Analytic Integrated Assessment and Uncertainty

Christian Traeger

Department of Agricultural & Resource Economics, UC Berkeley

April 2015

Abstract: I develop an analytic integrated assessment model of climate change. The paper closes a gap between complex numeric models used in policy advising and stylized models built for analytic insight. The climate component contains an explicit carbon cycle, the CO_2 radiative forcing equation, and a atmosphere-ocean temperature dynamics. Economic production relies on various clean and dirty, and potentially scarce energy sources. I derive a closed-form solution of the optimal carbon tax and the welfare implications from uncertainty over the carbon cycle and the climate's sensitivity to CO_2. The formula pinpoints the different contributions from social preference, production, damages, and temperature and carbon dynamics. I discuss unforeseen persistent shocks to the system (vector autoregressive uncertainty) as well as epistemological uncertainty such as a Bayesian learning. The analysis employs non-logarithmic risk attitude and distinguishes risk aversion from a decision maker's propensity to smooth consumption over time.

JEL Codes: Q54, H43 ,E13 ,D81 ,D90 ,D61

Keywords: climate change, uncertainty, integrated assessment, dynamic programming, risk aversion, recursive utility, learning

1 Introduction

Much of the policy advice provided by climate economists uses complex numeric models that are conceived as black boxes by outsiders. Analytic models provide qualitative insight, but usually lack the detail needed for a quantitative understanding of climate policy. This paper helps to bridge the two approaches, introducing an analytic model that has the same structure and descriptive power as numeric integrated assessment models (IAMs) used in policy advising. The closed-form expressions identify and quantify the different contributions to the optimal carbon tax and measure the cost of global warming and the benefit of carbon sequestration.

Despite several decades of research, the future temperature response to emissions remains largely uncertain. I explore the welfare implications of uncertainty over the carbon cycle and over the temperature response to a given carbon concentration (climate sensitivity). I equip the model with general non-logarithmic risk attitude and use Epstein-Zin-Weil (EZW) preferences to disentangle risk attitude and consumption smoothing incentives. I discuss persistent shocks to the climate dynamics (vector autoregressive uncertainty) as well as epistemological uncertainty such as a Bayesian learning.

Numeric deterministic climate models, including DICE, FUND, PAGE, WHITCH, MERGE, REMIND, and IMAGE, provide quantitative policy recommendations. Recent stochastic numeric climate models study a combination of uncertainty, learning, and catastrophic irreversibility, using either discounted expected utility or EZW preferences.[1] A few stylized analytic models address specific questions, e.g., the choice of taxes versus quotas (Newell & Pizer 2003, Karp & Zhang 2006). Recently, Golosov et al. (2014) developed an analytically tractable IAM that combines an explicit model of the energy sectors with a linear impulse response of economic production to carbon emissions. Gerlagh & Liski (2012) introduce an additional delay accounting for the missing temperature equations, improve the calibration of the impulse response function, and introduce β, δ discounting. In tribute to the pioneering work by Golosov, Hassler, Krusell, Tsyvinski, Gerlagh, and Liski, I call the model in this paper GAUVAL, taking the ith letter of the ith author. Anderson et al. (2014) employ a (slightly simplified) version of the Golosov et al. (2014) framework to analyze robust control.[2]

[1] Notable work addresses climate and growth uncertainty in dynamic programming implementations of the DICE model (Kelly & Kolstad 1999, Kelly 2005, Leach 2007, Heutel 2011, Fischer & Springborn 2011, Traeger 2012*b*, Lemoine & Traeger 2014, Jensen & Traeger 2014, Cai et al. 2012, Lontzek et al. 2012, Kelly & Tan 2013). In addition, deterministic models are employed to simulate the consequences of given policies using a Monte-Carlo averaging of sensitivity runs (Richels et al. 2004, Hope 2006, Nordhaus 2008, Dietz 2009, Anthoff et al. 2009, Ackerman et al. 2010, Interagency Working Group on Social Cost of Carbon 2010, Pycroft et al. 2011, Kopp et al. 2012).

[2] Anderson et al. (2014) use a linear relation between the economic growth rate, temperature increase, and cumulative historic emissions, and combine the simpler analytic model with a more complex numeric IAM for quantitative simulation. The assumption that temperatures respond linear to cumulative emissions is based on the scientific literature's finding that the medium to long-term temperature response in (deterministic)

GAUVAL is the first closed-form IAM incorporating a full climate change model comprising (i) a standard carbon cycle model, (ii) the physical non-linearity in the response of radiative forcing to atmospheric CO_2 accumulation ("greenhouse effect"), and (iii) a model of the atmosphere-ocean temperature dynamics. I derive general conditions under which such a model solves in closed form, and I show that its calibration tracks the climate system at least as well as the wide-spread DICE model. As compared to its analytic predecessors, the full climate model allows me a state of the art model calibration of climate dynamics. The model structure allows me to separate tax and welfare contributions from carbon dynamics and temperature dynamics, two major research areas of climate change science. I show that the payoffs to a better understanding of the temperature dynamics (for given CO_2 concentrations) is much higher than the payoff from reducing the uncertainty over the carbon flows.

The logarithmic utility function employed in these analytic IAMs is a reasonable approximation to the intertemporal substitutability of consumption.[3] However, the implied unit elasticity largely understates risk aversion. In particular, logarithmic risk aversion in GAUVAL's predecessors implies the absence of any welfare impact of a mean-preserving spread – a serious short-coming considering that numeric and stylized models generally find uncertainty to be highly relevant. I introduce non-logarithmic risk attitude and EZW preferences to improve the market calibration of the economic model, leading to non-trivial uncertainty implications while retaining analytic tractability.[4] In a world of uncertainty, analytic models have an additional advantage: they overcome the curse of dimensionality plaguing the numeric stochastic dynamic programming implementations of IAMs.

Notable pioneering work delivering general analytic insights in stylized integrated assessment models under non-trivial risk-attitude or robust control include, Hennlock (2009), Valentini & Vitale (2014), Jensen & Traeger (2014), Anderson et al. (2014), some of which contain as well more complex numeric quantification. These papers emphasize the relevance of carefully integrating risk and its evaluation into IAMs. The present paper distinguishes itself by delivering general quantitative closed-form expressions for the welfare loss under risk, and by using a full-blown IAM that explicitly models transient temperature dynamics, radiative forcing, and the carbon cycle in the same way as the standard numeric IAMs used in policy advising.

climate models is approximately proportional to cumulative historic emissions (Allen et al. 2009, Matthews et al. 2009).

[3]Historically, most assessments of the intertemporal elasticity were (at times significantly) smaller than unity. Over the recent years, many studies measured intertemporal elasticities of substitution larger than unity, in particular, when disentangling intertemporal sustainability from risk aversion. A value of unity lies well within the reasonable range of estimates, see Bansal et al. (2012) for a detailed review.

[4]The discounted expected utility standard model assumes that aversion to risk and to deterministic consumption change over time are governed by the curvature of the same utility function and, thus, coincide. This entanglement is not a consequence of the common rationality assumptions and challenged by market observations (see section 2.5).

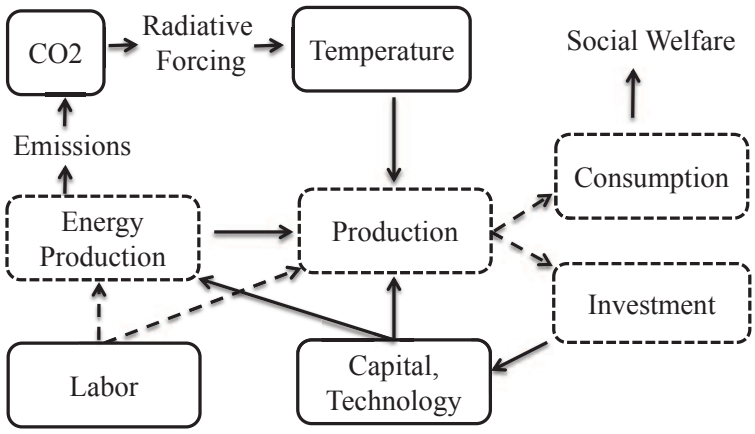

Figure 1: The structure of GAUVAL and most Integrated Assessment Models. Solid boxes characterize the model's state variables, dashed boxes are flows, and dashed arrows mark choice variables.

2 The Model

GAUVAL's structure follows that of most IAMs (Figure 1). Labor, capital, and technology create production that is either consumed or invested. Production relies on energy inputs which cause emissions. Emissions accumulate in the atmosphere, cause radiative forcing (greenhouse effect), and increase global temperature(s), which has a negative repercussion on production. This section introduces the basic model governing the economy, the energy sector, and the climate system. It derives the necessary and sufficient assumptions to solve the model in closed form, discusses the underlying calibration, and introduces preferences that disentangle risk aversion from intertemporal consumption smoothing.

2.1 The Economy

Utility is logarithmic and the social planner's time horizon is infinite. 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. Gross production derives, in Cobb-Douglas form, from technology level $A_{0,t}$, capital K_t, the energy composite E_t, and the amount of labor $N_{0,t}$ employed in the final consumption good sector

$$Y_t = A_{0,t} K_t^{\kappa} N_{0,t}^{1-\kappa-\nu} E_t^{\nu} \ .$$

The aggregate energy input E_t is a function

$$E_t = g(\boldsymbol{E}_t(\boldsymbol{A}_t, \boldsymbol{N}_t)) \tag{1}$$

of $I \in \mathbb{N}$ different energy sources, whose production levels $E_{i,t}$ are collected in the vector $\boldsymbol{E}_t \in \mathbb{R}_+^I$. These decomposed energy inputs are produced using technologies $\boldsymbol{A}_t \in \mathbb{R}_+^I$ and labor input levels $\boldsymbol{N}_t \in \mathbb{R}_+^I$. Total labor supply is normalized to unity, $\sum_{i=0}^{I} N_{i,t} = 1$. The

first I^d energy sources are fossil fuel based and emit CO_2 ("dirty"). I measure these energy sources in units of their carbon content. Their extraction is costly, they are potentially scarce, and I denote this subset of energy inputs by the vector $\boldsymbol{E_t^d} \in \mathbb{R}_+^{I^d}$. Total emissions from production $\sum_{i=1}^{I^d} E_{i,t}$. Renewable energy sources, labeled I^d+1 to I, are costly but not scarce and their production does not emit CO_2 ("clean").

For example, Golosov et al. (2014) suggest three energy sectors. Oil is scarce and extracted freely. Coal and a renewable aggregate are produced linearly using technology $A_{i,t}$ and labor inputs $N_{i,t}$: $E_{i,t} = A_{i,t}N_{i,t}$. These energy sources transform into the energy composite used in final production under constant elasticity of substitution $E_t = \left(\sum_{i=1}^{3} \alpha_i E_{i,t}^s\right)^{\frac{1}{s}}$. More generally, the present model can take detailed account of different renewable and non-renewable energy sources. Substitutability can be limited between liquid fuels and electric energy sources, but highly substitutable between coal and renewable electricity sources. Given present issues with the volatility of renewable electricity production and limitations in storage, the substitutabilities can change over time. For the purpose of the present paper, I only assume a system of energy sectors of the general form (1) that is sufficiently smooth and well-behaved to let the value function converge and to avoid boundary solutions.[5]

The dirty fossil fuel energy sources are (potentially) scarce and their resource stock in the ground $\boldsymbol{R_t^d} \in \mathbb{R}_+^{I^d}$ follows the equation of motion

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

together with the non-negativity constraint $\boldsymbol{R}_t \geq 0$ and the initial boundary condition $\boldsymbol{R}_0 \in \mathbb{R}_+^{I^d}$ given. The next section explains how the energy sector's carbon emissions increase the global atmospheric temperature $T_{1,t}$ measured as the increase over the preindustrial temperature level. This temperature increase causes damages, which destroy a fraction $D_t(T_{1,t})$ of production, $D_t(0) = 0$. Proposition 1 in section 2.3 characterizes the class of damage functions $D_t(T_{1,t})$ that permit an analytic solution of the model.

Following Golosov et al.'s (2014) assumption of full depreciation in the second period, the capital stock's equation of motion becomes

$$K_{t+1} = Y_t[1 - D_t(T_{1,t})] - C_t \ . \tag{2}$$

The model's time step of 10 years makes the capital depreciation assumption more reasonable than it might appear: instead of an annual decay that leaves some 30%-40% after 10 years, the model utilizes all of the capital after 10 years, and none after 20 years. Appendix A extends the model to allow for more sophisticated capital persistence by interacting an exogenous capital growth rate approximation with capital depreciation. This step enables GAUVAL to match the empirical capital accumulation and it makes the representative agent aware of the additional investment payoff from higher capital persistence. The extension does

[5] A sufficient but not necessary condition are smoothness, interior solutions for the controls, and convexity of the energy production set.

not affect the equations describing the optimal carbon tax or the welfare equations. The crucial implication of equation (2), as well as the empirically better founded extension in Appendix A, is that the investment rate will be independent of the system state.

2.2 The Climate System

The energy sector's CO_2 emissions enter the atmosphere and, thus, the carbon cycle. In addition to these emissions from fossil fuel burning, we emit CO_2 through land conversion, forestry, and agriculture. I denote these additional anthropogenic emission by E_t^{exo} and follow the wide-spread integrated assessment model DICE in treating E_t^{exo} as exogenous. Carbon released into the atmosphere does not decay, it only cycles through different carbon reservoirs. Let $M_{1,t}$ denote the atmospheric carbon content and $M_{2,t}, ..., M_{m,t}$, $m \in \mathbb{N}$, the carbon content of a finite number of non-atmospheric carbon reservoirs. These reservoirs continuously exchange carbon. For example, 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. Scientific climate models generally use larger numbers of reservoirs. Let $\boldsymbol{M_t}$ denote the vector of the carbon content in each reservoir and let the matrix $\boldsymbol{\Phi}$ capture the transfer coefficients. Then

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

captures the carbon dynamics. The first unit vector $\boldsymbol{e}_1$ channels new emissions from fossil fuel burning $\sum_{i=1}^{I^d} E_{i,t}$ and from land use change, forestry, and agriculture E_t^{exo} into the atmosphere $M_{1,t}$.

An increase in atmospheric carbon causes a change in our planet's energy balance. In equilibrium, the planet radiates the same amount of energy out into space that it receives from the sun. Atmospheric carbon $M_{1,t}$ and other greenhouse gases (GHGs) "trap" some of this outgoing infrared radiation, which causes the (additional, anthropogenic) radiative forcing

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

The exogenous process G_t captures non-CO_2 greenhouse gas forcing measured in CO_2 equivalents. There is no anthropogenic radiative forcing if $G_t = 0$ and $M_{1,t}$ is equal to the preindustrial atmospheric CO_2 concentration M_{pre}. We can think of radiative forcing as a small flame turned on (or up) to heat a big pot of soup, our planet. The parameter η captures the strength of this flame for a doubling of CO_2 with respect to the preindustrial concentration M_{pre}. Whereas radiative forcing is immediate, the planet's temperature (the big pot of soup) only reacts with delay. After several centuries, the new equilibrium[6] temperature correspond-

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

ing to a new level of radiative forcing F^{new} will eventually be $T_{1,eq}^{new} = \frac{s}{\eta}F^{new} = s\frac{\log\frac{M_t+G_t}{M_{pre}}}{\log 2}$. The parameter s is known as climate sensitivity. It measures the medium to long-term temperature response to a doubling of preindustrial CO_2 concentrations. Its best estimates lie currently around 3C, but the true temperature response to a doubling of CO_2 is highly uncertain.

Next period's atmospheric temperature is determined by the present temperature in the atmosphere and in the upper ocean as well as the prevailing radiative forcing. Next period's temperature change in the upper ocean is determined by the present temperature in the atmosphere and the next lower ocean layer. I denote the temperature of a finite number of ocean layers by $T_{i,t}, i \in \{2, ..., L\}, L \in \mathbb{N}$. For notational convenience, I abbreviate the atmospheric equilibrium temperature corresponding to radiative forcing level F_t by $T_{0,t} = \frac{s}{\eta}F_t$. Then, next period's temperature change in layer $i \in \{1, ..., L\}$ is determined by the present period's temperature in the adjacent layers. Frequently, this heat exchange is modeled linearly, making $T_{i,t+1}$ a weighted arithmetic mean of $T_{i-1,t}, T_{i,t}$, and $T_{i+1,t}$. Anticipating that the usual linear model defies an analytic solution to GAUVAL, I model next period's temperature as a generalized rather than arithmetic mean.[7] I denote the generalized mean's $L \times L$ weight matrix by $\boldsymbol{\sigma}$. In determining layer i's next period temperature, The weights $\sigma_{i,i+1}$ and $\sigma_{i,i-1}$ in the equation of motion for temperature $T_{i,t+1}$ specify the weights given to the temperature in the adjacent layers. As only adjacent layers affect next period's temperature, it is $\sigma_{i,j} = 0$ for $j < i-1$ and $j > i+1$. The temperature persistence in a given ocean layer $i \in \{2, ..., L-1\}$ is $\sigma_{i,i} = 1 - \sigma_{i,i-1} - \sigma_{i,i+1} > 0$. The lowest ocean layer only exchanges heat with the next upper layer implying $\sigma_{L,L} = 1 - \sigma_{L,L-1} > 0$. In the atmospheric temperature layer, the additional weight $\sigma^{forc} = \sigma_{1,0}$ determines the heat change through radiative forcing, and implies an atmospheric temperature persistence $\sigma_{1,1} = 1 - \sigma^{forc} - \sigma_{1,2} > 0$. I denote the generalized mean describing the temperature adjustments of temperature layer i by $\mathfrak{M}_i^{\boldsymbol{\sigma}}$. The equations of motion for temperature are

$$T_{i,t+1} = \mathfrak{M}_i^{\boldsymbol{\sigma}}(T_{i,t}, w_{i-1}T_{i-1,t}, w_i^{-1}T_{i+1,t}) \text{ for } i \in \{1, ..., L\}, \tag{5}$$

where the equilibrium temperature ratios w_i are empirical adjustments reflecting that the equilibrium warming does not coincide across all layers. In particular, in a warmer equilibrium the oceans lose more energy through evaporation, keeping them cooler relative to the atmosphere. Based on the data, my empirical calibration in section 2.4 assumes $w_1 = \frac{T_{1,eq}}{T_{2,eq}}$ and $w_i = 1$ for $i \neq 1$, i.e., it adjusts only for the equilibrium warming difference between atmosphere and oceans.

[7] The generalized mean of two values x_1 and x_2 with weights σ_1 and $\sigma_2 = 1 - \sigma_1$ resembles an arithmetic mean with an additional monotonic weighting function f: $\mathfrak{M}(x_1, x_2) = f^{-1}[\sigma_1 f(x_1) + \sigma_2 f(x_2)]$. Hardy et al. (1964) axiomatize the generalized mean based on the general properties of a mean value.

2.3 Solving GAUVAL

Appendix B solves GAUVAL by transforming it into an equivalent linear-in-state model (Karp 2013). This transformation also flashes out what extensions maintain (or destroy) its analytic tractability. Linear-in-state models rely on equations of motions that are linear in the state variable, and on control variables that are additively separable from the states. GAUVAL is not linear in the original state variables, but under appropriate transformations of capital and temperatures. The present paper assumes that the optimal labor allocation has an interior solution and that the scarce resources are stretched over the infinite time horizon along the optimal path, avoiding boundary value complications. For example, a complementarity assumption of scarce oil with other energy sources assures that the resources condition is met in Golosov et al. (2014). Linear-in-state models are solved by an affine value function. The following proposition summarizes the main result of Appendix B.

Proposition 1 *An affine value function of the form*

$$V(k_t, \boldsymbol{\tau}_t, \boldsymbol{M}_t, \boldsymbol{R}_t, t) = \varphi_k k_t + \boldsymbol{\varphi}_M^\top \boldsymbol{M}_t + \boldsymbol{\varphi}_\tau^\top \boldsymbol{\tau}_t + \boldsymbol{\varphi}_{R,t}^\top \boldsymbol{R}_t + \varphi_t \tag{6}$$

solves GAUVAL if, and only if, $k_t = \log K_t$, $\boldsymbol{\tau}_t$ *is a vector composed of the generalized temperatures* $\tau_i = \exp(\xi_i T_i)$, $i \in \{1, ..., L\}$, *the damage function takes the form*

$$D(T_{1,t}) = 1 - \exp[-\xi_0 \exp[\xi_1 T_{1,t}] + \xi_0],\ \xi_0 \in \mathbb{R},$$

the mean in the equation of motion (5) *for temperature layer* $i \in \{1, ..., L\}$ *takes the form*

$$\mathfrak{M}_i^{\boldsymbol{\sigma}}(T_{i,t}, w_i^{-1} T_{i-1,t}, w_{i+1} T_{i+1,t}) = \frac{1}{\xi_i} \log \Big((1 - \sigma_{1,i-1} - \sigma_{1,i+1}) \exp[\xi_i T_{i,t}] + \sigma_{i,i-1} \exp[\xi_i w_{i-1} T_{i-1,t}] + \sigma_{i,i+1} \exp[\xi_i w_i^{-1} T_{i+1,t}] \Big),$$

and the parameters ξ_i *take the values* $\xi_1 = \frac{\log 2}{s} \approx \frac{1}{4}$ *and* $\xi_{i+1} = \frac{\xi_i}{w_i}$ *for* $i \in \{1, ..., L-1\}$.

The coefficients φ are the shadow value of the respective state variables, and $\top$ denotes the transpose of a vector of shadow values. The coefficient on the resource stock has to be time-dependent: the shadow value of the exhaustible resource increases (endogenously) over time following the Hotelling rule. The process φ_t captures the value contribution of the exogenous processes including technological progress. The damage function has to be of a double-exponential form with a free parameter ξ_0, which scales the severity of damages at a given temperature level. The damage parameter ξ_0 is the semi-elasticity of net production with respect to a change of transformed atmospheric temperature $\tau_{1,t} = \exp(\xi_1 T_{1,t})$. The generalized mean $\mathfrak{M}_i^{\boldsymbol{\sigma}}$ uses the non-linear weighting function $\exp[\xi_i \cdot]$. Section 2.4 analyzes how well these assumptions match the actual climate dynamics and current assumptions about economic damages, calibrating the weight matrix $\boldsymbol{\sigma}$, the atmosphere-ocean equilibrium temperature difference w_1, and the damage parameter ξ_0.

Expressed in terms of the vector of transformed temperature states $\boldsymbol{\tau}$, the temperatures' equations of motion (5) take the simpler form

$$\boldsymbol{\tau}_{t+1} = \boldsymbol{\sigma}\boldsymbol{\tau}_t + \sigma^{forc}\frac{M_{1,t} + G_t}{M_{pre}}\boldsymbol{e}_1 \ . \tag{7}$$

I remind the reader that the parameter σ^{forc} is the weight on radiative forcing in the atmospheric temperature's equation of motion. To achieve additive separability between controls and states, the consumption rate $x_t = \frac{C_t}{Y_t[1-D_t(T_t)]}$ replaces absolute consumption as the consumption-investment control. Under the assumptions of proposition 1, the optimal consumption rate is

$$x_t^* = 1 - \beta\kappa \ . \tag{8}$$

The other controls depend on the precise form of the energy sector. The shadow values (value function coefficients) are

$$\varphi_k = \frac{\kappa}{1-\beta\kappa} \tag{9}$$

$$\boldsymbol{\varphi}_\tau^\top = -\xi_0(1+\beta\varphi_k)\boldsymbol{e}_1^\top(\mathbf{1}-\beta\boldsymbol{\sigma})^{-1} \tag{10}$$

$$\boldsymbol{\varphi}_M^\top = \frac{\beta\varphi_{\tau,1}\sigma^{forc}}{M_{pre}}\boldsymbol{e}_1^\top(\mathbf{1}-\beta\boldsymbol{\Phi})^{-1} \tag{11}$$

$$\boldsymbol{\varphi}_{R,t}^\top = \beta^t\boldsymbol{\varphi}_{R,0}^\top \ .$$

I will discuss these shadow values and their economic implications in detail in section 3. Note that the initial values $\boldsymbol{\varphi}_{R,0}^\top$ of the scarce resources depend on the precise evolution of the economy and, thus, further assumptions on the energy sector as well as the chosen climate policy.

2.4 Calibration

I employ the carbon cycle of DICE 2013. Running the model in a 10 year time step, I double the transition coefficients. Figure 4 in section 4.2 confirms that the rescaled 10 year transition matrix yields an evolution of the carbon stock indistinguishable from that of the original 5 year step of DICE 2013. I employ the usual capital share $\kappa = 0.3$ and use the International Monetary Fund's (IMF) 2015 investment rate forecast $1-x^* = 25\%$ to calibrate pure time preference. Equation (8) implies $\beta = \frac{1-x^*}{\kappa} = \frac{0.25}{0.3}$ and an annualized discount rate of $\rho = \frac{1}{10}\log\beta = 1.75\%$. The conversion of utility values into 2015 USD relies on the log utility's implication that $dC = C\,du = x\,Y\,du$, where the consumption rate is $x = 75\%$ and Y is equal to ten times (time step) the IMF's global economic output forecast of $Y_{2015} = 81.5$ trillion USD.

Economic damage functions are crucial and yet hard to determine. The most wide-spread IAM DICE uses the form $D(T) = \frac{1}{1+0.0028T^2}$. Nordhaus (2008) calibrates the coefficient 0.0028 based on a damage survey for a 2.5C warming. I calibrate GAUVAL's damage coefficient to match Nordhaus' calibration points of 0 and 2.5°C exactly, delivering $\xi_0 = 0.0222$. Figure 2 compares the resulting damage curve to that of the DICE-2007 model. The figure also depicts the damage curve $D(T) = 1 - 1/\left((1 + \frac{T}{20.46})^2 + (\frac{T}{6.081})^{6.754}\right)$ suggested by Weitzman (2010), who argues that little is known about damages at higher temperature levels, and that a more convex damage curve passing through Nordhaus' calibration point at 2.5°C is just as likely. GAUVAL's damage function initially generates damages that are slightly higher as compared to DICE-2007, matches it exactly at 2.5°C, delivers slightly lower damages up to a 12°C warming, and finally generates higher damages for global warming above 12°C, warming levels that imply a hard-to-conceive change of life on the planet. Figure 2 also depicts two dashed versions of GAUVAL's damage function. The lower curve reduces the damage parameter by 50%, resulting in a damage function that lies almost everywhere below DICE. The higher curve increases the damage parameter by 50%, resulting in a damage function that lies everywhere above that of DICE. Section 3 discusses how such changes affect welfare and the optimal carbon tax.

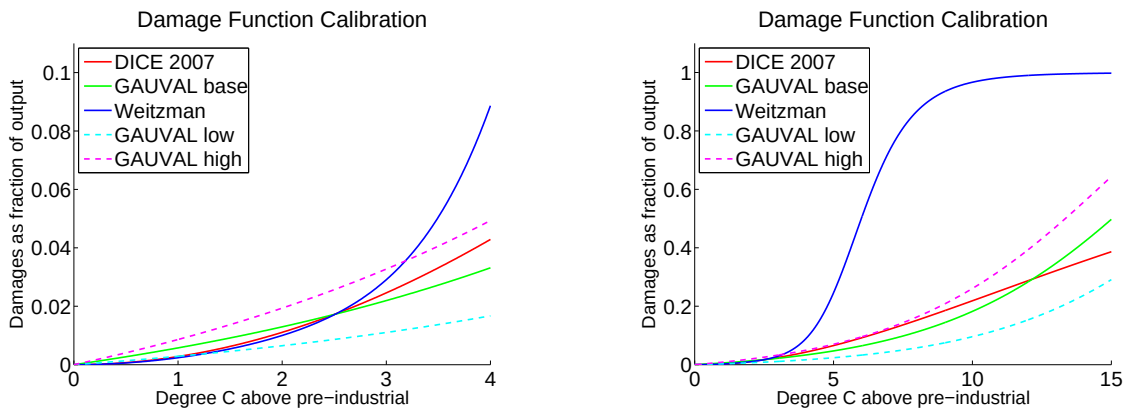

Figure 2: GAUVAL's damage function compared to that of DICE-2007 and a highly convex damage function suggested by Weitzman (2010). All three lines coincide for a 2.5°C warming, the common calibration point based on Nordhaus (2008). The dashed curves depict GAUVAL's damage function for a ±50% variation of the base case damage coefficient $\xi_0 \approx 0.022$.

The common linear approximation of the heat transfer between the atmosphere and the ocean layers defies an analytic solution of the model. I now calibrate the alternative system of equations (5) based on the generalized means derived in Proposition 1. I use the emission scenarios of the recent assessment report of the Intergovernmental Panel on Climate Change IPCC (2013). These so-called Representative Concentration Pathways (RCP) replace the (SRES-) scenarios of the earlier assessment reports. They are labeled by the approximate radiative forcing levels they produce by the end of the century (measured in W/m^2). These new RCP scenarios are better suited for calibration than the earlier SRES scenarios because

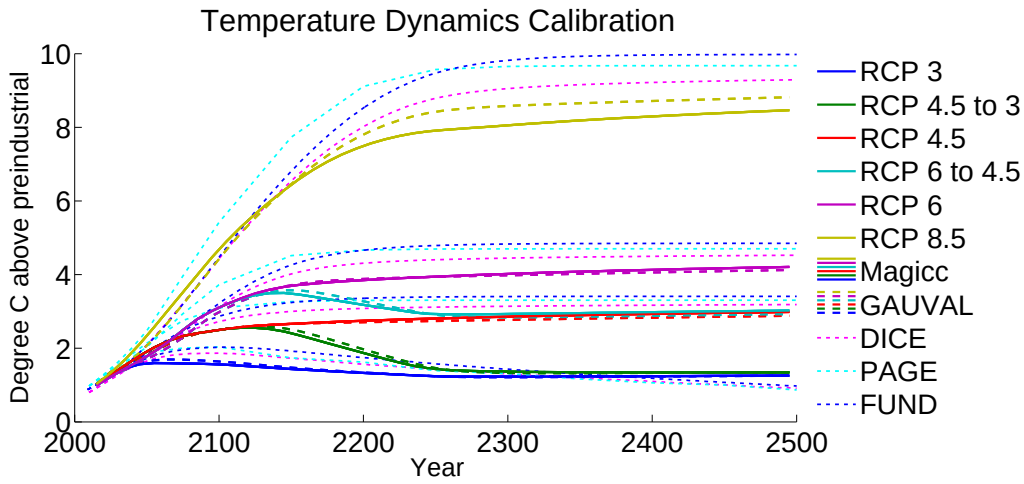

Figure 3: GAUVAL's response vis a vis Magicc's response to the color coded radiative forcing scenarios used in the latest IPCC assessment report. RCP 3 is the strongest stabilization scenario and RCP 8.5 is a business as usual scenario. The Magicc model emulates the large scientific models and is used in the IPCC's assessment reports. GAUVAL matches Magicc's temperature response very well for the "moderate" warming scenarios and reasonably well for RCP 8.5. By courtesy of R. Calel, the figure presents as well the corresponding temperature response of DICE 2013, PAGE 09, and FUND 3.9, all of which do worse as compared to GAUVAL. The scenarios RCP 4.5 to 3 and RCP 6 to 4.5 are scenarios switching from a higher to a lower forcing trajectory. They are not part of the IPCC's official scenario selection but useful to calibrate the model across a temperature peak.

they are defined for longer time horizons (Moss et al. 2007). I use the Magicc6.0 model by Meinshausen et al. (2011) to simulate the RCP scenarios over a time horizon of 500 years. The model emulates the results of the large atmospheric ocean general circulation models (AOGCMs) and is employed in the IPCC's assessment report. DICE was calibrated to one of the old SRES scenarios using an earlier version of Magicc. My calibration of GAUVAL uses three ocean layers (upper, middle, and deep) as compared to Magicc's 50 and DICE's single ocean layer(s).

Figure 3 shows the calibration results. The solid lines represent Maggic's response to the radiative forcing of the RCP scenarios, whereas the dashed lines represent GAUVAL's atmospheric temperature response. In addition to the original RCP scenarios, I include two scenarios available in Magicc6.0 that initially follow a higher radiative forcing scenario and then switch over to a lower scenario (RCP 4.5 to 3 and RCP6 to 4.5). These scenarios would be particularly hard to fit in a model only tracing atmospheric temperature. The ability to fit temperature dynamics across a peak is important for optimal policy analysis. GAUVAL's temperature model does an excellent job in reproducing Magicc's temperature response for the scenarios up to a radiative forcing of $6W/m^2$. It performs slightly worse for the high business as usual scenario RCP8.5, but still well as compared to other IAMs.

2.5 Uncertainty

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).[8] Merely increasing the utility function's curvature would result in a much larger risk-free discount rate than observed in the markets (risk-free rate puzzle). From a different perspective, the market rejects the assumption that the intertemporal eslasticity of substitution fully determines risk attitude. This assumption is built into the standard expected utility model and implies a form of risk neutrality in intertemporal choice (Traeger 2014). I follow the asset pricing literature, an increasing strand of macroeconomic literature, and some recent numeric approaches to climate change assessment in using Epstein-Zin-Weil preferences to accommodate a realistic coefficient of risk aversion, which I disentangle from the unit elasticity of intertemporal substitution.

The extension results in a Bellman equation with a non-linear risk aggregation

$$V(k_t, \boldsymbol{\tau}_t, \boldsymbol{M}_t, \boldsymbol{R}_t, t) = \max_{x_t, \boldsymbol{N}_t} \log c_t + \frac{\beta}{\alpha} \log\Big(\mathbb{E}_t \exp\big[\alpha\big(V(k_{t+1}, \boldsymbol{\tau}_{t+1}, \boldsymbol{M}_{t+1}, \boldsymbol{R}_{t+1}, t)\big)\big]\Big). \quad (12)$$

Expectations $\mathbb{E}_t$ are conditional on time t information. In general, consumption and next period's states are uncertain. The non-linear uncertainty aggregator is a generalized mean $f^{-1}\,\mathbb{E}_t\, f$ with $f(\cdot) = \exp(\alpha\ \cdot)$. A negative parameter α characterizes risk aversion in the intertemporal sense, axiomatically defined as intrinsic risk aversion in Traeger (2014). The limit $\alpha \to 0$ recovers the usual Bellman equation where risk aversion is merely generated by aversion to intertemporal inequality. Appendix C explains the relation between equation (12) and Epstein & Zin's (1991) original formulation for this special case of an intertemporal elasticity of substitution of unity. The coefficient of constant relative risk aversion in Epstein & Zin's (1991) definition of Arrow-Pratt risk aversion is RRA$= 1-\alpha^* = 1-\frac{\alpha}{(1-\beta)}$. The asset pricing literature estimates RRA in the range $[6, 9.5]$ corresponding to $\alpha \in [-1, -1.5]$. Note that Epstein-Zin preferences face the same issue as standard expected utility theory when it comes to calibrating risk aversion in the small and in the large (Rabin 2000): calibrating aversion on small bets requires higher degrees of risk aversion than seem reasonable for large bets. In consequence, I use $\alpha = -1$ for quantitative examples with high uncertainty and $\alpha = -1.5$ for quantitative examples with low uncertainty. Figure 5 in Appendix C illustrates the corresponding risk aversion for a small and a large binary lottery. The analytic formulas will make it easy for the reader to vary the degree of risk aversion for the quantitative results.

[8] Perhaps the lowest estimate is obtained combining the long-run risk model and the Barro-Riesz model, still resulting in an estimated coefficient of relative risk aversion of 6.4 (Nakamura et al. 2013).

3 Results from the Deterministic Model

The social cost of carbon (SCC) is the money-measured present value welfare loss from emitting an additional ton of CO_2. The economy in section 2.1 decentralizes in the usual way and the Pigovian carbon tax is the SCC along the optional trajectory of the economy. In the present model, the SCC is independent of the precise path of the economy and, thus, this unique SCC is the optimal carbon tax. The present section discusses the interpretation and quantification of its closed-form solution. It explores the social cost of global warming and the social benefits of carbon sequestration. A proposition establishes that mass conservation of CO_2 makes the SCC highly sensitive to pure time preference (not to the consumption discount rate in general).

3.1 The Price of Atmospheric Carbon

The social cost of carbon derives from equations (9-11) converted into money-measured consumption equivalents (see Appendix B). The general formula and its monetary value for 2015 are

$$SCC_t = \underbrace{\frac{\beta Y_t}{M_{pre}} \underbrace{\xi_0}_{2.2\%}}_{25.5\frac{\$}{tC}} \underbrace{\left[(\mathbf{1} - \beta\boldsymbol{\sigma})^{-1}\right]_{1,1}}_{1.4} \underbrace{\sigma^{forc}}_{0.42} \underbrace{\left[(\mathbf{1} - \beta\boldsymbol{\Phi})^{-1}\right]_{1,1}}_{3.7} = 56.5 \ \$/tC \ , \tag{13}$$

where $[\cdot]_{1,1}$ denotes the first element of the inverted matrix in squared brackets. As emphasized by Golosov et al. (2014), the SCC is proportional to production Y_t and increases over time at the rate of economic growth. In the present formula, the ratio of production to pre-industrial carbon emissions sets the units of the carbon tax. The concentration M_{pre} comes into play because anthropogenic radiative forcing and, thus, temperature increase is relative to the atmospheric saturation of CO_2 that prevailed in pre-industrial times. The discount factor β reflects a one period delay between temperature increase and production impact. The damage parameter ξ_0 represents the constant semi-elasticity of net production to a transformed temperature increase, i.e., to an increase of $\tau_1 = \exp(\xi_1 T_1)$. These terms together would imply a carbon tax of 25.5$ per ton of carbon.

The subsequent terms paint a detailed picture of the climate dynamics. Appendix D provides a simple illustrations for a two layer carbon and temperature system. I start with an interpretation of the term $[(\mathbf{1} - \beta\boldsymbol{\Phi})^{-1}]_{1,1}$ expanding the expression in its Neumann series ($\beta\boldsymbol{\Phi}$ is a bounded operator):

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

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

states the fraction of carbon leaving the atmosphere for layers $j \in \{1, ..., m\}$ in the first time step and arriving back to the atmosphere in the second time step. In summary, the term $[(\mathbf{1} - \beta\mathbf{\Phi})^{-1}]_{1,1}$ characterizes in closed form the discounted sum of CO_2 persisting in and returning to the atmosphere in all future periods. Its quantification states that the persistence of carbon increases the earlier value of 25.5\$/tC by a factor of 3.7. Without warming delay and the temperature's atmosphere ocean interaction the carbon tax would be 95\$/tC.

The terms $[(\mathbf{1} - \beta\boldsymbol{\sigma})^{-1}]_{1,1}\, \sigma^{forc}$ capture the atmosphere-ocean temperature delay dynamics. Analogously to the interpretation in the case of carbon, the expression $[(1 - \beta\boldsymbol{\sigma})^{-1}]_{1,1}$ characterizes the generalized heat flow that enters, stays, and returns to the atmospheric layer. Note that the simple closed-form expression in equation (13) captures an infinite double-sum stating that an additional ton of carbon emissions today causes radiative forcing in all future periods, and that the resulting radiative forcing in any given period causes warming in all subsequent periods. The parameter σ^{forc} captures the speed at which atmospheric temperature responds to radiative forcing. The response delay implied by its value around 0.4 significantly reduces the SCC. However, at the same time, the ocean implied temperature persistence increases the SCC by a factor of 1.4. Together, the ocean-atmosphere temperature dynamics reduce the carbon tax to a factor of 0.6 and its value of 56.5 USD per ton of carbon. Expressed in tons of CO_2, this SCC accounts to 15.5 USD. At the gas pump, it translates into 14 cent per gallon or 4 cent per liter. The (dashed) variation of the damage function in Figure 2 implies a $\pm 50\%$ variation of the semi-elasticity ξ_0 and, thus, the SCC. Ignoring the transitory atmosphere-ocean temperature dynamics calibrated in Figure 3 would overestimate the carbon tax by 70%. Ignoring carbon persistence would result in a carbon tax that is only 27% of its actual value.

Embedded in equation (13) is the social cost of a marginal temperature increase (SCT) in degree Celsius

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

Unlike the SCC and the transformed temperature state's shadow value, the cost of a marginal temperature increase in degree Celsius depends on the prevailing temperature level. This observation reflects the convexity of damages in temperature. Integrating the shadow value of a temperature increase from pre-industrial to present temperature levels yields the welfare cost of the present temperature increase

$$\Delta W^{Temp}_{USD\ 2015}(T \approx 0.77C) = Y\xi_0 \left[(1 - \beta\boldsymbol{\sigma})^{-1}\right]_{1,1} (\exp(\xi_1 T) - 1) \approx \$5 \text{ trillion},$$

or 6% of world output. What does this number summarize? It is larger than the 0.43% of immediate output loss per period characterized by the damage function and depicted in Figure 2. It is smaller than the present value welfare loss of keeping temperatures perpetually at a 0.77 C temperature increase, which amounts to 27% of world output. The number captures the welfare loss induced by present atmospheric warming, assuming that the planet

would return to its pre-industrial equilibrium with the delay captured by the heat transfer matrix $\boldsymbol{\sigma}$. The number neglects that we have already warmed the oceans and that warming is caused by persistent CO_2 emissions that will keep radiative forcing above the pre-industrial level. The social cost of the present atmospheric CO_2 increase is

$$\Delta W^{CO_2}_{USD\ 2015}(M_1 \approx 397ppm) = SCC\ (M - M_{pre}) \approx \$14 \text{ trillion} ,$$

or 17% of world output. This number reflects the damage already in the pipeline from present atmospheric CO_2. It does not include the CO_2 increase in the oceans or the non-CO_2 greenhouse gases, and the damage is additional to the above cited social cost of the temperature increase that already took place.

A much discussed geoengineering "solution" to climate change sequesters carbon into the oceans. E.g., engineers are currently exploring mechanisms to extract CO_2 from the exhaustion pipes of coal power plants, planning to pump it into the deep ocean. Pumping the CO_2 into layer i, instead of emitting into the atmosphere, results in the welfare gain

$$\Delta W^{seq} = \varphi_{M,i} - \varphi_{M,1} = \frac{\beta\varphi_{\tau,1}\sigma^{forc}}{M_{pre}} \Big(\big[(\mathbf{1} - \beta\mathbf{\Phi})^{-1}\big]_{1,i} - \big[(\mathbf{1} - \beta\mathbf{\Phi})^{-1}\big]_{1,1} \Big). \tag{14}$$

The bracket on the right hand side captures the (discounted and aggregated) difference in the amount of carbon prevailing in the atmosphere over time when an emission unit is injected into layer i instead of the atmosphere.[9] The shadow value difference in equation (14) will reappear in several settings of carbon cycle uncertainty. GAUVAL evaluates this welfare gain from pumping a ton of carbon into the upper ocean layer to $57 - 16 = 41$ USD, and to almost the full 57 when pumping the carbon into the deep ocean.[10] Appendix D.3 illustrates equation (14) for a two layer carbon cycle and discusses more generally the relation between carbon prices in different reservoirs.

3.2 The Optimal Carbon Tax: A Declaration of Independence

In general, the optimal carbon tax is the SCC along the optimal emission trajectory. The SCC in equation (13) is independent of the absolute stock of carbon in the atmosphere. In consequence, the SCC in GAUVAL is independent of the future emission trajectory, and

[9] This intuition is more easily observed using the Neumann series for the expression: $\Delta W^{seq} = \frac{\beta\varphi_{\tau,1}\sigma_1^{up}}{M_{pre}} \Big(\beta\left[\mathbf{\Phi}_{1,i} - \mathbf{\Phi}_{1,1}\right] + \sum_{n=2}^{\infty}\sum_{j,l}(\beta)^n\ \mathbf{\Phi}_{1,j}\big(\mathbf{\Phi}^{n-2}\big)_{j,l}\left[\mathbf{\Phi}_{l,i} - \mathbf{\Phi}_{l,1}\right]\Big)$. The first term in the brackets captures the difference between carbon flow from the ocean into the atmosphere $\mathbf{\Phi}_{1,i}$ and the persistence of carbon in the atmosphere $\mathbf{\Phi}_{1,1}$. The second term captures the fraction of carbon reaching the atmosphere after n periods if the carbon initially enters ocean layer i as opposed to entering the atmosphere directly (read right to left). The matrix entry $(\mathbf{\Phi}^{n-2})_{j,l}$ captures the overall carbon flow and persistence from layer l to j after $n-2$ periods. It approaches the stationary distribution given by its (right) eigenvectors (in all columns).

[10] Note that the present damage function does not explicitly model damages from ocean acidification. A follow-up paper will focus on modeling individual damage channels in more detail.

the SCC directly specifies the optimal carbon tax. This finding already prevails in Golosov et al. (2014). It raises a certain degree of discomfort: our optimal effort to reduce a ton of carbon is independent of whether we are in a world of high or low carbon concentrations, and independent of whether we are in a world of high or low prevailing temperatures. The discomfort only increases when we learn that a fraction of any emitted ton of carbon stays up in the atmosphere for at least thousands of years. A common argument governing climate change action is: if we delay mitigation today, we have to do even more tomorrow. The model tells us: if we delay policy today, we have to live with the consequences, but we do not have to compensate in our future mitigation effort.

The idea of mitigating more at higher levels of atmospheric CO_2 is based on the convexity of damages in global temperature increase. Figure 2 shows that GAUVAL has such a convex damage function, yet, optimal mitigation does not increase in the prevailing CO_2 concentration. The reason lies in the radiative forcing equation (4): the higher the CO_2 concentration, the less does an additional ton of emissions contribute to further forcing and, thus, warming. The precise physics of the underlying logarithm is slightly more complicated, but a simple intuition is as follows. CO_2 traps (absorbs) a certain spectrum of the wavelength that our planets radiates out into space, thereby warming the planet. If there is already a high concentration of CO_2 in the atmosphere, most of the energy leaving the planet in this wavelength is already trapped. As a consequence, an additional unit of CO_2 emissions has a much lower warming impact than the first unit of anthropogenic emissions. GAUVAL models explicitly the implicit assumptions of Golosov et al. (2014) and Gerlagh & Liski (2012) that the convexity of the damage curve and the concavity of the radiative forcing equation "cancel" each other. In contrast to the earlier papers, GAUVAL directly employs the carbon cycle of one of the most wide-spread integrated assessment models, explicitly uses the physical radiative forcing equation, and matches the forcing induced temperature dynamics better than most integrated assessment models. I conclude that the finding might be uncomfortable, but not unreasonable.

In addition, the optimal mitigation policy does not depend on the prevailing temperature level, despite higher temperatures causing higher marginal damages. The reason is that the long-term equilibrium temperature is determined entirely by the GHG concentrations, and a higher temperature level at a given CO_2 concentration implies less warming in the future. GAUVAL shows that this finding prevails in a model that nicely replicates the temperature dynamics of state of the art climate models (Figure 3). These findings connects immediately to the debate on the slope of the marginal damage curve in the "taxes versus quantities" literature (Weitzman 1974, Hoel & Karp 2002, Newell & Pizer 2003). GAUVAL states that the social damage curve for CO_2 emissions is flat. In consequence, taxes not only minimize the welfare cost under technological uncertainty and asymmetric information as compared to a cap and trade system, but they even eliminate these welfare costs. The marginal damage curve would gain a non-trivial slope if the model was to depart from the

assumption of an intertemporal elasticity of substitution of unity. Deterministic estimates usually suggest values smaller than unity. However, the long-run risk literature forcefully argues for an intertemporal elasticity of substitution larger than unity (and disentangled from risk attitude). The logarithmic middle ground stays reasonable. In particular, it is just as easy to argue for a slightly falling marginal damage curve as it is to argue for a slightly increasing marginal damage curve.[11]

3.3 Discounting and Mass Conservation

Optimal economic policy implies that we have to live with the consequences of historic overindulgence in carbon because our mitigation effort is independent of past emissions. What makes it worse: carbon does not decay. Carbon only cycles through the different reservoirs; the fact that some of it eventually turns into limestone is negligible for human planning horizons. A model comparison of scientific carbon cycle models found that on average 18% of a 100Gt emission pulse, approximately 10 years of present emissions, still remain in the atmosphere after 3000 years (Joos et al. 2013). In DICE 2013's carbon cycle adopted here, 6% of an anthropogenic emission unit stays in the atmosphere forever.[12]

This implication of mass conservation of carbon has an immediate and important impact on the optimal carbon tax.

Proposition 2 *A carbon cycle (equation 3) satisfying mass conservation of carbon implies a factor* $(1-\beta)^{-1}$*, approximately proportional to* $\frac{1}{\rho}$*, in the closed-form solution of the SCC (equation 13).*

In particular, the SCC approaches infinity as the rate of pure time preference approaches zero.[13] The proof runs as follows. Mass conservation of carbon implies that the columns of Φ add to unity. In consequence, the vector with unit entry in all dimensions is a left and, thus, right eigenvector. The corresponding eigenvalue is one and the determinant of $\mathbf{1}-\beta\mathbf{\Phi}$ has the root $1-\beta$. It follows from Cramer's rule (or as an application of the Cayley-Hamilton theorem) that the entries of the matrix $(\mathbf{1}-\beta\mathbf{\Phi})^{-1}$ are proportional to $(1-\beta)^{-1}$.

I briefly point out how the result changes if I had not normalized population to unity. I assume that the social welfare function is population weighted per capita consumption and that population grows at the factor $G = \exp(g)$. Then, the root and the factor in equation

[11]The marginal damage curve would also be negatively sloped if the damage function was less convex resembling more closely that of the DICE model. The intuition is that the logarithm in the radiative forcing equation is very strong, and that the underlying saturation in the CO_2 absorption spectrum can outweigh the damage convexity.

[12]The maximal eigenvalue of the transition matrix Φ is unity. The corresponding eigenvector governs the long-run distribution as the transitions corresponding to all other eigenvectors are damped. I obtain the 0.06 as the first entry of the corresponding eigenvector.

[13]The present objective function and the dynamic programming equation are not well-defined in the limit of a zero rate of pure time preference. However, the statement holds in that for any $n \in \mathbb{N}$ there exists a strictly positive pure rate of time preference ρ such that $SCC(\rho) > N$.

(13) change to $(1-\beta G)^{-1} \approx \frac{1}{\rho - g}$. The SCC becomes even more sensitive to the rate of pure time preference.[14] Note that, in contrast to the SCC, the temperature's shadow value does not have a root $(1-\beta)^{-1}$. The matrix $\boldsymbol{\sigma}$ does not have a unit eigenvalue because the planet exchanges heat with outer space.[15] The long-run temperature responds to changes of radiative forcing without hysteresis, i.e., without path dependence. Appendix D.1 illustrates Proposition 2 for a two-layer carbon cycle, and the absence of such sensitivity for a two-layer atmosphere-ocean temperature system. It also shows how a frequently used decay approximation of the carbon cycle misses the sensitivity to pure time preference.

It is well-known that the consumption discount rate plays a crucial role in valuing long-run impacts. The present finding is different. GAUVAL's damages are proportional to economic output, and the economic impact of climate change grows with output. Given the logarithmic utility specification, the usual consumption discount rate argument does not apply: economic growth does not effect the SCC. Yet, the SCC is extremely sensitive to the rate of pure time preference. It is a widely held believe in the integrated assessment community that it is of little importance how we calibrate the constituents of the consumption discount rate, as long as pure time preference and the growth-based component add up to the same overall consumption discount rate (Nordhaus 2007). The present finding fleshes out the shortcoming of this consumption discount rate based reasoning.

Illustrating the SCC's sensitivity to pure time preference, I reduce the investment-rate-implied annual rate $\rho = 1.75\%$ to a value of $\rho = 1\%$. The SCC increases to 93 USD per ton C or 25.5 USD per ton CO_2. Further reducing the rate of pure time preference to the value of $\rho = 0.1\%$ employed in the Stern (2008) Review results in an optimal carbon tax of 660 USD per ton C and 180 USD per ton CO_2. The Stern Review justified its low pick of the rate of pure time preference by normative reasoning, market failure, and a dual role of individuals who might behave differently on the market as compared to large-picture policy decisions (Hepburn 2006).

Schneider et al. (2013) show in a continuous time overlapping generations model how the common infinitely-lived-agent based calibration of IAMs overestimates time preference under limited altruism. In addition, the present model, like other IAMs, does not explicitly model the actual portfolio of risky investments and, yet, calibrates to overall investment and the Ramsey equation. In an asset pricing context, Bansal et al. (2012) calibrate the pure rate of time preference to $\rho = 0.11\%$ carefully disentangling risk attitude and risk premia from consumption smoothing and the risk-free discount rate. Their model explains observed asset prices significantly better than any asset pricing approach based on the standard economic

[14] The intuition is that population weighted per-capita consumption puts additional weight on future generations that are more numerous, acting as a reduction of time preference. As is well-known, the value function no longer converges as $\rho \to g$.

[15] Temperature is an intensive quantity and a conservation of "heat" would imply that the rows of the matrix $\boldsymbol{\sigma}$ added to unity. However, the atmospheric layer constantly exchanges heat with outer space. Formally, the subtraction of σ^{forc} which implies that the first row of the matrix $\boldsymbol{\sigma}$ does not add to unity, implying that the largest eigenvalue of the matrix is smaller than unity and historic influences are damped.

model with higher time preference. Traeger (2012*a*) shows how uncertainty-based discounting of an agent whose risk aversion does not coincide with her consumption smoothing preference (falsely) manifests as pure time preference in the economic standard model, and he discusses some implications for climate change evaluation. Disentangling the different contributions to the SCC for the $\rho = 0.1\%$ case delivers

$$SCC_t = \underbrace{\frac{\beta Y_t}{M_{pre}}}_{\cancel{26}\frac{\$}{tC}\,30\frac{\$}{tC}} \underbrace{\xi_0}_{2.2\%} \underbrace{\left[(\mathbf{1}-\beta\boldsymbol{\sigma})^{-1}\right]_{1,1}}_{\cancel{1.4}\,2} \underbrace{\sigma^{forc}}_{0.42} \underbrace{\left[(\mathbf{1}-\beta\boldsymbol{\Phi})^{-1}\right]_{1,1}}_{\cancel{3.7}\,26} = \cancel{\mathbf{57}}\mathbf{660}\ \frac{\$}{tC}\ .$$

The large part of the increase, a factor 7, arises from the carbon cycle's contribution.

4 Carbon Cycle Uncertainty, Learning, and Welfare

This section analyzes the welfare implications of uncertainty governing the carbon cycle. I present simple formulas quantifying the welfare impact and the value of uncertainty reduction. First, I model uncertainty as a sequence of persistent shocks drawn from an arbitrary distribution (vector autoregressive model). Second, I present a Bayesian model incorporating the anticipation of learning.

4.1 Carbon Sink Uncertainty

Over 10% of the annual flow of anthropogenic carbon emissions leave the atmosphere into an unidentified sink. These missing 1Gt+ in the carbon budget are over twice the weight of all humans walking the planet. Current research is not conclusive, but a likely candidate for at least part of the "missing sink" are boreal or tropical forests. The limitations in understanding the carbon flows and whether the uptake of the missing sink is permanent or temporary implies major uncertainties in predicting future carbon dynamics. The scientific community and governments invest significant sums into the reduction of these uncertainties, including the launching of satellites and new supercomputing facilities. GAUVAL can produce a simple estimate of the welfare costs of these uncertainties and serve as a formal model for quantifying the benefits of uncertainty reduction.

A useful approximation of the carbon cycle uncertainty adds a stochastic flow between the atmosphere and other layers. I modify the carbon cycle's equation of motion (3) to the form

$$\boldsymbol{M}_{t+1} = \boldsymbol{\Phi}\boldsymbol{M}_t + (\epsilon_t, -\epsilon_t, 0, ..., 0)^\top + \boldsymbol{e}_1(\textstyle\sum_{i=1}^{I^d} E_{i,t} + E_t^{exo})\ , \tag{15}$$

where ϵ_t, $t \in \{0, ..., \infty\}$ is a sequence of random variables and ϵ_0 has a zero mean. Using DICE 2013's three-layer carbon cycle, the random variable ϵ_t captures the uncertainty in

the carbon flow between the atmosphere and the joint upper ocean and biosphere carbon reservoir.

I suggest two conceptually different ways to think about the uncertainty. In the first interpretation, we are merely worried about changes in the carbon flows over time that cannot be predicted with certainty. A small persistent shock to ϵ moves the carbon flow, either increasing or decreasing the sink uptake. Over time, these shocks accumulate and so does the uncertainty in forecasting future carbon levels, temperatures, and economic damages. In the second interpretation, the physical carbon flows are merely unknown to the scientists, but over time scientists expect to learn the carbon flows, making the system more predictable. I will treat these two cases in the subsequent two sections and compare their economic implications.

The underlying model comparison study is subject to common bias and short-comings, and this standard deviation is best interpreted as a lower bound of actual uncertainty.

4.2 Vector Autoregressive (VAR) Uncertainty

I assume a VAR(1) process to model persistent shocks that change the carbon flows in an unpredicted way

$$\epsilon_{t+1} = \gamma\epsilon_t + \chi_t, \tag{16}$$

where $\gamma \leq 1$ and the sequence $\chi_{t,t\in\{0,\ldots,\infty\}}$ is independently distributed (and $\epsilon_0 = 0$). I calibrate persistence based on the model comparison study by Joos et al. (2013) to $\gamma = 0.997$. The new shadow value φ_ϵ of the persistent random variable in the carbon cycle is

$$\varphi_\epsilon = \frac{\beta}{1-\gamma\beta}[\varphi_{M_1} - \varphi_{M_2}] \ . \tag{17}$$

The higher the persistence γ, and the larger the shadow value difference between carbon in the two reservoirs, the larger is the welfare cost of a carbon cycle shock ϵ_t. Appendix E.1 derives the general welfare difference between the deterministic and the uncertain scenario. Here, I focus on a sequence of identically distributed shocks χ_t implying the welfare cost of uncertainty

$$\Delta W^{VAR,iid} = \frac{\beta}{\alpha(1-\beta)} \log\left[\mathbb{E}\exp\left[\alpha\varphi_\epsilon\chi\right]\right] = \frac{\beta}{\alpha(1-\beta)} \sum_{i=1}^{\infty} \kappa_i \frac{(\alpha\varphi_\epsilon)^i}{i!}. \tag{18}$$

The first expression for the welfare loss involves the cumulant generating function (cgf) $G_\chi(z) = \log\left[\mathbb{E}\exp(z\chi)\right]$ of the random variable χ. The cgf is the logarithm of the moment generating function. The welfare loss from carbon cycle uncertainty is the shock's cg f evaluated at the product of the persistent flow's shadow value φ_ϵ and the risk aversion parameter α. The factor $\frac{1}{1-\beta}$ reflects the infinite sum over future shocks, and the factor β reflects the one period delay between the shock and its welfare impact.

The second expression for the welfare loss expands the function in terms of the random variable's cumulants $\kappa_i, i \in \mathbb{N}$. The first cumulant is the expected value of χ. In the present case where $\kappa_1 = \mathbb{E}\,\chi = 0$, the second cumulant is the shock's variance and the third cumulant is the third central moment specifying skewness. The higher order cumulants do not coincide exactly with the central moments, but relate closely. An advantage of cumulants as compared to central moments is that they are additive for independent random variables and cumulant i is homogenous of degree i under scalar multiplication (cumulant i of $\lambda\chi$ is $\lambda^i\kappa_i$).

The welfare loss is the sum of the stochastic shock's cumulants, each weighted with the flow's shadow value taken to the power of the cumulant's order. The expected value is valued independently of risk aversion, the variance proportional to risk aversion, and skewness proportional to risk aversion squared. This basic structure for evaluating the welfare loss of different uncertainties will reappear for all uncertainties. For most distributions, the cumulant (or the moment) generating functions are tabled and the closed-form solution for the welfare loss follows directly. Alternatively, section 5.2 embraces numeric scientific estimates of temperature uncertainty and employs the cumulant expansion to evaluate the welfare loss.

In the case of a mean-zero normally distributed iid shock $\chi \sim N(0, \sigma_\chi^2)$ only the second cumulant $\kappa_2 = \sigma_\chi^2$ differs from zero, and the welfare impact is

$$\Delta W^{VAR,normal} = \frac{\alpha\beta}{1-\beta}\,\varphi_\epsilon^2\,\frac{\sigma_\chi^2}{2} = \frac{\alpha\beta}{1-\beta}\Big(\frac{\beta}{1-\gamma\beta}\Big)^2\,(\varphi_{M_1} - \varphi_{M_2})^2\,\frac{\sigma_\chi^2}{2}\ . \tag{19}$$

This welfare loss is proportional to the square of the shadow value φ_ϵ and, thus, proportional to the squared difference between the shadow values of carbon in the different reservoirs and proportional to the squared inverse of $1 - \gamma\beta$. This latter term makes the welfare loss from an individual shock approach infinity as both persistence and discount factor approach unity. For high persistence, this term makes the welfare loss highly sensitive to pure time preference. For an independently distributed shock, where $\gamma = 0$, the welfare loss is not sensitive to time preference. Recall that the disentangled Arrow-Pratt risk aversion parameter is $\frac{\alpha}{1-\beta}$. Hence, this term does not change with time preference when keeping Arrow-Pratt risk aversion constant.

4.3 Bayesian Uncertainty and Anticipated Learning

This section replaces the simple VAR(1) uncertainty by a Bayesian model, capturing the decision maker's subjective uncertainty as well as her anticipation of learning over time. Closely related Bayesian learning models have first been used in integrated assessment of climate change by Kelly & Kolstad (1999) in a numeric application to climate sensitivity uncertainty and by Karp & Zhang (2006) in a stylized semi-analytic application to damage uncertainty. The uncertain carbon flow ϵ_t is now governed by a subjective prior, which I

assume to be normally distributed with unknown mean but known variance

$$\epsilon_t \sim N(\mu_{\epsilon,t}, \sigma^2_{\epsilon,t}) \ , \ \mu_{\epsilon,0} = 0.$$

In addition, the equations of motion are subject to an objective stochastic shock $\nu_t \sim N(0, \sigma^2_{\nu,t})$, which can also be interpreted as measurement error. This stochasticity prevents the decision maker from learning the prior's mean from a single observation. The new equation of motion for the atmospheric and the biosphere-and-upper-ocean carbon reservoirs take the form

$$M_{1,t+1} = (\mathbf{\Phi} \boldsymbol{M}_t)_1 + \textstyle\sum_{i=1}^{I^d} E_{i,t} + E_t^{exo} + \epsilon_t + \nu_t, \tag{20}$$

$$M_{2,t+1} = (\mathbf{\Phi} \boldsymbol{M}_t)_2 - \epsilon_t - \nu_t. \tag{21}$$

I will model the learning process based on atmospheric carbon observation.[16] Rearranging equation (20), the decision maker derives information on ϵ_t from the realizations

$$\hat{\epsilon}_t = M_{1,t+1} - (\mathbf{\Phi} \boldsymbol{M}_t)_1 - \textstyle\sum_{i=1}^{I^d} E_{i,t} - E_t^{exo} - \nu_t \ .$$

Recently, three satellites were launched to reduce the carbon flow measurement errors ν_t, one of whom dropped straight into the Artic sea. But learning is not limited to future observation. Given the availability of historic data, learning also takes place through the advances in fundamental scientific knowledge and supercomputing. Thus, I interpret ν_t merely as a parameter determining the speed of learning.

The equations of motion for the Bayesian prior's mean and variance are

$$\mu_{\epsilon,t+1} = \frac{\sigma^2_{\epsilon,t}\hat{\epsilon}_t + \sigma^2_{\nu,t}\mu_{\epsilon,t}}{\sigma^2_{\epsilon,t} + \sigma^2_{\nu,t}} \quad \text{and} \quad \sigma^2_{\epsilon,t+1} = \frac{\sigma^2_{\nu,t}\sigma^2_{\epsilon,t}}{\sigma^2_{\nu,t} + \sigma^2_{\epsilon,t}} \ . \tag{22}$$

This standard Bayesian updating equation characterizes the posterier mean as a weighted average of the new observation and its prior mean. The weight of the new observation is inversely proportional to the variance of the measurement error (or proportional to its precision). The weight on the prior's mean is inversely proportional to its variance. The variance of the carbon cycle uncertainty in this Bayesian learning model falls exogenously over time. The smaller the ratio of stochasticity to overall uncertainty $\frac{\sigma^2_{\nu,t}}{\sigma^2_{\nu,t}+\sigma^2_{\epsilon,t}}$, the faster the learning.

The new shadow value φ_μ of a mean carbon transfer shift is

$$\varphi_\mu = \frac{\beta}{1-\beta}[\varphi_{M_1} - \varphi_{M_2}] \ . \tag{23}$$

[16] In principle, the decision-maker could simultaneously learn from observing the carbon concentration in the combined biosphere and upper ocean reservoir. However, as I explained earlier, measurement errors in the non-atmospheric carbon reservoir are so much larger that an observation-based learning model can comfortably ignore these additional measurements.

This shadow value coincides with the shadow value of a perfectly persistent shock (equation 17). The welfare cost of epistemological uncertainty in the Bayesian setting is (see Appendix E.1)

$$\Delta W^{Bayes} = \sum_{t=0}^{\infty} \beta^{t+1} \frac{\sigma^2_{\epsilon,t}+\sigma^2_{\nu,t}}{2} \alpha \left(\varphi_{M_1} - \varphi_{M_2}\right)^2 \left(\frac{1}{1-\beta}\right)^2 \underbrace{\left(\frac{\sigma^2_{\epsilon,t}}{\sigma^2_{\nu,t}+\sigma^2_{\epsilon,t}} + (1-\beta)\frac{\sigma^2_{\nu,t}}{\sigma^2_{\nu,t}+\sigma^2_{\epsilon,t}}\right)^2}_{\equiv \Omega_t}. \quad (24)$$

As in the case of a normally distributed persistent shock, the welfare loss is proportional to the squared difference of the shadow values of carbon in the atmosphere and the biosphere-ocean sink. However, the loss is now composed of two distinct contributions. The first arises from the epistemological uncertainty prevailing already in the current period (equations of motion 20 and 21). The second arises from the "updating shock" to the prior's mean in equation (22), which has a similar effect as the persistent shock to ϵ_t in the earlier setting, and whose effect is proportional to φ_μ. These two contributions combine in the weighted mean Ω_t (see Appendix E.1). The relevant variance in the case of learning is the prior's variance plus the variance of the periodic shock $(\sigma^2_{\epsilon,t} + \sigma^2_{\nu,t})$. The prior's variance is declining as the decision maker improves her estimate of the true carbon flows.

Ignoring the time dependence of the variances for a moment, the infinite sum becomes $\frac{\alpha\beta}{1-\beta}(\frac{1}{1-\beta})^2 \left(\varphi_{M_1} - \varphi_{M_2}\right)^2 \frac{\sigma^2_{\nu,t}+\sigma^2_{\epsilon,t}}{2} \Omega_t$, closely resembling equation (19) for the persistent shock. The two differences are, first, the absence of the persistence parameter γ in the denominator and of the discount factor β in the numerator of the factor $(\frac{1\cancel{\beta}}{1-\cancel{\gamma}\beta})^2$ and, second, the new term Ω_t. The first change evokes the impression that the case of Bayesian learning corresponds to a scenario of fully persistent shocks ($\gamma = 1$). Indeed, this difference amplifies the early contributions of the infinite sum in equation (24) and makes them even more sensitive to pure time preference (approximately proportional to $\frac{1}{\rho^2}$). In addition, epistemological uncertainty affects carbon flows without delay eliminating the discount factor in the numerator. However, in the long-run, the second change introduced by Ω_t more than offsets this apparent persistence. The term is a weighted mean of unity and $1 - \beta$. Initially, the variance of the prior $\sigma^2_{\epsilon,t}$ dominates, and $\Omega \approx 1$. Over time, the decision maker learns the subjectively uncertain part of the carbon flow uncertainty. As $\sigma^2_{\epsilon,t}$ falls, the weight on the term $(1 - \beta)$ increases and $\Omega \to (1 - \beta)^2$, offsetting the term $(\frac{1}{1-\beta})^2$. As the decision maker learns the subjective distribution of ϵ_t, the remaining uncertainty of the shock $\sigma^2_{\nu,i}$ is independently distributed. The welfare loss of independent shocks is small and not sensitive to pure time preference.

4.4 Quantifying the Welfare Impact of Carbon Cycle Uncertainty

This section derives a quantitative estimate of the welfare loss from carbon cycle uncertainty and the willingness to pay for uncertainty reduction. The uncertainty calibration is based on a business as usual scenario. Joos et al. (2013) subject 18 different carbon cycle models to a

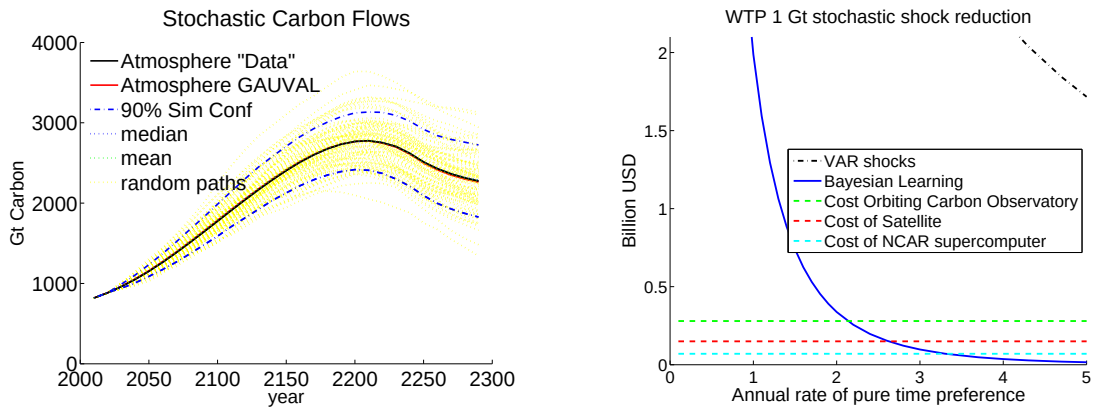

Figure 4: The graph on the left shows the evolution of atmospheric carbon for the DICE 2013 business as usual emission scenario. Decadal shocks with standard deviation of $\sigma_\chi = 20$ Gt per decade change the flow between the atmosphere and the carbon sinks with a persistence of $\gamma_M = 0.997$ calibrated to carbon cycle comparison study be Joos et al.'s (2013). The deterministic DICE evolution (5 year time steps, "Data"), the deterministic GAUVAL evolution (10 year time steps), and the mean and the median of 1000 uncertain trajectories are hardly distinguishable. The right graph depicts the willingness to pay for a 1Gt uncertainty reduction. In the Bayesian learning case, the reduction is in the measurement error, increasing the speed of learning. In the case of the vector autoregressive shock model ("VAR"), the willingness to pay is based on a physical reduction of carbon flow stochasticity, e.g., as a co-benefit of emission reductions.

100Gt and 5000Gt carbon pulse and track their responses for up to 3000 years. The larger shock corresponds to 500 years of present day emissions and it implies a medium run (100-1000 years) cross-model standard deviation of atmospheric carbon of approximately 500Gt. The underlying model comparison study is subject to common bias and short-comings, and this standard deviation is best interpreted as a lower bound of actual uncertainty. In the VAR model, Joos et al.'s (2013) simulations suggest a persistence in equation (19) of $\gamma = 0.997$, i.e., highly intertemporally correlated uncertainty governing the flow of carbon between the atmosphere and the sinks. All carbon cycle uncertainty scenarios employ a risk aversion coefficient of $\alpha = -1.5$ (see Appendix C).

For the auto-regressive shock scenario of section 4.2, I assume a standard deviation of the decadal shock of $\sigma_\chi = 20$ Gt per decade, which builds up into the 500 Gt standard deviation suggested by Joos et al. (2013) around the year 2300. Figure 4, left panel, evaluates the resulting carbon concentration along the business as usual scenario of DICE 2013. Appendix E.2 discusses variations. This uncertainty translates into a welfare loss of 110 billion USD. For the case of epistemological uncertainty, I assume that the Bayesian prior has a standard deviation of $\sigma_{\epsilon,0} = 20$Gt (per decade), and that the "measurement error" is $\sigma_\nu = 10$Gt (corresponding approximately to the missing carbon). This combination of prior and measurement error implies a remaining epistemological uncertainty of 4.4Gt after 50 years and of 2.6Gt per decade after 150 years. The resulting welfare loss from carbon cycle uncertainty in this Bayesian learning scenario is approximately 29 billion USD. It is most sensitive to the initial prior. Lowering initial uncertainty to $\sigma_{\epsilon,0} = 10$Gt lowers the welfare

loss to 9 billion USD. To obtain a welfare loss of 110 billion USD as in the case of the VAR shocks, I would have to raise initial uncertainty of the prior to a 40Gt standard deviation. For a comparison of magnitude, the annual NASA budget is about 20 billion USD.

In the case of an annual rate of pure time preference of $\rho = 0.1$ the uncertainty contributions increase to 2 trillion USD in the case of VAR shocks and to 60 trillion USD, or 73% of world output, in the case of Bayesian learning. The significantly higher sensitivity to pure time preference in the case of Bayesian learning reflects the permanent impact that learning has in the initial periods and which acts like a fully persistent VAR shock (see section 4.3). It is fair to conclude that the absolute welfare costs from uncertainty over the carbon flows are small to moderate as compared to the deterministic contributions discussed in section 3. The reason is the increasing satiation of the wave spectrum in CO_2's absorption spectrum which implies a logarithmic relation between CO_2 concentrations and warming. This logarithm eliminates the convexity arising from damages and logarithmic utility. The remaining welfare costs are born directly from risk aversion.

Figure 4, right panel, states the willingness to pay for a 1Gt reduction of the decadal standard deviation as a function of pure time preference. In the case of the VAR model, the uncertainty reduction lowers the physical stochasticity of the carbon flows (σ_χ = 10Gt→ 9Gt).[17] In the case of the Bayesian model, the uncertainty reduction lowers the measurement error and increases the speed of learning (σ_ν = 10Gt→ 9Gt). The figure compares the welfare gain from better measurement and faster learning to the costs of a common satellite (∼150 million USD), NASA's Orbiting Carbon Observatory (∼ 280 million), and the National Center for Atmospheric Research's recent supercomputer (∼ 70 million). For the standard calibration of the time preference these (American) investments are worth the (global) welfare gain. For an annual rate of time preference around 3% even the global welfare gain might no longer outweigh their costs in GAUVAL.[18] In contrast, following a normatively or long-run risk founded rate around $\rho = 0.1\%$ increases the willingness to pay for a decadal 1Gt stochasticity reduction to approximately 110 billion USD, the same amount as the complete risk elimination in the VAR shock scenario for the standard time preference calibration.

5 Temperature Uncertainty

So far, I have assumed that a doubling of the CO_2 concentrations from its pre-industrial concentration of 280 ppm to 560 ppm yields a long-term temperature increase of 3C. At present, CO_2 levels are up to almost 400 ppm. Including the CO_2 equivalent forcing of

[17]Reducing the carbon flow's decadal standard deviation σ_χ by 1 Gt reduces the welfare loss by the fraction $\frac{2}{\sigma_\chi} + \frac{1}{\sigma_\chi^2}$. The formula solves $x = \frac{\Delta W_{\sigma_\chi}^{VAR,normal} - \Delta W_{\sigma_\chi - 1}^{VAR,normal}}{\Delta W_{\sigma_\chi}^{VAR,normal}}$ for x. The graph is only visible in the upper right corner: the payoff of the physical stochastic shock reduction is much more valuable than a reduction of measurement error that accelerates learning.

[18]NASA's Orbiting Carbon Observatory is the investment closest to a direct reduction of measurement error to improve learning. Slowly coming out of its calibration phase, the ultimate precision is still unclear.

other GHGs the level is already close to 480 ppm and 70% of the way toward doubling pre-industrial concentrations. The implied 3C equilibrium warming is little more than a best guess. The value depends on a set of uncertain feedback processes that either increase or decrease the initial warming. For example, higher temperatures imply more evaporation, and water vapor itself is a powerful GHG. The value of 3C was cited as the best guess in the first four IPCC assessment reports. The latest report deleted this best guess and only cites a likely range of 1.5-4.5C (IPCC 2013).

The present section discusses the basics of temperature uncertainty. I explain why the normal-normal learning model would not be reasonable in the case of temperature uncertainty and construct and alternative model that comprises epistemological uncertainty as well as persistent shocks. I quantify a lower bound for the welfare loss from uncertainty over the temperature response to a given CO_2 concentration.

5.1 On Weighty Tails

Meinshausen et al. (2009) collect 20 estimates of the probability distributions governing the temperature increase from a doubling of the CO_2 concentration, the so-called climate sensitivity. These estimates derive from different groups and use a variety of methodological approaches. My evaluation of temperature uncertainty will rely on the average distribution assigning equal weight to each approach. I refer to Appendix F.2 and, in particular, Figure 7 for details. The support of these distributions spans from no warming to a 10C warming. A warming above 10C is possibly but the probabilistic estimates are conditional on not exceeding 10C for a doubling of pre-industrial CO_2 concentrations.

The climate sensitivity distribution governing global warming is positively skewed, i.e., it exhibits a pronounced right tail. In addition, the equations of motion are exponential in temperature (linear in $\tau_1 = \exp(\xi_1 T_1)$). Thus, even a normal distribution of temperature can translate into a log-normal distribution in the linear equations of motion. By section 4.2, the resulting welfare loss is proportional to the cumulant generating function. The cumulant generating function of the log-normal distribution is infinite. Hence, I can easily set up a model that delivers an infinite welfare loss from climate sensitivity uncertainty. This result takes Weitzman's (2009) "dismal theorem" and Millner's (2013) extension from their sylized frameworks into a full-fledged and well-calibrated integrated assessment model. Here, even the thin-tailed normal distribution as opposed to the fat-tailed prior in Weitzman (2009) can blow up the welfare loss through its translation into economic damages.[19] The present result is particularly interesting in the context of Kelly & Tan (2013), who show that fat tails should also imply a relatively quick learning with major updates that have not been

[19]Figure 2 shows that for temperature increases up to 12C GAUVAL's base case damage specification delivers damages lower than DICE. More than that, the "dismal result" holds for any $\xi_0 > 0$, implying that I can make damages at any given temperature level arbitrarily small and still find an infinite welfare loss from temperature uncertainty.

observed in the past.

The "dismal results" rely on the extrapolation of different functional forms that nicely capture our best guess of what happens at moderate degrees of warming. However, no economic model to date has incorporated a reasonable quantitative characterization of what happens to life and welfare on planet Earth for a 20C or 30C global temperature increase. Should we concluded that we commit all of our economic resources to fighting global warming? No.[20] It is already daring to evaluate damages from a warming of up to perhaps 10-20C with state of the art integrated assessment models. To evaluate even higher temperature and damage scenarios, we should *not* rely on the extrapolation of functional forms, but build a separate model aiming directly at the quantification of scenarios that lie far from experience and even imagination. What either version of the "dismal theorem" can do is to raise awareness about the importance of integrating uncertainty into climate change assessments. What GAUVAL can do is estimate a lower bound of the welfare loss from uncertainty given a warming range for which damages and temperature dynamics seem somewhat reasonable.

5.2 A VAR Model with Epistemological Uncertainty

A model of temperature uncertainty has to incorporate an epistemological uncertainty component. Epistemological uncertainty is characterized by endogenous informational states. Usually, the informational states are a small set of parameters that uniquely pin down the distributions of a given class. In the present context of a linear-in-state system, the equations of motion have to be linear in (suitable transformations of) the endogenous informational states. Analytically tractable Bayesian learning models rely on distributional assumptions that ensure that the posterior lies in the same distributional class as the prior (conjugate prior model). The only Bayesian (conjugate prior) model compatible with the linear in state model seems to be the normal-normal learning model that I employed in section 4.3.

For temperature uncertainty, I cannot reasonably assume a normal distribution.[21] In consequence, I introduce a combined epistemological and vector autoregressive shock model of uncertainty that disposes of the straight-jacket imposed by a normal distribution at the expense of no longer being strictly Bayesian. Instead of relying on a particular class of probability distributions whose (endogenous) parameters enter the equations of motion linearly, I characterize an arbitrary probability distribution through its (countably infinite) sequence

[20] Also technically, the result hinges on a crucial assumption that makes it inadequate for the evaluation of high temperature tails. Representations of rational preferences by expected utility, including the present representation, require bounded utility functions or alternative assumptions with similar consequences (von Neumann & Morgenstern 1944, Kreps 1988). The "dismal result" depends crucially on *not* meeting this rationality assumption when the damage distribution approaches full output loss with high density.

[21] A normal on temperature in degree Celsius would ignore positive skewness and, yet, the tails far outside of what I can reasonably calibrate would drive the results. See discussion in previous section for details. A normal on transformed temperature τ_1 would assign significant weight to large negative values of the actual atmospheric temperature increase T_1.

of cumulants. This approach allows me later to employ the particular numeric distribution derived in the scientific literature for a quantitative estimate of the welfare impact.

I represent uncertainty governing the temperature's equation of motion as a random contribution ϵ_t^τ to incremental warming, changing equation (7) to the form

$$\boldsymbol{\tau}_{t+1} = \boldsymbol{\sigma}\boldsymbol{\tau}_t + \sigma^{forc}\frac{M_{1,t} + G_t}{M_{pre}}\boldsymbol{e}_1 + \epsilon_t^\tau \boldsymbol{e}_1 \ . \tag{25}$$

The random variable ϵ_t^τ captures both epistemological uncertainty as well as stochastic changes. I characterize a general distribution of ϵ_t^τ by its cumulants $\kappa_{i,t}, i \in \mathbb{N}$. Initial epistemological uncertainty is given by $\kappa_{i,0}, i \in \mathbb{N}$, and the cumulants follow the equations of motion

$$\kappa_{i,t+1} = \gamma^i \kappa_{i,t} + \chi_{i,t}^\tau \ , \tag{26}$$

$0 \leq \gamma \leq 1$, for all $i \in \mathbb{N}$. The special case where $\kappa_i = 0$ for $i > 2$ relates closely to the Bayesian learning model in section 4.3.[22] In the absence of shocks ($\chi_{i,t} = 0$), epistemological uncertainty decays at rate γ: ϵ_{t+1} is then distributed as $\gamma\epsilon_t$. I assume that the stochastic shocks $\chi_{i,t}$ are independently and identically distributed. They introduce new uncertainty in every period, and they delay or even undo the decision maker's learning process. In the normally distributed Bayesian learning model, the uncertainty over the mean falls quickly. In the VAR model, the shocks remain constant over time and build up slowly into forecast uncertainty. The present model combines these features. In addition, the model in equation (26) permits shocks to the variance $\chi_{2,t}^\tau$, making the speed of learning itself uncertain (a feature also present in "non-normal" Bayesian learning models). In the context without learning, such non-zero shocks $\chi_{2,t}^\tau$ are considered in stochastic volatility models.[23]

Appendix E.1 derives the resulting shadow value of cumulant κ_i to

$$\varphi_{\kappa,i} = \frac{\beta}{1-\beta\gamma^i}\frac{(\alpha\varphi_{\tau,1})^i}{i!\ \alpha} \ . \tag{27}$$

The shadow value $\varphi_{\kappa,i}$ of a mean shift in (transformed) temperature flow is similar to those in the case of carbon cycle uncertainty (equations 17 and 23). The factor $\frac{1}{1-\beta\gamma^i}$ reflects the permanent impact of a shift in the mean and, here, the shadow value of a (transformed) temperature increase replaces the difference of the carbon price between the atmosphere and the sinks in the previous sections. For the variance and higher order cumulants, risk aversion

[22] The Bayesian model would exhibit a time dependent parameter $\gamma_{i,t}$ in the equation of motion $\kappa_{i,t+1} = \gamma_{i,t}\kappa_{i,t} + \chi_{i,t}^\tau$ and require an additional non-persistent iid shock in equation (25) (which has little impact on the welfare evaluation). The parameter $\gamma_{2,t} = \sigma_{\epsilon,t}^2$ can be chosen to reproduce the learning of the Bayesian model (plus stochasticity ν_t), and the shocks $\chi_{1,t}^\tau$ would have to match the Bayesian updates of the mean.

[23] Not all combinations of shocks to the cumulants will lead to well-defined probability distributions of ϵ_t. However, in such a case, the cumulants can still be considered a reasonable approximation of a closely related distribution.

and higher orders of the temperature's shadow value come into play as well. The resulting welfare loss from temperature uncertainty is composed of two parts

$$\Delta W^{temp} = \underbrace{\sum_{i=1}^{\infty} \varphi_{\kappa,i}\kappa_{i,0}}_{\text{Epistemological}} + \underbrace{\frac{\beta}{\alpha(1-\beta)} \sum_{i=1}^{\infty} \log\left[\mathbb{E}_t \exp\left[\alpha\varphi_{\kappa,i}\chi_i^{\tau}\right]\right]}_{\text{Future Shocks}} \tag{28}$$

First, initial epistemological uncertainty translates into the shadow-value-weighted sum of the cumulants. The welfare loss implied by any non-zero cumulant (variance, skewness,...) is proportional to its shadow value discussed above. Second, the future VAR shocks $\chi_{i,t}$ to the cumulants result in a welfare loss that is evaluated by the moment generating function of the shock and the corresponding cumulant's shadow value. In particular, shocks to the first cumulant, the mean, imply an identical formula for the welfare loss as the VAR shocks to the carbon flow in equation (18). Shocks to the higher order cumulants (volatility, skewness,...) depend on higher orders of risk aversion and the temperature's shadow value through the shadow values of the corresponding cumulants.

In the case of perfectly persistent epistemological uncertainty, I can compare the two different contributions in equation (28) directly. Expressed as the cgf of ϵ_t^{τ}, the epistemological contribution almost coincides with the cgf of $\chi_{1,t}$ evaluating shocks to the mean.[24] Both are evaluated proportional to risk aversion and the temperature's shadow value, but the cgf evaluating shocks to the mean contains the additional factor $\frac{\beta}{1-\gamma\beta}$. The discount factor β arises because the shock only affects welfare with a one period delay, whereas epistemological uncertainty is immediate. The factor $\frac{1}{1-\gamma\beta}$ translates the impact of the periodic and persistent shock χ_1^{τ} into the corresponding stationary uncertainty of the temperature flow ϵ_t. In determining its welfare impact, the persistence is discounted by time preference. It is this time preference that makes the shock contributions more sensitive to small discount rates as compared to the epistemological component: a patient decision maker is increasingly concerned with the cumulative uncertainty over the long-run future implied by periodic shocks.

The higher sensitivity of the "future shock" contribution to time preference seems to differ from the comparison between the Bayesian learning model and the VAR model in the case of the carbon cycle. However, in the present stylized learning model the future shocks simultaneously capture stochastic shocks to warming as well as updates to the mean (by moving next period's $\kappa_{1,t+1}$). Thus, the present model complements the earlier comparison by emphasizing that the updates to the mean (or higher moments) drive the high sensitivity to time preference of the Bayesian learning model, not merely the presence of epistemological uncertainty as captured by ϵ_0^{τ}.

[24] $\sum_{i=1}^{\infty} \varphi_{\kappa,i}\kappa_{i,t} = \frac{\beta}{\alpha(1-\beta)} \log\left[\mathbb{E}_t \exp\left[\alpha\varphi_{\tau,1}\epsilon_t^{\tau}\right]\right]$. This expression compares directly to the "Future Shock" contributions. In particular, the shock χ_1^{τ} moving the expected value ($i = 1$) takes the form: $\frac{\beta}{\alpha(1-\beta)} \log\left[\mathbb{E}_t \exp\left[\frac{\alpha\beta\varphi_{\tau,1}}{1-\beta\gamma}\chi_1^{\tau}\right]\right]$.

5.3 Quantifying the Welfare Impact of Temperature Uncertainty

For a quantitative assessment I employ again DICE's business as usual (BAU) emission scenario leading to the solid (deterministic) line in Figure 4. It employs the average of Meinshausen et al.'s (2009) survey of climate sensitivity distributions, whose mean I shift to the 3C best guess focussing on the impact of uncertainty only. Along the DICE BAU scenario greenhouse gas concentrations exceed the doubling of CO_2 significantly and I scale the temperature shocks accordingly. Initial epistemological uncertainty ϵ_0^τ and the distribution of the shocks to the mean $\chi_{1,t}$ are both chosen to reproduce climate sensitivity uncertainty if concentrations are double the pre-industrial level. Epistemological uncertainty ϵ_0^τ prevails immediately, whereas the shock uncertainty only builds up over time.

In the spirit of finding a lower bound of the welfare loss, I make the following two assumptions. First, I calibrate the shock uncertainty so as to produce the climate sensitivity distribution in the infinitely long run. Second, I split the overall climate sensitivity into an epistemological fraction ζ and a shock-based long-run fraction $1 - \zeta$. At any given point in time the actual forecast uncertainty will then be lower than the climate sensitivity distribution because epistemological uncertainty falls over time and the shocks only build up the fraction $1 - \zeta$ in the long-run. Third, I omit possible contributions from stochastic volatility or shocks to higher order cumulants. Also in the spirit of a lower bound, I pick $\alpha = -1$ at the lower end of measured risk aversion.

The first or "baseline" scenario assumes a persistence $\gamma = 0.9$ of both epistemological uncertainty and shocks to the mean, an equal split of overall climate sensitivity uncertainty between the epistemological and the shock contributions ($\zeta = \frac{1}{2}$), and the standard discount rate calibration to IMF 2015 data ($\rho = 1.75\%$). These assumption result in an overall welfare loss from climate sensitivity uncertainty of 16 trillion USD, approximately one year of US output. Initial epistemological uncertainty and the stochastic shocks contribute almost equal shares to this loss. As a consequence, attributing e.g. a larger share of the uncertainty to shocks and future updating hardly changes the welfare loss ($1 - \zeta = \frac{2}{3}$ increases the loss by half a trillion).

Reducing the persistence of shocks and of epistemological uncertainty to $\gamma = 0.7$ slightly reduces the epistemological contribution and significantly increases the welfare loss caused by the shocks. The shocks then contribute 4 times as much as the epistemological uncertainty to an overall welfare loss of 20 trillion USD. Analogously, an increase of the persistence to the value calibrated for carbon cycle uncertainty of $\gamma = 0.997$ reduces the overall welfare loss to 11 trillion USD deriving from initial epistemological uncertainty plus 0.5 trillion USD deriving from future shocks. The significant reduction (rather than increase) of the welfare loss from future shocks under a higher shock persistence is a result of the calibration: under a higher persistence a smaller decadal shock will result in the same long-run climate sensitivity uncertainty and an impatient decision maker prefers such a postponement of uncertainty.

A reduction of pure time preference to $\rho = 0.1\%$ in the "baseline" scenario increases the

loss to over 700 trillion USD or 8.5 years of world output. This factor 40 increase is even larger than the response of the carbon tax to the change in pure time preference. 95% of the contribution result from the future shocks. The robust quantitative insight is that a lower bound of the welfare loss from uncertainty over the climate's sensitivity to CO_2 is 2-3 orders of magnitude higher than the best guess of the welfare loss from uncertainty over the carbon flows. A clear quantitative message from economics to science to shift more attention to the feedback processes on the temperature side.

6 Conclusions

I introduce an integrated assessment model of climate change that matches scientific climate models as well as do numeric integrated assessment models used in policy advising. Yet, the GAUVAL model solves for the optimal carbon tax and welfare loss in closed form. GAUVAL merges Golosov et al.'s (2014) framework with a standard carbon cycle, radiative forcing, temperature dynamics, risk attitude, and different uncertainty frameworks. The resulting model compares closely to the wide-spread integrated assessment model DICE. GAUVAL exhibits a simpler capital depreciation structure, but has a better temperature tracking. The damage functions differ slightly without a clear winner. GAUVAL allows for a better market calibration by disentangling intertemporal substitution and the low risk free discount rate from risk aversion and risk premia. Finally, GAUVAL's results hold for a general class of energy production sectors that contain DICE's abatement formulation as a special case.

Policy measures have to rely on quantitative advice. GAUVAL delivers quantitative advice without losing closed-form analytic traction or brushing over major features of the climate and its interaction with the economy. The deterministic model finds a market-based social cost of carbon of 57 USD per ton of carbon (15 USD per ton of CO_2), using a standard calibration approach. The closed-form solution sizes the different contributions to this optimal carbon tax and establishes back-of-the-envelope formulas for the welfare cost of a temperature increase or the welfare gains from carbon sequestration. Like Golosov et al.'s (2014) model, GAUVAL implies a flat marginal benefit curve from mitigation. The finding underpins the advantages of a carbon tax over a cap and trade mechanism to regulate the climate externality. Another implication of the flat marginal benefit curve is that the optimal mitigation effort is independent of whether we followed business as usual or optimal policy in the past. If we "sinned" in the past, the optimal policy will not tell us to repent, but to live with the (perpetually) persisting consequences in the future.

It is a wide-spread believe that the optimal carbon tax is sensitive to the overall consumption discount rate, but not to its individual constituents. In contrast, I prove that mass conservation in the carbon cycle makes the optimal carbon tax in the present setting is highly sensitive to the rate of pure time preference, whereas proportionality of damages to output make it insensitive to growth related discounting. The sensitivity to pure time

preference weights particularly strong as recent asset pricing approaches as well as overlapping generations based calibration formulas suggest much lower rates of pure time preference than the 1.75% calibrated here. These approaches give support to rates as low as the 0.1% used as well in the Stern Review, increasing the optimal carbon tax tenfold.

I employ GAUVAL to advance our understanding of the welfare implications of uncertainty in climate change. I evaluate the impact of vector autoregressive shocks in substantial generality based on a distribution's moment generating function, or a sum of (intertemporal) risk aversion weighted cumulants. I solve a Bayesian model with anticipated learning over carbon flows under the assumption of normally distributed uncertainty. The anticipation of updated information has a similar effect as the VAR shocks, however, the updates initially act as fully persistent shocks. In consequence, the Bayesian learning model is more sensitive to pure time preference. In the standard calibration, the welfare loss from carbon cycle uncertainty is in the order of a hundred billion USD, and the willingness to pay for reducing measurement error and accelerating learning is in the order of half a billion (a few satellites and a supercomputer) per Gt C of decadal resolution.

Uncertainty about the temperature response to a given CO_2 path causes a lower bound welfare loss that is 2 to 3 orders of magnitudes larger (about one year of US output). Here, a reduction of pure time preference to 0.1% raises the loss even 40fold creating a lower bound welfare loss that is over 8 years of world output. The temperature uncertainty caused by climate sensitivity is not normally distributed and, for the assessment, I developed a stylized model combining VAR shocks with an epistemological uncertainty. Instead of tracking a few endogenous informational states characterizing a particular distribution, the model tracks the cumulant representation of an arbitrary distribution. Governments and research institutions are spending large amounts to better understand the carbon cycle. An immediate conclusion is that better assessments of the temperature feedback response has a significantly higher social payoff.

The intuition for this finding is as follows. Every additional ton of carbon in the atmosphere traps less energy than the preceding ton. This decreasing "harmfullness" of CO_2 to temperature partially offsets the convexity of damages from the implied warming. The remaining part of the damage convexity is offset by decreasing marginal utility (governing intertemporal trade-offs). Thus, if it was not for (disentangled) risk aversion, negative and positive shocks in the carbon flow would exactly offset each other. Risk aversion implies a willingness to pay for a risk reduction, but it is not large. In contrast, temperature feedbacks operate directly on temperatures. Because of the convex damage function, which is only partially offset by log-utility, high realizations cause more loss than low realizations return. In expectation, the shocks reduce overall welfare, an effect that is only amplified by risk aversion.

The present paper paves the way for a wide array of analytic and quantitative research. An accompanying paper analyzes the response of the optimal carbon tax to climate and

economic uncertainties. GAUVAL can be generalized for regional analysis, to examine adaptation, to analyze detailed damage channels like ocean-acidification or sea level rise, and to evaluate benefits from climate engineering projects. The present paper specifies the optimal carbon tax for a large class of energy sectors. A sequel to this paper will specify the details of the energy sector and analyze the sectoral and emission response to policy under technological uncertainty. Climate change is an intergenerational problem. Even the market-based approach contains hidden normative assumptions (Schneider et al. 2013). The present paper focuses on market-based evaluation, following common practice of policy advising in the US. GAUVAL also lends itself to a normatively motivated analysis. GAUVAL major virtues is to combine quantitative analysis with analytic insight. Any analytic approach has its limitations in the non-linearities and interactions it can handle. The model serves best as a benchmark, guiding and helping to enlighten fine-tuned quantitative numeric research.

References

Ackerman, F., Stanton, E. A. & Bueno, R. (2010), 'Fat tails, exponents, extreme uncertainty: Simulating catastrophe in DICE', *Ecological Economics* **69**(8), 1657–1665.

Allen, M. R., Frame, D. J., Huntingford, C., Jones, C. D., Lowe, J. A., Meinshausen, M. & Meinshausen, N. (2009), 'Warming caused by cumulative carbon emissions towards the trillionth tonne', *Nature* **485**, 1163–66.

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 No.13-05* .

Anthoff, D., Tol, R. S. J. & Yohe, G. W. (2009), 'Risk aversion, time preference, and the social cost of carbon', *Environmental Research Letters* **4**, 1–7.

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.

Cai, Y., Judd, K. L. & Lontzek, T. S. (2012), DSICE: A dynamic stochastic integrated model of climate and economy, Working Paper 12-02, RDCEP.

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

Dietz, S. (2009), 'High impact, low probability? An empirical analysis of risk in the economics of climate change', *Climatic Change* **108**(3), 519–541.

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.

Fischer, C. & Springborn, M. (2011), 'Emissions targets and the real business cycle: Intensity targets versus caps or taxes', *Journal of Environmental Economics and Management* **62**(3), 352–366.

Gerlagh, R. & Liski, M. (2012), Carbon prices for the next thousand years, Cesifo working paper series, CESifo Group Munich.

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

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

Hennlock, M. (2009), 'Robust control in global warming management: An analytical dynamic integrated assessment', *RFF Workding Paper* (DP 09-19).

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

Heutel, G. (2011), 'How should environmental policy respond to business cycles? Optimal policy under persistent productivity shocks', *University of North Carolina, Greensboro, Department of Economics Working Paper Series* **11**(08).

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

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

Interagency Working Group on Social Cost of Carbon, U. S. G. (2010), 'Technical support document: Social cost of carbon for regulatory impact analysis under executive order 12866', *Department of Energy* .

IPCC (2013), *Climate Change 2013: The Physical Science Basis*, Intergovernmental Panel on Climate Change, Geneva. Preliminary Draft.

Jensen, S. & Traeger, C. P. (2014), 'Optimal climate change mitigation under long-term growth uncertainty: Stochastic integrated assessment and analytic findings', *European Economic Review* **69**(C), 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. (2013), 'Provision of a public good with altruistic overlapping generations and many triibes', *Working Paper* .

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

Kelly, D. L. (2005), 'Price and quantity regulation in general equilibrium', *Journal of Economic Theory* **125**(1), 36 – 60.

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. (2013), 'Learning and climate feedbacks: Optimal climate insurance and fat tails', *University of Miami Working paper* .

Kopp, R. E., Golub, A., Keohane, N. O. & Onda, C. (2012), 'The influence of the specification of climate change damages on the social cost of carbon', *Economic E-Journal* **6**, 1–40.

Kreps, D. (1988), *Notes on the Theory of Choice*, Westview Press, Boulder.

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

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

Lontzek, T. S., Cai, Y. & Judd, K. L. (2012), 'Tipping points in a dynamic stochastic IAM', *RDCEP Working paper series* **No.12-03**.

Matthews, D., Gillett, N. P., Stott, P. A. & Zickfeld, K. (2009), 'The proportionality of global warming to cumulative carbon emissions', *Nature* **459**, 829–33.

Meinshausen, M., Meinshausen, N., Hare, W., Raper, S., Frieler, K., Knutti, R., Frame, D. & Allen, M. (2009), 'Greenhouse-gas emission targets for limiting global warming to 2c', *Nature* **458**, 1158–1162.

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.

Millner, A. (2013), 'On welfare frameworks and catastrophic climate risks', *Journal of Environmental Economics and Management* **65**(2), 310–325.

Moss, R., Babiker, M., Brinkman, S., Calvo, E., Carter, T., Edmonds, J., Elgizouli, I., Emori, S., Erda, L., Hibbard, K., Jones, R., Kainuma, M., Kelleher, J., Lamarque, J. F., Manning, M., Matthews, B., Meehl, J., Meyer, L., Mitchell, J., Nakicenovic, N., O'Neill, B., Pichs, R., Riahi, K., Rose, S., Runci, P., Stouffer, R., van Vuuren, D., Weyant, J., Wilbanks, T., van Ypersele, J. P. & Zurek, M. (2007), *Towards New Scenarios for Analysis of Emissions, Climate Change, Impacts, and Response Strategies.*, Intergovernmental Panel on Climate Change, Geneva.

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

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

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

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

Pycroft, J., Vergano, L., Hope, C. W., Paci, D. & Ciscar, J. C. (2011), 'A tale of tails: Uncertainty and the social cost of carbon dioxide', *Economic E-Journal* **5**, 1–29.

Rabin, M. (2000), 'Risk aversion and expected-utility theory: A calibration theorem', *Econometrica* **68**(5), 1281–1292.

Richels, R. G., Manne, A. S. & Wigley, T. M. (2004), 'Moving beyond concentrations: The challange of limiting temperature change', *Working Paper* .

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

Stern, N. (2008), 'The Economics of Climate Change', *American Economic Review* **98**(2), 1 – 37.

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

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

Traeger, C. P. (2014), 'Capturing intrinsic risk aversion', *Working Paper* .

Valentini, E. & Vitale, P. (2014), 'Optimal climate policy for a pessimistic social planner', *Working Paper* .

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

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

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

Weitzman, M. (1974), 'Prices versus quantities', *Review of Economic Studies* **41**(4), 477–91.

Weitzman, M. (2009), 'On modeling and interpreting the economics of catastrophic climate change', *The Review of Economics and Statistics* **91**(1), 1–19. 06.

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

Appendix

A General Capital Depreciation

Equation (2) assumes full capital depreciation. In this appendix, I show how to avoid the full capital depreciation assumption and match observed capital-output ratios through an exogenous adjustment of the growth rate. The model extension keeps the structural assumptions that imply a constant investment rate. Under a depreciation rate δ_k the capital accumulation equation (2) changes to

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

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

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

Defining the capital growth rate $g_{k,t} = \frac{K_{t+1}}{K_t} - 1$, I obtain the equivalent equation of motion for capital

$$K_{t+1} = Y_t[1 - D_t(T_{1,t})](1 - x_t)\left[\frac{1 + g_{k,t}}{\delta_k + g_{k,t}}\right] \ . \tag{29}$$

For full depreciation $\delta_k = 1$ the squared bracket is unity and equation (29) coincides with equation (2) in the main text. For $\delta_k < 1$ the squared bracket states an empirical correction multiplier larger unity. First, this multiplier can be used to match the model's capital accumulation to the empirical capital accumulation. Second, this multiplier makes the representative agent realize the additional capital value deriving from its persistence beyond its end of period value for production.

Treating the growth and depreciation correction in squared brackets as exogenous remains an approximation. The extension shows that the model is robust against the immediate criticism of not being able to represent the correct capital evolution and capital output ratio, and against the agent's neglecting of capital value beyond the time step. However, equation (29) with $g_{k,t}$ treated as exogenous remains an approximation. It is the price to pay for an analytic solution.

B Solution of the Linear-in-State Model

To obtain the equivalent linear-in-state-system, I first replace capital K_{t+1} by logarithmic capital $k_t \equiv \log K_t$. Second, I replace temperature levels in the atmosphere and the different ocean layers by the transformed exponential temperature states $\tau_{i,t} \equiv \exp(\xi_i T_{i,t})$,

$i \in \{1, ..., L\}$. I collected these transformed temperature states in the vector $\boldsymbol{\tau}_t \in \mathbb{R}^L$. Third, I use the consumption rate $x_t = \frac{C_t}{Y_t[1-D_t(T_t)]}$, rather than absolute consumption, as the consumption-investment control. Only the rate will be separable from the system's states. Finally, I define $a_t = \log A_{0,t}$ and express utility in terms of the consumption rate

$$u(C_t(x_t)) = \log C_t(x_t) = \log x_t + \log Y_t + \log[1 - D_t(T_t)] = \log x_t + a_t$$
$$+\kappa k_t + (1 - \kappa - \nu) \log N_{0,t} + \nu \log E_t - \xi_0 \exp[\xi_1 T_t] + \xi_0.$$

The Bellman equation in terms of the transformed state variables is

$$V(k_t, \boldsymbol{\tau}_t, \boldsymbol{M}_t, \boldsymbol{R}_t, t) = \max_{x_t, \boldsymbol{N_t}} \log x_t + a_t + \kappa k_t + (1 - \kappa - \nu) \log N_{0,t} \tag{30}$$
$$+\nu \log E_t - \xi_0 \tau_t + \xi_0 + \beta V(k_{t+1}, \boldsymbol{\tau}_{t+1}, \boldsymbol{M}_{t+1}, \boldsymbol{R}_{t+1}, t+1) \ ,$$

and is subject to the linear equations of motion

$$k_{t+1} = a_t + \kappa k_t + (1-\kappa-\nu) \log N_{0,t} + \nu \log g(\boldsymbol{E}_t(\boldsymbol{A}_t, \boldsymbol{N}_t))$$
$$-\xi_0 \tau_{1,t} + \xi_0 + \log(1-x_t) \tag{31}$$
$$\boldsymbol{M_{t+1}} = \boldsymbol{\Phi} \boldsymbol{M}_t + \left(\sum_{i=1}^{I^d} E_{i,t} + E_t^{exo} \right) \boldsymbol{e}_1 \tag{32}$$
$$\boldsymbol{\tau}_{t+1} = \boldsymbol{\sigma} \boldsymbol{\tau}_t + \sigma^{forc} \frac{M_{1,t} + G_t}{M_{pre}} \boldsymbol{e}_1 \tag{33}$$
$$\boldsymbol{R}_{t+1} = \boldsymbol{R}_t - \boldsymbol{E}_t^d, \tag{34}$$

and the constraints

$\sum_{i=0}^{I} N_{i,t} = N_t, \ N_{i,t} \geq 0,$

$\boldsymbol{R}_t \geq 0$ and $\boldsymbol{R}_0$ given.

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

I now show that the affine value function

$$V(k_t, \boldsymbol{\tau}_t, \boldsymbol{M}_t, \boldsymbol{R}_t, t) = \varphi_k k_t + \boldsymbol{\varphi}_M^\top \boldsymbol{M}_t + \boldsymbol{\varphi}_\tau^\top \boldsymbol{\tau}_t + \boldsymbol{\varphi}_{R,t}^\top \boldsymbol{R}_t + \varphi_t \tag{35}$$

solves the above linear-in-state system. The coefficients φ are the shadow value of the respective state variables, and $\top$ denotes the transpose of a vector of shadow values. The coefficient on the resource stock has to be time-dependent: the shadow value of the exhaustible resource increases (endogenously) over time following the Hotelling rule. The controls in the equations of motion (31)-(34) are additively separated from the states. Therefore, maximizing

the right hand side of the resulting Bellman equation delivers optimal control rules that are independent of the state variables. These controls are functions of the shadow values.

In detail, inserting the value function's trial solution in equation (35) and the next period states (equations 31-34) into the (deterministic) Bellmann equation (30) delivers

$$
\begin{aligned}
\varphi_k k_t + \boldsymbol{\varphi}_M^\top \boldsymbol{M}_t + \boldsymbol{\varphi}_\tau^\top \boldsymbol{\tau}_t + \boldsymbol{\varphi}_{R,t}^\top \boldsymbol{R}_t + \varphi_t = \\
&\max_{x_t, \boldsymbol{N}_t} \log x_t + \beta\varphi_k \log(1-x_t) + (1+\beta\varphi_k)\kappa k_t + (1+\beta\varphi_k) a_t \\
&\quad +(1+\beta\varphi_k)(1-\kappa-\nu)\log N_{0,t} \\
&\quad +(1+\beta\varphi_k)\nu \log g(\boldsymbol{E}_t(\boldsymbol{A}_t, \boldsymbol{N}_t)) \\
&\quad -(1+\beta\varphi_k)\xi_0 \tau_{1,t} + (1+\beta\varphi_k)\xi_0 \\
&\quad +\beta\boldsymbol{\varphi}_M^\top \left(\boldsymbol{\Phi}\boldsymbol{M}_t + \left(\textstyle\sum_{i=1}^{I^d} E_{i,t} + E_t^{exo} \right) \boldsymbol{e}_1 \right) \\
&\quad +\beta\boldsymbol{\varphi}_\tau^\top \left(\boldsymbol{\sigma}\boldsymbol{\tau}_t + \sigma^{forc} \frac{M_{1,t} + G_t}{M_{pre}} \boldsymbol{e}_1 \right) \\
&\quad +\beta\boldsymbol{\varphi}_{R,t+1}^\top \left(\boldsymbol{R}_t - \boldsymbol{E}_{1,t} \right) \\
&\quad +\beta\varphi_{t+1} \\
&\quad +\lambda_t \left(N_t - \textstyle\sum_{i=0}^{I} N_{i,t} \right)
\end{aligned}
$$

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

$$
\frac{1}{x} - \beta\varphi_k \frac{1}{1-x} = 0 \quad \Rightarrow \quad x^* = \frac{1}{1+\beta\varphi_k} \ . \tag{36}
$$

The labor input into the various sector's depends on the precise assumptions governing the energy sector, i.e., the specification of $g(\boldsymbol{E}_t(\boldsymbol{A}_t, \boldsymbol{N}_t))$. For a well-defined energy system, I obtain unique solutions as functions of the technology levels in the energy sector and shadow values of the endogenous state variables $\boldsymbol{N}_t^*(\boldsymbol{A}_t, \varphi_k, \boldsymbol{\varphi}_M, \boldsymbol{\varphi}_{R,t+1})$. Knowing these solutions is crucial to determine the precise output path and energy transition under a given policy regime. However, the SCC and, thus, the carbon tax do not depend on these solutions.

Inserting the (general) control rules into the maximized Bellman equation delivers the value function coefficients. In detail, I collect terms that depend on the state variables on

the left hand side of the resulting Bellman equation

$$
\begin{aligned}
&\big(\varphi_k - (1+\beta\varphi_k)\kappa\big)k_t + \left(\boldsymbol{\varphi}_M^\top - \beta\boldsymbol{\varphi}_M^\top\boldsymbol{\Phi} - \beta\varphi_{\tau,1}\frac{\sigma^{forc}}{M_{pre}}\boldsymbol{e}_1^\top\right)\boldsymbol{M}_t \\
&+\big(\boldsymbol{\varphi}_\tau^\top - \beta\boldsymbol{\varphi}_\tau^\top\boldsymbol{\sigma} + (1+\beta\varphi_k)\xi_0\boldsymbol{e}_1^\top\big)\boldsymbol{\tau}_t + \big(\boldsymbol{\varphi}_{R,t}^\top - \beta\boldsymbol{\varphi}_{R,t+1}^\top\big)\boldsymbol{R}_t \\
&+\varphi_t \;=\; \log x_t^*(\varphi_k) + \beta\varphi_k\log(1-x_t^*(\varphi_k)) + (1+\beta\varphi_k)\xi_0 + (1+\beta\varphi_k)a_t \qquad (37)\\
&\qquad +(1+\beta\varphi_k)(1-\kappa-\nu)\log \boldsymbol{N}_{0,t}^*(\boldsymbol{A}_t,\varphi_k,\boldsymbol{\varphi}_M,\boldsymbol{\varphi}_{R,t+1}) \\
&\qquad +(1+\beta\varphi_k)\nu\log g\big(\boldsymbol{E}_t(\boldsymbol{A}_t,\boldsymbol{N}_t^*(\boldsymbol{A}_t,\varphi_k,\boldsymbol{\varphi}_M,\boldsymbol{\varphi}_{R,t+1}))\big) \\
&\qquad +\beta\varphi_{M,1}\textstyle\sum_{i=1}^{I^d} E_{i,t}(\boldsymbol{A}_t,\boldsymbol{N}_t^*(\boldsymbol{A}_t,\varphi_k,\boldsymbol{\varphi}_M,\boldsymbol{\varphi}_{R,t+1})) + E_t^{exo} \\
&\qquad -\beta\boldsymbol{\varphi}_{R,t+1}^\top\boldsymbol{E}_t^d(\boldsymbol{A}_t,\boldsymbol{N}_t^*(\boldsymbol{A}_t,\varphi_k,\boldsymbol{\varphi}_M,\boldsymbol{\varphi}_{R,t+1})) \\
&\qquad +\beta\varphi_{t+1}\ .
\end{aligned}
$$

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

$$
\begin{aligned}
&\varphi_k - (1+\beta\varphi_k)\kappa = 0 && \Rightarrow \varphi_k = \frac{1}{1-\beta\kappa} \\
&\boldsymbol{\varphi}_M^\top - \beta\boldsymbol{\varphi}_M^\top\boldsymbol{\Phi} - \beta\varphi_{\tau,1}\frac{\sigma^{forc}}{M_{pre}}\boldsymbol{e}_1^\top = 0 && \Rightarrow \boldsymbol{\varphi}_M^\top = \frac{\beta\varphi_{\tau,1}\sigma^{forc}}{M_{pre}}\boldsymbol{e}_1^\top(\boldsymbol{1}-\beta\boldsymbol{\Phi})^{-1} \\
&\boldsymbol{\varphi}_\tau^\top + (1+\beta\varphi_k)\xi_0\boldsymbol{e}_1^\top - \beta\boldsymbol{\varphi}_\tau^\top\boldsymbol{\sigma} = 0 && \Rightarrow \boldsymbol{\varphi}_\tau = -\xi_0(1+\beta\varphi_k)\boldsymbol{e}_1^\top(1-\beta\boldsymbol{\sigma})^{-1} \\
&\boldsymbol{\varphi}_{R,t}^\top - \beta\boldsymbol{\varphi}_{R,t+1}^\top = 0 && \Rightarrow \boldsymbol{\varphi}_{R,t} = \beta^t\boldsymbol{\varphi}_{R,0}\ .
\end{aligned}
$$

Using the shadow value of log capital in equation (36) results in the optimal investment rate $x = 1-\beta\kappa$. From line (37) onwards, the maximized Bellman equation defines recursively the time-dependent affine part of the value function φ_t. Everything discussed in this paper is independent of the process φ_t and only assumes convergence of the value function. For most choices of $g(\boldsymbol{E}_t(\boldsymbol{A}_t,\boldsymbol{N}_t))$, the process φ_t has to be solved numerically together with the initial value of shadow price vectors of the scarce resources.

The SCC is the negative of the shadow value of atmospheric carbon expressed in money-measured consumption units. Inserting equation (9) for the shadow value of log-capital (see as well above) and (10) for the shadow value of atmospheric temperature (first entry of the vector) into equation (11) characterizing the shadow value of carbon in the different reservoirs delivers

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

As a consequence of logarithmic utility, this marginal welfare change translates into a consumption change as $du = \frac{1}{c}dc = \frac{1}{xY}dc \Rightarrow dc = (1-\beta\kappa)Y du$. Thus, the SCC is

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

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

I show the equivalence of the Bellman equation (12) and 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 GAUVAL's intertemporal elasticity of unity, the usual formulation reads

$$V_t^* = \exp\left((1-\beta)\log c_t + \beta \log\left[\mathbb{E}_t\, {V_{t+1}^*}^{\alpha^*}\right]^{\frac{1}{\alpha^*}}\right) . \tag{38}$$

Defining $V_t = \frac{\log V_t^*}{1-\beta}$ and rearranging equation (39) delivers

$$V_t = \log c_t + \frac{\beta}{1-\beta}\log\left[\mathbb{E}_t \exp\left((1-\beta)V_{t+1}\right)^{\alpha^*}\right]^{\frac{1}{\alpha^*}} . \tag{39}$$

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 (12) stated in the text.

Figure 5 illustrates the quantitative implications of a choice of risk aversion RRA$= 1-\alpha$ in the model.[25] In the basline, an agent consumes a constant level $\bar{c}$ in perpetuity. In a coin toss lottery, she loses 5% of her consumption in the upcoming decade (left) or 25% (right) in case of tails (probability 1/2). The graph presents, as a function of her risk aversion RRA, the percentage gain over the baseline that the agent requests if heads comes up to be indifferent between the lottery and the baseline. It is important to realize that these losses and gains are direct consumption changes. The numeric illustrations in the paper are based on the range RRA$^* = 1-\alpha^* \in [6, 9.5]$ found in the long-run risk literature. The bounds translate approximately into $\alpha = (1-\beta)\alpha^* \in \{1, 1.5\}$ in the present model's equation (12) and into RRA$\in \{2, 2.5\}$ in Figure 5.

[25] I directly illustrate risk aversion for the choice of $1-\alpha$ as opposed to Epstein-Zin's $1-\alpha^* = 1-\frac{\alpha}{1-\beta}$. This illustration is independent of time preference. A similar time preference independent illustration of Epstein-Zin's $1-\alpha^*$ would involve a lottery over infinite consumption streams. The argument why $1-\alpha^*$ as opposed to $1-\alpha$ would be time preference invariant relies on the idea that the lottery payoffs in the current period have less significance for a more patient agent.

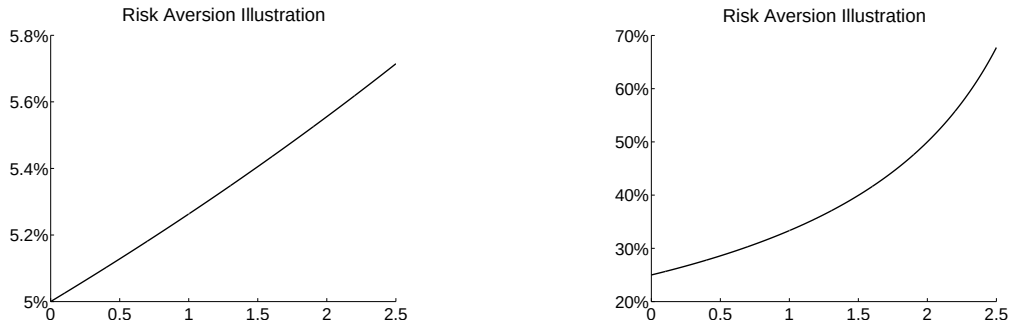

Figure 5: The graphs illustrate the relation between the risk aversion RRA= $1-\alpha$ and the relative consumption gains and losses that leave an agent indifferent to her original position. With probability 1/2, the agent loses 5% of her decadal consumption (left) or 25% (right). The graphs show how much of a relative consumption gain she requires for being indifferent to here initial deterministic position under different degrees of risk aversion.

D Illustrating the "Climate Matrices"

D.1 A Two Layer Carbon Cycle

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

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

The shadow value becomes less negative if more carbon flows from the atmosphere into the ocean (higher $\delta_{\text{Atm}\to\text{Ocean}}$). It becomes more negative for a higher persistence of carbon in the ocean (higher $1-\delta_{\text{Ocean}\to\text{Atm}}$). These impacts on the SCC are straight forward: the carbon in the ocean is the "good carbon" that does not contribute to the greenhouse effect. In round brackets, I find Proposition 2's root $(1-\beta)^{-1}$ that makes the expression so sensitive to a low rate of pure time preference.

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

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

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

Temperature is an intensive variable: it does not scale proportional to mass or volume (as is the case for the extensive variable carbon). The columns of the matrix $\boldsymbol{\sigma}$ do not sum to unity. As a consequence of the mean structure in equation (5), however, the rows in

the ocean layers' transition matrix sum to unity. The first row determining next period's atmospheric temperature sums to a value smaller than unity: it "misses" the weight that the mean places on radiative forcing. The "missing weight" is a consequence of the permanent energy exchange with outer space. Radiative forcing characterizes a flow equilibrium of incoming and outgoing radiation.

D.2 A Two Layer Atmosphere-Ocean Temperature System

The two layer example of atmosphere-ocean temperature dynamics has the transition matrix $\boldsymbol{\sigma} = \begin{pmatrix} 1-\sigma_1^{up}-\sigma_1^{down} & \sigma_1^{down} \\ \sigma_2^{up} & 1-\sigma_2^{up} \end{pmatrix}$. The corresponding term of the SCC (equation 13) takes the form

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

Persistence of the warming in the atmosphere or in the oceans increases the shadow cost. Persistence of warming in the oceans increases the SCC proportional to the terms σ_1^{down} routing the warming into the oceans and σ_1^{up} routing the warming back from the oceans into the atmosphere. The discount factor β accompanies each weighting factor, as each of them acts as a transition coefficient causing a one period delay. Taking the limit of $\beta \to 1$ confirms that (an analogue to) Proposition 2 does not hold

$$\lim_{\beta\to 1} \varphi_{\tau,1} = -\xi_0(1+\varphi_k)\left[(1-\boldsymbol{\sigma})^{-1}\right]_{11} = -\frac{\xi_0(1+\varphi_k)}{\sigma_1^{up}} \neq \infty. \tag{40}$$

As the discount rate approaches zero, the transient temperature dynamics characterized by σ_1^{down} and σ_2^{up} becomes irrelevant for evaluation, and only the weight σ_1^{up} reducing the warming persistence below unity contributes.[26]

D.3 The Price of Carbon and the Different Reservoirs

The carbon price in the atmosphere is immediately connected to its exchange with the remaining reservoirs. In fact, it can also be expressed the shadow value of carbon in any reservoir as the following function of the shadow prices in the remaining reservoirs

$$\varphi_{M,i} = \beta\frac{\sum_{j\neq i}\varphi_{M,j}\Phi_{j,i} + 1_{i,1}\frac{\varphi_{\tau,1}\sigma_1^{up}}{M_{pre}}}{1-\beta\Phi_{i,i}}. \tag{41}$$

[26]The carbon cycle lacks the reduction in persistence deriving from the forcing weight σ_1^{up} and equation (40) gives another illustration of the impact of mass conservation in the case of carbon: "$\sigma_1^{up} \to 0 \Rightarrow \varphi_{\tau,1} \to \infty$". Note that in the SCC formula σ_1^{up} cancels, as it simultaneously increases the impact of a carbon change on temperature. This exact cancellation (in the limit $\beta \to 1$) is a consequence of the weights σ_1^{up} on forcing and $1-\sigma_1^{up}$ on atmospheric temperature summing to unity. The result that a warming pulse has a transitional impact and an emission pulse has a permanent impact on the system is independent of the fact that these weights sum to unity.

The carbon price in layer i is the sum of carbon prices in the other layers times the flow coefficient capturing the carbon transition into that other layer (generally only positive for the two adjacent layers). The atmospheric carbon price has as an additional contribution ($1_{i,1}$ denotes the Kronecker delta) the shadow value of the atmospheric temperature increase. Finally, the denominator implies that the carbon price increases over the stated sum according to the persistence $\Phi_{i,i}$ of carbon in that given layer. Equation (41) resembles the carbon pricing formula for the one layer "carbon cycle" model discussed at the end of section D.1 where the atmospheric carbon persistence is $\Phi_{i,i} = 1 - \delta_{decay}$, and the present equation adds the pricing contributions from the other carbon absorbing layers as, unfortunately, the carbon leaving the atmosphere does not decay.

Finally, I illustrate the value of carbon sequestration in equation (14) for the case of the two layer carbon cycle introduced in section D.1

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

The value of carbon sequestration into the ocean falls in the stated manner in the transition parameter $\delta_{\text{Ocean}\to\text{Atm}}$ that captures the carbon diffusion from the ocean back into the atmosphere and increases with the transition parameter $1 - \delta_{\text{Atm}\to\text{Ocean}}$ that characterizes the persistence of carbon in the atmosphere. In the DICE carbon cycle, the value of sequestering carbon into the intermediate ocean and biosphere corresponding is \$40 per ton and the value of pumping carbon into the deep ocean is \$56 per ton.[27]

E Carbon Cycle Uncertainty: Details

E.1 Solving the Bellman Equation

In the case of persistent carbon sink uncertainty, the adjustments in the equations of motion (15) and (16) modify or add the following terms to the Bellman equation (12)

$$\varphi_\epsilon \epsilon_t + \varphi_t + ... = ... + \beta \varphi_{t+1} + \beta \varphi_\epsilon \gamma \epsilon_t + \beta[\varphi_{M_1} - \varphi_{M_2}]\epsilon_t + \frac{\beta}{\alpha} \log\left(\mathbb{E}_t \exp\left[\alpha \varphi_\epsilon \chi_t\right]\right) .$$

It is easily observed that these changes do not affect the optimal investment rate and labor distribution. Matching the coefficients of the flow adjustment ϵ_t to make the Bellman equation independent of its level delivers equation (17) for the shadow value φ_ϵ. The remaining terms imply $\varphi_t = \beta \varphi_{t+1} + \frac{1}{\alpha} \log\left(\mathbb{E}_t \exp\left[\alpha \beta \varphi_\epsilon \chi_t\right]\right) + const_t$, where $const_t$ is a term that is independent of the uncertainty. Given $\epsilon_0 = 0$, the welfare difference between the deterministic and the uncertain scenario is determined by the difference of the affine value function

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

contributions

$$\Delta W^{VAR} = V_0^{unc} - V_0^{det} = \varphi_0^{unc} - \varphi_0^{det} = \beta(\varphi_1^{unc} - \varphi_1^{det}) + \frac{\beta}{\alpha}\log\Big(\mathbb{E}_0 \exp\left[\alpha\varphi_\epsilon\chi_0\right]\Big) \ .$$
$$= \textstyle\sum_{i=0}^{\infty} \frac{\beta^{i+1}}{\alpha}\log\Big(\mathbb{E}_i \exp\left[\alpha\varphi_\epsilon\chi_i\right]\Big) + \lim_{i\to\infty}\beta^i(\varphi_i^{unc} - \varphi_i^{det}) \ .$$

For a well-defined dynamic system $\lim_{i\to\infty}\beta^i(\varphi_{t+i}^{unc} - \varphi_{t+i}^{det}) = 0$ and I obtain the general welfare loss equation for non-stationary shocks

$$\Delta W^{VAR} \tfrac{1}{\alpha}\textstyle\sum_{t=0}^{\infty}\beta^{t+1}\log\left[\mathbb{E}\exp\left[\alpha\varphi_\epsilon\chi_t\right]\right] \ . \tag{42}$$

For a sequence of identically distributed shocks χ_t, I obtain the welfare cost of uncertainty stated in (18) by evaluating implied geometric sum in equation (42).

In the case of anticipated learning, the adjustments in the equations of motion imply modifications of the Bellman equation (12) captured by the terms

$$\varphi_\mu\mu_{\epsilon,t} + \varphi_t + ... = ...+\beta\varphi_{t+1} + \beta\varphi_\mu\frac{\sigma_{\nu,t}^2}{\sigma_{\nu,t+1}^2 + \sigma_{\epsilon,t}^2}\mu_{\epsilon,t} \tag{43}$$
$$+\frac{\beta}{\alpha}\log\Big(\mathbb{E}_t\exp\Big[\alpha\Big(\varphi_{M_1} - \varphi_{M_2} + \varphi_{\mu,t}\frac{\sigma_{\epsilon,t}^2}{\sigma_{\nu,t+1}^2 + \sigma_{\epsilon,t}^2}\Big)(\epsilon_t + \nu_t)\Big]\Big).$$

Matching the coefficients of the informational state $\mu_{\epsilon,t}$ to make the Bellman equation independent of its level delivers equation (23) for the shadow value φ_μ. Solving inductively the remaining state-independent terms in equation (43) for the welfare difference between the uncertain and the deterministic scenario as above delivers the welfare loss

$$\Delta W^{Bayes} = \textstyle\sum_{t=0}^{\infty}\beta^{t+1}\alpha\frac{\sigma_{\epsilon,t}^2+\sigma_{\nu,t}^2}{2}\left[\varphi_{M_1} - \varphi_{M_2} + \varphi_{\mu,t}\frac{\sigma_{\epsilon,t}^2}{\sigma_{\nu,t}^2+\sigma_{\epsilon,t}^2}\right]^2 \ .$$
$$= \textstyle\sum_{t=0}^{\infty}\beta^{t+1}\alpha\frac{\sigma_{\epsilon,t}^2+\sigma_{\nu,t}^2}{2}\left(\varphi_{M_1} - \varphi_{M_2}\right)^2\left[\frac{(1-\beta)\sigma_{\epsilon,t}^2+(1-\beta)\sigma_{\nu,t}^2+\beta\sigma_{\epsilon,t}^2}{(1-\beta)(\sigma_{\epsilon,t}^2+\sigma_{\nu,t}^2)}\right]^2 \ ,$$

where I inserted the shadow value φ_μ from equation (23). Canceling terms in the numerator of the expression in squared brackets delivers equation (24) in the main text.

E.2 Quantitative Analysis of Carbon Cycle Uncertainty

The quantification of carbon cycle uncertainy in section 4.4 is an informed guess based on Joos et al.'s (2013) model comparison study and the measurement error implied by the missing sink. Here, I attempt to bound the welfare impact using a somewhat reasonable lower and upper bound for carbon cycle uncertainty. In the VAR model of section 4.2, the left panel of Figure 6 reduces the shock's standard deviation to 10 Gt per decade. It builds up to a 200Gt standard deviation after about 300 years, which is significantly lower than the 500Gt standard deviation in Joos et al.'s (2013) model comparison study. The

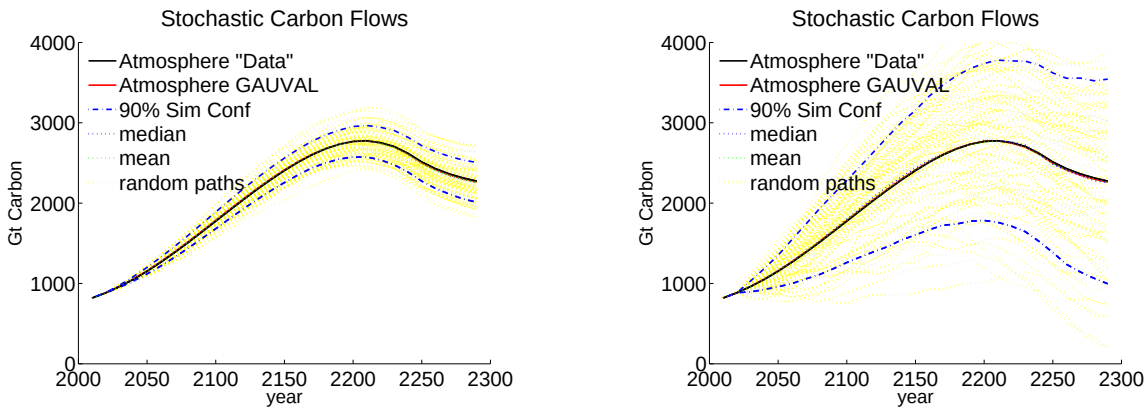

Figure 6: shows the evolution of atmospheric carbon under the low and the high specifications of the carbon cycle shock in equation (16), $\sigma_\chi = 10$ Gt on the left and $\sigma_\chi = 50$ Gt on the right. The shock's persistence of $\gamma_M = 0.997$ is calibrated to Joos et al.'s (2013) model comparison study. The underlying emission scenario is DICE's business as usual. The deterministic DICE evolution (5 year time steps, "Data"), the deterministic GAUVAL evolution (10 year time steps), and the mean and the median of 1000 uncertain trajectories are hardly distinguishable.

resulting welfare loss is approximately 28 billion USD. The right panel of Figure 6 increases the shock's standard deviation to 50 Gt per decade. It builds up to the suggested 500Gt standard deviation after 125 years, but implies double that value after around 350 years. The resulting welfare loss is approximately 700 billion USD.

The willingness to pay for a stochasticity reduction in Figure 4 also quantifies the sensitivity to a change in the measurement error or speed of learning, which is approximately half a billion USD per Gt decadal standard deviation. If the initial measurement error σ_ν is already down to 5Gt instead of 10Gt per decade, then this willingness to pay is also cut into half to approximately 260 million USD. If the initial measurement error is doubled ($\sigma_\nu = 20$Gt), then the willingness to pay increases to 750 million USD.

F Temperature Uncertainty: Details

F.1 Solving the Bellman Equation

In the combined model of persistent epistemological and VAR uncertainty over the temprature increase in section 5.2 the Bellman equation gains the following terms

$$\sum_{i=1}^{\infty} \varphi_{\kappa,i}\kappa_{i,t} + \varphi_t + ... = ... + \beta\varphi_{t+1} + \frac{\beta}{\alpha}\log\Big(\mathbb{E}_t \exp\Big[\alpha\Big(\sum_{i=1}^{\infty}\varphi_{\kappa,i}(\gamma_i\kappa_{i,t} + \chi_{i,t}^{\tau}) + \varphi_{\tau,1}\epsilon_t^{\tau}(\kappa_1, \kappa_2, ...)\Big)\Big]\Big).$$

$$\Rightarrow \varphi_t + ... = ... + \beta\varphi_{t+1} + \beta\sum_{i=1}^{\infty}\left[\varphi_{\kappa,i}(\beta\gamma_i - 1) + \beta\frac{(\alpha\varphi_{\tau,1})^i}{i!\alpha}\right]\kappa_{i,t} + \frac{\beta}{\alpha}\log\Big(\mathbb{E}_t \exp\Big[\alpha\Big(\sum_{i=1}^{\infty}\varphi_{\kappa,i}\chi_{i,t}^{\tau}\Big)\Big]\Big).$$

Matching the coefficients of the new states $\kappa_i, i \in \mathbb{N}$, eliminates the squared bracket in from of the cumulants and delivers the shadow values stated in equation (27). Moreover, analogously to the earlier scenarios, the difference in the uncertain and the deterministic value function's affine components is

$$\begin{aligned}\varphi_0^{unc} - \varphi_0^{det} &= \beta(\varphi_1^{unc} - \varphi_1^{det}) + \frac{\beta}{\alpha}\log\Big(\mathbb{E}_0 \exp\Big[\alpha\Big(\textstyle\sum_{i=1}^{\infty}\varphi_{\kappa,i}\chi_{i,t}^{\tau}\Big)\Big]\Big) \\ &= \textstyle\sum_{t=0}^{\infty}\frac{\beta^{t+1}}{\alpha}\log\Big(\mathbb{E}_t \exp\Big[\alpha\Big(\textstyle\sum_{i=1}^{\infty}\varphi_{\kappa,i}\chi_{i,t}^{\tau}\Big)\Big]\Big)\ .\end{aligned}$$

The welfare difference between the uncertain and the deterministic scenario is now comprised of a state (cumulant) dependent part $\sum_{i=1}^{\infty}\varphi_{\kappa,i}\kappa_{i,t}$ and the affine part of the value functions

$$\begin{aligned}\Delta W^{temp} &= V_0^{unc} - V_0^{det} = \sum_{i=1}^{\infty}\varphi_{\kappa,i}\kappa_{i,t} + \varphi_0^{unc} - \varphi_0^{det} \\ &= \textstyle\sum_{i=1}^{\infty}\varphi_{\kappa,i}\kappa_{i,t} + \sum_{i=0}^{\infty}\frac{\beta^{i+1}}{\alpha}\log\big(\mathbb{E}_t \exp\big[\alpha\big(\sum_{i=1}^{\infty}\varphi_{\kappa,i}\chi_{i,t}^{\tau}\big)\big]\big).\end{aligned}$$

In the case of identically distributed shocks over time, the second sum characterizes a geometric series giving rise to the factor $\frac{\beta}{1-\beta}$, turning the welfare loss into the form stated in equation (28) in the main text.

F.2 Quantitative Analysis of Temperature Uncertainty

Figure 7 illustrates the uncertainty governing the temperature increase from a doubling of the CO_2 concentration, the so-called climate sensitivity. On the left, the figure depicts 20 probability distributions of climate sensitivity derived by different groups and using different methodological approaches (Meinshausen et al. 2009). These probability densities are conditional on the temperature increase not exceeding 10 C. On the right, the figure depicts the average distribution assigning equal weight to each approach. The distribution is positively skewed and exhibits more weight in the right tail as compared to a (truncated) normal distribution. It serves as the starting point for my numeric estimates of the welfare loss from temperature uncertainty.

This average climate sensitivity distribution on the right of Figure 7 has an expected value of 3.4 C, differing from the common best guess of 3 C employed so far. Focusing on the uncertainty contribution, I shift Meinshausen et al.'s (2009) distribution to the left to conserve the 3 C warming expectation. I denote the implied distribution of the generalized temperature state by $\tilde{\tau}^{\infty}$. By equation (25), the temperature flow uncertainty $\epsilon^{\tau} = [\mathbf{1} - \boldsymbol{\sigma}]_{1,1}\tilde{\tau}^{\infty} - 2\sigma^{forc}$ generates this long-run temperature uncertainty under the assumption of a doubling of preindustrial CO_2 concentrations. I start by assuming only the VAR model, which corresponds to autoregressive shocks χ_1 to the mean. Such shocks build up over time, and for a doubling of CO_2 concentrations a stationary shock $\chi_1 = (1-\gamma)\epsilon^{\tau}$ generates the depicted distribution of climate sensitivity. More than two decades of IPCC assessment reports have not tightened

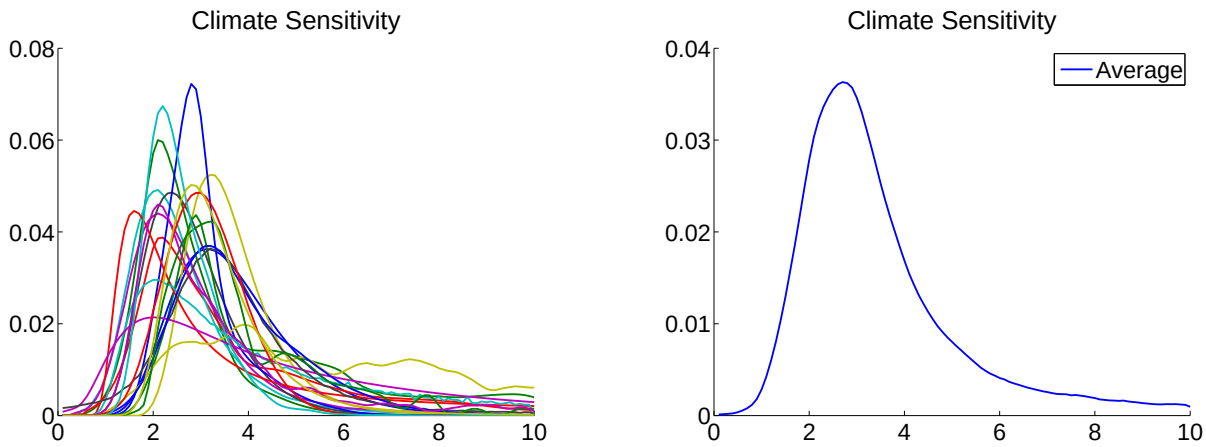

Figure 7: shows estimates of the probability distribution of a temperature increase in degree Celsius resulting from a doubling of CO_2 concentrations with respect to industrial levels. On the left, the figure depicts 20 probability distributions of climate sensitivity derived by different groups and using different methodological approaches (Meinshausen et al. 2009). On the right, the figure depicts the average distribution assigning equal weight to each approach. The Figure is to be interpreted as the probability density of a temperature increase conditional on not exceeding 10C.

the confidence interval on climate sensitivity. Therefore, I assume a persistence of epistemological uncertainty of $\gamma = 0.9$ in my "baseline" scenario. In evaluating the welfare loss from temperature uncertainty along the DICE business as usual scenario, I scale the exogenous shocks $\chi_{1,t}$ proportional to the atmospheric CO_2 concentrations along the business as usual path (thick black 'data' line in Figure 4).[28]

[28]The scaling of the shock is proportional to the CO_2 concentration because the shock affects transformed temperature, which translates logarithmically into real temperature, accounting for falling radiative forcing from an additional ton of CO_2.