Carbon prices for the next thousand years

Reyer Gerlagh and Matti Liski*

March 4, 2013

Abstract

An open puzzle for climate-policy analysis is how policies could be made sensitive to climate change impacts spanning over centuries while keeping the shorter-term macroeconomic policies connected to the descriptive facts. We develop a tractable general-equilibrium model for climate-economy interactions with time-declining pure discounting. The model resolves the puzzle: preferences over long-term climate outcomes can be expressed without sacrificing the description of the economy. The optimal carbon price shows a striking departure from the externality cost obtained from the economy's aggregate statistics — the equilibrium carbon price exceeds the imputed externality cost by multiple factors.

(JEL classification: H43; H41; D61; D91; Q54; E21. Keywords: carbon tax, discounting, climate change, inconsistent preferences)

---

*Gerlagh <r.gerlagh@uvt.nl> is at the economics department of the Tilburg University. Liski <matti.liski@aalto.fi> is at the economics department of the Aalto University, Helsinki. We thank Geir Asheim, Larry Goulder, Baard Harstad, John Hassler, Michael Hoel, Terry Iverson, Larry Karp, Dave Kelly, Per Krusell, Thomas Michielsen, Rick van der Ploeg, Tony Smith, Sjak Smulders, Christian Traeger, Cees Withagen, participants at Cowles Foundation 2012, NBER Summer Institute 2012, and SURED 2012 meetings, and at seminars in Helsinki, Stockholm, Toulouse, and Oxford for many useful comments and discussions.

1 Introduction

The essence of climate policy is a price for carbon, following from an evaluation of social costs arising from climate change. The unusual feature of climate change is the arrival delay of impacts, with persistent consequences spanning over centuries or possibly millennia into the future.[1] But the applied climate-economy models, commonly used to evaluate the carbon price, ignore the majority of climate externalities when discounting time as needed to describe shorter-term macroeconomic choices.[2] This fundamental puzzle of the climate-policy analysis has been fiercely debated, some emphasizing that discounting and thus policies should respect the shorter-term time preferences consistent with historical consumption choices (Nordhaus, 2007) while others put more weight on longer-term climate outcomes (Stern, 2006). The carbon tax recommendations can depart by a factor of ten. The discussion on how policies could be made sensitive to climate outcomes while keeping shorter-term economic decisions realistic has been inconclusive and confrontational.[3] Are we forced to either ignore the climate impacts or disconnect from the descriptive facts? Perhaps surprisingly, there has been no attempt to incorporate both views, high discounting for shorter-term macroeconomic decisions and lower discounting for longer-term trade-offs, in an equilibrium framework. We develop a tractable climate-economy model with such discounting to show that there is no climate-policy puzzle in general equilibrium: any description of the economy that is deemed realistic can be reconciled with carbon pricing policies that are sensitive to the climate outcomes.

There is evidence for treating the far-distant future differently from the short term. Layton and Brown (2000) surveyed 376 subjects and found no difference in their willingness to pay to prevent future ecosystem losses if these appeared after 60, or after 150 years. Weitzman (2001) surveyed 2,160 economists for their best estimate of the appropriate real discount rate to be used for evaluating environmental projects over a long time horizon, and used the data to argue that the policy maker should use a discount rate that declines over time — coming close to zero after 300 years.

[1]The impacts involve an intricate delay structure for atmospheric and ocean carbon dioxide diffusion, and land surface and ocean temperature adjustments. See, for example, Maier-Reimer and Hasselmann (1987), and Hooss, Voss, Hasselmann, Maier-Reimer, Joos (2001)

[2]The climate-economy models are commonly called integrated assessment models (IAMs) put forward by Peck and Teisberg (1992), Nordhaus (1993), and Manne and Richels (1995).

[3]For constructive contributions to the debate, see, e.g., Nordhaus (2007), Weitzman (2007), Dasgupta (2008).

While the relevance of time-declining pure discounting for climate policy evaluations has been long recognized, the general-equilibrium implications of such discounting have gone unnoticed. We derive a tractable general-equilibrium carbon price formula building on an explicit carbon cycle representation for climate impacts and the equilibrium implications of hyperbolic discounting. The formula allows a transparent quantitative assessment, and shows that the hyperbolic-discounting carbon price deviates from the standard Pigouvian principle: the current optimal carbon price exceeds the net present value of the future externality costs of emissions by multiple factors in our quantitative assessment. Apart from the time structure of preferences and our explicit carbon cycle for climate impacts, the framework for quantitative analysis is a general-equilibrium growth framework, following the Nordhaus' approach and its recent gearing towards the macro traditions by Golosov, Hassler, Krusell, and Tsyvinsky (2011).

Table 1 contains the gist of the quantitative assessment. The technology parameters are calibrated to 25 per cent gross savings, when both the short- and long-term annual time discount rate is 2 per cent. This is consistent with the Nordhaus' DICE 2007 baseline scenario (Nordhaus, 2007), giving 8.4 Euros per ton of CO_2 as the optimal carbon price in the year 2010 (i.e., 40 Dollars per ton C). When the longer-term receives a higher weight (roughly consistent with Weitzman's survey results), the shorter-term preferences can be matched so that the model remains observationally equivalent to Nordhaus in terms of macroeconomic performance, savings in particular. But carbon prices increase: for very low long-term discounting, carbon prices ultimately approach those suggested by Stern (2006).[4]

	discount rate			
	short-term	long-term	savings	carbon price
"Nordhaus"	.02	.02	.25	8.4
Equilibrium	.026	.001	.25	116.4
"Stern"	.001	.001	.30	151.8

Table 1: Equilibrium carbon prices in EUR/tCO_2 year 2010.

The aggregate statistics of the macroeconomy become distorted under hyperbolic discounting: there is a shortage of future savings leading to higher capital returns than what

[4]Under "Stern" the capital-share of output is fully saved (30 per cent); increasing the capital-share leads to unrealistic savings as discussed, e.g., in Weitzman 2007 and Dasgupta 2008.

the current policy maker would like to see.[5] As is well-known in cost-benefit analysis, a distorted capital return is not the social rate of return for public investments.[6] Our carbon price formula gives the optimal current policy, given the return distortions in the economy. Note that such distortions cannot be avoided with time-changing discount rates since policy decisions are *de facto* made in the order of appearance of policy makers in the time line.

The carbon price under hyperbolic discounting comes from a Markov equilibrium where each generation sets its self-interested savings and climate policies understanding how the future generations respond to current choices — time-changing discounting leads to a policy game between generations even when the current and all future policy makers internalize all climate impacts of emissions. Distortions arise from the lack of commitment to actions that we would like to implement in the future, as the future decision makers control their own capital savings and emissions but discount differently. But the future decision makers face the same dilemma — they value future savings and emission reductions, after their time, relatively more than the subsequent actual polluters. Therefore, also the future policy makers would value commitment to long-run actions. The extreme persistence of climate impacts provides commitment: the current climate policies alter directly the utilities of future agents. Also, future agents have no reason to undermine past climate investments as they value the climate capital for the same reason. The mechanism is similar to that delivering value for commitment devices in self-control problems (Laibson, 1997);[7] it explains why actions in climate protection are valued above the level implied by the pure Pigouvian externality pricing.

The quantitative significance of the commitment value follows from the unusual delays of the consequences of climate change. We develop and calibrate a novel representation of the carbon cycle, with the peak impact lagging 60-70 years behind the date of emissions. The analytics allows us to decompose the contribution of the different layers of the climate system to the carbon price: ignoring the delay of impacts — as in Golosov et al. (2011) — misses the correct price levels by a factor of 2, even when preferences are consistent. Getting the carbon price right is not merely an academic exercise; such prices are increasingly factored into the policy decisions, for example, into those that favor

[5]This distortion is the same as in Barro (1999); and Krusell, Kuruscu, and Smith (2002)

[6]See Lind (1982), or, e.g., Dasgupta (2008).

[7]However, self-control at the individual level is not the interpretation of the "behavioral bias" in our economy; we think of decision makers as generations as in Phelps and Pollak (1968). In this setting, the appropriate interpretation of hyperbolic discounting is that each generation has a social welfare function that expresses altruism towards long-term beneficaries (see also Saez-Marti and Weibull, 2005).

particular electricity generation technologies.[8]

The relevance of time-declining pure discounting for climate policy evaluations has been recognized before. Nordhaus (1999) and Mastrandrea and Schneider (2001) include hyperbolic discounting in an integrated-assessment model assuming that the current decision makers can choose also the future policies; they do not analyze how optimal policies evolve in an economy where future generations' policies cannot be dictated today. As a result, these papers show that decreasing discount rates lead to more aggressive climate policies as savings and climate policy are both targeted to long-term preferences; this way, the description of the shorter-term decisions becomes unrealistic, leading back to the climate-policy puzzle. Karp (2005), Fujii and Karp (2008) and Karp and Tsur (2011) consider Markov equilibrium climate policies under hyperbolic discounting without commitment to future actions, but these studies employ a very stylized setting without a model for intertemporal consumption choices. None of the above models can answer the question how optimal climate policy should be designed when concerns for long-term outcomes imply a preference for deep emission reductions, but when shorter-term macroeconomic choices should also be respected. Our tractable general-equilibrium model features a joint inclusion of macro and climate policy decisions. The key features of our model are hyperbolic time preferences, an equilibrium structure for consumption-climate policy decisions, and a solid carbon cycle description — these features are all essential for addressing the carbon price-discount rate puzzle.

We take the time-structure of preferences as given and focus on their general equilibrium climate policy implications, but multiple recent arguments can justify the deviation from geometric discounting. First, if we accept that the difficulty of distinguishing long-run outcomes describes well the climate-policy decision problem, then our decision procedure can imply a lower long-term discount factor than that for the short-term decisions (see Rubinstein 2003 for the procedural argument). Second, climate investments are public decisions requiring aggregation over heterogenous individual time-preferences, leading again to a non-stationary aggregate time-preference pattern, typically declining with the length of the horizon, for the group of agents considered (Gollier and Zeckhauser 2005; Jackson and Yariv 2011). We can also interpret Weitzman's (2001) study based on the survey of experts' opinions on discount rates as an aggregation of persistent views. Third, the long-term valuations must by definition look beyond the welfare of the imme-

[8]See Muller, Mendelsohn, and Nordhaus (2011) for the dramatic effect that carbon prices can have for the value-added evaluation of the US electricity sector. Greenstone, Kopits, and Wolverton (2011) is an overview of the values and estimates for the social cost of carbon used in the US federal rulemakings.

diate next generation; any pure altruism expressed towards the long-term beneficiaries implies changing utility-weighting over time (Phelps and Pollak 1968 & Saez-Marti and Weibull 2005).

The paper is organized as follows. The next section introduces the infinite-horizon climate-economy model, and develops the main results. Section 3 provides the quantitative assessment of the conceptual results. To obtain sharp results in a field dominated by simulation models, we make specific functional assumptions. Section 4 discusses those assumptions, and some robustness analysis as well as extensions to uncertainty and learning. Section 5 concludes.

2 An infinite horizon climate-economy model

2.1 Technologies and preferences: general setting

Consider a sequence of periods $t \in \{1, 2, ...\}$. The economy's production possibilities, captured by function $f_t(k_t, l_t, z_t, s_t)$, depend on capital k_t, labour l_t, current fossil-fuel use z_t, and the emission history (i.e., past fossil-fuel use),

$$s_t = (z_1, z_2, ..., z_{t-2}, z_{t-1}).$$

History s_t enters in production for two reasons. First, climate-change that follows from historical emissions changes production possibilities, as usual in climate-economy models. Second, the current fuel use is linked to historical fuel use through energy resources whose availability and the cost of use depends on the past usage. In the specific model that we detail below, we abstract from the latter type of history dependence, because the scarcity of conventional fossil-fuel resources is not binding when the climate policies are in place. The economy has one final good. The closed-form solutions require that capital depreciates in one period, leading to the following budget accounting equation between period t and $t + 1$:

$$c_t + k_{t+1} = y_t = f_t(k_t, l_t, z_t, s_t), \tag{1}$$

where c_t is the total consumption, k_{t+1} is capital built for the next period, and y_t is gross output. In each period, the representative consumer makes the consumption, fuel use, and investment decisions. Let per-period utility be u_t and define generation t welfare generated by sequence $\{c_t, z_t, k_t\}_{t=1}^{\infty}$ as

$$w_t = u_t + \beta \sum\nolimits_{\tau=t+1}^{\infty} \delta^{\tau-t} u_\tau \tag{2}$$

where $0 < \delta < 1$ is the long-term discount factor, and $0 < \beta$. As, for example, in Krusell et al. 2002, this β, δ –formulation implies that when $\beta < 1$, decision-makers use lower discount rates for long- than for short-term evaluations.[9] Furthermore, this increasing patience implies altruistic weights on future generations' welfare levels; see Saez-Marti and Weibull (2005) for the explicit derivation for the generation-specific welfare functionals.[10]

2.2 The specific climate-economy model

Golosov et al. (2011) marks a deviation from the Nordhaus' approach (e.g., 1993) to climate-economy modeling: abatement does not enter as a separate decision but is implied by the energy input choices. We follow this approach but our modeling of the climate dynamics, in addition to preferences, departs substantially from both Golosov et al. and Nordhaus. We pull together the production structure as follows:

$$y_t = k_t^{\alpha} A_t(l_{y,t}, e_t)\omega(s_t) \tag{3}$$

$$e_t = E_t(z_t, l_{e,t}) \tag{4}$$

$$l_{y,t} + l_{e,t} = l_t \tag{5}$$

$$\omega(s_t) = \exp(-\Delta_y D_t), \tag{6}$$

$$D_t = \sum\nolimits_{\tau=1}^{\infty} \theta_\tau z_{t-\tau} \tag{7}$$

Production. The gross production consists of: (i) the Cobb-Douglas capital contribution k_t^{α} with $0 < \alpha < 1$; (ii) function $A_t(l_{y,t}, e_t)$ for the energy-labour composite in the final-good production with $l_{y,t}$ denoting labor input and e_t the total energy use in the economy; (iii) total energy $e_t = E_t(z_t, l_{e,t})$ using fossil fuels z_t and labour $l_{e,t}$, and (iv) the damage part given by function $\omega(s_t)$ capturing the output loss of production depending on the history of emissions from the fossil-fuel use. The functional forms for the capital contribution and damages allow a Markov structure for policies, and thus the determination of the currently optimal policies as function of the state of the economy, say, at year 2010. Our quantitative assessment focuses on the currently optimal policies, and therefore we leave the detailed elaboration of functions A_t and E_t to our longer

[9]The formal analysis is not restricted to this quasi-hyperbolic setting. For interpretations, the quasi-hyperbolic case is the most natural to keep in mind, but we will state explicitly the formal results that require $\beta < 1$. Moreover, in Section 3 we discuss how the analysis extends to an arbitrary sequence of discount factors.

[10]These preferences are specific for generation t, and in that sense, w_t is different from the generation-independent social welfare function (SWF) as discussed, e.g., in Goulder and Williams (2012) and Kaplow et al. (2010).

working paper Gerlagh&Liski (2012).[11] Historical emissions, z_t for $t < 1$, affect future damages. We assume that the final-good and energy-sector outputs are differentiable and increasing in labor, energy, and carbon inputs.

Damages and carbon cycle. Equations (6)-(7) show that climate damages are interpreted as reduced output, and depend on the history of emissions through the state variable D_t that measures the global mean temperature increase; below, with slight abuse of terminology, we refer to D_t as damages. The exponential form for output losses combined with linear dependence on the past emissions is the same as in Golosov et al. (2011). But the specification of the parameters θ_τ in our model is very different: in the Appendix, we develop a closed-form representation for the global carbon-climate cycle, allowing a transparent and detailed calibration. The pattern of the temperature anomaly that follows from this representation shows a delay between the cause (emissions) and the effect (output losses); the pattern that results from our calibration is qualitatively similar to that in DICE (Nordhaus, 2007). The delay pattern has substantial implications for policies; we contrasting our calibrated emissions-damage response with both Nordhaus and Golosov et al. (2011) below.

The weighting of past emissions in (7) is obtained from a model for the global carbon cycle that refers to a diffusion process of carbon between various reservoirs of the atmosphere, oceans and biosphere (Maier-Reimer and Hasselman 1987). In the Appendix we describe this diffusion as a Markov process. Emissions z_t enter the atmospheric CO_2 reservoir, and slowly diffuse to the other reservoirs. The deep ocean is the largest reservoir, and acts as the major sink of atmospheric CO_2. We calibrate this reservoir system, and, for ease of analysis, by linear transformation obtain an isomorphic decoupled system of "atmospheric boxes" where the diffusion pattern between the boxes is eliminated. We describe the carbon cycle in terms of such a system of independent atmospheric boxes. Let $\mathcal{I}$ denote the set of boxes, with share $0 < a_i < 0$ of annual emissions entering each box $i \in \mathcal{I}$, and $\eta_i < 1$ its carbon depreciation factor. A three-box representation will be sufficiently rich to capture the analytical essence of the carbon cycle dynamics.

After an emissions impulse, carbon concentrations rise but temperatures and thus

[11]Emissions can decline through energy savings, obtained by substituting labor $l_{y,t}$ for total energy e_t. Emissions can also decline through "de-carbonization" that involves substituting non-carbon inputs for carbon energy inputs z_t in energy production; de-carbonization is obtained by allocating the total energy labor $l_{e,t}$ further between carbon and non-carbon energy sectors. Typically, the climate-economy adjustment paths feature early emissions reductions through energy savings, whereas de-carbonization is necessary for achieving long-term reduction targets. See Gerlagh&Liski (2012).

damages increase with delay. We assume a linear relationship between concentrations and damages. We describe the sensitivity of the damages to increases in atmospheric CO_2 stocks by parameter π: one-unit increase in the steady-state atmospheric CO_2 stock leads to π-unit increase in the steady-state D_t.

Outside steady state, there is delay in the effect from concentrations to damages, and this delay is captured by parameter $0 < \varepsilon < 1$: one-unit increase in emissions increases next period concentrations one-to-one but damages only $\varepsilon\pi$ -units. This description leads to a closed-form for an emissions-damages response (see the Appendix for the derivation): the impact of emissions at time t on damages at time $t + \tau$ is

$$\frac{dD_{t+\tau}}{dz_t} = \theta_\tau = \sum\nolimits_{i \in \mathcal{I}} a_i \pi \varepsilon \frac{(1-\eta_i)^\tau - (1-\varepsilon)^\tau}{\varepsilon - \eta_i} > 0,$$

where the geometric terms $(1 - \eta_i)^\tau$ and $(1 - \varepsilon)^\tau$ characterize the delays in carbon concentration and temperature adjustments; η_i is the calibrated carbon depreciation in each climate box.

Function θ_τ, when calibrated, is hump-shaped with a peak around 60-70 years after emissions. The essence of the response is very intuitive. Parameter η_i captures, for example, the carbon uptake from the atmosphere by forests and other biomass, and oceans. The term $(1-\eta_i)^\tau$ measures how much of carbon z_t still lives in box i, and the term $-(1 - \varepsilon)^\tau$ captures the slow temperature adjustment in the earth system. The limiting cases are revealing. Consider one CO_2 box, so that the share parameters are $a = 1$. If atmospheric carbon-dioxide does not depreciate at all, $\eta = 0$, then the temperature slowly converges at speed $\theta_\tau = \pi[1 - (1 - \varepsilon)^\tau]$ to the long-run equilibrium damage sensitivity π. If atmospheric carbon-dioxide depreciates fully, $\eta = 1$, the temperature immediately adjusts to $\pi\varepsilon$, and then slowly converges to zero, $\theta_\tau = \pi\varepsilon(1 - \varepsilon)^{\tau-1}$. If temperature adjustment is immediate, $\varepsilon = 1$, then the temperature response function directly follows the carbon-dioxide depreciation $\theta_\tau = \pi(1 - \eta)^{\tau-1}$. If temperature adjustment is very slow, $\varepsilon = 0$, there is no response, $\theta_\tau = 0$.

The physical data on carbon emissions, stocks in various boxes, and the observed concentration developments are used to calibrate a 3-box carbon cycle representation leading to the following emission shares and depreciation factors per decade:[12]

$$\begin{aligned} a &= (.163, .184, .449) \\ \eta &= (0, .074, .470). \end{aligned}$$

[12]Some fraction of emissions enters the ocean and biomass within a decade, so the shares a_i do not sum to unity.

Thus, about 16 per cent of carbon emissions does not depreciate while about 45 per cent has a half-time of one decade. For a climate sensitivity of $3K$ (Kelvin), we proxy its value at about 4.25 $[K^2/GtCO_2]$. We assume $\varepsilon = .183$ per decade, implying a global temperature adjustment speed of 2 per cent per year. These choices are within the ranges of scientific evidence (Solomon et al. 2007).

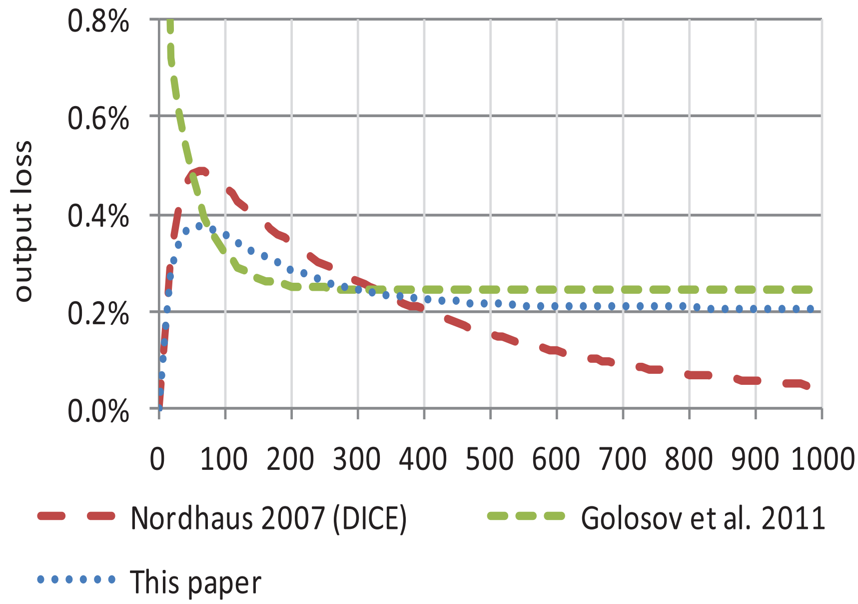

Figure 1: Emissions-Damage responses.

Figure 1 shows the life path of damages (percentage of total output) caused by inserting one TtonC in the first period, and then contrasting the impact with a counterfactual path without the carbon impulse.[13] Golosov et al.'s (2011) specification can be understood as one where the temperature adjustment is immediate: $\varepsilon = 1$ so that an emission impulse leads to an immediate temperature shock that slowly decays. The specification following Golosov et al. produces an immediate peak but a fat tail, while the DICE model shows an emissions-damage peak after 60 years with a thinner tail. Our model, that we calibrate with data from the natural sciences literature, produces a combination of the effects: a peak in the emission-damage response function after about 60 years and

[13]One TtonC equals about 50 years of global CO_2 emissions at current levels (40 GtCO_2/yr.) At current growth rates, it is about what we expect to emit between 2010 and 2050. See Appendix for the details of the experiment.

a fat tail; about 16 per cent of emissions do not depreciate within the horizon of thousand years.[14] We note that the reduced-form model above can be calibrated very precisely to approximate the DICE model (Nordhaus 2007).

Periodic utility. We assume that the utility function is logarithmic, and through a separable linear term we also include the possibility of intangible damages associated with climate change:

$$u_t = \ln(\frac{c_t}{l_t}) - \Delta_u D_t. \tag{8}$$

The utility loss $\Delta_u D_t$ is not necessary for the substance matter of this paper, but it proves useful to explicate how it enters the carbon price formulas. In calibration, we let $\Delta_u = 0$ to maintain an easy comparison with the previous studies.[15] Note that we consider average utility in our analysis.[16]

Strategies. Our focus is the symmetric and stationary Markov equilibrium.[17] Symmetry means that all generations use the same policy functions — even though there can be technological change and population growth, the form of the objective in (8) ensures that there will be an equilibrium where the same policy rule will be used for all t. The Markov restriction means that the policy does not condition on the history of past behavior: strategies are identical in states where the continuation payoffs are identical (see Maskin and Tirole, 2001).[18] In equilibrium, the policy will take the form $k_{t+1} = g_t(k_t, \Theta_t)$, $z_t = h_t(k_t, \Theta_t)$, where Θ_t collects the vector of climate state variables.

[14]The main reason for the deviation from DICE is that DICE assumes an almost full CO_2 storage capacity for the deep oceans, while large-scale ocean circulation models point to a reduced deep-ocean overturning running parallel with climate change (Maier-Reimer and Hasselman 1987). The positive feedback from temperature rise to atmospheric CO_2 through the ocean release is essential to explain the large variability observed in ice cores in atmospheric CO_2 concentrations.

[15]See Tol (2009) for a review of the existing damage estimates. While the estimates for intangible damages are mostly missing, our carbon pricing formulas can help to transform output losses into equivalent intangible losses to gauge the relative magnitudes of such losses that can be associated with a given carbon price level.

[16]Alternatively, we can write aggregate utility within a period by multiplying utility with population size, $u_t = l_t \ln(c_t/l_t) - l_t \Delta_u D_t$. The latter approach is feasible but it leads to considerable complications in the formulas below. Scaling the objective with labor rules out stationary strategies — they become dependent on future population dynamics —, and also impedes a clear interpretation of inconsistencies in discounting. While the formulas in the Lemmas depend on the use of an average utility variable, the substance of the Propositions is not altered. The expressions for this case are available on request

[17]We describe an equilibrium in symmetric but non-stationary strategies in Gerlagh&Liski (2012) too see the implications for carbon pricing; the differences are not large.

[18]We will construct a natural Markov equilibrium where policies have the same functional form as when $\beta = 1$. For multiplicity of equilibria in this setting, see Krusell and Smith (2003) and Karp (2007).

However, since the climate affects the continuations payoffs only through the weighted sum of past emissions, as expressed in (7), we will replace Θ_t by history s_t below, treating it as a state variable.

Structure of equilibria. Given policies $g_t(k_t, s_t)$ and $h_t(k_t, s_t)$, we can write welfare in (2) as follows

$$\begin{aligned} w_t &= u_t + \beta\delta W_{t+1}(k_{t+1}, s_{t+1}), \\ W_t(k_t, s_t) &= u_t + \delta W_{t+1}(k_{t+1}, s_{t+1}) \end{aligned}$$

where $W_{t+1}(k_{t+1}, s_{t+1})$ is the (auxiliary) value function. All equilibria considered in this paper will be of the form where a constant share $0 < g < 1$ of the gross output is invested,

$$k_{t+1} = gy_t, \tag{9}$$

whereas the climate policy defines emissions through a constant h that defines the carbon policies through

$$f_{t,z} = h(1-g)y_t, \tag{10}$$

where $f_{t,z}$ is the marginal product of fossil fuel use, the carbon price. Equilibrium policies will be characterized simply by a pair of constants (g, h). That a constant fraction of output is saved should not be surprising, given the log utility and Cobb-Douglas contribution of capital in production.[19] Condition (10) implies that the marginal carbon price per consumption is a constant, $h = f_{t,z}/c_t$ where $c_t = (1-g)y_t$. This may seem surprising given the complicated delay structure (7), and changing productivities in (3)-(7), and preference inconsistencies.[20]

Postponing the verification that the policies actually take the above form, it proves useful to state the properties of the value function implied by (g, h) policies (the proofs, unless helpful in the text, are in the Appendix).

Lemma 1 *(separability) Given the model (3)–(8), assuming that future policies $g_\tau(\cdot)$ and $h_\tau(\cdot)$ for $\tau = t+1, t+2, ...$ satisfy (9) and (10), then the value function is separable in capital and historical emissions*

$$W_{t+1}(k_{t+1}, s_{t+1}) = V_{t+1}(k_{t+1}) - \Omega(s_{t+1}).$$

[19]See, e.g., Barro 1999, for the analysis of the one-capital good case.

[20]Golosov et al. find emission policies that have the same features; our policies exploit the same functional assumptions, despite the added complexity.

with parametric form

$$
\begin{aligned}
V_{t+1}(k_{t+1}) &= \xi \ln(k_{t+1}) + \tilde{A}_{t+1} \\
\Omega(s_{t+1}) &= \sum_{\tau=1}^{t-1} \zeta_\tau z_{t+1-\tau},
\end{aligned}
$$

where $\xi = \frac{\alpha}{1-\alpha\delta}$, $\zeta_1 = \Delta \sum_{i\in\mathcal{I}} \frac{a_i \pi \varepsilon}{[1-\delta(1-\eta_i)][1-\delta(1-\varepsilon)]}$, $\Delta = (\frac{\Delta_y}{1-\alpha\delta} + \Delta_u)$ *and* $\tilde{A}_{t+1}$ *is independent of* k_{t+1} *and* s_{t+1}*.*

The result that the value of savings k_{t+1} and the costs from fossil-fuel use z_t can be obtained separately hinges on the strong functional assumptions; we discuss these in detail in Section 3.

2.3 The Markov equilibrium policies

The functional forms and the capital depreciation assumption imply that the consumption choice model is effectively Brock-Mirman (1972). Krusell et al. (2002) describe the savings policies for this model with quasi-hyperbolic preferences. Each generation takes the future policies, captured by constants (g, h) in (9)-(10), as given and chooses its current savings to satisfy

$$u'_t = \beta\delta V'_{t+1}(k_{t+1}),$$

where u'_t denotes marginal consumption utility and function $V(\cdot)$ from Lemma 1 captures the continuation value implied by the equilibrium policy.

Lemma 2 *(savings) The equilibrium investment share* $g = k_{t+1}/y_t$ *is*

$$g^* = \frac{\alpha\beta\delta}{1 + \alpha\delta(\beta - 1)}. \tag{11}$$

The proof of the Lemma is a straightforward verification exercise following from the first-order condition. If future savings could be dictated today, then $g^{\beta=1} = \alpha\delta$ for future decision-makers would maximize the wealth as captured by $W_{t+1}(k_{t+1}, s_{t+1})$; however, equilibrium g^* with $\beta < 1$ falls short of $g^{\beta=1} = \alpha\delta$ because each generation has an incentive to deviate from this long-term plan due to higher impatience in the short run (Krusell et al., 2002).

Consider then the equilibrium carbon price $f_{t,z}$, that is, the marginal product of the fossil-fuel use z_t, satisfying

$$u'_t f_{t,z} = \beta\delta \frac{\partial\Omega(s_{t+1})}{\partial z_t},$$

where function $\Omega(.)$ gives the future costs of emissions implied by the equilibrium policy. The optimal policy thus equates the marginal current utility gain from fuel use with the change in equilibrium costs on future agents. Given Lemma 1, the equilibrium carbon price and the fossil-fuel use can be obtained:

Proposition 1 *Equilibrium emissions $z_t = z_t^*$ depend only on the current technology at period t as captured through $A_t(.)$ and $E_t(.)$. The equilibrium carbon price is*

$$MCP_t = f_{t,z} = h^*(1-g)y_t \tag{12}$$

$$h^* = \Delta \sum_{i\in\mathcal{I}} \frac{\beta\delta a_i \pi\varepsilon}{[1-\delta(1-\eta_i)][1-\delta(1-\varepsilon)]}. \tag{13}$$

When y_t is known, say $y_{t=2010}$, the carbon policy for $t = 2010$ can be obtained from (13), by reducing fossil-fuel use to the point where the marginal product of z equals the consumption-weighted externality cost of carbon, as expressed in (13). For the functional form of policy h^*, note that Δ is the total damage, measured in utility, per unit of increase in D_t at time t. To obtain the carbon price intuitively, that is, the social cost of carbon emissions z_t as seen by the current generation, consider the effect of emissions at t on period $t+\tau$ utility:[21]

$$\frac{du_{t+\tau}}{dz_t} = \Delta\frac{dD_{t+\tau}}{dz_t} = \Delta\sum_{i\in\mathcal{I}} a_i\pi\varepsilon\frac{(1-\eta_i)^\tau - (1-\varepsilon)^\tau}{\varepsilon-\eta_i}.$$

Summing over all future periods and discounting with factor $0 < \beta, \delta < 1$ gives the present-value utility cost of emissions:

$$\begin{aligned}\beta\sum_{\tau=1}^{\infty}\delta^\tau\frac{du_{t+\tau}}{dz_t} &= \Delta\sum_{i\in\mathcal{I}}\frac{\beta a_i\pi\varepsilon}{\varepsilon-\eta_i}\sum_{\tau=1}^{\infty}\delta^\tau(1-\eta_i)^\tau - \delta^\tau(1-\varepsilon_j)^\tau \\ &= \Delta\sum_{i\in\mathcal{I}}\frac{\beta\delta\pi a_i\varepsilon}{[1-\delta(1-\eta_i)][1-\delta(1-\varepsilon)]}.\end{aligned}$$

This is exactly the value of h^*.

The Markov equilibrium carbon price, as indicated by (13) and Lemma 1, depends on the delay structure in the carbon cycle captured by parameters η_i and ε. Carbon prices increase with the damage sensitivity ($\partial h/\partial\pi > 0$), slower carbon depreciation ($\partial h/\partial\eta_i < 0$), and faster temperature adjustment ($\partial h/\partial\varepsilon > 0$). Higher short- and long-term discount rates both decrease the carbon price ($\partial h/\partial\beta > 0$; $\partial h/\partial\delta > 0$). The carbon price rises sharply if the discount factor comes close to one, $\delta \to 1$, and if some box the

[21] Recall that $\Delta = (\frac{\Delta_y}{1-\alpha\delta} + \Delta_u)$. The adjustment of the output loss Δ_y by $(1-\alpha\delta)^{-1}$ is to account for the decrease in the future capital stock caused by a current drop in output.

depreciation is low, $\eta_i \to 0$. If carbon depreciates quickly, $\eta_i > 0$, then the carbon price will be less sensitive to the discount factor δ.[22]

2.4 The imputed Pigouvian tax

The equilibrium defines a utility-discount factor $0 < \gamma < 1$ for consumption that is obtained from

$$u'_t = \gamma u'_{t+1} R_{t,t+1}$$

where $R_{t,t+1}$ is the capital return between t and $t+1$. Thus,

$$\gamma = \frac{u'_t}{u'_{t+1} R_{t,t+1}} = \frac{c_{t+1}}{c_t R_{t,t+1}} = \frac{c_{t+1}}{c_t} \frac{k_{t+1}}{\alpha y_{t+1}} = \frac{g}{\alpha}. \tag{14}$$

In the Markov equilibrium where $g = g^*$, we have

$$\gamma^* = \frac{\beta\delta}{1 + \alpha\delta(\beta - 1)}. \tag{15}$$

This is the geometric utility discount factor that is consistent with the efficiency of the equilibrium consumption stream: a fictitious planner who has consistent preferences and discounts with γ^* would find the equilibrium policy g optimal. We can also find the social cost of carbon for a planner who discounts with γ^*. Since this defines the full externality cost of equilibrium actions for such a planner, we arrive at the definition of the Pigouvian tax imputed to the consistent preferences discounting γ^*.[23]

Proposition 2 *(Imputed Pigouvian tax) The net present value of marginal damages of emissions τ_t^γ, discounted with $\gamma = \gamma^*$, is given by*

$$\begin{aligned} \tau_t^\gamma &= h^\gamma (1-g) y_t && (16) \\ h^\gamma &= \Delta^\gamma \textstyle\sum_{i \in \mathcal{I}} \frac{\gamma \pi a_i \varepsilon}{[1-\gamma(1-\eta_i)][1-\gamma(1-\varepsilon)]} && (17) \\ \Delta^\gamma &= \frac{\Delta_y}{1-\alpha\gamma} + \Delta_u. \end{aligned}$$

[22]This feature explains the finding by Fujii and Karp (2008) who conclude that the mitigation level is not very sensitive to the discount rate. Their representation of climate change can be interpreted as one in which CO_2 depreciates at more than 25% per decade, well above the estimates in the natural sciences literature.

[23]The adjustment of the output loss Δ_y by $(1-\alpha\gamma)^{-1}$ is to account for the decrease in the future capital stock caused by a current drop in output for the planner show discounts with γ and thus has $g = \alpha\gamma$; see also footnote 21

We have now two definitions for the social cost of carbon. The one in Proposition 1 is the current best response to future policies; it foresees the distortions in the economy, from the current preferences perspective. The other carbon price in Proposition 2 uses the aggregate statistics of the economy to measure the costs imposed on future agents from increases in current emissions. If the current carbon price does not equal the imputed Pigouvian tax, it is possible to strictly increase consumption (i.e., utility) in one period, without decreasing consumption in another period. However, it is not immediate that there are gains to be obtained by imposing the imputed carbon pricing rule on the economy — this latter exercise is artificial but useful as it reveals whether institutions that enforce Pigouvian carbon pricing based on the aggregate statistics only should be established.

We address first the conditions when the two carbon prices differ:

Proposition 3 *For quasi-hyperbolic preferences, $\beta < 1$, the equilibrium carbon price strictly exceeds the imputed Pigouvian tax if climate change delays are sufficiently long. Formally, the ratio of the equilibrium carbon price and the efficient carbon price, $f_{t,z}/\tau_t^\gamma$, is continuous in parameters $\beta, \delta, \eta_i, \varepsilon_j, a_i, b_j$, and γ. Evaluating at $\gamma = \gamma^*$, $\beta < 1$, $\eta_i = \varepsilon_j = 0$,*

$$f_{t,z} > \tau_t^\gamma.$$

If preferences are quasi-hyperbolic and the climate system is sufficiently persistent, then the current generation uses the climate asset as a commitment device to transfer wealth to future generations, and therefore it values the external climate costs above the imputed Pigouvian level. It is well known that when $\beta < 1$ the future equilibrium savings are lower than preferred from the current generation's point of view (Laibson 1997; Krusell et al. 2002). There is thus a capital market distortion, implying higher future capital returns than what the current generation would like to see. The imputed Pigouvian tax uses those distorted returns to obtain the present value of climate impacts, and thus identifies a wrong cost-benefit ratio for the current emissions; this links with the well-know result in cost-benefit analysis that the distorted capital returns do not identify the correct social returns for public investments (Lind, 1982; Dasgupta, 2008). The true return on climate policies is higher if the climate asset is sufficiently persistent; the equilibrium carbon price formula incorporates the social value of this persistence.

We ask next if there are potential gains to be achieved from enforcing the pricing rule in Proposition 2 as an institutional constraint. Strong welfare conclusions can be obtained for this model, if we treat agents in different periods as distinct generations (as in Phelps and Pollak, 1968). Then, the multi-generation Pareto optimality is a natural welfare concept (as, e.g., in Caplin and Leahy, 2004) for considering whether policy measures can improve welfare above that in the Markov equilibrium.[24] We provide such a comprehensive welfare analysis for an equivalent model in Gerlagh&Liski (2011); here we bring the essence of the welfare impacts.[25]

The requirement that emissions should follow the rule $f_{t,z} = \tau_t^{\gamma}$ has a seemingly clear justification: it implements efficiency. Only if carbon prices imputed to consistent preferences decisions with discounting $\gamma = \gamma^*$ are imposed, it is not possible to increase utility at any t without decreasing utility at some $t' \neq t$.[26] Yet such a carbon price rule does not imply Pareto optimality; not even a Pareto improvement can be achieved as we will now demonstrate.[27] Note that from the perspective of the current generation, future savings and emission levels are optimal if they are consistent with the long-term time preference δ, that is, if $g = \alpha\delta$ and $h^{\gamma=\delta}$ where h^{γ} is defined in Proposition 2; then future agents would behave as if they were consistent with present long-term preferences. This thought-experiment gives a clear benchmark against which we can test how policy proposals affect current welfare through future policies.

Lemma 3 *For $\beta \neq 1$ and $\tau > t$,*

$$\begin{aligned} \frac{\partial w_t}{\partial g_\tau} &> 0 \textit{ iff } g_\tau < \alpha\delta \\ \frac{\partial w_t}{\partial h_\tau} &> 0 \textit{ iff } h_\tau < h^\delta. \end{aligned}$$

Since the equilibrium policies depart from those optimal for the long-run preference δ, any policy that manages to take the decision variables closer to the long-run optimal levels increases current welfare. It turns out that imposing the stand-alone Pigouvian carbon tax principle implies a correction in the wrong direction.

---

[24]See Bernheim and Rangel (2009) for an alternative concept, and its relationship to the Pareto criterion. The Pareto criterion may not be reasonable when the focus is on the behavioral anomalies at the individual level.

[25]For completeness, these results are reproduced for the current climate-economy model in our longer working paper version Gerlagh&Liski (2012).

[26]It is interesting to note that while in the Markov equilibrium the decision-makers internalize all future impacts of current actions, the equilibrium is observationally distinct from any planner's optimum, unless the Pigouvian rule is imposed exogenously. This contrasts Barro (1999) where observational equivalence with an allocation chosen by a fictitious planner follows without restrictions on the actions of the decision-makers with hyperbolic preferences. This shows that when there is more than one capital-good the observational-equivalence does not hold in general. See Gerlagh&Liski (2011).

[27]In a different context, Bernheim and Ray (1987) also show that, in the presence of altruism, efficiency does not imply Pareto optimality.

Proposition 4 *For slow climate change, implementing $f_{t,z} = \tau_t^{\gamma}$ in Proposition 3 from period t onwards implies a welfare loss for generation t vis-a-vis the Markov equilibrium.*

Proof. By Lemma 1, the change to the imputed Pigouvian price does not affect policy g; thus, we can focus on the change in current welfare w_t due to the effect of future carbon prices. Let $\beta < 1$ so that $\beta\delta < \gamma < \delta$, and let climate change be a slow process such that $\tau_t^{\delta} > f_{t,z} > \tau_t^{\gamma}$; see Proposition 3. Imposing the imputed carbon price will then decrease the future carbon price, taking it further away from τ_t^{δ}, decreasing current welfare as shown in Lemma 3. The same mechanism applies for $\beta > 1$, when we have $\tau_t^{\delta} < f_{t,z} < \tau_t^{\gamma}$. Moreover, imposing the imputed carbon prices on current policies implies a deviation from the current best response, so the present welfare decreases due to changes both in today's and future's actions. ■

The remarkable feature of the above proposition is that the carbon pricing policy guided by the economy's aggregate statistics strictly decreases welfare, not as a second-order effect, but as a first-order effect.[28]

2.5 Quantitative assessment

To evaluate the quantitative significance of the conceptual results, we exploit the closed-form price formulas — given the Markov structure for policies, the initial carbon price level is a function of the income level and the carbon cycle parameters. Reasonable choices for the climate-economy parameters and consistent preferences ($\beta = 1$) can reproduce the carbon price levels of the more comprehensive climate-economy models such as DICE (Nordhaus, 2007).[29] We then introduce a difference between short- and long-term discounting, $\beta < 1$, while keeping savings decisions unaltered. The exercise shows how the sensitivity to climate outcomes can be reconciled with a positive description of the macroeconomy.

The model is decadal (10-year periods),[30] and year '2010' corresponds to period 2006-2015. We calibrate the damage parameter $\Delta_y = 0.003$ so that 2.7 per cent of output

[28]Lemma 3 suggests that we can achieve self-enforcing and welfare-improving policies in the infinite horizon setting. Such advanced policies are based on non-stationary strategies; see our working paper Gerlagh&Liski (2012).

[29]We can also reproduce the carbon tax time path of DICE when the energy sector of our model is specified and calibrated in detail; see our working paper Gerlagh&Liski (2012).

[30]The period length could be longer, e.g., 20-30 years to better reflect the idea that the long-term discounting starts after one period for each generation. We have these results available on request.

loss follows at a temperature rise of 3 Kelvin, as in Nordhaus (2001).[31] We set $\Delta_u = 0$. Population in 2010 is set at 6.9 [billion]. We take the Gross Global Product as 600 Trillion Euro [$Teuro$] for the first decade, 2006-2015 (World Bank, using PPP). The capital elasticity α follows from the assumed time-preference structure β and δ, and observed historic gross savings g. As a base-case, we consider net savings of 25% ($g = .25$), and a 2 per cent annual pure rate of time preference ($\beta = 1$,$\delta = 0.817$), resulting in $\alpha = g/\rho = 0.306$.

These parameter choices together with our carbon cycle result in a consistent-preferences Pigouvian, i.e., efficient carbon price, of 8.4 Euro/tCO_2, equivalent to 40 USD/tC, for 2010.[32] This number is very close to the level found by Nordhaus.[33] Consider then the determinants of this number in detail.

We can decompose the carbon price (13) into three contributing parts. First, consider the one-time costs if damages were immediate (ID) but only for one period,

$$ID = \Delta\pi(1 - g)y_t,$$

This value is multiplied by a factor to correct for the persistence of climate change, the persistence factor (PF),

$$PF = \sum\nolimits_{i \in \mathcal{I}} \frac{a_i}{[1 - \delta(1 - \eta_i)]},$$

which we then multiply by a factor to correct for the delay in the temperature adjustment, the delay factor (DF),

$$DF = \frac{\beta\delta\varepsilon}{1 - \delta(1 - \varepsilon)}.$$

Table 2 below presents the decomposition of the carbon tax for a set of short- and long-term discount rates such that the economy's macroeconomic statistics remain the same. The first row reproduces the efficient carbon price case assuming consistent preferences when the annual utility discount rate is set at 2 per cent: this row presents the carbon price under the same assumptions as in Nordhaus (2007). Keeping the equilibrium time-preference rate at 2, thus maintaining the savings rate at a constant level (reported also in Table 1 of the Introduction), we move to the Markov equilibrium by departing the short- and long-term discount rates, presented in the first and second columns.[34]

[31] In Appendix, to obtain the linear relationship between damages and carbon concentrations, we define D_t to be the global mean temperature squared (GMTS). Thus, when $GMTS = 3^2 = 9 = D_t$, $\Delta_y = 0.003$ implies about 2.7 per cent output loss.

[32] Note that 1 tCO2 = 3.67 tC, and 1 Euro is about 1.3 USD.

[33] Minor differences are caused by a correction for the price index, and a somewhat more persistent

	annual discount rate						
	short-term	long-term	equilibrium	ID	PF	DF	Carbon price
"Nordhaus/Pigou"	.02	.02	.02	7.61	2.44	.45	8.4
Equilibrium	.0235	.01	.02	7.89	3.70	.55	16.1
Equilibrium	.0255	.005	.02	8.05	5.79	.63	29.5
Equilibrium	.0271	.001	.02	8.19	19.55	.73	116.4
"Stern"	.001	.001	.001	8.19	19.55	.95	151.8

Table 2: Decomposition of the carbon price, MCP [Euro/tCO2]. ID=immediate costs, PF=persistence factor, DF=delay factor, $MCP = ID \times PF \times DF$. Parameter values in text.

We obtain a radical increase in the carbon price as the long-term discounting decreases, while savings remain unchanged from one set of preferences to the next. Note that by construction the Nordhaus number 8.4 EUR/tCO_2 becomes the imputed Pigouvian and thus the non-optimal tax for the hyperbolic discounting cases; because all equilibria have the same equilibrium discount rate, the imputed tax remains constant. The highest equilibrium carbon tax, 116.4 EUR/tCO_2, corresponds to the case where the long-run discounting is as proposed by Stern (2006); this case also best matches the Weitzman's values. For reference, we report the Stern case where the long-term discounting holds throughout, the carbon price takes a value of 151.8 EUR/tCO_2, and gross savings cover about 30 per cent of income. Thus, the Markov equilibrium closes considerably the gap between Stern's and Nordhaus' carbon prices, without having unrealistic by-products for the macroeconomy.[35]

damage structure in our reduced model

[34] Weitzman's (2001) survey led to discount rates declining from 4 per cent for the immediate future (1-5 years) to 3 per cent for the near future (6-25 years), to 2 per cent for medium future (26-75 years), to 1 per cent for distant future (76-300), and then close to zero for far-distant future. Roughly consistent with Weitzman and our 10-year length of one period, we use the short-term discount rate close to 3 per cent, and the long-term rate at or below 1 per cent. This still leaves degrees of freedom in choosing the two rates $\beta\delta$ and δ; we choose them to match the savings rate of 25 per cent and thus the macroeconomic performance in Nordhaus (2007). That is, we choose β and δ to maintain the equilibrium utility discount factor at $\gamma = 0.817$ (2 per cent annual discount rate). Since the equilibrium utility discount rate remains at 2 per cent, the macroeconomy remains observationally equivalent to that in Nordhaus ($g = .25$).

[35] The deviation between the Markov (thus Nordhaus) and Stern savings can be made extreme by sufficiently increasing the capital share of the output that gives the upper bound for the fraction of y_t saved; close to all income is saved under Stern preferences as this share approaches unity (Weitzman,

The decomposition shows that leaving out the time lag between CO_2 concentrations and the temperature rise amounts to replacing the column DF by $\beta\delta$: abstracting from the delay in temperature adjustments, as in Golosov et al. (2011), almost doubles the efficient carbon price level. But, as expected, the persistence of impacts, capturing the commitment value of climate policies, contributes most to the deviation between the efficient and Markov equilibrium prices.

Table 2 quantifies the economic substance of appropriately accounting for the distortions in the economy's aggregate statistics: when the long-run discount rate declines, the future equilibrium saving rate falls below the one the current generation would like to see. The greater is this distortion, the larger is the gap between the equilibrium, currently optimal, and the imputed Pigouvian tax.

3 Discussion

To obtain transparent analytical and quantitative results in a field that has been dominated by simulation models, we exploit strong functional assumptions. First, building on Golosov et al. (2011) we assume that income and substitution effects in consumption choices over time cancel out, leading to policies for savings and carbon prices that are separable. For more general functional forms, climate policies generate income effects influencing future savings, thereby creating deeper linkages between the two policies.[36] The quantitative importance of the interdependencies between the policies are probably best analyzed using a numerical approach. But, while the quantitative significance depends on the functional assumptions, the main observation is general: the optimal policy should correct for the return distortions that arise from time-declining discount rates.

Second, quasi-hyperbolic discount factors are only rough approximations for the discount rate paths estimated in the literature (e.g., Weitzman, 2001). It is possible to solve this model for an arbitrary sequence of discount factors. Again, this will affect the quantitative evaluations but not the essence of the carbon pricing formulas; Iverson (2012) partly builds on our setting to elaborate the implications of more flexible discounting.

Third, there is a concern that our linearized model for carbon diffusion might not well describe the relevant dynamics when the system is far off the central path — that is, nonlinearities captured by more complicated climate simulation models may be important. To address this concern, we devised a Monte Carlo experiment to test consistency of our

2007). However, with reasonable parameters such extreme savings do not occur, as in Table 2.

[36]We explicate these effects in the longer working paper version Gerlagh&Liski (2012, Section 2.2)

closed-form carbon price with that predicted by a benchmark simulation model, DICE 2007. Drawing parameters from distributions for all key parameters in DICE, including those that appear in our formula as well as those not in our formula, we found that the formula explains 99% of the DICE variation in the carbon price. The deviation due to the non-linearities of climate change is inconsequential compared to variation that is captured by the formula.[37]

Fourth, our reduced-form carbon cycle and damage representations assume no uncertainty, although great uncertainties describe both the climate system parameters as well as the impacts of climate change on our economies. Golosov et al. (2011) show that optimal polices are robust to impact uncertainty - rewriting the carbon price formula in expected terms. Iverson (2012) shows robustness of the Markov equilibrium policy rules, and of the gap between optimal carbon prices and efficient carbon prices, for a stochastic Markov equilibrium with multiple stochastic parameters.

The tractable climate-economy models can prove very useful in the further analysis, especially for addressing the consequences of uncertainty and learning. Since the closed-form carbon price formula captures well the essence of computational climate-economy models, it allows a transparent mapping from climate system and normative uncertainties to carbon price distributions and, thereby, it potentially offers a sharp disentanglement of subjective and objective determinants of carbon prices.

4 Concluding remarks

September 2011, the U.S. Environmental Protection Agency (EPA) sponsored a workshop to seek advice on how the benefits and costs of regulations should be discounted for projects with long horizons; that is, for projects that affect future generations. The EPA invited 12 academic economists to address the following overall question: "What principles should be used to determine the rates at which to discount the costs and benefits of regulatory programs when costs and benefits extend over very long horizons?" In the background document, the EPA prepared the panelists for the question as follows: "Social discounting" in the context of policies with very long time horizons involving multiple generations, such as those addressing climate change, is complicated by at least three factors: (1) the "investment horizon" is significantly longer than what is reflected in observed

[37]The online Appendix summarizes the experiment, and contains a note on the surprising prediction power of the reduced form carbon pricing formula for the more comprehensive simulation model results. See www.hse-econ.fi/liski/.

interest rates that are used to guide private discounting decisions; (2) future generations without a voice in the current policy process are affected; and (3) compared to shorter time horizons, intergenerational investments involve greater uncertainty. Understanding these issues and developing methodologies to address them is of great importance given the potentially large impact they have on estimates of the total benefits of policies that impact multiple generations."

In this paper, we have developed a methodology for addressing the over-arching question posed above and a quantitative evaluation. The resulting tool for policy purposes is a carbon pricing formula that compresses the relevant elements of the climate and the economy — while it is not a substitute for the comprehensive climate-economy models, the formula identifies the contributions of the key elements to optimal carbon prices and allows discussing them transparently. The formula incorporates the practical program evaluation principle that the time-discounting rate should depend on the time horizon considered. In general equilibrium, which is the approach needed for climate policy evaluations, time-changing discount rates distort the economy, stipulating a correction to carbon pricing. The formula allows policy-makers to experiment with their prescriptive views on longer-term discounting to see the effect on the optimal carbon tax. We used discount factors from the literature to show that the equilibrium correction above the standard efficient carbon tax is significant, and that this requires no loss of descriptive realism regarding the economy.

References

[1] Barro R.J. (1999), Ramsey meets Laibson in the neoclassical growth model, Quarterly Journal of Economics 114: 1125-1152.

[2] Bernheim, D.B., and A. Rangel (2009), Beyond Revealed Preference: Choice-Theoretic Foundations for Behavioral Welfare Economics, The Quarterly Journal of Economics, MIT Press, vol. 124(1), pages 51-104.

[3] Bernheim, D.B., and D. Ray (1987), Economic growth with intergenerational altruism, The Review of Economic Studies, Vol. 54, No. 2: 227-241.

[4] Boden, T.A., G. Marland, and R.J. Andres. 2011. Global, Regional, and National Fossil-Fuel CO2 Emissions. Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, U.S. Department of Energy, Oak Ridge, Tenn., U.S.A.

[5] Brock, W. A., and Mirman, L. J. (1972), Optimal economic growth and uncertainty: The discounted case, Journal of Economic Theory, Elsevier, vol. 4(3), pages 479-513, June.

[6] Caldeira K. and M. Akai (eds) 2005, Ocean storage, Ch 6 in IPCC special report on carbon dioxide capture and storage, edited by Metz B., O. Davidson, H. de Coninck, M. Loos, and L. Meyer, Cambridge University Press.

[7] Caplin A, and Leahy J (2004), The social discount rate, Journal of Political Economy 112: 1257-1268.

[8] Dasgupta, P., Discounting Climate Change, Journal of Risk and Uncertainty, 2008, 37(2-3), 141-169.

[9] Fujii T., and L. Karp (2008), Numerical analysis of non-constant pure rate of time preference: A model of climate policy, J. of Environmental Economics and Management 56: 83-101.

[10] Gerlagh, R., and M. Liski, Public investment as commitment, CESifo Working Paper Series No. 3330, available at hse-econ.fi/liski/papers

[11] Gerlagh, R., and M. Liski (June 28, 2012), Carbon prices for the next thousand years, CESifo Working Paper Series No. 3855. Available at hse-econ.fi/liski/papers/

[12] Gerlagh, R., and M. Liski (2013), No news is good news in climate change?, mimeo, Available at hse-econ.fi/liski/papers/

[13] Greenstone, M., E. Kopits, A. Wolverton (2011), Estimating the Social Cost of Carbon for Use in U.S. Federal Rulemakings: A Summary and Interpretation, NBER Working Paper No. 16913.

[14] Gollier, C., and R. Zeckhauser (2005), Aggregation of Heterogeneous Time Preferences, Journal of Political Economy, 2005, vol. 113, issue 4, ppp. 878-896.

[15] Golosov, M., J. Hassler, P. Krusell, A. Tsyvinski (2011), Optimal taxes on fossiel fuel in general equilibrium, NBER WP 17348.

[16] Goulder, L.H, and R.C. Williams III (2012), The choice of discount rate for climate policy change policy evaluation, NBER Working Paper No. 18301.

[17] Hasselmann, K., S. Hasselmann, R Giering, V Ocana, H v Storch (1997), Sensitivity of optimal CO_2 emissions paths using a simplified structural integrated assessment model (SIAM), Climatic Change 37: 345-387

[18] Hassler, J., P. Krusell, and C. Olovsson (2011), Energy-saving technical change, mimeo.

[19] Hooss G, R. Voss, K Hasselmann, E Maier-Reimer, F Joos (2001), A nonlinear impulse response model of the coupled carbon cycle-climate system (NICCS), Climate Dynamics 18: 189-202

[20] Houghton, R.A. 2003.Revised estimates of the annual net flux of carbon to the atmosphere from changes in land use and land management 1850-2000. Tellus55B(2):378-390.

[21] IPCC, Intergovernmental Panel on Climate Change (2000), Special Report on Emissions Scenarios, edited by N. Nakicenovic and R. Swart, Cambridge Univ. Press, Cambridge, U.K.

[22] Iverson, T. (Dec 13, 2012), Optimal Carbon Taxes with Non-Constant Time Preference, Munich Personal RePEc Archive, Working paper 43264.

[23] Jackson, O. M., and L. Yariv, Collective Dynamic Choice: The Necessity of Time Inconsistency, WP-2011

[24] Kaplow L., E. Moyer, and D. A. Weisbach (2010) ÒThe Social Evaluation of Intergenerational Policies and Its Application to Integrated Assessment Models of Climate Change,Ó The B.E. Journal of Economic Analysis and Policy: Vol. 10: Iss. 2 (Symposium), Article 7.

[25] Karp L. (2005), Global warming and hyperbolic discounting, Journal of Public Economics 89: 261-282.

[26] Karp L. (2007), Non-constant discounting in continuous time, Journal of Economic Theory 132: 557-568.

[27] Karp, L., and Y. Tsur (2011), Time perspective and climate change policy, Journal of Environmental Economics and Management 62, 1-14.

[28] Krieglera, Hall, Helda, Dawson, and Schellnhuber (2009), “Imprecise probability assessment of tipping points in the climate system”, Proceedings of the National Academy of Sciences, vol 106: 5041-5046.

[29] Krusell P., Kuruscu B, Smith A.A. (2002), Equilibrium welfare and government policy with quasi-geometric discounting, Journal of Economic Theory 105: 42-72.

[30] Krusell P. and A.A. Smith (2003), Consumption-savings decisions with quasi-geometric discounting, Econometrica 71: 365-375.

[31] Laibson D. (1997), Golden eggs and hyperbolic discounting, Quarterly Journal of Economics 112: 443-477.

[32] Lind, R. C. (1982), ”A Primer on the Major Issues Relating to the Discount Rate for Evaluating National Energy Options: Discounting for Time and Risk in Energy Policy, RC Lind, ed., Johns Hopkins University Press, Washington, DC.”

[33] Layton, D.F., and G. Brown, Heterogeneous Preferences Regarding Global Climate Change, The Review of Economics and Statistics, Vol. 82, No. 4 (Nov., 2000), pp. 616-624.

[34] Maier-Reimer E. and K.Hasselman (1987), Transport and storage of CO2 in the ocean - an inorganic ocean-circulation carbon cycle model, Climate Dynamics 2, 63-90.

[35] Manne A.S., R.Mendelsohn, and R.Richels (1995), MERGE, A model for evaluating regional and global effects of GHG reduction policies, Energy Policy 23, 17-34.

[36] Nordhaus, W. D.(1993), Optimal greenhouse-gas reductions and tax policy in the ”DICE” model, The American Economic Review, 83 (2), pp. 313-317.

[37] Nordhaus, W.D. (1997), Discounting in economics and climate change, an editorial, in Climatic Change 37: 315-328.

[38] Nordhaus, W.D. (1999), Discounting and public policies that affect the distant future, in: P.R. Portney, J.P. Weyant (Eds.), Discounting and Intergenerational Equity, Resources for the Future, Washington, DC.

[39] Nordhaus, W. D.(2007), A Review of The Stern Review on the Economics of Climate Change, Journal of Economic Literature, 45 (3), 686-702.

[40] Maskin, E., and J. Tirole. 2001. “Markov Perfect Equilibrium I: Observable Actions”, Journal of Economic Theory 100, 191-219.

[41] Mastrandrea, M.D., and S.H. Schneider, Integrated assessment of abrupt climatic changes, Climate Policy 1 (2001) 433-449.

[42] Muller, N.Z, R. Mendelsohn, and W. Nordhaus (2011), Environmental accounting for pollution in the United States economy, American Economic Review, 101 (5): 1649-75

[43] Peck S.C. and T.J.Teisberg (1992), CETA: a model for carbon emissions trajectory assessment, Energy Journal 13, 55-77.

[44] Phelps E.S. and R.A. Pollak (1968), On second-best national saving and game-equilibrium growth, Review of Economic Studies 35(2): 185-199.

[45] Rubinstein, A. (2003), Economics and psychology? The case of hyperbolic discounting, International Economic Review 44: 1207-1216.

[46] Saez-Marti M. and J.W. Weibull (2005), Discounting and altruism to future decision makers, Journal of Economic Theory 122: 254-266

[47] Solomon, S., D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M. Tignor and H.L. Miller (eds), ”Climate Change 2007: The Physical Science Basis”, Technical Summary, Table TS2.5, footnote b.

[48] Stern, N. (2006), ”The economics of climate change: the Stern review”, Cambridge, UK: Cambride University Press.

[49] Tol, R. (2009), The economic effects of climate change, Journal of economic perspectives, 23(2): 29-51.

[50] Weitzman, M. (2001). Gamma Discounting, American Economic Review, Vol. 91, No. 1, pp. 260-271.

[51] Weitzman, M. (2007). A Review of The Stern Review on the Economics of Climate Change, Journal of Economic Literature, 45 (3), 703-724.

Appendix

Lemma 1

The proof is by induction. Assume that (9) and (10) hold for all future periods $t+1, t+2, ...$, and that the lemma holds for $t+2$. We can thus construct the value function for the next period, as

$$W_{t+1}(k_{t+1}, s_{t+1}) = u_{t+1} + \delta W_{t+2}(k_{t+2}, s_{t+2}).$$

Substitution of the investment decision at time $t+1$, $k_{t+2} = gy_{t+1}$ and emissions $z_{t+1} = z^*_{t+1}$, gives

$$\begin{aligned} W_{t+1}(k_{t+1}, s_{t+1}) \quad = \quad & [\ln(1-g_{t+1}) + \ln(A_{t+1}) + \alpha \ln(k_{t+1}) + \ln(\omega(s_{t+1}))] - \Delta_u D_{t+1} \\ & + \delta \widetilde{A}_{t+2} + \delta\xi[\ln(g_{t+1}) + \ln(A_{t+1}) + \alpha \ln(k_{t+1}) + \ln(\omega(s_{t+1}))] + \delta\Omega(s_{t+2}) \end{aligned}$$

Collecting the coefficients that only depend on future policies g_τ and z_τ for $\tau > t$, and that do not depend on the next-period state variables k_{t+1} and s_{t+1}, we get the constant part of $V_{t+1}(k_{t+1})$:

$$\widetilde{A}_{t+1} = \ln(1-g_{t+1}) + \delta\xi \ln(g_{t+1}) + (1+\delta\xi)\ln(A_{t+1}) - \delta\zeta_1 z_{t+1} + \delta\widetilde{A}_{t+2}. \tag{18}$$

Collecting the coefficients in front of $\ln(k_{t+1})$ yields the part of $V_{t+1}(k_{t+1})$ depending k_{t+1} with the recursive determination of ξ,

$$\xi = \alpha(1+\delta\xi).$$

so that $\xi = \frac{\alpha}{1-\alpha\delta}$ follows.

Collecting the terms with s_{t+1} yields $\Omega(s_{t+1})$ through

$$\Omega(s_{t+1}) = \ln(\omega(s_{t+1}))(1+\delta\xi) - \Delta_u D_{t+1} + \delta\Omega(s_{t+2}).$$

where $z_{t+1} = z^*_{t+1}$ appearing in $s_{t+2} = (z_1, ...z_t, z_{t+1})$ is independent of k_{t+1} and s_{t+1} (by Lemma 1 that holds by the induction hypothesis) so that we only need to consider the values for $z_1, ..., z_t$ when evaluating $\Omega(s_{t+1})$. The values for ζ_τ can be calculated by collecting the terms in which $z_{t+1-\tau}$ appear. Recall that $\ln(\omega(s_{t+1})) = -\Delta_y D_{t+1}$ so that

$$\zeta_\tau = ((1+\delta\xi)\Delta_y + \Delta_u)\sum\nolimits_{(i,j)} a_i b_j \pi\varepsilon_j \frac{(1-\eta_i)^\tau - (1-\varepsilon_j)^\tau}{\varepsilon_j - \eta_i} + \delta\zeta_{\tau+1}$$

Substitution of the recursive formula, for all subsequent τ, gives

$$\zeta_\tau = (\frac{\Delta_y}{1-\alpha\delta} + \Delta_u)\sum\nolimits_{(i,j)}\sum\nolimits_{t=\tau}^{\infty} a_i b_j \pi\varepsilon_j \delta^{t-\tau} \frac{(1-\eta_i)^t - (1-\varepsilon_j)^t}{\varepsilon_j - \eta_i}$$

To derive the value of ζ_1, we consider

$$\begin{aligned}
& \sum\nolimits_{t=1}^{\infty} \delta^{t-1} \frac{(1-\eta_i)^t - (1-\varepsilon_j)^t}{\varepsilon_j - \eta_i} \\
= & \frac{\sum\nolimits_{t=1}^{\infty} [\delta(1-\eta_i)]^t - \sum\nolimits_{t=1}^{\infty} [\delta(1-\varepsilon_j)]^t}{\delta(\varepsilon_j - \eta_i)} \\
= & \frac{\frac{\delta(1-\eta_i)}{1-\delta(1-\eta_i)} - \frac{\delta(1-\varepsilon_j)}{1-\delta(1-\varepsilon_j)}}{\delta(\varepsilon_j - \eta_i)} \\
= & \frac{1}{[1-\delta(1-\eta_i)][1-\delta(1-\varepsilon_j)]}
\end{aligned}$$

Finally, we notice that a careful examination shows that the final equation still holds when $\eta_i = \varepsilon_i$, even though we then cannot follow the same derivation. Q.E.D.

Proposition 1

The first-order conditions for fossil-fuel use z_t, and the labor allocations over the final goods $l_{y,t}$ and the energy sectors $l_{e,t}$ give:

$$u_t' \frac{\partial y_t}{\partial z_t} = \beta\delta \frac{\partial \Omega_{t+1}}{\partial s_{t+1}} \frac{\partial s_{t+1}}{\partial z_t} \Rightarrow \frac{1}{1-g} \frac{1}{A_t} \frac{\partial A_t}{\partial e_t} \frac{\partial E_t}{\partial z_t} = \beta\delta\zeta_1 \tag{19}$$

$$\frac{\partial A_t}{\partial l_{y,t}} = \frac{\partial A_t}{\partial e_t} \frac{\partial E_t}{\partial l_{e,t}} \tag{20}$$

The second part of the Lemma follows immediately from (19):

$$\frac{\partial y_t}{\partial z_t} = f_{t,z} = \beta\delta\zeta_1(1-g)y_t.$$

The second equation equates the marginal product of labor in the final good sector with the indirect marginal product of labor in energy production. We have thus four equations, energy production (4), labour market clearance (5), and the above two first-order conditions, that jointly determine four variables: $z_t, l_{y,t}, l_{e,t}, e_t$, only dependent on technology at time t through $A_t(l_{y,t}, e_t)$ and $E_t(z_t, l_{e,t})$. Thus, $z_t = z_t^*$ can be determined independent of the state variables k_t and s_t.Q.E.D.

Proposition 2

To determine the efficient carbon tax as the net present value of marginal damages, note that from the emissions-damage response function we have

$$\begin{aligned}\frac{du_{t+\tau}}{dz_t} &= (\frac{\Delta_y}{1-g}+\Delta_u)\frac{dD_{t+\tau}}{dz_t} \\ &= (\frac{\Delta_y}{1-g}+\Delta_u)\sum\nolimits_i a_i b_j \pi \varepsilon_j \frac{(1-\eta_i)^\tau-(1-\varepsilon_j)^\tau}{\varepsilon_j-\eta_i}\end{aligned}$$

Furthermore, we note that the marginal rate of substitution for utility, between two periods, in equilibrium, is γ, so that the net present value of future damages associated with one extra unit of emissions, in current utility terms, h^{Pig}, is given by

$$\begin{aligned}h^{Pig} &= \sum\nolimits_{\tau=1}^{\infty}\gamma^\tau \frac{du_{t+\tau}}{dz_t} \\ &= (\frac{\Delta_y}{1-g}+\Delta_u)\sum\nolimits_i \frac{a_i b_j \pi \varepsilon_j}{\varepsilon_j-\eta_i}\sum\nolimits_{\tau=1}^{\infty}\gamma^\tau(1-\eta_i)^\tau-\gamma^\tau(1-\varepsilon_j)^\tau \\ &= (\frac{\Delta_y}{1-g}+\Delta_u)\sum\nolimits_i a_i b_j \pi \varepsilon_j \gamma \sum\nolimits_{\tau=0}^{\infty}\gamma^\tau(1-\eta_i)^\tau-\gamma^\tau(1-\varepsilon_j)^\tau \\ &= (\frac{\Delta_y}{1-g}+\Delta_u)\sum\nolimits_i \frac{\gamma\pi a_i b_j \varepsilon_j}{[1-\gamma(1-\eta_i)][1-\gamma(1-\varepsilon_j)]} \\ &= h^\gamma\end{aligned}$$

Q.E.D.

Proposition 3

We consider the ratio between the carbon price and the efficient carbon price for very long climate change delays, $\eta_i = \varepsilon_j = 0$, and quasi-hyperbolic preferences, $\beta < 1$:

$$\begin{aligned}\frac{f_{t,z}}{\tau_t^\gamma} &= \frac{(1-\gamma)^2}{(1-\delta)^2}\frac{\beta\delta}{\gamma} \\ &= \frac{\left(1-\frac{\beta\delta}{1-\alpha\delta+\alpha\beta\delta}\right)^2}{(1-\delta)^2}(1-\alpha\delta+\alpha\beta\delta) \\ &= \frac{(1-\delta(\alpha+(1-\alpha)\beta))^2}{(1-\delta)^2(1+\alpha\delta(\beta-1))} \\ &> 1\end{aligned}$$

The first equality follows from substitution of $\eta_i = \varepsilon_j = 0$ in the equation for the equilibrium carbon price and efficient carbon price. The second equality substitutes the value for γ. The final inequality follows as for $\beta < 1$, we have that $\alpha + (1-\alpha)\beta < 1$, and thus the numerator exceeds $1-\delta$, while $\beta < 1$ also ensures that the second term in the denominator falls short of 1. Q.E.D.

Damages and carbon cycle

The carbon reservoirs contain physical carbon stocks measured in Teratons of carbon dioxide [$TtCO_2$]. These quantities are denoted by a $n \times 1$ vector $L_t = (L_{1,t}, ..., L_{n,t})$. In each period, share b_j of total emissions z_t enters layer j, and the shares sum to 1. The diffusion between the layers is described through a $n \times n$ matrix $\mathbf{M}$ that has real and distinct eigenvalues $\lambda_1, ..., \lambda_n$. Dynamics satisfy

$$L_{t+1} = \mathbf{M}L_t + bz_t.$$

No CO_2 leaves the system, so that row elements of $\mathbf{M}$ sum to one. Using the eigen-decomposition theorem of linear algebra, we can define the linear transformation of co-ordinates $H_t = \mathbf{Q}^{-1}L_t$ where $\mathbf{Q} = [\ v_1 \quad ... \quad v_n\ ]$ is matrix of linearly independent eigenvectors v_λ such that

$$\mathbf{Q}^{-1}\mathbf{M}\mathbf{Q} = \mathbf{\Lambda} = \mathbf{diag}[\lambda_1, ..., \lambda_n].$$

We obtain

$$\begin{aligned} H_{t+1} &= \mathbf{Q}^{-1}L_{t+1} = \mathbf{Q}^{-1}\mathbf{M}\mathbf{Q}H_t + \mathbf{Q}^{-1}bz_t \\ &= \mathbf{\Lambda}H_t + \mathbf{Q}^{-1}bz_t, \end{aligned}$$

which enables us to write the (uncoupled) dynamics of the vector H_t as

$$H_{i,t+1} = \lambda_i H_{i,t} + c_i z_t$$

where λ_i are the eigenvalues, and $c = \mathbf{Q}^{-1}b$. This defines the vector of climate units (boxes) H_t that have independent dynamics but that can be reverted back to L_t to obtain the original physical interpretation.

For the calibration, we consider only three climate system layers: atmosphere and upper ocean layer ($L_{1,t}$), biomass ($L_{2,t}$), and deep oceans ($L_{3,t}$). For the greenhouse effect, we are interested in the total atmospheric CO_2 stock. Layer $L_{1,t}$ contains both atmosphere and upper ocean carbon that almost perfectly mix within one period of ten years; we can find the atmospheric stock by correcting for the amount that is stored in the upper oceans. Let μ is the amount of CO_2 stored in the upper ocean layer, relative to the amount in the atmosphere. Then, the total the atmospheric CO_2 stock is

$$S_t = \frac{L_{1,t}}{1 + \mu}.$$

Let $\sum_i q_{1,i}$ denote the first row of $\mathbf{Q}$. Then, we can solve for the development of the atmospheric CO_2 as

$$S_t = \frac{\sum_i q_{1,i} H_{i,t}}{1+\mu}$$

This allows the following breakdown: Redefine $S_{i,t} = \frac{q_{1,i}}{1+\mu} H_{i,t}$, $a = \frac{q_{1,i}}{1+\mu}\mathbf{Q}^{-1}b$, and $\eta_i = 1 - \lambda_i$, to obtain

$$\begin{aligned} S_{i,t+1} &= (1-\eta_i)S_{i,t} + a_i z_t \\ S_t &= \textstyle\sum_{i \in \mathcal{I}} S_{i,t}. \end{aligned}$$

This is now a system of carbon stocks where depreciation factors are defined by eigenvalues of the original physical representation. Notice that we know one eigenvalue $\lambda = 1$, as no carbon can leave the system. From this it follows that we have one box i with no depreciation, $\eta_i = 0$.[38]

We follow Hooss et al (2001, table 2) and assume an asymptotic climate sensitivity $\varphi(S_t)$ function that describes the pressure on temperatures caused increases in concentrations. Typically, the relationship between the asymptotic temperature sensitivity and the atmospheric CO_2 stock, $\varphi(S)$, is concave; the logarithmic form is frequently assumed. Constant $0 < \varepsilon < 0$ captures the adjustment speed in temperatures:

$$T_t = T_{t-1} + \varepsilon(\varphi(S_t) - T_{t-1})$$

In steady state, we have $T = \varphi(S)$, but elsewhere temperature T_t changes depending on the atmospheric CO_2 stock. Damages, in turn, are a function of the temperature

$$D_t = \psi(T_t)$$

where $\psi(T)$ is convex. To be explicit we assume that $\psi(T) = T^2$. It has been noted in the literature that in the relevant domain of atmospheric CO_2 concentrations between 400 and 1000 ppmv,[39] the composition of the typical convex damage and concave climate sensitivity functions returns an almost linear function through the origin:[40]

$$\psi(\varphi(S_t)) \approx \pi S_t$$

[38]Note also that if the model is run in almost continuous time, that is, with short periods so that most of the emissions enter the atmosphere, $b_1 = 1$, it follows that $\sum_i a_i = 1/(1+\mu)$. Otherwise, we have $\sum_i a_i < 1/(1+\mu)$.

[39]ppmv=parts per million by volume.

[40]Indeed, the early calculations by Nordhaus (1991) based on local linearization, are surprisingly close to later calculations based on his DICE model with a fully-fledged carbon-cycle temperature module, apart from changes in parameter values based on new insights from the natural science literature.

with $\pi > 0$, a constant sensitivity of damages to atmospheric CO_2.[41] Using the approximation, we can rewrite the damage dynamics directly as dependent on stocks:

$$D_t \;=\; D_{t-1} + \varepsilon_j(\pi S_t - D_{t-1}).$$

Given the two layers of climate variables — one for carbon stocks $S_{i,t}$, and the other for damages D_t — it is a straightforward matter of verification that future damages depend on past emissions as follows:

$$S_{i,t} \;=\; (1-\eta_i)^{t-1} S_{i,1} + \sum\nolimits_{\tau=1}^{t-1} a_i (1-\eta_i)^{\tau-1} z_{t-\tau} \tag{21}$$

$$\begin{aligned} D_t \;=\; & (1-\varepsilon)^{t-1} D_1 + \sum\nolimits_{i\in\mathcal{I}} \pi\varepsilon \frac{(1-\eta_i)^t - (1-\eta_i)(1-\varepsilon)^{t-1}}{\varepsilon-\eta_i} S_{i,1} + \\ & \sum\nolimits_{i\in\mathcal{I}} \sum\nolimits_{\tau=1}^{t-1} a_i \pi\varepsilon \frac{(1-\eta_i)^\tau - (1-\varepsilon)^\tau}{\varepsilon-\eta_i} z_{t-\tau}, \end{aligned} \tag{22}$$

where $S_{i,1}$ and D_1 are taken as given at $t = 1$, and then values for $t > 1$ are defined by the expressions. The model can be applied to a situation where some climate change has taken place at the start of time $t = 1$, so we write the system dependent on $S_{i,1}$,$D_1 > 0$ — however, interpreting $t = 1$ as the beginning of the industrial era, say 1850, we can set $S_{i,0} = D_0 = 0$. Collecting terms allows us to express D_t as in (7). This defines the emissions-damage function θ_τ in the text.

For calibration, we take data from Houghton (2003) and Boden et al. (2011) for carbon emissions in 1751–2008; the data and calibration is available in the online Appendix (www.hse-econ.fi/liski.fi). We calibrate the original multi-layer model parameters $\mathbf{M}$, b, μ, to minimize the error between the atmospheric concentration prediction from the 3-layer model and the Mauna Loa observations under the constraint that CO_2 stocks in the various layers and flows between the layers should be consistent with scientific evidence as reported in Fig 7.3 from the IPCC fourth assessment report from Working Group I (Solomon et. al. 2007). There are 4 parameters we calibrate. We set $b = (1, 0, 0)$ so that emissions enter the first layer. The matrix $\mathbf{M}$ has 9 elements. The condition that the rows sum to one removes 3 parameters. We assume no diffusion between the biosphere and the deep ocean, removing 2 other parameters. We fix the steady state share of the deep ocean at 4 times the atmospheric share. This leaves us with 3 elements of $\mathbf{M}$ to be calibrated, plus μ. In words, we calibrate: (1) the CO_2 absorption capacity of the "atmosphere plus upper ocean"; (2) the CO_2 absorption capacity of the biomass layer

[41] Multiplying the constants Δ_y and π gives the damage sensitivity: the asymptotic percentage loss of output per TtCO2 in the atmosphere. Inversely, $1/\Delta_y\pi$ is the amount of atmospheric CO_2 that leads to an asymptotic 63 per cent (e^{-1}) loss of output.

relative to the atmosphere, while we fix the relative size of the deep ocean layer at 4 times the atmophere, based on the IPCC special report on CCS, Fig 6.3 (Caldeira and Akai, 2005); (3) the speed of CO_2 exchange between the atmosphere and biomass, and (4) between the atmosphere and the deep ocean.

We transform this annual 3-layer model into a decadal layer model adjusting the exchange rates within a period between the layers and the shares of emissions that enter the layers within the period of emissions. Then, we transform the decadal 3-layer model into the decadal 3-box model, as described above. The transformed box model has no direct physical meaning other than this: box 0 measures the amount of atmospheric carbon that never depreciates; box 1 contains the atmospheric carbon with a depreciation of about 7 per cent in a decade; while carbon in box 2 depreciates 50 per cent per decade.[42] About 20 per cent of emissions enter either the upper ocean layer, biomass, or the deep ocean within the period of emissions. In the box representation, they do not enter the atmospheric carbon stock, so that the shares a_i sum to 0.8. Our procedure provides an explicit mapping between the physical carbon cycle and the reduced-form model for atmospheric carbon with varying deprecation rates; the Excel file available as supplementary material contains these steps and allows easy experimentation with the model parameters. The resulting boxes, their emission shares, and depreciation factors are as reported in the text.

Comparison of climate response functions

We compare our response function for damages, as percentage of output, resulting from emissions, with those in Nordhaus (2007) and Golosov et al. (2011). The GAMS source code for the DICE model provides a large variety of scenarios with different policies such as temperature stabilization, concentration stabilization, emission stabilization, the Kyoto protocol, a cost-benefit optimal scenario, and delay scenarios. For each of these scenarios we calculated the damage response function by simulating an alternative scenario with equal emissions, apart from a the first period when we decreased emissions by 1GtC. Comparison of the damages, in terms relative of output, then defines the response function for that specific scenario. It turns out that the response functions are very close, and we took the average over all scenarios. To interpret the response function in Nordhaus (2007), we notice that the average DICE carbon cycle and damage

[42] As explained above, the decay rates in the final model come from the eigenvalues of the original model.

response can very accurately be described by our reduced form using the parameters $a = (0.575, 0.395, 0.029)$, $\eta = (0.310, 0.034, 0)$, which give a perfect fit for the carbon cycle of DICE2007, and $\varepsilon = 0.183$, $\pi = 4.09$ for the temperature delay. That is, the carbon-cycle in DICE (Nordhaus 2007) is characterized by a very large long-term uptake of CO2 in the oceans. The reduced model in Golosov et al. is represented by $a = (0.2, 0.486, 0.314)$, $\eta = (0, 0.206, 1)$, which implies a similar carbon cycle model to ours, but Golosov et al. have no temperature delay structure, $\varepsilon = 1$. Figure 1 presents the emissions damage responses.