Carbon prices for the next thousand years

Reyer Gerlagh and Matti Liski*

December 12, 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 climate-economy model that resolves the puzzle. The optimal carbon price diverges from the imputed externality cost, obtained from the economy's aggregate statistics, when pure discounting is time-declining. The effect is quantitatively large enough to solve any conceivable "description puzzle" when the climate-system representation is explicit, with impacts peaking several decades after the date of emissions. We identify a "prescription puzzle"; policies based on discounting below the descriptive rate improve welfare. A solution to this puzzle is to sustain low discounting as a self-enforcing contract over time.

(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, Bård 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, NBER Summer Institute, and SURED meetings, and at seminars in Helsinki, LSE, Stockholm, Toulouse, Oxford, and Basel 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) and Layton and Levine (2003) used a survey of 376 non-economists, and found a small or no difference in the willingness to pay to prevent future climate change impacts appearing after 60 or 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

[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).

to zero after 300 years.[4]

To incorporate the equilibrium implications of declining time discounting, we build on a general-equilibrium growth framework, following Nordhaus' approach and its recent gearing towards the macro traditions by Golosov, Hassler, Krusell, and Tsyvinsky (2011). We develop and calibrate an explicit carbon cycle representation for climate impacts to derive a tractable carbon price formula under hyperbolic discounting. The formula allows a transparent quantitative assessment, and shows that the equilibrium carbon price deviates from the standard Pigouvian principle where the externality pricing is based on the economy's returns.

Table 1 contains the gist of the quantitative assessment. The model is calibrated to 25 per cent gross savings, when both the short- and long-term annual time discount rate is 2.7 per cent. This is consistent with Nordhaus' DICE 2007 baseline scenario (Nordhaus, 2007),[5] giving 7.1 Euros per ton of CO_2 as the optimal carbon price in the year 2010 (i.e., 34 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).[6]

	discount rate			
	short-term	long-term	savings	carbon price
"Nordhaus"	.027	.027	.25	7.1
Equilibrium	.037	.001	.25	133
"Stern"	.001	.001	.33	174

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

The Nordhaus' number is the first-best Pigouvian tax for the 2.7 per cent annual

[4]For revealed-preference evidence, Giglio, Maggiore, and Stroebel (2013) compare infinite-maturity ownership and leaseholds of various lengths in housing markets to infer long-run discount rates. The estimated rates are found to be considerably below those routinely assumed in economic analysis, yet higher than those in Stern (2006).

[5]Nordhaus uses an annual pure rate of time preference of 1.5 per cent; our value 2.7 is the equivalent number when adjusting for the difference in the consumption smoothing parameter, and labor productivity growth. See Nordhaus (2008) for detailed documentation of DICE 2007.

[6]Under "Stern" the capital-share of output is fully saved (33 per cent); increasing the capital-share leads to unrealistic savings as discussed, e.g., in Weitzman (2007) and Dasgupta (2008).

discount rate. For distinct short- and long-term discounting, the experiment keeps the economy observationally unchanged, so an outsider looking at the economy's aggregate statistics could conclude that the present-value of externality costs also remain unchanged. However, 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 the current policy maker would like to see.[7] As is well-known in cost-benefit analysis, a distorted capital return is not the social rate of return for public investments.[8] For this reason, the equilibrium carbon price is higher than the Nordhaus' number; 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 that solves the "description puzzle" comes from a Markov equilibrium where each generation sets its self-interested savings and climate policies understanding the future generations' policies — 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 more commitment than shorter-term capital savings: the current climate policies alter directly the utilities of far-future agents. This explains why climate investments are valued above the level implied by the externality pricing based on the economy's returns. Also, future agents have no reason to undermine past climate investments as they value the climate capital for the same reason.[9]

A Markov policy is descriptive; but more can be achieved by coordination over time. We define sustainable policies as savings and climate policy prescriptions that (i) apply

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

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

[9] Although the commitment problem is similar to that in Laibson (1997), 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 beneficiaries (see also Saez-Marti and Weibull, 2005).

symmetrically to all generations; (ii) are mutually agreeable so that the prescription coincides with what each generation would like to propose for itself and the future; and (iii) are self-enforcing. The Markov policy serves as the default in case no other policy prescription exists that satisfies the conditions. In our setting, the emerging contract is unique and leads to coordinated savings, affecting the imputed time-discounting in the economy, and thereby the carbon price. In our quantitative assessment, the sustainable time-discounting comes close to values assumed by Stern. In this sense, we solve the "prescription puzzle": Stern's proposed low discount rates can be sustained as an equilibrium outcome.

The quantitative significance of the conceptual results follows from the unusual delays of the consequences of climate change. We develop and calibrate a novel representation of the global carbon cycle, with the peak impact lagging 60-70 years behind the date of emissions. The analytics allow us to decompose the contribution of the immediate, delayed, and persistent climate impacts to the carbon price: ignoring the delay of impacts — as in Golosov et al. (2011) and the follow-up literature — misses the correct price levels by a factor of 2. 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 particular electricity generation technologies.[10]

The relevance of hyperbolic discounting in the climate policy analysis has been acknowledged before; however, the equilibrium implications have been overlooked. Mastrandrea and Schneider (2001) and Guo, Hepburn, Tol, and Anthoff (2006) include hyperbolic discounting in simulation models assuming that the current decision makers can choose also the future policies. That is, these papers do not analyze if the policies targeted towards the long-term preferences are sustainable in equilibrium; we introduce such policies in a well-defined sense. 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 stylized setting without intertemporal consumption choices. That is, these papers consider a descriptive Markov equilibrium but cannot include climate policies that are sensitive to long-term climate outcomes and reconcilable with shorter-term macroeconomic choices. Our tractable general-equilibrium model features a joint inclusion of macro and climate policy decisions, with hyperbolic time preferences, and a detailed carbon cycle description — these fea-

[10]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.

tures are all essential for a credible resolution of the carbon price-discount rate puzzle as well as for identifying the sustainable prescriptive plans.

We take no stand on what the time-structure of preferences should be; we only present a framework for expressing preferences over long-term outcomes without disconnecting the economy from the descriptive facts. However, several recent conceptual arguments 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 rate than that for the short-term decisions; see Rubinstein (2003) for the procedural argument.[11] 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 immediate 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 climate system representation. Section 3 proceeds to the equilibrium analysis and presents the main conceptual results. Section 4 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 5 discusses those assumptions, and some robustness analysis as well as extensions to uncertainty and learning. Section 6 concludes. The supplementary material cited in the text is available in a public folder.[12]

[11]From the current perspective, generations living after 400 or, alternatively, after 450 years look the same. That being the case, no additional discounting arises from the added 50 years, while the same time delay commands large discounting in the near term.

[12]Follow the link `https://www.dropbox.com/sh/q9y9l12j3l1ac6h/dgYpKVoCMg`

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 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. Sequence $\{c_t, z_t, k_t\}_{t=1}^{\infty}$ generates per-period utilities, denoted by u_t, whose discounted sum defines the welfare of generation t as

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

where discounting is geometric and defined by factor $0 < \delta < 1$ for all dates excluding the current date when $\beta \neq 1$. As, for example, in Krusell et al. (2002), this quasi-hyperbolic formulation implies that when $\beta < 1$, the decision-maker's patience increases over time: one-period postponement of a utility gain is first discounted with $\beta\delta$ and then with δ. The increasing patience captures the idea of utility discount rates declining over time.[13] Furthermore, the 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.[14]

[13]The formal analysis is not restricted to this quasi-hyperbolic setting. In Section 5 we discuss how the analysis extends to an arbitrary sequence of discount factors.

[14]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

2.2 The economy

We build on Brock-Mirman (1972) for the climate-economy interactions, following Golosov et al. (2011), but our approaches to the climate dynamics, preferences, and equilibrium interactions are substantially different. We pull together the production structure as follows:

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

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

$$l_{y,t} + l_{e,t} = l_t \quad (5)$$

$$\omega(s_t) = \exp(-D_t), \quad (6)$$

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

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 given by function $\omega(s_t)$ capturing the output loss of production depending on the history of emissions from 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 depending only on the state of the economy, say, at year 2010. In this sense, the currently optimal policies become free of the details of the energy sector as captured by A_t and E_t, although the future development of the economy and thus future policies depend on these details. We study the future scenarios and specify A_t and E_t in detail in our longer working paper Gerlagh&Liski (2012).[15] Here, we merely assume that the final-good and energy-sector outputs are differentiable, increasing, and strictly concave in labor, energy, and carbon inputs.

The utility function is logarithmic in consumption and, through a separable linear term, we also include the possibility of intangible damages associated with climate change:

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

et al. (2010).

[15] 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).

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 the calibration, we let $\Delta_u = 0$ to maintain an easy comparison with the previous studies.[16] [17]

2.3 Damages and carbon cycle

Equations (6)-(7) show that climate damages are interpreted as reduced output, and they depend on the history of emissions through the state variable D_t that measures the global mean temperature increase. The weight structure of past emissions in (7) is derived from a Markov diffusion process of carbon between various carbon reservoirs in the atmosphere, oceans and biosphere (see Maier-Reimer and Hasselman 1987). Emissions z_t enter the atmospheric CO_2 reservoir, and slowly diffuse to the other reservoirs. The deep ocean is the largest reservoir, and the major sink of atmospheric CO_2. We calibrate this reservoir system, and, in the analysis below, by a linear transformation obtain an isomorphic decoupled system of "atmospheric boxes" where the diffusion pattern between the boxes is eliminated. The 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 reservoir j, and the shares sum to 1. The diffusion between the reservoirs is described through a $n \times n$ matrix $\mathbf{M}$ that has real and distinct eigenvalues $\lambda_1, ..., \lambda_n$. Dynamics satisfy

$$L_{t+1} = \mathbf{M} L_t + b z_t. \tag{9}$$

Definition 1 *(closed carbon cycle) No CO_2 leaves the system: 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 a matrix of linearly

[16] See Tol (2009) for a review of the existing damage estimates; the estimates for intangible damages are mostly missing. The carbon pricing formulas 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.

[17] Note that we consider average utility in our analysis. 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

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 reservoirs: atmosphere and upper ocean reservoir ($L_{1,t}$), biomass ($L_{2,t}$), and deep oceans ($L_{3,t}$). For the greenhouse effect, we are interested in the total atmospheric CO_2 stock. Reservoir $L_{1,t}$ contains both atmosphere and upper ocean carbon that almost perfectly mix within a ten-year period; we can find the atmospheric stock by correcting for the amount that is stored in the upper oceans. Let μ be the factor that corrects for the CO_2 stored in the upper ocean reservoir, so that the total atmospheric CO_2 stock is

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

Let $q_{1,i}$ denote the first row of $\mathbf{Q}$, corresponding to reservoir $L_{1,t}$. Then, the development of the atmospheric CO_2 in terms of the climate boxes is

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

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

$$\begin{aligned} S_{i,t+1} &= (1-\eta_i)S_{i,t} + a_i z_t && (10) \\ S_t &= \textstyle\sum_{i\in\mathcal{I}} S_{i,t}. && (11) \end{aligned}$$

This is now a system of atmospheric carbon stocks where depreciation factors are defined by eigenvalues from the original physical representation. When no carbon can leave the system, we know one eigenvalue, $\lambda_i = 1$,[18]

[18] 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 $\sum_i a_i < 1/(1+\mu)$.

Remark 1 *For a closed carbon cycle, one box $i \in \mathcal{I}$ has no depreciation, $\eta_i = 0$*

This observation will have important economic implications when the discount rate is small.

The carbon cycle description is well-rooted in natural science; however, the dependence of temperatures on carbon concentrations and the resulting damages are more speculative.[19] Following Hooss et al (2001, table 2), assume a steady-state relationship between temperatures, T, and steady-state concentrations $T = \varphi(S)$. Typically, the assumed relationship is concave; for example, logarithmic. Damages, in turn, are a function of the temperature $D_t = \psi(T_t)$ where $\psi(T)$ is convex. The composition of a convex damage and concave climate sensitivity is approximated by a linear function:[20]

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

with $\pi > 0$, a constant sensitivity of damages to atmospheric CO_2.[21]

Let ε be the adjustment speed of temperatures and damages, so that we can write for the dynamics of damages:

$$D_t = D_{t-1} + \varepsilon(\pi S_t - D_{t-1}). \tag{12}$$

This representation of carbon cycle and damages leads to the following analytical emissions-damage response.[22]

Theorem 1 *For the multi-reservoir model with linear damage sensitivity (9)-(12), the time-path of the damage response following emissions at time t 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

$$\eta_i = 1 - \lambda_i$$
$$a_i = \frac{q_{1,i}}{1+\mu}\mathbf{Q}^{-1}b$$

For $n = 1$, the maximum impact occurs at time between $1/\varepsilon$ and $1/\eta$.

[19] See Pindyck (2013) for a critical review

[20] 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.

[21] Section 5 reports our sensitivity analysis of the results to this approximation.

[22] All proofs, unless helpful in the text, are in the Appendix.

Theorem 1 describes the carbon cycle in terms of a system of independent atmospheric boxes, where $\mathcal{I}$ denotes the set of boxes, each with share $0 < a_i < 0$ of annual emissions entering each box $i \in \mathcal{I}$, and $\eta_i < 1$ its carbon depreciation factor. 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 parameter is $a = 1$. If atmospheric carbon-dioxide does not depreciate at all, $\eta = 0$, then the temperature slowly converges at speed ε to the long-run equilibrium damage sensitivity π, giving $\theta_\tau = \pi[1 - (1 - \varepsilon)^\tau]$. 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 absent, $\varepsilon = 0$, there is no response, $\theta_\tau = 0$.

Figure 1 shows the life path of damages (percentage of total output) caused by an impulse of one Teraton of Carbon [TtCO_2] in the first period, contrasted with a counterfactual path without the carbon impulse.[23] The output loss is thus measured per TtCO_2, and it equals $1 - \exp(-\theta_\tau)$, τ periods after the impulse. The graphs are obtained by calibrating this damage-response, that is, weights $(\theta_\tau)_{\tau \geqslant 1}$ in (7), to three cases.[24] Matching Golosov et al.'s (2011) specification produces an immediate damage peak and a fat tail of impacts, while calibrating to 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 a fat tail; about 16 per cent of emissions do not depreciate within the horizon of thousand years.

The emissions-damage response in Figure 1 has three boxes calibrated as follows. The physical data on carbon emissions, stocks in various reservoirs, and the observed concentration developments are used to calibrate a three-box carbon cycle representation

[23]One TtCO_2 equals about 25 years of global CO_2 emissions at current levels (40 GtCO_2/yr.)

[24]See Appendix for the details of the experiment.

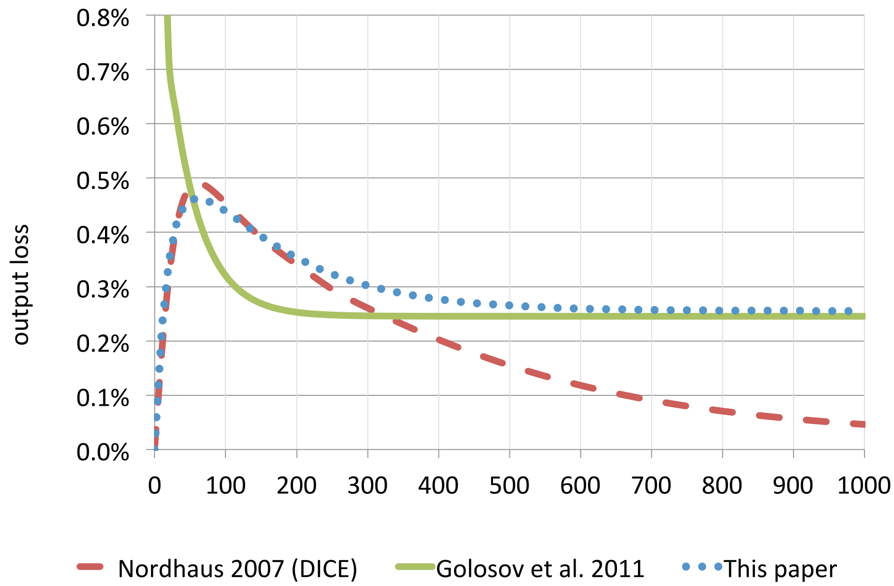

Figure 1: Emissions-damage response for three specifications

leading to the following emission shares and depreciation factors per decade:[25]

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

Thus, about 16 per cent of carbon emissions does not depreciate while about 45 per cent has a half-time of one decade. As in Nordhaus (2001), we assume that doubling the steady state CO_2 stock leads to 2.6 per cent output loss. This implies a value $\pi = .0156$ [per TtC02].[26] We assume $\varepsilon = .183$ per decade, implying a global temperature adjustment speed of 2 per cent per year. This choice is within the range of scientific evidence (Solomon et al. 2007).[27] See the Appendix for further details.

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

[26] Adding one TtCO2 to the atmosphere, relative to preindustrial levels, leads to steady-state damages that are about 0.79% of output. Adding up to 2.13 TtCO2 relative to the preindustrial level, leads to about 2.6% loss of output. The equilibrium damage sensitivity is then readily calculated as $(2.56 - 0.79)/(2.13 - 1) = 1.56\%/TtCO2$.

[27] In Figure 1, 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

3 Equilibrium

Policies take the form $k_{t+1} = \mathcal{G}_t(k_t, \Theta_t)$, $z_t = \mathcal{H}_t(k_t, \Theta_t)$, where $\Theta_t = (D_t, S_{1,t}, ..., S_{n,t})$ collects the vector of climate state variables. However, the climate affects the continuations payoffs only through the weighted sum of past emissions, as expressed in (7); for a more convenient algebra, we will replace Θ_t by s_t below, keeping in mind that the current state is equivalently described by either Θ_t or s_t.

replacing Θ_t by history s_t below, treating it as a state variable. For some given policies $\mathcal{G}_t(k_t, s_t)$ and $\mathcal{H}_t(k_t, s_t)$, 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 policies of interest in this paper will be of the form where share $0 < g_t < 1$ of the gross output is invested,

$$k_{t+1} = g_t y_t, \tag{13}$$

whereas the climate policy defines emissions through a constant h_t that defines the marginal product of the fossil fuel use, the carbon price,

$$\frac{\partial y_t}{\partial z_t} = h_t(1 - g_t)y_t. \tag{14}$$

Before the equilibrium analysis, where we explicate the restrictions under which the policies actually take the above form, it is useful to state a few general implications of such policies. First, similarly as g_t measures the stringency of the savings policy, h_t measures the stringency of the climate policy. In particular, the carbon price, $\frac{\partial y_t}{\partial z_t}$, is monotonic in policy h_t, which allows an interchangeable use of these two concepts.[28] Second, for any sequence of constants $(g_\tau, h_\tau)_{\tau \geqslant t}$ such that (13) and (14) are satisfied, we have a representation of welfare:

---

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. We note that our closed-form model can be calibrated very precisely to approximate the DICE model (Nordhaus 2007); Section 5 discusses further on the surprising prediction power of our carbon pricing formula for the DICE results. We also include a more detailed note on this issue in the supplementary material.

[28]We show this in Lemma 3 of the Appendix.

Theorem 2 *It holds for every policy sequence $(g_\tau, h_\tau)_{\tau>t}$ that*

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

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

The future cost of the emission history is thus given by $\Omega(s_{t+1})$, giving also the marginal cost of the current emissions as ζ_1 that is a compressed expression for the climate-economy impacts. But, we can immediately see from Remark 1 that a closed carbon cycle leads to persistent impacts ($\eta_i = 0$ for one i), implying thus unbounded future marginal losses when the long-term discounting vanishes:

Proposition 1 *For a closed carbon cycle,* $\frac{\partial \Omega(s_{t+1})}{\partial z_t} \to \infty$ *as* $\delta \to 1$.

The result has strong implications for the comparison of the policies considered next.

3.1 Markov equilibrium policies

We focus on symmetric and stationary Markov equilibrium. Symmetry means that all generations use the same policy functions.[29] The Markov policies do not condition on the history of past behavior: strategies are identical in states where the continuation payoffs are identical (see Maskin and Tirole, 2001).[30]

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 (13)-(14), as given and chooses its current savings to satisfy

$$u_t' = \beta\delta V_{t+1}'(k_{t+1}),$$

[29] Even though there can be technological change and population growth, the form of the objective, (8) combined with (2), ensures that there will be an equilibrium where the same policy rule will be used for all t.

[30] 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).

where u'_t denotes marginal consumption utility and function $V(\cdot)$ from Theorem 2 captures the continuation value implied by the equilibrium policy.

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

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

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 choice for the fossil-fuel use, z_t, satisfying

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

The optimal policy thus equates the marginal current utility gain from fuel use with the change in equilibrium costs on future agents. Denote the equilibrium carbon price by $\tau_t^{\beta\delta}(=\partial y_t/\partial z_t)$. Given Theorem 2, carbon price $\tau_t^{\beta\delta}$ can be obtained:

Proposition 2 *The equilibrium carbon price is*

$$\tau_t^{\beta\delta} = h^*(1-g)y_t \tag{16}$$

$$h^* = \Delta \sum\nolimits_{i\in\mathcal{I}} \frac{\beta\delta a_i \pi\varepsilon}{[1-\delta(1-\eta_i)][1-\delta(1-\varepsilon)]} \tag{17}$$

$$\Delta = \left(\frac{1}{1-\alpha\delta} + \Delta_u\right)$$

When y_t is known, say $y_{t=2010}$, the carbon policy for $t = 2010$ can be obtained from (16), by reducing fossil-fuel use to the point where the marginal product of z equals the externality cost of carbon.

To obtain the current externality cost of carbon intuitively, that is, the social cost of carbon emissions z_t as seen by the current generation, consider the effect of damages $D_{t+\tau}$ on utility in period $t+\tau$. Recall that the consumption utility is $\ln(c_{t+\tau}) = \ln((1-g)y_{t+\tau}) = \ln(1-g)+\ln(y_{t+\tau})$ so that, through the exponential output loss in (6), $\partial \ln(c_{t+\tau})/\partial D_{t+\tau} = -1$. As there is also the direct utility loss, captured by Δ_u in (8), the full loss in utils at $t+\tau$ is

$$-\frac{du_{t+\tau}}{dD_{t+\tau}} = 1 + \Delta_u.$$

But, the output loss at $t+\tau$ propagates through savings to periods $t+\tau+n$ with $n>0$,

$$-\frac{du_{t+\tau+n}}{dD_{t+\tau}} = \alpha^n,$$

leading to the full stream of losses in utils, discounted to $t+\tau$,

$$-\sum_{n=0}^{\infty} \delta^n \frac{du_{t+\tau+n}}{dD_{t+\tau}} = \frac{1}{1-\alpha\delta} + \Delta_u = \Delta.$$

The full loss of utils per increase in temperatures as measured by $D_{t+\tau}$ is thus a constant given by Δ for any future τ, giving the social cost of carbon emissions z_t at time t, appropriately discounted to t, as

$$\begin{aligned}
-\beta \sum_{\tau=1}^{\infty} \delta^\tau \frac{du_{t+\tau}}{dz_t} &= \sum_{\tau=1}^{\infty} \sum_{n=0}^{\infty} \beta\delta^{\tau+n} \frac{du_{t+\tau+n}}{dD_{t+\tau}} \frac{dD_{t+\tau}}{dz_t} \\
&= \Delta \sum_{\tau=1}^{\infty} \beta\delta^\tau \frac{dD_{t+\tau}}{dz_t} \\
&= \Delta \sum_{i\in\mathcal{I}} \frac{\beta a_i \pi \varepsilon}{\varepsilon - \eta_i} \sum_{\tau=1}^{\infty} \delta^\tau (1-\eta_i)^\tau - \delta^\tau (1-\varepsilon_j)^\tau \\
&= \Delta \sum_{i\in\mathcal{I}} \frac{\beta\delta\pi a_i \varepsilon}{[1-\delta(1-\eta_i)][1-\delta(1-\varepsilon)]}.
\end{aligned}$$

This is exactly the value of h^*. Thus, in equilibrium, the present-value utility costs of current emissions remain constant at level h^*. However, since this cost is weighted by income in (16), the equilibrium carbon price increases over time in a growing economy.

The Markov equilibrium carbon price 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 ($\partial h/\partial\beta > 0$; $\partial h/\partial\delta > 0$). Consistent with Proposition 1, the carbon price rises sharply if the discount factor comes close to one, $\delta \to 1$, and if some box has slow depreciation, $\eta_i \to 0$.[31]

[31] If carbon depreciates quickly, $\eta_i >> 0$, then the carbon price will be less sensitive to the discount factor δ. Fujii and Karp (2008) 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 the effect of CO_2 on the economy depreciates more than 25 per cent per decade. This rate is well above the estimates for CO_2 depreciation in the natural-science literature; however, induced adaptation may lead to similar reduction in damages.

3.2 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,

$$\gamma = \frac{u_t'}{u_{t+1}' R_{t,t+1}} = \frac{c_{t+1}}{c_t R_{t,t+1}} = \frac{c_{t+1}}{c_t} \frac{k_{t+1}}{\alpha y_{t+1}} = \frac{g}{\alpha}. \tag{18}$$

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

$$\gamma^* = \frac{\beta\delta}{1 + \alpha\delta(\beta - 1)}. \tag{19}$$

This is the geometric utility discount factor that is consistent with the efficiency of the equilibrium consumption stream: a fictitious Ramsey 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 actions for such a planner, we arrive at the definition of the Pigouvian tax for the consistent preferences discounting γ^*.

Proposition 3 *(Imputed Pigouvian tax) The optimal carbon price, τ_t^γ, for a Ramsey planner who discounts utilities with geometric discount factor $\gamma = \gamma^*$ equals the consumption-weighted net present value of future marginal utility losses from emissions z_t. Value τ_t^γ is given by*

$$\begin{aligned} \tau_t^\gamma &= h^\gamma (1-g) y_t & (20)\\ h^\gamma &= \Delta^\gamma \textstyle\sum_{i \in \mathcal{I}} \dfrac{\gamma \pi a_i \varepsilon}{[1-\gamma(1-\eta_i)][1-\gamma(1-\varepsilon)]} & (21)\\ \Delta^\gamma &= \frac{1}{1-\alpha\gamma} + \Delta_u. \end{aligned}$$

We have now two definitions for the social cost of carbon. The one in Proposition 2 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 3 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 equals the imputed Pigouvian tax, a weighted sum of the utility sequence is maximized, and in this sense the economy is on the efficiency frontier. However, it is not immediate that there are welfare 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 for the two carbon prices to differ:

Proposition 4 *For quasi-hyperbolic preferences, $\beta, \delta < 1$, 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, $\tau_t^{\beta\delta}/\tau_t^{\gamma}$, is continuous in parameters $\beta, \delta, \eta_i, \varepsilon, a_i$, and γ. Evaluating at $\gamma = \gamma^*$, $\beta < 1$, $\eta_i = \varepsilon = 0$,*

$$\tau_t^{\beta\delta} > \tau_t^{\gamma}.$$

If the climate system is sufficiently persistent, then the current generation uses the climate asset as a commitment device to transfer wealth to future generations; the external climate costs are valued 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. The distortion identified here has no bound in the following sense:

Proposition 5 *For a closed carbon cycle and $\beta < 1$: $\tau_t^{\beta\delta}/\tau_t^{\gamma} \to \infty$ as $\delta \to 1$.*

When no carbon leaves the system, a fraction of the climate impacts is persistent, which drives the Markov price to infinity while the imputed price remains bounded.

We ask next if there are potential gains to be achieved from enforcing the pricing rule in Proposition 3 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.[32] We provide such a comprehensive welfare analysis for an equivalent model in Gerlagh&Liski (2011); here we bring the essence of the welfare impacts.[33]

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 $h = h^{\gamma=\delta}$ where h^{γ} is defined in Proposition 3; 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 2 *For $\beta \neq 1$ and any given $\tau > t$,*

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

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.

Proposition 6 *For slow climate change, implementing $\partial y_t/\partial z_t = \tau_t^{\gamma}$ from period t onwards implies a welfare loss for generation t vis-à-vis the Markov equilibrium.*

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.[34]

3.3 Sustainable policies

Consider a policy pair $(\widehat{g}, \widehat{h})$ that the current generation would like to propose for all generations, including itself. The proposal is required to be symmetric: $(\widehat{g}, \widehat{h})$ is the same for

[32]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.

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

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

all affected generations. From Lemma 2, generation t would like to propose to all future generations the decision rule g^δ and h^δ but achieving this requires that these policies are followed also at t, which is ruled out by the current incentive constraints: (g^δ, h^δ) does not maximize w_t. But, also from Lemma 2, the current generation is willing to give up part of its consumption, by increasing g and h, beyond their Markov equilibrium levels, anticipating that all subsequent decision-makers will follow suit when facing the same decision.[35]

Policy $(\widehat{g}, \widehat{h})$ that we define is *sustainable*; it is self-enforcing, and identifies a well-defined optimal symmetric "contract" that can be proposed for future at each t. We define policies $\widehat{g}$ and $\widehat{h}$ independently. Formally, at time t, the proposal for savings, for any sequence of h_τ $(\tau \geq t)$, is

$$\widehat{g}_t \in \{\arg\max_{g\in[0,1]} w_t | g_\tau = g \text{ for all } \tau \geq t\}.$$

Building on Theorem 2, we show in Appendix that the set is well defined. The equilibrium proposal comes from

$$\widehat{g} \in \cap_{t=1}^{\infty}\{\widehat{g}_t\}, \tag{22}$$

or if no joint proposal emerges, $\widehat{g}$ is taken to be the Markov rule g^*. Similarly, without conditioning on g, the proposal for the climate policy is

$$\widehat{h}_t \in \{\arg\max_{h\in\mathbb{R}_+} w_t | h_\tau = h \text{ for all } \tau \geq t\} \Rightarrow \widehat{h} \in \cap_{t=1}^{\infty}\{\widehat{h}_t\}. \tag{23}$$

If the set is empty, we take $\widehat{h} = h^*$.

Proposition 7 *Sustainable policy $(\widehat{g}, \widehat{h})$ satisfies*

$$g^\delta > \widehat{g} > g^*$$
$$h^\delta > \widehat{h} \geqslant h^*$$

where $\widehat{g}$ is unique. Moreover, $\widehat{h} = h^*$ if $A_t \neq A_\tau$ for $\tau > t$. The sustainable imputed discount factor is

$$\widehat{\gamma} = \frac{\beta\delta}{1 + \alpha\delta(\beta - 1) + (1 - \alpha\delta)(\beta - 1)\delta}$$

such that $\gamma^* < \widehat{\gamma} < \delta$.

[35] Roemer (2010) defines a Kantian equilibrium where each subject presumes that *all* other subjects follow the same rule; for a static economy, Roemer shows that the Kantian conjecture leads to an efficient outcome. In our context, a Kantian policy rule would demand that each generation sets a policy pair (g, h) that it would like to see for both *future and past* generations. But in our setting, time flows only forward so we derive only forward-looking policy rules. Moreover, our policy is self-enforcing.

The unique sustainable saving rule is $\widehat{g} = \alpha\widehat{\gamma}$; all agents agree on these higher-than-Markov savings. In contrast, if technology A_t changes over time, it is generally not possible to find any other "contractible" symmetric policy rule $\widehat{h}$ than the Markov rule h^*. But there is still an impact on the carbon price: sustainable savings mitigate the capital market distortion, thereby reducing the distortion correction in the carbon price. If savings can be co-ordinated in the sense discussed here, the Pigouvian tax based on revealed discounting (i.e., the imputed tax using $\widehat{\gamma}$) and the equilibrium carbon price are close to equal, as our quantitative assessment shows. They are also equal in the following limit:

Proposition 8 *For low long-term discounting, $\delta \to 1$, zero time-discounting is sustainable: $\widehat{\gamma} \to 1$.*

For a closed carbon cycle, the sustainable imputed carbon price becomes then unbounded, $\tau_t^{\widehat{\gamma}} \to \infty$, as is the case for the Markov price.

4 Quantitative assessment

To evaluate the quantitative significance of the conceptual results, we exploit the closed-form price formulas — given the structure of 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).[36] 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),[37] and year '2010' corresponds to period 2006-2015. We set $\Delta_u = 0$. We take the Gross Global Product as 600 Trillion Euro [$Teuro$] for the 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 2.7 per cent annual pure rate of time preference ($\beta = 1$,$\delta = 0.761$), resulting in $\alpha = g/\delta = 0.329$.

[36]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).

[37]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.

4.1 Assessment of Markov policies

The parameter choices together with our carbon cycle result in a consistent-preferences Pigouvian carbon price of 7.1 Euro/tCO_2, equivalent to 34 USD/tC, for 2010.[38] This number is very close to the level found by Nordhaus. Consider then the determinants of this number in detail.

We can decompose the carbon price (18) into three contributing parts. First, consider the one-time costs assuming full immediate damages (ID) taking place in the immediate next period,

$$ID = \beta\delta\Delta\pi(1-g)y_t.$$

This value is multiplied by a factor to correct for the persistence of climate change due to slow depreciation of carbon in the atmosphere, the persistence factor (PF),

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

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

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

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 2.7 per cent: this row presents the carbon price under the same assumptions as in Nordhaus (2007). Keeping the equilibrium time-preference rate at $\gamma^* = 2.7$, 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.[39]

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

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

[39] 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 $\gamma = 0.76$ (2.7 per cent annual discount rate). Since the equilibrium utility discount rate remains at 2.7 per cent, the macroeconomy remains observationally equivalent to that in Nordhaus ($g = .25$).

annual discount rate						carbon price	
short-term	long-term	equilibrium	ID	PF	DF	equilibrium	imputed
027	.027	.027	7.12	2.06	.48	7.1	7.1
.033	.01	.027	7.12	3.70	.70	18.5	7.1
.035	.005	.027	7.12	5.79	.82	33.8	7.1
.037	.001	.027	7.12	19.6	.96	133	7.1
.001	.001	.001	9.27	19.6	.96	174	174

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

that by construction the Nordhaus number 7.1 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, 133 EUR/tCO$_2$, corresponds to the case where the long-run discounting is as proposed by Stern (2006); this case also best matches Weitzman's values.[40] For reference, we report the Stern case where the long-term discounting at .1 per cent holds throughout; the carbon price takes a value of 174 EUR/tCO$_2$, and gross savings cover about 33 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.[41]

The decomposition of the carbon price is revealing. Leaving out the time lag between CO_2 concentrations and the temperature rise amounts to replacing the column DF by 1. When preferences are consistent (the first line), abstracting from the delay in temperature adjustments, as in Golosov et al. (2011), doubles the carbon price level. For hyperbolic discounting, as expected, the persistence of impacts, capturing the commitment value of climate policies, contributes significantly to the deviation between the imputed and Markov equilibrium prices.

Table 2 quantifies the economic substance of appropriately accounting for the distor-

[40]See footnote 39.

[41]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, 2007). However, with reasonable parameters such extreme savings do not occur, as in Table 2.

tions 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.

4.2 Assessment of sustainable policies

The "savings contract" from Proposition 7 deals with the main distortion arising from time-changing discounting in our intergenerational policy game. Coordination of savings has strong implications for equilibrium discounting as Table 3 shows: the sustainable imputed discount rate falls quickly below the Markov level that stays at 2.7 per cent (as in Table 2). The imputed Pigouvian carbon price, reported in the last column of Table 3, is based on this new lower equilibrium rate and thus increases strongly as a result of savings co-ordination. The assessment confirms that, if savings are coordinated, using aggregate statistics for obtaining the appropriate externality price gives close to the correct value, that is, the best-responding equilibrium carbon tax.[42]

annual discount rate			sustainable	carbon price	
short-term	long-term	equilibrium	savings rate	equilibrium	imputed
.027	.027	.027	0.25	7.1	7.1
.033	.01	.018	0.29	17	19
.035	.005	.0062	0.31	31	35
.037	.001	.0013	0.32	120	137
.001	.001	.001	0.33	174	174

Table 3: Sustainable carbon prices based on $\widehat{g} = \alpha\widehat{\gamma}$ and $\widehat{h} = h^*$ and the imputed Pigouvian externality price [Euro/tCO2]. Parameter values in text.

5 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

[42] We assume for the sustainable equilibrium carbon prices that $\widehat{h} = h^*$, which from Proposition 7, is the equilibrium when technology A_t changes over time.

Brock-Mirman (1972) we assumed 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.[43] Second, we assumed a linearized model for carbon diffusion that might not well describe the relevant dynamics when the system is far off the central path — that is, non-linearities captured by more complicated climate simulation models may be important.

To address the sensitivity of the results to the above two sets of assumptions, we devised a Monte Carlo experiment for testing how well the closed-form carbon price predicts the carbon price of a benchmark simulation model, DICE 2007, that features more flexible policies and non-linearities of the climate system. Assuming geometric discounting and 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 per cent of the DICE variation in the carbon price. This suggests that the loss of generality from not including (i) the deeper linkages between policies and (ii) the non-linearities of climate change is inconsequential for the carbon price results.[44]

The evidence from the above experiment is suggestive that tractable climate-economy models can be very good approximations for the more comprehensive models. Since the carbon price is a closed-form function of the climate-economy fundamentals, such models can prove useful in addressing the consequences of uncertainties related to the fundamentals in a direct way. Our reduced-form carbon cycle and damage representations assumed no uncertainty, although great uncertainties describe both the climate system parameters as well as the impacts of climate change. Golosov et al. (2011) make progress in this direction showing that the optimal polices are robust to impact uncertainty; this effectively leads to rewriting of the carbon price formula in expected terms. Iverson (2012) shows the robustness of the Markov equilibrium policy rules in a stochastic Markov equilibrium with multiple stochastic parameters. Furthermore, it is possible to construct carbon price distributions directly using the tractable carbon price formula and the primitive uncertainties regarding its determinants to decompose their respective contributions to carbon price uncertainty.[45]

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

[44] See Figure 2 in the note titled "SCC Formula" of the supplementary material. The note summarizes the results and documents the experiment. For conciseness, the note presents a continuous-time version of the carbon price formula; the note is self-contained.

[45] The analysis of uncertainties is beyond the scope of the current paper; however, see Figure 3 in the

Finally, the quasi-hyperbolic discount factors are only rough approximations for the discount rate paths estimated in the literature (e.g., Weitzman, 2001). However, 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) builds on our setting to elaborate the implications of more flexible discounting.

6 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 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 principle, often invoked in practical program evaluations, that the time-discounting rate should depend on the time horizon of the project. In general equilibrium, which is the approach needed for climate policy evaluations, time-changing discount rates distort the economy, stipu-

supplementary note "SCC Formula" for an illustration of the approach.

lating 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 price. We used discount factors from the literature to show that the equilibrium correction to the standard Pigouvian pricing principle 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] Giglio, S. M. Matteo, J. Stroebel (2013), Very long-run discount rates, Available at SSRN: http://dx.doi.org/10.2139/ssrn.2346049.

[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. Forthcoming: Econometrica.

[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] Guo, J., C. J. Hepburn, R.S.J. Tol, D. Anthoff (2006), Discounting and the social cost of carbon: a closer look at uncertainty, Environmental Science and Policy 9, 205-216

[18] 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

[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] Layton, D.F., and R. A. Levine (2003), How Much Does the Far Future Matter? A Hierarchical Bayesian Analysis of the Public's Willingness to Mitigate Ecological Impacts of Climate Change, Journal of the American Statistical Association, Vol. 98, No. 463, pp. 533- 544.

[35] 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.

[36] 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.

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

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

[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] Nordhaus, W. D. (2008), A Question of Balance: Weighing the Options on Global Warming Policies. (Yale University Press, New Haven, CT).

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

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

[43] 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

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

[45] Pindyck, Robert S. (2013), "Climate Change Policy: What Do the Models Tell Us?" Journal of Economic Literature, 51(3): 860-72.

[46] 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.

[47] Roemer J.E. (2010), Kantian equilibrium, Scandinavian Journal of Economics 112(1): 1-24.

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

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

[50] 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.

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

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

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

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

Appendix

Proof of Theorem 1

Given the sequence of climate variables — carbon stocks $S_{i,t}$ and damages D_t — that we developed in the text, it is a straightforward matter of verification that future damages depend on past emissions as follows:

$$\begin{aligned} S_{i,t} &= (1-\eta_i)^{t-1} S_{i,1} + \sum\nolimits_{\tau=1}^{t-1} a_i (1-\eta_i)^{\tau-1} z_{t-\tau} && (24) \\ D_t &= (1-\varepsilon)^{t-1} D_1 + \sum\nolimits_{i\in\mathcal{I}} \pi\varepsilon \frac{(1-\eta_i)^t - (1-\eta_i)(1-\varepsilon)^{t-1}}{\varepsilon-\eta_i} S_{i,1} + && (25) \\ & \quad \sum\nolimits_{i\in\mathcal{I}} \sum\nolimits_{\tau=1}^{t-1} a_i \pi\varepsilon \frac{(1-\eta_i)^\tau - (1-\varepsilon)^\tau}{\varepsilon-\eta_i} z_{t-\tau}, \end{aligned}$$

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. If some climate change has taken place at the start of time $t = 1$, we can write the system dependent on $S_{i,1}, D_1 > 0$ — however, we can also rewrite the model to start at $t = T$, possibly $T < 0$, indicating the beginning of the industrial era, say 1850; we set $z_t = 0$ for $t < T$, and $S_{i,T} = D_T = 0$. It is then immediate that the equation reduces to (7). This defines the emissions-damage function θ_τ in Theorem 1. Q.E.D.

Lemma 3

We state first the following Lemma that will be used in other proofs and is also cited in the main text. The first item of Lemma 3 is an independence property following from the functional assumptions: the energy sector choices do not depend on the current state of the economy (k_t, s_t). The latter item in Lemma 3 allows us to interpret the policy stringency as measured by h directly as the stringency of the carbon price τ.

Lemma 3 *For all t:*

(i) *Given policy sequence $(g_t, h_t)_{t \geqslant 0}$, emissions $z_t = z_t^*$ at t implied by the policy are independent of the current state (k_t, s_t), but depend only on the current technology at t as captured by $A_t(.)$ and $E_t(.)$;*

(ii) *Given the current state (k_t, s_t) at t, the carbon price, $\tau_t = \partial y_t / \partial z_t$, satisfying $\tau_t = h_t(1 - g_t) y_t$, is monotonic in the policy variable: $d\tau_t / dh_t > 0$.*

Proof: For given state and labour supply, (k_t, s_t, l_t), output $y_t = f_t(k_t, l_t, z_t, s_t)$ is increasing and concave in emissions z_t, so that if the carbon price equals the marginal carbon product $\tau_t = f_{t,z} = \partial y_t / \partial z_t$, we have $dy_t/dz_t > 0$ and $d\tau_t/dz_t < 0$. For a policy pair (g_t, h_t) at time t, we also derive $dh_t/dz_t = [f_{t,zz} - (f_{t,z})^2]/(1 - g)f_t < 0$, so that the carbon price measured in units h_t and the carbon price measured in units τ_t are monotonically related, $d\tau_t/dh_t > 0$.

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:

$$\frac{1}{y_t} \frac{\partial y_t}{\partial e_t} \frac{\partial E_t}{\partial z_t} = h_t(1 - g_t), \tag{26}$$

$$\frac{\partial A_t}{\partial l_{y,t}} = \frac{\partial A_t}{\partial e_t} \frac{\partial E_t}{\partial l_{e,t}} \tag{27}$$

Equation (27) balances 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 two first-order conditions (26)-(27), 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 $E_t(z_t, l_{e,t})$, but independent of the state variables k_t and s_t. Thus, $z_t = z_t^*$ can be determined independently of (k_t, s_t). Q.E.D.

Proof of Theorem 2

The proof is by induction. Induction hypothesis: assume (i) that future policies are given by a sequence of constants $(g_\tau, h_\tau)_{\tau > t}$ such that

$$\begin{aligned} k_{\tau+1} &= g_\tau y_\tau, && (28) \\ \frac{\partial y_\tau}{\partial z_\tau} &= h_\tau (1 - g_\tau) y_\tau, && (29) \end{aligned}$$

and (ii) that Theorem 2 holds for $t+2$. We can thus construct the value function for the next period, as

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

Consider policies at $t+1$. From (28), $k_{t+2} = g_{t+1} y_{t+1}$. Emissions $z_{t+1} = z_{t+1}^*$ can be determined independently of the state variables k_{t+1} and s_{t+1} as shown in Lemma 3. Substituting the policies at $t+1$ gives:

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

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 $V_{t+1}(k_{t+1})$:

$$\widetilde{A}_{t+1} = \ln(1 - g_{t+1}) + \delta \xi \ln(g_{t+1}) + (1 + \delta \xi) \ln(A_{t+1}) - \delta \zeta_1 z_{t+1} + \delta \widetilde{A}_{t+2}. \qquad (30)$$

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} 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 $\ln(\omega(s_{t+1})) = -D_{t+1}$ so that

$$\zeta_\tau = ((1+\delta\xi) + \Delta_u) \sum\nolimits_{i\in\mathcal{I}} a_i \pi \varepsilon \frac{(1-\eta_i)^\tau - (1-\varepsilon)^\tau}{\varepsilon - \eta_i} + \delta\zeta_{\tau+1}$$

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

$$\zeta_\tau = (\frac{1}{1-\alpha\delta} + \Delta_u) \sum\nolimits_{i\in\mathcal{I}} \sum\nolimits_{t=\tau}^{\infty} a_i \pi \varepsilon \delta^{t-\tau} \frac{(1-\eta_i)^t - (1-\varepsilon)^t}{\varepsilon - \eta_i}$$

To derive the value of ζ_1, we consider

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

(When $\eta_i = \varepsilon$, ζ_1 still has a closed-form solution; this derivation is available on request) Q.E.D.

Proof of Proposition 1

In text.

Proof of Proposition 2

In text.

Proof of Proposition 3

To prove the result, set $\delta = \gamma^*$ and $\beta = 1$, and optimal policy for the Ramsey planner follows from Proposition 2. For such a planner, Theorem 2 defines the present-value future marginal utility losses from emissions through

$$\frac{\partial\Omega(s_{t+1})}{\partial z_t} = \Delta^{\gamma} \sum_{i\in\mathcal{I}} \frac{\gamma\pi a_i\varepsilon}{[1-\gamma(1-\eta_i)][1-\gamma(1-\varepsilon)]}$$

where $\gamma = \gamma^*$. Since the planner sets

$$u_t'\frac{\partial y_t}{\partial z_t} = \gamma^*\frac{\partial\Omega(s_{t+1})}{\partial z_t},$$

the optimal policy has the interpretation given in Proposition 3. Q.E.D.

Proof of Proposition 4

Because of Lemma 3 (ii), the proposition, stated as $\frac{\tau^{\beta\delta}}{\tau_t^{\gamma}} > 1$, can be rewritten, equivalently, as one where carbon prices are measured in utility units: $\frac{h^{\beta\delta}}{h_t^{\gamma}} > 1$. We consider the latter ratio for very long climate change delays, $\eta_i = \varepsilon = 0$, and, $\beta < 1$:

$$\begin{aligned}
\frac{h^{\beta\delta}}{h_t^{\gamma}} &= \frac{(1-\gamma)^2}{(1-\delta)^2}\frac{\beta\delta}{\gamma} \\
&= \frac{\left(1-\frac{\beta\delta}{1-\alpha\delta+\alpha\beta\delta}\right)^2}{(1-\delta)^2}(1-\alpha\delta+\alpha\beta\delta) \\
&= \frac{(1-\delta(\alpha+(1-\alpha)\beta))^2}{(1-\delta)^2(1+\alpha\delta(\beta-1))} \\
&> 1
\end{aligned}$$

The first equality follows from substitution of $\eta_i = \varepsilon = 0$ 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.

Proof of Proposition 5

From Proposition 2, $\tau^{\beta\delta} \to \infty$ as $\delta \to 1$. From Proposition, 4, we see that τ^{γ} remains bounded. Q.E.D.

Proof of Lemma 2

Consider a given policy path $(g_\tau, z_\tau)_{\tau\geq t}$. We look at variations of policies at time τ, and consider the effect on welfare at time t. All effects are captured by W_{t+1} in Theorem 2.

The analysis in the proof of Theorem 2 implies: the value function at time t is separable in states and the parameters ξ and ζ do not depend on future polices (g_τ, z_τ), but term $\widetilde{A}_t$ does. Technically, we need to show that, for some given $\tau > t$,

1. $\widetilde{A}_t$ increases in g_τ for $g_\tau < \alpha\delta$,
2. $\widetilde{A}_t$ decreases in z_τ for $z_\tau > z_\tau^\delta$,

where z_τ^δ is the emission level that is consistent with the policy variable h^δ and z_τ is the emission level consistent with some $h < h^\delta$. In the proof of Theorem 2, consider (30). Term $\widetilde{A}_t$ increases with $\widetilde{A}_\tau$ for some $\tau > t$. Moreover, $\widetilde{A}_\tau$ is strictly concave in g_τ, and maximal when g_τ maximizes $\ln(1-g_\tau)+\delta\xi\ln(g_\tau)$, that is, for $g_\tau = \frac{\delta\xi}{1+\delta\xi} = \alpha\delta$. This proves item 1. Furthermore, notice that A_τ depends on z_τ, that $\widetilde{A}_\tau$ is strictly concave in z_τ and maximal when z_τ maximizes $(1+\delta\xi)\ln(A_\tau(z_\tau)) - \delta\zeta_1 z_\tau$, that is, for $\frac{d\ln A_t}{A_\tau dz_\tau} = \delta\zeta_1(1-\alpha\delta)$. This is the value of z_τ consistent with h^δ. This proves item 2. We have now shown the "if" part of Lemma 2. The "only if" follows from the strict concavity of $\widetilde{A}_\tau$ with respect to (g_τ, z_τ). Q.E.D.

Proof of Proposition 6

By Theorem 2, the change of the carbon price 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. Also, by Lemma 3, a higher policy h implies a higher carbon price τ. Let $\beta < 1$ so that $\beta\delta < \gamma < \delta$, and let climate change be a slow process such that $\tau_t^\delta > \tau^{\beta\delta} > \tau_t^\gamma$; see Proposition 4. 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 2. The same mechanism applies for $\beta > 1$, when we have $\tau_t^\delta < \tau^{\beta\delta} < \tau_t^\gamma$. Moreover, imposing the imputed carbon price on current policies implies a deviation from the current best response. That is, both changes induced, those in the present and future polices, decrease the present welfare.

Proof of Proposition 7

The proposal $\widehat{g}_t$, as defined in the text, maximizes $w_t = u_t + \beta\delta W_{t+1}(k_{t+1}, s_{t+1})$. From the proof of Theorem 2, we see that $W_{t+1}(k_{t+1}, s_{t+1})$ depends on g_{t+1} only through $\widetilde{A}_{t+1}$ so that

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

$$(\forall\tau > t, \widehat{g}_\tau = g) \Rightarrow$$

$$\widetilde{A}_{t+1} = \frac{1}{1-\delta}\ln(1-g) + \frac{\delta\xi}{1-\delta}\ln(g) + (1+\delta\xi)\textstyle\sum_{\tau=t+1}^{\infty}\delta^{\tau-t-1}[\ln(A_\tau) - \delta\zeta_1 z_\tau]$$

$$\Rightarrow$$

$$\arg\max_g w_t = \arg\max_g \ln((1-g)y) + \beta\delta\widetilde{A}_{t+1}$$

$$= \arg\max_g \ln(1-g) + \beta\delta\xi\ln(g) + \frac{\beta\delta}{1-\delta}\ln(1-g) + \frac{\beta\delta^2}{1-\delta}\xi\ln(g)$$

$$\Rightarrow$$

$$\widehat{g}_t = \widehat{g} = \frac{\alpha\beta\delta}{1+\alpha\delta(\beta-1)+(1-\alpha\delta)(\beta-1)\delta}.$$

Since $\widehat{g}$ it is independent of t, this same proposal is optimal for any agent at $\tau > t$. It remains to be shown that the policy is self-enforcing, so that for any t there is no one-shot deviation $g_t \neq \widehat{g}$ implying a higher payoff w_t. From Theorem 2, the one-shot deviation at t is independent of the policy sequence $(g_\tau, h_\tau)_{\tau>t}$; that is, it is the Markov policy, $g_t = g^*$. Anticipating that $g_t = g^*$ triggers $(g_\tau = g^*, h_\tau = h^*)_{\tau>t}$, gives then the Markov payoff as the deviation payoff. But since, by construction of $\widehat{g}$, $w_t^* < w_t$, where the latter is supported by the continuation policy, the deviation leads to a strict loss. (Since w_t is separable in (g, h), the argument holds independently of whether a deviation for one policy triggers the Markov response for the other policy.)

Consider then the proposal $\widehat{h}_t$ that maximizes $w_t = u_t + \beta\delta W_{t+1}(k_{t+1}, s_{t+1})$. From Theorem 2, we see that $W_{t+1}(k_{t+1}, s_{t+1})$ depends on h_τ through z_τ for $\tau > t$; from Lemma 3, z_t is a monotonic (strictly decreasing) function of h_t. Thus, from Theorem 2, $W_{t+1}(k_{t+1}, s_{t+1})$ depends on z_τ and h_τ for $\tau > t$ only through $\widetilde{A}_{t+1}$,

$$\beta\delta(1+\delta\xi)\textstyle\sum_{\tau=t+1}^{\infty}\delta^{\tau-t-1}[\ln(A_\tau(z_\tau)) - \delta\zeta_1 z_\tau],$$

where we have written $A_\tau = A_\tau(z_\tau)$ to emphasize dependence on z_τ. Hence,

$$\arg\max_h w_t = \arg\max_h \ln((1-g)y) - \beta\delta\zeta_1 z_t + \beta\delta\widetilde{A}_{t+1}$$

$$= \arg\max_h \ln(A_t(z_t)) - \beta\delta\zeta_1 z_t + \beta\delta(1+\delta\xi)\textstyle\sum_{\tau=t+1}^{\infty}\delta^{\tau-t-1}[\ln(A_\tau(z_\tau)) - \delta\zeta_1 z_\tau].$$

Then, if $(\forall\tau \geqslant t, \widehat{h}_t = h)$ is such a proposal at t, it solves

$$[\frac{d\ln A_t(z_t)}{dz_t} - \beta\delta\zeta_1]\frac{dz_t}{dh} = -\beta\delta(1+\delta\xi)\sum_{\tau=t+1}^{\infty}\delta^{\tau-t-1}[\frac{d\ln A_\tau(z_\tau)}{dz_\tau} - \delta\zeta_1]\frac{dz_\tau}{dh}$$

where $dh = dz_t[f_{t,zz} - (f_{t,z})^2]/(1-g)f_t$ from Lemma 3. The proposal is independent of t, that is, for $t \neq t'$, $\widehat{h}_t = \widehat{h}_{t'} = h$ if and only if both $A_t = A_{t'}$ and $dz_t/dh = dz_{t'}/dh$. Thus, non-stationary technology $A_t \neq A_{t'}$ rules out a joint proposal $\widehat{h}$ that is independent of time. Thus, Markov default follows, $\widehat{h} = h^*$. Q.E.D.

Proof of Proposition 8

Follows directly from the expression for $\widehat{\gamma}$. Q.E.D.

Calibrating carbon cycle

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 supplementary material.[46] We calibrate the model parameters $\mathbf{M}$, b, μ, to minimize the error between the atmospheric concentration prediction from the three-reservoir model and the Mauna Loa observations under the constraint that CO_2 stocks in the various reservoirs and flows between them 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 to be calibrated. We set $b = (1, 0, 0)$ so that emissions enter the first reservoir (athmospere). 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 reservoir relative to the atmosphere, while we fix the relative size of the deep ocean reservoir at 4 times the atmosphere, 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 three-reservoir model into a decadal reservoir model by adjusting the exchange rates within a period between the reservoirs and the shares of

[46] Follow the link `https://www.dropbox.com/sh/q9y9l12j3l1ac6h/dgYpKVoCMg`

emissions that enter the reservoirs within the period of emissions. Then, we transform the decadal three-reservoir model into the decadal three-box model, following linear algebra steps described above. The transformed box model has no direct physical meaning other than this: box 1 measures the amount of atmospheric carbon that never depreciates; box 2 contains the atmospheric carbon with a depreciation of about 7 per cent in a decade; while carbon in box 3 depreciates 50 per cent per decade.[47] About 20 per cent of emissions enter either the upper ocean reservoir, 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.

Figure 1: calibrating damage-response functions

For Figure 1, we calibrate our response function for damages, presented as a percentage drop of output, to those in Nordhaus (2007) and Golosov et al. (2011). The GAMS source code for the DICE2007 model provides a precise description of the carbon cycle through a three-reservoir model. We use the linear algebra from the previous Appendix to convert the DICE reservoir model into a three-box model, using Matlab (the code is available in the supplementary package). This gives the parametric representation of the DICE2007 carbon cycle through $a = (0.575, 0.395, 0.029)$, $\eta = (0.306, 0.034, 0)$. To find the two remaining parameters π and ε for calibrating our representation to DICE2007, we consider a series of scenarios presented in Nordhaus (2008), each with a different policy 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 counterfactual scenario with equal emissions, apart from a the first period when we decreased emissions by 1GtCO_2 (Gigaton rather than Teraton used in the tex to keep the impulse marginal for the purposes here). Comparison of the damages, relative of output, then defines the response function θ_τ for that specific scenario. It turns out that the response functions are very close, and we take the average over all scenarios. Finally, we search for the

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

values of π and ε that approximate the average response θ_τ as closely as possible. We find $\varepsilon = 0.156$ [decade^{-1}], $\pi = 0.0122$ [TtCO2^{-1}].

Golosov et al. is matched by setting $a = (0.2, 0.486)$, $\eta = (0, 0.206)$; they have no temperature delay structure, so that $\varepsilon = 1$. Figure 1 presents the emissions damage responses.