by using a linear relation between the economic growth rate, temperature increase, and cumulative historic emissions. Both Li et al. (2016) and Anderson et al. (2014) combine a simpler analytic model with a more complex numeric IAM for quantitative simulation.

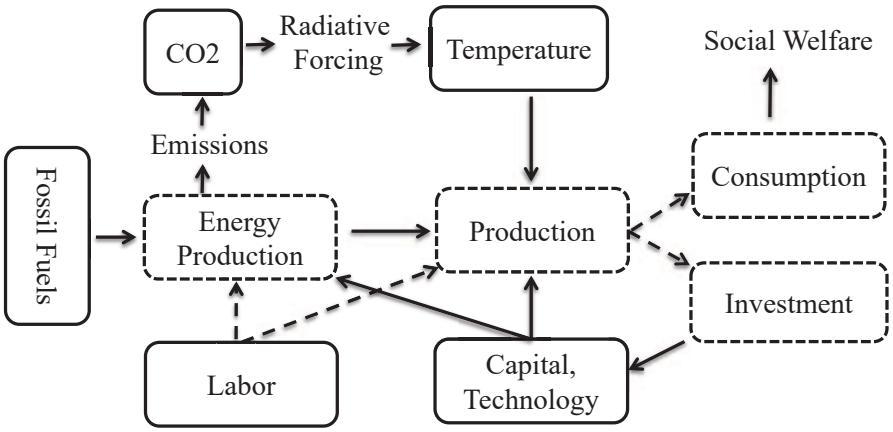


Figure 1: The structure of ACE and most IAMs. Solid boxes characterize the model's state variables, dashed boxes are flows, and dashed arrows mark choice variables.

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. 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 like the common Cobb–Douglas final good production with a constant-elasticity-of-substitution energy sector and, in contrast to earlier papers, 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.

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 like water, wind, or sunlight, which I assume to be abundant. In 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 $\boldsymbol{R}_0$ given and $\boldsymbol{R}_0 \in \mathbb{R}_+^{I^d}$.

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. Weakening Golosov et al.'s (2014) assumption of full depreciation, I assume the capital stock's equation of motion

$$K_{t+1} = \underbrace{Y_t[1 - D_t(T_{1,t})]}_{\equiv Y_t^{net}}(1 - x_t)\left[\frac{1 + g_{k,t}}{\delta_k + g_{k,t}}\right] , \tag{4}$$

where $x_t = \frac{C_t}{Y_t^{net}}$ is the endogenous consumption rate, δ_k the rate of capital depreciation, and $g_{k,t}$ is an exogenous approximation of the growth rate of capital. Appendix A shows that equation (4) coincides with the standard assumption on 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. The limited depreciation correction $\frac{1+g_{k,t}}{\delta_k+g_{k,t}}$ is larger under slow capital depreciation and slow capital growth. It makes the decision maker aware of the additional capital available in the next period and can adjust ACE's capital-output dynamics to macroeconomic observation.[6] [7]

2.2 ACE's Climate System

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

[6]Limited capital persistence changes output growth and, thereby, also affects the SCC. Note, however, that already the model's time step of 10 years makes the capital depreciation assumption more reasonable than it might appear: instead of an annual decay that leaves 30%–40% after 10 years, the model uses all the capital during 10 years, and none afterwards.

[7]The limited depreciation factor has no impact on the optimal carbon policy, given current world output. It is relevant only for the evolution of the model over time. The relevant implication of the capital accumulation in equation (4), similarly to the full depreciation assumption, is that the investment rate is independent of the system states. Consequently, climate policy will not operate through the consumption rate. Appendix A shows that the consumption rate is approximately 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).

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 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

$$\boldsymbol{M}_{t+1} = \boldsymbol{\Phi}\boldsymbol{M}_t + \boldsymbol{e}_1(\textstyle\sum_{i=1}^{I^d} E_{i,t} + E_t^{exo}) \tag{5}$$

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 $M_{1,t}$. 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. In equilibrium, the planet radiates the same amount of energy out into space as 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 that is commonly referred to as the greenhouse effect, causing a "heating mechanism" known as anthropogenic radiative forcing

$$F_t = \eta \frac{\log \frac{M_{1,t}+G_t}{M_{pre}}}{\log 2} \ . \tag{6}$$

The exogenous process G_t captures the greenhouse contribution from other non-CO_2 GHGs (measured in CO_2 equivalents). Anthropogenic radiative forcing was absent in preindustrial times where $G_t = 0$ 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[8] temperature of the (lower) atmosphere caused by a new level of radiative forcing F^{new} will be $T_{1,eq}^{new} = \frac{s}{\eta}F^{new} =$

[8]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.

$\frac{s}{\log 2} \log \frac{M_{1,eq}+G_{eq}}{M_{pre}}$. The parameter s is known as *climate sensitivity*. It measures 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 highly uncertain.

Temperature Dynamics. The next period's atmospheric temperature depends on current atmospheric temperature, the current temperature in the upper ocean, and on radiative forcing. 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 $T_{0,t} = \frac{s}{\eta} F_t$. Then, 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 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]

$$T_{i,t+1} = \mathfrak{M}_i^{\boldsymbol{\sigma}}(T_{i,t}, T_{i-1,t}, T_{i+1,t}) \text{ for } i \in \{1, ..., l\}. \tag{7}$$

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.

2.3 Uncertainty

I focus on the uncertainty generated by the evolution of carbon and temperature. This uncertainty turns (all) the equations of motion stochastic. Formally, the state variables become functions on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, whose filtration $\mathcal{F}$ is generated by the stochastic climate evolution (I suppress this dependence). Scenarios with persistent shocks or structural learning require additional informational state variables that I denote by the vector $\boldsymbol{I}_t$, itself an endogenous stochastic process with one-step-ahead uncertainty, and like all processes adapted to the filtration $\mathcal{F}$: we can learn only from what we already observed.

[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

Different sections will make different assumptions regarding the processes governing the conditional expectations over the future. They all have in common that they assume *affine stochastic processes*. Let $\boldsymbol{X}_t$ denote an n-dimensional vector of one-step-ahead uncertain state variables.[10] The process is affine if its conditional cumulant generating function is linear in the state

$$\log\left[\mathbb{E}\left(\exp(\boldsymbol{z}^\top \boldsymbol{X}_{t+1})|\boldsymbol{X}_t\right)\right] = a(\boldsymbol{z},\boldsymbol{A}_t,\boldsymbol{N}_t,\boldsymbol{\mathcal{K}}_t,\boldsymbol{E}_t) + \boldsymbol{b}(\boldsymbol{z})\boldsymbol{X}_t, \tag{8}$$

where $a(\cdot) \in \mathbb{R}$ and $\boldsymbol{b}(\cdot) \in \mathbb{R}^n$. The cumulant-generating function is the logarithm of the moment-generating function. In the applications, $z \in \mathbb{R}^n$ will relate closely to the the state's shadow values and equation (8) should hold for a domain incorporating the relevant state space and shadow-value domains. The $^\top$ denotes transposition. Both functions $a(\cdot)$ and $\boldsymbol{b}(\cdot)$ depend on the model's parametrization.

Equation (8) implies that the cumulant generating function is separable between the stochastic states and the labor and capital distributions as well as energy inputs and emissions. The right-hand side of equation (8) resembles the assumptions governing the equations of motion in linear-in-state models, including the AIAMs building on Golosov et al. (2014). However, here the condition is imposed on the cumulant generating function, not the equation of motion. Most of my stochastic specifications will be non-linear-in-state.

2.4 Objective Function

Utility governing deterministic outcomes is logarithmic in consumption C_t and the social planner's time horizon is infinite with discount factor β. I assume a stable population normalized to unity, but the approach generalizes to a population-weighted sum of logarithmic per capita consumption with population growth. Logarithmic utility provides a reasonable description of intertemporal substitutability. However, the

$\sigma_{i,i} = 1-\sigma_{i,i-1}-\sigma_{i,i+1} > 0$. 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 (7) 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] For are the states generating the uncertainty as well as informational state variables. Variables that are stochastic only in response to other states are not one-step-ahead uncertain, e.g., capital becoming stochastic only because next period temperature will change next period production and the resulting investment.

assumption performs poorly in capturing risk attitude. The long-run risk literature estimates the coefficient of relative risk aversion of a representative household closer to 10 than to unity (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).[11] Merely increasing the utility function's curvature would result in high risk-free discount rates that cannot be reconciled with market observation (risk-free rate puzzle) and would unwarrantedly discount away worries about the future climate. Moreover, the market rejects the assumption that the intertemporal elasticity of substitution fully determines risk attitude, which is an assumption built into the intertemporally additive expected utility (standard) model but which is not implied by the von Neumann & Morgenstern (1944) axioms.

I follow the asset pricing literature, an increasing strand of macroeconomic literature, and some recent numeric approaches to climate change assessment in using Epstein–Zin–Weil preferences. This approach accommodates a realistic coefficient of risk aversion, disentangling it from the unit elasticity of intertemporal substitution.[12] Then the Bellman equation is

$$V(k_t, \boldsymbol{\tau}_t, \boldsymbol{M}_t, \boldsymbol{R}_t, \boldsymbol{I}_t, t) = \max_{x_t, \boldsymbol{N_t}, \boldsymbol{K}_t, \boldsymbol{E_t}} \quad \log C_t \tag{9}$$
$$+\frac{\beta}{\alpha} \log\Big(\mathbb{E}_t \exp\big[\alpha\, V(k_{t+1}, \boldsymbol{\tau}_{t+1}, \boldsymbol{M}_{t+1}, \boldsymbol{R}_{t+1}, \boldsymbol{I}_{t+1}, t)\big]\Big).$$

Expectations $\mathbb{E}_t$ are conditional on time t information ($\mathbb{E}_t(\cdot) \equiv \mathbb{E}(\cdot|\, \mathcal{F}_t)$). I note that also the Bellman equation uses a generalized mean, here with the nonlinear weighting function $\exp(\alpha\ \cdot)$. A positive parameter α characterizes intrinsic risk loving, and a negative parameter characterizes intrinsic risk aversion. I use this sign convention because the risk attitude parameter will act on negative shadow values φ and the positive terms $\alpha\varphi > 0$ will correspond to risk-*aversion*-weighted shadow *costs*.

The parameter α characterizes risk attitude above and beyond the desire to smooth consumption over time. In the present model, it is mostly this parameter rather than Arrow–Pratt risk aversion that drives the risk premia. Figure 6 in the Appendix characterizes the total aversion deriving from the curvature of utility and the intrinsic attitude to risk α for a simple coin toss consumption lottery. For example, an agent with log-utility and no intrinsic risk aversion ($\alpha = 0$) will be indifferent between a

[11] Nakamura et al. (2013) obtain one of the lowest estimates by combining the long-run risk model and the Barro–Riesz model, still resulting in a coefficient of relative risk aversion of 6.4.

[12] The unit elasticity version of the Epstein–Zin–Weil preferences was first employed by Tallarini (2000), see Traeger (2012*a*) for an axiomatization of this special case.

lottery where she loses 5% of consumption if heads comes up, or gains 5.26% in case tails comes up. Say she consumes the equivalent of USD 1000 over some period. Then a fifty percent probability loss of USD 50 will be compensated by a fifty percent probably gain of USD 52.60. A calibration procedure described in Appendix E translates the asset pricing literature's estimates of Epstein & Zin's (1991) Arrow–Pratt risk aversion measure in the range of $[6, 10]$ into the range $\alpha \in [-1.2, -0.7]$. I pick the baseline value $\alpha = -1$ and will present sensitivity variations for the values $\alpha = -1.25$ and $\alpha = -.5$. In the lottery described above, the baseline choice of α increases the compensating gain to USD 55.60; the high and low sensitivity variations change this compensating gain to USD 56.30 and 54.10, respectively.

2.5 Deterministic Solution and Calibration

This subsection characterizes the class of damage functions and ocean-atmosphere temperature dynamics that permit an analytic solution (Proposition 1). To make this solution relevant, these functions have to permit a reasonable calibration, which I discuss subsequently to the proposition. Then, I characterize a class of stochastic models for which the high-dimensional dynamic programming problem in the state space reduces to a simple system of simultaneous equations for the shadow values (Theorem 2).

Deterministic solution. The *deterministic ACE* model is characterized by equations (1-7) (sections 2.1 and 2.2). The policy maker optimizes energy and labor inputs as well as consumption and investments to maximize the recursively defined objective (9) over the infinite time horizon. Appendix B transforms ACE into an equivalent linear-in-state model (Karp 2017). Such a transformation simplifies solving the model and fleshes out which changes would maintain (or destroy) analytic tractability. Linear-in-state models are solved by an affine value function.

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[13] $k_t = \log K_t$, $\boldsymbol{\tau}_t$ *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*

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

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

with $\xi_0 \in \mathbb{R}$ *and the mean in the equation of motion* (7) *for temperature layer* $i \in \{1, ..., l\}$ *takes the form*

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

with parameter $\xi_1 = \frac{\log 2}{s} \approx \frac{1}{4}$.

Appendix C provides the proof. The coefficients φ in the value function are the shadow values of the respective state variables. The coefficient vector on the resource stock, $\boldsymbol{\varphi}_{R,t}^{\top}$, must be time dependent: the shadow values of the exhaustible resources increase over time following the endogenously derived Hotelling rule (see Appendix B). The process φ_t captures the value contribution of the exogenous processes, including technological progress.

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 calibrate ACE's damage coefficient to match DICE 2013's calibration points of 0 and 2.5°C exactly, delivering the damage semi-elasticity $\xi_0 = 0.0212$ (Nordhaus 2008, Nordhaus & Sztorc 2013). Figure 2 compares the resulting damage curve (solid green), variations with plus/minus 50% of the damage semi-elasticity (dashed), Nordhaus' original DICE damage function (solid red), and a damage function by Weitzman (2010) suggesting that little is known about damages at higher temperature levels and that damages might turn out much more convex.[14] As compared to DICE, the base calibration of ACE's damage function generates slightly higher damages below a 2.5°C warming, slightly lower damages above a 2.5°C warming until 12°C warming, and higher damages at a warming above 12°C (a hard-to-conceive change of life on the planet). The plus/minus 50% variations lie almost everywhere above/below DICE's damage curve and the analytic solution permits a simple evaluation of such changes in the calibration.

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

[14] The functional forms are $D(T) = \frac{1}{1+0.00267T^2}$ for the DICE 2013 model and $D(T) = 1 - 1/\left((1 + \frac{T}{20.46})^2 + (\frac{T}{6.081})^{6.754}\right)$ for Weitzman (2010).

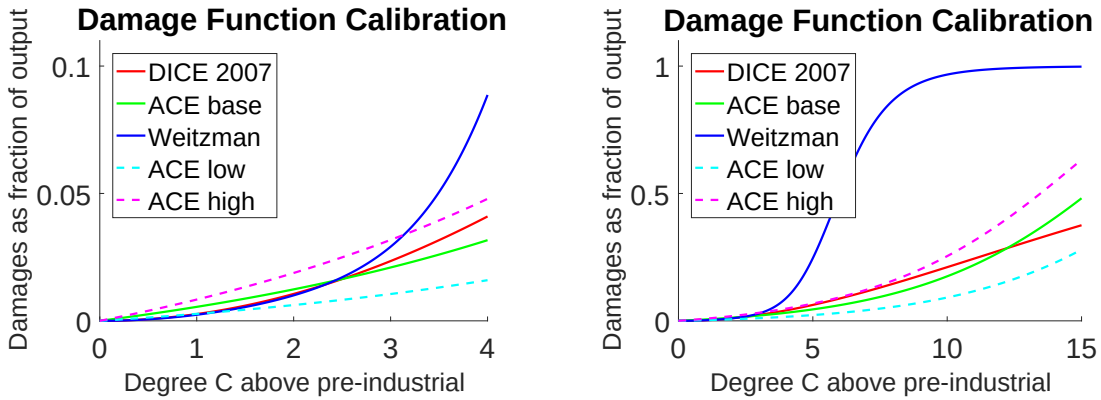


Figure 2: ACE's damage function compared to that of DICE 2013 and to a highly convex damage function suggested by Weitzman (2012). All three solid lines coincide for a 2.5°C warming, the common calibration point based on Nordhaus & Sztorc (2013). The dashed curves depict ACE's damage function for a ±50% variation of the base-case damage coefficient $\xi_0 \approx 0.021$.

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 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. It performs slightly worse for the high "business as usual" scenario RCP 8.5, but still well compared to other IAMs.[15] Transformed to the vector of generalized temperatures $\boldsymbol{\tau}_t$, the temperatures' equations of motion (11) take the linear

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

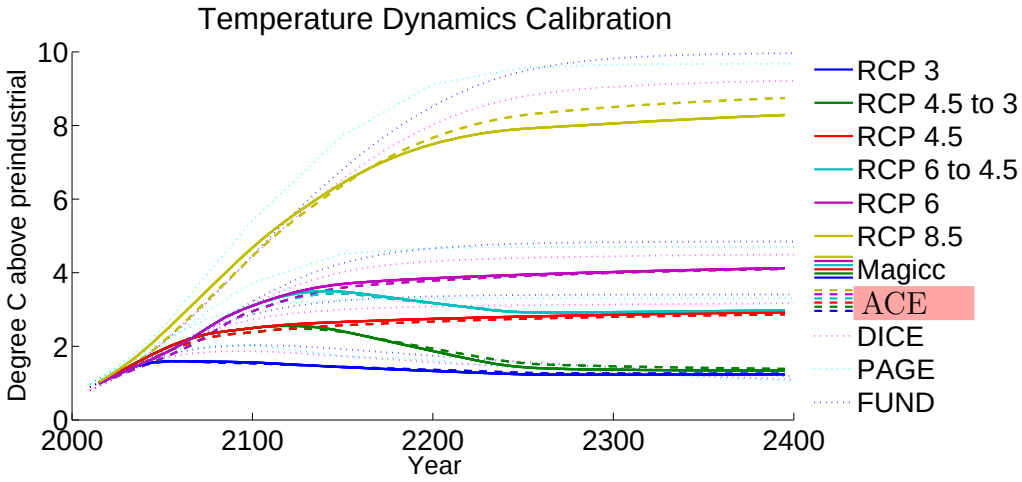


Figure 3: ACE's temperature response compared to MAGICC6.0 using the color-coded radiative forcing scenarios of the latest IPCC assessment report. RCP 3 is the strongest stabilization scenario and RCP 8.5 is a business as usual scenario. The MAGICC model (solid lines) emulates the large AOGCMs and is 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 presents as well 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.

vector form

$$\boldsymbol{\tau}_{t+1} = \boldsymbol{\sigma}\boldsymbol{\tau}_t + \sigma^{forc}\frac{M_{1,t} + G_t}{M_{pre}}\boldsymbol{e}_1 \ . \tag{12}$$

Further characterization and calibration. The optimal consumption rate is $x_t^* = 1 - \beta\kappa$. 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 energy sector, but they are not needed to determine the optimal carbon tax.

I calibrate the remaining parts of ACE as follows. A capital share of $\kappa = 0.3$ and the International Monetary Fund's (IMF) 2018 investment rate forecast of $1 - x^* = 26\%$ pin down an annual rate of pure time preference of $\rho = 1.42\%$. Present world output Y is 10 times (time step) the IMF's global economic output forecast of $Y_{2018}^{annual} = 135$ trillion USD (purchasing power parity). I use the carbon cycle of DICE 2013.

2.6 General Solution of the Stochastic ACE

The present study focuses on uncertain climate dynamics. The *stochastic ACE* is characterized by equations (1-4) and equation (8) with $\boldsymbol{X}_t = (\boldsymbol{M}_t^\top, \boldsymbol{\tau}_t^\top, \boldsymbol{I}_t^\top)^\top$: the affine process governs one-step-ahead uncertainty of carbon, temperature, and information. The carbon and temperature dynamics in equations (5) and (12) are (degenerate) cases of equation (8). The specific stochastic models introduced later collapse to their deterministic counter parts in the case of vanishing uncertainty.

Proposition 2 *Define k_t, $\boldsymbol{\tau}_t$, and $D(T_{1,t})$ as in Proposition 1. Let $\boldsymbol{X}_t = (\boldsymbol{M}_t^\top, \boldsymbol{\tau}_t^\top, \boldsymbol{I}_t^\top)^\top \in \mathbb{R}^n$ follow an affine stochastic process (equation 8) with $\boldsymbol{b}(\cdot) = (\boldsymbol{b}^M(\cdot), \boldsymbol{b}^\tau(\cdot), \boldsymbol{b}^I(\cdot))$.*

Then, an affine value function solves the stochastic ACE if and only if the set of shadow values $\boldsymbol{\varphi}_M, \boldsymbol{\varphi}_\tau, \boldsymbol{\varphi}_I$ solve the algebraic equations

$$\begin{aligned}
\varphi_{M,i} &= \frac{\beta}{\alpha}\, b_i^M(\alpha\boldsymbol{\varphi}_M, \alpha\boldsymbol{\varphi}_\tau, \alpha\boldsymbol{\varphi}_I) && \forall i = 1, ..., m \\
\varphi_{\tau,i} &= \frac{\beta}{\alpha}\, b_i^\tau(\alpha\boldsymbol{\varphi}_M, \alpha\boldsymbol{\varphi}_\tau, \alpha\boldsymbol{\varphi}_I) - \delta_{i,1}\frac{\xi_0}{1-\beta\kappa} && \forall i = 1, ..., l \\
\varphi_{I,i} &= \frac{\beta}{\alpha}\, b_i^I(\alpha\boldsymbol{\varphi}_M, \alpha\boldsymbol{\varphi}_\tau, \alpha\boldsymbol{\varphi}_I) && \forall i = 1, ..., n-l-m
\end{aligned}$$

where $\delta_{i,1}$ denotes the Kronecker delta (one if $i = 1$ and zero otherwise). The shadow value $\varphi_{M,1}$ determines the optimal carbon tax.

Appendix F provides the proof. Proposition (8) transforms a high dimensional and difficult- or — depending on the precise specification — impossible-to-solve stochastic dynamic optimization problem on the state space into a simple root-finding problem for the shadow values, which solves either in closed form or easily on any computer.

Most analytic assessments of climate change under uncertainty have focused on the welfare impact under uncertainty. Proposition 2 points out that even major welfare losses may not change the optimal carbon tax if the uncertainty operates through the affine part $a(\boldsymbol{z},\boldsymbol{A}_t,\boldsymbol{N}_t,\boldsymbol{\kappa}_t,\boldsymbol{E}_t)$ of equation (8). The impact of uncertainty on the optimal policy depends on its interaction with the state variables. The set of affine stochastic processes is large and includes the autoregressive shock model with almost arbitrary distributions, the normal-normal Bayesian learning model, the Gaussian square root process, and the autoregressive gamma model (Gourieroux & Jasiak 2006, Le et al. 2010).

3 Policy Results

The SCC is the money-measured present value welfare loss from adding a ton of CO_2 to the atmosphere. The economy in section 2.1 decentralizes in the usual way and the Pigovian carbon tax is the SCC along the optional 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.

3.1 The Carbon Tax under Certainty

Appendix C 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, emissions' impulse responses, or reservoirs, giving a simpler and yet richer description of the dynamic characteristics driving the optimal carbon tax.

Proposition 3 *(1) Under certainty and the assumptions of section 2, the SCC in (USD-2018-) money-measured consumption equivalents is*

$$SCC_t^{det} = \underbrace{\frac{\beta Y_t^{net}}{M_{pre}} \underbrace{\xi_0}_{2.1\%}}_{11\frac{USD}{tCO_2}} \underbrace{\left[(\mathbf{1}-\beta\boldsymbol{\sigma})^{-1}\right]_{1,1}}_{1.1} \underbrace{\sigma^{forc}}_{0.54} \underbrace{\left[(\mathbf{1}-\beta\boldsymbol{\Phi})^{-1}\right]_{1,1}}_{4.3} = 30\frac{USD}{tCO_2} \tag{13}$$

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 **??**.
(2) A carbon cycle (equation 5) satisfying mass conservation of carbon implies a factor $(1-\beta)^{-1}$, approximately proportional to $\frac{1}{\rho}$, in the SCC (equation 13).

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 constant semi-elasticity of net output to a transformed temperature increase, i.e., to an increase of $\tau_1 = \exp(\xi_1 T_1)$. In the absence of any interesting climate dynamics, these terms would imply a carbon tax of $11\frac{USD}{tCO_2}$, or 10 cents per gallon at the pump ($\approx$ 2 €-cents per liter). The interesting stuff making the carbon tax more serious happens in the climate dynamics (and in the uncertainty).

Contribution of carbon dynamics. Analytic models of climate change so far captured carbon dynamics in a decay formulation, whereas numeric IAMs incorporate

a carbon cycle respecting that CO_2 does not decay. ACE integrates a carbon cycle and a von Neumann series expansion of $\beta\mathbf{\Phi}$ interprets the implications for the carbon tax

$$(1-\beta\mathbf{\Phi})^{-1} = \sum_{i=0}^{\infty} \beta^i \mathbf{\Phi}^i \ .$$

The element $[\mathbf{\Phi}^i]_{1,1}$ of the transition matrix characterizes how much of the carbon injected into the atmosphere in the present remains in or returns to the atmospheric layer in period i, after cycling through the different carbon reservoirs. E.g., $[\mathbf{\Phi}^2]_{1,1} = \sum_j \Phi_{1,j}\Phi_{j,1}$ 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

$$carb_1 \equiv \left[(\mathbf{1}-\beta\mathbf{\Phi})^{-1}\right]_{1,1}$$

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. Similarly $carb_2 \equiv [(\mathbf{1}-\beta\mathbf{\Phi})^{-1}]_{1,2}$ would characterize the carbon cycle contribution if our CO_2 emissions were pumped into the shallow ocean instead of the atmosphere. This measure has been suggested as a geoengeneering solution. More importantly, the difference between the value contributions of atmospheric and shallow ocean carbon

$$\Delta carb \equiv carb_1 - carb_2 \tag{14}$$

will play an important role when integrating uncertainty. Quantitatively, the persistence of carbon increases the earlier value by a factor of $carb_1 = 4.3$ and the resulting carbon tax would be almost $50\frac{USD}{tCO_2}$ or over 40 cents per gallon – ignoring temperature dynamics.

Contribution of temperature dynamics. The terms $\left[(\mathbf{1}-\beta\boldsymbol{\sigma})^{-1}\right]_{1,1}\sigma^{forc}$ capture the atmosphere-ocean temperature dynamics resulting in both delay and persistence: 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, stays, and returns to our atmosphere. Thus, the simple closed-form expression for the carbon tax in equation (13) 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 26 cents per gallon (≈ 6 €-cents per liter).[16] Appendix D discusses benefits of temperature cooling and geoengeneering. It also illustrates two-reservoir carbon and temperature models, and compares the carbon cycle's impact on the SCC to frequently used models assuming simple carbon decay.

A conceptual implication for policy making. The SCC in equation (13) is independent of the atmospheric carbon concentration and of the prevailing temperature level. A corresponding independence of past emissions prevails already in Golosov et al. (2014), and it opposes 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: 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 of only a certain range of wavelengths. As CO_2 accumulates in the atmosphere, more and more of this energy spectrum is already trapped and most of the equilibrium energy exchange happens through other parts of the spectrum. As a result, warming is logarithmic in the prevailing atmospheric CO_2 concentration (equation C.10), and the marginal ton's warming impact is proportional to the inverse of the prevailing concentration.

3.2 Discounting and Time Preference

It is well known that the consumption *discount rate* plays a crucial role in valuing long-run impacts. Finding (2) in Proposition 3 is different. It states that the interaction of *pure time preference* and carbon cycle dynamics is the main sensitivity when it comes

[16] Golosov et al. (2014) and Gerlagh & Liski (2018*b*) use an emission response model similar to 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.

to discounting. The proposition ties this sensitivity directly to the fact that carbon dioxide does not decay but only cycles through the different reservoirs, implying that a fraction of our current emissions remains in or returns to the atmosphere in the long run. The growth contribution to the discount rate is less relevant because the damages from climate change grow approximately proportionally to consumption, offsetting the reduction in marginal utility caused by economic growth. This extreme sensitivity weakens as we move away from log utility as van den Bijgaart et al. (2016) and Rezai & 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 less about the long run (along the optimal trajectory). ACE normalizes population size to unity. Part (2) of Proposition 3 might be the most interesting change if population grows instead at a constant rate g. Assuming the common population-weighted average utility objective, the factor $\frac{1}{\rho}$ changes to $\frac{1}{\rho-g}$, making the SCC even more sensitive already at higher rates: IAMs put additional weight on future generations that are more numerous, which acts as a reduction of *pure* time preference. The contribution of temperature dynamics is less sensitive to time preference because heat is not conserved and constantly exchanged with outer space.[17]

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%.[18] It does not give any direct guidance regarding its composition, leaving a huge degree of freedom to the modelers of the Interagency Working Group on the US' federal SCC. The standard calibration adopted above implied a rate of pure time preference of $\rho = 1.42\%$. More sophisticated long-run-risk models in asset pricing deliver much lower rates while matching risk-free rate and risk premia substantially better, e.g., Bansal et al. (2012) calibrate a rate of 0.11% closely matching the Stern's (2007) Review's normative choice.[19] Similarly, much lower rates of pure time preference arise

[17] Already Golosov et al.'s (2014) SCC formula shows a related sensitivity to pure time preference, arising from modeling emissions as a convex combination 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 is the first to directly factor out the sensitivity factor $(1-\beta)^{-1}$ resulting from mass conservation in the carbon cycle formulation.

[18] 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% reflects pre-tax capital interest. In addition, a third rate between 1-3% can be suggested if future generations are affected.

[19] Traeger (2012*a*) shows how uncertainty-based discounting by an agent whose risk aversion does

if observed market equilibria are interpreted as reflecting individual life-cycle choices (Schneider et al. 2013) and when distinguishing consumption choice from political long-term decision making and voting (Hepburn 2006).

I illustrate the response of the different carbon tax components to the choice of time preference using the median response of Drupp et al.'s (2018) recent expert survey for the rate of pure time preference of $\rho = 0.5\%$

$$SCC_t^{det} = \underbrace{\frac{\beta Y_t^{net}}{M_{pre}}}_{\cancel{1}12\frac{USD}{tCO_2}} \underbrace{\xi_0}_{2.1\%} \underbrace{\left[(\mathbf{1}-\beta\boldsymbol{\sigma})^{-1}\right]_{1,1}}_{\cancel{1.1}1.3} \underbrace{\sigma^{forc}}_{0.54} \underbrace{\left[(\mathbf{1}-\beta\boldsymbol{\Phi})^{-1}\right]_{1,1}}_{\cancel{4.3}8.4} = \cancel{30}72 \frac{USD}{tCO_2}.$$

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).

3.3 Optimal Emission Response

The optimal carbon tax (SCC) goes along with an optimal emission response.

Proposition 4 *Under the assumptions of section 2, the optimal emissions from a dirty resource $i \in \{1, ..., I^d\}$ satisfy*

$$E_{i,t}^* = \frac{\sigma_{Y,E_i}(\boldsymbol{A}_t, \boldsymbol{N}_t^*, \boldsymbol{K}_t^*, \boldsymbol{E}_t^*) Y_t^{net}}{HOT_{i,t} + \beta SCC} \tag{15}$$

where $\sigma_{Y,E_i}(\cdot) = \frac{\partial F(\cdot)}{\partial E_i}\frac{E_i}{Y}$ is the production elasticity of the resource, stars denote the optimal (or decentralized equilibrium) allocation (under the optimal tax), and $HOT_{i,t}$ is the Hotelling rent of resource i in period t consumption equivalents.

Given the Hotelling rent of a resource, its optimal use increases in the production level and in the production elasticity σ_{Y,E_i}, and it decreases in the SCC. Unless the elasticity is constant, equation (15) is an implicit equation that has to be solved together with the other optimality conditions (see Appendix C.1).

The Hotelling rent arises from the intertemporal scarcity of a resource. It is zero if we do not expect to exhaust the resource over time. A higher SCC tends to reduce emissions. As a result, demand for fossil fuel decreases in all periods and so

not coincide with her consumption-smoothing preference (falsely) manifests as pure time preference in the standard economic model.

does resource scarcity: the Hotelling rent falls, which partly crowds out the emission reduction resulting from a higher carbon tax. Assuming that the production elasticity of the resource only depends on exogenous technological progress, I show in the proof of Proposition 4 that a higher SCC always lowers emissions in the present (it lowers emissions conditional on the same resource stock). If the resource remains scarce, lower emissions today imply higher emissions in the future. Equation (15) points out that optimal emissions from a scarce resource with a high Hotelling rent (e.g. oil) respond much less to the SCC than those from a sector with a low Hotelling rent (e.g. coal): for resources with a high scarcity value, the carbon tax largely replaces scarcity value. The simple equation illustrates the literature's finding, including in Golosov et al.'s (2014) simulations, that coal use will respond substantially stronger to optimal emission policies than oil use. Finally, equation (15) illustrates that technological progress reduces optimal emissions if (and only if) it reduces the production elasticity of the resource. For scarce resources, this improvement in the production elasticity will, once more, be partly crowded out by a reduction in the Hotelling rent.

The present finding relates to, but does not constitute the so-called "green paradox" (Sinn 2012, van der Ploeg & Withagen 2012, Jensen et al. 2015). In the above setting, a partial crowding out of the SCC's effect on emissions merely reflects that the value of using the resource is high even after adding the tax that replaces a high scarcity rent. The standard green paradox would arise if the policy maker announces a carbon tax for the future without taxing the fossil fuel immediately. Then, the future reduction in scarcity lowers $HOT_{i,t}$ already today when the SCC is not yet in place. Today's emissions would increase even further above the optimal trajectory than without the announcement. Similarly, if a large fraction of countries sets the optimal carbon tax, the resulting reduction in $HOT_{i,t}$ makes it more attractive for a non-participating country to emit. Another "green paradox" arises if a new technology is expected to reduce or replace carbon fuel usage in the medium to long-run future ($\sigma_{Y,E_i}(\boldsymbol{A}_t)$ lower for some $t > T$). Then, demand and $HOT_{i,t}$ fall with respect to a world in which such a technology is absent (or not expected), and near-term emissions under the optimal tax tend to be higher in a world with the cleaner expected technology.

4 Long-Run Risk and Climate Change

The causal origin the greenhouse effect can be measured in the laboratory: carbon dioxide absorbs the planet's outgoing radiation. Yet, quantifying the warming result-

ing from a given CO_2 trajectory is difficult and results in a highly uncertain temperature response. First, over 10% of the annual flow of anthropogenic carbon emissions leave the atmosphere into an unidentified sink. Our *lack of understanding current carbon flows* implies major uncertainties in predicting future carbon dynamics. Second, even if we knew future atmospheric carbon concentrations, we would remain highly uncertain about the implied warming. The deterministic ACE assumed that a doubling of the preindustrial CO_2 concentration increases medium- to long-run temperature by 3C (*climate sensitivity*).[20] The value of 3°C was the best guess in the first four IPCC assessment reports. The latest report deleted this best guess and only cites a "likely" range of 1.5-4.5°C (IPCC 2013), where "likely" characterizes a 66% probability interval, and the distribution is strongly skewed towards higher values. ACE integrate at least some of the huge uncertainties explicitly into the models' decision-making and evaluation process. For this purpose, I borrow heavily from the finance literature.

One of the most popular (and descriptively successful) asset pricing models is the so-called *long-run risk model* based on Bansal & Yaron (2004). It combines the aversion to risk generated by Epstein-Zin-Weil preferences with small but highly persistent (long-run risk) shocks. In the asset pricing literature, these shocks directly govern the consumption process. Croce (2014) generates the consumption process from shocks to factor productivity in a macroeconomic model. Jensen & Traeger (2014) introduce such a long-run risk model to the integrated assessment of climate change.[21] Here, the long-run risk governs productivity and, thus, growth of the economy. In addition, the earlier applications accounted only for shocks to the expected mean, whereas the long-run risk model's application in finance also relies crucially on stochastic volatility.

ACE applies the long-run risk framework to climate risk, and employs affine processes to obtain closed-form solutions. Asset pricing models are generally concerned with the valuation of assets under exogenous risk. In contrast, optimally controlling the climate has to acknowledge the endogeneity of climate risk: the more we perturb the climate system with our CO_2 emissions, the higher the uncertainty about the future climate, and the higher the risk premium for the SCC. The next section explains the

[20] We are currently over 40% towards such a doubling of CO_2, and if we include other GHGs' CO_2 equivalent forcing, we are over 60% of the way. The present warming is still much below the corresponding equilibrium increase because of the atmosphere-ocean temperature interaction discussed in sections 2.2 and 3.1 and because of the response time of some feedback processes.

[21] Jensen & Traeger (2014) use a simplified climate model to derive both an approximate analytic as well as numeric solutions. Cai & Lontzek (forthcoming) recently redo their numeric analysis with additional climate states and a more detailed calibration of the resulting consumption process. For this purpose, Cai & Lontzek (forthcoming) tame the tail-risk in Jensen & Traeger's (2014) infinite time horizon stochastic fix-point problem by using a discrete Markov chain (and a finite time horizon).

standard long-run risk model, applies it to climate, and extends it to endogenous risk. For this purpose, it is easiest to start with uncertainty about future carbon flows, which permits extending the canonical long-run risk framework. The second section develops a novel version of the long-run risk model that permits a reasonable application to the temperature process, where the canonical model fails.

4.1 Uncertainty about Carbon Flows

The long-run risk model adds two stochastic processes to the equation of motion, here equation (5) governing carbon flows. The first process $\boldsymbol{x}_t^M$ governs the conditional expectations, i.e., the one-step-ahead expectations given the current state of the system. The second process $\boldsymbol{\sigma}_t^M$ governs the conditional volatility of carbon flows. Importantly, both are highly persistent and, thus, describe long-run uncertainty. The resulting equation for the carbon stock becomes

$$\boldsymbol{M}_{t+1} = \boldsymbol{\Phi}\boldsymbol{M}_t + (\textstyle\sum_{i=1}^{I^d} E_{i,t} + E_t^{exo})\boldsymbol{e}_1 + \boldsymbol{x}_t^M + \boldsymbol{\sigma}_t^M \mu_t^M \tag{16}$$

where $\mu_t^M \sim N(0,1)$ is a serially uncorrelated white noise, implying that $\boldsymbol{\sigma}_t^M$ characterizes the conditional variance of carbon flows. Because carbon does not decay, conditional expectation and variance have to be vectors balancing the flow between the different reservoirs. The vector $\boldsymbol{x}_t^M = (1,-1,0)^\top \; x_t^M$ redirects x_t^M tons of carbon from the shallow oceans and biosphere into the atmosphere – where it enhances the greenhouse effect. A negative realization of x_t^M goes along with a better-than-expected carbon removal from the atmosphere into the other reservoirs. Similarly $\boldsymbol{\sigma}_t^M = (1,-1,0)^\top \; \sigma_t^M$ characterizes the one-step ahead stochasticity of the carbon flows between the atmosphere and its neighboring sink.[22] The persistent processes governing conditional expectations and variance are

$$x_{t+1}^M = \gamma^x x_t^M + \delta^{Mx}\sqrt{\tfrac{M_{1,t}}{M_{pre}} - \eta_M}\; \chi_t^M + \delta^{\sigma x}\; \sigma_t^M \; \omega_t^M \tag{17}$$

$$\sigma_{t+1}^{M\;2} = \gamma^\sigma \sigma_t^{M2} + \delta^{M\sigma}\left(\tfrac{M_{1,t}}{M_{pre}} - \eta_M\right) + \bar{\sigma}^M \nu_t^M \tag{18}$$

The γ-*parameters* characterize the persistence of the shocks to the mean (first equation) and the volatility process (second equation). As I show in section 4.4, epistemological uncertainty corresponds to a high persistence. Epistemological uncertainty expresses

[22] I focus on the exchange of atmospheric carbon, which determines the greenhouse effect, with the land sinks and sources and the shallow ocean. Using the DICE model's carbon cycle both of these adjacent sinks and sources are combined into the second reservoir.

the "ignorance" of the modeler and, here, the scientific community. In the canonical long-run risk model, the risk is exogenous to the decision maker. The decision-maker merely values the risk, but does not control it.

In climate change, the risk is *endogenous* to our decision problem. The more we perturb the climate system the higher the uncertainty about its future evolution. To capture the endogeneity of the risk, I introduce the second terms on the right hand side. The term $\frac{M_{1,t}}{M_{pre}} - \eta_M$ grows as we deviate further from the pre-industrial level where the climate was relatively stable; $\eta_M < 1$ is a free calibration parameter. The *δ-parameters* in these second terms characterize the strength of the endogenous contribution to climate risk. For the stochastic volatility (equation 18), the last term on the right-hand side specifies the exogenous uncertainty, where $\nu_t^M \sim N(0,1)$ implies that $\bar{\sigma}^M$ characterize the corresponding variance. Finally, the parameter $\delta^{\sigma x}$ in equation (17) permits coupling the stochastic volatility to the conditional expectations process, where $\omega_t^M \sim N(0,1)$ is once again serially uncorrelated white noise. This common third channel in long-run risk models permits that the long-run uncertainty about the volatility of carbon flows also affects the movement of conditional expectations.

Proposition 5 *The stochastic processes (16-18) governing carbon flows change the deterministic SCC^{det} stated in Proposition 3 to*

$$SCC_t = SCC_t^{det}\Big(1 + \theta_M(\theta_M^*)\Big) \approx \frac{SCC^{det}}{1-\theta_M^*} \tag{19}$$

with the uncertainty contributions $\theta_M(\theta_M^*) = \frac{1-\sqrt{1-4\theta_M^*}}{1+\sqrt{1-4\theta_M^*}} > 0$ and

$$\theta_M^* = \frac{\alpha}{2}\frac{\beta\varphi_{M,1}^{det}}{M_{pre}}\Big[A^{M\to x^2} + A^{M\to\sigma} + A^{M\to\sigma}A^{\sigma\to x^2}\Big]\frac{(\Delta carb)^2}{carb_1} \tag{20}$$

where $A^{M\to x} = \frac{\delta^{Mx}\beta}{1-\gamma^x\beta}$, $A^{M\to\sigma} = \frac{\delta^{M\sigma}\beta}{1-\gamma^\sigma\beta}$ and $A^{\sigma\to x} = \frac{\delta^{\sigma x}\beta}{1-\gamma^x\beta}$ characterize the individual risk channels and $\varphi_{M,1}^{det}$ denotes the shadow value of atmospheric carbon under certainty.[23] *The uncertainty contribution $\theta_M(\theta_M^*)$ is convex in θ_M^*.*

Equation (19) expresses the SCC's premium for carbon flow uncertainty as a proportionality factor to the deterministic SCC. It always increases the SCC and is convex in the contributions characterized by equation (20) as θ_M^*. This contribution is proportional to intrinsic risk aversion α and the discounted shadow value of atmospheric

[23] That is the SCC expressed in utils: $\varphi_{M,1}^{det} = -\frac{SCC_t}{(1-\beta\kappa)Y_t^{net}}$. Recall that $\alpha < 0$ so that $\alpha\varphi_{M,1}^{det} > 0$.

carbon under certainty. First, it is not merely Arrow–Pratt risk aversion that matters for the SCC, but by how much Arrow–Pratt risk aversion exceeds the desire to smooth consumption over time (α measures this difference and characterizes intrinsic aversion to risk, see section 2.4). Second, the proposition implies that the share of the overall SCC contributed by uncertainty ($\frac{SCC_t}{SCC_t^{det}}$) increases convexly in its deterministic base value SCC_t^{det} because the shadow value of atmospheric carbon is itself proportional SCC^{det}. The more serious the climate problem, the larger the relative importance of uncertainty.

The three risk channels abbreviated $A^{M\to x}$, $A^{M\to\sigma}$, and $A^{\sigma\to x}$ disentangle the contributions from the different aspect's of long-run risk. Each of these channels increases in the corresponding δ-parameter scaling the endogeneity of climate risk. The proposition shows that only this *endogenous uncertainty*, which results from changes in the carbon concentrations, affects the optimal policy.[24] In addition, each channel's contribution increases in its corresponding γ-parameters scaling the *shock persistence*. We do not have to worry about serially uncorrelated short-term fluctuations in the mean or volatility of carbon flows. If uncertainty matters, then it is because of persistent long-run risk. The conditional expectations channel $A^{M\to x}$ contributes quadratically. This contribution represents the long-term uncertainty about expected carbon flows. The stochastic volatility channel $A^{M\to\sigma}$ enters twice. It reflects the harm from an increase in volatility that results from a (further) perturbation of the climate system. Its first appearance captures the long-run uncertainty about the volatility of carbon flows. For $\delta^{\sigma x} > 0$, stochastic volatility also governs the conditional expectations. Then, the combined channel $A^{M\to\sigma}{A^{\sigma\to x}}^2$ reflects that an increase in carbon increases stochastic volatility which increases the variance of conditional expectations (in a persistent manner).

Finally, the uncertainty contribution is convex in its driving force: carbon in the atmosphere is more harmful than carbon in the adjacent sinks. Defined in equation (14), $\Delta carb$ captures this difference in the value contribution and enters relative to the value contribution of atmospheric carbon $carb_1$. These two terms interact carbon cycle characteristics and time preference with the long-run risk.

[24] The exogenous uncertainties affect welfare and "the climate asset's value." Yet, they do not affect the optimal climate policy.

4.2 Temperature Uncertainty

The Gaussian model of the previous section cannot adequately represent temperature uncertainty. Temperature $T_{i,t} = \frac{1}{\xi_i}\log(\tau_{i,t})$ is a logarithmic transformation of the state $\tau_{i,t}$ and the expected temperature dynamics has to approximately equal the deterministic dynamics. As a result, generalized temperature $\tau_{1,t}$ has to be governed by a positively skewed distribution with a suitable lower bound.[25] For this purpose, I use the autoregressive gamma process introduced by Gourieroux & Jasiak (2006), which has been applied to the long-run risk literature in asset pricing by Le et al. (2010) and Creal (2017). Again, I extend the model to capture the endogeneity of climate risk.[26] The one-step-ahead state in the autoregressive gamma process is governed by a gamma distribution whose shape parameter is modulated by a realization of a Poisson distribution. I refer to Appendix F.2 for details. The canonical Gaussian long-run risk model separates the long-run risk between the risks governing conditional expectations and stochastic volatility. The current model merges these two long-run uncertainties into a single process.

I model the long-run risk by the autoregressive gamma process y_t, which I shift by a deterministic process y_t^o that adjusts the expectations. Then, the equation of motion (12) changes to

$$\boldsymbol{\tau}_{t+1} = \boldsymbol{\sigma}\boldsymbol{\tau}_t + \left(\sigma^{forc}\frac{M_{1,t}+G_t}{M_{pre}} + h(\underbrace{y_{t+1}-y_{t+1}^o}_{\equiv z_{t+1}})\right)\boldsymbol{e}_1. \tag{21}$$

The parameter h scales the uncertainty relative to the deterministic contribution and the first unit vector $\boldsymbol{e}_1$ ensures that the feedbacks are driving atmospheric temperature. In contrast to the earlier model, I use the long-run risk process in $t+1$ on the right side of equation (21): the process governs both conditional expectations of τ_{t+1} and

[25] Otherwise, we would discuss the policy impact of changes in expected temperature dynamics rather than the impact of uncertainty. In addition, negative realization of generalized temperature $\tau_{i,t}$ would imply nonsensical realization of real temperature. The long-run risk model in asset pricing gives rise to some nonsensical negative realizations of the variance ${\sigma_{t+1}^M}^2$. This fact is well known, yet it is widely used as an approximate model with a closed-form solution, assuming that the actual calibration of the model makes these realizations of second-order importance. The issue with temperature is more serious. To keep temperature expectations (log expectations of $\tau_{i,t}$) close to the deterministic evolution, the model has to be de-biased. Yet, any realization of $\tau_{i,t} = 0$ would cause an infinitely negative expectation. Therefore, the Gaussian model cannot be de-biased in a meaningful way.

[26] In contrast to the mentioned applications, the model below makes both autoregression and the shape parameter of the underlying gamma distribution state dependent. I thereby use the fact that the underlying cumulant-generating function is linear not only in last period's state but also in the shape parameter.

the one-step-ahead conditional volatility.

The relevant information about the process is summarized in the conditional expectation and variance

$$\mathbb{E}_t\, z_{t+1} \quad = \gamma^z z_t + \epsilon(c)\left(\tfrac{M_{1,t}+G_t}{M_{pre}} - \eta_\tau\right) \tag{22}$$

$$\mathbb{V}\mathrm{ar}_t\, z_{t+1} = \mathbb{V}\mathrm{ar}_t\, y_{t+1} = c\Big[2\gamma^z y_t + \left(\tfrac{M_{1,t}+G_t}{M_{pre}} - \eta_\tau\right)\Big]. \tag{23}$$

Equation (F.2) governs the conditional volatility of atmospheric temperature. As in the previous section, it grows in the perturbation of the climate system captured by $\frac{M_{1,t}+G_t}{M_{pre}} - \eta_\tau$ and is autoregressive. The calibration parameter c scales the variance. Equation (22) characterizes the conditional expectations of atmospheric temperature. It is autoregressive with persistence γ^z; a high persistence will once again capture long-run risk mimicking epistemological uncertainty. In contrast to the earlier model, expectations are biased upwards by a term proportional to $\epsilon(c)$ and the deviation from the pre-industrial equilibrium. This upward bias adjusts the temperature expectations $T_{1,t} = \frac{1}{\xi_1}\log(\tau_{1,t})$ to their deterministic trajectory.[27]

In summary, temperature uncertainty is governed by a skewed stochastic long-run risk process that will deliver a reasonable fit to scientific data about temperature uncertainty. The parameter γ^z captures the persistence of the uncertainty and the parameter c scales its variance. Uncertainty increases endogenously with the perturbation of the climate system. The term $\epsilon(c)$ reflects the non-linearity of the temperature process and adjusts the expected temperature evolution to the deterministic evolution.

Proposition 6 *The stochastic processes for carbon and temperature defined in section 4 change the deterministic SCC^{det} stated in Proposition 3 to*

$$SCC^{unc} = SCC^{det}\Big(1+\theta_\tau\Big)\Big(1+\theta_M\Big(\theta_M^*(1+\theta_\tau)\Big)\Big)$$

$$\text{with } \theta_\tau = \frac{h}{\sigma^{forc}}\,\frac{\epsilon(c)+\theta_\tau^*}{1-\beta\gamma^z} \;\approx\; \frac{h}{\sigma^{forc}}\frac{1}{1-\beta\gamma^z}\left(\epsilon(c)+\frac{1}{2}\,\frac{1+\beta\gamma^z}{1-\beta\gamma^z}\,F\right) \tag{24}$$

and $F = \alpha\varphi_{\tau,1}^{det}\frac{ch}{1-\beta\gamma^z}$. Here, $\varphi_{\tau,1}^{det}$ is the shadow value of $\tau_{1,t}$ under certainty[28] *and $\theta_M(\cdot)$*

[27] Temperature is a concave transformation of the state τ. Thus, a mean-zero shock to $\tau_{1,t+1}$ would reduce the expectation of $T_{1,t+1}$ below its deterministic value. The higher the uncertainty, the higher the bias. Therefore, the de-biasing increases proportional to $\frac{M_{1,t}+G_t}{M_{pre}}-\eta_\tau$, which increases the variance of the process as a consequence of perturbing the climate system. In addition, the exogenous parameter c scales the variance and I write $\epsilon(c)$ to make explicit that a different calibration of c also changes the calibration of ϵ achieving an approximate de-biasing of the temperature expectations.

[28] It is $\varphi_{\tau,1}^{det} = -\frac{\xi_0}{1-\beta\kappa}[(1-\beta\boldsymbol{\sigma})^{-1}]_{1,1}$.in utils. Note that $\alpha\varphi_{\tau,1}^{det} >$ because $\alpha < 0$.

and θ_M^* are defined in Proposition 5. For the exact solution

$$\theta_\tau^* = \frac{-\log\left(1-F\ (1+\theta_\tau^\dagger)\right)}{F} - 1 \quad \textit{with} \quad \theta_\tau^\dagger = \beta\gamma^z \frac{1+F-\sqrt{(1-F)^2-4F\frac{\beta\gamma^z}{1-\beta\gamma^z}}}{1-F+\sqrt{(1-F)^2-4F\frac{\beta\gamma^z}{1-\beta\gamma^z}}} \ .$$

The long-run risk about the temperature further increases the SSC. The increase $\theta_\tau > 0$ enters the SCC twice. First, it directly increases the SCC by the fraction θ_τ over its deterministic value. Second, it increases the contribution resulting from the long-run risk over carbon flows because the long-run risk about temperature changes the valuation of the carbon flows that drive the valuation of carbon risk. We know from Proposition 6 that the overall contribution from carbon flow uncertainty θ^M increases convexly in its argument, now the interaction of carbon flow and temperature risk. In addition, the policy impacts of long-run risk over carbon flows and temperature reinforce each other because they appear as multiplicative factors: $(1+\theta_\tau)\,(1+\theta_M(\cdot)) = 1+\theta_\tau+\theta_M(\cdot)+\theta_\tau\theta_M(\cdot)$. Thus, a first conclusion is that both risks are mutually aggravating.

Equation (24) spells out the components driving the policy impact of temperature risk. The factor $\frac{h}{\sigma^{forc}}$ weighs the stochastic forcing contribution relative to the deterministic contribution; the higher the stochastic relative to the deterministic contribution, the higher the factor increasing the stochastic over the deterministic SCC. The first core component of the temperature contribution is $\epsilon(c)$. It results directly from the non-linearity of the temperature dynamics: a high temperature realization is more painful than a low temperature realization, even in the absence of (intrinsic) risk aversion. This contribution increases in the uncertainty level c, where both c and ϵ will be calibrated to scientific data. The second core contribution θ_τ^* results from the interactions of risk and risk aversion. Both contributions are amplified by the persistence γ^z of the stochastic process, and more so for a more patient decision maker (β large).

Equation (24) also presents an approximation of the risk and risk aversion contribution θ_τ^* that captures its essence. First, an additional dependence on the discount factor weighted persistence emphasizes that this contribution will be highly sensitive to the combination of the decision maker's patience and the long-run risk's persistence. Second, at the heart of the contribution lies the term $F = \alpha\varphi_{\tau,1}^{det}\frac{ch}{1-\beta\gamma^z}$, which is once again driven by the risk aversion weighted shadow value under certainty. The (intrinsic) risk aversion parameter α reflects that the excess of Arrow–Pratt risk aversion over consumption smoothing drives the risk premium. As in the case of carbon risk, whenever the shadow cost of a temperature increase is high under certainty ($\varphi_{\tau,1}^{det}$ large), the long-run risk contribution to the SCC will grow relatively more relevant than the de-

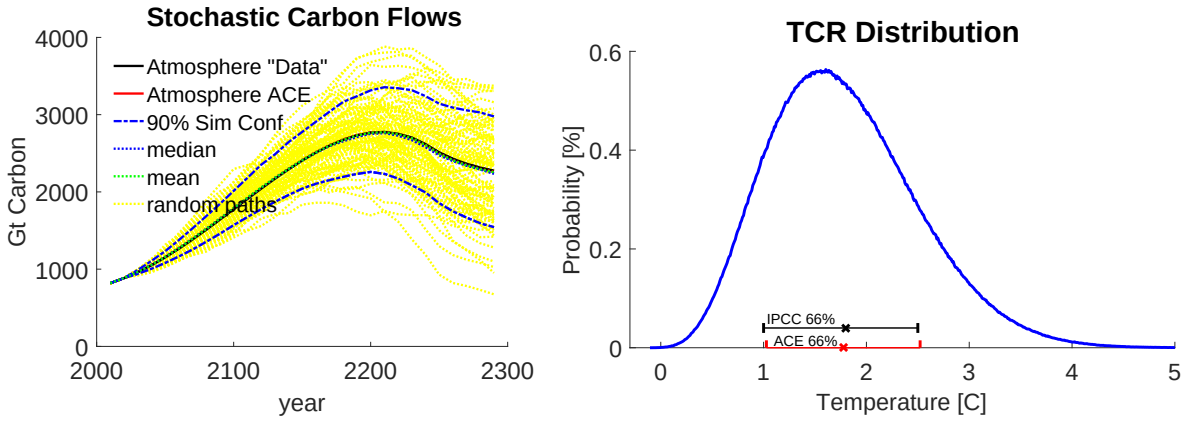


Figure 4: left panel shows atmospheric carbon under DICE's business as usual scenario given carbon flow uncertainty ($\delta^{\sigma x} = 1$, $\delta^{Mx} = \delta^{M\sigma} = 25$, $\gamma^x = \gamma^\sigma = 0.95$, $\eta_M = 0.8$). The deterministic DICE evolution (5-year time steps, "Data"), the deterministic ACE evolution (10-year time steps), and the mean and the median of 1000 uncertain trajectories are hardly distinguishable. The right panel gives the probability distribution of the transient climate response (TCR) in ACE and compares its mean ("x") and 66% probability interval to the IPCC (2013) data. The parameters are $\gamma^z = 0.95$, $h = 0.23$, $\eta = 0.8$, $c = 0.21$, and $\epsilon = 0.05$, and the resulting TCR distribution exhibits the typical moderate positive skew.

terministic contribution ($\frac{SCC_t^{unc}}{SCC_t^{det}}$ increases). Furthermore, the contribution increases in the variance scaling parameter c and the scaling parameter of the stochastic feedbacks h, making the contribution convex in h which already appears in equation (24). More generally, the exact solution shows that the contribution θ_τ^* is itself convex rather than linear in F.

4.3 Quantitative Insights

Appendix G calibrates the stochastic processes. The right panel of Figure 4 closely resembles the probability distributions and statistical information about the transient climate response (TCR) provided by the IPCC (2013). The TCR better characterizes the climate response for the coming century than does the climate sensitivity, which characterizes the medium- to long-term response over a few centuries. It is harder to find good probabilistic information on carbon flow uncertainty. My calibration uses the magnitude of the missing sink and Joos et al.'s (2013) study subjecting 18 different carbon cycle models to a 5000Gt carbon pulse. The approach is rule of thumb. The left graph in Figure 4 translates the calibration into the resulting uncertainty over atmospheric carbon dioxide concentrations along the DICE 2013 business as usual scenario. Section 4.4 discusses the choice of autoregressive persistence in view of epistemological uncertainty.

Carbon flow uncertainty adds a negligible risk premium ($< 1\%$). Uncertainty over conditional expectations contributes most of this premium, followed by stochastic volatility of conditional expectations ($\delta^{\sigma x} = 1$). The direct channel of stochastic volatility on carbon flows hardly contributes ($< 1\%$ of the already tiny premium). This result differs from the long-run risk's application to asset pricing, where the stochastic volatility channel is generally considered the most important. Here, the most relevant information is the long-run expected carbon flow, not its long-run volatility. In contrast, temperature uncertainty adds a substantial risk premium of 25%. The nonlinearity of temperature dynamics captured by $\epsilon(c)$ accounts for about half of this premium. The risk-aversion-sensitive contribution θ_τ^* contributes the other half.

Proposition 6 found that the temperature contribution is highly sensitive to the discount factor weighted persistence of the long-run risk. Drupp et al.'s (2018) expert survey suggests a rate of pure time preference of $\rho = 0.5\%$ (median value) instead of the 1.42% resulting from the standard calibration. As explained in section 3.2 such a lower choice is also supported by emprical application of the long-run risk model to asset pricing (or for normative purposes). This reduction of impatience increases the risk premium to 180% and the tax to $200\frac{USD}{tCO_2}$. Now the risk-aversion-sensitive part, θ_τ^*, contributes the lion's share of the premium (150%). The carbon flow uncertainty multiplier θ_M adds 2% or $4\frac{USD}{tCO_2}$. The difference in impact of uncertainty over carbon versus temperature dynamics is a result of the highly concave transformation of emissions into radiative forcing and warming. The welfare change is strongly convex in temperature, but hardly convex in changes of CO_2. The finding emphasizes the great importance of distinguish temperature dynamics from carbon dynamics when evaluating the impact of uncertainty. Lowering or increasing risk aversion to $\alpha = -0.5$ or $\alpha = -1.25$ (see section 2.4) decreases the optimal tax to $125\frac{USD}{tCO_2}$ or increases it to $300\frac{USD}{tCO_2}$ in the setting with the reduced risk-free rate, and to $35\frac{USD}{tCO_2}$ or $39\frac{USD}{tCO_2}$ in the base case.

Relating these findings to earlier numeric findings, Jensen & Traeger (2013) find a climate sensitivity risk premium of 25% for a pure rate of time preference of 1.5%, similar to my base case, and Kelly & Tan (2015) find a risk premium of 24% using a rate of time preference of 5% in combination with fat tails. Both models use a simplified climate model including a carbon decay approximation to the carbon cycle to save numerically expensive state variables. Rudik & Lemoine (2017) find a risk premium of 10% using a Smolyak grid in an implementation without state space reduction.[29]

[29]Kelly & Tan (2015) find that, in the presence of fat tails, the ability to learn can matter substantially for the SCC, and the risk premium can be up to 60% in scenarios without learning. In

4.4 Welfare, Uncertainty, and Learning

The section discusses the relation between two different conceptualizations of uncertainty. First, nature's stochastic physical processes lead to an uncertain evolution of the future climate. Second, epistemological uncertainty reflects the limited understanding of natural processes by the scientific community. The most important difference is that epistemological uncertainty can potentially be reduced over time as scientists gain a better understanding. The section will also flesh out an analytic formula for the welfare loss, showing a discounting sensitivity that increases in the power of the moments of the uncertainty distribution (variance, skewness, kurtosis,...).

For a simpler closed-form tractability, this section simplifies the stochastic evolution of the climate variables.

$$\boldsymbol{M}_{t+1} = \boldsymbol{\Phi}\boldsymbol{M}_t + (\textstyle\sum_{i=1}^{I^d} E_{i,t} + E_t^{exo})\boldsymbol{e}_1 + \boldsymbol{\epsilon}_t^M + \boldsymbol{\nu}_t^M \tag{25}$$

$$\boldsymbol{\tau}_{t+1} = \boldsymbol{\sigma}\boldsymbol{\tau}_t + \sigma^{forc}\frac{M_{1,t} + G_t}{M_{pre}}\boldsymbol{e}_1 + \boldsymbol{\epsilon}_t^\tau + \boldsymbol{\nu}_t^\tau \ . \tag{26}$$

The vectors $\boldsymbol{\epsilon}_t^M$ and $\boldsymbol{\epsilon}_t^\tau$ reflect uncertainty, the vectors $\boldsymbol{\nu}_t^M$ and $\boldsymbol{\nu}_t^\tau$ reflect measurement error and are non-zero only in the Bayesian learning model, where they determine the speed of learning. Henceforth, the climate state j will label any of the carbon reservoirs or temperature layers $M_1, ..M_m, \tau_1, ...\tau_l$.

I compare the uncertainty dynamics of an autoregressive shock process to that of a Bayesian learning model. To ease the comparison, I write these processes in a slightly unusual way. A first-order autoregressive shock introduces one-step-ahead uncertainty for the random variable ϵ_t^j. The mean of ϵ_t^j, denoted μ_t^j, follows the equation of motion

$$\mu_{t+1}^j = \gamma^j \mu_t^j + \chi_t^j \ , \tag{27}$$

where $0 \leq \gamma \leq 1$ and χ_t^j is a sequence of iid mean-zero shocks. The one-step-ahead variance of ϵ_t^j is given by the variance of χ_t^j (and similarly for higher moments).

A Bayesian learning model with normally distributed measurement error $\nu_t^j \sim N(0, \sigma_\nu^2)$ (likelihood) and prior $\epsilon_t^j \sim N(\mu_t^j, \sigma_{\epsilon,t}^2)$ gives rise to the following dynamics

contrast, Jensen & Traeger (2013) find that anticipated learning, in their model without fat tails, has no impact on the present SCC. Rudik & Lemoine (2017), who also use a time preference of 1.5%, find that without anticipated learning their uncertainty premium falls to 1%. See section Section 4.4 for a discussion of learning.

of the mean[30]

$$\mu^j_{t+1} = \mu^j_t + \chi^j_t \quad \text{with } \chi^j_t \sim N\left(0, \tfrac{\sigma^4_{\epsilon,t}}{\sigma^2_{\epsilon,t}+\sigma^2_{\nu,t}}\right). \tag{28}$$

Writing the updating equation in the form of equation (28) emphasizes the close similarity between learning and a (persistent) AR(1) shock. The important conceptual difference from the autoregressive model is that the variance of ϵ^j_t does not vanish with period t information: $\sigma^j_{\epsilon,t} \equiv \mathbb{V}\text{ar}[\epsilon^j_t | I_t] > 0$. Yet, what matters to the decision maker is the one-step-ahead forecast uncertainty, which is similar for both settings. The only difference between equations (27) and (28) is that the conditional expectation of the Bayesian model exhibits full persistence and a prescribed evolution of the shock variance that falls over time, $\sigma^2_{\epsilon,t+1} = \frac{\sigma^2_{\nu,t}\sigma^2_{\epsilon,t}}{\sigma^2_{\nu,t}+\sigma^2_{\epsilon,t}}$.[31]

Proposition 7 *Let uncertainty in equations (25-26) affect state j.*
(1) A normally distributed first-order autoregressive process ϵ_t with one-step-ahead variance σ^2 implies the welfare loss

$$\Delta W^{AR}_{normal} = \textstyle\sum_{t=0}^{\infty} \beta^{t+1} \left(\frac{\beta}{1-\gamma^j\beta}\right)^2 \alpha\varphi_j^2 \quad \frac{\sigma^2}{2} \quad = \frac{\beta}{1-\beta}\left(\frac{\beta}{1-\gamma^j\beta}\right)^2 \alpha\varphi_j^2 \, \frac{\sigma^2}{2}.$$

(2) A Bayesian learning model with normally distributed prior $\epsilon_t \sim N(\mu_{\epsilon,t}, \sigma^2_{\epsilon,t})$ and measurement error $\nu_t \sim N(0, \sigma^2_{\nu,t})$ implies the welfare loss

$$\Delta W^{Bayes} = \textstyle\sum_{t=0}^{\infty} \beta^{t+1} \left(\frac{\Omega_t}{1-\beta}\right)^2 \alpha\varphi_j^2 \, \frac{\sigma^2_{\epsilon,t}+\sigma^2_{\nu,t}}{2}$$

with $\Omega_t \equiv \frac{\sigma^2_{\epsilon,t}}{\sigma^2_{\nu,t}+\sigma^2_{\epsilon,t}} + (1-\beta)\, \frac{\sigma^2_{\nu,t}}{\sigma^2_{\nu,t}+\sigma^2_{\epsilon,t}}$.
(3) A first-order autoregressive process ϵ_t with arbitrarily[32] distributed iid shocks χ_t implies the welfare loss

$$\Delta W^{AR}_{general} = \tfrac{\beta}{1-\beta}\, \tfrac{1}{\alpha}\, G_\chi\left(\tfrac{\beta}{1-\gamma^j\beta}\alpha\varphi_j\right) = \tfrac{\beta}{1-\beta}\tfrac{1}{\alpha}\textstyle\sum_{l=1}^{\infty} \kappa_l \frac{1}{l!}\left(\frac{\beta}{1-\gamma^j\beta}\alpha\varphi_j\right)^l, \tag{29}$$

where κ_l are the cumulants of the iid shock χ_t.

[30] The standard way of writing the Bayesian updating equation for the mean is $\mu^j_{t+1} = \frac{\sigma^2_{\epsilon,t}}{\sigma^2_{\epsilon,t}+\sigma^2_{\nu,t}}\mu^j_t + \frac{\sigma^2_{\nu,t}}{\sigma^2_{\epsilon,t}+\sigma^2_{\nu,t}} z_t$ with observation $z \sim N(\mu_{\epsilon,t}, \sigma^2_{\epsilon,t} + \sigma^2_{\nu,t})$. Defining $\chi^j_{1,t} = \frac{\sigma^2_{\nu,t}}{\sigma^2_{\epsilon,t}+\sigma^2_{\nu,t}}(z_t - \mu_{\epsilon,t})$ delivers equation (28). Note that the observational variable z is defined in equations (25-26). For example, in the case of uncertain atmospheric carbon content, the observation z is $M_{t+1} - \mathbf{\Phi}_{1,\cdot}\boldsymbol{M}_t - (\sum_{i=1}^{I^d} E_{i,t} + E^{exo}_t)$.

[31] Kelly & Kolstad (1999) and Karp & Zhang (2006) employ such a simple Bayesian learning model for the assessment of climate change feedbacks and damages. Kelly & Tan (2015) analyze learning speed when climate sensitivity is fat tailed.

[32] I assume that the shock χ_t has a finite cumulant-generating function.

In both models (i) and (ii), the welfare loss is proportional to (intrinsic) risk aversion α and the square of the state's shadow value under certainty, e.g., $\varphi_j = \varphi_{\tau,1}$ if uncertainty governs atmospheric temperature. The welfare loss is also proportional to the variance. Assumed constant in the AR model, this variance falls over time in the Bayesian learning model, where it is the sum of measurement error and Bayesian prior. In the AR model, the assumption of constant variance collapses the infinite sum into a factor $(1-\beta)^{-1}$, yielding a sensitivity to time preference similar to the one observed for the carbon tax under certainty (Proposition 3). In addition, the welfare loss is proportional to the factor $(1-\gamma\beta)^{-2}$: a high uncertainty persistence γ makes the result even more sensitive to the choice of pure time preference.

The *Bayesian learning* model swaps this factor[33] against the factor $\left(\frac{\Omega_t}{1-\beta}\right)^2 = (1-\beta)^{-2}\left(\frac{\sigma^2_{\epsilon,t}}{\sigma^2_{\nu,t}+\sigma^2_{\epsilon,t}} + (1-\beta)\,\frac{\sigma^2_{\nu,t}}{\sigma^2_{\nu,t}+\sigma^2_{\epsilon,t}}\right)^2$. The term Ω_t is a weighted mean of unity (weighted by prior uncertainty) and $1-\beta$ (weighted by the measurement error). Initially, when the prior uncertainty is large ($\sigma_{\epsilon,t} \gg \sigma_{\nu,t}$), the time preference sensitivity is that of a fully persistent AR shock with $\gamma = 1$: every update implies a revision of the long-run future. If the decision maker is patient, this long-term update moves her welfare substantially. As she becomes more assertive of her environment, uncertainty reduces to the prior and, once $\sigma_{\epsilon,t} \ll \sigma_{\nu,t}$, the term Ω_t cancels the time sensitivity $1-\beta$: post-learning the iid error σ^2_ν has no more long-term repercussions.

This comparison between an autoregressive uncertainty model and a Bayesian learning model also informs my choice of the $\gamma-$parameters specifying uncertainty persistence in the models of section 4.

$$(1-\gamma\beta)^{-2} \overset{!}{=} \left(\frac{\Omega_t}{1-\beta}\right)^2 \quad \Leftrightarrow \quad \gamma \overset{!}{=} \left(1+(1-\beta)\,\frac{\sigma^2_{\nu,t}}{\sigma^2_{\epsilon,t}}\right)^{-1} \approx 1-\rho\frac{\sigma^2_{\nu,t}}{\sigma^2_{\epsilon,t}}. \tag{30}$$

Evaluating this back-of-the-envelope correspondence for early periods suggests $\gamma \approx 0.99$: a 10-year time step with a 1% annual discount rate and a (decadal) measurement error that is 10% of the prior uncertainty. Once the magnitude of the prior uncertainty approaches that of the measurement error, the corresponding γ drops to 0.9. As a result, I chose $\gamma = 0.95$ as an intermediate value for my quantitative assessments in section 4.[34]

[33] I ignore the term β^2 in this comparison as I am interested in the sensitivity for $\beta \to 1$.

[34] The relation in equation (30) could suggest that a lower rate of time preference should imply a higher corresponding γ. This insight would further increase the Bayesian learning model's time preference sensitivity. Yet, with a lower rate of pure time preference, the comparison in equation (30) should also pay increasing attention to later periods where the corresponding γ is smaller than in earlier periods because of learning.

The *general autoregressive* shock model in case (iii) of Proposition 7 shows that higher-order moments of the uncertainty contribute proportionally to higher orders of the risk-aversion-weighted shadow value of the state. It also shows that for high uncertainty persistence, the sensitivity to time preference of the welfare loss increases in the power of the contributing moment: the contribution of the kurtosis is more sensitive to time preference and (intrinsic) risk aversion than the contribution of skewness, which is more sensitive than the contribution of the variance, which is more sensitive than the contribution of the mean. In the colloquial use of "fat tailedness" prevalent in the climate change debate, equation (29) suggests that the fatter the tail the more relevant the calibration of time preference.

Finally, carbon does not decay; it merely travels between different reservoirs. Therefore, we cannot look at independent individual shocks across states. The stochastic carbon flow between the atmosphere and a sink is a perfectly negatively correlated shock to adjacent layers. Such uncertainty gives rise to the formula in Proposition 7 with the shadow value $\varphi_j = \varphi_{M,1} - \varphi_{M,2}$ (see the proof of the proposition for details).

5 Conclusions

ACE is a stochastic analytic IAM closely resembling the core components of numeric models used in policy advising, in particular, the widespread DICE model. ACE fleshes out the quantitative policy contributions of the different parts of IAMs in a simple formula, and it complements current studies with a careful uncertainty analysis.

Under certainty, a standard calibration delivers an optimal carbon tax of $30\frac{USD}{tCO_2}$ or 63 cents per gallon. Here, carbon dynamics contribute a factor of 4 and temperature dynamics contribute a reduction of 40%. The standard calibration overestimates the risk-free discount rate. The median response of Drupp et al.'s (2018) expert survey suggests a pure time preference of 0.5%, almost 1% lower than in the standard calibration. The long-run risk literature finds that even lower rates explain observed asset prices well once risk aversion is disentangled from intertemporal consumption smoothing. Such a reduction in time preference increases the optimal carbon tax under certainty to over $70\frac{USD}{tCO_2}$, with carbon dynamics contributing a factor of 8. I show that the high sensitivity to pure time preference derives from the fact that carbon does not decay but instead cycles through different reservoirs. This sensitivity increases further under population growth with population-weighted average consumption, but it decreases when moving away from ACE's assumptions of a unit elasticity for the

intertemporal elasticity of substitution and for the elasticity of damages to output.

The stochastic ACE merges the linear-in-state systems of recent analytic IAMs with affine stochastic (nonlinear-in-state) processes used in asset pricing. In contrast to much of the asset pricing literature, climate change policy must be concerned with the endogeneity of the risk, i.e., the implications of our current policies on future uncertainty. ACE shows that the huge uncertainties surrounding future atmospheric carbon accumulation hardly affect the optimal policy. The marginal effect of a ton of carbon on temperature increase is inversely proportional to the amount of carbon already present. Together with decreasing marginal utility, this marginal decrease of carbon-induced warming balances the convexity of temperature damages. Risk aversion still contributes a premium, but it is small. The uncertainty over conditional expectations of carbon buildup contributes more than does the uncertainty over carbon flow fluctuations (stochastic volatility).

The main quantitative concern is the uncertain temperature response to a given carbon concentration. In the standard calibration of time preference, uncertainty contributes an additional 25% to the deterministic carbon tax. A little more than half of the effect derives from the decision maker's risk aversion. The remainder derives from the mere nonlinearity of welfare in temperature and would prevail even without risk aversion. The applicable measure of risk aversion is not the Arrow–Pratt measure, but a measure of intertemporal or intrinsic risk aversion. This measure characterizes how much more averse a decision maker is to risk than to deterministic consumption fluctuations. The risk premium increases convexly in this risk aversion and in the magnitude and persistence of uncertainty. The relative risk premium also increases convexly in all the contributions to the deterministic carbon tax: the risk premium makes up a larger *share* of the optimal tax when climate change is judged as more severe to start with. Reducing the risk-free discount rate by applying the 0.5% pure rate of time preference increases the risk premium to 180%. That is 180% of the already higher SCC under certainty, delivering an overall optimal tax of $200\frac{USD}{tCO_2}$. Varying risk aversion within reasonable bounds varies this tax between $125\frac{USD}{tCO_2}$ and $300\frac{USD}{tCO_2}$.

Taking the structure and delays of IAMs seriously, ACE finds that the risk premium is even more sensitive to time preference than the deterministic tax is, and this sensitivity seems to increase with the fatness of the tail. Unfortunately, this result contradicts more stylized assessments that suggested we could leave the difficult field of discounting sensitivities behind by focusing on the dominant importance of climate change uncertainty. ACE also points out that the US government's prescription of the consumption discount rate in deriving the federal SCC leaves a flexibility in pick-

ing pure time preference that easily changes the optimal carbon tax by an order of magnitude.

Finally, ACE delivers a simple formula for the emission response to a carbon tax under general assumptions about the production sectors. It shows how a carbon tax trades off against the Hotelling rent of a scarce resource in a full-fledged IAM. Technological progress will only reduce emissions if it reduces the production elasticity of emissions, and the formula shows that also such an innovation will be partly crowded out by changes in the Hotelling rent. Thus, innovations reducing resource dependence will be most effective in low scarcity resource sectors like coal, especially if we do not implement taxes in the magnitude implied by a low risk-free rate calibration.

ACE is a quantitative analytic model. As such, it has limitations in the choice of functional forms. I have shown that the chosen functional forms fit the real-world climate change problem sufficiently well to have a serious quantitative discussion of the implications of model assumptions and sensitivities, and to enable a broad audience to familiarize itself with a full-fledged IAM of climate change. ACE permits uncertainty analysis in closed form and, more generally, by reducing the curse of dimensionality in dynamic programming to a simple system of simultaneous equations. Extensions of ACE can deal with regional production and temperature and, thus, can integrate uncertainty and solve games in complex regional settings of geoengineering and climate negotiations. ACE's major virtue is to combine quantitative analysis with analytic insight, and to enable stochastic modeling in complex dynamic environments. Any quantitative analytic approach has its limitations in the nonlinearities and interactions it can handle.

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

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.

Cai, Y. & Lontzek, T. S. (forthcoming), 'The social cost of carbon with economic and climate risks', *Journal of Political Economy* .

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

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.

Creal, D. (2017), 'A class of non-gaussian state space models with exact likelihood inference', *Journal of Business and Economic Statistics* **35**(4), 585–597.

Croce, M. M. (2014), 'Long-run productivity risk: A new hope for production-based asset pricing?', **66**, 13–31.

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

Epstein, L. G. & Zin, S. E. (1991), 'Substitution, risk aversion, and the temporal behavior of consumption and asset returns: An empirical analysis', *Journal of Political Economy* **99**(2), 263–86.

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.

Gourieroux, C. & Jasiak, J. (2006), 'Autoregressive gamma processes', *Journal of Forecasting* **25**, 129–152.

Ha-Duong, M. & Treich, N. (2004), 'Risk aversion, intergenerational equity and climate change', *Environmental and Resource Economics* **28**(2), 195–207.

Hambel, C., Kraft, H. & Schwartz, E. S. (2018*a*), 'The carbon abatement game', *SSRN working paper* **3122261**.

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

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* .

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. (2001), 'Taxes and quotas for a stock pollutant with multiplicative uncertainty', *Journal of Public Economics* **82**, 91–114.

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.

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. (2017), 'Carbon taxes and commitment with non-constant time preference', *Working Paper* .

Jensen, S., Mohlin, K., Pittel, K. & Sterner, T. (2015), 'An introduction to the green paradox: The unintended consequences of climate policies', *Review of Environmental Economics and Policy* **9**(2), 246–265.

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

Jensen, S. & Traeger, C. (2014), 'Optimal climate change mitigation under long-term growth uncertainty - stochastic integrated assessment and analytic findings', *European Economic Review* **69**, 104–125.

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.

Kelly, D. L. & Tan, Z. (2015), 'Learning and climate feedbacks: Optimal climate insurance and fat tails', *Journal of Environmental Economics and Management* **72**, 98–122.

Le, A., Singleton, K. J. & Dai, Q. (2010), 'Discrete-time affineq term structure models with generalized market prices of risk', *Review of Financial Studies* **23**, 2184–2227.

Li, X., Narajabad, B. & Temzelides, T. (2016), 'Robust dynamic energy use and climate change', *Quantitative Economics* **7**.

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.

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.

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.

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

Rezai, A. & der Ploeg, F. V. (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.

Rudik, I. & Lemoine, D. (2017), 'Managing climate change under uncertainty: Recursive integrated assessment at an inflection point', *Annual Review of Resource Economics* **9**, 117–42.

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

Sinn, H.-W. (2012), *The Green Paradox*, MIT Press, Cambridge.

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.

Tallarini, T. D. (2000), 'Risk-sensitive real business cycles', *Journal of Monetary Economics* **45**(3), 507 – 532.

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. (2014*a*), 'Why uncertainty matters - discounting under intertemporal risk aversion and ambiguity', *Economic Theory* **56**(3), 627–664.

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. (2014*b*), 'Capturing intrinsic risk aversion', *CUDARE Working Paper 1102* .

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. & Withagen, C. (2012), 'Is there really a green paradox?', *Journal of Environmental Economics and Management* **64**(3), 342–363.

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

von Neumann, J. & Morgenstern, O. (1944), *Theory of Games and Economic Behaviour*, Princeton University Press, Princeton.

Weil, P. (1990), 'Nonexpected utility in macroeconomics', *The Quarterly Journal of Economics* **105**(1), 29–42.

Weitzman, M. L. (2010), 'GHG targets as insurance against catastrophic climate damages', *NBER Working Paper* (16136).

Weitzman, M. L. (2012), 'GHG targets as insurance against catastrophic climate damages', *Journal of Public Economic Theory* **14**(2), 221–244.

Appendix

Part I - Deterministic ACE

A Capital Depreciation

This section derives the capital equation of motion (4) and discusses the implied result that the consumption *rate* is unaffected by climate states. The usual capital accumulation equation, enriched by climage 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[35]

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

Defining the capital growth rate $g_{k,t} = \frac{K_{t+1}}{K_t} - 1$, I obtain the equation of motion for capital (4) stated in the main text.

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 the time step. The crucial result from the assumptions underlying equation (4) is that the investment rate is independent of the climate states. It is the price to pay for an analytic solution. The remainder of this section shows that this price seems small.

Figure 5 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*).[36] It also abandons the assumption of logarithmic utility, further stacking the cards against ACE's

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

[36] The recursive implementation based on the Bellman equation solves for the optimal control rule as a function of the states. Thus, solving the model once immediately delivers 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.

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 (see section 2.4).

The qualitative behavior is the same for all graphs in Figure 5. 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 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 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.

B Transformation to Linear-in-State Model (helpful for insight & preparation for proofs)

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

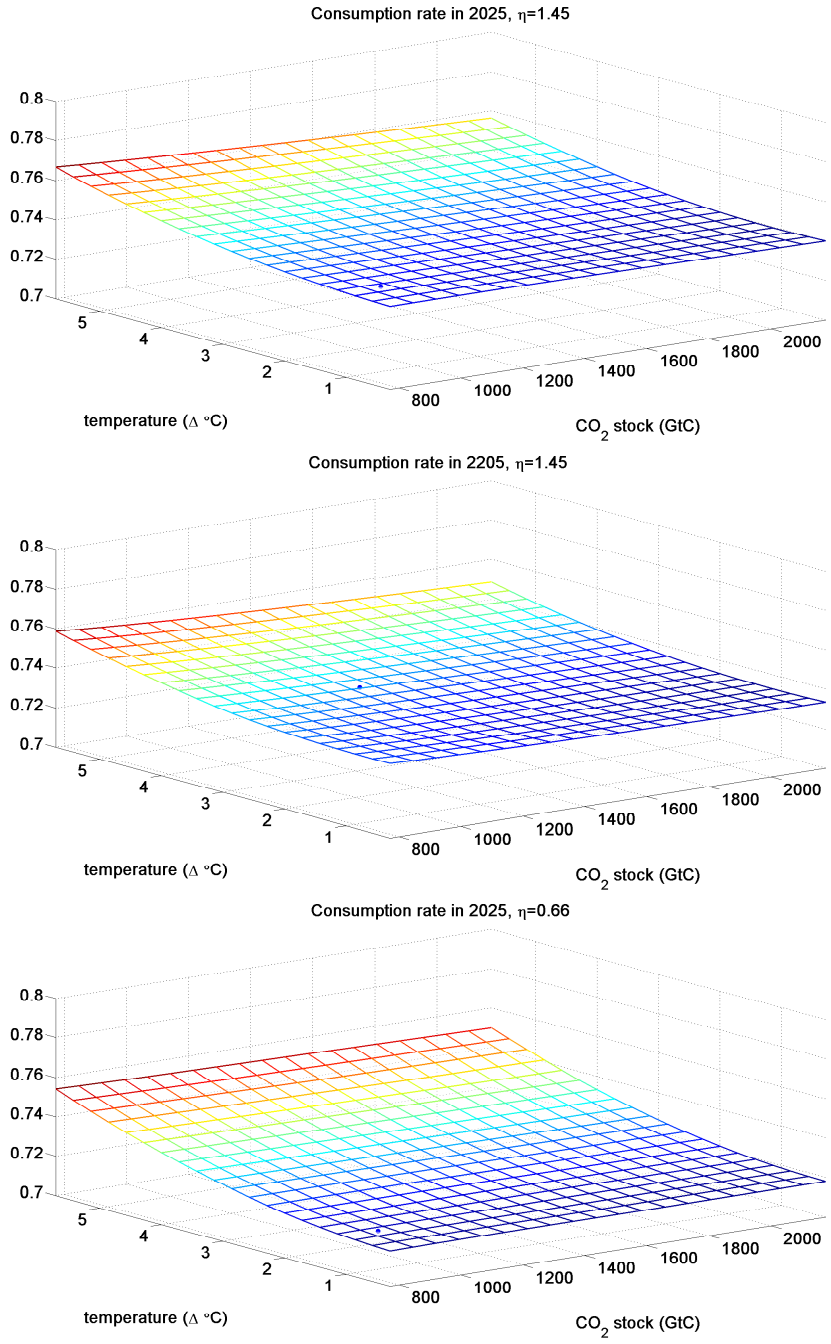


Figure 5: The graphs analyze the climate (in-)dependence of the optimal consumption rate x^* in the wide-spread DICE model, relying on the control rules of the recursive implementation by Traeger (2012*b*). The first two graphs assume the DICE-2013 value $\eta = 1.45$, the third graph follows the long-run 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.

$\boldsymbol{\tau}_t \in \mathbb{R}^L$. 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

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

The Bellman equation in terms of the transformed state variables is

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

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

$$k_{t+1} = a_t + \kappa k_t + \log F(\boldsymbol{A}_t, \boldsymbol{N}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{E}_t) - \xi_0 \tau_{1,t} + \xi_0 + \log(1-x_t)$$
$$+ \log[1 + g_{k,t}] - \log[\delta_k + g_{k,t}] \tag{B.2}$$
$$\boldsymbol{M_{t+1}} = \boldsymbol{\Phi} \boldsymbol{M}_t + \left(\textstyle\sum_{i=1}^{I^d} E_{i,t} + E_t^{exo} \right) \boldsymbol{e}_1 \ . \tag{B.3}$$

Using equation (11), I transform the temperature's equation of motion (7) 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

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

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

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

Finally, the equation of motion for the resource stock is

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

The underlying constraints 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 initial states 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.

C Proofs (Certainty)

C.1 Proof of Proposition 1

1) Sufficiency: I show that the affine value function

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

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 B transformed these assumptions into the linear-in-state-system summarized by equations (B.1-B.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 (B.2)-(B.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 C.1) and the next period states (equations B.2-B.5) into the (deterministic) Bellmann equation (B.1) delivers

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

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

$$\frac{1}{x} - \beta\varphi_k \frac{1}{1-x} = 0 \quad \Rightarrow \quad x^* = \frac{1}{1+\beta\varphi_k} \; . \tag{C.2}$$

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

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

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

$$\Rightarrow \lambda_t^{\mathcal{K}} = \sum_{i=1}^{I_K}(1+\beta\varphi_k)\sigma_{Y,\mathcal{K}_i}(\boldsymbol{A}_t,\boldsymbol{N}_t,\boldsymbol{\mathcal{K}}_t,\boldsymbol{E}_t)$$

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

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

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

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

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

$$(1+\beta\varphi_k)\frac{\frac{\partial F(\boldsymbol{A}_t,\boldsymbol{N}_t,\boldsymbol{\mathcal{K}}_t,\boldsymbol{E}_t)}{\partial E_{i,t}}}{F(\boldsymbol{A}_t,\boldsymbol{N}_t,\boldsymbol{\mathcal{K}}_t,\boldsymbol{E}_t)} = \beta(\varphi_{R,i,t+1}-\varphi_{M,1}) \tag{C.3}$$

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

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

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

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

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

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

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

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

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 (C.2) results in the optimal consumption rate $x^* = 1-\beta\kappa$. Then equation equation (C.4) turns into the condition

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

This condition will be satisfied by picking the sequence $\varphi_0, \varphi_1, \varphi_2, ...$ Equation (C.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 $\varphi_0, \varphi_1, \varphi_2, \varphi_3, ...$.

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

$$T_{0,t} = \frac{s}{\eta}F_t = \frac{s}{\log 2}\log\frac{M_{1,t}+G_t}{M_{pre}}\ , \tag{C.10}$$

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

$$\begin{aligned} &T_{1,t+1} = \mathfrak{M}_1(T_{0,t}\ ,\ T_{1,t}\ ,\ T_{2,t}) = f^{-1}[\sigma_{1,0}f(T_{0,t}) + \sigma_{1,1}f(T_{1,t}) + \sigma_{1,2}f(T_{2,t})] \\ &\Leftrightarrow f(T_{1,t+1}) = \sigma_{1,0}f\left(\frac{s}{\log 2}\log\frac{M_{1,t}+G_t}{M_{pre}}\right) + \sigma_{1,1}f(T_{1,t}) + \sigma_{1,2}f(T_{2,t}). \end{aligned} \tag{C.11}$$

First, equation (C.11) 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 (C.11) 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 $\xi_1 = \log\frac{2}{s}$.[37] Consequently, the generalized 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

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

[37]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.

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 (4) 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 $T_{1,t} = 0$.

C.2 Proof of Proposition 3

Proof of Part (1): The SCC is the negative of the shadow value of atmospheric carbon expressed in money-measured consumption units. Inserting equation (C.5) for the shadow value of log-capital and (C.7) for the shadow value of atmospheric temperature (first entry of the vector) into equation (C.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 $\mathbf{1} - \beta\boldsymbol{\Phi}$ 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 $(1-\beta)^{-1}$.

C.3 Proof of Proposition 4

Extending numerator and denominator of the expression for optimal emissions in equation (C.3) by consumption $x^* Y_t^{net}$ (the inverse of marginal utility) yields

$$E_{i,t} = \frac{(1+\beta\varphi_k) x^* Y_t^{net} \sigma_{Y,E_i}(\boldsymbol{A}_t, \boldsymbol{N}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{E}_t)}{\beta(\varphi_{R,i,t+1} x^* Y_t^{net} - \varphi_{M,1} x^* Y_t^{net})} = \frac{Y_t^{net} \sigma_{Y,E_i}(\boldsymbol{A}_t, \boldsymbol{N}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{E}_t)}{HOT_{i,t} + \beta SCC}$$

with $HOT_{i,t} \equiv \beta\varphi_{R,i,t+1} x^* Y_t^{net} = \varphi_{R,i,t} x^* Y_t^{net} = \frac{\varphi_{R,i,t}}{u'(C_t)}$ being the marginal consumption value of a unit of the resource $E_{i,t}$.

I now proof the statement in the text. Let the resource elasticities be constant apart from a potential dependence on exogenous technological change: $\sigma_{Y,E_i}(\boldsymbol{A}_t)$. I compare the two emission scenarios with with differing shadow price of carbon $\varphi_{M,1}$, e.g., because of a different damage coefficient or a different climate dynamics. I denote the variables of the scenario with the higher absolute shadow value of carbon by a tilde ($\tilde{}$): the shadow value of carbon changes by assumption and the shadow value of the resource as well as emissions change in response (all else equal). I assume that resource i remains scarce also in the economy with the higher shadow cost of carbon (otherwise emissions response will only be stronger). Then, the resource constraint has to be satisfied in both scenarios and imposes

$$\begin{aligned} R_0 &= \sum_{t=0}^{\infty} E_{i,t} = \sum_{t=0}^{\infty} \frac{(1+\beta\varphi_k)\sigma_{Y,E_i}(\boldsymbol{A}_t)}{\beta^{-t}\varphi_{R,i,0} - \beta\varphi_{M,1}} \\ &= \sum_{t=0}^{\infty} \tilde{E}_{i,t} = \sum_{t=0}^{\infty} \frac{(1+\beta\varphi_k)\sigma_{Y,E_i}(\boldsymbol{A}_t)}{\beta^{-t}\tilde{\varphi}_{R,i,0} - \beta\tilde{\varphi}_{M,1}}. \end{aligned}$$

I proof $\tilde{E}_{i,0} < E_{i,0}$ by contradiction. Assume

$$\begin{aligned}
\tilde{E}_{i,0} > E_{i,0} &\Leftrightarrow \frac{(1+\beta\varphi_k)\sigma_{Y,E_i}(\boldsymbol{A}_0)}{\tilde{\varphi}_{R,i,0} - \beta\tilde{\varphi}_{M,1}} > \frac{(1+\beta\varphi_k)\sigma_{Y,E_i}(\boldsymbol{A}_0)}{\varphi_{R,i,0} - \beta\varphi_{M,1}} \\
&\Leftrightarrow \varphi_{R,i,0} - \beta\varphi_{M,1} > \tilde{\varphi}_{R,i,0} - \beta\tilde{\varphi}_{M,1} \\
&\Leftrightarrow \varphi_{R,i,0} - \tilde{\varphi}_{R,i,0} > \beta(\varphi_{M,1} - \tilde{\varphi}_{M,1}) \quad [> 0 \text{ by assumption}] \\
&\Leftrightarrow \varphi_{R,i,0} - \tilde{\varphi}_{R,i,0} > \beta^{t+1}(\varphi_{M,1} - \tilde{\varphi}_{M,1}) \quad \forall t \in \mathbb{N} \\
&\Leftrightarrow \beta^{-t}(\varphi_{R,i,0} - \tilde{\varphi}_{R,i,0}) > \beta(\varphi_{M,1} - \tilde{\varphi}_{M,1}) \quad \forall t \in \mathbb{N} \\
&\Leftrightarrow \beta^{-t}\varphi_{R,i,0} - \beta\varphi_{M,1} > \beta^{-t}\tilde{\varphi}_{R,i,0} - \beta\tilde{\varphi}_{M,1} \quad \forall t \in \mathbb{N} \\
&\Leftrightarrow \frac{(1+\beta\varphi_k)\sigma_{Y,E_i}(\boldsymbol{A}_t)}{\beta^{-t}\tilde{\varphi}_{R,i,0} - \beta\tilde{\varphi}_{M,1}} > \frac{(1+\beta\varphi_k)\sigma_{Y,E_i}(\boldsymbol{A}_t)}{\beta^{-t}\varphi_{R,i,0} - \beta\varphi_{M,1}} \quad \forall t \in \mathbb{N} \\
&\Leftrightarrow \tilde{E}_{i,t} > E_{i,t} \quad \forall t \in \mathbb{N}.
\end{aligned}$$

Thus, if emissions and resource extraction were higher in the first period under the higher shadow cost of carbon, they would have to be higher in all subsequent periods as well, which violates the resource constraint. As a result, $\tilde{E}_{i,0} < E_{i,0}$ and there exists some time $t^* \in \mathbb{N}$ where emissions become larger in the regime with the higher shadow cost of carbon. Note that the higher future emissions are an immediate result of saving resources in the earlier periods and assuming that the resource remained scarce. If the resource is no longer scarce, then $\tilde{\varphi}_{R,i,0} = 0$ and $\tilde{E}_{i,0} < E_{i,0}$ is obviously satisfied.

D Additional Results (Certainty)

D.1 Social benefits of atmospheric cooling and geoengeneering

The social cost of an atmospheric temperature increase follows similarly from the shadow value of the generalized temperature state $\tau_{1,t}$

$$SC\mathcal{T} = -(1-\beta\kappa)Y_t^{net}\varphi_{\tau,1} = Y_t^{net}\ \xi_0\left[(1-\beta\boldsymbol{\sigma})^{-1}\right]_{1,1}\ .$$

A marginal increase in generalized temperature relates to a temperature increase in degree Celsius as $d\tau_{1,t} = \xi_1 \exp(\xi_1 T_{1,t}) dT_{1,t}$ implying the social cost of a temperature increase in degree Celsius of

$$SCT(T_{1,t}) = Y_t^{net}\ \xi_0\left[(1-\beta\boldsymbol{\sigma})^{-1}\right]_{1,1} \xi_1 \exp(\xi_1 T_{1,t})\ .$$

In contrast to the SCC, the cost of a marginal temperature increase in degree Celsius depends on the prevailing temperature level. This level-dependence reflects the convexity of damages in temperature. The SCT characterizes the marginal cost of warming or the benefit from a marginal cooling of the atmosphere. At a 3°C warming, the SCT in the deterministic standard calibration is about 15 trillion USD for a marginal degree of cooling.

One of many geoengineering suggestions is to mitigate climate change by sequestering carbon into the oceans or other reservoirs like the biosphere. ACE gives a simple back of the envelope calculation of the benefits of rerouting CO_2 emissions into other reservoirs. Pumping a ton of CO_2 into layer i, instead of emitting it into the atmosphere, results in the welfare gain

$$\Delta W^{seq} = \varphi_{M,i} - \varphi_{M,1} = \frac{\beta \varphi_{\tau,1} \sigma^{forc}}{M_{pre}} \Big(\big[(\mathbf{1} - \beta \mathbf{\Phi})^{-1} \big]_{1,i} - \big[(\mathbf{1} - \beta \mathbf{\Phi})^{-1} \big]_{1,1} \Big). \tag{D.1}$$

The bracket on the right hand side captures the discounted sum of the differences in the amount of carbon prevailing in the atmosphere over time when an emission unit is injected into layer i instead of the atmosphere. This intuition is more easily observed using the Neumann series for the expansion

$$\Delta W^{seq} = \frac{\beta \varphi_{\tau,1} \sigma_1^{up}}{M_{pre}} \left(\beta \left[\mathbf{\Phi}_{1,i} - \mathbf{\Phi}_{1,1} \right] + \sum_{n=2}^{\infty} \sum_{j,l} (\beta)^n \mathbf{\Phi}_{1,j} \big(\mathbf{\Phi}^{n-2} \big)_{j,l} \left[\mathbf{\Phi}_{l,i} - \mathbf{\Phi}_{l,1} \right] \right).$$

The first term in the brackets captures the difference between carbon flow from the ocean into the atmosphere $\mathbf{\Phi}_{1,i}$ and the persistence of carbon in the atmosphere $\mathbf{\Phi}_{1,1}$. The second term captures the fraction of carbon reaching the atmosphere after n periods if the carbon initially enters ocean layer i as opposed to entering the atmosphere directly (read right to left). The matrix entry $(\mathbf{\Phi}^{n-2})_{j,l}$ captures the overall carbon flow and persistence from layer l to j after $n-2$ periods. It approaches the stationary distribution given by its (right) eigenvectors (in all columns). In the DICE carbon cycle, the value of sequestering carbon into the intermediate ocean and biosphere is $20 \frac{USD}{tCO_2}$ for the standard time preference calibration, and $30 \frac{USD}{tCO_2}$ for the case of $\rho = 0.5\%$.[38] Note that the reduction of pure time preference increases the value of sequestering carbon into the ocean relatively less than it increases the value of the SCC: sequestering carbon dioxide into the ocean is largely a "delay the problem" solution.

[38]Note that the present model does not explicitly model damages from ocean acidification, which would be an interesting and feasible extension.

I briefly note how the shadow cost of carbon in some reservoir is a simple function of its exchange with the remaining reservoirs

$$\varphi_{M,i} = \beta \frac{\sum_{j \neq i} \varphi_{M,j} \Phi_{j,i} + 1_{i,1} \frac{\varphi_{\tau,1} \sigma_1^{up}}{M_{pre}}}{1 - \beta \Phi_{i,i}}. \tag{D.2}$$

The carbon price in layer i is the sum of carbon prices in the other layers times the flow coefficient capturing the carbon transition into that other layer (generally only positive for the two adjacent layers). The atmospheric carbon price has as an additional contribution ($\delta_{i,1}$ denotes the Kronecker delta): the shadow value of the atmospheric temperature increase. The denominator weighs the sum by the reservoirs discounted persistence $\beta \Phi_{i,i}$.

D.2 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, I find Proposition 3(2)'s root $(1-\beta)^{-1}$ that 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

$$\varphi_{M,1} = \beta \varphi_{\tau,1} \sigma^{forc} M_{pre}^{-1} \big(1 - \beta(1 - \delta_{decay})\big)^{-1}, \tag{D.3}$$

which misses the long-run carbon impact and the SCC's sensitivity to pure time preference. Equation (D.3) somewhat resembles the carbon pricing formula (D.2), where atmospheric carbon persistence is $\Phi_{i,i} = 1 - \delta_{decay}$. Yet, the full equation (D.3) adds

the pricing contributions from the other carbon absorbing layers as, unfortunately, the carbon leaving the atmosphere does not decay.

Finally, I illustrate the value of carbon sequestration in equation (D.1) for the case of the two layer carbon cycle

$$\Delta W^{seq} = \beta\varphi_{\tau,1}\sigma_1^{up} M_{pre}^{-1}\left[1+\beta\delta_{\text{Ocean}\to\text{Atm}} - \beta(1-\delta_{\text{Atm}\to\text{Ocean}})\right]^{-1}.$$

The value of carbon sequestration into the ocean falls in the stated manner in the transition parameter $\delta_{\text{Ocean}\to\text{Atm}}$ that captures the carbon diffusion from the ocean back into the atmosphere and increases with the transition parameter $1-\delta_{\text{Atm}\to\text{Ocean}}$ that characterizes the persistence of carbon in the atmosphere.

D.3 Illustrating a Two Layer Atmosphere-Ocean Temperature System

The two layer example of atmosphere-ocean temperature dynamics has the transition matrix $\boldsymbol{\sigma} = \begin{pmatrix} 1-\sigma_1^{up}-\sigma_1^{down} & \sigma_1^{down} \\ \sigma_2^{up} & 1-\sigma_2^{up} \end{pmatrix}$. The corresponding term of the SCC (equation 13) 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 3(2) does not hold for the temperature system

$$\lim_{\beta\to 1}\varphi_{\tau,1} = -\xi_0(1+\varphi_k)\left[(1-\boldsymbol{\sigma})^{-1}\right]_{11} = -\frac{\xi_0(1+\varphi_k)}{\sigma_1^{up}} \neq \infty. \tag{D.4}$$

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.[39]

[39] I note that the carbon cycle lacks the reduction in persistence deriving from the forcing weight σ_1^{up}. With this observation, equation (D.4) gives another illustration of the impact of mass conservation

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 (7), 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.

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.

Part II - Uncertainty

E Equivalence to Epstein-Zin-Weil Utility and Illustration of Risk Aversion

This section presents a quantitative illustration of the adopted risk aversion and derives the equivalence to Epstein-Zin-Weil preferencs. I start by showing the equivalence of the Bellman equation (9) to the wide-spread formulation of recursive utility going back to Epstein & Zin (1991) and Weil (1990). Keeping isoelastic risk aggregation and using the logarithmic special case for intertemporal aggregation reflecting ACE's intertemporal elasticity of unity, the usual formulation reads

$$V_t^* = \exp\Big((1-\beta)\log c_t + \beta \log\big[\,\mathbb{E}_t\, {V_{t+1}^*}^{\alpha^*}\big]^{\frac{1}{\alpha^*}}\Big)\ . \tag{E.1}$$

Defining $V_t = \frac{\log V_t^*}{1-\beta}$ and rearranging equation (E.1) delivers

$$V_t = \log c_t + \frac{\beta}{1-\beta}\log\Big[\,\mathbb{E}_t \exp\big((1-\beta)V_{t+1}\big)^{\alpha^*}\Big]^{\frac{1}{\alpha^*}}\ . \tag{E.2}$$

Defining $\alpha = (1-\beta)\alpha^*$ and pulling the risk aversion coefficient α^* of the Epstein-Zin setting to the front of the logarithm and into the exponential yields equation (9) stated in the text.

The renormalization of the Bellman equation from equation (E.1) to equations (E.2) and (9) renormalizes utility such that marginal utility in the present is invariant to the choice of discount factor. This insight underlies the interpretation of the welfare losses in section (4.4). This renormalization and equation (9) suggest the natural measure of relative risk aversion RRA$= 1-\alpha$, which differs from the normalization suggested by Epstein & Zin (1991) leading to RRA$^* = 1-\alpha^* = 1-\frac{\alpha}{1-\beta}$. Only the measure RRA$= 1-\alpha$ is normalized so that RRA$= 0$ indeed corresponds to risk neutrality.[40] As importantly, the risk aversion measure RRA$= 1-\alpha$ is time preference invariant in that the lottery choice depicted in Figure 6, which I will use to illustrate the strength of a given risk aversion, depends only on the choice of risk aversion and not on time preference.

[40] The reader can convince herself of this statement by either substituting V_{t+1} recursively into equations (E.1) or (9), or by looking at Figure 6 for the special case of the lottery that I will introduce below. See Traeger (2014*b*) for more on the normalization of risk aversion measures in the Epstein-Zin-Weil setting.

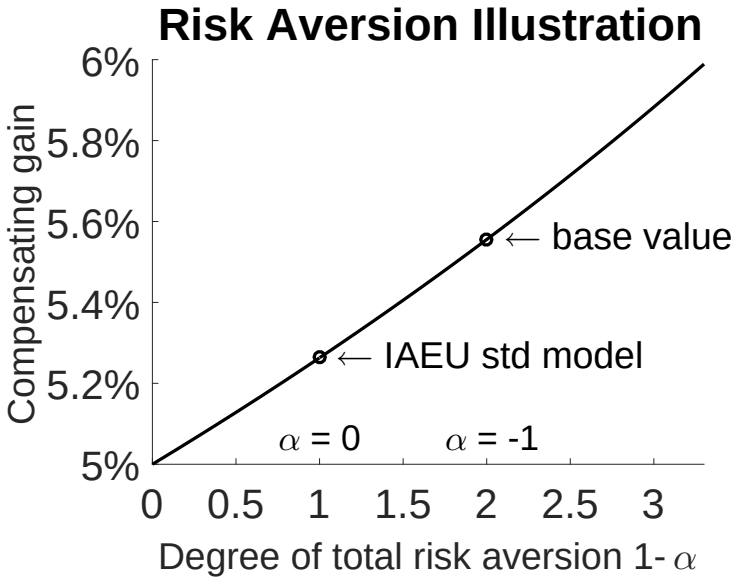


Figure 6: The graphs shows the compensating risk premium that an agent requires with probability one half to compensate a 5% loss occurring as well with probability one half. The intertemporally additive expected utility (IAEU) model corresponds to $\alpha = 0$ (no risk aversion beyond the desire to smooth consumption over time) and to a risk aversion of unity. The base calibration in ACE corresponds to $\alpha = -1$, and a total risk aversion of 2 (see text for normalization of the risk measure).

Figure 6 illustrates in a simple lottery the strength of risk aversion implied by the numeric choices of the parameters α and RRA$= 1 - \alpha$. In the baseline, an agent consumes a constant level $\bar{c}$ in perpetuity. Now I offer the agent a lottery where she either loses 5% of her baseline consumption $\bar{c}$ in the upcoming decade or gains the fraction z of consumption, each with probability one half. The graph presents, as a function of her risk aversion RRA, the percentage gain z that leaves the agent indifferent between the lottery and the baseline. Note that these losses and gains are direct consumption changes.[41] The asset pricing literature usually finds RRA$^* = 1 - \alpha^* \in [6, 10]$. In ACE's baseline calibration, these values translate approximately into the range $\alpha \in [-1.2, -0.7]$ and I pick $\alpha = -1$ as the baseline (RRA$= 2$), also presenting results for $\alpha = -1.25$ and $\alpha = -.5$ (RRA$= 1.5$ and RRA$= 2.25$) just outside

[41]The underlying calculation comes down to comparing the welfare for the deterministic path $\exp\big(\alpha\big[\log\bar{c}+\beta\sum\cdot\big]\big)$ with that for the lottery $\frac{1}{2}\exp\big(\alpha\big[\log(1-5\%)+\log\bar{c}+\beta\sum\cdot\big]\big)+\frac{1}{2}\exp\big(\alpha\big[\log(1+z)+\log\bar{c}+\beta\sum\cdot\big)\big]$, where $\sum\cdot$ is the coinciding future utility from future consumption. Equating the welfare resulting from the deterministic path and the lottery implies the formula $z = (2-(1-5\%)^\alpha)^{\frac{1}{\alpha}} - 1$ depicted in the figure. Note that the Bellman formulation of welfare in equation (9) assesses uncertainty only in the next period. One can either use the terms subsequent to β in equation (9) to evaluate an immediate lottery, or one can interpret the lottery as taking place over next period's consumption level where current period consumption is at the deterministic level $\bar{c}$.

of these bounds as a sensitivity range.

F Proofs (Uncertainty)

F.1 Proof of Proposition 2

Inserting an affine trial solution of the value function into the Bellman equation (9) and using the same transformations as in the deterministic case delivers

$$\varphi_k k_t + \boldsymbol{\varphi}_M^\top \boldsymbol{M}_t + \boldsymbol{\varphi}_\tau^\top \boldsymbol{\tau}_t + \boldsymbol{\varphi}_{R,t}^\top \boldsymbol{R}_t + \varphi_t + \boldsymbol{\varphi}_I^\top \boldsymbol{I}_t = \max_{x_t, \boldsymbol{N_t}, \boldsymbol{\mathcal{K}}_t, \boldsymbol{E_t}} \log x_t + \beta\varphi_k \log(1-x_t)$$
$$+A(\cdot) + \frac{\beta}{\alpha} \log\Big(\mathbb{E}_t \exp\Big[\alpha\Big(\boldsymbol{\varphi}_M^\top \boldsymbol{M}_{t+1} + \boldsymbol{\varphi}_\tau^\top \boldsymbol{\tau}_{t+1} + \boldsymbol{\varphi}_I^\top \boldsymbol{I}_{t+1}\Big)\Big]\Big) \ , \tag{F.1}$$

where $A(\cdot)$ summarizes the terms defined on page 47, already appearing in the deterministic solution. Using assumption (8) on the conditional expectation $\boldsymbol{X}_t = (\boldsymbol{M}_t, \boldsymbol{\tau}_t, \boldsymbol{I}_t)$ with $\boldsymbol{z} = (\alpha\boldsymbol{\varphi}_M^\top, \alpha\boldsymbol{\varphi}_\tau^\top, \alpha\boldsymbol{\varphi}_I^\top)$ yields

$$\varphi_k k_t + \boldsymbol{\varphi}_M^\top \boldsymbol{M}_t + \boldsymbol{\varphi}_\tau^\top \boldsymbol{\tau}_t + \boldsymbol{\varphi}_{R,t}^\top \boldsymbol{R}_t + \varphi_t + \boldsymbol{\varphi}_I^\top \boldsymbol{I}_t = \max_{x_t, \boldsymbol{N_t}, \boldsymbol{\mathcal{K}}_t, \boldsymbol{E_t}} \log x_t + \beta\varphi_k \log(1-x_t)$$
$$+A(\cdot) + \frac{\beta}{\alpha}\Big(a\big(\alpha\boldsymbol{\varphi}_M^\top, \alpha\boldsymbol{\varphi}_\tau^\top, \alpha\boldsymbol{\varphi}_I^\top, \boldsymbol{A}_t, \boldsymbol{N}_t, \boldsymbol{K}_t, \boldsymbol{E}_t\big) + \sum_{i=1}^{N} b_i(\alpha\boldsymbol{\varphi}_M^\top, \alpha\boldsymbol{\varphi}_\tau^\top, \alpha\boldsymbol{\varphi}_I^\top)) X_{t,i}\Big).$$

Maximizing the right hand side of the Bellman equation implies the same optimal consumption rate $x^* = \frac{1}{1+\beta\varphi_k}$ as in the deterministic case, and an analogous set of 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 $\boldsymbol{E}_t^*(\boldsymbol{A}_t, \varphi_k, \boldsymbol{\varphi}_M, \boldsymbol{\varphi}_{R,t+1})$. Inserting the optimal control rules and collecting the state-dependent terms (ordered by state) on the left hand side of the equality yields

$$\begin{aligned}
&\big(\varphi_k - (1+\beta\varphi_k)\kappa\big)k_t + \textstyle\sum_{i=1}^{m}\Big(\varphi_{M,i} - \frac{\beta}{\alpha}\, b_i^M(\alpha\boldsymbol{\varphi}_M^\top, \alpha\boldsymbol{\varphi}_\tau^\top, \alpha\boldsymbol{\varphi}_I^\top)\Big) M_{i,t} \\
&\quad + \textstyle\sum_{i=1}^{l}\Big(\varphi_{\tau,i} + (1+\beta\varphi_k)\xi_0\delta_{i,1} - \frac{\beta}{\alpha}\, b_i^\tau(\alpha\boldsymbol{\varphi}_M^\top, \alpha\boldsymbol{\varphi}_\tau^\top, \alpha\boldsymbol{\varphi}_I^\top)\Big)\tau_{i,t} \\
&\quad + \big(\boldsymbol{\varphi}_{R,t}^\top - \beta\boldsymbol{\varphi}_{R,t+1}^\top\big)\boldsymbol{R}_t + \varphi_t \\
&\quad + \textstyle\sum_{i=1}^{N-l-m}\big(\varphi_{I,i} - \frac{\beta}{\alpha}\, b_i^I(\alpha\boldsymbol{\varphi}_M^\top, \alpha\boldsymbol{\varphi}_\tau^\top, \alpha\boldsymbol{\varphi}_I^\top)\big)\, I_{i,t} \\
=&B(\cdot) + \beta\varphi_{t+1} + \tfrac{\beta}{\alpha}\, a(\alpha\boldsymbol{\varphi}_M^\top, \alpha\boldsymbol{\varphi}_\tau^\top, \alpha\boldsymbol{\varphi}_I^\top) \ ,
\end{aligned}$$

where $B(\cdot)$ summarizes the terms defined on page 49 that coincide with those of the deterministic solution. where Moreover, $(b_1^M, ..., b_m^M, b_1^\tau, ..., b_l^\tau, b_1^I, ..., b_{N-l-m}^I) = (b_1, ..., b_N)$

and $\delta_{i,j}$ denotes the Kronecker-delta (one if $i = j$ and zero otherwise) as defined in the proposition. As in the deterministic case, the trial solution solves the stochastic optimization problem if (and only if) all the coefficients in front of the state variables vanish, and the affine terms are eliminated by appropriate choice of the sequence $\varphi_0, \varphi_1, \varphi_2, \ldots$. The coefficient on log capital results in $\varphi_k = \frac{\kappa}{1-\beta\kappa}$ (as in equation C.5). The coefficient in front of the resource vector implies Hotelling's rule $\boldsymbol{\varphi}_{R,t} = \beta^t \boldsymbol{\varphi}_{R,0}$ (as in equation C.8). For sufficiency of the proposition, note that the remaining equations needed to eliminate the coefficients on the states in the Bellman equation are those equations stipulated in the proposition. For necessity, note that any solution that does not satisfy the stated equations will imply non-zero coefficients on the state variables and, thus, implies that the Bellman equation cannot be satisfied for all relevant state realizations.

F.2 Proof of Propositions 5 and 6

Propositions 5 is the special case of Proposition 6 focusing on only carbon flow uncertainty. To avoid repetition, I go straight to proving the general case with joint uncertainty, after introducing the details of the autoregressive gamma process.

The autoregressive gamma process by Gourieroux & Jasiak (2006) is as a Poisson mixture of gamma distributions,

$$\frac{X_{t+1}}{c}|(Z, X_t) \sim gamma(\nu_t + Z), \text{ where } Z|X_t \sim Poisson\left(\frac{\gamma X_t}{c}\right)$$

for $c, \gamma, \nu_t > 0$ in all periods. The random variable Z is drawn from a Poisson distribution and modulates the shape parameter of the standard gamma distribution (with scale c). The expectation and variance of this process are

$$\begin{aligned}\mathbb{E}_t(X_{i,t+1}|X_t) &= \nu_t c + \gamma X_t \\ \mathbb{V}\text{ar}_t(X_{i,t+1}|X_t) &= \nu_t c^2 + 2c\gamma X_t.\end{aligned}$$

and the cumulant generating function is

$$G_{X_{t+1}}(u) = \log\left[\mathbb{E}\left(\exp(uX_{t+1})|X_t\right)\right] = -\nu_t \log(1-uc) + \tfrac{u}{1-uc}\gamma X_t \ .$$

Applying the model to the temperature-carbon feedback, I specify the gamma autoregressive process y_t choosing

$$\nu_t = \frac{1}{c}\left(\tfrac{M_{1,t}+G_t}{M_{pre}} - \eta_\tau\right) \ ,$$

which results in the expectation and variance

$$\begin{aligned}\mathbb{E}\, y_{t+1} \quad &= \gamma^z y_t + \left(\tfrac{M_{1,t}+G_t}{M_{pre}} - \eta_\tau\right)\\ \mathbb{V}\mathrm{ar}\, y_{t+1} &= c\Big[2\gamma^z y_t + \left(\tfrac{M_{1,t}+G_t}{M_{pre}} - \eta_\tau\right)\Big].\end{aligned} \tag{F.2}$$

I define the deterministic process neutralizing the temperature expectations to those of the deterministic model[42] as

$$y^o_{t+1} = \gamma^z y^o_t + (1-\epsilon(c))\left(\tfrac{M_{1,t}+G_t}{M_{pre}} - \eta_\tau\right)$$

Then the expectation adjusted process $z_t \equiv y_t - y^o_t$ has the expectation and variance stated as equations (22) and (23) in the main text. To apply Proposition 2, I calculate (one over α times) the cumulant generation function of $\boldsymbol{X}_t = (\boldsymbol{M}_t, \boldsymbol{\tau}_t, \boldsymbol{I}_t)$ with $\boldsymbol{I}_t = (x^M_t, \sigma^M_t, y_t, y^o_t)$ and $\boldsymbol{z} = \alpha\boldsymbol{\varphi}^\top$:

$$\begin{aligned}\frac{1}{\alpha}\log\left(\mathbb{E}\exp(\alpha\boldsymbol{\varphi}^\top \boldsymbol{X}_{t+1})|\boldsymbol{X}_t\right) = \\ &\boldsymbol{\varphi}_M^\top \boldsymbol{\Phi}\boldsymbol{M}_t + ... + (\varphi_{M,1}-\varphi_{M,2})x^M_t + \frac{\alpha}{2}(\varphi_{M1}-\varphi_{M2})^2{\sigma^M_t}^2\\ &+\boldsymbol{\varphi}_\tau^\top \boldsymbol{\sigma}\boldsymbol{\tau}_t + \frac{\sigma^{forc}}{M_{pre}}\varphi_{\tau,1}M_{1,t} - h\varphi_{\tau,1}\gamma^z y^o_t - h\varphi_{\tau,1}(1-\epsilon(c))\tfrac{M_{1,t}}{M_{pre}} + ...\\ &+\varphi^M_x\gamma^x x^M_t + \frac{\alpha}{2}{\varphi^M_x}^2{\delta^{Mx}}^2\frac{M_{1,t}}{M_{pre}} + \frac{\alpha}{2}{\varphi^M_x}^2{\delta^{\sigma x}}^2{\sigma^M_t}^2\\ &+\varphi^M_\sigma\gamma^\sigma {\sigma^M_t}^2 + ... + \varphi^M_\sigma\delta^{M\sigma}\frac{M_{1,t}}{M_{pre}}\\ &-\tfrac{1}{\alpha c}\tfrac{M_{1,t}}{M_{pre}}\log(1-\alpha[\varphi^\tau_y + h\varphi_{\tau,1}]c) + \tfrac{\varphi^\tau_y + h\varphi_{\tau,1}}{1-\alpha[\varphi^\tau_y+h\varphi_{\tau,1}]c}\gamma^z y_t\\ &+\varphi^\tau_{y^o}\gamma^z y^o_t + \varphi^\tau_{y^o}(1-\epsilon(c))\tfrac{M_{1,t}}{M_{pre}} + ...\end{aligned} \tag{F.3}$$

I abbreviate by "..." affine terms that are independent of the states. Sorting the r.h.s. of equation (F.3) by states identifies Proposition 2's linear terms $b^M_i(\alpha\boldsymbol{\varphi}^\top_M, \alpha\boldsymbol{\varphi}^\top_\tau, \alpha\boldsymbol{\varphi}^\top_I)$, $b^\tau_i(\alpha\boldsymbol{\varphi}^\top_M, \alpha\boldsymbol{\varphi}^\top_\tau, \alpha\boldsymbol{\varphi}^\top_I)$, and $b^I_i(\alpha\boldsymbol{\varphi}^\top_M, \alpha\boldsymbol{\varphi}^\top_\tau, \alpha\boldsymbol{\varphi}^\top_I)$ where $\boldsymbol{\varphi}_I = (\varphi^M_x, \varphi^\tau_y, \varphi^M_\sigma, \varphi^\tau_\sigma)$. Then,

[42] When combining carbon flow uncertainty with temperature uncertainty, y^o_t also becomes stochastic, but accounts only for the stochastic evolution of carbon, not for the persistent shocks to the temperature response to CO_2 concentrations.

Proposition 2 implies the following system of equations for the shadow values

$$\boldsymbol{\varphi}_M^\top = \beta\boldsymbol{\varphi}_M^\top\boldsymbol{\Phi}+\beta\Big(\tfrac{\sigma^{forc}}{M_{pre}}\varphi_{\tau,1}+\tfrac{\alpha}{2}\tfrac{\delta^{Mx^2}}{M_{pre}}{\varphi_x^M}^2+\tfrac{\delta^{M\sigma}}{M_{pre}}\varphi_\sigma^M+\tfrac{1}{M_{pre}}(\varphi_{y^o}^\tau-h\varphi_{\tau,1})(1-\epsilon(c)) - \tfrac{1}{M_{pre}}\tfrac{\log(1-\alpha c(\varphi_y^\tau+h\varphi_{\tau,1})}{\alpha c}\Big)\boldsymbol{e}_1^\top \tag{F.4}$$

$$\boldsymbol{\varphi}_\tau^\top = \beta\boldsymbol{\varphi}_\tau^\top\boldsymbol{\sigma} - (1+\beta\varphi_k)\xi_0\boldsymbol{e}_1^\top \tag{F.5}$$

$$\varphi_{y^o}^\tau = \beta(\varphi_{y^o}^\tau - h\varphi_{\tau,1})\gamma^z \tag{F.6}$$

$$\varphi_y^\tau = \beta\tfrac{\varphi_y^\tau+h\varphi_{\tau,1}}{1-\alpha c(\varphi_y^\tau+h\varphi_{\tau,1})}\gamma^z \tag{F.7}$$

$$\varphi_x^M = \beta(\varphi_{M,1}-\varphi_{M,2})+\beta\varphi_x^M\gamma^x \tag{F.8}$$

$$\varphi_\sigma^M = \beta\frac{\alpha}{2}\left((\varphi_{M,1}-\varphi_{M,2})^2+\delta^{\sigma x^2}{\varphi_x^M}^2\right)+\beta\varphi_\sigma^M\gamma^\sigma \tag{F.9}$$

Temperature related shadow values:
The temperature's shadow value is as before by equation (F.5)

$$\boldsymbol{\varphi}_\tau^\top = -(1+\beta\varphi_k)\xi_0\boldsymbol{e}_1^\top(\mathbb{1}-\beta\boldsymbol{\sigma})^{-1}.$$

The feedback operates through the carbon's shadow value and through the persistent shock shadow value φ_y^τ for which equation (F.7) delivers the quadratic equation

$$\begin{aligned}&\varphi_y^\tau-\alpha c{\varphi_y^\tau}^2-\alpha ch\varphi_{\tau,1}\varphi_y^\tau=\beta\varphi_y^\tau\gamma^z+\beta h\varphi_{\tau,1}\gamma^z\\ \Leftrightarrow\ &\underbrace{\alpha c}_{\equiv\tilde a}{\varphi_y^\tau}^2+\underbrace{(\beta\gamma^z+\alpha ch\varphi_{\tau,1}-1)}_{\equiv\tilde b}\varphi_y^\tau+\underbrace{\beta h\varphi_{\tau,1}\gamma^z}_{\equiv\tilde c}=0.\end{aligned}$$

Instead of using the common *abc*-formula I use the solution arrived at by Mullers method, which solves $\tilde a x^2+\tilde b x+\tilde c=0$ by the roots $x=\frac{-2\tilde c}{\tilde b\pm\sqrt{\tilde b^2-4\tilde a\tilde c}}$. The solution is advantageous because it yields a valid root for the case $\tilde a=0$, which corresponds to the deterministic case.[43] Then

$$\begin{aligned}\varphi_y^\tau &= \frac{-2\tilde c}{\tilde b\pm\sqrt{\tilde b^2-4\tilde a\tilde c}} = \varphi_{\tau,1}\frac{2\beta h\gamma^z}{(1-\beta\gamma^z-\alpha ch\varphi_{\tau,1})\pm\sqrt{(1-\beta\gamma^z-\alpha ch\varphi_{\tau,1})^2-4\alpha ch\varphi_{\tau,1}\beta\gamma^z}}\\ &= \frac{\beta\gamma^z}{1-\beta\gamma^z}\underbrace{\frac{2}{1-\frac{\alpha ch\varphi_{\tau,1}}{1-\beta\gamma^z}\pm\sqrt{\left(1-\frac{\alpha ch\varphi_{\tau,1}}{1-\beta\gamma^z}\right)^2-4\frac{\alpha ch\varphi_{\tau,1}}{1-\beta\gamma^z}\frac{\beta\gamma^z}{1-\beta\gamma^z}}}}_{\equiv T}h\varphi_{\tau,1}\end{aligned} \tag{F.10}$$

[43] The common *abc*-formula yields a fraction $\frac{0}{0}$ for $\tilde a=0$. Having a well-defined root for the deterministic special case permits me to connect the uncertain SCC directly to the deterministic SCC.

To identify the economically meaningful root, I take $c \to 0$. The negative root diverges and identifies the positive root as the correct root (the root with $+\sqrt{\ }$). The correct deterministic limit delivers $\varphi_y^\tau \to \varphi_{\tau,1}\frac{\beta h\gamma^z}{(1-\beta\gamma^z)}$ for $c \to 0$. The shadow value in the deterministic limit coincides with the (negative of the) shadow value $\varphi_{y^o}^\tau$ that results from equation (F.7) as

$$\varphi_{y^o}^\tau = -\frac{\beta h\gamma^z}{1-\beta\gamma^z}\varphi_{\tau,1}.$$

Carbon-flow uncertainty:
Equation (F.4) delivers the shadow value vector equation

$$\boldsymbol{\varphi}_M^\top = \beta\Bigg(\frac{\sigma^{forc}}{M_{pre}}\varphi_{\tau,1} + \frac{\alpha}{2}\frac{\delta^{Mx^2}}{M_{pre}}\varphi_x^{M^2} + \frac{\delta^{M\sigma}}{M_{pre}}\varphi_\sigma^M + \frac{1}{M_{pre}}(\varphi_{y^o}^\tau - h\varphi_{\tau,1})(1-\epsilon(c)) \\ -\frac{1}{M_{pre}}\frac{\log(1-\alpha c(\varphi_y^\tau + h\varphi_{\tau,1}))}{\alpha c}\Bigg)\left[(\mathbb{1}-\beta\boldsymbol{\Phi})^{-1}\right]_{1,\cdot}\,. \tag{F.11}$$

Dividing the second through the first shadow value entry I obtain

$$\varphi_{M,2} = \underbrace{\frac{\left[(\mathbb{1}-\beta\boldsymbol{\Phi})^{-1}\right]_{1,2}}{\left[(\mathbb{1}-\beta\boldsymbol{\Phi})^{-1}\right]_{1,1}}}_{\equiv r}\,\varphi_{M,1}\,. \tag{F.12}$$

Equation (F.8) delivers the shadow value

$$\varphi_x^M = \frac{\beta}{1-\gamma^x\beta}(\varphi_{M,1}-\varphi_{M,2}) = \underbrace{\frac{\beta}{1-\gamma^x\beta}(1-r)}_{\equiv A}\varphi_{M,1}\,, \tag{F.13}$$

where the second equality uses equation (F.12). Substituting these results into equation (F.9) delivers

$$\varphi_\sigma^M = \beta\frac{\alpha}{2}\frac{(\varphi_{M,1}-\varphi_{M,2})^2 + \delta^{\sigma x^2}\varphi_x^{M^2}}{1-\gamma^\sigma\beta} = \underbrace{\beta\frac{\alpha}{2}\frac{(1-r)^2+\delta^{\sigma x^2}A^2}{1-\gamma^\sigma\beta}}_{\equiv B}\varphi_{M,1}{}^2. \tag{F.14}$$

Inserting equation (F.13) and (F.14) into the atmospheric shadow value component of

equation (F.11) results in the quadratic equation

$$\begin{aligned}
\varphi_{M,1} &= \beta\left(\frac{\alpha}{2}\frac{\delta^{Mx^2}}{M_{pre}}{\varphi_x^M}^2 + \frac{\delta^{M\sigma}}{M_{pre}}\varphi_\sigma^M\right)\left[(\mathbb{1}-\beta\mathbf{\Phi})^{-1}\right]_{1,1} \\
&\quad + \frac{\beta}{M_{pre}}\left(\sigma^{forc}\varphi_{\tau,1} + (\varphi_{y^o}^\tau - h\varphi_{\tau,1})(1-\epsilon(c)) - \frac{\log(1-\alpha c(\varphi_y^\tau + h\varphi_{\tau,1})}{\alpha c}\right)\left[(\mathbb{1}-\beta\mathbf{\Phi})^{-1}\right]_{1,1} \\
&= \underbrace{\beta\left(\frac{\alpha}{2}\frac{\delta^{Mx^2}}{M_{pre}}A^2 + \frac{\delta^{M\sigma}}{M_{pre}}B\right)\left[(\mathbb{1}-\beta\mathbf{\Phi})^{-1}\right]_{1,1}}_{\equiv\hat{a}}{\varphi_{M,1}}^2 \\
&\quad + \underbrace{\frac{\beta}{M_{pre}}\left(\sigma^{forc}\varphi_{\tau,1} + (\varphi_{y^o}^\tau - h\varphi_{\tau,1})(\delta_\tau - \epsilon(c)) - \frac{\log(1-\alpha c(\varphi_y^\tau + h\varphi_{\tau,1})}{\alpha c}\right)\left[(\mathbb{1}-\beta\mathbf{\Phi})^{-1}\right]_{1,1}}_{\equiv\hat{c}}.
\end{aligned}$$

Using once more the quadratic formula deriving from Muller's method I obtain the solution

$$\varphi_{M,1} = \frac{2\hat{c}}{1 \pm \sqrt{1-4\hat{a}\hat{c}}}$$

and once again the positive root is the one that is economically meaningful as it converges for $\hat{a} = 0$ to the correct solution (including the deterministic special case if all uncertainty is absent). I transform the expression for $\varphi_{M,1}$ and, in the last step, do a first order Taylor approximation in both numerator and denominator

$$\varphi_{M,1} = \frac{2\hat{c}}{1+\sqrt{1-4\hat{a}\hat{c}}} = \hat{c}\left(1 + \underbrace{\frac{1-\sqrt{1-4\hat{a}\hat{c}}}{1+\sqrt{1-4\hat{a}\hat{c}}}}_{\equiv\theta_M}\right) \approx \hat{c}\left(1 + \frac{\hat{a}\hat{c}}{1-\hat{a}\hat{c}}\right), \tag{F.15}$$

where the approximation is first order around $\hat{a}\hat{c} = 0$ in both numerator and denomi-

nator. The term $\hat{a}$ is

$$\begin{aligned}
\hat{a} &= \beta\left(\frac{\alpha}{2}\frac{\delta^{Mx^2}}{M_{pre}}A^2 + \frac{\delta^{M\sigma}}{M_{pre}}B\right)\left[(\mathbb{1}-\beta\mathbf{\Phi})^{-1}\right]_{1,1} \\
&= \beta\frac{\alpha}{2}\frac{1}{M_{pre}}\left[\left(\frac{\beta\delta^{Mx}}{1-\gamma^x\beta}\right)^2(1-r)^2\right. \\
&\qquad \left.+\frac{\beta\delta^{M\sigma}}{1-\gamma^\sigma\beta}\left((1-r)^2+\left(\frac{\beta\delta^{\sigma x}}{1-\gamma^x\beta}\right)^2(1-r)^2\right)\right]\left[(\mathbb{1}-\beta\mathbf{\Phi})^{-1}\right]_{1,1} \\
&= \frac{\alpha}{2}\frac{\beta}{M_{pre}}\Big[A^{M\to x^2}+A^{M\to\sigma}A^{\sigma\to x^2}+A^{M\to\sigma}\Big](1-r)^2\left[(\mathbb{1}-\beta\mathbf{\Phi})^{-1}\right]_{1,1} \\
&= \frac{\alpha}{2}\frac{\beta}{M_{pre}}\Big[A^{M\to x^2}+A^{M\to\sigma}A^{\sigma\to x^2}+A^{M\to\sigma}\Big]\frac{\left(\left[(\mathbf{1}-\beta\mathbf{\Phi})^{-1}\right]_{1,1}-\left[(\mathbf{1}-\beta\mathbf{\Phi})^{-1}\right]_{1,2}\right)^2}{\left[(\mathbf{1}-\beta\mathbf{\Phi})^{-1}\right]_{1,1}} \qquad \text{(F.16)}
\end{aligned}$$

with $A^{M\to x}=\dfrac{\delta^{Mx}\beta}{1-\gamma^x\beta}$, $\quad A^{M\to\sigma}=\dfrac{\delta^{M\sigma}\beta}{1-\gamma^\sigma\beta}$, $\quad A^{\sigma\to x}=\dfrac{\delta^{\sigma x}\beta}{1-\gamma^x\beta}$.

Temperature-carbon feedback:
Evaluating the term $\hat{c}$ requires the evaluation of

$$\varphi^\tau_{y^o}-h\varphi_{\tau,1} = -\left(\frac{\beta\gamma^z}{1-\beta\gamma^z}+1\right)h\varphi_{\tau,1} = -\frac{h}{1-\beta\gamma^z}\varphi_{\tau,1} = -\bar{h}\varphi_{\tau,1},$$

where I defined $\bar{h}=\frac{h}{1-\beta\gamma^z}$, and, using equation (F.10), the evaluation of

$$\varphi^\tau_y+h\varphi_{\tau,1} = \left(\frac{\beta\gamma^z}{1-\beta\gamma^z}T+1\right)h\varphi_{\tau,1} = \frac{1+\beta\gamma^z(T-1)}{1-\beta\gamma^z}h\varphi_{\tau,1} = \Big(1+\beta\gamma^z(T-1)\Big)\bar{h}\varphi_{\tau,1}.$$

Using the definition $F\equiv\alpha c\frac{h}{1-\beta\gamma^z}\varphi_{\tau,1}=\alpha c\bar{h}\varphi_{\tau,1}$, I define the expression

$$\begin{aligned}
\theta^\dagger_\tau &\equiv \beta\gamma^z(T-1) = \beta\gamma^z\left(\frac{2}{1-F+\sqrt{(1-F)^2-4F\frac{\beta\gamma^z}{1-\beta\gamma^z}}}-1\right) \\
&= \beta\gamma^z\frac{1+F-\sqrt{(1-F)^2-4F\frac{\beta\gamma^z}{1-\beta\gamma^z}}}{1-F+\sqrt{(1-F)^2-4F\frac{\beta\gamma^z}{1-\beta\gamma^z}}} \approx \frac{\beta\gamma^zF}{1-\beta\gamma^z-F} \qquad \text{(F.17)}
\end{aligned}$$

Using this definition and denoting the shadow value of atmospheric carbon under cer-

tainty, see equation (C.6), by $\varphi^{det}_{M,1}$ the term $\hat{c}$ becomes

$$\begin{aligned}\hat{c} &= \frac{\beta\sigma^{forc}\varphi_{\tau,1}}{M_{pre}}\left[(\mathbb{1}-\beta\mathbf{\Phi})^{-1}\right]_{1,1}\left(1+\frac{\bar{h}}{\sigma^{forc}}\left(-1\right.\right.\\ &\qquad\left.\left.-\frac{\log(1-\alpha c\big(1+\beta\gamma^z(T-1)\big)\bar{h}\varphi_{\tau,1}}{\alpha c\bar{h}\varphi_{\tau,1}}+\epsilon(c)\right)\right)\\ &= \varphi^{det}_{M,1}\left(1+\frac{\bar{h}}{\sigma^{forc}}\left(\frac{-\log\left(1-\alpha c\bar{h}\varphi_{\tau,1}(1+\theta^{\dagger}_{\tau})\right)}{\alpha c\bar{h}\varphi_{\tau,1}}-1+\epsilon(c)\right)\right)\\ &= \varphi^{det}_{M,1}\left(1+\frac{\bar{h}}{\sigma^{forc}}\left(\frac{-\log\left(1-F(1+\theta^{\dagger}_{\tau})\right)}{F}-1+\epsilon(c)\right)\right).\end{aligned} \tag{F.18}$$

The joined first order approximation in $\theta^{\dagger}_{\tau}$ and F (first approximation), and a first order approximation in F using the definition of $\theta^{\dagger}_{\tau}$ (second approximation) deliver

$$\theta^{*}_{\tau}\equiv\frac{-\log\left(1-F(1+\theta^{\dagger}_{\tau})\right)}{F}-1 \;\approx\; \theta^{\dagger}_{\tau}+\frac{1}{2}F \;\approx\; \frac{1}{2}\,\frac{1+\beta\gamma^z}{1-\beta\gamma^z}\,F. \tag{F.19}$$

Summarizing the results for Proposition 5

I define $\theta^{*}_{M}=\hat{a}\hat{c}$. The main text uses the definitions of *persist* and Δ*persist* defined on page **??** simplifying the representation of $\hat{a}$ in equation (F.16). Moreover, equations (F.18) and (F.19) imply

$$\hat{c}=\varphi^{det}_{M,1}\left(1+\frac{\bar{h}}{\sigma^{forc}}\left(\theta^{*}_{\tau}+\epsilon(c)\right)\right) \text{ with } \theta^{*}_{\tau} \text{ as in equation (F.19).}$$

Equation (F.15) delivers $\varphi_{M,1}=\hat{c}\big(1+\theta_M\big)$ with

$$\theta_M=\frac{1-\sqrt{1-4\theta^{*}_{M}}}{1+\sqrt{1-4\theta^{*}_{M}}}\approx\frac{\theta^{*}_{M}}{1-\theta^{*}_{M}}. \tag{F.20}$$

Therefore, $\varphi_{M,1}=\varphi^{det}_{M,1}\big(1+\theta_M\big)\Big(1+\frac{\bar{h}}{\sigma^{forc}}\big(\theta^{*}_{\tau}+\epsilon(c)\big)\Big)$ and transformed to consumption units

$$SCC_t=SCC^{det}_{t}\big(1+\theta_M\big)\left(1+\frac{\bar{h}}{\sigma^{forc}}\Big(\theta^{*}_{\tau}+\epsilon(c)\Big)\right). \tag{F.21}$$

In the case of Proposition 5, $\bar{h}=0$ so that the second bracket is unity. The approximation in equation (F.20) delivers the approximation in equation (19) of the proposition.

Summarizing the results for Proposition 6

Equation (F.21) summarizes the general SCC of Proposition 6. The expressions $\bar{h}=$

$\frac{h}{1-\beta\gamma^z}$ and $F = \alpha c\bar{h}\varphi_{\tau,1}$ are as defined in the text above. The expression for θ_τ^* is stated in equation (F.19) and equation (F.17) defines $\theta_\tau^\dagger$. The approximation in equation (24) follows from the approximation in equation (F.19).

F.3 Proof of Proposition 7

Let ϵ_t^j be distributed with existing cumulant generating function, and let $\nu_{t,j}$ be iid white noise (and let other shocks be zero for the moment). Making use of Proposition 2, the informational state variable is the current realization of the autoregressive shock μ_t^j, which is known at time t. Then equation (F.1) becomes

$$\begin{aligned}
&\varphi_k k_t + \boldsymbol{\varphi}_M^\top \boldsymbol{M}_t + \boldsymbol{\varphi}_\tau^\top \boldsymbol{\tau}_t + \boldsymbol{\varphi}_{R,t}^\top \boldsymbol{R}_t + \varphi_t + \varphi_\mu^j \mu_t^j = \max_{x_t, \boldsymbol{N_t}, \boldsymbol{\mathcal{K}_t}, \boldsymbol{E_t}} \log x_t + \beta\varphi_k \log(1-x_t) \\
&\qquad + A(\cdot) + \frac{\beta}{\alpha} \log\Big(\mathbb{E}_t \exp\Big[\alpha\Big(\boldsymbol{\varphi}_M^\top \boldsymbol{M}_{t+1} + \boldsymbol{\varphi}_\tau^\top \boldsymbol{\tau}_{t+1} + \varphi_\mu^j \mu_{t+1}^j\Big)\Big]\Big) \\
\Leftrightarrow\ &\varphi_k k_t + \boldsymbol{\varphi}_M^\top \boldsymbol{M}_t + \boldsymbol{\varphi}_\tau^\top \boldsymbol{\tau}_t + \boldsymbol{\varphi}_{R,t}^\top \boldsymbol{R}_t + \varphi_t + \varphi_\mu^j \mu_t^j = \max_{x_t, \boldsymbol{N_t}, \boldsymbol{\mathcal{K}_t}, \boldsymbol{E_t}} \log x_t + \beta\varphi_k \log(1-x_t) \\
&\qquad + A(\cdot) + \beta \boldsymbol{\varphi}_M^\top \Big(\boldsymbol{\Phi} \boldsymbol{M}_t + \big(\textstyle\sum_{i=1}^{I^d} E_{i,t}(\boldsymbol{A}_t, \boldsymbol{N}_t) + E_t^{exo}\big)\boldsymbol{e}_1 \Big) \\
&\qquad + \beta \boldsymbol{\varphi}_\tau^\top \left(\boldsymbol{\sigma} \boldsymbol{\tau}_t + \sigma^{forc} \frac{M_{1,t} + G_t}{M_{pre}} \boldsymbol{e}_1 \right) \\
&\qquad + \frac{\beta}{\alpha} \log\Big(\mathbb{E}_t \exp\Big[\alpha\Big(\varphi_j \epsilon_t^j + \varphi_j \nu_t^j + \varphi_\mu^j \mu_{t+1}^j\Big)\Big]\Big) \,. \qquad \text{(F.22)}
\end{aligned}$$

In the **autoregressive shock model**, ϵ_t^j is known to be μ_t^j in period t and $\mu_{t+1}^j = \gamma^j \mu_t^j + \chi_t^j$. Moreover, $\nu_t^j = 0$ (by assumption). Therefore, in the autoregressive shock model, I obtain the Bellman equation

$$\begin{aligned}
&\varphi_k k_t + \boldsymbol{\varphi}_M^\top \boldsymbol{M}_t + \boldsymbol{\varphi}_\tau^\top \boldsymbol{\tau}_t + \boldsymbol{\varphi}_{R,t}^\top \boldsymbol{R}_t + \varphi_t + \varphi_\mu^j \mu_t^j = \max_{x_t, \boldsymbol{N_t}, \boldsymbol{\mathcal{K}_t}, \boldsymbol{E_t}} \log x_t + \beta\varphi_k \log(1-x_t) \\
&\qquad + A(\cdot) + \beta \boldsymbol{\varphi}_M^\top \Big(\boldsymbol{\Phi} \boldsymbol{M}_t + \big(\textstyle\sum_{i=1}^{I^d} E_{i,t}(\boldsymbol{A}_t, \boldsymbol{N}_t) + E_t^{exo}\big)\boldsymbol{e}_1 \Big) \\
&\qquad + \beta \boldsymbol{\varphi}_\tau^\top \left(\boldsymbol{\sigma} \boldsymbol{\tau}_t + \sigma^{forc} \frac{M_{1,t} + G_t}{M_{pre}} \boldsymbol{e}_1 \right) \\
&\qquad + \beta\varphi_j \mu_t^j + \beta \varphi_\mu^j \gamma^j \mu_t^j + \frac{\beta}{\alpha} \log\Big(\mathbb{E}_t \exp\Big[\alpha \varphi_\mu^j \chi_t^j\Big]\Big).
\end{aligned}$$

Moreover

$$\log\Big(\mathbb{E}_t \exp\Big[\alpha \varphi_\mu^j \chi_t^j\Big]\Big) = G_\chi\left(\alpha \varphi_\mu^j \chi_t^j\right) = \sum_{l=1}^{\infty} \frac{(\alpha \varphi_\mu^j)^l}{l!} \kappa_l$$

The Bellman equation delivers in close analogy to the deterministic setting (see equation C.4)

$$
\begin{aligned}
&\big(\boldsymbol{\varphi}_M^\top - \beta\boldsymbol{\varphi}_M^\top\boldsymbol{\Phi} - \beta\varphi_{\tau,1}\frac{\sigma^{forc}}{M_{pre}}\boldsymbol{e}_1^\top\big)\boldsymbol{M}_t + \big(\boldsymbol{\varphi}_\tau^\top - \beta\boldsymbol{\varphi}_\tau^\top\boldsymbol{\sigma} + (1+\beta\varphi_k)\xi_0\boldsymbol{e}_1^\top\big)\boldsymbol{\tau}_t \\
&\big(\varphi_k - (1+\beta\varphi_k)\kappa\big)k_t + \big(\boldsymbol{\varphi}_{R,t}^\top - \beta\boldsymbol{\varphi}_{R,t+1}^\top\big)\boldsymbol{R}_t \\
&+ \Big(\varphi_\mu^j - \beta(\varphi_j + \gamma^j\varphi_\mu^j)\Big)\mu_t^j \\
&+\varphi_t \;=\; \beta\varphi_{t+1} \\
&B(\cdot) + \beta\varphi_{M,1}\left(\textstyle\sum_{i=1}^{I^d} E_{i,t}^* + E_t^{exo}\right) + \beta\varphi_{\tau,1}\frac{\sigma^{forc}}{M^{pre}}G_t, + \frac{\beta}{\alpha}\textstyle\sum_{l=1}^{\infty}\frac{(\alpha\varphi_\mu^j)^l}{l!}\kappa_l
\end{aligned}
$$

where the coefficient on the new state μ_t has to vanish implying

$$\varphi_\mu^j = \frac{\beta}{1-\gamma^j\beta}\varphi_j \tag{F.23}$$

The value function difference between the deterministic and the autoregressive shock model is determined by the contribution from the informational state and the contribution of the affine parts of the value function. The informational state is zero in the present by assumption $\varphi_\mu^j\mu_0^j = 0$ (same expected motion as under certainty). By equation (C.9) we know that the affine part of the value function evolves as

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

in the deterministic model, and similarly as

$$\varphi_t^{AR} = \beta\varphi_{t+1}^{AR} + B(\cdot) + \beta\varphi_{M,1}\left(\textstyle\sum_{i=1}^{I^d} E_{i,t}^* + E_t^{exo}\right) + \beta\varphi_{\tau,1}\frac{\sigma^{forc}}{M^{pre}}G_t + \frac{\beta}{\alpha}\textstyle\sum_{l=1}^{\infty}\frac{(\alpha\varphi_\mu^j)^l}{l!}\kappa_l$$

in the autoregressive shock model. Therefore, the value function difference is

$$\begin{aligned}
\Delta W^{AR}_{general} &= V^{AR} - V^{det} = \varphi^j_\mu \mu^j_0 + \varphi^{unc}_0 - \varphi^{det}_0 \\
&= 0 + \beta(\varphi^{AR}_1 - \varphi^{det}_1) + \frac{\beta}{\alpha}\sum_{l=1}^{\infty}\frac{(\alpha\varphi^j_\mu)^l}{l!}\kappa_l \\
&= \beta(\varphi^{AR}_2 - \varphi^{det}_2) + \sum_{p=1}^{2}\frac{\beta^p}{\alpha}\sum_{l=1}^{\infty}\frac{(\alpha\varphi^j_\mu)^l}{l!}\kappa_l \\
&= \lim_{t\to\infty}\beta(\varphi^{AR}_t - \varphi^{det}_t) + \lim_{t\to\infty}\sum_{p=1}^{t}\frac{\beta^p}{\alpha}\sum_{l=1}^{\infty}\frac{(\alpha\varphi^j_\mu)^l}{l!}\kappa_l \\
&= \lim_{t\to\infty}\beta(\varphi^{AR}_t - \varphi^{det}_t) + \left(\frac{1}{1-\beta} - 1\right)\frac{1}{\alpha}\sum_{l=1}^{\infty}\frac{(\alpha\varphi^j_\mu)^l}{l!}\kappa_l \\
&= \sum_{l=1}^{\infty}\frac{\beta}{1-\beta}\frac{1}{\alpha}\frac{(\alpha\varphi^j_\mu)^l}{l!}\kappa_l = \frac{\beta}{1-\beta}\frac{1}{\alpha}G_\chi\big(\alpha\varphi^j_\mu\big),
\end{aligned}$$

which, together with equation (F.23) delivers the result stated in part (3) of the proposition. For a normally distributed mean-zero shock χ^j_t the first cumulant (expectation) is zero and only the second cumulant contributes (all others being zero), delivering part (1) of the proposition. Part (1) of the proposition follows from the

In the case of **anticipated learning**, the informational state evolves as $\mu^j_{t+1} = \frac{\sigma^2_{\epsilon,t}\tilde{z}+\sigma^2_{\nu,t}\mu^j_t}{\sigma^2_{\epsilon,t}+\sigma^2_{\nu,t}}$, where $\tilde{z}$ is the observation, which is distributed as the sum of measurement error and Bayesian prior $z \sim N(\mu^j_t, \sigma^2_{\epsilon,t} + \sigma^2_{\nu,t})$ (see as well footnote 30). The variance of the normal-normal Bayesian learning model evolves deterministically as $\sigma^2_{\epsilon,t+1} = \frac{\sigma^2_{\nu,t}\sigma^2_{\epsilon,t}}{\sigma^2_{\nu,t}+\sigma^2_{\epsilon,t}}$. Therefore, the Bellman equation (F.22) becomes

$$\begin{aligned}
\varphi_k k_t + \boldsymbol{\varphi}^\top_M \boldsymbol{M}_t + \boldsymbol{\varphi}^\top_\tau \boldsymbol{\tau}_t + \boldsymbol{\varphi}^\top_{R,t}\boldsymbol{R}_t + \varphi_t + \varphi^j_\mu \mu^j_t = \max_{x_t, \boldsymbol{N_t}, \boldsymbol{\mathcal{K}_t}, \boldsymbol{E_t}} &\log x_t + \beta\varphi_k \log(1-x_t) \\
&+ A(\cdot) + \beta\boldsymbol{\varphi}^\top_M\left(\boldsymbol{\Phi}\boldsymbol{M}_t + \big(\textstyle\sum_{i=1}^{I^d} E_{i,t}(\boldsymbol{A}_t, \boldsymbol{N}_t) + E^{exo}_t\big)\boldsymbol{e}_1\right) \\
&+ \beta\boldsymbol{\varphi}^\top_\tau\left(\boldsymbol{\sigma}\boldsymbol{\tau}_t + \sigma^{forc}\frac{M_{1,t}+G_t}{M_{pre}}\boldsymbol{e}_1\right) \\
&+ \frac{\beta}{\alpha}\log\Big(\mathbb{E}_t \exp\Big[\alpha\Big(\varphi_j(\epsilon^j_t + \nu^j_t) + \varphi^j_\mu\frac{\sigma^2_{\epsilon,t}\tilde{z} + \sigma^2_{\nu,t}\mu^j_t}{\sigma^2_{\epsilon,t}+\sigma^2_{\nu,t}}\Big)\Big]\Big)\ ,
\end{aligned}$$

Moreover

$$\log\Big(\mathbb{E}_t \exp\Big[\alpha\Big(\varphi_j(\epsilon_t^j+\nu_t^j)+\varphi_\mu^j\frac{\sigma_{\epsilon,t}^2\tilde{z}+\sigma_{\nu,t}^2\mu_t^j}{\sigma_{\epsilon,t}^2+\sigma_{\nu,t}^2}\Big)\Big]\Big)$$

$$=\varphi_\mu^j\frac{\sigma_{\nu,t}^2\mu_t^j}{\sigma_{\epsilon,t}^2+\sigma_{\nu,t}^2}+\log\Big(\mathbb{E}_t\exp\Big[\alpha\Big(\varphi_j+\varphi_\mu^j\frac{\sigma_{\epsilon,t}^2}{\sigma_{\epsilon,t}^2+\sigma_{\nu,t}^2}\Big)(\epsilon_t^j+\nu_t^j)\Big]\Big)$$

$$=\varphi_\mu^j\frac{\sigma_{\nu,t}^2\mu_t^j}{\sigma_{\epsilon,t}^2+\sigma_{\nu,t}^2}+\alpha\Big(\varphi_j+\varphi_\mu^j\frac{\sigma_{\epsilon,t}^2}{\sigma_{\epsilon,t}^2+\sigma_{\nu,t}^2}\Big)\mu_t^j+\alpha^2\Big(\varphi_j+\varphi_\mu^j\frac{\sigma_{\epsilon,t}^2}{\sigma_{\epsilon,t}^2+\sigma_{\nu,t}^2}\Big)^2\ \frac{\sigma_{\epsilon,t}^2+\sigma_{\nu,t}^2}{2}$$

Then, the Bellman equation delivers in close analogy to above or the deterministic setting (see equation C.4)

$$\begin{aligned}
&\big(\boldsymbol{\varphi}_M^\top-\beta\boldsymbol{\varphi}_M^\top\boldsymbol{\Phi}-\beta\varphi_{\tau,1}\frac{\sigma^{forc}}{M_{pre}}\boldsymbol{e}_1^\top\big)\boldsymbol{M}_t+\big(\boldsymbol{\varphi}_\tau^\top-\beta\boldsymbol{\varphi}_\tau^\top\boldsymbol{\sigma}+(1+\beta\varphi_k)\xi_0\boldsymbol{e}_1^\top\big)\boldsymbol{\tau}_t\\
&\big(\varphi_k-(1+\beta\varphi_k)\kappa\big)k_t+\big(\boldsymbol{\varphi}_{R,t}^\top-\beta\boldsymbol{\varphi}_{R,t+1}^\top\big)\boldsymbol{R}_t\\
&+\Big(\varphi_\mu^j-\beta\varphi_\mu^j\frac{\sigma_{\nu,t}^2}{\sigma_{\epsilon,t}^2+\sigma_{\nu,t}^2}-\beta\Big(\varphi_j+\varphi_\mu^j\frac{\sigma_{\epsilon,t}^2}{\sigma_{\epsilon,t}^2+\sigma_{\nu,t}^2}\Big)\Big)\mu_t^j\\
&+\varphi_t\ =\ \beta\varphi_{t+1}\\
&B(\cdot)+\beta\varphi_{M,1}\Big(\textstyle\sum_{i=1}^{I^d}E_{i,t}^*+E_t^{exo}\Big)+\beta\varphi_{\tau,1}\frac{\sigma^{forc}}{M^{pre}}G_t,+\alpha\beta\Big(\varphi_j+\varphi_\mu^j\frac{\sigma_{\epsilon,t}^2}{\sigma_{\epsilon,t}^2+\sigma_{\nu,t}^2}\Big)^2\frac{\sigma_{\epsilon,t}^2+\sigma_{\nu,t}^2}{2}
\end{aligned}$$

where the coefficient on the informational state μ_t has to vanish implying

$$\varphi_\mu^j\Bigg(1-\beta\underbrace{\Big(\frac{\sigma_{\nu,t}^2}{\sigma_{\epsilon,t}^2+\sigma_{\nu,t}^2}+\frac{\sigma_{\epsilon,t}^2}{\sigma_{\epsilon,t}^2+\sigma_{\nu,t}^2}\Big)}_{=1}\Bigg)=\beta\varphi_j$$

$$\Leftrightarrow\ \varphi_\mu^j=\frac{\beta}{1-\beta}\varphi_j \tag{F.24}$$

Analogously to the autoregressive model, the value function difference between the deterministic and the Bayesian learning model is determined by the contribution from the informational state, which is zero in the present by assumption (or rather calibration), and the contribution of the affine parts of the value function. Here, the affine part of the Bayesian learning model is

$$\begin{aligned}
\varphi_t^{Bayes}=\ &\beta\varphi_{t+1}^{Bayes}+B(\cdot)+\beta\varphi_{M,1}\Big(\textstyle\sum_{i=1}^{I^d}E_{i,t}^*+E_t^{exo}\Big)+\beta\varphi_{\tau,1}\frac{\sigma^{forc}}{M^{pre}}G_t\\
&+\alpha\beta\Big(\varphi_j+\varphi_\mu^j\frac{\sigma_{\epsilon,t}^2}{\sigma_{\epsilon,t}^2+\sigma_{\nu,t}^2}\Big)^2\frac{\sigma_{\epsilon,t}^2+\sigma_{\nu,t}^2}{2}
\end{aligned}$$

where the last term is new with respect to the deterministic equation. Therefore, the value function difference is

$$\begin{aligned}
\Delta W^{Bayes} &= V^{Bayes} - V^{det} = \varphi_\mu^j \mu_0^j + \varphi_0^{unc} - \varphi_0^{det} \\
&= 0 + \beta(\varphi_1^{Bayes} - \varphi_1^{det}) + \alpha\beta\Big(\varphi_j + \varphi_\mu^j \frac{\sigma_{\epsilon,0}^2}{\sigma_{\epsilon,0}^2+\sigma_{\nu,0}^2}\Big)^2 \frac{\sigma_{\epsilon,0}^2 + \sigma_{\nu,0}^2}{2} \\
&= \beta(\varphi_2^{AR} - \varphi_2^{det}) + \sum_{\tau=0}^{1} \alpha\beta^{\tau+1}\Big(\varphi_j + \varphi_\mu^j \frac{\sigma_{\epsilon,\tau}^2}{\sigma_{\epsilon,\tau}^2+\sigma_{\nu,\tau}^2}\Big)^2 \frac{\sigma_{\epsilon,\tau}^2 + \sigma_{\nu,\tau}^2}{2} \\
&= \lim_{t\to\infty} \beta(\varphi_t^{AR} - \varphi_t^{det}) + \lim_{t\to\infty} \sum_{\tau=0}^{t-1} \alpha\beta^{\tau+1}\Big(\varphi_j + \varphi_\mu^j \frac{\sigma_{\epsilon,\tau}^2}{\sigma_{\epsilon,\tau}^2+\sigma_{\nu,\tau}^2}\Big)^2 \frac{\sigma_{\epsilon,\tau}^2 + \sigma_{\nu,\tau}^2}{2} \\
&= \sum_{\tau=0}^{\infty} \alpha\beta^{\tau+1}\Big(\varphi_j + \varphi_\mu^j \frac{\sigma_{\epsilon,\tau}^2}{\sigma_{\epsilon,\tau}^2+\sigma_{\nu,\tau}^2}\Big)^2 \frac{\sigma_{\epsilon,\tau}^2 + \sigma_{\nu,\tau}^2}{2} \\
&= \sum_{\tau=0}^{\infty} \alpha\beta^{\tau+1}\Big(\varphi_j + \frac{\beta}{1-\beta}\varphi_j \frac{\sigma_{\epsilon,\tau}^2}{\sigma_{\epsilon,\tau}^2+\sigma_{\nu,\tau}^2}\Big)^2 \frac{\sigma_{\epsilon,\tau}^2 + \sigma_{\nu,\tau}^2}{2},
\end{aligned}$$

where the last line uses equation (F.24). Moreover, the term in brackets is equivalent to

$$\varphi_j \frac{(1-\beta)(\sigma_{\epsilon,\tau}^2+\sigma_{\nu,\tau}^2) + \beta\sigma_{\epsilon,\tau}^2}{(1-\beta)(\sigma_{\epsilon,\tau}^2+\sigma_{\nu,\tau}^2)} = \varphi_j \frac{\sigma_{\epsilon,\tau}^2 + (1-\beta)\sigma_{\nu,\tau}^2}{(1-\beta)(\sigma_{\epsilon,\tau}^2+\sigma_{\nu,\tau}^2)}$$

$$= \frac{\varphi_j}{1-\beta} \underbrace{\Big(\frac{\sigma_{\epsilon,\tau}^2}{\sigma_{\epsilon,\tau}^2+\sigma_{\nu,\tau}^2} + (1-\beta)\frac{\sigma_{\nu,\tau}^2}{\sigma_{\epsilon,\tau}^2+\sigma_{\nu,\tau}^2}\Big)}_{\equiv\Omega_t}$$

delivering

$$\sum_{\tau=0}^{\infty} \alpha\beta^{\tau+1}\varphi_j^2\Big(\frac{\Omega_t}{1-\beta}\Big)^2 \frac{\sigma_{\epsilon,\tau}^2 + \sigma_{\nu,\tau}^2}{2},$$

as stated in part (2) of the proposition.

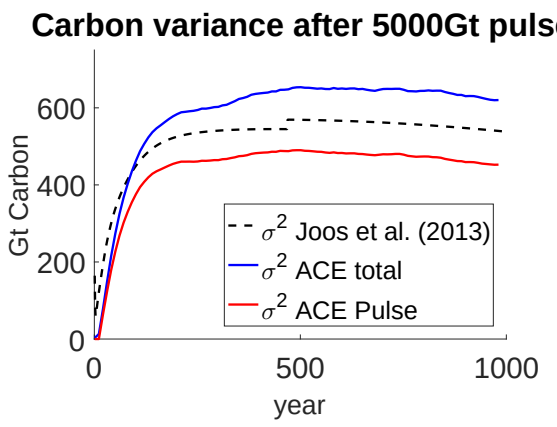
G Calibration of the stochastic processes

The climate sensitivity measures the medium to long-term response of global warming to a doubling of the carbon dioxide concentrations. Because this medium to long-term response takes a few centuries, integrated assessment modelers increasingly calibrate

their models to the transient climate response (TCR). TCR measures the warming response to a scenario that increases atmospheric CO_2 concentrations from preindustrial levels by one percent yearly until concentrations doubled w.r.t. preindustrial (7 decades), keeping concentrations constant afterwards. TCR measures the average temperature increase during the two decades centered at the year when concentrations doubled. TFE.6 Figure 2 IPCC (2013) shows a set of different probabilistic TCR distributions that share the slight positive skewness of ACE's TCR depicted in Figure 4. The IPCC (2013) summarizes the mean TCR prediction of 30 models (CMIP5) as $1.8°C$ and reports the 66% probability interval of TCR as $1°C$ to $2.5°C$, see black "x" and black bars in the right panel of Figure 4. In addition, the calibrating of ACE's TCR distribution was guided by my argument discussed in section 4.4 that the persistence of the autoregressive shock model should lie somewhere between 0.9 and 0.99 to reflect epistemological uncertainty over the coming decades, and that $\eta \in \{0.5, 1\}$ reflects well that the uncertainty should increase somewhat more steeply in the perturbation of the climate system than a lower η would suggest. I varied all parameters substantially, and the calibration $\gamma = 0.95$, $h = 0.23$, $\eta = 0.8$, $c = 0.21$, and $\epsilon = 0.05$ depicted in Figure 4 most closely resembled the look and stated moments of the TCR information provided by the IPCC (2013).

It proved more difficult to find good probabilistic information governing carbon dynamics. The current carbon budget has a missing sink of a little over 1GtC *per year*. It has been somewhat stable, but the fear is that we cannot predict how it responds to further perturbation of the climate system. In addition, there is uncertainty how the known flows between sinks and sources respond to climate change. I assume that a *change* in carbon flows governed by a volatility of 5GtC per *decade* in response to doubling atmospheric CO_2 is a reasonable order of magnitude proxy of carbon flow uncertainty. Once more, I set $\eta = 0.8$ to reflect that the uncertainty should increase somewhat more steeply in the perturbation of the climate system. Then, setting $\delta^{Mx} = \delta^{M\sigma} = 25$ implies that a doubling of preindustrial concentrations adds a 5Gt variance shock per decade to the conditional mean. I set $\delta^{\sigma x} = 1$ to explore the impact of stochastic volatility affecting not only carbon flows directly, but also their conditional expectations with a comparable order of magnitude. Again, I choose a persistence of the autoregressive shock models as 0.95 to reflect epistemological uncertainty and to make it comparable to the uncertainty in the temperature response.

I compare the resulting stochastic carbon dynamics to the model comparison study by Joos et al. (2013) who subject 18 different carbon cycle models to a 5000Gt carbon pulse. The study calculates the variance of the pulse evolution across models. ACE



- Black dashed: Uncertainty (variance) across 18 deterministic carbon cycle models responding to a 5000Gt carbon pulse as reported by Joos et al. (2013). After year 500 several models leave the ensemble causing a discontinuity in the variance.
- Blue: Overall uncertainty (variance) in ACE when subjected to the same pulse.
- Red: Variance measure of uncertainty attributable to the carbon pulse itself.

Figure 7: Uncertainty of response to a 5000Gt carbon pulse.

is a single stochastic model rather than a set of different deterministic models. Thus, not only the evolution of the pulse will be uncertain, but also the baseline evolution of atmospheric carbon dioxide. For my order of magnitude comparison, I calculate the overall uncertainty in ACE as well as a measure of the uncertainty of the pulse itself.[44] Figure 7 presents the result. Overall, the magnitude of the uncertainty is close.

[44]To obtain a measure of the uncertainty adhering to the pulse itself in an overall stochastic scenario, I draw 2000 paths for the random variables and simulate the model with and without the carbon pulse. I then calculate the variance in the difference of the carbon evolution with and without the pulse, taking differences between the paths with coinciding draws of the shock.