DISCUSSION PAPER SERIES

DP16065

Uncertainty in the Analytic Climate Economy

Christian Traeger

CLIMATE CHANGE RPN

CEPR

ISSN 0265-8003

Uncertainty in the Analytic Climate Economy

Christian Traeger

Discussion Paper DP16065
Published 22 April 2021
Submitted 16 April 2021

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

Uncertainty in the Analytic Climate Economy

Abstract

The paper analyzes optimal climate policy under uncertainty. It endows a recent quantitative analytic integrated assessment model (IAM) with long-run risk, adapting methods from the asset pricing literature to deal with endogenous climate risk. The model solves in closed-form for general degrees of risk aversion, stochastic climate feedbacks, and a stochastic damage-adaptation process. The model permits an exact solution of the infinite horizon stochastic fixed-point problem of a complex IAM. The approach facilitates new quantitative evidence for the role of uncertainty as well as analytic insights into the drivers and sensitivities of the optimal carbon tax facing an uncertain future.

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

Keywords: climate change, Integrated assessment, uncertainty, risk aversion, Social cost of carbon, Damages, Adaptation, Long-run risk, endogenous risk, stochastic volatility

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

Uncertainty in the Analytic Climate Economy*

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

March 2020

Abstract: The paper analyzes optimal climate policy under uncertainty. It endows a recent quantitative analytic integrated assessment model (IAM) with long-run risk, adapting methods from the asset pricing literature to deal with endogenous climate risk. The model solves in closed-form for general degrees of risk aversion, stochastic climate feedbacks, and a stochastic damage-adaptation process. The model permits an exact solution of the infinite horizon stochastic fixed-point problem of a complex IAM. The approach facilitates new quantitative evidence for the role of uncertainty as well as analytic insights into the drivers and sensitivities of the optimal carbon tax facing an uncertain future.

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

Keywords: climate change, integrated assessment, uncertainty, risk aversion, recursive utility, social cost of carbon, carbon tax, carbon cycle, temperature, damages, adaptation, climate sensitivity, long-run risk, endogenous risk, stochastic volatility, autoregressive gamma

*The present paper emerged out of Traeger (2015), which I recently split into the accompanying paper Traeger (2021*a*), introducing the deterministic base model, and the present paper focusing on uncertainty. I am grateful for feedback and inspiration from Larry Karp, Terry Iverson, Svenn Jensen, Bard Harstad, Kjetil Storesletten, Armon Rezai, Buzz Brock, Lars Hansen, Rick van der Ploeg, Michael Greenstone, Andre Butz, Fortunat Joos, David Anthoff, Valentina Bossetti, Richard Tol, Christian Gollier, Drew Creal, Nour Meddahi, Ravi Bansal, Till Requate, Andreas Lange, Grischa Perino, Rob Nicholas, Klaus Keller, Nancy Tuana and participants of the PET 2015, AERE 2015, SURED 2016, CESifo 2016, SITE 2017, ASSA 2018, WCERE 2018, ESEM 2018, and of department seminars at Stanford University, University of Amsterdam, University of Miami, University of Oslo in 2014, College de France, University of Toulouse, London School of Economics, Ohio State, University of Cambridge in 2015, University of California, Berkeley, University of Colorado, Iowa State, University of Chicago, CESifo Munich in 2016, University of Gothenburg, 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.

1 Introduction

Climate change is one of our major challenges directly linking long-term economic activity with the need for regulation. The World Economic Forum's (2019) Global Risk Report places climate risks at the top of the list. Understanding optimal policy requires long-term structural analysis. The present paper builds on the Analytic Climate Economy (ACE) to examine the relation between climate risks and optimal climate policy. In doing so it combines quantitative and analytic results.

ACE is an integrated assessment model of climate change (IAM) that is on a par with the typical numeric IAMs used in policy advising, relying on a state of the art climate dynamics representation (Traeger 2021*a*). It endorses a general production system relying on a variety of renewable and (potentially scarce) non-renewable energy inputs that are of limited and time-changing substitutability. In contrast to other models of similar complexity and descriptive power, ACE solves a general IAM in closed form. The present analysis takes this model into a stochastic setting, acknowledging a set of crucial real world uncertainties and risks. It solves for today's optimal policy in a forward looking infinite horizon stochastic fixed-point problem that describes the complex interaction of economic activity and climate change. The main rigidity of the deterministic ACE is an assumption of log-utility. Maintaining a unit elasticity of intertemporal substitution, I solve the model for arbitrary levels of relative Arrow-Pratt risk aversion in a framework disentangling risk aversion from the desire to smooth consumption over time (Epstein-Zin preferences).

Nordhaus's (2017) "Nobel-awarded" DICE model recommends slightly more than a doubling of atmospheric carbon dioxide (CO_2) concentrations w.r.t. preindustrial levels along an optimal trajectory. The warming resulting from a doubling of atmospheric CO_2 is known as the *climate sensitivity*. Its best guess is a 3°C (degree Celsius) warming. Yet, the latest assessment report by the Intergovernmental Panel on Climate Change eliminates this best guess and merely states that climate sensitivity lies with 66% confidence between 1.5°C and 4.5°C, and with a probability of up to 10% above 6°C (IPCC 2013). Already a 1.5°C warming takes us beyond the temperature range that our planet experienced over the past 100 000 years. A warming above 3°C moves us beyond the range experienced over the past million years. As a result, it is unsurprising that current estimates of the damages caused by future climate change at such a level of warming vary widely with, e.g., both 2% (Nordhaus 2017) and 10% (Howard & Sterner 2017, Pindyck 2020) of world output

being somewhat popular estimates. I let ACE tackle these uncertainties explicitly. As a result, I derive the benchmark policy where every policy maker anticipates the vast set of possible futures and future responses.

ACE is the first analytic IAM that models carbon and temperature dynamics explicitly. As Traeger (2021*a*) demonstrates, under certainty, the carbon cycle delivers the main carbon tax multiplier whereas temperature dynamics slightly reduces the tax because of ocean cooling and a resulting warming delay. The present paper shows that uncertainty over carbon flows has only a minor impact on the optimal tax level. In contrast, uncertainty about the temperature feedbacks causes a major "risk premium". The reason lies in the non-linearity of the greenhouse effect that governs the relation between atmospheric CO_2 and the resulting temperature increase. The direct greenhouse effect (radiative forcing) is logarithmic in the carbon concentration. This concavity reduces the impact of the uncertainty governing emission flows, but does not affect the impact of the uncertainty governing temperature feedbacks.

Weitzman (2009) demonstrates in a simple, stylized, and widely cited model that uncertainty potentially has a huge impact on the SCC, and he suggests that uncertainty can render redundant the notorious debate over time preference. ACE takes the structure and timing of the economy-climate interaction seriously, requiring a somewhat more complex and calibrated model. As a result, I find that (i) the uncertainty contribution adds about 50% to the deterministic SCC under standard calibration procedures, (ii) this risk premium is convex in almost all SCC components, (iii) time preference is particularly crucial to the relevance of the risk premium, (iv) fat-tails do not make time preference less, but rather more important; the sensitivity to pure time preference grows in the order of the moments of the uncertainty contributions. In contrast to Weitzman's (2009) hope, the latter finding makes the calibration of time preference even more relevant under uncertainty than it already is under certainty. A standard IAM calibration procedure suggests an annual rate of 1.4%. In contrast, a recent expert survey finds that over 50% of the experts recommend a rate less than half this number (Drupp et al. 2018). Cutting the time preference by half to $\rho = 0.7\%$ increases the SCC under uncertainty from almost $150\frac{USD}{tCO_2}$ to over $600\frac{USD}{tCO_2}$; here, the risk premium makes up close to three quarters of this total value.

Climate change is an uncertain long-run problem. Guided by the long-run risk literature in asset pricing, ACE disentangles risk aversion from consumption smoothing, so as to separately calibrate the risk-free discount rate and risk premia. Models lacking this feature either discount the future too highly, or disregard the risk premia

(equity premium and risk-free rate puzzles). Jensen & Traeger (2014) introduced a simplified model of long-run *growth* risk based on Bansal & Yaron (2004) and Croce (2014) into the IAM literature, recently also adopted and recalibrated by Cai & Lontzek (2019). The long-run risk literature is typically concerned with asset *valuation*, where the risk is exogenous to the individual decision maker. By contrast, I focus on climate uncertainty and climate *policy* where the *risk is endogenous*, and I extend the long-run risk model to endogenous climate change risk.

I find that uncertainty in the present model always raises the optimal carbon tax. I explain that such positive risk premia result from the interaction of the risk's endogeneity with the temperature non-linearity and (intrinsic) risk aversion. A temperature non-linearity (or risk aversion) by itself causes a *risk premium* that reduces welfare. Yet, a welfare loss does not necessarily affect policy. The *policy*'s risk premium is generated by the following mechanism. As we emit more CO_2, uncertainty about the future climate increases. This increase in uncertainty increases the welfare loss and implies an additional *incentive to strengthen the mitigation policy*. I show how different risks are mutually aggravating in their policy impact. The formulas reveal that, compared to the deterministic SCC, the risk premium increases much more strongly in the factors characterizing climate change, e.g., the (semi-) elasticity of production to global warming ("damage parameter") and the climate related multipliers. Thus, if the climate problem becomes or is judged more severe, then also the importance of incorporating uncertainty becomes more serious.

For a wide class of stochastic processes and production models, Lemma 1 breaks the "curse of dimensionality", translating a high-dimensional dynamic programming problem that is difficult or impossible to solve on a computer into a simple system of nonlinear equations. I achieve this result by moving from the state space, describing the physical characteristics of the system, to the space of shadow values describing their marginal values (for a set of affine stochastic processes). I then analyze a subspace of solutions in closed-form. An accompanying paper discusses the consequences of the present changes in the SCC for the optimal CO_2 emissions in IAMs (Traeger 2021*b*).

Literature. A first generation of analytic integrated assessment models (AIAMs) examines uncertainty in the context of climate policy instrument choice (Hoel & Karp 2002, Newell & Pizer 2003, Karp & Zhang 2006, Karp & Zhang 2012). These papers use a linear quadratic model where *welfare* responds to uncertainty, but *policy* remains unaffected by risk (weak certainty equivalence). An early exception to this

weak equivalence is Hoel & Karp (2001) who use multiplicative instead of additive shocks. More recently, Valentini & Vitale (2019) use a linear-quadratic Gaussian control model to re-examine the shape of the "mitigation ramp" under risk sensitive preferences. Karydas & Xepapadeas (2019) connect the framework to asset markets and show that climate change should reduce the holding of carbon intensive portfolios. A disadvantage of these linear quadratic approaches is the highly stylized representation of the economy and the climate system. In particular, these models have no production or energy sector.

Golosov et al. (2014) sparked a second generation of integrated assessment models by amending the log-utility and full-depreciation version of Brock & Mirman's (1972) stochastic growth model with an explicit energy sector producing CO_2 emissions. Applications include multi-regional settings (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*). Gerlagh & Liski (2018*b*) add the empirically important delay between emission accumulation and damages and the accompanying Traeger (2021*a*) builds an explicit model of temperature dynamics and introduces the non-linear greenhouse effect into the ACE model adopted here. Golosov et al. (2014) framework imposes *strong certainty equivalence*: not even welfare responds to uncertainty. I show that this feature arises from the absence of the non-linearities in the climate system and from simultaneously setting the intertemporal elasticity of substitution and Arrow–Pratt risk aversion to unity. Whereas unity might be a somewhat reasonable estimate 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 increasing absolute Arrow–Pratt risk aversion or risk neutrality.

Karp (2017) explains that Golosov et al.'s (2014) model solves analytically because it can be transformed to a system that is linear in the states' equations of motion. The endogenous uncertainty in the present paper no longer follows this linearity in the states and merges the original approach with affine stochastic processes developed for asset pricing (Gourieroux & Jasiak 2006, Le et al. 2010). 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, normative, or policy choice. Alternatively, Li et al. (2016), Anderson et al. (2014), and von zur Muehlen (2018) depart from the world of von Neumann & Morgenstern's (1944) axioms and introduce a preference for robustness, also leaving behind the strong certainty equivalence of Golosov et al.'s (2014) framework.

In more stylized climate economic settings, Ha-Duong & Treich (2004) disentangle risk aversion from intertemporal substitutability in climate change evaluation, using a simple two-period model. Traeger (2014) analyzes the role of baseline and payoff uncertainty of mitigation and adaptation projects and their (potentially ambiguous) correlation through the lens of the discount rate. Bansal et al. (2019) enrich the consumption-based long-run-risk asset pricing model by temperature-triggered catastrophes, estimating the SCC from the capital market's response to temperature fluctuations. Lemoine (2021) analyzes climate uncertainties and the roles of insurance versus precautionary motives in a related setting without decision makers. Dietz et al. (2018) and van den Bremer & van der Ploeg (2018) translate these reasonings into integrated assessment models defining "climate betas". Dietz et al. (2018) use a simple two-period model in combination with a DICE-based Monte-Carlo simulation and van den Bremer & van der Ploeg (2018) derive an approximate analytic perturbation solution. With their more stylized AK-based IAM, van den Bremer & van der Ploeg (2018) also complement the present analysis by providing approximate solutions for the SCC in a setting where the elasticity of intertemporal solution differs from unity. Hambel et al. (2018*a*) develop a simple analytic continuous time IAM, which they employ for a regional game showing that trade induces countries to incorporate some of the emission externalities they impose on other regions.

Kelly & Kolstad's (1999) and Pizer (1999) build stochastic implementations of Nordhaus' DICE model, pioneering the field of numeric stochastic integrated assessment. Pizer (1999) log-linearizes DICE to find optimal CO_2 abatement under a variety of stochastic ("one-shot") realizations of economic, climate, and preference parameters. Kelly & Kolstad's (1999) introduce Bayesian learning over the climate's sensitivity into a recursive stochastic dynamic programming implementation of DICE. The paper focuses on the speed of learning rather than optimal climate policy. Leach (2007) shows in a closely related IAM that learning the true climate sensitivity is even slower if the persistence of climate dynamics is also unknown. Crost & Traeger (2014) use a simplified recursive stochastic version of DICE to analyze the implica-

tions of stochastic damages under Epstein-Zin preferences. Given their shocks are not persistent (long-run), they only have a minor impact on optimal climate policy.

Most closely related to the present paper, Jensen & Traeger (2013) and Kelly & Tan (2015) analyze climate sensitivity uncertainty in a simplified DICE model and find an increase of the optimal carbon tax in response to temperature uncertainty similar to that of my base calibration. Kelly & Tan (2015) also show that "learning away" fat tails is faster than the learning about the mean climate sensitivity. Daniel et al. (2019) find in a discretized finite horizon implementation of DICE with uncertainty over climate sensitivity that Epstein-Zin preferences can lead to a decreasing rather than increasing tax trajectory. Like the DICE-based Jensen & Traeger's (2014), also the present ACE-based analysis cannot confirm such a finding in the infinite time-horizon setting. Rudik & Lemoine (2017) use a numeric approximation (Smolyak grid) to analyze the policy response to climate sensitivity in the full DICE and find a substantially lower risk premium. Their paper provides an excellent survey over the literature of numeric stochastic integrated assessment modeling. Hambel et al. (2018*b*) build their own numeric stochastic AK-based IAM featuring a more stylized climate system and analyze the SCC under a variety of uncertainties and damage functions. Rudik (2020) introduces learning and robust control into a DICE model with stochastic damages; the paper confirms that the shock's persistence generated by learning proves crucial to generate risk premia. Barnett et al. (2020) develop a continuous time reduced-form integrated assessment model with closed-form solution and explore a variety of different decision criteria including disentanglement between risk aversion and intertemporal substitution, ambiguity aversion, and preferences for robustness.

Keller et al. (2004) introduce an uncertain climate threshold into the DICE model, which potentially triggers large damages. Lemoine & Traeger (2014, 2016) analyze the implications of a variety of such thresholds and their possible interactions in a recursive stochastic implementation of DICE with Bayesian updating. Lontzek et al. (2015), Cai et al. (2015), Cai et al. (2016), and Cai & Lontzek (2019) analyze optimal policy in DICE when rising temperatures increase the probability of triggering regime shifts in the damage function and Van der Ploeg & de Zeeuw (2018) discuss precautionary savings in the context of such a climate tipping point. Golub et al. (2014), Heal & Millner (2014), Brock & Hansen (2018), and Berger & Marinacci (2020) survey climate uncertainties and possible decision criteria. Kotlikoff et al. (2021) analyze intergenerational financing constraints of optimal climate policy in a

stochastic overlapping generations integrated assessment model.

2 Model and General Solution

This section summarizes the deterministic "base ACE", introduces a wide class of generic stochastic processes, explains the Bellman equation, and shows how to translate the complex stochastic dynamic programming problem into a set of simple algebraic equations.

2.1 Summary of the Deterministic Base Model

ACE's structure follows (and generalizes) that of most IAMs, see Figure 1. This section summarizes the "base ACE" or *D*eterministic *ACE*. For a detailed discussion of the corresponding equations, interpretations, calibration and for various extensions I refer to Traeger (2021*a*), henceforth DACE. As I show in DACE, population weighting, a distinction between investment and consumption goods, and a combination of log-utility with CES preferences over a variety of goods slightly modify the formula for the optimal tax under certainty stated here (Section 2.5). Yet, they generally do not affect the relation between the stochastic SCC and the deterministic SCC derived in the present paper. Thus, the model below is a bare bones version of a larger class of models to which the results derived in the present paper apply.

Technology, capital, labor, and energy produce (gross) output Y_t that is either consumed or invested. The only requirement on the production function $F(\cdot)$ is homogeneity in capital, i.e., the existence of κ such that

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

and that the resulting optimization problem is well-defined. The technology vector $\boldsymbol{A}_t \in \mathbb{R}^I$ evolves exogenously, and total labor N_t and capital K_t are optimally distributed across the different sectors at the beginning of the period, giving rise to the vectors $\boldsymbol{N}_t, \boldsymbol{K}_t \in \mathbb{R}^I$ characterizing this sectoral distribution. I define the normalized capital distribution over sectors by the vector $\boldsymbol{\mathcal{K}}_t \equiv \frac{\boldsymbol{K}_t}{K_t}$. See DACE for specific production examples. The first I^d energy sources $E_1, ..., E_{I^d}$ are fossil fuels and measured by their CO_2 content so that CO_2 emissions from production are $\sum_{i=1}^{I^d} E_{i,t}$. If essential for production, fossil resources can be scarce and the model endogenously

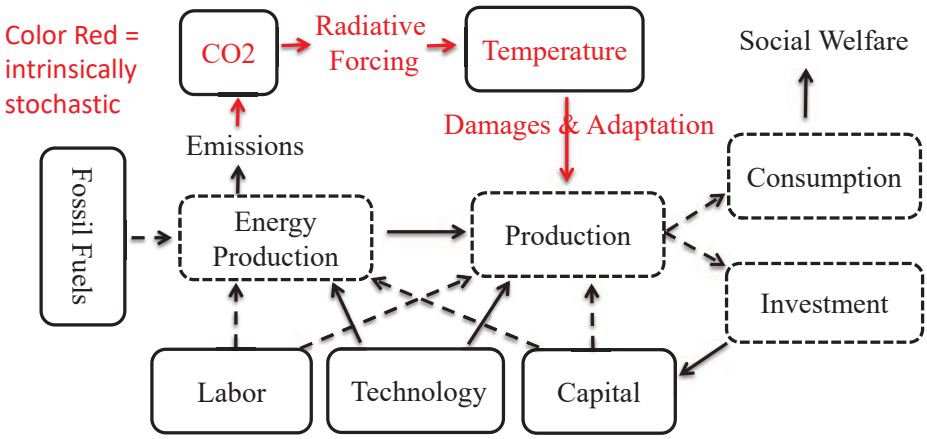


Figure 1: The structure of ACE and most IAMs. In red intrinsically stochastic variables that drive the uncertainty (all future outcomes are stochastic). Solid boxes characterize the model's state variables, dashed boxes are flows, and dashed arrows mark choice variables.

implies a scarcity rent following Hotelling's rule. The level of fossil resources in the ground is summarized by the vector $\boldsymbol{R}_t \in \mathbb{R}^{I_d}$. It evolve as

$$\boldsymbol{R}_{t+1} = \boldsymbol{R}_t - \boldsymbol{E}_t^d \tag{2}$$

and has to remain non-negative. In addition, land conversion, forestry, and agriculture emit smaller quantities of CO_2, labeled E_t^{exo}, that ACE treats as exogenous. Emissions accumulate in the atmosphere following a standard carbon cycle model

$$\boldsymbol{M}_{t+1} = \boldsymbol{\Phi}\boldsymbol{M}_t + \boldsymbol{e}_1(\textstyle\sum_{i=1}^{I^d} E_{i,t} + E_t^{exo}) \tag{3}$$

where $M_{1,t}$ denotes atmospheric CO_2 and $M_{2,t}, ..., M_{m,t}$, $m \in \mathbb{N}$, denote the carbon content of other reservoirs such as the oceans, plant-matter, and soils, and the first unit vector $\boldsymbol{e}_1$ reflects that we emit CO_2 into the atmosphere. For a quantification I will use the DICE 2013 carbon cycle; it 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.[1] Atmospheric CO_2, together with other (exogenous) greenhouse gases G_t, causes a greenhouse effect referred to as anthropogenic radiative forcing

$$F_t = \eta \frac{\log \frac{M_{1,t}+G_t}{M_{pre}}}{\log 2} \ . \tag{4}$$

[1]Van der Ploeg et al. (2020) show that DICE's 2013 carbon cycle performs better than the 2016 vintage for calculating the SCC. DACE shows that it gives almost the same results as using Joos et al.'s (2013) frequently promoted model employed, e.g., in IPCC (2013).

I emphasize the concavity of the greenhouse effect in atmospheric CO_2 because of a "saturation".

The planet's long-term atmospheric temperature increase $T_{1,eq}$ is proportional to the radiative forcing F_{eq}. However, temperatures respond with a delay to the greenhouse effect because the oceans keep cooling us for a while. Let $T_{i,t}$, $i \in \{2, ..., l\}$, denote the temperature in different ocean layers. DACE shows that the climate's temperature dynamics can be approximated using a non-linear temperature "diffusion" model. Defining transformed temperatures $\tau_{i,t} = \exp(\xi_1 T_{i,t})$, $i \in \{1, ..., l\}$, and stacking them into the vector $\boldsymbol{\tau}_t$, the transformed diffusion model becomes linear and the corresponding equation of motion is

$$\boldsymbol{\tau}_{t+1} = \boldsymbol{\sigma}\boldsymbol{\tau}_t + \sigma^{forc}\frac{M_{1,t} + G_t}{M_{pre}}\boldsymbol{e}_1 \ , \tag{5}$$

where the matrix $\boldsymbol{\sigma}$ specifies the heat transfer coefficients and σ^{forc} characterizes the responsiveness to radiative forcing.

An increase in atmospheric temperature causes economic damages

$$D(T_{1,t}) = 1 - \exp(-\xi_0 \exp[\xi_1 T_{1,t}] + \xi_0) \tag{6}$$

that specify a percentage loss of world output in double-exponential form. The generic capital stock follows the equation of motion

$$K_{t+1} = \underbrace{Y_t[1 - D_t(T_{1,t})]}_{\equiv Y_t^{net}}(1 - x_t)\left[\frac{1 + g_{k,t}}{\delta_k + g_{k,t}}\right] \ , \tag{7}$$

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. DACE shows that equation (7) coincides with the standard equation for capital accumulation $K_{t+1} = Y_t[1 - D_t(T_{1,t})] - C_t + (1 - \delta_k)K_t$ if the exogenously calibrated capital growth rate is correct, $g_{k,t} = \frac{K_{t+1}}{K_t} - 1$, or if $\delta_k = 1$ (full depreciation). Based on the Penn World tables, the correction factor is $\left[\frac{1+g_{k,t}}{\delta_k+g_{k,t}}\right] \approx 1.8$ for a decadal time step and allows ACE to match observed capital dynamics and the capital-output ratio. The model will be solved using log-capital $k_t \equiv \log K_t$.

2.2 General Uncertainty

I focus on the uncertainty generated by the evolution of carbon, temperature, and damages (including uncertainty over potential adaptation). This uncertainty turns

(all) the equations of motion stochastic. Formally, the state variables become functions on a probability space, whose filtration $\mathcal{F}$ is generated by the stochastic climate and damage evolution (my notation will suppress this dependence). Scenarios with persistent shocks or structural learning require additional informational state variables that I denote by the vector $\boldsymbol{I}_t \in \mathbb{R}^n$, 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 have already observed.

Different sections of the paper will make different assumptions regarding the specific processes governing the uncertain future. They all have in common that they assume *affine stochastic processes*. Let $\boldsymbol{X}_t$ denote the $1 + l + m + n$ dimensional state vector stacking the state vectors k_t, $\boldsymbol{\tau}_t$, $\boldsymbol{M}_t$, and $\boldsymbol{I}_t$.[2] The stochastic process governing the state vector $\boldsymbol{X}_t$ is affine if its conditional cumulant generating function is linear in the state. The cumulant-generating function is the logarithm of the moment-generating function. In detail, there has to exist an $\mathbb{R}$-valued function $a(\cdot)$ and a $\mathbb{R}^{1+l+m+n}$-valued function $\boldsymbol{b}^\top(\cdot)$ such that

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

for all $\boldsymbol{z} \in G \subset \mathbb{R}^{1+l+m+n}$. The $^\top$ denotes transposition. So far, equation (8) merely defines a generic structure of a stochastic process. Lemma 1 will show how ACE solves for this class of stochastic processes. When solving the model, $\boldsymbol{z}$ will relate closely to the the state's shadow values, and equation (8) will have to hold on a domain incorporating the relevant state space and shadow-value domains. Both functions $a(\cdot)$ and $\boldsymbol{b}^\top(\cdot)$ depend on the model's parametrization and the functional forms chosen for a particular specification, e.g., those of Section 3.

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

[2]The resources state vector $\boldsymbol{R}_t$ is not part of $\boldsymbol{X}_t$. X_t will contain states that are one-step-ahead stochastic (generating the uncertainty); including the resources state would complicate dealing with the non-negativity constraint of a resource.

2.3 Objective Function

Utility governing deterministic outcomes is logarithmic in consumption C_t and the social planner's time horizon is infinite with discount factor $\beta < 1$. 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 (see DACE). Logarithmic utility provides a reasonable description of intertemporal substitutability. However, the 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).[3] 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 concerns 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 address these issues following the asset pricing literature, an increasing strand of macroeconomic literature, and some recent numeric approaches to climate change assessment by using Epstein–Zin–Weil preferences. This approach accommodates a realistic coefficient of risk aversion and disentangles it from the unit elasticity of intertemporal substitution.[4] Then, the Bellman equation is

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

Expectations $\mathbb{E}_t$ are conditional on time t information ($\mathbb{E}_t(\cdot) \equiv \mathbb{E}(\cdot \,|\, \mathcal{F}_t)$). The Bellman equation uses a generalized expectation 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

[3] 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.

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

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. It will mostly be this parameter rather than Arrow–Pratt risk aversion that drives the risk premia. Traeger (2019) provides a direct axiomatic foundation of $-\alpha$ as a measure of intrinsic risk aversion. Figure 4 in Appendix B characterizes the total aversion deriving from the curvature of utility and the intrinsic risk attitude 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 lottery where she loses 5% of consumption if heads comes up, or gains 5.26% if tails comes up. Assume she consumes the equivalent of USD 1000 over some period. Then, a fifty percent probability loss of USD 50 will be compensated for by a fifty percent probably gain of USD 52.60. A calibration procedure described in Appendix B 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 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.

2.4 General Solution of the Stochastic ACE

The *stochastic ACE* is characterized by the production equation (1), the equations of motion for the resources (2) and the other states (8), the Bellman equation (9) and the constraints on labor, capital, and the resources.[5] The equation of motion (8) is generic, but the lemma below establishes that the deterministic ACE presented in equations (1-7) is a special case. As a result, Lemma 1 establishes how to extend the deterministic ACE to a large class of stochastic problems, some of which I will explore in the subsequent sections. ACE contains a large (possibly infinite, see Appendix A where information is captured by a series of moments) set of state variables, which makes it hard to impossible to solve the model numerically.[6] Lemma 1 transforms

[5] I also assume that the stochastically extended damage function keeps using the semi-elasticity of damages ξ_0 from equation (6), see e.g. the stochastic extension in equation (23). Otherwise, the final term in equation (10) would have to be adapted.

[6] Currently, Cai & Lontzek (2019) is probably the largest IAM solved numerically under uncertainty. It has substantially fewer states than even the least-states-version of the stochastic ACE. There are several approaches emerging in the context of machine learning that promise the ability to

this high dimensional stochastic dynamic programming 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.

Lemma 1 *Let the resources vector follow equation* (2)*. Let the state vector* $\boldsymbol{X}_t$ *comprising the remaining states follow an affine stochastic process of the form stated in equation* (8)*.*
1) The deterministic ACE summarized in Section 2.1 satisfies these assumptions (as a degenerate stochastic process).
2) An affine value function solves the stochastic ACE if the shadow value vector $\boldsymbol{\varphi}_X^\top = (\varphi_k, \boldsymbol{\varphi}_\tau^\top, \boldsymbol{\varphi}_M^\top, \boldsymbol{\varphi}_I^\top)$ *solves the algebraic system of equations*

$$\boldsymbol{\varphi}_X^\top = \frac{\beta}{\alpha}\boldsymbol{b}^\top(\alpha\boldsymbol{\varphi}_X) + \kappa\boldsymbol{e}_1^\top - \xi_0\boldsymbol{e}_2^\top \tag{10}$$

where $\boldsymbol{e_1^\top}$ *and* $\boldsymbol{e_2^\top}$ *denote the first and the second unit vectors (corresponding to the capital and the atmospheric temperature entries). The shadow value* $\varphi_{M,1}$ *determines the optimal carbon tax.*

Appendix C.1 provides the proof. 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), all of which will be explored in different sections (including the Appendix). The present paper focuses on closed-form solutions. More generally, equation (10) permits simple numeric solutions to an even wider class of stochastic integrated assessment models.

Lemma 1 has another general insight to offer. The affine stochastic processes in equation (8) are characterized by two functions, $a(\cdot)$ and $\boldsymbol{b}^\top(\cdot)$. The lemma shows that only the part characterized by $\boldsymbol{b}^\top(\cdot)$ enters the social shadow cost of carbon determining the optimal tax. As the proof points out, the part of the stochastic process entering $a(\cdot)$ only affects welfare but not optimal policy. The uncertainty conveyed by $a(\cdot)$ can generate large welfare losses and, yet, have no impact on the optimal policy. The specific formulations explored in subsequent sections will illustrate this point and emphasize that the risk premium on optimal policy is mostly driven by the interaction between the classical risk premia affecting welfare and the endogeneity of uncertainty, an interaction that operates through the function $\boldsymbol{b}^\top(\cdot)$.

solve models with much larger dimensionality, e.g., Brumm & Scheidegger (2017). I hope ACE can eventually serve as a benchmark for those approaches offering an exact solution of the base model.

2.5 The SCC in a Deterministic World

The SCC is defined as the money-measured present value welfare loss from releasing a ton of CO_2 into the atmosphere. ACE has a unique SCC that corresponds to the optimal Pigovian carbon tax. In a deterministic world, DACE derives

$$SCC_t^{det} = \frac{\beta Y_t^{net}}{M_{pre}} \xi_0 \left[(\mathbf{1} - \beta\boldsymbol{\sigma})^{-1}\right]_{1,1} \sigma^{forc} \left[(\mathbf{1} - \beta\boldsymbol{\Phi})^{-1}\right]_{1,1}, \tag{11}$$

where $[\cdot]_{1,1}$ denotes the first element of the inverted matrix in square brackets. The "static base" of the tax multiplies discounted net production per unit of preindustrial carbon with the semi-elasticity ξ_0 of output w.r.t. an exponential temperature increase $\tau_1 = \exp(\xi_1 T_1)$. These terms already carry the proper units of USD per ton of CO_2 emitted. This base value substantially increases as a result of a carbon cycle based multiplier $[(\mathbf{1} - \beta\boldsymbol{\sigma})^{-1}]_{1,1}$ and slightly decreases as a result of the temperature dynamics multipliers σ^{forc} and $[(\mathbf{1} - \beta\boldsymbol{\Phi})^{-1}]_{1,1}$ (see DACE for a quantification).

The transition matrix $\boldsymbol{\Phi}$ reflects that carbon does not decay but merely changes reservoirs. This stylized fact about the carbon cycle will be mirrored in the impact of carbon flows uncertainty on the SCC. Anticipating the result, it is worthwhile to follow DACE in interpreting the carbon multiplier in more detail. A Neumann series expansion of $\beta\boldsymbol{\Phi}$ delivers the identity $(1 - \beta\boldsymbol{\Phi})^{-1} = \sum_{i=0}^{\infty} \beta^i \boldsymbol{\Phi}^i$. Taking the i^{th} power of the transition matrix, the element $[\boldsymbol{\Phi}^i]_{j,1}$ characterizes how much of the carbon in the atmosphere (reservoir 1) travels into reservoir j after i time steps. E.g., $[\boldsymbol{\Phi}^2]_{1,1} = \sum_j \Phi_{1,j}\Phi_{j,1}$ characterizes the fraction of carbon leaving the atmosphere for layers $j \in \{1, ..., m\}$ in the first time step and arriving back to (or staying in) the atmosphere in the second time step. Thus, the term

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

represents a closed-form expression for the discounted sum of CO_2 persisting in and returning to the atmosphere over the course of an infinite future. The discount factor accounts for the delay between the act of emitting CO_2 and the resulting impact on temperature forcing (the greenhouse effect). Similarly, $carb_2 \equiv [(\mathbf{1} - \beta\boldsymbol{\Phi})^{-1}]_{1,2}$ characterizes the long-term greenhouse effect contribution from CO_2 that is currently in the second reservoir, e.g., the shallow ocean and biosphere. Given major uncertainties regarding the flow between the atmosphere and its adjacent reservoirs, the difference

$$\Delta carb \equiv carb_1 - carb_2 = \left[(\mathbb{1} - \beta\boldsymbol{\Phi})^{-1}\right]_{1,1} - \left[(\mathbb{1} - \beta\boldsymbol{\Phi})^{-1}\right]_{1,2} \tag{12}$$

will play a role in quantifying the SCC under uncertainty over carbon flows; the term will account for the fact that carbon is conserved and merely changes reservoirs.

3 Climate Change and Uncertainty

In the stochastic setting, the SCC is the welfare loss from releasing a ton of CO_2 anticipating all possible future scenarios and consequences. These consequences include the possibility that current emissions change the probabilities of different future scenarios. The present section analyzes the SCC under a stochastic evolution of carbon, temperature, and damages. I start with a brief introduction of the long-run risk model, a wide-spread tool in asset pricing, motivated in the context of climate change. The climate change application borrows heavily from the asset pricing literature. Yet, climate change requires a few extensions including functional forms of the stochastic processes and, most importantly, endogenizing uncertainty. Whereas the individual investor does not affect the market risk, global climate policy directly affects future climate and economic risk. The subsequent section will tackle one uncertainty at a time. They *build on the deterministic ACE model summarized in Section 2.1, replacing the corresponding equations governing carbon flows, temperature dynamics, and damages by a set of (affine) stochastic processes.*

3.1 Long-Run Risk and Climate Change

The causal origin of the greenhouse effect can be measured in the laboratory; carbon dioxide absorbs the planet's outgoing radiation. Yet, quantifying the warming resulting from a given CO_2 trajectory is difficult, and results in a relatively 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 concentrations. Second, the paper's introduction emphasized that, even if we knew future atmospheric carbon concentrations, we would remain highly uncertain about the implied warming (climate sensitivity). Finally, given we have not experienced the resulting temperature levels in the past hundreds of thousands of years, we are hugely uncertain about the implied damages. This section integrates carbon flow, temperature, and damage uncertainty into the ACE model. Given climate change is about the long-term damages of current actions, it is important to carefully model

long-run uncertainties. For this purpose, I build on insights and advances from the asset pricing 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 risk aversion 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 numeric integrated assessment of climate change.[7] Their long-run risk governs the exogenous growth of the economy, questioning that future generations are certainly much richer than present generations. This exogenous economic risk is limited to shocks to the expected mean, whereas asset pricing applications of the long-run risk model also rely 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. By contrast, optimally controlling the climate must 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 standard long-run risk model, applies it to climate, and extends it to *endogenous risk*. I start with uncertainty about future carbon flows, which permits building on the canonical long-run risk framework. The second section develops a novel version of the long-run risk process that permits a reasonable application to temperature, where the canonical model fails. The final section discusses damage uncertainty.

3.2 Uncertainty about Carbon Flows

The long-run risk model adds two stochastic processes to the equation of motion, here equation (3) governing carbon flows. The first process characterizes the long-term risk about the *level* of – here – carbon removal, whereas the second process governs the long-term risk about period-to-period *volatility* of these carbon flows. Formally, the first process $\boldsymbol{x}_t^M$ governs the conditional expectations, i.e., one-step-

[7] Jensen & Traeger (2014) use a simplified climate model to derive both approximate analytic and numeric solutions. Cai & Lontzek (2019) recently redo their numeric analysis with additional climate states and higher uncertainty of the consumption process. For this purpose, Cai & Lontzek (2019) tame the tail-risk in Jensen & Traeger's (2014) infinite time horizon stochastic fixed-point problem by using a discrete Markov chain (and a finite time horizon).

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 processes are highly persistent and thus describe long-run uncertainty. The resulting equation for the carbon stock, replacing equation (3) of the "base ACE", becomes

$$\boldsymbol{M}_{t+1} = \boldsymbol{\Phi}\boldsymbol{M}_t + (\textstyle\sum_{i=1}^{I^d} E_{i,t} + E_t^{exo})\boldsymbol{e}_1 + \boldsymbol{x}_t^M + \boldsymbol{\sigma}_t^M \mu_t^M \tag{13}$$

where $\mu_t^M \sim N(0,1)$ is a serially uncorrelated white noise, implying that $\boldsymbol{\sigma}_t^M$ indeed 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 is associated 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 carbon flows between the atmosphere and its neighboring reservoir.[8] The stochastic processes governing conditional expectations and variance are

$$x_{t+1}^M \;=\; \gamma^x x_t^M \;+ \delta^{Mx}\sqrt{\tfrac{M_{1,t}}{M_{pre}} - \eta_M}\;\chi_t^M \;+ \delta^{\sigma x}\;\sigma_t^M\;\omega_t^M \tag{14}$$

$${\sigma_{t+1}^M}^2 = \gamma^\sigma {\sigma_t^M}^2 + \quad \delta^{M\sigma}\left(\tfrac{M_{1,t}}{M_{pre}} - \eta_M\right) \quad + \quad \bar{\sigma}^M \nu_t^M. \tag{15}$$

The γ-*parameters* characterize the persistence of the shocks to the mean (first equation) and the volatility process (second equation) and satisfy $\gamma^x, \gamma^\sigma < 1$. As I show in Appendix A, epistemological uncertainty corresponds to a high persistence. Epistemological uncertainty expresses the "ignorance" of the modeler and, here, the scientific community. In the canonical long-run risk model, the overall risk is exogenous to the decision maker.

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 this crucial endogeneity of the risk, I introduce the second terms on the right. Their component $\frac{M_{1,t}}{M_{pre}} - \eta_M$ 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

[8] 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.

endogenous contribution to climate risk. For the stochastic volatility (equation 15), the last term on the right 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 (14) enables 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 enables long-run uncertainty about the volatility of carbon flows to affect the movement of conditional expectations.

Proposition 1 *The stochastic processes (13-15) governing carbon flows, replacing the deterministic equation* (3)*, change the deterministic* SCC^{det} *stated in equation* (11) *to*

$$SCC_t = SCC_t^{det}\Big(1 + \theta_M(\theta_M^*)\Big) \approx \frac{SCC^{det}}{1-\theta_M^*} \tag{16}$$

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

$$\theta_M^* = \frac{\alpha}{2}\frac{\beta\varphi_{M,1}^{det}}{M_{pre}}\Big[A^{M\to x^2} + A^{M\to\sigma} + A^{M\to\sigma}A^{\sigma\to x^2}\Big]\frac{(\Delta carb)^2}{carb_1} \tag{17}$$

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.*[9] *The uncertainty contribution* $\theta_M(\theta_M^*)$ *is convex in* θ_M^**.*[10]

Risk attitude and the relative importance of uncertainty. Equation (16) 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 (17) as θ_M^*. The basic intuition is the following. Risk causes a *welfare loss* as a result of risk aversion (and the difference in damage caused by CO_2 in the atmosphere versus the ocean). Because the risk is endogenous, more emissions raise this risk and the corresponding welfare loss, which increases the optimal *policy incentive* to reduce emissions. This contribution is proportional to intrinsic risk aversion α and the discounted shadow value of atmospheric carbon under certainty. Thus, it is the extent to which Arrow–Pratt risk aversion exceeds the (unit-) desire to smooth consumption over time that drives the SCC; α measures this difference

[9] 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 $\alpha\varphi_{M,1}^{det} > 0$.

[10] The solution is well-defined for $\theta_M^* < \frac{1}{4}$ (it solves a quadratic equation), which is met for all quantitative results in this paper.

and characterizes intrinsic aversion to risk, see section 2.3 and Appendix B. The proposition implies that the *uncertainty share* of the overall SCC (and thereby also $\frac{SCC_t}{SCC_t^{det}}$) *increases convexly in its deterministic base value* SCC_t^{det} because the shadow value of atmospheric carbon is itself proportional to SCC^{det}. *The more serious the climate problem, the larger the relative importance of uncertainty.*

The long-run risk channels. The three risk channels abbreviated $A^{M\to x}$, $A^{M\to\sigma}$, and $A^{\sigma\to x}$ disentangle the contributions from the different aspects of long-run risk. The quantification of these channels in Section 4.3 will clarify whether the uncertainty about flow levels ($A^{M\to x}$), flow volatility ($A^{M\to\sigma}$), or their interaction ($A^{\sigma\to x}$) are most relevant in the context of climate change (see Table in Figure 3). Each of these channels increases in the strength of the endogenous climate risk characterized by the δ-parameters. In addition, each channel's contribution increases in the *shock persistence* characterized by the γ-parameters. The formula shows that 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.[11]

Endogenous risk. All three risk channels are proportional to the δ-parameters, which scale the endogenous contribution of climate risk. Thus, the proposition shows that only the *endogenous uncertainty*, resulting from our climate perturbation, affects the optimal policy. The variance of the exogenous risk, e.g. $\bar{\sigma}^M$, has no impact on the optimal carbon price. As Appendix A shows, such exogenous risk reduces the welfare. Yet, to matter for the SCC, the risk has to interact with the carbon stock; *the damaging impact of uncertainty has to increase when releasing an additional ton of* CO_2. The proposition demonstrates that it is crucial how uncertainty enters the equations. In particular, the standard long-run risk model, featuring only exogenous risk, would miss this policy premium to the carbon tax. Lemma 1 shows this result more generally by stating that the affine components of the risk process (the $a(\cdot)$ function), here the risk independent of the carbon stock, do not affect the shadow values and thus optimal policy. The result is also a warning to be careful when

[11] The long-term uncertainty about expected carbon flows contributes quadratically (conditional expectations channel $A^{M\to x}$). The stochastic volatility channel $A^{M\to\sigma}$ enters twice; it reflects the harm imposed by an increase in future volatility that results from a perturbation of the climate system. Its first appearance in equation (17) captures long-run uncertainty about the volatility of carbon flows. Its second appearance, when turning on the interaction $\delta^{\sigma x}$, captures stochastic volatility governing the long-run level of carbon flows (volatility of long-term conditional expectations); here the combined channel $A^{M\to\sigma}A^{\sigma\to x^2}$ reflects the fact that an increase in carbon increases stochastic volatility, which in turn increases the uncertainty over long-run expectations.

using (overly) simple cost-benefit analysis to assess the impact of climate change uncertainty. By Lemma 1 and Proposition 1, an affine stochastic process (8) can have an affine term $a(\cdot)$, here e.g. $\bar{\sigma}^M \nu_t^M$, that takes the welfare loss to infinity without affecting $\boldsymbol{b}(\cdot)$ and, thus, optimal climate policy.

Carbon reservoirs and driving force. Finally, the uncertainty contribution is convex in its driving force. The model acknowledges that released CO_2 emissions do not simply decay (or multiply), but that carbon only moves between different sinks. Carbon in the atmosphere is more harmful than carbon in the adjacent sinks. It is this difference in harmfulness, $\Delta carb$ defined in equation (12), that drives the uncertainty contribution. It enters relative to the value contribution of atmospheric carbon $carb_1$ that is responsible for the deterministic SCC. These two terms interact carbon cycle characteristics and time preference with the long-run risk.

In summary, the optimal tax on CO_2 increases in the endogenous risk over carbon flows that derives from perturbing the climate system. This contribution is convex in all of relevant parameters. The uncertainty contribution grows faster than the deterministic contribution as climate change becomes (or is judged) more severe. Proposition 1 characterizes the contributions of the different risk channels corresponding to the extended long-run risk model and, in particular, shows that all channels increase in a product of patience and uncertainty persistence.

3.3 Temperature Uncertainty

The Gaussian model of the previous section cannot adequately represent temperature uncertainty. First, the temperature distribution governing the future global warming is right skewed. Second, temperature $T_{i,t} = \frac{1}{\xi_i} \log(\tau_{i,t})$ is a logarithmic transformation of the state $\tau_{i,t}$. The expected temperature has to approximate the deterministic temperature dynamics. Therefore, the state $\tau_{1,t}$ has to be governed by an (even more) positively skewed distribution with a suitable lower bound.[12] For this purpose, I use

[12]Otherwise, we would discuss the policy impact of changes in expected temperature dynamics, rather than the impact of uncertainty. In addition, negative realizations of generalized temperature $\tau_{i,t}$ imply nonsensical realizations of real temperature. The long-run risk model in asset pricing can give rise to some nonsensical negative realizations of the variance. This fact is well known, yet the model 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.

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). The one-step-ahead state in the autoregressive gamma process is governed by a gamma distribution whose shape parameter is modulated by a Poisson distribution. I extend the model to capture the endogeneity of climate risk and modify the state dependence.[13] I refer to Appendix C.2 for details. The canonical Gaussian model separates the long-run risk governing conditional expectations and stochastic volatility. The present model merges these two long-run uncertainties into a single process.

I model long-run temperature risk by the autoregressive gamma process y_t, which I shift by a deterministic process y_t^o to adjust expectations. The equation of motion (5) changes to

$$\boldsymbol{\tau}_{t+1} = \boldsymbol{\sigma}\boldsymbol{\tau}_t + \left(\sigma^{forc}\frac{M_{1,t}+G_t}{M_{pre}} + h(\underbrace{y_{t+1}-y_{t+1}^o}_{\equiv z_{t+1}})\right)\boldsymbol{e}_1. \tag{18}$$

The parameter h scales the uncertainty relative to the deterministic contribution, and the first unit vector $\boldsymbol{e}_1$ ensures that the feedback drives atmospheric temperature.[14] The relevant information about the process z_t is summarized in the conditional expectation and variance

$$\mathbb{E}_t\, z_{t+1} \quad = \gamma^z z_t + \epsilon(c)\left(\tfrac{M_{1,t}+G_t}{M_{pre}} - \eta_\tau\right) \tag{19}$$

$$\mathbb{V}\mathrm{ar}_t\, z_{t+1} = \mathbb{V}\mathrm{ar}_t\, y_{t+1} = c\Big[2\gamma^z y_t + \left(\tfrac{M_{1,t}+G_t}{M_{pre}} - \eta_\tau\right)\Big]. \tag{20}$$

Equation (20) governs the conditional uncertainty 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 it is autoregressive. The calibration parameter c scales the variance. Equation (19) characterizes the conditional expectations of atmospheric temperature. It is autoregressive with persistence $\gamma^z < 1$ and 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 $\epsilon(c)$

[13] In contrast to the above-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.

[14] In contrast to the earlier carbon flow model, I use the long-run risk process in period $t+1$ on the right side of equation (18). As a result, the single process governs both conditional expectations of τ_{t+1} and the one-step-ahead conditional volatility.

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.[15]

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 to the deterministic evolution.

Proposition 2 *The stochastic process* (18) *governing temperature dynamics, replacing the deterministic equation* (5)*, changes the deterministic* SCC^{det} *stated in equation* (11) *to*

$$SCC^{unc} = SCC^{det}\Big(1+\theta_\tau(F)\Big) \qquad \textit{with}$$

$$\theta_\tau(F) = \frac{h}{\sigma^{forc}}\,\frac{\epsilon(c)+\theta^*_\tau(F)}{1-\beta\gamma^z} \;\approx\; \frac{h}{\sigma^{forc}}\frac{1}{1-\beta\gamma^z}\left(\epsilon(c)+\frac{1}{2}\,\frac{1+\beta\gamma^z}{1-\beta\gamma^z}\,F\right) \tag{21}$$

and $F = \alpha\varphi^{det}_{\tau,1}\frac{ch}{1-\beta\gamma^z}$*. Here,* $\varphi^{det}_{\tau,1}$ *is the shadow value of* $\tau_{1,t}$ *under certainty.*[16] *For the exact solution*[17]

$$\theta^*_\tau(F) = \frac{-\log\left(1-F\;(1+\theta^\dagger_\tau(F))\right)}{F} - 1 \;\; \textit{with} \;\; \theta^\dagger_\tau(F) = \beta\gamma^z\frac{1+F-\sqrt{(1-F)^2-4F\frac{\beta\gamma^z}{1-\beta\gamma^z}}}{1-F+\sqrt{(1-F)^2-4F\frac{\beta\gamma^z}{1-\beta\gamma^z}}}\,. \tag{22}$$

The uncertainty premium from temperature risk is composed of the two terms $\epsilon(c)$ and θ^*_τ (equation 21).

Temperature non-linearity. The first contribution $\epsilon(c)$ results from the non-linearity of the temperature impact. This contribution arises even in the absence of intrinsic risk aversion. Such an (intrinsic) risk-aversion-independent contribution was

[15]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 proportionally 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 render explicit that a different calibration of c also changes the calibration of ϵ, achieving an approximate de-biasing of the temperature expectations.

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

[17]The solution is well-defined for $F\;(1+\theta^\dagger_\tau) < 1$ and $(1-F)^2 > 4F\frac{\beta\gamma^z}{1-\beta\gamma^z}$. A sufficient but not necessary condition for these inequalities to hold is $F < \frac{1}{3}$ and $\beta\gamma^z < \frac{(1-F)^2}{(1-F)^2+4F}$. This condition is met for all quantitative applications in this paper.

absent in the case of carbon dynamics. This absence was caused by the logarithm in the greenhouse effect (radiative forcing equation 4), which effectively offsets the relevant damage convexity. The contribution $\epsilon(c)$ 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 (see next paragraph). Both contributions are multiplied by the factor $\frac{h}{\sigma^{forc}}$ weighing the stochastic forcing contribution relative to the deterministic contribution. And both contributions are amplified by the persistence of uncertainty γ^z, and more so for a patient decision maker ($\frac{1}{1-\beta\gamma^z}$).

Risk aversion, time preference, and the low-level drivers. The r.h.s. of equation (21) presents the approximate composition of the risk aversion based contribution θ_τ^*. At the heart of the contribution lies the term $F = \alpha\varphi_{\tau,1}^{det}\frac{ch}{1-\beta\gamma^z}$, which is driven by the risk-aversion-weighted shadow value of atmospheric temperature under certainty. Again, the relevant risk-aversion is captured by $|\alpha|$, reflecting the extent to which Arrow–Pratt risk aversion exceeds the unit desire to smooth consumption over time. As in the case of carbon risk, the *temperature-based uncertainty contribution grows faster than the deterministic SCC in all contributing factors* ($\frac{SCC_t^{unc}}{SCC_t^{det}}$ increases in $\varphi_{\tau,1}^{det}$). 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*. The contribution increases in the variance-scaling parameter c and the scaling parameter of the stochastic feedback h, making the contribution convex in h (because h already appears in equation 21). The left graph in Figure 3 in the next section shows that the exact solution for θ_τ^*, equation (22), turns convex also in F, in particular, for patient decision makers.

3.4 Damage Uncertainty and Adaptation

The other major uncertainty surrounding climate change governs the temperature's impact on the economy. The damage functions of IAMs assume some degree of adaptation (often implicitly so as in DICE). Both the temperature impact and the effectiveness of adaptation measures are uncertain. The present section incorporates a stochastic component of the damage function, as well as adaptation with stochastic payoffs. For this purpose, I modify the damage function from equation (6) to the form

$$D(\tau_{1,t}) = 1 - \exp[\xi_0(1 - \tau_{1,t} + \pi_t)] \tag{23}$$

where the new term π_t introduces a stochastic damage-adaptation level. Letting $\pi_0 = 0$, it is a stochastic process that captures the deviations from our current best guess damages and evolves as

$$\pi_{t+1} = d_t\sqrt{\tau_{1,t} - \eta_d} + h_t(i_t) + \gamma^d \pi_t. \tag{24}$$

The first entry $d_t\sqrt{\tau_{1,t} - \eta_d}$ is a stochastic damage component with the coefficient $d_t \sim N(0, \sigma_d^2)$ so that $\mathbb{E}\, d_t\sqrt{\tau_{1,t} - \eta_d} = 0$; ξ_0 already captures our best guess, and the expectation of the stochastic damage component is zero. The parameter $0 < \eta_d < 1$ is once again a degree of freedom in calibrating the increase in risk with a deviation from the pre-industrial climate. The process π_t also permits active investment in adaptation; I denote such investment in period t by i_t. This investment trades off with other production, thus changing the production function from equation (1) to the form

$$\begin{aligned} &Y_t = F(\boldsymbol{A}_t, \boldsymbol{N}_t, \boldsymbol{K}_t, \boldsymbol{E}_t, i_t) \quad \text{with} \\ &F(\boldsymbol{A}_t, \boldsymbol{N}_t, \gamma\boldsymbol{K}_t, \boldsymbol{E}_t, i_t) = \gamma^\kappa F(\boldsymbol{A}_t, \boldsymbol{N}_t, \boldsymbol{K}_t, \boldsymbol{E}_t, i_t)\ \forall \gamma \in \mathbb{R}_+. \end{aligned} \tag{25}$$

The adaptation investment's damage reduction payoff $h_t(i_t)$ in equation (24) is generally stochastic. For consistency with adaptation inclusive damage estimates for ξ_0, I require that $\mathbb{E}_t\, h_t(i_t) = 0$ along the expected trajectory. Moreover, I assume that the functions $h_t(\cdot)$ and $F(\cdot)$ deliver a well-defined optimization problem for the adaptation investment. The adaption process' persistence γ^d renders damages and the adapation level a slow-moving uncertain process, introducing long-run risk as a proxy for epistemological uncertainty as well as stochastic adaptation.

Proposition 3 *Let damages and production be extended to the forms in equations* (23) *and* (25)*. The stochastic damage-adaptation process* (24) *changes the deterministic* SCC^{det} *stated in equation* (11) *to*

$$SCC^{unc} = SCC^{det}\Big(1 + \theta_d\Big) \qquad \textit{with} \qquad \theta_d = \xi_0\beta\frac{-\alpha\sigma_d^2}{2(1-\beta\gamma^d)^2}x^*\ . \tag{26}$$

Damage uncertainty further increases the optimal carbon tax, noting that α is negative for the risk averse decision maker.

The drivers. Equation (26) identifies the drivers of the carbon tax contribution from damage uncertainty. They scale the relative SCC increase θ_d proportional to the

base damage level ξ_0, the decision maker's patience β, his or her (intrinsic) risk aversion $-\alpha > 0$, and the long-run damage uncertainty, which increases in the variance σ_d^2 and the time preference weighted persistence γ^d. The uncertainty contribution also increases in the consumption rate x^*. The stochastic and costly adaptation investment itself only affects welfare and the evolution of net production (and thereby the future SCC). Somewhat similar to the earlier uncertainties, damage uncertainty turns the SCC convex in the expected damage level ξ_0 (semi-elasticity of production) because it appears in both the SCC^{det} and the damage multiplier. If expected damages are worse, then also the uncertainty contribution gains relative to the deterministic contribution.

4 Joint Uncertainty & Quantification

4.1 Joint Uncertainty & Interactions

Proposition 4 *Let damages and production be extended to the forms in equations* (23) *and* (25)*. The stochastic processes (13-15) governing carbon flows,* (18) *governing temperature dynamics, and* (24) *governing damages and adaptation change the deterministic* SCC^{det} *stated in equation* (11) *to*

$$SCC^{unc} = SCC^{det} \cdot \Big(1 + \theta_d\Big) \cdot \Big(1 + \theta_\tau\big(F \cdot (1 + \theta_d)\big)\Big) \cdot \Bigg(1 + \theta_M\Big(\theta_M^* \cdot \Big(1 + \theta_\tau\big(F \cdot (1 + \theta_d)\big)\Big)\Big)\Bigg) \tag{27}$$

where the functions $\theta_M(\cdot)$ *and* $\theta_\tau(\cdot)$ *are defined in Propositions 1 and 2, and the terms* θ_M^**,* F *and* θ_d *are defined in Propositions 1 and 2 and 3, respectively.*

The risks are mutually aggravating. The long-run risks over carbon flow, temperature, and damages interact and reinforce each other in two qualitatively different ways. The first reinforcement is asymmetric and corresponds to a causal chain of events causing damages.[18] Damage uncertainty directly amplifies the other uncertainty multipliers because uncertain damages make fluctuations in temperature more costly $\theta_\tau\big(F \cdot (1 + \theta_d)\big)$ and, similarly, temperature uncertainty makes carbon flow

[18]The multipliers affect each other in what might be considered "inverse causality". Physically, carbon flow affects temperatures and temperatures affect damages. But we are interested in the shadow values. The carbon's shadow value is amplified by the additional fluctuation of temperatures and damages before affecting consumption and production.

fluctuations more costly $\theta_M\big(...\theta_\tau(...\theta_d)\big)$. Thus, even a moderate damage uncertainty multiplier θ_d can substantially raise the overall risk premium. The second reinforcement is symmetric and merely a consequence of having damage multipliers that – as the name suggests – interact multiplicatively rather than additively. If each uncertainty multiplier adds 10% to the SCC, then the total risk premium would be 33%. This interaction effect is negligible for small uncertainty premia but can become quite large whenever the individual multipliers are large; for mere illustration think of the case where each uncertainty adds 100% to the SCC, then the total uncertainty premium would be $2^3 - 1 = 700\%$.

4.2 Summary of the Calibration

Traeger (2021*a*) discusses the calibration of the base model. The temperature system is calibrated to MAGICC6.0, a model by Meinshausen et al. (2011) used to emulate the big atmosphere-ocean general circulation models (AOGCMs) and employed abundantly in various IPCC assessment reports. The carbon cycle is taken from DICE 2013, which Traeger (2021*a*) shows to deliver SCC results very close to those when combining ACE with more recently promoted scientific models. Otherwise, the model uses a capital share $\kappa = 0.3$ borrowed from DICE and calibrates to 2020 IMF data including an investment rate of 26% and annual world output of approximately 130 trillion USD in PPP (IMF 2020). The calibration implies a rate of pure time preference of $\rho = 1.4\%$. I now summarize the calibration of the stochastic processes, which Appendix D spells out in detail.

The left graph in Figure 2 closely resembles the probability distributions and statistical information about the transient climate response (TCR) provided by the IPCC (2013). The TCR more accurately 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. Good probabilistic information on carbon flow uncertainty is scarce and the calibration is rule of thumb using Joos et al.'s (2013) study subjecting 18 different carbon cycle models to a 5000Gt carbon pulse as a point of reference. The right graph in Figure 2 translates the calibration into the resulting uncertainty over atmospheric carbon dioxide concentrations along the DICE 2013 business-as-usual scenario. Damage uncertainty σ_d might be the hardest to calibrate. Traeger (2021*a*) presents results based on two different damage calibrations. The first is based on the various DICE vintages that produce damages just above 2% of world output at a 3C warming. The second is based on a recent meta-analysis by Howard

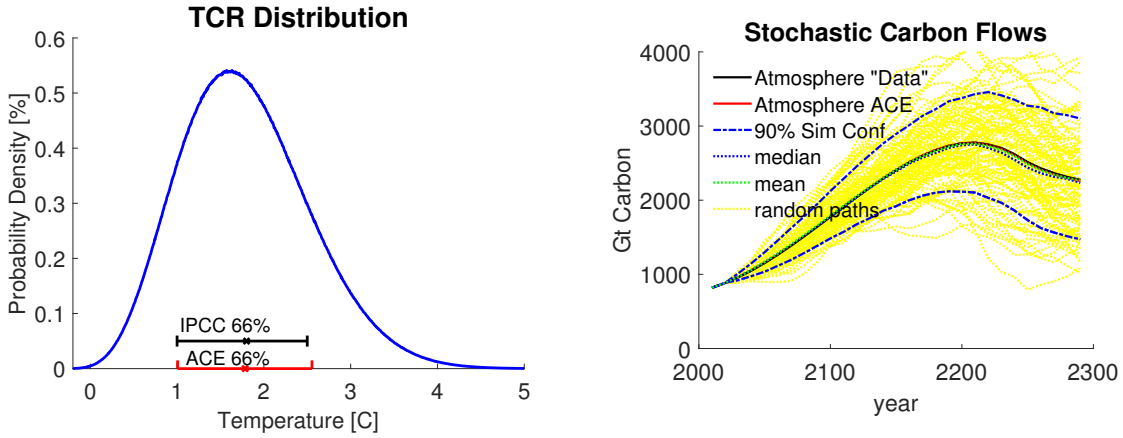


Figure 2: The left panel gives the probability distribution of the transient climate response (TCR) in ACE and compares its mean ("x") and 66% probability interval to the IPCC (2013) data. The parameters are $\gamma^z = 0.95$, $h = 0.23$, $\eta = 0.8$, $c = 0.21$, and $\epsilon = 0.05$, and the resulting TCR distribution exhibits the typical moderate positive skew. The right 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 median of 1000 uncertain trajectories are barely distinguishable.

& Sterner (2017) as well as a survey by Pindyck (2020) both suggesting damages of almost 10% of global output at a 3C warming. I call this second damage scenario HSP and interpret the Nordhaus and HSP scenario as covering opposite ends of a broad interval of reasonably likely scenarios. I pick the expected semi-elasticity of production ξ_0 as an average of the Nordhaus and the HSP scenarios and calibrate the uncertainty using a scenario that increases temperatures to 3C by 2100. I then calibrate the 2100 damages at a 3C warming so that the 10^{th} percentile of the damage distribution implies a 2.0% loss of world output and the 90^{st} percentile implies a 10.0% loss. This calibration implies a one sigma interval of damages of $[3.0\%, 9.1\%]$, which seems somewhat reasonable given the widely different damage estimates and the fact that most of the more recent findings suggest that damages are rather a bit higher than assumed in DICE.

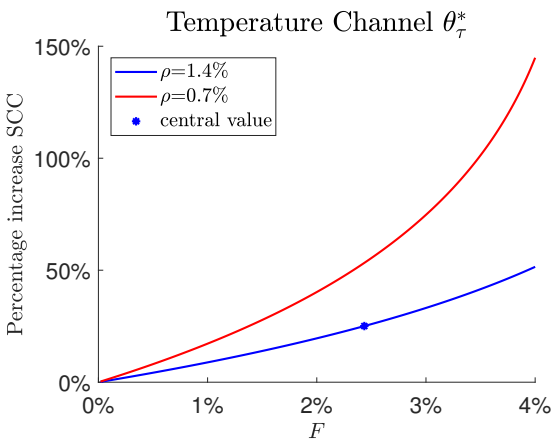
The information laid out above is not sufficient to fix the γ-parameters specifying persistence of the uncertainty nor the η-parameters specifying how steeply uncertainty increases in the deviation from preindustrial. Appendix A discusses the meaning of the autoregressive persistence parameters ($\gamma \equiv \gamma^x = \gamma^\sigma = \gamma^z = \gamma^d$) in view of epistemological uncertainty and learning shocks. This discussion motivates a choice of $\gamma \approx 0.9$, and variations of the baseline scenario will increase the shock persistence to $\gamma = 0.95$ and reduce it to $\gamma = 0.8$. Moreover, it seems fair to assume that the described

uncertainties are mostly born out of deviations from the preindustrial equilibrium, suggesting a choice of $\eta \equiv \eta_M = \eta_\tau = \eta_d$ closer to unity than to zero. I pick a somewhat conservative value of $\eta = \frac{2}{3}$ for the baseline scenario and offer a variation that pins uncertainty more sharply onto the deviation from preindustrial ($\eta = 0.9$). Note that η itself does not show in the SCC formula. However, its choice affects how steeply the endogenous uncertainty increases with deviations from the preindustrial equilibrium and, thereby, affects the calibration of the relevant uncertainty parameters entering the SCC's risk premium. Given the baseline choices and stated variations of γ and η, the calibration approaches of the previous paragraph pin down the parameters governing temperature, carbon flow, and damage uncertainty.

4.3 Quantitative Insights

Carbon-flow uncertainty. The table in Figure 3 presents the SCC's risk premium from carbon flow uncertainty and analyzes the corresponding long-run risk channels for both the baseline specification and a scenario cutting the pure rate of time preference by half. The subsequent Table 1 offers additional variations (risk aversion, persistence γ, steepness of the endogenous risk increase η). All scenarios show that the risk premium caused by carbon flow uncertainty is negligibly small. The largest premium arises under the reduced time preference and is $10\frac{USD}{tCO_2}$ or a good 2% if I incorporate the amplifications from temperature and damage uncertainty using $\theta_M(...\theta_\tau(...\theta_d))$ as done in the table of Figure 3. It is even less in a model with only carbon flow uncertainty (first column of Table 1). Thus, the first interesting result is that *carbon flow uncertainty hardly contributes to the SCC*, which stands in stark contrast to the finding that the deterministic carbon multiplier $[(\mathbf{1} - \beta\mathbf{\Phi})^{-1}]_{1,1}$ in equation (11) delivers a huge multplier of 4.3 in the baseline (Traeger 2021*a*). The reason for this small impact of carbon flow uncertainty is the highly concave transformation of atmospheric carbon dioxide (4) into warming expressed in the radiative forcing equation (4). It is essentially a saturation effect because an additional ton of CO_2 at high concentrations does not trap as much of the outgoing energy as done by the first ton of CO_2 (eventually most of the outgoing energy in the spectrum where CO_2 traps outgoing radiation is already trapped). This concavity in translating CO_2 into warming mostly neutralizes the convexity of damages.[19]

[19]The well-established logarithm in radiative forcing effectively neutralizes the second exponential function in the damage equation (6). The functions cannot be literally canceled with each other because the dynamic equations governing (stochastic) temperature stand between radiative forcing



Carbon Channels

Scenario	RP	Channel contribution in %		
	USD	$M \to x$	$M \to \sigma$	$M \to \sigma \to x$
Base	0.7	90 %	1 %	9 %
$\delta^{Mx} \uparrow$	0.8	99.8 %	0.0 %	0.2 %
$\delta^{M\sigma} \uparrow$	0.3	0.1 %	5.8 %	94.0 %
$\delta^{\sigma x} \uparrow$	0.4	0.1%	0.0%	99.9%
$\rho = 0.7\%$	9.5	87%	0.4%	12.6%
$\delta^{Mx} \uparrow$	11	99.7 %	0.0 %	0.3 %
$\delta^{M\sigma} \uparrow$	4.8	0.1 %	2.9 %	97.0 %
$\delta^{\sigma x} \uparrow$	6.5	0.1%	0.0%	99.9%

Figure 3: Left: risk averse θ_τ^*-contribution to the temperature SCC premium for base scenario ($\rho = 1.4\%$, blue) and low discounting ($\rho = 0.7\%$, red) as a function of F, defined in Proposition 2, which is proportional to the damage semi-elasticity ξ_0, risk aversion, and the climate multipliers in equation (11). Right: relative contributions of long-run risk channels to the carbon SCC risk premia (RP) in the base and low discounting scenarios. $M \to x$: conditional expectations, $M \to \sigma$: stochastic volatility, $M \to \sigma \to x$: stochastic volatility in conditional expectations. The variation $\delta^{Mx} \uparrow$ pushes the uncertainty onto the conditional expectation channel, the variation $\delta^{M\sigma} \uparrow$ pushes the uncertainty onto the stochastic volatitlity channel, and the variation $\delta^{\sigma x} \uparrow$ pushes the uncertainty into the interaction between stochastic volatility and conditional expectations.

The carbon flow uncertainty permits a disentanglement of the different long run risk channels. The table in Figure 3 breaks up the risk premium into the contributing channels. In the baseline, the uncertainty about expected flows (conditional expectations, $M \to x$) clearly dominates the other contributions. Stochastic volatility of carbon flows ($M \to \sigma$) is negligible relative to the other contributions. The scenarios $\delta^{Mx} \uparrow$, $\delta^{M\sigma} \uparrow$, and $\delta^{\sigma x} \uparrow$ change the uncertainty calibration by pushing most of the uncertainty into one of the channels, $\delta^{Mx} \uparrow$ pushes almost all the carbon flow uncertainty into conditional expectations, $\delta^{M\sigma} \uparrow$ pushes almost all the uncertainty into the stochastic volatility channel, and $\delta^{\sigma x} \uparrow$ pushes almost all the uncertainty into the interaction channel where stochastic volatility directly affects the long-run conditional expectations.[20] When pushing most of the uncertainty onto the stochastic volatil-

and damages. In the absence of (intrinsic) risk aversion, the other exponential function in the damage equation would neutralize the logarithmic utility; note the opposite roles of curvature in this second reasoning, absent the minus sign these are benefits not damages. This second exponential function normalizes damages so that they cannot exceed total output; this normalization offsets the aversion from log-utility. This reasoning covers only the first part of the intuition, the welfare loss under uncertainty. Interacting this welfare loss with the endogeneity of uncertainty closes the argument.

[20]These different scenarios maintain a comparable overall uncertainty in the sense described above by calibrating to Joos et al.'s (2013) study. See Appendix D for details.

Uncertainty / Scenario	Carbon		Temperature		RP deriving from		Damages		Joint	
	USD	RP	USD	RP	ϵ	θ_T^*	USD	RP	USD	RP
Certain	97	-	97	-	-	-	97	-	97	-
Base unc	98	0.5%	132	35%	10%	25%	103	6%	143	46%
RA high	98	0.7%	140	44%	10%	34%	105	8%	156	60%
RA low	98	0.3%	118	21%	10%	11%	100	3%	122	25%
$\gamma = 0.95$, $\eta \uparrow$	98	0.8%	170	75%	12%	62%	107	9%	201	106%
$\gamma = 0.95$, Var $\downarrow$	98	0.3%	157	61%	12%	49%	105	7%	176	80%
$\gamma = 0.8$, Var $\uparrow$	98	0.7%	113	16%	7%	9%	102	4%	120	23%
$\eta = 0.9$, Var $\uparrow$	98	0.6%	147	51%	12%	39%	105	8%	165	70%
$\rho = 0.7\%$ - cert	181	-	181	-	-	-	181	-	181	-
$\rho = 0.7\%$ - unc	185	1.7%	416	129 %	13%	116%	204	12%	602	232%

Table 1: Optimal tax in USD/tCO$_2$ and risk premium (RP) for base calibration and several variations. The columns Carbon, Temperature, and Damages report the results if only that particular uncertainty prevails. RA low and high vary intrinsic risk aversion $|\alpha|$ from 1 to 0.5 and 1.25. The subsequent three scenarios vary the long-run risk persistence of 0.9 in the baseline and either recalibrate η or the corresponding shock's variance accordingly. The scenario $\eta = 0.9$ makes the risk more sensitive to deviations from the preindustrial equilibrium (recalibrating the shock's variance). The final scenario cuts pure time preference from 1.4% to half.

ity channel, the overall premium falls substantially and the dominating channel is (still) not the stochastic volatility channel itself, but its interaction with conditional expectations. Consistently, when pushing most of the uncertainty onto the interaction channel, it remains the dominant channel. Pushing the uncertainty into the conditional expectations channel delivers the largest premium. None of these results change when I cut the time preference into half, the scenario that causes the biggest quantitative changes of the premia in Table 1. It is safe to conclude that, in contrast to many asset pricing contexts, *it is the long-run expectations that matter much more than uncertainty about future volatility.*

Temperature uncertainty. In the baseline, temperature uncertainty adds a risk premium of 35% (Table 1). 25% of this premium rely on the interaction with intrinsic risk aversion (the term θ_T^* in equation 21). The remaining 10% would also prevail in a model with simple log-utility (no Epstein-Zin preferences) as a result of the damage convexity's interaction with the endogeneity of the uncertainty. Risk aversion at the high end of the estimates raises this premium by another 10 percentage points and estimates at the low end lower it by 15 percentage points. Rows five and six of Table 1 raise the shock persistence of the long-run risk from $\gamma = 0.9$ to $\gamma = 0.95$. Raising

the persistence reflects more persistent shocks or slower learning and approximately doubles the risk premium. Such an increase of persistence requires a recalibration of temperature uncertainty explained in Appendix D. The first row (more than doubling the premium) recalibrates η from $\frac{2}{3}$ to 0.8, pinning uncertainty more steeply on deviations from the preindustrial equilibrium (reducing the "base uncertainty" prevailing already in equilibrium). The second row instead recalibrates the shock's variance by reducing c from 0.21 to 0.18. Then, the impact of the increase in persistence γ is partially compensated by a reduction in variance and the risk premium slightly less than doubles, still confirming the premium's high sensitivity to the persistence of the long-run risk. Reducing the persistence to $\gamma = 0.8$ corresponds to less persistent long-run risk and mimics faster learning. It is not possible to recalibrate the TCR with a reasonable choice of η; therefore I only present the scenario increasing the variance to compensate for the reduced shock persistence.[21] The temperature risk premium drops to 16%, less than half of its original value.

The scenario "$\eta = 0.9$, Var $\uparrow$" in row 8 pins uncertainty most sharply onto deviations from the preindustrial equilibrium. Leaving the persistence unchanged, it recalibrates the TCR by increasing the long-run risk's shock variance c (from 0.21 to 0.29).[22] Temperature risk alone contributes a premium of 50%, an increase that is substantial, but far less than the increase resulting from an increase in persistence. In all scenarios, the temperature risk premium deriving from the interaction with intrinsic risk aversion is larger and more sensitive to the parameter variations than the part prevailing also in the absence of intrinsic aversion to risk (no Epstein-Zin preferences). The graph on the left of Figure 3 shows the increase of this risk-sensitive contribution in F, defined in Proposition 2, which is proportional to the damage semi-elasticity ξ_0, risk aversion, and the climate multipliers of equation (11). The blue line representing the base calibration characterizes an almost linear increase, underlying the approximation in equation (21), which helps to interpret the formula. The red line shows that the premium turns much more convex in these contributions under a reduction of pure time preference, where the full formula is essential for a quantitative assessment.

[21] Merely reducing persistence to $\gamma = 0.8$ without increasing the shock variance results in a quite similar SCC, implying a risk premium of $7\% + 8\% = 15\%$ for the temperature premium and 19% for the joint risk premium.

[22] The substantial increase in the variance also requires a recalibration of the mean adjustment ϵ, which increases from 0.05 in the other scenarios to 0.06. As a result, the risk premium unrelated to intrinsic risk aversion increases to 12%, which is similar to the impact observed under increasing the long-run risk's persistence.

Damage uncertainty. The uncertain evolution of damages and adaptation create a risk premium that is notable but smaller than the one resulting from temperature uncertainty. Table 1 shows a direct risk premium of 6% in the baseline ±3 percentage points for the variations (not touching time preference)6. Qualitatively, the premium responds to the different parameter variations in the same way as the temperature-based premium. It is slightly less sensitive to the variations in both persistence and pure time preference, which confirms a higher sensitivity of the premium to higher moments of the uncertainty distribution: Given good knowledge of the distribution, ACE's temperature uncertainty replicates the skewness in the data; in the absence of similar knowledge, ACE's damage uncertainty evolves according to a normal distribution. In addition to the direct impact specified in Proposition 3, uncertainty over the resulting damages also amplifies the contributions of carbon and temperature uncertainty (see Proposition 4). As a result, the total impact of damage uncertainty on the *joint* uncertainty premium is larger and amounts to 10% in the baseline and even 29% in the scenario "$\gamma = 0.95$, $\eta \uparrow$" with slow learning and pinning uncertainty more steeply onto deviations from the preindustrial equilibrium.

Joint uncertainty. In the baseline, the joint uncertainty premium on the SCC is $45\frac{USD}{tCO_2}$ or 46%. The variations that leave time preference untouched change the total risk premium between $20\frac{USD}{tCO_2}$ and over $100\frac{USD}{tCO_2}$. In the baseline, the interaction between the different uncertainties plays a somewhat minor role, making up $5\frac{USD}{tCO_2}$ of the total risk premium. The interaction effect becomes more pronounced under an increase in the long-run risks' persistence or for a more patient decision maker. In the high persistence scenario "$\gamma = 0.95$, $\eta \uparrow$," the joint uncertainty premium exceeds the sum of the individual premia by over 20 percentage points. Proposition 4 identifies two qualitatively different interaction channels causing this mutual aggravation of risks. First, damage uncertainty directly increases all uncertainty multipliers because it amplifies the impact of carbon flow and temperature variations, and temperature uncertainty similarly increases the premium on carbon flow uncertainty. In the high persistence scenario, this direct interaction doubles the carbon premium from 0.8% to 1.6% and increase the temperature premium from 75% to 85% causing a good 10 percentage point increase of the joint risk premium relative to the sum of the individual premia. Second, the three uncertainty multipliers in equation (27) interact – as the name suggests – multiplicatively rather than additively. This second interaction effect adds the other 10 percentage points to the joint uncertainty premium. Also these uncertainty interaction effects are highly sensitive to the uncertainty's persistence and

to time preference. In the baseline, the direct effect and the multiplier effect each increase the premium by 2 percentage points; in the high persistence scenario just discussed, each adds 10 percentage points; and in the low discounting scenario, the direct interaction effect adds 55 percentage points and the multiplier effect another 34 percentage points.[23]

Time preference. The propositions in Section 3 establish a high sensitivity of the SCC's risk premium to (persistence weighted) time preference. The magnitude of the impact demonstrated by Table 1 might still surprise. The joint risk premium (more than) quadruples when cutting the rate of pure time preference by half. In contrast, the deterministic SCC (less than) doubles. This finding forcefully illustrates the theoretic findings that (i) the uncertainty premium increases more strongly than the deterministic SCC in the factors making climate change judged more severe and (ii) that the uncertainty premium is even more sensitive to time preference than the already sensitive deterministic SCC. Appendix A demonstrates that the sensitivity to (persistence weighted) time preferences increases with the moment of the distribution. Given that the TCR distribution governing temperatures exhibits both skew and some kurtosis, the temperature premium is the most sensitive to the reduction in time preference and its direct premium quadruples. The direct damage risk premium only doubles, but its particularly strong interaction with the other uncertainties make it almost as potent a contributor to the overall quadrupling of the premium.

Traeger (2021*a*) shows that the value of abatement relative to other consumption and, thus, the SCC goes to infinity as pure time preference goes to zero; the present study shows that this sensitivity even increases under uncertainty. Thus, the sensitivity to a given reduction in time preference should be much higher for an already patient planner. Indeed, a reduction of pure time preference from the calibrated baseline value of 1.4% to 1% increases the joint risk premium from 46% to 87% ($SCC^{det} \approx 130\frac{USD}{tCO_2}$ and $RP = 115\frac{USD}{tCO_2}$). This 0.4% reduction of pure time preference has a much more moderate impact than the subsequent 0.3% reduction boosting the risk premium all the way to 230% and the SCC to over $600\frac{USD}{tCO_2}$. Is such a low rate of pure time preference reasonable? The median of Drupp et al.'s (2018) recent expert survey on discounting suggests an even lower rate of time preference of 0.5% (and

[23]Figure 3 reports the carbon risk premium under joint uncertainty whereas Table 1 shows the premia under carbon uncertainty only. In the low discounting scenario, the premium under carbon flow uncertainty is boosted from $4\frac{USD}{tCO_2}$ to $10\frac{USD}{tCO_2}$ in a model with joint uncertainty or, in percent, from 2% to 5%. In the same low discounting scenario, the temperature uncertainty premium is boosted from 129% to 181%.

logarithmic utility). Traeger (2021*a*) discusses the calibration of time preference in IAMs in much more detail and surveys arguments that imply even lower rates based on ethical grounds, overlapping generations, and the long-run risk model in asset pricing. Yet, once we reach optimal carbon prices exceeding 400 or even $600\frac{USD}{tCO_2}$ direct carbon capture will become an attractive technology and interact with the climate risks in ways not easily incorporated into the present model. I would phrase the practical implication of the high SCC values as follows. The most stringent policies in a few countries and sectors currently correspond to prices of 150-200$\frac{USD}{tCO_2}$. In economic models, these values have generally been motivated based on low rates of pure time preference. The present results tell us that, under uncertainty, we do not need Stern (2007)-type rates of 0.1% to rationalize such stringent mitigation efforts, but much more moderate rates of time preference already call for substantially stronger efforts.

4.4 Related Findings

Jensen & Traeger (2013) find a climate-sensitivity risk premium of just over 20% for a pure rate of time preference of 1.5%. The authors explicitly model Bayesian learning, but do not use Epstein-Zin preferences. Their Arrow-Pratt risk aversion is 2, so that I would indeed expect results closer to my low risk aversion scenario.[24] Kelly & Tan (2015) find a risk premium of 24% using a rate of time preference of 5% in combination with fat tails, using a joint coefficient of 1.5 for relative risk aversion and aversion to intertemporal substitution. 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 contrast, Jensen & Traeger (2013) find that anticipated learning, in their model without fat tails, has no impact on the present SCC. Both models use a simplified climate model including a carbon decay approximation to the carbon cycle, in order to save numerically expensive state variables. Rudik & Lemoine (2017) find a risk premium of 10% using a numerically less expensive Smolyak grid in an implementation without state-space reduction, also using a time preference of 1.5%. In their setting, learning drops this premium to 1%. See Appendix A for a simple but instructive analytic discussion of the effects of

[24]The authors also look at a scenario with smooth ambiguity aversion. While related to Epstein-Zin preferences, the higher coefficient of (ambiguity) aversion only affects part of the distribution. As a result, ambiguity aversion implies a much more moderate increase of the premium than my increase of relative risk aversion. A revision of the paper will include a scenario with Epstein-Zin preferences, still yielding a slightly lower premium than the present paper, and add an analytic discussion of the drivers of risk premia in slightly less complex IAM.

learning.

Lemoine & Traeger (2016) find that a combination of climate and economic tipping points can double the SCC, but point out that merely doubling the expected damage coefficient as done by Nordhaus to adjust the DICE model for missing tipping points has almost the same effect on the present SCC. Cai & Lontzek (2019) find the same doubling of the SCC under tipping points in a slightly different setting where damage can move up irreversibly in different tipping regimes. Tipping points operate mostly through increasing expected damages and are, therefore, of a different nature than the present uncertainty analysis. Cai & Lontzek (2019) also introduce growth uncertainty, which reduces the SCC (Jensen & Traeger 2014). What further increases the SCC in Cai & Lontzek (2019) under Epstein-Zin preferences is a (deterministic) calibration effect resulting from the lower risk-free rate, first quantified in DICE by Crost & Traeger (2014), analyzing stochastic damages. The SCC impact of the magnitude of damage stochasticity in Crost & Traeger (2014) is negligibly small because the authors only model stochastic shocks without long-run risk persistence. Recently, Rudik (2020) adds persistence (learning) to this stochastic damage analysis and finds a risk premium of 3-4% from damage uncertainty;[25] again using the full DICE specification under a Smolyak grid approximation. Rudik (2020) also finds a slightly larger premium for a robust control specification. All of the above papers are based on Nordhaus' DICE model. While somewhat comparable in relative terms, ACE's base SCC and risk premia are substantially larger in absolute terms for three reasons. First, I cut the excessive delays of the DICE model (Van der Ploeg et al. 2020), which holds partly as well for Jensen & Traeger (2014) and Kelly & Tan (2015). Second, I use updated purchasing power parity weights in calculating world output as compared to the DICE model underlying the other studies, which increases the SCC by over 40%. Third, and maybe most importantly, I use higher expected damages than the DICE model as discussed in Section 4.2.

The recent study by van den Bremer & van der Ploeg (2018) finds an approximate analytic solution to a more stylized integrated assessment model and, like ACE, addresses several uncertainties jointly. It confirms ACE's results that carbon uncertainty adds a negligible premium and finds risk premia of about 40% for both damage

[25]The 3-4% is a guesstimate based on the learning premium graph in Figure 7(a) for 2020 relative to the base value taken from Figure 5(a). The most important difference between the graph labeled "uncertainty" and "learning" is that those labeled uncertainty have no shock persistence as in Crost & Traeger (2014) (and consequently no premium) whereas the learning model introduces shock persistence.

and temperature uncertainty in their corresponding calibration.[26] In a related continuous time AK model with emissions and temperature Hambel et al. (2021) find much more moderate risk premia for smooth temperature risk (diffusion process) but huge premia for jump risk (which also raises the expected damage). For smooth risk, which is more comparable to the present study, the temperature risk premium is approximately 4% under DICE damages but increases to 18% when combining more extreme "Weitzman-type" risk with growth rather than level impacts. The authors use log-utility in combination with a (disentangled) Arrow-Pratt risk aversion similar to my "high RA" scenario and a pure time preference of 1.5%. In a stylized model without decision maker, Lemoine (2021) simulates the SCC under uncertainty finding risk premia of about 3% for temperature uncertainty and 50% for damage uncertainty under the same time preference.

5 Conclusions

The paper discusses the impact of long-run climate and damage risks on the optimal carbon policy. Quantitatively, the temperature feedback uncertainty delivers the largest risk premium to the SCC, whereas uncertain carbon uptake by the oceans and biosphere results in a negligible risk premium. Damage uncertainty contributes a notable risk premium by itself, and it increases the contributions of all other risks. The SCC's risk premium is almost 50% in the baseline. This value is particularly sensitive to the persistence of the long-run risk, doubling under an increase in the risk's persistence and falling to half for a reduction. The premium quadruples when cutting pure time preference from its base value of 1.4% into half, a value still exceeding the median of a recent expert elicitation. In the base calibration, the mutual reinforcement of the different risks is moderate; merely adding risk premia across

[26] I am referring to their "ethics-based calibration", which uses a 1.5% rate of pure time preference close to the present analysis and most of the papers cited above. Their alternative specification uses a rate of 5.8% cutting the temperature risk premium to 9% and the damage risk premium to 18%. The study also uses Epstein-Zin preferences and their dominating risk premium is that of growth uncertainty causing a (positive) 160% risk premium. I purposefully avoid growth uncertainty in the present study and focus on climate and damage uncertainties because (i) the growth uncertainty premium most crucially depends on the log-assumption and (ii) deviating from it the way van den Bremer & van der Ploeg (2018) calibrate to macroeconomic observation (elasticity of intertemporal substitution $EIS < 1$) causes a huge positive risk premium whereas calibrating the model relative to the asset pricing literature ($IES > 1$) causes a huge negative risk premium (Jensen & Traeger 2014). See Traeger (2019) for an explanation of this conflict when calibrating Epstein-Zin preferences; a good solution to this conflict has yet to be developed.

different models would result in a good approximation. In more severe scenarios, the mutual aggravation of the risks becomes a substantial part of the overall premium.

The optimal carbon tax under certainty is linear in a variety of factors including the damage semi-elasticity and a set of climate multipliers. Uncertainty turns the SCC convex in all of these contributions including the expected damage level. As a result, the risk premium grows much faster than the deterministic SCC in the relevant contributions; the more serious the climate change problem, the larger the relative importance of uncertainty.

It is well-established that climatic change causes a positively skewed probability distribution over future temperatures. This skew (and higher order moments) make temperature uncertainty particularly important for the risk premium. The paper develops a novel implementation of the autoregressive gamma process to adequately capture this skew and the transient climate response in an analytic integrated assessment model quantifying its impact on the optimal carbon tax. In a simplified model, I explain how the sensitivity to time preference, (intrinsic) risk aversion, and shock persistence increases in the moments of the uncertainty distribution. In particular, the stochastic model is more sensitive to time preference than the deterministic model, and skewed uncertainty (or a distribution with kurtosis) has a more time preference and risk aversion sensitive impact than normally distributed uncertainty. Intrinsic risk aversion is crucial for the magnitude of the risk premia. It captures by how much (disentangled) Arrow-Pratt risk aversion exceeds the unit aversion to intertemporal change. Mere log-utility would result in risk premia of only 10%.

The low impact of carbon flow uncertainty on the optimal carbon tax results from "saturation effects" in the atmosphere. The first unit of atmospheric CO_2 causes substantially more warming than the current unit. Thus, despite of the convexity of climate damages, a higher realization of atmospheric carbon does not necessarily cause more damage than a low realization alleviates. This finding emphasizes the importance of distinguishing temperature dynamics from carbon dynamics, which is usually neglected in the analytic literature. Moreover, the long-run risk model for carbon flows demonstrates that, in contrast to a generic asset pricing application, the climate planner should be more concerned about future expected levels than about possible stochastic volatility. This finding is convenient because it is currently still – or "even more" – unclear whether or how climate change will cause changes in volatility.

Finally, the paper explains that risk aversion and uncertainty by themselves reduce

expected welfare but do not necessarily affect the optimal policy level. The present framework focuses on the interaction of risk aversion and non-linearities in climate dynamics with the endogeneity of climate risk. We are uncertain about the future precisely because we are perturbing the climate system. Thus, by reducing emissions, we can reduce the uncertainty and increase our welfare. Therefore the optimal carbon tax always increases under uncertainty in the present setting.

It should be needless to say that, despite major advances in developing and calibrating the underlying model and stochastic processes, the model is far from perfect. There are many non-linearities and potential extensions that the model cannot handle and that eventually will have to be analyzed numerically, likely soon given the current advances of numeric methods and computing power. In particular, more thorough assessments of the role of learning would be valuable. At present, the model pushes the envelope of quantitative assessments of the SCC under various interacting uncertainties. As – or maybe most – importantly, the model permits insights, some easy and some more intricate, into the workings of uncertainty in a complex integrated assessment model. I hope that the model can serve as a benchmark for a rich set of variations analyzing how and why different variations affect optimal climate policy or, more generally, economic policy in macro models of ACE's complexity.

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* .

Bansal, R., Kiku, D. & Ochoa, M. (2019), 'Climate change risk', *Working Paper* .

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.

Barnett, M., Brock, W. & Hansen, L. P. (2020), 'Pricing uncertainty induced by climate change', *The Review of Financial Studies* **33**(3), 1024–1066.

Berger, L. & Marinacci, M. (2020), 'Model uncertainty in climate change economics: A review and proposed framework for future research', *Environmental and Resource Economics* pp. 1–27.

Brock, W. A. & Hansen, L. P. (2018), 'Wrestling with uncertainty in climate economic models', *University of Chicago, Becker Friedman Institute for Economics Working Paper* (2019-71).

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

Brumm, J. & Scheidegger, S. (2017), 'Using adaptive sparse grids to solve high-dimensional dynamic models', *Econometrica* **85**(5), 1575–1612.

Cai, Y., Judd, K. L., Lenton, T. M., Lontzek, T. S. & Narita, D. (2015), 'Environmental tipping points significantly affect the cost- benefit assessment of climate policies', *Proceedings of the National Academy of Sciences* **112**(15), 4606–4611.

Cai, Y., Lenton, T. M. & Lontzek, T. S. (2016), 'Risk of multiple interacting tipping points should encourage rapid co2 emission reduction', *Nature Climate Change* **6**(5), 520–525.

Cai, Y. & Lontzek, T. S. (2019), 'The social cost of carbon with economic and climate risks', *Journal of Political Economy* **127**(6), 2684–2734.

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.

Crost, B. & Traeger, C. P. (2014), 'Optimal co2 mitigation under damage risk valuation', *Nature Climate Change* **4**, 631–636.

Daniel, K. D., Litterman, R. B. & Wagner, G. (2019), 'Declining co2 price paths', *Proceedings of the National Academy of Sciences* **116**(42), 20886–20891.

Dietz, S., Gollier, C. & Kessler, L. (2018), 'The climate beta', *Journal of Environmental Economics and Management* **87**, 258–278.

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.

Golub, A., Narita, D. & Schmidt, M. G. (2014), 'Uncertainty in integrated assessment models of climate change: Alternative analytical approaches', *Environmental Modeling & Assessment* **19**(2), 99–109.

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. (2021), 'Optimal carbon abatement in a stochastic equilibrium model with climate change', *European Economic Review* **132**, 103642.

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

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. & Millner, A. (2014), 'Reflections: Uncertainty and decision making in climate change economics', *Review of Environmental Economics and Policy* **8**(1), 120–137.

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.

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

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

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. & 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.

Karydas, C. & Xepapadeas, A. (2019), 'Pricing climate change risks: Capm with rare disasters and stochastic probabilities', *CER-ETH Working Paper Series Working Paper* **19**, 311.

Keller, K., Bolker, B. M. & Bradford, D. F. (2004), 'Uncertain climate thresholds and optimal economic growth', *Journal of Environmental Economics and Management* **48**, 723–741.

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.

Kotlikoff, L., Kubler, F., Polbin, A., Sachs, J. & Scheidegger, S. (2021), 'Making carbon taxation a generational win win', *International Economic Review* **62**(1), 3–46.

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.

Leach, A. J. (2007), 'The climate change learning curve', *Journal of Economic Dynamics and Control* **31**, 1728–1752.

Lemoine, D. (2021), 'The climate risk premium: how uncertainty affects the social cost of carbon', *Journal of the Association of Environmental and Resource Economists* **8**(1), 27–57.

Lemoine, D. & Traeger, C. (2016), 'Economics of tipping the climate dominoes', *Nature Climate Change* **6**, 514–519.

Lemoine, D. & Traeger, C. P. (2014), 'Watch your step: Optimal policy in a tipping climate', *American Economic Journal: Economic Policy* **6**(1), 1–31.

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

Lontzek, T. S., Cai, Y., Judd, K. L. & Lenton, T. M. (2015), 'Stochastic integrated assessment of climate tipping points indicates the need for strict climate policy', *Nature Climate Change* **5**(5), 441–444.

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

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

Pizer, W. A. (1999), 'The optimal choice of climate change policy in the presence of uncertainty', *Resource and Energy Economics* **21**(3-4), 255–287.

Rudik, I. (2020), 'Optimal climate policy when damages are unknown', *American Economic Journal: Economic Policy* **12**(2), 340–73.

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.

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

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

Traeger, C. P. (2015), 'Closed-form integrated assessment and uncertainty'.

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

Traeger, C. P. (2021*a*), ACE - analytic climate economy, Technical report, Mimeo.

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

Valentini, E. & Vitale, P. (2019), 'Optimal climate policy for a pessimistic social planner', *Environmental and Resource Economics* **72**(2), 411–443.

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

Van der Ploeg, F. & de Zeeuw, A. (2018), 'Climate tipping and economic growth: precautionary capital and the price of carbon', *Journal of the European Economic Association* **16**(5), 1577–1617.

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

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

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

von zur Muehlen, P. (2018), 'Confronting climate skepticism: Ramsey carbon taxation and pricing under ambiguity', *SSRN 3291641* .

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

Weitzman, M. L. (2009), 'On modeling and interpreting the economics of catastrophic climate change', *The Review of Economics and Statistics* **91**(1), 1–19.

Appendix

A Welfare, Uncertainty, Learning, and Sensitivity to the Distributional Moments

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 also fleshes 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 and assumes that the damage function is known

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

$$\boldsymbol{\tau}_{t+1} = \boldsymbol{\sigma}\boldsymbol{\tau}_t + \sigma^{forc}\frac{M_{1,t} + G_t}{M_{pre}}\boldsymbol{e}_1 + \boldsymbol{\epsilon}_t^\tau + \boldsymbol{\nu}_t^\tau \ . \tag{A.2}$$

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

$$\mu_{t+1}^j = \gamma^j \mu_t^j + \chi_t^j \ , \tag{A.3}$$

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

of the mean[27]

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

Writing the updating equation in the form of equation (A.4) 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 (A.3) and (A.4) is that the conditional expectation of the Bayesian model exhibits full persistence and a prescribed evolution of the shock variance that falls over time, $\sigma^2_{\epsilon,t+1} = \frac{\sigma^2_{\nu,t}\sigma^2_{\epsilon,t}}{\sigma^2_{\nu,t}+\sigma^2_{\epsilon,t}}$.[28]

Proposition 5 *Let uncertainty in equations (A.1-A.2) 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[29] distributed iid shocks χ_t implies the welfare loss

$$\Delta W^{AR}_{general} = \tfrac{\beta}{1-\beta}\, \tfrac{1}{\alpha}\, G_\chi\big(\tfrac{\beta}{1-\gamma^j\beta}\alpha\varphi_j\big) = \tfrac{\beta}{1-\beta}\tfrac{1}{\alpha}\textstyle\sum_{l=1}^{\infty} \kappa_l \frac{1}{l!}\left(\frac{\beta}{1-\gamma^j\beta}\alpha\varphi_j\right)^l, \tag{A.5}$$

[27] 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 (A.4). Note that the observational variable z is defined in equations (A.1-A.2). For example, in the case of uncertain atmospheric carbon content, the observation z is $M_{t+1} - \boldsymbol{\Phi}_{1,\cdot}\boldsymbol{M}_t - (\sum_{i=1}^{I^d} E_{i,t} + E^{exo}_t)$.

[28] 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.

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

where κ_l are the cumulants of the iid shock χ_t.

In both models (i) and (ii), the welfare loss is proportional to (intrinsic) risk attitude α and the square of the state's shadow value under certainty, e.g., $\varphi_j = \varphi_{\tau,1}$ if uncertainty governs atmospheric temperature.[30] 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. 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[31] 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 is helpful to gauge the $\gamma-$parameters specifying uncertainty persistence in the models of Section 3. Merely comparing the two squared weighting factors would, however, be misleading. Initially, the prior uncertainty is much larger than the shock's uncertainty. For a back of the envelope calibration of γ, I take the following values. I assume that in the first few decades, the predictive distribution of the Bayesian learning model has a variance $\sigma^2_{\nu,t} + \sigma^2_{\epsilon,t}$ that is about four times as large as

[30] For uncertainty over carbon flows, keep in mind that 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 5 with the shadow value $\varphi_j = \varphi_{M,1} - \varphi_{M,2}$ (see the proof of the proposition for details).

[31] More precisely, it swaps the factor times $\beta^2(1-\gamma\beta)^{-2}$ against $\left(\frac{\Omega_t}{1-\beta}\right)^2$, in a comparison where β is close to unity dropping this factor is of minor relevance.

the variance of the shocks in the AR(1) process σ^2.[32] I also assume that the (decadal) measurement error, $\sigma_{\nu,t}$, is about 10% of the prior's standard deviation $\sigma_{\epsilon,t}$, so that $\frac{\sigma_{\nu,t}}{\sigma_{\epsilon,t}} = 10\%$. Using the baseline's discount factor, I find

$$\left(\frac{\beta}{1-\gamma\beta}\right)^2 \sigma^2 \stackrel{!}{=} \left(\frac{\Omega_t}{1-\beta}\right)^2 (\sigma_{\nu,t}^2 + \sigma_{\epsilon,t}^2) \qquad \Rightarrow \qquad \gamma \approx 0.89.$$

This equation and value allow me to root the autoregression factor in the underlying epistemologic nature of the uncertainty. That said, it is but a back of the envelope reasoning how to calibrate a shock-based model to a learning-based model and by no means is meant to replace models that thoroughly analyze and incorporate learning into the contexts of the present paper's SCC analysis. In the truly long-run, the shock model will always exaggerate the uncertainty and the corresponding gamma would be smaller.[33]

The *general autoregressive* shock model in case (iii) of Proposition 5 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 (A.5) suggests that the fatter the tail the more relevant the calibration of time preference.

B 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 equiv-

[32] The outcome of the calibration below results in $\gamma \approx 0.89$, which in turn implies that a simple AR(1) process' unconditional variance is $(1-\gamma^2)^{-1} \approx 4.8$ times the shock's variance. If we associate the first period's prior with the unconditional variance, the initial prior would be almost five times the shock's variance but falling over time.

[33] As an illustration, say at some point when most of the uncertainty has been learned, it holds that $\sigma_{\nu,t} \approx \sigma_{\epsilon,t} \approx \sigma$, then $\frac{\sigma_{\nu,t}}{\sigma_{\epsilon,t}} = 1$ and $\frac{\sigma^2}{\sigma_{\nu,t}+\sigma_{\epsilon,t}} = \frac{1}{2}$ and the equation gives $\gamma = 0.82$ close to the low persistence scenario.

alence 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

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

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

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

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 (B.1) to equations (B.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 (A). 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.[34] As importantly, the risk aversion measure RRA$= 1-\alpha$ is time preference invariant in that the lottery choice depicted in Figure 4, 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.

Figure 4 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

[34]The reader can convince herself of this statement by either substituting V_{t+1} recursively into equations (B.1) or (9), or by looking at Figure 4 for the special case of the lottery that I will introduce below. See Traeger (2019) for more on the normalization of risk aversion measures in the Epstein-Zin-Weil setting.

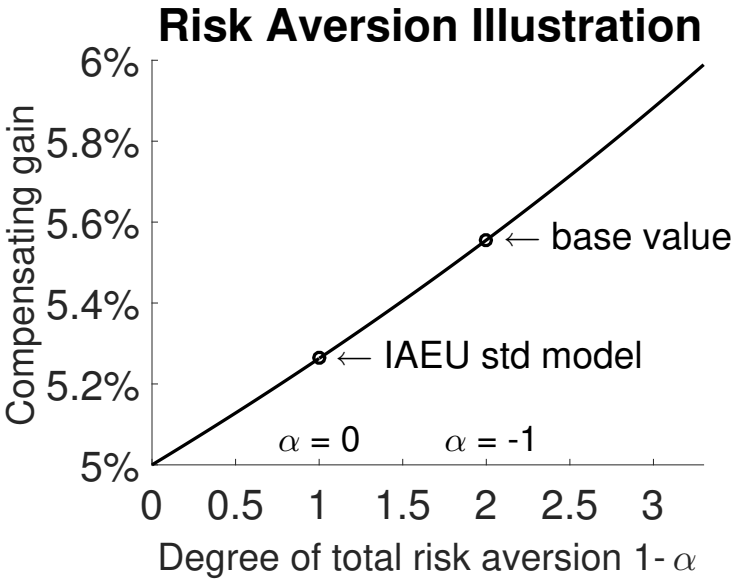


Figure 4: The graphs shows the compensating risk premium that an agent requires with probability one half to compensate a 5% loss occurring as well with probability one half. The intertemporally additive expected utility (IAEU) model corresponds to $\alpha = 0$ (no risk aversion beyond the desire to 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).

and gains are direct consumption changes.[35] The asset pricing literature usually finds $\text{RRA}^* = 1 - \alpha^* \in [6, 10]$. 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 of these bounds as a sensitivity range.

[35]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 $\bar{c}$.

Online Appendix

C Proofs

C.1 Proof of Lemma 1

1) Under certainty, equation (8) requires the existence of functions $a(\cdot)$ and $\boldsymbol{b}(\cdot)$ such that

$$\boldsymbol{z}^\top \boldsymbol{X}_{t+1} = a(\boldsymbol{z},\boldsymbol{A}_t,\boldsymbol{N}_t,\boldsymbol{\mathcal{K}}_t,\boldsymbol{E}_t, x_t) + \boldsymbol{b}(\boldsymbol{z})\boldsymbol{X}_t,$$

which is merely a requirement that the equations of motions are linear in the states and separable between the states and the variables $\boldsymbol{A}_t, \boldsymbol{N}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{E}_t$ and x_t. Moreover, the deterministic model has no additional informational vectors. Thus, the condition is satisfied for X_t if each of the equations of motion for $\boldsymbol{M}_{t+1}, \boldsymbol{\tau}_{t+1}$, and k_{t+1} satisfy such linearity and separability. The equations of motion of the carbon cycle (3) are linear in the states and satisfy the separation condition – the control $\boldsymbol{E}_t$ is additively separated from the states. The equation of motion (5) governing temperature dynamics was already introduced in its transformed form, also satisfying the separation requirement (see DACE for underlying calibrated temperature system and its transformation). The capital stock's equation of motion (7) transforms into log-capital (see DACE) as

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

also satisfying the requirement.

2) First, I express a period's welfare $u(C_t) = \log(C_t)$ in terms of the transformed variables as (see DACE)

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

yielding the Bellman equation in terms of the transformed state variables

$$\begin{aligned} V(k_t, \boldsymbol{\tau}_t, \boldsymbol{M}_t, \boldsymbol{R}_t, \boldsymbol{I}_t, t) = \max_{x_t, \boldsymbol{N_t}, \boldsymbol{\mathcal{K}}_t, \boldsymbol{E_t}} \log x_t + \kappa k_t + \log F(\boldsymbol{A}_t, \boldsymbol{N}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{E}_t) \\ -\xi_0\tau_{1,t} + \xi_0 + \frac{\beta}{\alpha}\log\Big(\mathbb{E}_t \exp\big[\alpha\ V(k_{t+1}, \boldsymbol{\tau}_{t+1}, \boldsymbol{M}_{t+1}, \boldsymbol{R}_{t+1}, \boldsymbol{I}_{t+1}, t)\big]\Big). \end{aligned} \tag{C.1}$$

which is subject to the underlying constraints

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

the initial states, and the equations of motion (2) for the resource and (8) for the combined state vector $\boldsymbol{X}_t$. Using the affine trial solution

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

and adding the constraints binding in every period with a Lagrange multiplier delivers the Bellman equation

$$\begin{aligned}
&\boldsymbol{\varphi}_X^\top \boldsymbol{X}_t + \boldsymbol{\varphi}_{R,t}^\top \boldsymbol{R}_t + \varphi_t = \\
&\quad \max_{x_t, \boldsymbol{N_t}, \boldsymbol{\mathcal{K}}_t, \boldsymbol{E_t}} \log x_t + \kappa k_t + \log F(\boldsymbol{A}_t, \boldsymbol{N}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{E}_t) - \xi_0 \tau_{1,t} + \xi_0 \\
&\qquad\quad + \lambda_t^N \big(1 - \textstyle\sum_{i=1}^{I_N} N_{i,t}\big) + \lambda_t^{\mathcal{K}} \big(1 - \textstyle\sum_{i=1}^{I_K} \mathcal{K}_{i,t}\big) \\
&\qquad\quad + \frac{\beta}{\alpha} \log\Big(\mathbb{E}_t \exp\Big[\alpha\Big(\boldsymbol{\varphi}_X^\top \boldsymbol{X}_t\Big)\Big]\Big) + \beta \boldsymbol{\varphi}_{R,t+1}^\top \left(\boldsymbol{R}_t - \boldsymbol{E}_t^d\right) + \beta \varphi_{t+1}.
\end{aligned}$$

Employing equation (8) with $\boldsymbol{z} = \alpha \boldsymbol{\varphi}_X$ delivers

$$\begin{aligned}
&\boldsymbol{\varphi}_X^\top \boldsymbol{X}_t + \boldsymbol{\varphi}_{R,t}^\top \boldsymbol{R}_t + \varphi_t = \\
&\quad \max_{x_t, \boldsymbol{N_t}, \boldsymbol{\mathcal{K}}_t, \boldsymbol{E_t}} \log x_t + \kappa k_t + \log F(\boldsymbol{A}_t, \boldsymbol{N}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{E}_t) - \xi_0 \tau_{1,t} + \xi_0 \\
&\qquad\quad + \lambda_t^N \big(1 - \textstyle\sum_{i=1}^{I_N} N_{i,t}\big) + \lambda_t^{\mathcal{K}} \big(1 - \textstyle\sum_{i=1}^{I_K} \mathcal{K}_{i,t}\big) + \beta \varphi_{t+1} \\
&\qquad\quad + \frac{\beta}{\alpha} \left[a(\alpha \boldsymbol{\varphi}_X, \boldsymbol{A}_t, \boldsymbol{N}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{E}_t, x_t) + \boldsymbol{b}^\top(\alpha \boldsymbol{\varphi}_X) \boldsymbol{X}_t\right] + \beta \boldsymbol{\varphi}_{R,t+1}^\top \left(\boldsymbol{R}_t - \boldsymbol{E}_t^d\right).
\end{aligned}$$

Let $\boldsymbol{e}_1^\top$ denote the first (row-)unit vector, corresponding to the capital entry of $\boldsymbol{X}_t$, and let $\boldsymbol{e}_2^\top$ denote the second (row-)unit vector, corresponding to the atmospheric temperature entry. I rearrange the Bellman equation to bring all state-dependent terms to the left

$$\begin{aligned}
&\big(\boldsymbol{\varphi}_X{}^\top - \frac{\beta}{\alpha} \boldsymbol{b}^\top(\alpha \boldsymbol{\varphi}_X) - \kappa \boldsymbol{e}_1^\top + \xi_0 \boldsymbol{e}_2^\top\big) \boldsymbol{X}_t + \big(\boldsymbol{\varphi}_{R,t}^\top - \beta \boldsymbol{\varphi}_{R,t+1}^\top\big) \boldsymbol{R}_t + \varphi_t - \beta \varphi_{t+1} = \\
&\qquad \max_{x_t, \boldsymbol{N_t}, \boldsymbol{\mathcal{K}}_t, \boldsymbol{E_t}} \log x_t + \log F(\boldsymbol{A}_t, \boldsymbol{N}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{E}_t) + \xi_0 + \lambda_t^N \big(1 - \textstyle\sum_{i=1}^{I_N} N_{i,t}\big) \\
&\qquad\qquad + \frac{\beta}{\alpha} a(\alpha \boldsymbol{\varphi}_X, \boldsymbol{A}_t, \boldsymbol{N}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{E}_t, x_t) - \beta \boldsymbol{\varphi}_{R,t+1}^\top \boldsymbol{E}_t^d + \lambda_t^{\mathcal{K}} \big(1 - \textstyle\sum_{i=1}^{I_K} \mathcal{K}_{i,t}\big).
\end{aligned}$$

This Bellman equation has to hold for all feasible states (a continuum) in any feasible combination (independently). Thus, a necessary condition is that the coefficients of the states vanish, i.e.,

$$\boldsymbol{\varphi}_X^\top = \frac{\beta}{\alpha}\boldsymbol{b}^\top(\alpha\boldsymbol{\varphi}_X) + \kappa\boldsymbol{e}_1^\top - \xi_0\boldsymbol{e}_2^\top \tag{C.2}$$

$$\boldsymbol{\varphi}_{R,t}^\top = \beta\boldsymbol{\varphi}_{R,t+1}^\top. \tag{C.3}$$

Let $\boldsymbol{\varphi}_X{}^*$ denote the solution to equation (C.2). Equation (C.3) merely produces a set of Hotelling equations $\boldsymbol{\varphi}_{R,t+1}^\top = \beta^{-1}\boldsymbol{\varphi}_{R,t}^\top$ implying that the (utility measured) shadow value of each resource stock has to increase at the (utility) discount rate, which is easily satisfied. For resources that are not scarce, the shadow value is simply zero in all periods. For resources that are scarce, the initial shadow value $\varphi_{R,0}$ has to be found ex-post by solving the problem below as a function (or for a sequence) of $\varphi_{R,0}$ until the boundary solution is satisfied. Given these shadow values, the Bellman equation is satisfied if (and only if) in addition

$$\begin{aligned}\varphi_t - \beta\varphi_{t+1} = \max_{x_t,\boldsymbol{N_t},\boldsymbol{\mathcal{K}}_t,\boldsymbol{E_t}} &\log x_t + \log F(\boldsymbol{A}_t,\boldsymbol{N}_t,\boldsymbol{\mathcal{K}}_t,\boldsymbol{E}_t) + \xi_0 + \lambda_t^N\big(1-\textstyle\sum_{i=1}^{I_N} N_{i,t}\big) \\ &+\frac{\beta}{\alpha}a(\alpha\boldsymbol{\varphi}_X^*,\boldsymbol{A}_t,\boldsymbol{N}_t,\boldsymbol{\mathcal{K}}_t,\boldsymbol{E}_t,x_t) - \beta\boldsymbol{\varphi}_{R,t+1}^\top\boldsymbol{E}_t^d + \lambda_t^{\mathcal{K}}\big(1-\textstyle\sum_{i=1}^{I_K}\mathcal{K}_{i,t}\big).\end{aligned} \tag{C.4}$$

I have assumed that the problem is well-defined in that the optimizaton problem on the right has a well-defined solution. Therefore, for any φ_t, equation (C.4) merely defines a φ_{t+1} and there exists a sequence $\varphi_0, \varphi_1, \varphi_2, ...$ such that equation (C.4) and the Bellman equation are satisfied (assuming equation (C.2) has a solution). The apparent degree of freedom in choosing φ_0 is pinned down by requiring that the solution to the dynamic programming problem coincides with the underlying optimization of the sum of utility over time. Precisely, the 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). Only the absolute welfare level depends on the sequence $\varphi_0, \varphi_1, \varphi_2, \varphi_3, ...$, not the optimal policy. □

Note 1 (SCC): The social cost of carbon is the (negative of the) shadow value of atmospheric carbon $\varphi_{M,1}$ translated from utils into consumption equivalents (measured, e.g., in USD). Given logarithmic utility, this conversion uses $du_t = \frac{1}{C_t}dC_t \Rightarrow dC_t = C_t du_t = x_t Y_t du_t$ so that $SCC = -C_t\varphi_{M,1} = x_t Y_t du_t$. In some instances, the

optimal consumption rate is $x_t = 1-\beta\kappa$, a constant which can cancel a corresponding occurrence of $1-\beta\kappa$ in $\varphi_{M,1}$, making the SCC directly proportional to output.

C.2 Proof of Propositions 1 and 2

To avoid repetition, I derive the SCC under joint uncertainty about carbon flows and temperature feedbacks. Propositions 1 and 2 are then summarized as the special cases where only one of the uncertainties is active. The shocks and, thus, one-step-ahead uncertainties in the carbon uncertainty model are all independently normally distributed. The cumulant generating function of a normally distributed variable $x \sim N(\mu, \sigma^2)$ is $\log\left[\mathbb{E}\exp(ux)\right] = \mu u + \frac{u^2\sigma^2}{2}$.

The autoregressive gamma process by Gourieroux & Jasiak (2006) used in the temperature uncertainty model 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. Here the light X_t refers to a generic state and not the bold state vector $\boldsymbol{X}_t$. 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

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

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

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

I define the deterministic process neutralizing the temperature expectations to those of the deterministic model[36] 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 (19) and (20) in the main text. To apply Lemma 1, I calculate (one over α times) the cumulant generation function of $\boldsymbol{X}_t = (k_t, \boldsymbol{\tau}_t, \boldsymbol{M}_t, \boldsymbol{I}_t)$ with $\boldsymbol{I}_t = (x^M_t, \sigma^M_t, y_t, y^o_t)$ and $\boldsymbol{z} = \alpha\boldsymbol{\varphi}^\top$; I only keep track of the terms proportional to the states ignoring those ("...") that make up the function $a(\cdot)$ and that are not needed to calculate the shadow values:

$$\begin{aligned}
\frac{1}{\alpha}\log\left(\mathbb{E}\exp(\alpha\boldsymbol{\varphi}^\top\boldsymbol{X}_{t+1})|\boldsymbol{X}_t\right) &= \varphi_k(\kappa k_t - \xi_0\tau_{1,t}) \\
&+\boldsymbol{\varphi}_M^\top\boldsymbol{\Phi}\boldsymbol{M}_t + ... + (\varphi_{M,1}-\varphi_{M,2})x^M_t + \frac{\alpha}{2}(\varphi_{M1}-\varphi_{M2})^2{\sigma^M_t}^2 \\
&+\boldsymbol{\varphi}_\tau^\top\boldsymbol{\sigma}\boldsymbol{\tau}_t + \frac{\sigma^{forc}}{M_{pre}}\varphi_{\tau,1}M_{1,t} - h\varphi_{\tau,1}\gamma^z y^o_t - h\varphi_{\tau,1}(1-\epsilon(c))\tfrac{M_{1,t}}{M_{pre}} + ... \\
&+\varphi^M_x\gamma^x x^M_t + \frac{\alpha}{2}{\varphi^M_x}^2\delta^{Mx^2}\frac{M_{1,t}}{M_{pre}} + \frac{\alpha}{2}{\varphi^M_x}^2\delta^{\sigma x^2}{\sigma^M_t}^2 \\
&+\varphi^M_\sigma\gamma^\sigma{\sigma^M_t}^2 + ... + \varphi^M_\sigma\delta^{M\sigma}\frac{M_{1,t}}{M_{pre}} \\
&-\tfrac{1}{\alpha c}\tfrac{M_{1,t}}{M_{pre}}\log(1-\alpha[\varphi^\tau_y + h\varphi_{\tau,1}]c) + \tfrac{\varphi^\tau_y + h\varphi_{\tau,1}}{1-\alpha[\varphi^\tau_y+h\varphi_{\tau,1}]c}\gamma^z y_t \\
&+\varphi^\tau_{y^o}\gamma^z y^o_t + \varphi^\tau_{y^o}(1-\epsilon(c))\tfrac{M_{1,t}}{M_{pre}} + ...
\end{aligned} \tag{C.5}$$

The right hand side identifies Lemma 1's function $\boldsymbol{b}(\cdot)$. Sorting by states, equa-

[36] 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.

tion (10) delivers the equations for the shadow values

$$\varphi_k = (1+\beta\varphi_k)\kappa \quad \Rightarrow \quad \varphi_k = \frac{\kappa}{1-\beta\kappa} \tag{C.6}$$

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

$$\begin{aligned}\boldsymbol{\varphi}_M^\top = \beta\boldsymbol{\varphi}_M^\top\boldsymbol{\Phi} + \beta\Big(&\tfrac{\sigma^{forc}}{M_{pre}}\varphi_{\tau,1} + \tfrac{\alpha}{2}\tfrac{\delta^{Mx^2}}{M_{pre}}{\varphi_x^M}^2 + \tfrac{\delta^{M\sigma}}{M_{pre}}\varphi_\sigma^M + \tfrac{1}{M_{pre}}(\varphi_{y^o}^\tau - h\varphi_{\tau,1})(1-\epsilon(c)) \\ &- \tfrac{1}{M_{pre}}\tfrac{\log(1-\alpha c(\varphi_y^\tau + h\varphi_{\tau,1}))}{\alpha c}\Big)\boldsymbol{e}_1^\top\end{aligned} \tag{C.8}$$

$$\varphi_{y^o}^\tau = \beta(\varphi_{y^o}^\tau - h\varphi_{\tau,1})\gamma^z \tag{C.9}$$

$$\varphi_y^\tau = \beta\tfrac{\varphi_y^\tau + h\varphi_{\tau,1}}{1-\alpha c(\varphi_y^\tau + h\varphi_{\tau,1})}\gamma^z \tag{C.10}$$

$$\varphi_x^M = \beta(\varphi_{M,1} - \varphi_{M,2}) + \beta\varphi_x^M\gamma^x \tag{C.11}$$

$$\varphi_\sigma^M = \beta\frac{\alpha}{2}\left((\varphi_{M,1} - \varphi_{M,2})^2 + \delta^{\sigma x^2}{\varphi_x^M}^2\right) + \beta\varphi_\sigma^M\gamma^\sigma. \tag{C.12}$$

Ultimately, I am interested in the shadow value of atmospheric carbon. This shadow value depends on all the other shadow values above and I am going to solve the expressions successively. Before doing so, I note that under certainty – see DACE or work out the special case of the equations above – it is

$$\boldsymbol{\varphi_M}^{\top\,det} - \beta\boldsymbol{\varphi}_M^\top\boldsymbol{\Phi} - \beta\varphi_{\tau,1}\frac{\sigma^{forc}}{M_{pre}}\boldsymbol{e}_1^\top = 0 \Rightarrow \boldsymbol{\varphi}_M^{\top\,det} = \frac{\beta\varphi_{\tau,1}\sigma^{forc}}{M_{pre}}\boldsymbol{e}_1^\top(\mathbf{1}-\beta\boldsymbol{\Phi})^{-1} \tag{C.13}$$

$$\boldsymbol{\varphi}_\tau^{\top det} = -\xi_0(1+\beta\varphi_k)\boldsymbol{e}_1^\top(1-\beta\boldsymbol{\sigma})^{-1} \quad \Rightarrow \varphi_{\tau,1}^{det} = -\xi_0(1+\beta\varphi_k)[1-\beta\boldsymbol{\sigma})^{-1}]_{1,1} \tag{C.14}$$

which together with equation (C.6) above delivers

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

The subsequent proof relates this shadow value under certainty to the general shadow value of atmospheric carbon under uncertainty.

Note 2: Some of the extended ACE versions in Traeger (2021*a*) affect the shadow value of capital and the transformation of $\varphi_{M,1}^{det}$ from utils into consumption equivalents (i.e., the conversion in Note 1). The expressions about to be derived for the stochastic shadow value only rely on the relation between the expression $\varphi_{M,1}$ and its deterministic counterpart. These relations hold for the base ACE just as well as for the extensions derived in Traeger (2021*a*).

Temperature related shadow values:
The temperature's shadow value solves equation (C.7) to the form

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

The feedback operates through the carbon's shadow value and through the persistent shock shadow value φ_y^τ for which equation (C.10) delivers the quadratic equation

$$\begin{aligned} &\varphi_y^\tau - \alpha c {\varphi_y^\tau}^2 - \alpha c h \varphi_{\tau,1}\varphi_y^\tau = \beta\varphi_y^\tau\gamma^z + \beta h\varphi_{\tau,1}\gamma^z \\ \Leftrightarrow\ &\underbrace{\alpha c}_{\equiv\tilde{a}} {\varphi_y^\tau}^2 + \underbrace{(\beta\gamma^z + \alpha c h\varphi_{\tau,1} - 1)}_{\equiv\tilde{b}}\varphi_y^\tau + \underbrace{\beta h\varphi_{\tau,1}\gamma^z}_{\equiv\tilde{c}} = 0. \end{aligned} \tag{C.15}$$

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.[37] Then

$$\begin{aligned} \varphi_y^\tau &= \frac{-2\tilde{c}}{\tilde{b} \pm \sqrt{\tilde{b}^2 - 4\tilde{a}\tilde{c}}} \\ &= \varphi_{\tau,1}\frac{2\beta h\gamma^z}{(1-\beta\gamma^z - \alpha c h\varphi_{\tau,1}) \pm \sqrt{(1-\beta\gamma^z - \alpha c h\varphi_{\tau,1})^2 - 4\alpha c h\varphi_{\tau,1}\beta\gamma^z}} \\ &= \frac{\beta\gamma^z}{1-\beta\gamma^z}\underbrace{\frac{2}{1 - \frac{\alpha c h\varphi_{\tau,1}}{1-\beta\gamma^z} \pm \sqrt{\left(1 - \frac{\alpha c h\varphi_{\tau,1}}{1-\beta\gamma^z}\right)^2 - 4\frac{\alpha c h\varphi_{\tau,1}}{1-\beta\gamma^z}\frac{\beta\gamma^z}{1-\beta\gamma^z}}}}_{\equiv T} h\varphi_{\tau,1} \end{aligned} \tag{C.16}$$

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 (C.10) as

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

Carbon-flow uncertainty:

[37] 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 connecting the uncertain SCC directly to the deterministic SCC.

Equation (C.8) delivers the shadow value vector equation

$$\boldsymbol{\varphi}_M^\top = \beta\Bigg(\frac{\sigma^{forc}}{M_{pre}}\varphi_{\tau,1} + \frac{\alpha}{2}\frac{\delta^{Mx^2}}{M_{pre}}{\varphi_x^M}^2 + \frac{\delta^{M\sigma}}{M_{pre}}\varphi_\sigma^M + \frac{1}{M_{pre}}(\varphi_{y^o}^\tau - h\varphi_{\tau,1})(1-\epsilon(c)) \\ - \frac{1}{M_{pre}}\frac{\log(1-\alpha c(\varphi_y^\tau + h\varphi_{\tau,1})}{\alpha c}\Bigg)\left[(\mathbb{1}-\beta\boldsymbol{\Phi})^{-1}\right]_{1,\cdot}\,. \tag{C.17}$$

Dividing the second through the first shadow value entry I obtain

$$\varphi_{M,2} = \underbrace{\frac{\left[(\mathbb{1}-\beta\boldsymbol{\Phi})^{-1}\right]_{1,2}}{\left[(\mathbb{1}-\beta\boldsymbol{\Phi})^{-1}\right]_{1,1}}}_{\equiv r}\,\varphi_{M,1}\,. \tag{C.18}$$

Equation (C.11) delivers the shadow value

$$\varphi_x^M = \frac{\beta}{1-\gamma^x\beta}(\varphi_{M,1}-\varphi_{M,2}) = \underbrace{\frac{\beta}{1-\gamma^x\beta}(1-r)}_{\equiv A}\,\varphi_{M,1}\,, \tag{C.19}$$

where the second equality uses equation (C.18). Substituting these results into equation (C.12) delivers

$$\varphi_\sigma^M = \beta\frac{\alpha}{2}\frac{(\varphi_{M,1}-\varphi_{M,2})^2 + \delta^{\sigma x^2}{\varphi_x^M}^2}{1-\gamma^\sigma\beta} = \underbrace{\beta\frac{\alpha}{2}\frac{(1-r)^2+\delta^{\sigma x^2}A^2}{1-\gamma^\sigma\beta}}_{\equiv B}{\varphi_{M,1}}^2. \tag{C.20}$$

Inserting equation (C.19) and (C.20) into the atmospheric shadow value component of equation (C.17) results in the quadratic equation

$$\begin{aligned}\varphi_{M,1} &= \beta\left(\frac{\alpha}{2}\frac{\delta^{Mx^2}}{M_{pre}}{\varphi_x^M}^2 + \frac{\delta^{M\sigma}}{M_{pre}}\varphi_\sigma^M\right)\left[(\mathbb{1}-\beta\boldsymbol{\Phi})^{-1}\right]_{1,1} + \frac{\beta}{M_{pre}}\Bigg(\sigma^{forc}\varphi_{\tau,1} \\ &\quad +(\varphi_{y^o}^\tau - h\varphi_{\tau,1})(1-\epsilon(c)) - \frac{\log(1-\alpha c(\varphi_y^\tau+h\varphi_{\tau,1})}{\alpha c}\Bigg)\left[(\mathbb{1}-\beta\boldsymbol{\Phi})^{-1}\right]_{1,1} \\ &= \underbrace{\beta\left(\frac{\alpha}{2}\frac{\delta^{Mx^2}}{M_{pre}}A^2 + \frac{\delta^{M\sigma}}{M_{pre}}B\right)\left[(\mathbb{1}-\beta\boldsymbol{\Phi})^{-1}\right]_{1,1}}_{\equiv\hat{a}}{\varphi_{M,1}}^2 + \end{aligned} \tag{C.21}$$

$$\underbrace{\frac{\beta}{M_{pre}}\left(\sigma^{forc}\varphi_{\tau,1}+(\varphi_{y^o}^\tau-h\varphi_{\tau,1})(\delta_\tau-\epsilon(c))-\frac{\log(1-\alpha c(\varphi_y^\tau+h\varphi_{\tau,1})}{\alpha c}\right)\left[(\mathbb{1}-\beta\boldsymbol{\Phi})^{-1}\right]_{1,1}}_{\equiv\hat{c}}.$$

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

$$\varphi_{M,1} = \frac{2\hat{c}}{1 + \sqrt{1 - 4\hat{a}\hat{c}}} = \hat{c}\Bigg(1 + \underbrace{\frac{1 - \sqrt{1 - 4\hat{a}\hat{c}}}{1 + \sqrt{1 - 4\hat{a}\hat{c}}}}_{\equiv \theta_M}\Bigg) \approx \hat{c}\left(1 + \frac{\hat{a}\hat{c}}{1 - \hat{a}\hat{c}}\right), \tag{C.22}$$

where the approximation is first order around $\hat{a}\hat{c} = 0$ in both numerator and denominator. The term $\hat{a}$ is

$$\begin{aligned}
\hat{a} &= \beta\left(\frac{\alpha}{2}\frac{\delta^{Mx^2}}{M_{pre}}A^2 + \frac{\delta^{M\sigma}}{M_{pre}}B\right)\left[(\mathbb{1} - \beta\mathbf{\Phi})^{-1}\right]_{1,1} \\
&= \beta\frac{\alpha}{2}\frac{1}{M_{pre}}\Bigg[\left(\frac{\beta\delta^{Mx}}{1 - \gamma^x\beta}\right)^2(1 - r)^2 \\
&\qquad + \frac{\beta\delta^{M\sigma}}{1 - \gamma^\sigma\beta}\left((1 - r)^2 + \left(\frac{\beta\delta^{\sigma x}}{1 - \gamma^x\beta}\right)^2(1 - r)^2\right)\Bigg]\left[(\mathbb{1} - \beta\mathbf{\Phi})^{-1}\right]_{1,1} \\
&= \frac{\alpha}{2}\frac{\beta}{M_{pre}}\left[A^{M\to x^2} + A^{M\to\sigma}A^{\sigma\to x^2} + A^{M\to\sigma}\right](1 - r)^2\left[(\mathbb{1} - \beta\mathbf{\Phi})^{-1}\right]_{1,1} \\
&= \frac{\alpha}{2}\frac{\beta}{M_{pre}}\left[A^{M\to x^2} + A^{M\to\sigma}A^{\sigma\to x^2} + A^{M\to\sigma}\right]\frac{\left(\left[(\mathbf{1} - \beta\mathbf{\Phi})^{-1}\right]_{1,1} - \left[(\mathbf{1} - \beta\mathbf{\Phi})^{-1}\right]_{1,2}\right)^2}{\left[(\mathbf{1} - \beta\mathbf{\Phi})^{-1}\right]_{1,1}}
\end{aligned} \tag{C.23}$$

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 (C.16), the evaluation of

$$\begin{aligned}\varphi_y^\tau + h\varphi_{\tau,1} &= \left(\frac{\beta\gamma^z}{1-\beta\gamma^z}T + 1\right) h\varphi_{\tau,1} = \frac{1+\beta\gamma^z(T-1)}{1-\beta\gamma^z} h\varphi_{\tau,1} \\ &= \Big(1+\beta\gamma^z(T-1)\Big)\bar{h}\varphi_{\tau,1}.\end{aligned}$$

Using the definition $F \equiv \alpha c \frac{h}{1-\beta\gamma^z}\varphi_{\tau,1} = \alpha c \bar{h}\varphi_{\tau,1}$ and recalling the abbreviation T from equation (C.16), I define the expression

$$\begin{aligned}\theta_\tau^\dagger \equiv \beta\gamma^z(T-1) &= \beta\gamma^z \left(\frac{2}{1-F+\sqrt{(1-F)^2 - 4F\frac{\beta\gamma^z}{1-\beta\gamma^z}}} - 1\right) \\ &= \beta\gamma^z \frac{1+F-\sqrt{(1-F)^2 - 4F\frac{\beta\gamma^z}{1-\beta\gamma^z}}}{1-F+\sqrt{(1-F)^2 - 4F\frac{\beta\gamma^z}{1-\beta\gamma^z}}} \approx \frac{\beta\gamma^z F}{1-\beta\gamma^z - F} \qquad \text{(C.24)}\end{aligned}$$

Using this definition (in the second step) and denoting the shadow value of atmospheric carbon under certainty by $\varphi_{M,1}^{det}$ (see equation C.13), the term $\hat{c}$ defined in equation (C.21) becomes

$$\begin{aligned}\hat{c} &= \frac{\beta\sigma^{forc}\varphi_{\tau,1}}{M_{pre}} \left[(\mathbb{I} - \beta\mathbf{\Phi})^{-1}\right]_{1,1} \left(1 + \frac{\bar{h}}{\sigma^{forc}}\Bigg(-1 \right. \\ &\qquad \left. - \frac{\log(1-\alpha c\big(1+\beta\gamma^z(T-1)\big)\bar{h}\varphi_{\tau,1}}{\alpha c\bar{h}\varphi_{\tau,1}} + \epsilon(c)\Bigg)\right) \\ &= \varphi_{M,1}^{det}\left(1 + \frac{\bar{h}}{\sigma^{forc}}\left(\frac{-\log\big(1-\alpha c\bar{h}\varphi_{\tau,1}(1+\theta_\tau^\dagger)\big)}{\alpha c\bar{h}\varphi_{\tau,1}} - 1 + \epsilon(c)\right)\right) \\ &= \varphi_{M,1}^{det}\left(1 + \frac{\bar{h}}{\sigma^{forc}}\left(\frac{-\log\big(1-F(1+\theta_\tau^\dagger)\big)}{F} - 1 + \epsilon(c)\right)\right). \qquad \text{(C.25)}\end{aligned}$$

The joint first order approximation in $\theta_\tau^\dagger$ and F (first approximation), and a first order approximation in F using the definition of $\theta_\tau^\dagger$ (second approximation) deliver

$$\theta_\tau^* \equiv \frac{-\log\big(1-F(1+\theta_\tau^\dagger)\big)}{F} - 1 \;\approx\; \theta_\tau^\dagger + \frac{1}{2}F \;\approx\; \frac{1}{2}\,\frac{1+\beta\gamma^z}{1-\beta\gamma^z}\,F. \qquad \text{(C.26)}$$

Summarizing the case of joint carbon flow and temperature uncertainty: I define $\theta_M^* = \hat{a}\hat{c}$. The main text uses the definitions $carb_1$ and $\Delta carb$ defined on

page 14 simplifying the representation of $\hat{a}$ in equation (C.23). Moreover, equations (C.25) and (C.26) imply

$$\hat{c} = \varphi_{M,1}^{det}\left(1 + \frac{\bar{h}}{\sigma^{forc}}\left(\theta_\tau^* + \epsilon(c)\right)\right) \text{ with } \theta_\tau^* \text{ as in equation (C.26).}$$

Equation (C.22) delivers $\varphi_{M,1} = \hat{c}\big(1 + \theta_M\big)$ with

$$\theta_M = \frac{1 - \sqrt{1 - 4\theta_M^*}}{1 + \sqrt{1 - 4\theta_M^*}} \approx \frac{\theta_M^*}{1 - \theta_M^*}. \tag{C.27}$$

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

$$SCC_t = SCC_t^{det}\big(1 + \theta_M\big)\left(1 + \frac{\bar{h}}{\sigma^{forc}}\Big(\theta_\tau^* + \epsilon(c)\Big)\right). \tag{C.28}$$

Summarizing Result for Proposition 1:
In the case of Proposition 1, $\bar{h} = 0$ so that the second bracket in equation (C.28) is unity. The approximation in equation (C.27) delivers the approximation in equation (16) of the proposition.

The convexity of θ_M in θ_M^* follows from equation (C.27):

$$\theta_M(\theta_M^*) = \frac{1 - (1 - 4\theta_M^*)^{.5}}{1 + (1 - 4\theta_M^*)^{.5}}$$

$$\theta'_M(\theta_M^*) = 2(1 - 4\theta_M^*)^{-.5}\frac{1 + (1 - 4\theta_M^*)^{.5} + 1 - (1 - 4\theta_M^*)^{.5}}{(1 + (1 - 4\theta_M^*)^{.5})^2} = 4\frac{(1 - 4\theta_M^*)^{-.5}}{(1 + (1 - 4\theta_M^*)^{.5})^2}$$

$$\theta''_M(\theta_M^*) = 8\frac{(1 - 4\theta_M^*)^{-1.5}(1 + (1 - 4\theta_M^*)^{.5})^2 + 2(1 - 4\theta_M^*)^{-1}(1 + (1 - 4\theta_M^*)^{.5})}{(1 + (1 - 4\theta_M^*)^{.5})^4}$$

which is positive for $1 - 4\theta_M^* > 0 \Leftrightarrow \theta_M^* < \frac{1}{4}$, a condition that has to hold for the solution to be well-defined in the first place (see below).

Summarizing Results for Proposition 2:
In the absence of uncertainty over carbon flow also the endogenous carbon flow uncertainty vanishes and

$$\delta^{M\sigma} = \delta^{\sigma x} = \delta^{Mx} = 0 \quad \Rightarrow \quad \hat{a} = 0 \quad \Rightarrow \quad \theta_M^* = 0 \quad \Rightarrow \quad \theta_M = 0.$$

Thus, the result is stated by equation (C.28) where the first multiplier of the deterministic SCC is unity. 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 (C.26) and equation (C.24) defines $\theta_\tau^\dagger$. The approximation in equation (21) follows from the approximation in equation (C.26).

Well-definedness (domains of the propositions):
The solution to the quadratic equation (C.21) for carbon flow uncertainty is well-defined for $\theta_M^* \equiv \hat{a}\hat{c} < \frac{1}{4}$. The solution to the quadratic equation (C.15) is well-defined for $\tilde{b}^2 > 4\tilde{a}\tilde{c} \Leftrightarrow (1-F)^2 > 4F\frac{\beta\gamma^z}{1-\beta\gamma^z}$ as best observed from the transformed equations (C.16) or (C.24). This inequality is equivalent to $\beta\gamma^z < \frac{(1-F)^2}{(1-F)^2+4F}$. Moreover, the cumulative generating functions of the autoregressive gamma process first observed in equation (C.5) requires that the argument of the logarithm remains positive. After solving for the shadow values, this logarithm reappears in equation (C.25) and requires $F\ (1+\theta_\tau^\dagger) < 1$. A sufficient but not necessary condition for this inequality to hold is $F < \frac{1}{3}$; using the definition of $\theta_\tau^\dagger$ in equation (C.24) find that $F\ (1+\theta_\tau^\dagger) < F\big(1+\frac{1+F}{1-F}\big) < \frac{1}{3}\big(1+\frac{2}{4}\big) = 1$.

C.3 Proof of Propositions 3 and 4

Damage uncertainty moves slightly beyond Lemma 1, but in a somewhat trivial sense. First, the formulation of the stochastic damage-adaptation process introduces an additional control variale, i_t, and changes the production function from the form $F(\boldsymbol{A}_t, \boldsymbol{N}_t, \boldsymbol{K}_t, \boldsymbol{E}_t)$ to the form $F(\boldsymbol{A}_t, \boldsymbol{N}_t, \boldsymbol{K}_t, \boldsymbol{E}_t, i_t)$. Second, the damage-adaptation state π_t directly affects consumption in a given period and, therefore, not only affects the next period state but also the expression for the current period's welfare. As a result, I restate the Bellman equation (C.1) in the slightly extended form

$$V(k_t, \boldsymbol{\tau}_t, \boldsymbol{M}_t, \boldsymbol{R}_t, \boldsymbol{I}_t, t) = \max_{x_t, \boldsymbol{N_t}, \boldsymbol{\mathcal{K}}_t, \boldsymbol{E_t}, i_t} \log x_t + \kappa k_t + \log F(\boldsymbol{A}_t, \boldsymbol{N}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{E}_t, i_t) \qquad \text{(C.29)}$$
$$-\xi_0\tau_{1,t} + \xi_0 + \xi_0\pi_t + \frac{\beta}{\alpha}\log\Big(\mathbb{E}_t \exp\big[\alpha\ V(k_{t+1}, \boldsymbol{\tau}_{t+1}, \boldsymbol{M}_{t+1}, \boldsymbol{R}_{t+1}, \boldsymbol{I}_{t+1}, t)\big]\Big).$$

The informational state vector $\boldsymbol{I}_t$ now contains the additional component π_t. Employing the affine value function, the expectation on the right of equation (C.29) yields

the novel terms

$$\frac{\beta}{\alpha}\log\Big(\mathbb{E}_t\exp\Big(\alpha\varphi_\pi\big(-d\sqrt{\tau_{1,t}-\eta_d}+h(i)+\gamma_d\pi_t\big)+\alpha\varphi_k\xi_0\pi_t\Big)\Big) \tag{C.30}$$

$$=\beta\alpha\varphi_\pi^2\frac{\sigma_d^2}{2}(\tau_{1,t}-\eta_d)+\beta\underbrace{\frac{1}{\alpha}\log\big(\mathbb{E}_t\exp\big(\alpha\varphi_\pi h(i)\big)\big)}_{\equiv\tilde{h}(\alpha,\varphi_\pi,\eta)}+\beta\gamma_d\pi_t+\beta\varphi_k\xi_0\pi_t$$

deriving form the damage-adaption process' equation of motion (24) and the damage equation (23)'s direct impact on next period's capital stock.

First, I obtain a new coefficient matching condition for the new state π_t. Collecting the terms proportional to Π_t in equation (C.29), equation (C.30), and from the l.h.s. value function delivers the condition for the new shadow value

$$\begin{aligned}&\varphi_\pi=\beta\varphi_\pi\gamma_d+(1+\beta\varphi_k)\xi_0\\ \Rightarrow\ &\varphi_\pi=\frac{(1+\beta\varphi_k)\xi_0}{1-\beta\gamma_d}=\frac{\xi_0}{(1-\beta\gamma_d)x^*}.\end{aligned}$$

Second, the term $\tilde{h}(\alpha,\varphi_\pi,\eta)$ in equation (C.30) only affects the optimization problem and welfare, but not the shadow value of carbon. Third, the remaining term in equation (C.30) affects the equation for the atmospheric temperature's shadow value, turning equation (C.7) in the proof in Section C.2 into

$$\begin{aligned}&\boldsymbol{\varphi}_\tau^\top=\beta\boldsymbol{\varphi}_\tau^\top\boldsymbol{\sigma}-(1+\beta\varphi_k)\xi_0\boldsymbol{e}_1^\top+\beta\alpha\varphi_\pi^2\frac{\sigma_d^2}{2}\tau_{1,t}\\ \Rightarrow\ &\boldsymbol{\varphi}_\tau^\top=\left(-(1+\beta\varphi_k)\xi_0+\beta\alpha\varphi_\pi^2\frac{\sigma_d^2}{2}\right)\boldsymbol{e}_1^\top(\mathbb{I}-\beta\boldsymbol{\sigma})^{-1}\\ \Rightarrow\ &\boldsymbol{\varphi}_\tau^\top=-(1+\beta\varphi_k)\xi_0\left(1++\frac{\beta(-\alpha)}{\xi_0}\left(\frac{(1+\beta\varphi_k)\xi_0}{1-\beta\gamma_d}\right)^2\frac{\frac{\sigma_d^2}{2}}{1+\beta\varphi_k}\right)\boldsymbol{e}_1^\top(\mathbb{I}-\beta\boldsymbol{\sigma})^{-1}\\ \Rightarrow\ &\boldsymbol{\varphi}_\tau^\top=-(1+\beta\varphi_k)\xi_0\Big(1+\underbrace{\beta(-\alpha)\frac{\xi_0}{(1-\beta\gamma_d)^2}\frac{\sigma_d^2}{2}}_{\equiv\theta_d}\Big)\boldsymbol{e}_1^\top(\mathbb{I}-\beta\boldsymbol{\sigma})^{-1}=(1+\theta_d)\varphi_{\tau,1}^{det}\\ \Rightarrow\ &\varphi_{\tau,1}=(1+\theta_d)\varphi_{\tau,1}^{det},\end{aligned}$$

where $\varphi_{\tau,1}^{det}$ refers to the deterministic shadow value defined in equation (C.14), coinciding with the earlier result under uncertainty that only affected $\boldsymbol{\varphi}_M$. Thus, the shadow value $\varphi_{\tau,1}$ picks up the new factor $(1+\theta_d)$ as compared to the settings, either certainty or joint uncertainty over carbon flow and temperature.

For the proof of Proposition 3, the shadow value of atmospheric carbon, equation (C.13), is proportional to $\varphi_{\tau,1}$ and, thus, $\varphi_{M,1}$ and the SCC merely get multiplied by $(1+\theta_d)$ with $\theta_d = \beta(-\alpha)\frac{\xi_0}{(1-\beta\gamma_d)^2}\frac{\sigma_d^2}{2}$ define in the equation above.

For the proof of Proposition 4, I have to identify multiple occurrences of the shadow value $\varphi_{\tau,1}$. First, it is part of the term $F \equiv \alpha c \frac{h}{1-\beta\gamma^z}\varphi_{\tau,1}$ and, second, it is part of $\hat{c}$ defined in equation (C.15). The first dependence adds the $(1+\theta_d)$ to the new factor $F^{new} = F^{orig}(1+\theta_d)$ as stated in the proposition. The second dependence adds the factor $(1+\theta_d)$ to equation (C.25) and, therefore, equation (C.28) defining the total uncertainty multiplier to the SCC.

C.4 Proof of Proposition 5

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 1, the informational state variable is the current realization of the autoregressive shock μ_t^j, which is known at time t. Then, the Bellman equation (9), using once again an affine trial solution, 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) \\
& \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) \\
& + \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) \\
& + 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) \\
& + \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) \\
& + \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{(C.31)}
\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

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

Moreover

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

Sorting coefficients in the Bellman equation by states gives

$$\begin{aligned}&\left(\boldsymbol{\varphi}_M^\top - \beta \boldsymbol{\varphi}_M^\top \boldsymbol{\Phi} - \beta \varphi_{\tau,1} \frac{\sigma^{forc}}{M_{pre}} \boldsymbol{e}_1^\top \right) \boldsymbol{M}_t + \left(\boldsymbol{\varphi}_\tau^\top - \beta \boldsymbol{\varphi}_\tau^\top \boldsymbol{\sigma} + (1 + \beta\varphi_k) \xi_0 \boldsymbol{e}_1^\top \right) \boldsymbol{\tau}_t \\ &\left(\varphi_k - (1 + \beta\varphi_k)\kappa \right) k_t + \left(\boldsymbol{\varphi}_{R,t}^\top - \beta \boldsymbol{\varphi}_{R,t+1}^\top \right) \boldsymbol{R}_t \\ &+ \left(\varphi_\mu^j - \beta(\varphi_j + \gamma^j \varphi_\mu^j) \right) \mu_t^j \\ &+ \varphi_t \ = \ \beta \varphi_{t+1} \\ &\qquad + \ \log x_t^*(\varphi_k) + \beta\varphi_k \log(1 - x_t^*(\varphi_k)) + (N_t + \beta\varphi_k)\xi_0 \\ &\qquad + (1 + \beta\varphi_k)\kappa k_t + (1 + \beta\varphi_k) \log F(\boldsymbol{A}_t, \boldsymbol{N}_t^*, \boldsymbol{\mathcal{K}}_t^*, \boldsymbol{E}_t^*) \\ &\qquad + \beta\varphi_k(\log[1 + g_{k,t}] - \log[\delta_k + g_{k,t}]) - \beta \boldsymbol{\varphi}_{R,t+1}^\top \boldsymbol{E}_t^{d*} \\ &\qquad + \ \beta\varphi_{M,1} \left(\textstyle\sum_{i=1}^{I^d} E_{i,t}^* + E_t^{exo} \right) + \beta\varphi_{\tau,1} \frac{\sigma^{forc}}{M^{pre}} G_t, + \frac{\beta}{\alpha} \sum_{l=1}^{\infty} \frac{(\alpha\varphi_\mu^j)^l}{l!} \kappa_l \ .\end{aligned} \quad \left. \begin{aligned} \\ \\ \\ \\ \\ \\ \end{aligned} \right\} \equiv B(\cdot) \tag{C.32}$$

Apart from the coefficients related to the new state, this equation mostly resembles the equation of the deterministic base model (see DACE). As before, also the coefficient on the new state μ_t has to vanish implying

$$\varphi_\mu^j = \frac{\beta}{1 - \gamma^j \beta} \varphi_j \ . \tag{C.33}$$

In the present proposition, I am interested in welfare differences. Thus, I also (and in particular) need to keep track of the affine terms of the value function. For the deterministic case, DACE derives

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

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.32) the affine part of the value function evolves as

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

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

$$\begin{aligned}
\Delta W_{general}^{AR} &= V^{AR} - V^{det} = \varphi_\mu^j\mu_0^j + \varphi_0^{unc} - \varphi_0^{det} \\
&= 0 + \beta(\varphi_1^{AR} - \varphi_1^{det}) + \frac{\beta}{\alpha}\sum_{l=1}^{\infty}\frac{(\alpha\varphi_\mu^j)^l}{l!}\kappa_l \\
&= \beta(\varphi_2^{AR} - \varphi_2^{det}) + \sum_{p=1}^{2}\frac{\beta^p}{\alpha}\sum_{l=1}^{\infty}\frac{(\alpha\varphi_\mu^j)^l}{l!}\kappa_l \\
&= \lim_{t\to\infty}\beta(\varphi_t^{AR} - \varphi_t^{det}) + \lim_{t\to\infty}\sum_{p=1}^{t}\frac{\beta^p}{\alpha}\sum_{l=1}^{\infty}\frac{(\alpha\varphi_\mu^j)^l}{l!}\kappa_l \\
&= \lim_{t\to\infty}\beta(\varphi_t^{AR} - \varphi_t^{det}) + \left(\frac{1}{1-\beta} - 1\right)\frac{1}{\alpha}\sum_{l=1}^{\infty}\frac{(\alpha\varphi_\mu^j)^l}{l!}\kappa_l \\
&= \sum_{l=1}^{\infty}\frac{\beta}{1-\beta}\frac{1}{\alpha}\frac{(\alpha\varphi_\mu^j)^l}{l!}\kappa_l = \frac{\beta}{1-\beta}\frac{1}{\alpha}G_\chi\big(\alpha\varphi_\mu^j\big),
\end{aligned}$$

which, together with equation (C.33) delivers the result stated in part (3) of the proposition. For a normally distributed mean-zero shock χ_t^j the first cumulant (expectation) is zero and only the second cumulant contributes (all others being zero), delivering part (1) of the proposition.

In the case of **anticipated learning**, the informational state evolves as $\mu_{t+1}^j = \frac{\sigma_{\epsilon,t}^2\tilde{z} + \sigma_{\nu,t}^2\mu_t^j}{\sigma_{\epsilon,t}^2 + \sigma_{\nu,t}^2}$, where $\tilde{z}$ is the observation, which is distributed as the sum of measurement error and Bayesian prior $z \sim N(\mu_t^j, \sigma_{\epsilon,t}^2 + \sigma_{\nu,t}^2)$ (see as well footnote 27). The variance of the normal-normal Bayesian learning model evolves deterministically as $\sigma_{\epsilon,t+1}^2 =$

$\frac{\sigma^2_{\nu,t}\sigma^2_{\epsilon,t}}{\sigma^2_{\nu,t}+\sigma^2_{\epsilon,t}}$. Therefore, the Bellman equation (C.31) becomes

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

Moreover

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

Then, the Bellman equation delivers in close analogy to above and the deterministic setting

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

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

$$\begin{aligned}
& \varphi_\mu^j \Big(1 - \beta \underbrace{\Big(\frac{\sigma^2_{\nu,t}}{\sigma^2_{\epsilon,t} + \sigma^2_{\nu,t}} + \frac{\sigma^2_{\epsilon,t}}{\sigma^2_{\epsilon,t} + \sigma^2_{\nu,t}} \Big)}_{=1} \Big) = \beta \varphi_j \\
\Leftrightarrow \ & \varphi_\mu^j = \frac{\beta}{1 - \beta} \varphi_j \ . \qquad \text{(C.34)}
\end{aligned}$$

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

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

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 (C.34). Moreover, the term in brackets is equivalent to

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

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.

D Calibration of the stochastic processes

At the outset of the calibration, I pick a common long-run risk persistence $\gamma \equiv \gamma^x = \gamma^\sigma = \gamma^z = \gamma^d$ for all stochastic processes for the baseline, and I pick a high and low variation. As explained in Appendix A, the autoregressive constants are a proxy for the speed of learning under epistemological uncertainty. The best researched and quantitatively most important uncertainty governs the climate sensitivity. Our progress in narrowing down this uncertainty over the course of three decades of research and two decades of IPCC reports is sobering if not deflating as the confidence interval hardly moved at all. We probably have not been much better when it came to carbon flow uncertainty and damages. In fact, the more recent high end damages might have actually increased the confidence interval again. Such highly persistent uncertainty motivates my pick of $\gamma = 0.9$ for the baseline. Appendix A motivates $\gamma \approx 0.9$ somewhat more formally comparing a simple autoregressive shock model to a Bayesian learning model. The autoregression coefficient of a Bayesian learning model would fall over time and my base choice will be a bit too low in the near future and too high in the further future. I will therefore add sensitivity scenarios with $\gamma = 0.8$ and $\gamma = 0.95$.

The parameters $\eta \equiv \eta_M = \eta_\tau = \eta_d$ characterize how steeply uncertainty increases when deviating from the preindustrial equilibrium versus how much uncertainty already prevails in equilibrium. While η does not directly affect the SCC, it indirectly affects the SCC through the calibration. A higher value of $\eta \in (0, 1)$ implies that more of the uncertainty is causal to the antropogenic perturbation of the climate system and the corresponding feedbacks as opposed to already present independently. For the baseline, I pick $\eta = \frac{2}{3}$, which might be a more conservative assignment of uncertainty to our perturbations and climate feedbacks. Therefore, I will also present results for a scenario with $\eta = 0.9$, which pins uncertainty more sharply on the deviation from preindustrial.

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 double 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 double. Given ACE's decadal time step, I simply average the values of the decade before and after doubling of the carbon concentration. 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 2. 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 2. The parameters of this baseline calibration are $\gamma = 0.9$, $h = 0.23$, $\eta = 0.8$, $c = 0.21$, and $\epsilon = 0.05$. The parameters and graphs for the scenario variations of Table 1 are presented in Figure 6.

It proves more difficult to find good probabilistic information governing carbon dynamics. Joos et al. (2013) provide a useful benchmark against which to calibrate and, in particular, recalibrate carbon flow uncertainty across different scenarios. Joos et al. (2013) subject 18 different carbon cycle models to a 5000Gt carbon pulse. The study calculates the variance of the pulse evolution across models. The base calibration opens the interaction channel between conditional expectations and conditional volatility by setting $\delta^{\sigma x} = 1$. I set the exogenous uncertainty, which has no impact on the SCC, to a nominal value of $\bar{\sigma} = 0.1\%$.[38] I then simultaneously vary $\delta^{Mx} = \delta^{M\sigma}$ until I approximately match the uncertainty suggested by the model comparison study, which is the case for $\delta^{Mx} = \delta^{M\sigma} = 40$ in the baseline calibration. ACE 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 so is 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 governing the mere pulse.[39] The upper left panel of Figure 5 presents the result. Overall, the mag-

[38] Given the already small impact of carbon uncertainty on the SCC, I rather err on the side of exaggerating the endogeneity of the uncertainty.

[39] 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

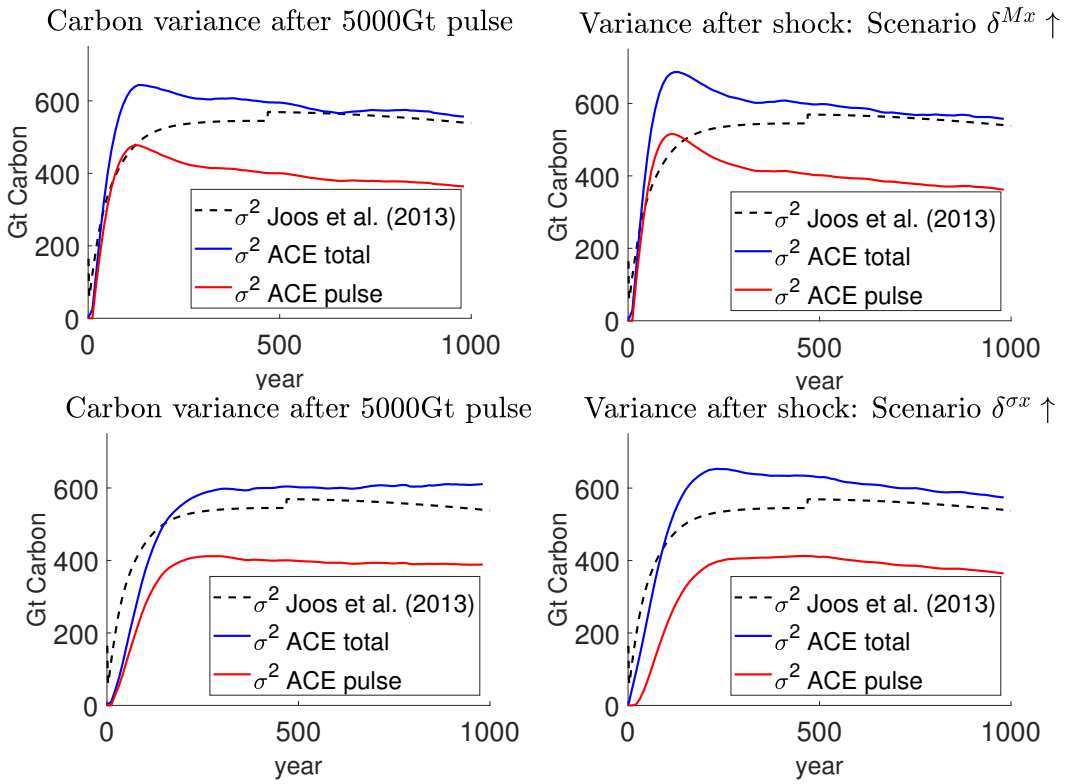


Figure 5: Carbon flow uncertainty compared to Joos et al. (2013). Black dashed: Uncertainty (variance) across 18 deterministic carbon cycle models responding to a 5000Gt carbon pulse as reported by Joos et al. (2013). After year 500 several models leave the ensemble causing a discontinuity in the variance. Blue: Overall uncertainty (variance) in ACE when subjected to the same pulse. Red: Variance measure of uncertainty attributable to the carbon pulse itself.
Upper left: Baseline. Upper right: Scenario $\delta^{Mx}\uparrow$ pushing uncertainty onto conditional expectations. Lower left: Scenario $\delta^{M\sigma}\uparrow$ pushing uncertainty onto stochastic volatility. Lower right: Scenario $\delta^{\sigma x}\uparrow$ pushing uncertainty onto the interaction of stochastic volatility and conditional expectations.

nitude of the uncertainty is fairly close. The parameters and graphs for the scenario variations of Table 1 are presented in Figure 6. The right panel of Figure 2 illustrates the implications of this uncertainty calibration in a more meaningful scenario. It uses the DICE 2013 business as usual scenario and depicts the original DICE evolution of atmospheric CO_2 ("data"), the evolution in the deterministic ACE model (using a 10 year rather than 5 year time step), and the evolution of the mean CO_2 concentration, all of which are basically indistinguishable. The figure then depicts the 90% simulated confidence intervals, the median which lies slightly below the mean, and a set of 100 random paths (yellow). The graph illustrates that the uncertainty about the future

the pulse, taking differences between the paths with coinciding draws of the shock.

carbon evolution resulting from my calibration is substantial.

The calibration above merely assumed $\delta^{Mx} = \delta^{M\sigma}$ and there is no data to meaningfully tell the different sources of uncertainty apart. Similarly, I merely assumed that the stochastic volatility enters the conditional expectations with equal magnitude ($\delta^{\sigma x} = 1$). Therefore, I introduce three additional scenarios, each of which fixes all but one of these parameters to unity, increasing the free parameter to approximately match the overall uncertainty suggested by my comparison with Joos et al. (2013). The remaining panels in Figure 5 shows the calibration results of the corresponding scenarios. Increasing the uncertainty over the conditional expectations while fixing $\delta^{M\sigma} = \delta^{\sigma x} = 1$ delivers an increase of $\delta^{Mx} = 50$ (upper right panel, scenario $\delta^{Mx} \uparrow$). Increasing the uncertainty over the stochastic volatility while fixing $\delta^{Mx} = \delta^{\sigma x} = 1$ delivers an increase of $\delta^{M\sigma} = 165$ (lower left panel, $\delta^{M\sigma} \uparrow$). Note that this value is much higher because the scenario essentially tries to match an overall uncertainty over levels by merely increasing stochastic volatility. And yet the SCC impact will turn out very small. Increasing the interaction channel while fixing $\delta^{Mx} = \delta^{M\sigma} = 1$ delivers a multiplier of the stochastic volatility in the conditional expectations channel of $\delta^{\sigma x} = 12$ (lower right panel, $\delta^{\sigma x} \uparrow$).

When it comes to damages, we even lack any consensus governing a best guess estimate. The DICE model in its various vintages suggests a good 2% damage at a 3C warming,[40] whereas a recent meta-analysis by Howard & Sterner (2017) as well as a survey by Pindyck (2020) suggest damages closer to 10% of global output at a similar warming.[41] See Traeger (2021*a*) for an extended discussion as well as plots of these various damages functions (Figure 2). That paper presents results based on two damage functions, one calibrated to DICE and a scenario labeled "HSP" calibrated to Howard & Sterner (2017) and Pindyck's (2020). Here, I essentially average the corresponding semi-elasticities of production ξ_0 across the two scenario for my expected damage level and then use their difference to inform the damage variance. In more detail, the "preferred estimate" of Howard & Sterner (2017) uses a quadratic dam-

[40] The DICE 2013 and 2016 models assumem damage of 2.35% and 2.08% based on the functional form $1 - \frac{1}{1+a*T^2}$ with $a = 0.00267$ and $a = 0.00236$, respectively. Note that usage of this functional form is not fully consistent across texts and model implementations and sometimes the implementations employ the non-normalized damage function $a * T^2$ instead, which results in slightly higher damages.

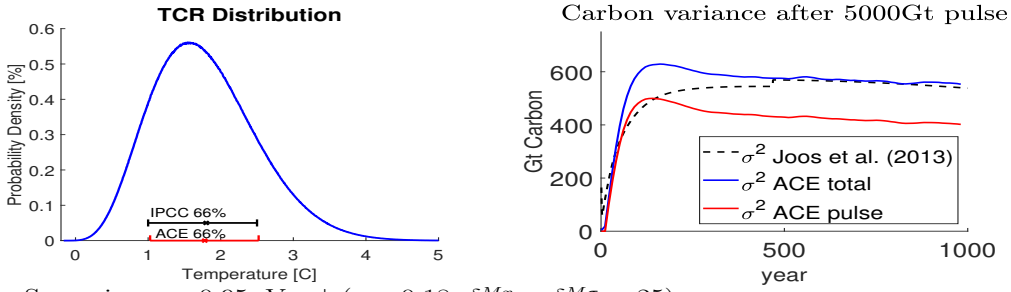
[41] Whereas Howard & Sterner (2017) use a similar quadratic damage function over temperature as Nordhaus (2017), Pindyck's (2020) asks for damage estimates based on a business as usual emission scenario 50 years into the future. Using the IPCC's RCP 8.5 scenario, the information translates to a similar 10% loss at a 3C warming as Howard & Sterner (2017).

age function without normalization and damages will eventually outgrow production and even the capital stock. By construction, ACE assumes that damages cannot exceed production. Therefore, Traeger (2021*a*) renormalizes Howard & Sterner's (2017) damage function and decides to match the renormalized function at a 3C warming ($x_0^{HSP} = 0.10$). Alternatively, the paper also shows the damage curve when matches the non-normalized damage curve at a 3C warming ($x_0^{HSP^*} = 0.11$). Given the present analysis explicitly introduces uncertainty, I will average the two values before taking the average with $\xi_0^{DICE} = 0.022$. As a result, I find $\xi_0 = 0.063$. I then use the following scenario to calibrate damage uncertainty. I let temperature increase linearly from today's 1C warming to an end of the century warming of 3C. This temperature increase is deterministic. I then calibrate the 2100 damages at a 3C warming so that the 10^{th} percentile of the damage distribution implies a 2.0% loss of world output and the 90^{st} percentile implies a 10.0% loss. This calibration implies $\sigma_d = 0.26$ and a one sigma interval of damages of $[3.0\%, 9.1\%]$, which seems somewhat reasonable given the widely different damage estimates and the fact that most of the more recent findings suggest that damages are rather a bit higher than assumed in DICE.[42] The variations for the scenario variations of Table 1 all recalibrate the σ_d to imply the same 10^{th} and 90^{st} percentiles.[43]

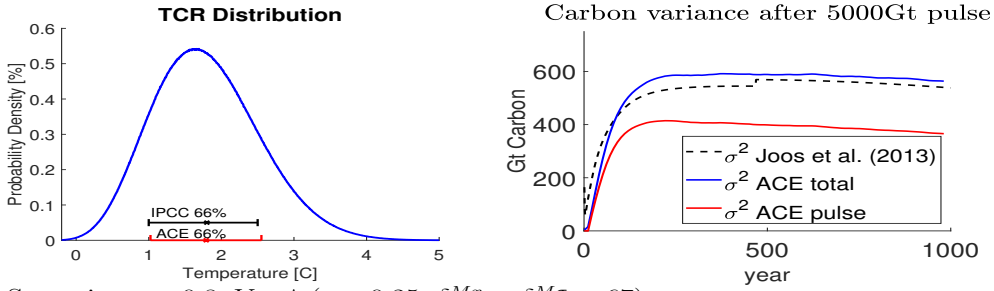
[42]By proposition 3, uncertainty over the function $h(\cdot)$ will not affect the optimal carbon tax and, thus, I do not calibrate h for the present paper. However, making h stochastic affects the cited simulation results informing σ_d. A stochastic h increases the variance and confidence intervals. Bearing this in mind, a stochastic h would slightly enlarge those uncertainty intervals going along with the cited presentation, just as a higher η_d would slightly reduce those uncertainty intervals. I note that there is a (very) small downward bias in the expected damages under uncertainty, where the mean damage of 6.034 lies slightly below the deterministic simulation outcome at expected damages of 6.087% (based on a simulation with 10 million draws). This miniscule difference could in principle be offset by an ever so slight adjustment of the the zero expectation of h.

[43]In the scenario $\gamma = 0.95$, $\eta \uparrow$ I find $\eta_d = 0.88$, in the scenario $\gamma = 0.95$, Var $\downarrow$ I find $\sigma_d = 0.23$, in the scenario $\gamma = 0.8$, Var $\uparrow$ I find $\sigma_d = 0.313$, and in the scenario $\eta = 0.9$, Var $\uparrow$ I find $\sigma_d = 0.295$.

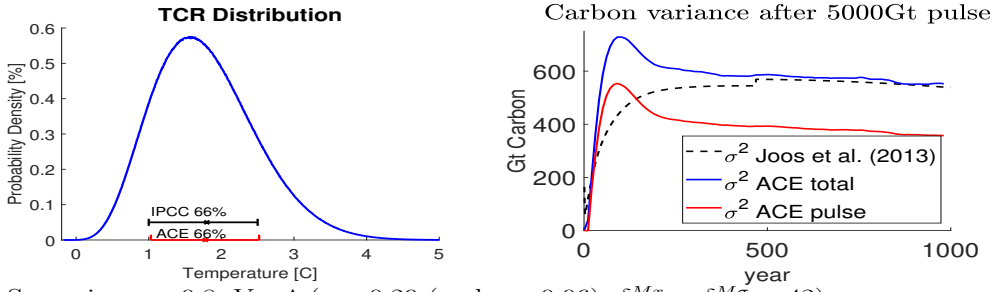
Scenario: $\gamma = 0.95$, $\eta \uparrow$ ($\eta_\tau = 0.8$, $\eta_M = 0.9$)



Scenario: $\gamma = 0.95$, Var $\downarrow$ ($c = 0.18$, $\delta^{Mx} = \delta^{M\sigma} = 25$)



Scenario: $\gamma = 0.8$, Var $\uparrow$ ($c = 0.25$, $\delta^{Mx} = \delta^{M\sigma} = 67$)



Scenario: $\eta = 0.8$, Var $\uparrow$ ($c = 0.29$ (and $\epsilon = 0.06$), $\delta^{Mx} = \delta^{M\sigma} = 42$)

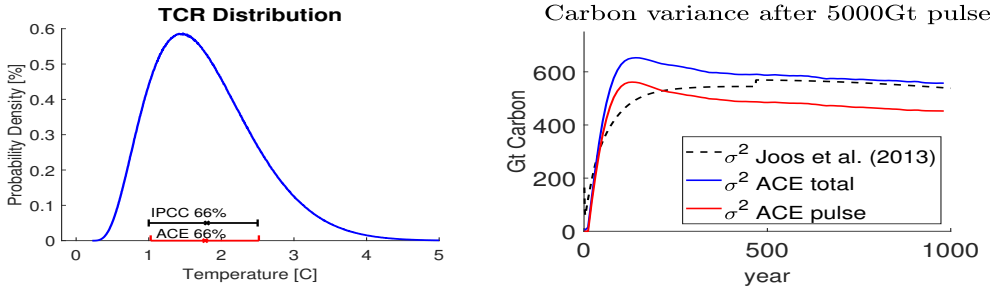


Figure 6: Calibrations of TCR distribution (left) and carbon flow uncertainty (right) for the the alternative scenarios of Table 1. Panels are the analogues to Figure 2 on the left and to Figure 5, which show the baseline calibration.