# Consistent climate policies

Reyer Gerlagh and Matti Liski*

August 2, 2016

**Abstract**

We consider climate policies when time preferences deviate from the standard exponential type and there is no commitment to future policies. The conceptual and quantitative results follow from the observation that, with time-declining discounting, the delay and persistence of climate impacts provide a commitment device to policy-makers. We quantify the commitment value in a climate-economy model by solving time-consistent Markov equilibrium capital and emission taxes explicitly. The equilibrium returns on capital and climate investments are no longer equal, leading to a large increase in emission taxes, compared to a benchmark with equalized returns.

(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. This paper was previously titled "Carbon prices for the next thousand years". We thank anonymous reviewers, the editor (Krueger), 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 choice of the long-run discount rate is central when evaluating public projects with very long-run impacts such as the optimal response to climate change. While there is no general consensus on the discount rate to be used for different time horizons, there is certainly little evidence for using the same constant rate for all horizons. For example, recent revealed-preference evidence suggest that "Households discount very long-run cash flows at low rates, assigning high present value to cash flows hundreds of years in the future" (Giglio et al., 2015), consistent with earlier findings based on stated preference surveys.[1]

It is not unreasonable to think that policy-makers discount utility gains within their lifetime differently from those after their time. Moreover, if policies have impacts at the level of the economy and if future policy-makers' decisions cannot be dictated today, the setting becomes an intergenerational game between agents who make decisions in the order they enter the time-line.

We consider the climate-policy implications of discounting that deviates from the standard geometric case in such a policy game. Our analysis is normative in the sense that we describe the best-responding policies for a representative aggregate planner, given the future decision rules which, of course, depend on the equilibrium concept. We start with a Markov equilibrium that does not condition on the behavior of the previous planners and, thereby, has certain appeal in the intergenerational context. In the Markov equilibrium, the extreme delays and persistence of climate impacts provide a commitment device for policy-makers. Climate-related variables are much more persistent than the economic variables that we are used to, and so the climate policies of today have a peak impact on future utilities with a considerable delay, that is, after 60-70 years in our quantitative model.

Climate policies, when responding to future policies, should exploit the commitment to future utility impacts. When doing so, they depart from the idea that the same return requirement holds for all investments in the economy. Intuitively, the climate asset,

[1]For stated-preference evidence, 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 to zero after 300 years. See Cropper et al. (2014) for various interpretations of declining discount rate schedules.

through its extreme persistence, provides a "golden egg" for present-day policies, with commitment value arising endogenously in the equilibrium.

We assess the commitment value by restricting attention to a parametric class for preferences and technologies, and solving for the time-consistent Markov equilibrium policies explicitly. We introduce quasi-geometric discounting in a general-equilibrium growth framework,[2] building on Nordhaus' approach to climate-economy modelling (2008) and its recent gearing towards the macro traditions by Golosov, Hassler, Krusell, and Tsyvinsky (2014). Following Krusell, Kurusçu, and Smith (2002), we also describe the fiscal instruments, that is, the capital and carbon emission tax policies that decentralize the outcome of the policy game.

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 utility discount rate is 2.7 per cent. This is consistent with Nordhaus' DICE 2007 baseline scenario (Nordhaus, 2007),[3] giving 7.1 Euros per ton of $CO_2$ as the optimal carbon tax in the year 2010 (i.e., 34 Dollars per ton C). The first row provides the optimal, consistent-preferences, benchmark carbon price.

In the second row of Table 1 we show the Markov equilibrium capital and carbon taxes that are the optimal best responses in the climate policy game.[4] The Markov planner introduces a distorting tax on capital, as in Krusell et al. (2002), but complements this with a carbon tax that is considerably higher than Nordhaus' benchmark. The planner differentiates between the persistence of capital and climate investments. The persistence gap is important for the planner as each asset has its own commitment value through its effect on future utilities. The large increase in the carbon tax reflects the policy-maker's willingness to pay for a commitment to long-lasting utility impacts.

A zero capital tax is a natural benchmark that removes the distortion between the planner's and private returns on capital (Krusell et al. 2002). Similarly, ensuring equal returns for investments both in capital and climate is another natural benchmark as it removes a distortion in the asset portfolio. This second benchmark would use the

[2]Formally, we consider quasi-geometric discount functions as defined by Krusell et al. (2002). They are quasi-hyperbolic in the sense that, for certain parameter values, they bear resemblance to hyperbolic functions.

[3]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 a detailed documentation of DICE 2007.

[4]For the sake of illustration, we choose the short- and long-run time discount rates so that savings in the Markov equilibrium remain the same as in the first row.

economy's return on capital savings as the return requirement for climate investments: it would reset the carbon tax to the level identified by Nordhaus.[5] We know from Krusell et al. (2002) that a commitment to a zero capital tax increases welfare for both present and future consumers, and in our analysis we show that policy-makers can coordinate on zero capital taxes in a subgame-perfect equilibrium — even negative capital taxes can be sustained. But, interestingly, we find that even if policy-makers had the option to commit to use capital returns when evaluating climate investments, the commitment is not welfare-improving: all consumers are better off under the Markov equilibrium carbon taxes.

| | discount rate | | | | |
|---|---|---|---|---|---|
| | short-term | long-term | savings | capital tax | carbon tax |
| "Nordhaus" | .027 | .027 | 25% | 0% | 7.1 |
| Markov Equilibrium | .037 | .001 | 25% | 29% | 133 |
| "Stern" | .001 | .001 | 25% | 0 | 174 |

Table 1: Carbon taxes in EUR/t$CO_2$ year 2010.

The capital and climate policies are closely related. The Markov equilibrium savings become distorted under non-geometric discounting: there is a wedge between the marginal rate of substitution (MRS) and the marginal rate of transformation (MRT). The wedge arises from a shortage of future savings leading to higher capital returns than what the current policy-maker would like to see.[6] As is well-known in cost-benefit analysis, a distorted capital return is not the social rate of return for public investments.[7] Coordination of capital taxes between subsequent planners can mitigate the return distortions, bringing the future capital returns closer to the ones preferred today. However, even with coordinated capital policies consumers are better off following the Markov carbon tax, which remains above the carbon price based on capital returns.

The Markov carbon tax of 133 €$/CO_2$ may seem surprisingly ambitious given that the actual policies fall short of, rather than exceed, the benchmark proposals (7.1 €$/CO_2$ by Nordhaus, and 174 €$/CO_2$ by Stern (2006) in the third row of Table 1). However, our analysis is not descriptive. Instead, the focus is on a global planning problem with

[5]Nordhaus advocates this approach as follows: "[...] As this approach relates to discounting, it requires that we look carefully at the returns of alternative investments —at the real interest rate— as the benchmarks for climate investments." Nordhaus (2007, p. 692).

[6]This distortion is the same as in Barro (1999); and Krusell, Kurusçu, and Smith (2002).

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

intergenerational distortions but without more immediate obstacles to policies such as those arising from international free-riding. The objective is to find rules or institutions that support consistency in global policy-making over time. The analysis thus suggests that the gap between observed, low or non-existent, carbon taxes and the optimal one is even larger as the gap that is already suggested by models based on time-consistent preferences.

The relevance of hyperbolic discounting in the climate policy analysis has been acknowledged before; however, the broader general 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 that can be sustained in equilibrium; we introduce such policies in a well-defined sense.[8] 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. Our tractable general-equilibrium model features a joint inclusion of macro and climate policy decisions, with quasi-hyperbolic time preferences, and a detailed carbon cycle description — these features are all essential for a credible quantitative assessment of the commitment value.

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 such lack of a precise long-term view can imply a lower long-term discount rate than that for the short-term decisions; see Rubinstein (2003) for the procedural argument.[9] 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, 2014). We may 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

[8] Iverson (December 2012), subsequent to our working paper (June 2012), shows that the Markov equilibrium policy identified in our paper is unique when the equilibrium is constructed as a finite-horizon limit.

[9] 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.

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. Section 2 introduces the infinite-horizon climate-economy model and develops the climate system representation that allows us to decompose the contributions of the size, delay, and persistence of climate impacts to the carbon tax, and their interaction with the time-structure of preferences. This structured quantification is a contribution that applies even with constant discounting. For example, adding the delay of impacts to the setting in Golosov et al. (2014) reduces the carbon tax level by a factor of two.

Section 3 proceeds to the Markov equilibrium analysis for the policy maker (planner) and presents the main conceptual results. The results in Section 3 are presented for a parametric class of preferences and technologies. Complementing Section 3, the Appendix explicates the implications of the assumptions using general functional forms in a three-period illustration. In general, savings and climate investments can be strategic substitutes or complements for future savings and climate investments and, thus, lower or above those implemented in the case of full commitment to future actions. For a widely used parametric specification, covering for example those in Golosov et al. (2014), we show that climate investments are not used for manipulating future savings and vice versa; the "over-investment" in the climate asset reflects purely the greater persistence of the utility impacts in comparison with shorter term capital savings.[10] Thus, for the parametric class considered, the generations "agree" that a lower rate of return should be used for climate investments, so that current climate investments are not undermined by reduced future actions, even though, in principle, such a response is available to agents in equilibrium.

Section 4 introduces the decentralization of the planner's Markov equilibrium, separately for capital and carbon taxes. Section 5 provides the quantitative assessment of the conceptual results. To obtain sharp results in a field dominated by simulation models, we make specific assumptions. Section 6 discusses those assumptions, and some robustness analysis as well as extensions to uncertainty and learning. Section 7 concludes. All

[10] 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).

proofs, unless helpful in the text, are in the Appendix. The supplementary material cited in the text is available in a public folder.[11]

# 2 An infinite horizon climate-economy model

## 2.1 Technologies

For a sequence of periods $t \in \{1, 2, 3, ...\}$, 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 since climate-change, that arises because of historical emissions, changes production possibilities.[12] The economy has one final good. Capital depreciates in one period, leading to the following resource constraint 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.

For closed-form solutions, we put more structure on the primitives. We pull together the production structure as follows:

$$\begin{aligned} y_t &= k_t^{\alpha} A_t(l_{y,t}, e_t)\omega(s_t) &(2)\\ e_t &= E_t(z_t, l_{e,t}) &(3)\\ l_{y,t} + l_{e,t} &= l_t &(4)\\ \omega(s_t) &= \exp(-D_t), &(5)\\ D_t &= \sum\nolimits_{\tau=1}^{\infty} \theta_\tau z_{t-\tau} &(6) \end{aligned}$$

Gross production consists of: (i) Cobb-Douglas capital contribution $k_t^{\alpha}$ 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 $e_t$ total energy use in the economy; (iii) total

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

[12] History can matter for production also because 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. We abstract from the latter type of history dependence; the scarcity of conventional fossil-fuel resources is not binding when the climate policies are in place (see also Golosov et al., 2014).

energy $e_t = E_t(z_t, l_{e,t})$ with fossil fuels $z_t$ and labour $l_{e,t}$; and (iv) the climate impact given by function $\omega(s_t)$ capturing the output loss of production depending on the history of emissions from fossil-fuel use. We assume that the final-good and energy-sector outputs are differentiable, increasing, and strictly concave in labor, energy, and carbon inputs. The key allocation problem determining emissions at given time $t$ is how total labor $l_t$ is allocated between the final-good and energy sectors. By $f_t(k_t, l_t, z_t, s_t) = \max_{l_{y,t}} k_t^\alpha A_t(l_{y,t}, E_t(z_t, l_t - l_{y,t}))\omega(s_t)$, the production structure is as reported in the right-hand side of eq. (1). To simplify the analysis of the decentralized economy, we assume that $f_t(k_t, l_t, z_t, s_t)$ has constant returns to scale in $(k_t, l_t, z_t)$.[13]

## 2.2 Preferences

The consumption, fuel use, labor allocation, and investment choices generate sequence $\{c_\tau, z_\tau, k_\tau, s_\tau\}_{\tau=t}^{\infty}$ and per-period utilities, denoted by $u_t$, whose discounted sum defines the welfare at time $t$ as

$$w_t = u_t + \beta \sum\nolimits_{\tau=t+1}^{\infty} \delta^{\tau-t} u_\tau \tag{7}$$

where discounting is quasi-geometric and defined by factor $0 < \delta < 1$ for all dates excluding the current date when $\beta \neq 1$.

Let us next introduce the agents in the economy. There is a representative consumer who lives all periods $t = 1, 2, 3, ...$ The consumer at each $t = 1, 2, 3, ...$ has distinct preferences, discounting the immediate-next postponement of utility gains with factor $\beta\delta$ and then later postponements with $\delta$. This is the standard quasi-geometric discount function formulation that, for $\beta < 1$, becomes the quasi-hyperbolic approximation of a generalized hyperbolic discount function as, for example, in Krusell et al. (2002). The infinite-lived consumer can also be interpreted as a dynasty, that is, a chain of generations $t$ who disagree about the weights given to future generations' welfare. In the dynastic chain of generations, there is no individual-level behavioral inconsistency but, rather, only a differential discounting of future agents' utilities at different points in the future (as in Phelps and Pollak, 1968). In fact, the parametric model below remains tractable for an arbitrary sequence of discount factors, under certain conditions for boundedness,

[13] The assumption is not needed before Section 4 where it simplifies the equilibrium fiscal rules by leaving out a redistribution of rents since the total value of output is exhausted by factor compensations. The assumption requires that the nesting structure in (2)-(3) has constant returns to scale, and that the energy-labor composite can be written as $A_t(l_{y,t}, e_t) = \mathcal{A}_t(l_{y,t}, e_t)^{1-\alpha}$.

but the quasi-hyperbolic approximation allows sharper analytical results.[14] We take the discount function as a primitive element but, equivalently, one can take altruistic weights on future welfares as the primitive element and construct a discount function for utilities.[15] Together with the consumer, there is also a representative planner, who has the same preferences as the consumer. The planner sets taxes on energy use and also on the capital savings, understanding how the tax policies impact the competitive equilibrium where consumers rent their capital holdings and labor services to firms who combine energy with capital and labor in production. We first consider the planner's equilibrium, and introduce the decentralized economy with prices and taxes in Section 4.

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:

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

where $\Delta_u \geqslant 0$ is a given parameter. We include $\Delta_u D_t$ for a flexible interpretation of climate impacts that we develop through a social cost formula covering both direct utility and output losses.[16] In the calibration, we let $\Delta_u = 0$ to maintain an easy comparison with the previous studies.[17]

This parametric class for technologies and preferences builds on Brock-Mirman (1972). With geometric discounting, $\beta = 1$, the parametric class for technologies and preferences leads to a consumption choice model that is essentially the same as in Brock-Mirman (1972); see Golosov et al. (2014). In particular, the currently optimal policies depend

[14] The qualitative effect of declining time preference can be understood by studying the quasi-geometric case. See Iverson (2012) for the extension of our analysis to the flexible discounting case.

[15] We explicate this in the Appendix with a three-period model; see Saez-Marti and Weibull (2005) for the general equivalence between generation-specific welfare functionals and discount functions. Note that the preferences are specific for generation $t$, and in that sense, $w_t$ is different from the generation-independent social welfare function (SWF) as discussed, e.g., in Goulder and Williams (2012) and Kaplow et al. (2010).

[16] See Tol (2009) for a review of the existing damage estimates; the estimates for intangible losses are very uncertain and mostly missing. Including such losses in the social cost formulas can be helpful if one is interested in gauging how large they should be to justify 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.

only on the state of the economy, say, at year 2015, so that policies become free of the details of the energy sector as captured by $A_t$ and $E_t$, although the full outcome path for the economy depends on these details.[18] In the policy game, with $\beta \neq 1$, we show that the Markov equilibrium has the same convenient properties.

## 2.3 Damages and carbon cycle

We now provide micro-foundations for equations (5)-(6) that formalize the productivity impact of climate change. Climate damages are interpreted as reduced output, depending on the history of emissions through state variable $D_t$ that measures the global mean temperature increase. The weight structure of past emissions in (6) 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

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

**Definition 1** *(closed carbon cycle) No carbon leaves the system: column 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 $\mathbf{Q} = [\ v_1 \ \ ... \ \ v_n\ ]$ is a matrix of linearly independent eigenvectors $v_\lambda$ such that

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

[18] We study the future scenarios and specify $A_t$ and $E_t$ in detail in Gerlagh&Liski (2016). 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", obtained by allocating 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; de-carbonization is necessary for achieving long-term reduction targets.

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 $\lambda_i$ 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 reconverted 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 (which is the period length assumed in the quantitative analysis). Let $\mu$ 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, $\lambda_i = 1$,[19]

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

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

This observation will have important economic implications when the discount rate is small. We say that the carbon cycle has incomplete absorption if this box is non-negligible:

**Definition 2** *(incomplete absorption) Some $CO_2$ remains forever in the atmosphere: there is one box $i \in \mathcal{I}$ that has no depreciation, $\eta_i = 0$ and is non-negligible, $a_i > 0$.*

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.[20] 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:[21]

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

with $\pi > 0$, a constant characterizing sensitivity of damages to the atmospheric $CO_2$.[22]

Let $\varepsilon$ be the adjustment speed of temperatures and damages, so that we can write for the dynamics of damages:[23]

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

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

**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*

$$\eta_i = 1 - \lambda_i$$
$$a_i = \frac{q_{1,i}}{1+\mu} c_i$$

[20]See Pindyck (2013) for a critical review.

[21]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.

[22]Section 6 reports our sensitivity analysis of the results to this approximation.

[23]The equation follows from an explicit gradual temperature adjustment process, as modeled in DICE also. See Gerlagh and Liski (2012) for details.

*For a one-box model (with no indexes i), the maximum impact occurs at time between the temperature lifetime $1/\varepsilon$ and the atmospheric $CO_2$ lifetime $1/\eta$.*[24]

Theorem 1 describes the carbon cycle in terms of a system of independent atmospheric boxes, where $\mathcal{I}$ denotes the set of boxes, with share $0 < a_i < 0$ of annual emissions entering box $i \in \mathcal{I}$, and $\eta_i < 1$ its carbon depreciation factor. The last line of the theorem informs us that long delays in climate change between emissions and damages are described through small values for $\varepsilon$ and $\eta$. The substantial implications of the delays become clear in Proposition 4. The essence of the response is very intuitive. Parameter $\eta_i$ 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 $\varepsilon$ to the long-run equilibrium damage sensitivity $\pi$, 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, $\theta_\tau = 0$.

# 3 Markov equilibrium of the planning game

In this Section, we assume that each planner at $t$ controls aggregates $(k_{t+1}, z_t)$ directly. The outcome of the planning game gives the equilibrium marginal social cost of using one more unit of carbon energy for each planner $t$. We call the social cost defined this way as the equilibrium carbon price. In Section 4, where we decentralize the planning equilibrium through a set of fiscal rules, the carbon price becomes the equilibrium tax on emissions.

[24]The $CO_2$ lifetime is the expected number of periods that an emitted CO2 particle remains in the atmosphere. The temperature life time is the average duration that a fictitious temperature shock persists.

## 3.1 The game

The game is played between a sequence of planners, indexed with $t = 1, 2, 3, ...$ Planner $t$ has a Markov strategy, mapping from the current state to savings and emissions. Before defining the Markov policies, we must identify the state relevant for the continuation payoffs. When written in full, the state reads as $(k_t, \Theta_t)$, where $\Theta_t = (D_t, S_{1,t}, ..., S_{n,t})$ collects the vector of climate state variables. However, the climate affects the continuation payoffs only through the weighted sum of past emissions, as expressed in (6); we replace $\Theta_t$ by $s_t$ since the history is the sufficient statistics for $\Theta_t$.

The Markov policies, denoted by $k_{t+1} = \mathcal{G}_t(k_t, s_t)$ and $z_t = \mathcal{H}_t(k_t, s_t)$, do not condition on the history of past behavior (see Maskin and Tirole, 2001).[25] Given the parametric class for preferences and technologies, a Markov equilibrium can be found from a particular parametric class for $k_{t+1} = \mathcal{G}_t(k_t, s_t)$ and $z_t = \mathcal{H}_t(k_t, s_t)$ that, together with the implied welfare, we define next.

## 3.2 Planner's welfare

For given policies $\mathcal{G}_t(k_t, s_t)$ and $\mathcal{H}_t(k_t, s_t)$, we can write welfare in (7) 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. More specifically, consider the payoff implications from a sequence of constants $(g_\tau, h_\tau)_{\tau \geqslant t}$ where $0 < g_t < 1$ is the share of the gross output invested,

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

and $h_t$ is the climate policy variable that measures the social cost of current emissions; it equals the current utility gain from increasing emissions marginally, $h_t = \frac{\partial y_t}{\partial z_t}\frac{\partial u_t}{\partial c_t}$. This measure, through the functional assumptions, defines the marginal product of the fossil fuel use, the carbon price, as

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

Similarly as $g_t$ measures the stringency of the savings policy, $h_t$ measures the stringency of the climate policy. In particular, the marginal product of carbon (the planner's

[25] We allow the policies to depend on time, which in turn allows us to analyze the payoff implications of changes in policies; the (symmetric) equilibrium Markov policies as defined below do not depend on time.

carbon price), $\frac{\partial y_t}{\partial z_t}$, is monotonic in policy $h_t$, which allows an interchangeable use of these two concepts.[26] Now, 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:

**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 $\zeta_1$ 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:

**Corollary 1** *For a closed carbon cycle with incomplete absorption, $\frac{\partial\Omega(s_{t+1})}{\partial z_t} \to \infty$ as the long-run time discount factor $\delta \to 1$.*

The result has strong implications for the policies.

## 3.3 Markov policies

Theorem 2 describes continuation welfares for a class of policies, and now we proceed to a Markov equilibrium that can be found from this class.

**Definition 3** *A Markov equilibrium is a sequence of savings and carbon price rules $(g_t, h_t)_{t \geq 1}$ satisfying (13) and (14) such that $(g_t, h_t)$ maximizes welfare at each $t$, given $(g_\tau, h_\tau)_{\tau > t}$.*

[26]We show this in Lemma 5 of the Appendix.

More precisely, just below, we look for a symmetric Markov equilibrium where all generations use the same policy $(g_\tau, h_\tau)_{\tau \geqslant t} = (g, h)$.[27] [28]

Krusell et al. (2002) describe the savings policies for a one-sector model in the same parametric class with quasi-geometric preferences. Our setting is more complicated since, with two-sectors, the policies for the sectors can be either strategic substitutes or complements; however, the Brock-Mirman (1972) structure for the consumption choice and exponential productivity shocks from climate change eliminates such interactions, and thus the savings and climate policies become separable.[29] Each generation takes the future policies, captured by constants $(g_\tau, h_\tau)_{\tau > t}$ in (13)-(14), as given and chooses its current savings to satisfy

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

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

**Lemma 1** *(savings) The planner's Markov 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 $g_{\tau>t} = g^{\beta=1} = \alpha\delta$ for future decision-makers would maximize the wealth as captured by $W_{t+1}(k_{t+1}, s_{t+1})$; however, equilibrium $g^*$ with $\beta < 1$ is less than $g^{\beta=1} = \alpha\delta$ 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

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

[27] There can be exogenous technological change and population growth, but the form of the objective, (8) combined with (7), ensures that there will be an equilibrium where the same policy rule will be used for all $t$.

[28] We will construct a natural Markov equilibrium where policies have the same functional form as when $\beta = 1$. Moreover, Iverson (2012) shows for this model that the Markov equilibrium considered here is the unique limit of a finite horizon equilibrium. For multiplicity of equilibria in related settings, see Krusell and Smith (2003) and Karp (2007).

[29] In the online Appendix, we develop a three-period model with general functional forms to explicate the interactions eliminated by the parametric assumptions.

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 $\tau_t^{z(\beta,\delta)}(=\partial y_t/\partial z_t)$. Given Theorem 2, carbon price $\tau_t^{z(\beta,\delta)}$ can be obtained:

**Proposition 1** *The planner's Markov equilibrium carbon price is*

$$\begin{aligned} \tau_t^{z(\beta,\delta)} &= h^*(1-g^*)y_t && (16)\\ h^* &= \Delta \textstyle\sum_{i\in\mathcal{I}} \dfrac{\beta\delta a_i \pi\varepsilon}{[1-\delta(1-\eta_i)][1-\delta(1-\varepsilon)]} && (17)\\ \Delta &= (\dfrac{1}{1-\alpha\delta}+\Delta_u) \end{aligned}$$

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. If future policies could be dictated today, the externality cost would be higher: $h^{\beta=1} > h^*$.[30]

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 (5), $\partial \ln(c_{t+\tau})/\partial D_{t+\tau} = -1$. As there is also the direct utility loss, captured by $\Delta_u$ 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 $t+\tau$,

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

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

[30]It is not difficult to verify that $h^*(1-g^*)$ is increasing in $\beta$. The current planner would like to see the future planners to save more and to choose a larger carbon price.

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 planner's Markov equilibrium carbon price depends on the delay structure in the carbon cycle captured by parameters $\eta_i$ and $\varepsilon$. 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 Corollary 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]

# 4 Decentralization

The Markov equilibrium for the planning game identifies an allocation $\{c_\tau, z_\tau, k_\tau, s_\tau\}_{\tau=t}^{\infty}$, but it is yet silent about the economic instruments implementing the outcome. Following Krusell et al. (2002), we now re-interpret the game as one where each planner chooses fiscal instruments (in our economy, current taxes on private savings and emissions), without ability to commit to future taxes. We take the taxes as functions of the state and derive them explicitly, after characterizing the recursive competitive equilibrium resulting from given tax functions. Denote the taxes on capital investments and emissions

[31] If carbon depreciates quickly, $\eta_i >> 0$, then the carbon price will be less sensitive to the discount factor $\delta$. 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.

by $(\tau_t^k, \tau_t^z) = (\tau_t^k(k_t, s_t), \tau_t^z(k_t, s_t))$, respectively.[32]

Factor markets are competitive. The representative firm maximizes profits given the price of capital capital $r_t$, the emissions price as given by policy $\tau_t^z$, and wages $q_t$:

$$\frac{\partial f_t(k_t, l_t, z_t, s_t)}{\partial k_t} = r_t \tag{18}$$

$$\frac{\partial f_t(k_t, l_t, z_t, s_t)}{\partial z_t} = \tau_t^z \tag{19}$$

$$\frac{\partial f_t(k_t, l_t, z_t, s_t)}{\partial l_t} = q_t. \tag{20}$$

The equilibrium price of capital $r_t$ is endogenous, equalizing the previous period's savings and current factor demand. Without climate policies, the competitive market factor price for emissions is zero. Policy-determined factor price $\tau_t^z$ sets a price on emissions in production. Labor $l_t$ is supplied inelastically and, through (20), its equilibrium factor compensation $q_t$ is endogenous. The consumer takes the aggregate law of motions for $k_t$ and $s_t$ as given, as well as future factor prices and tax rates as functions of aggregate variables $r_t = r_t(k_t, s_t), q_t = q_t(k_t, s_t)$.

Revenues from emission taxes and capital investment taxes are returned lump sum to households, denoted by $T_t = T_t(k_t, s_t)$. To separate the consumer's decisions from the planner's, we denote the former by superscript $i$. The consumer's budget constraint is

$$c_t^i + (1 + \tau_t^k)k_{t+1}^i = q_t l_t^i + r_t k_t^i + T_t, \tag{21}$$

$$T_t = \tau_t^k k_{t+1} + \tau_t^z z_t. \tag{22}$$

The consumer's only decision is to choose how much capital to save $k_{t+1}^i$, given total income consisting of factor service compensations and the lump sum transfer of tax returns $T_t$. The consumer maximizes utility $u_t^i = \ln(c_t^i) - \Delta_u D_t$ and future welfare

$$u_t^i + \beta\delta w_{t+1}^i \tag{23}$$

with the future values defined through[33]

$$w_t^i = W_t^i(k_t^i; k_t, s_t) = u_t^i + \delta W_{t+1}^i(k_{t+1}^i; k_{t+1}, s_{t+1}). \tag{24}$$

[32]For subscript $t$ in policies, consider Proposition 1, and conjecture that the implemented tax policy coincides with the Markov policy from the planning game, $\tau_t^z = h^*(1 - g^*)y_t$. Output $y_t$ depends on state $(k_t, s_t)$ but also on time since we do not restrict to stationary technologies and labor supply.

[33]We also write superscript $i$ for individual value functions, to separate these from the aggregate value functions. The individual functions are, though, not different between individuals.

**Definition 4** *Given tax rules* $(\tau_t^k(k_t, s_t), \tau_t^z(k_t, s_t))$*, complemented with budget neutral lump-sum transfers* $T_t(k_t, s_t)$*, a recursive competitive equilibrium consists of individual savings policies* $k_{t+1}^i = \mathcal{G}_t^i(k_t^i; k_t, s_t)$*, value function* $W_t^i(k_t^i; k_t, s_t)$*, price functions* $r_t = r_t(k_t, s_t), q_t = q_t(k_t, s_t)$*, and emissions* $z_t = z_t(k_t, s_t)$ *such that (i)* $k_{t+1}^i = \mathcal{G}_t^i(k_t^i; k_t, s_t)$ *solves the consumer's problem, (ii)* $W_t^i(k_t^i; k_t, s_t)$ *is generated by the consumer's policy, (iii) market clearing conditions (18)-(20) hold, and (iv) aggregate capital dynamics satisfy* $k_{t+1} = \mathcal{G}_t(k_t, s_t) = \mathcal{G}_t^i(k_t^i; k_t, s_t)$ *when* $k_t^i = k_t$*.*

We show now that a constant capital tax $\tau_t^k = \tau^k > 0$ and a carbon tax proportional to consumption (14), $\tau_t^z = h(1 - g_t)y_t$, can be used to decentralize the planner's Markov equilibrium. For the carbon tax, the planner's marginal product of carbon coincides with market clearing condition (19) if $\tau_t^z = h^*(1 - g^*)y_t$ as defined in Proposition 1. This tax on emissions is clearly part of the decentralization. For the fiscal rules for capital, we show, based on Krusell et al. (2002), that when facing a constant capital savings tax, the households decisions in the recursive competitive equilibrium have a simple parametric form:

**Lemma 2** *Consider a constant tax* $\tau^k$ *on capital investments, and an emission tax rule proportional to consumption* $\tau_t^z(k_t, s_t) = h_t(1-g_t)y_t$*, where* $g_t \equiv k_{t+1}/y_t$*. In the recursive competitive equilibrium, aggregate savings are a constant share of output,* $g_t = g$*, and the equilibrium is parametrically characterized through*

$$\mathcal{G}_t^i(k_t^i; k_t, s_t) = g\frac{k_t^i}{k_t}y_t$$
$$W_t^i(k_t^i; k_t, s_t) = a_t + b\ln(k_t) + c\ln(k_t + \varphi k_t^i)$$

*with* $a_t$ *independent of capital, and parameters satisfy*

$$\begin{aligned} b &= \frac{-(1-\alpha)}{(1-\alpha\delta)(1-\delta)}, \\ c &= \frac{1}{1-\delta}, \\ \varphi &= \frac{\alpha - g(1+\tau_k)}{1-\alpha+g\tau_k}, \\ g &= \frac{1}{1+\tau^k}\frac{\alpha\beta\delta}{1+\delta(\beta-1)}. \end{aligned} \tag{25}$$

While the carbon tax internalizes the climate externality, the planner's tax on savings has a more subtle reasoning. The current planner controls total resources for the economy; the decentralized decisions by consumers build on a linear income constraint given factor prices, in which savings provide more commitment in equilibrium. Without capital taxation $\tau^k = 0$, the competitive equilibrium savings share is given by

$$g^{\tau^k=0} = \frac{\alpha\beta\delta}{1+\delta(\beta-1)} > g^*$$

where $g^*$ is the Markov equilibrium savings fraction from the planning game (Lemma 1). As in Krusell et al. (2002), the laissez-faire decentralized savings exceed those in the planning game, and thus the implementation of the planning game policies requires a positive tax on savings $\tau^k > 0$, iff $\beta < 1$ :

**Theorem 3** *Consider the policy game where each planner $t$ controls taxes and the resulting lump-sum transfers, $(\tau^k, \tau_t^z, T_t)$, given future fiscal rules $\{\tau_\tau^k(k_\tau, s_\tau), \tau_\tau^z(k_\tau, s_\tau), T_t(k_\tau, s_\tau)\}_{\tau>t}$. Markov equilibrium taxes that decentralize the planning game outcome $(g^*, h^*)$ are*

$$\begin{aligned} \tau^{k*} &= \frac{\delta(1-\alpha)(1-\beta)}{1-\delta(1-\beta)}, \\ \tau_t^{z*} &= h^*(1-g^*)y_t. \end{aligned}$$

*complemented with lumpsum transfers (22).*

## 4.1 Taxes on capital: a closer look

Krusell et al. (2002) show that removing capital taxes increases welfare for all generations. Is it possible for the planners to sustain zero capital taxation as a subgame perfect Nash equilibrium? For this question, with the aid of Theorem 2, it is useful to state how changes in future savings impact current welfare:

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

$$\frac{\partial w_t}{\partial g_\tau} > 0 \text{ iff } g_\tau < \alpha\delta.$$

The future savings maximize current welfare if they are consistent with the long-term time preference $\delta$; that is, if $g = \alpha\delta$. An equilibrium policy that manages to take future savings closer to $g = \alpha\delta$ increases current welfare.

Consider constant capital tax $\widehat{\tau}^k$ that the current planner would like to propose for all planners, with the requirement that the planner at $t$ has to comply with the proposal as well. In view of Lemma 3, the planner would like to propose to all *future* planners a tax implementing $g = \alpha\delta$ (which requires subsidies to savings). This proposal is ruled out by the current planner's own incentive constraints. But, the current planner is willing to give up some consumption and increase savings, by lowering the capital tax below its Markov equilibrium level, if all subsequent decision-makers will follow suit when facing the same decision.

**Proposition 2** *For $\beta < 1$, zero capital tax $\tau^k = 0$ increases welfare for all generations. The constant capital tax that maximizes current welfare is negative:*

$$\widehat{\tau}^k = \frac{-\alpha\delta(1-\delta)(1-\beta)}{1-\delta(1-\beta)} < 0.$$

*Both $\tau^k = 0$ and $\widehat{\tau}^k$ can be sustained in subgame perfect equilibrium where a deviation triggers all future planners to revert to the Markov equilibrium tax $\tau^{k*}$.*

Note that the Proposition considers only capital tax coordination, keeping the carbon tax at the Markov level.[34] Yet, if planners coordinate on capital taxes, the Markov carbon tax changes according to

$$\tau_t^{z*} = h^*(1-g)y_t.$$

We have considered three potential capital tax rules that are all in principle equilibrium rules: subsidy $\widehat{\tau}^k$, zero tax $\tau^k = 0$, and the Markov equilibrium capital tax, $\tau^{k*}$. These taxes are ordered $\widehat{\tau}^k < 0 < \tau^{k*}$, and implemented savings satisfy $g^* < g^{\tau^k=0} < \widehat{g} < \alpha\delta$, respectively. The respective carbon taxes have a reverse ordering.

For the further analysis of carbon taxes, it proves useful to define a utility-discount factor $0 < \gamma < 1$ for consumption, 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,

$$\gamma = \frac{u'_t}{u'_{t+1}R_{t,t+1}} = \frac{c_{t+1}}{c_t R_{t,t+1}} = \frac{c_{t+1}}{c_t}\frac{k_{t+1}}{\alpha y_{t+1}} = \frac{g}{\alpha}. \tag{26}$$

Capital tax $\tau^{k*}$ in Theorem 3 implements savings policy $g = g^*$ and thus defines

$$\gamma^* = \frac{\beta\delta}{1+\alpha\delta(\beta-1)}. \tag{27}$$

With no capital tax, the utility discount factor would be

$$\gamma^{\tau^k=0} = \frac{\beta\delta}{1+\delta(\beta-1)}.$$

[34] We define the subgame-perfect equilibrium and strategies considered in the Proposition formally in the Appendix.

With savings subsidy $\widehat{\tau}^k$, the utility-discount factor obtained for $\widehat{g}$ from (26) becomes

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

with $\beta\delta < \gamma^* < \gamma^{\tau^k=0} < \widehat{\gamma} < \delta$. Thus, in this sense, coordination of capital policies increases the equilibrium patience. The potential welfare effects from the subgame perfect savings tax rule are particularly stark for vanishing long-run discounting; the welfare maximizing subgame-perfect equilibrium converges to the golden rule, while the Markov savings policy remains bounded away from the golden rule:

**Corollary 2** *For $\delta \to 1$,*

$$\widehat{\tau}^k \to 0, \widehat{g} \to \alpha, \widehat{\gamma} \to 1.$$

## 4.2 Taxes on carbon

What is a Pigouvian tax? The Markov outcome of the planning game implements tax $\tau_t^{z*}$ that conforms with the standard definition in the sense that the tax internalizes the discounted future utility costs of marginal increases in energy use today. But the discount factor used in this evaluation is not the same as the one used for capital investments, $\gamma^*$. For a benchmark, we now develop a carbon tax rule that is based on the capital returns, and then we show why and how the Markov equilibrium tax deviates from the principle that all investments in the economy should earn the same return.

**Proposition 3** *Consider carbon tax $\tau_t^{z(\gamma)}$ that internalizes all future costs of current emissions when utilities are discounted with geometric factor $\gamma$. It equals*

$$\tau_t^{z(\gamma)} = h^\gamma (1-g) y_t \tag{28}$$

$$h^\gamma = \Delta^\gamma \sum\nolimits_{i\in\mathcal{I}} \frac{\gamma\pi a_i \varepsilon}{[1-\gamma(1-\eta_i)][1-\gamma(1-\varepsilon)]} \tag{29}$$

$$\Delta^\gamma = \frac{1}{1-\alpha\gamma} + \Delta_u.$$

When $\gamma$ is taken from Euler equation $u_t' = \gamma u_{t+1}' R_{t,t+1}$, giving $\gamma = g/\alpha$ through (26), the tax will lead to today's marginal carbon product, $MCP_t$, to equal the sum of future marginal carbon damages caused by current emissions, $MCD_{t,T}$, for $T > t$, discounted to the present with equilibrium capital returns, $MCP_t = \sum_{T>t} MCD_{t,T}/R_{t,T}$.

The Markov planner deviates from the principle of equalized returns on investments since it looks at the real time-preference structure and understands that the equilibrium compound capital return between $t$ and some future period $T > t + 1$, no longer reflects how the current policy-maker sees the consumption trade-offs: $MRS_{t,T} < R_{t,T}$, future capital returns are excessive from the current point of view. As a result, the Markov outcome implies tighter carbon policies than the one using capital returns: $MCP_t = \sum_{T>t} MCD_{t,T}/MRS_{t,T} > \sum_{T>t} MCD_{t,T}/R_{t,T}$.[35] We establish now the precise conditions for the policies to differ:

**Proposition 4** *For $\beta, \delta < 1$, the Markov equilibrium carbon tax $\tau_t^{z*}$ strictly exceeds $\tau_t^{z(\gamma^*)}$ if climate change delays are sufficiently long. Formally, ratio $\tau_t^{z*}/\tau_t^{z(\gamma)}$ is continuous in parameters $\beta, \delta, \eta_i, \varepsilon, a_i$, and $\gamma$. Evaluating at $\gamma = \gamma^*$, and letting $\eta_i, \varepsilon \to 0$ gives*

$$\tau_t^{z*}/\tau_t^{z(\gamma)} > 1. \tag{30}$$

*The result also holds if $\gamma = \gamma^{\tau^k=0}$ (no capital tax), or if $\gamma = \widehat{\gamma}$ (capital tax coordination).*

If climate delays are long, as captured by $\eta_i$ and $\varepsilon$ (see the last line of Theorem 1), climate policies affect future utility levels for longer periods than capital investments (limit properties can be hard to assess, and for this purpose, we calculate the ratio for various scenarios in the last column of Table 2 bleow).[36] The equilibrium coordination of capital taxes (Proposition 2) increases savings and thus brings future capital returns closer to how the current planner sees the consumption-savings trade-offs. The last part of the proposition states that for very long climate delays, such coordination cannot fully eliminate the incentive to use the climate asset as a commitment device. The last part also states that if capital taxes are not in the set of instruments (perhaps because of policy frictions), so that planners have an institutional commitment to zero capital taxes and make only the carbon tax choices in equilibrium, the commitment value delivered by climate impacts will still be exploited in equilibrium.

The next proposition does not consider the long delays between emissions and impacts, but their persistence. If the climate system is sufficiently persistent, as in Remark 1, the Markov decision-maker values the commitment to future utility impacts. The commitment value has no bound in the following sense:

**Proposition 5** *For a closed carbon cycle with incomplete absorption and $\beta < 1$: $\tau_t^{z*}/\tau_t^{z(\gamma)} \to \infty$ for any $\gamma < 1$ as $\delta \to 1$.*

[35]In the Appendix we provide a three-period model with general functional forms to explicate the assumptions on preference and technologies that are needed for this result to follow. The parametric class considered in the infinite-horizon model satisfies the assumptions.

[36]Note that in the limit, $\varepsilon = 0$, and both carbon prices are zero, $\tau_t^{z*} = \tau_t^{\gamma} = 0$. The proposition states that there is a neighbourhood around $\eta_i = \varepsilon = 0$ in which $\tau_t^{z*} > \tau_t^{\gamma}$.

When no carbon leaves the system, a fraction of the temperature increase caused by current emissions never dies out. Then, with low long-run discounting, the difference between the two carbon taxes becomes unbounded. Yet, recall that if planners succeed in coordinating on the welfare maximizing capital-subsidy, $\widehat{\tau}^k$, the utility discount factor $\gamma$ converges to 1, and the proposition does not apply. That is, climate as a commitment device is not needed when generations can coordinate capital taxes and associate non-vanishing weights to far-future utilities.

We have seen that the Markov planner's equilibrium capital taxation moves future savings in the wrong direction, away from the coordinated optimum, $g^* < g^{\tau^k=0} < \widehat{g}$. Eliminating part of this distortion, that is, reducing capital taxes, improves welfare. This is why it is possible to sustain some coordination of capital taxation, $\widehat{\tau}_k$, in equilibrium. For carbon taxes, Proposition 4 suggests a distortion between capital and climate returns: $\tau_t^{z*} > \tau_t^{\gamma}$. Possibly, Markov carbon policies also distort the equilibrium in the wrong direction. Can we improve welfare by eliminating the gap between the returns on capital and climate investments? We find that this is not the case. First, with the aid of Theorem 2, we state how changes in future climate policies impact current welfare:

**Lemma 4** *For $\beta \neq 1$ and any given $\tau > t$,*

$$\frac{\partial w_t}{\partial h_\tau} > 0 \text{ iff } h_\tau < h^\delta.$$

Policy variable $h_t$ measures the strictness of the future climate policy so that any equilibrium policy change that takes $h_t$ closer to $h^\delta$ improves current welfare, where $h = h^{\gamma=\delta}$ is defined in Proposition 3. Proposition 4 shows that the carbon tax based on the market returns is lower than the equilibrium Markov carbon tax, and thus Lemma 4 tells us that such a carbon tax rule decreases the present welfare if applied in the future. Add to this insight that, by definition, the Markov carbon tax maximizes present welfare for fixed future tax rules. It then becomes clear that a move from a Markov carbon tax policy to externality pricing based on equalized returns for all assets must reduce welfare throughout. We state the result formally in:

**Proposition 6** *For given $(k_t, s_t)$, sufficiently slow climate change, and any given constant capital tax $\tau^k$: continuation policies $(\tau^k, \tau_\tau^{z(\gamma)})_{\tau \geqslant t}$ with $\gamma = \gamma^*$, or $\gamma = \gamma^{\tau^k=0}$, or $\gamma = \widehat{\gamma}$, all imply a lower welfare at $t$ than policies $(\tau^k, \tau_\tau^{z*})_{\tau \geqslant t}$ that price carbon according to the Markov rule $\tau_\tau^{z*}$.*

Policy $(\tau^k, \tau_\tau^{z(\gamma)})_{\tau \geqslant t}$ with $\gamma$ based on capital returns, conforms to the idea that all assets earn the same return in the economy. The remarkable feature of the above proposition is that the efficiency gain from equal returns on the capital and climate assets cannot prevent a decrease of welfare, not as a second-order effect, but as a first-order effect.[37] For this reason, such a cost-benefit requirement cannot be sustained as a welfare improving subgame-perfect Nash equilibrium. On reflection, the result is not surprising since a requirement for equal returns on capital and climate removes the equilibrium commitment that is provided by the persistent climate asset. The results holds even when the planners can coordinate the capital taxes.

# 5 Quantitative assessment

Our analysis is positive in the sense that the basis calibration of the parameters is consistent with the observed savings rate. But the climate policies we determine assume global coordination in the Markov equilibrium, and in addition intergenerational coordination in the subgame-perfect equilibrium. We abstract from both international free-riding and intertemporal political frictions. In that sense we provide a normative perspective on the level of carbon taxes that would maximize welfare, under a set of well-specified conditions for the policy game.

## 5.1 Emissions-damage response

Figure 1 shows the life-path of losses (percentage of total output) caused by an impulse of one Teraton of Carbon [Tt$CO_2$] in the first period, contrasted with a counterfactual path without the carbon impulse.[38] The output loss is thus measured per Tt$CO_2$, and it equals $1-\exp(-\theta_\tau)$, $\tau$ periods after the impulse. The graphs are obtained by calibrating the damage-response, that is, weights $(\theta_\tau)_{\tau\geqslant 1}$ in (6), to three cases.[39] Matching Golosov et al.'s (2014) 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 a thousand years.

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

[38] One Tt$CO_2$ equals about 25 years of global $CO_2$ emissions at current levels (40 Gt$CO_2$/yr.)

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

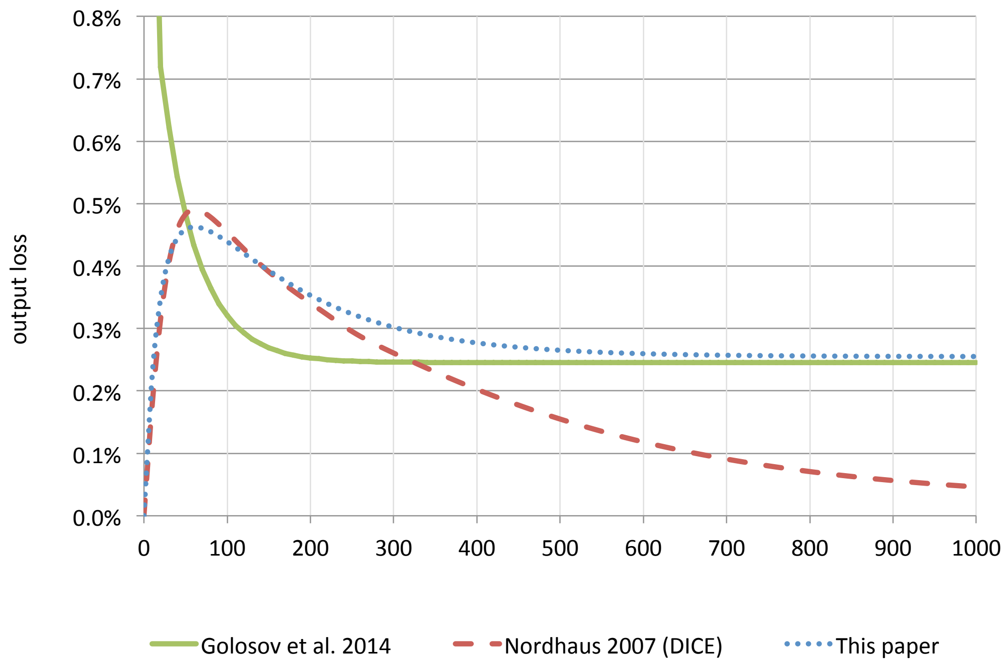


Figure 1: Emissions-damage response for three specifications

Our emissions-damage response, used in the quantitative part and depicted in Figure 1 ("this paper") 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 leading to the following emission shares and depreciation factors per decade:[40]

$$a = (.163, .184, .449)$$
$$\eta = (0, .074, .470).$$

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].[41] 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

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

[41]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 $(2.56 - 0.79)/(2.13 - 1) = 1.56\%/TtCO2$.

evidence (Solomon et al. 2007).[42] See the Appendix for further details.

## 5.2 Capital and carbon taxes

For the quantitative magnitudes of the results, we exploit the closed-form price formulas to evaluate the taxes that the model predicts the present day.

The model is decadal (10-year periods),[43] 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 $\alpha$ follows from the assumed time-preference structure $\beta$ and $\delta$, 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 ($\beta = 1$,$\delta = 0.761$), consistent with $\alpha = g/\delta = 0.329$. Choices for the climate-economy parameters are specified in Section 5.1.

With consistent preferences ($\beta = 1$), our model reproduces the carbon tax levels of the more comprehensive climate-economy models such as DICE (Nordhaus, 2008). We then introduce a difference between short- and long-term discounting, $\beta < 1$, while controlling for the effective discounting in the economy. The quantitative evaluation is thus structured such that we control for the capital savings, using the relationship between equilibrium savings $g$ and discount factors $\beta, \delta$— this allows keeping the Nordhaus case as a well-defined benchmark and exploring how the time preferences matter for the equilibrium tax structure.[44]

[42] In Figure 1, the main reason for the deviation from DICE 2007 is that DICE assumes an almost full $CO_2$ storage capacity for the deep oceans, while large-scale ocean circulation models point to a reduced deep-ocean overturning running parallel with climate change (Maier-Reimer and Hasselman 1987). The positive feedback from temperature rise to atmospheric $CO_2$ through the ocean release is essential to explain the large variability observed in ice cores in atmospheric $CO_2$ concentrations. We note that our closed-form model can be calibrated to very precisely approximate the DICE model (Nordhaus 2007). Section 6 discusses further on the surprising prediction power of our carbon pricing formula for the DICE results.The DICE 2013 model has updated the ocean carbon storage capacity.

[43] 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.

[44] One could also consider calibrating the short- and long-run discount rates. However, we are unaware of any empirical paper that reports revealed-preference data for the pure rate of time-preference over horizons such as 2-25, 26-50, 51-100 and so on years. Obviously, there is an extensive literature on the time structure of preferences in the context of self-control, but this literature looks at intra-personal short-term decisions and not the inter-generational trade-offs relevant for this paper. Giglio et al. (2015) measure the discount of leasehold property versus freehold property. We interpret Giglio et al.'s finding as a measure of the time-structure of returns on a specific private asset (houses). The time structure

The parameter choices result in a consistent-preferences Pigouvian tax of 7.1 Euro/tCO$_2$, equivalent to 34 USD/tC, for 2010.[45] 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 tax into three contributing parts: a base price that would apply if damages are *immediate* and temporary, an accumulation factor for the *persistence* of damages, and a discount factor for the *delay* in damages. First, consider the one-time costs assuming full immediate damages ($ID$) taking place in the immediate next period,

$$ID = \beta\delta\Delta\pi(1-g)y_t. \tag{31}$$

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 ($PF$),

$$PF = \sum_{i\in\mathcal{I}} \frac{a_i}{[1-\delta(1-\eta_i)]}, \tag{32}$$

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

$$DF = \frac{\varepsilon}{1-\delta(1-\varepsilon)}. \tag{33}$$

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 savings policy remains the same. The first row reproduces the efficient carbon tax case assuming consistent preferences when the annual utility discount rate is set at 2.7 per cent: this row presents the carbon tax under the same assumptions as in Nordhaus (2007). Keeping the equilibrium time-preference rate at 2.7 per cent per year, 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.

We invoke Weitzman's (2001) survey for obtaining some guidance in choosing the short- and long-run rates. In Weitzman, discount rates decline from 4 per cent for the

of returns on leaseholds follows from expectations about the duration of ownership, interacted with expectations on costs over this period of ownership, and expectations on the time of sale of the asset, interacted with the expected value of the asset at the time of sale of the asset and its net present equivalent value. Specifically, for a property with an above 100 years leasehold, we see no mechanism through which the price discount (for the finite leasehold) could measure the time-structure of preference of the property owner over such a horizon. The owner will not live after 100 years, and there is no evidence that the majority of owners expect the children to keep the asset over the full leasehold duration (in which case one could invoke altruism). That is, there is no indication that pricing a 100 years leasehold has any relation to the preferences of the owner concerning the possible lease costs after hundred years.

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

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 above 1 per cent. This still leaves degrees of freedom in choosing the two rates $\beta\delta$ and $\delta$ — we choose $\beta$ and $\delta$ to maintain the utility discount factor implied by the Euler equation for savings at $\gamma = 0.76$ (2.7 per cent annual discount rate).[46] In other words, the economy continues to choose savings $g^* = .25$ consistent with (15) in all experiments but the last.[47]

| annual discount rate | | | Markov Equilibrium | | | | |
|---|---|---|---|---|---|---|---|
| short-term | long-term | $g^*$ | ID | PF | DF | carbon tax | EF |
| 027 | .027 | .25 | 7.12 | 2.06 | .48 | 7.1 | 1 |
| .033 | .01 | .25 | 7.12 | 3.70 | .70 | 18.5 | 2.6 |
| .035 | .005 | .25 | 7.12 | 5.79 | .82 | 33.8 | 4.8 |
| .037 | .001 | .25 | 7.12 | 19.6 | .96 | 133 | 18.8 |
| .001 | .001 | .33 | 9.27 | 19.6 | .96 | 174 | 1 |

Table 2: Decomposition of the carbon price [Euro/tCO2] year 2010. $ID$=immediate damages, $PF$=persistence factor, $DF$=delay factor, Carbon price $= ID \times PF \times DF$. Parameter values in text. $EF$=excess factor, the ratio given in Proposition 4.

For the carbon tax, the last column indicates the excess factor ($EF$) that we have formalized in Proposition 4: it tells the multiple by which the Markov tax exceeds the benchmark tax using the capital returns to evaluate future impacts. The highest equilibrium carbon tax, 133 EUR/t$CO_2$, corresponds to the case where the long-run discounting is as proposed by Stern (2006); this case also best matches Weitzman's values. For reference, we report the Stern case where the long-term discounting at .1 per cent holds throughout; the carbon tax takes value 174 EUR/t$CO_2$, and gross savings cover about 33 per cent of income. Thus, the Markov equilibrium closes considerably the gap between Stern's and Nordhaus' carbon taxes, without having unrealistic by-products for the macroeconomy.[48]

[46]For example, 3 per cent short run and 1 per cent long run annual rates correspond to $\beta = .788$ and $\delta = .904$. See the supplementary material for all numerical values.

[47]Excluding the last row that is explained just below.

[48]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 $y_t$

The decomposition of the carbon tax 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. (2014), doubles the carbon tax level. For hyperbolic discounting, as expected, the persistence of impacts, capturing the commitment value of climate policies, contributes significantly to the deviation between the capital-market based and Markov equilibrium prices.

For the same preferences, we now consider the quantitative significance of coordinated capital taxes and their effect on savings and carbon taxes. Table 3 presents the Markov equilibirum capital tax and the best constant tax on capital that can be sustained in a subgame perfect equilibrium, defined in Proposition 2 as $\widehat{\tau}^k$. The Markov equilibrium capital tax is larger, the greater is the discrepancy between short- and long-run preferences. Arguably, the capital tax levels remain reasonable. The best achievable capital policy involves subsidizing capital at low rate, converging to zero when the long-run time preference involves no discounting. The increase in savings reduces the equilibrium carbon tax, although the quantitative difference to the values in Table 2 is not large.

| annual discount rate | | Markov equilibrium | | | Subgame Perfect equilibrium | | |
|---|---|---|---|---|---|---|---|
| short-term | long-term | capital tax | savings | carbon tax | capital tax | savings | carbon tax |
| 027 | .027 | 0 | 0.25 | 7.1 | 0 | 0.25 | 7.1 |
| .033 | .01 | 16% | 0.25 | 18.5 | $-.7\%$ | 0.29 | 17 |
| .035 | .005 | 23% | 0.25 | 34 | $-.5\%$ | 0.31 | 31 |
| .037 | .001 | 29% | 0.25 | 133 | $-.1\%$ | 0.32 | 120 |
| .001 | .001 | 0 | 0.33 | 174 | 0 | 0.33 | 174 |

Table 3: Capital and carbon taxes [Euro/tCO2] year 2010 for both the Markov equilibirum and the coordination subgame perfect equilibrium.

# 6 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 Brock-Mirman (1972) we assumed that income and substitution effects in consumption

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.

choices over time cancel out, leading to capital and climate policies that are separable. With general functional forms, climate policies can generate income effects influencing future savings, thereby creating interactions between the two policies. Based on a three-period extension to general functional forms, presented in the online Appendix, we discuss below the effects that are ruled out by the assumptions in the main analysis. 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. Finally, the quasi-hyperbolic discount functions are only rough approximations of more general discount functions. Given the parametric class for preferences and technologies, it is possible to solve this model for an arbitrary sequence of discount factors; this extension is provided in Iverson (2012). The flexible discounting does not change the conceptual substance matter in a material way, although the quantitative evaluations can depend on the added flexibility.[49] Yet, currently, there is no evident data for the path of the time-preferences that would call for the flexible formulation. We now briefly discuss how results may be expected to change for other functional forms of utility, production and climate change.

## 6.1 Sensitivity of policies under geometric discounting

Before assessing the changes in strategic interactions when the functional forms are more general, we discuss the sensitivity of policies in a context where the current and future planners do not strategically interact but where policies follow a time-consistent planning. Barrage (2014), in a supplement to Golosov et al. (2014), has numerically assessed the loss of generality implied by logarithmic utility and full capital depreciation. Log utility implies a relatively low preference for consumption smoothing over time: the decision maker becomes more "patient", increasing savings and the initial carbon price level when compared to a case where the utility function has more curvature. As long as we stay in the expected utility framework, this overshooting can be made to vanish by appropriate adjustment of the time discount rate. The one period depreciation assumption tends to decrease the carbon price level and its growth, since it implies a lower growth of the economy than in the case if some capital survives to the next period. But, again,

[49] Iverson et al. (2015, Table 1), show similarly to our Table 2, that the carbon tax increases from the 'Nordhaus value' close to the 'Stern value' when time discounting after the first 20 years moves to the Stern values.

adjustments in the calibration can almost exactly offset the full depreciation, closing the gap between the predictions of the numerical models with partial depreciation and analytical models with full depreciation.

A further study on the sensitivity is presented in van den Bijgaart et al. (2016). They devise a Monte Carlo experiment for testing how well the closed-form carbon price formula, slightly extended from the one that we have developed, predicts the social cost of carbon from a benchmark simulation model, DICE 2007. This benchmark model assumes a more general parametric class for preferences and technologies, and also features non-linearities of the climate system. Assuming geometric discounting and drawing parameters from pre-determined distributions for all key parameters in DICE, including those that appear in our formula as well as those not in our formula, they that the formula explains the DICE prediction without systematic bias. The largest gaps in outcomes are associated with situations where climate damages are either strongly concave or convex, and, at the same time, the discount rate takes extreme values (low or high). The results suggests that the loss of generality from not including the interactions between policies, that capture mainly income and substitution effects in consumption are not central when evaluating the social cost of carbon.[50]

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. (2014) 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. Arguably, the basic question is if "climate change unknowns" undermine the usefulness of the closed-form model outcomes such as the ones presented in the current paper. Gerlagh and Liski (2016) develop a tractable extension of the current paper's setting to allow for a quantitative assessment of the optimal carbon price when the impacts of climate change are unknown and can be learned only gradually over time. They find that the high-risk carbon price path need not be that different from the mainstream policy ramp.

[50]Rezai and van der Ploeg (2015) consider further extensions by allowing for mean reversion in global warming damages, negative effects of global warming on trend growth, and a non-unitary elasticity of damages with respect to aggregate output. They find minimal welfare losses if one applies the simple rule as the basis for the climate policy over time.

## 6.2 Sensitivity of strategic carbon policies

Moving to a general description of preferences, technologies and climate change, opens new opportunities to strategically influence the future policies by current decisions. In a stylized but general three-period model (see the online Appendix), we can show that a higher elasticity of marginal utility leads to laxer climate policy today, as current planners foresee that future planners will tend to compensate a current increase in emissions through changes in savings. Today's climate policies and future savings become strategic substitutes, which tends to lower the equilibrium carbon tax today. In addition, the strategic substitutability of policies depends on the interaction between damages and output. A less-than-proportional increase of damages implies that current emissions have less of an effect on future returns on capital investments, and, thus, current emissions become less of a substitute for future savings. Therefore, we find laxer climate policies when damages are less dependent on output levels.

Yet, these effects are indirect and we have no reason to believe that they are quantitatively substantial.[51] We believe there is more scope for strategically guiding future decision makers by investing in specific capital stocks, including technology. Generally, we expect that if capital and emissions are complementary in production, then planners with quasi-hyperbolic preferences will tend to invest less in capital, as it commits future planners to increased emissions. But, specific types of capital that substitute for emissions, such as investments in clean energy or clean energy R&D, will attract larger investments from a planner who wishes to commit future planners to lower emissions. In spirit of Gul and Pesendorfer (2001), decision-makers may want to expend resources to remove alternatives (such as cheap fossil fuels) from the future choice sets. That is, if possible policy makers will choose options that induce strategic complementarity, as these increase overall welfare, and technology choices provide a means for such policies.[52]

We can also assess how the details of climate change damages are expected to modify results. If marginal damages tend to increase with past emissions, emissions will be strategic substitutes over time, and the current generations can strategically increase their own emissions expecting future generations to reduce theirs in response. The extreme case of such a scenario is one in which there is a known catastrophe threshold. Suppose

[51] Iverson, in a revision of his (2012) manuscript, follows up on our analysis and develops a numerical model to conclude: "Nevertheless, in all cases the quantitative effect [of a different parametric form] is tiny - on the order of one one-thousandth the magnitude of the initial period perturbation." We note though that Iverson abstracts from productivity growth.

[52] Harstad (2015) formalizes some of these ideas in a setting with quasi-hyperbolic discounting.

that climate change is moderate up to levels of cumulative emissions in the range of four thousand Gigaton of $CO_2$, after which a trigger sets in dangerous climate change. Given such known threshold, each generation can freely add their emissions, as long as the threshold is not reached, as on the margin future policies will offset current emissions one-to-one. Similarly, we may consider different greenhouse gases having different lifetimes, and thereby, different commitment value. Long-lived gases, such as $N_2O$, will typically provide larger commitment, as compared to short-lived gases such as methane.

# 7 Concluding remarks

In 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 developed a methodology for addressing the over-arching question posed above and a quantitative evaluation. Our analysis provides one way to incorporate the idea, 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 drive a wedge between the marginal rate of substitution and transformation, stipulating a correction to the carbon tax resulting from the evaluation of future damages based on capital returns.

The resulting tool for policy purposes is a carbon pricing formula (Proposition 1) 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 and its decomposition (31)-(33) identifies the contributions of the key elements to optimal carbon prices and allows discussing them transparently. For discount factors consistent with those in the literature we show that the equilibrium correction to the standard Pigouvian pricing principle is quantitatively significant. However, there is very limited solid empirical evidence on the time-structure of social preferences over long time horizons. Our study shows the relevance of such information; it has a large effect on the evaluation of currently observed energy-use patterns. The carbon price directly impacts the estimate of "genuine savings" that are calculated by, for example, the World Bank. Currently valued at $20\$/tC$, the World Bank estimates the "negative savings" due to $CO_2$ emissions at 0.3 per cent of GDP for the US, and 1.1 per cent for China. Using a $100\$/tC$ carbon price ($21EUR/tCO_2$) derived from a moderate quasi-hyperbolic preference structure, will increase the estimate for the negative savings to above 1 percent for the US and above 5 per cent for China. More generally, the formula allows policy-makers to experiment with their prescriptive views on longer-term discounting to see the effect on the optimal carbon price.

# References


[1] Barrage, L., 2014. Sensitivity Analysis for Golosov, Hassler, Krusell, and Tsyvinski (2013): "Optimal Taxes on Fossil Fuel in General Equilibrium". Supplementary Material, Econometrica 82(1), 41–88.

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

[3] 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.

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

[5] 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.

[6] 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.

[7] 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.

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

[9] Cropper M.L., Freeman M.C., Groom B. and Pizer W. (2014). Declining Discount Rates. American Economic Review: Papers and Proceedings, 104(5): pp. 538-43.

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

[11] 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.

[12] 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/

[13] Gerlagh, R., and M. Liski (2016), Carbon prices for the next hundred years, The Economic Journal, forthcoming.

[14] Giglio, S. M. Matteo, J. Stroebel (2015), Very long-run discount rates, Quarterly Journal of Economics, 130(1), February.

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

[16] Golosov, M., J. Hassler, P. Krusell, A. Tsyvinski (2014), Optimal taxes on fossiel fuel in general equilibrium, Econometrica 82: 41-88.

[17] 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.

[18] Gul F. and W. Pesendorfer (2001), Temptation and self-control, Econometrica 69: 1403-1435.

[19] 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

[20] Harstad, B. (2015). Investment Policy for Time-Inconsistent Discounters, working paper, University of Oslo.

[21] 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

[22] 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

[23] 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.

[24] 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.

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

[26] Iverson, T. S. Denning, S. Zahran (2015), When the long run matters, Climatic Change 192: 57-72.

[27] Jackson, O. M., and L. Yariv (2015), Collective Dynamic Choice: The Necessity of Time Inconsistency, American Economic Journal: Microeconomics 7:4, 150-178.

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

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

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

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

[32] 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.

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

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

[35] Kydland F.E., and E.C. Prescott (1977), Rules rather than discretion: the inconsistency of optimal plans, J of Political Economy 85: 473-492

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

[37] 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.”

[38] 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.

[39] 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.

[40] 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.

[41] Nordhaus, W.D., (1991), To slow or not to slow: the economics of the greenhouse effect. Econonomic Journal, 101 (407), 920?937.

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

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

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

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

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

[47] Rezai, A., van der Ploeg, R., (2015) Intergenerational Inequality Aversion Growth and the Role of Damages: Occam?s Rule for the Global Carbon Tax. Oxcarre Research 150.

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

[49] 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.

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

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

[52] 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.

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

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

[55] van den Bijgaart I., R. Gerlagh, M. Liski (2016), A simple formula for the social costs of carbon, J. of Environm. Econ. and Management, 77:75-94.

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

[57] 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:

$$S_{i,t} = (1-\eta_i)^{t-1} S_{i,1} + \sum_{\tau=1}^{t-1} a_i (1-\eta_i)^{\tau-1} z_{t-\tau} \quad (34)$$

$$D_t = (1-\varepsilon)^{t-1} D_1 + \sum_{i\in\mathcal{I}} \pi\varepsilon \frac{(1-\eta_i)^t - (1-\eta_i)(1-\varepsilon)^{t-1}}{\varepsilon-\eta_i} S_{i,1} + \quad (35)$$
$$\sum_{i\in\mathcal{I}} \sum_{\tau=1}^{t-1} a_i \pi\varepsilon \frac{(1-\eta_i)^\tau - (1-\varepsilon)^\tau}{\varepsilon-\eta_i} z_{t-\tau},$$

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 (6). This defines the emissions-damage function $\theta_\tau$ in Theorem 1. Q.E.D.

## Lemma 5

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 5 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 5 allows us to interpret the policy stringency as measured by $h$ directly as the stringency of the carbon price $\tau$.

**Lemma 5** *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 $dh_t/dz_t = [f_{t,zz}f_t - (f_{t,z})^2]/(1-g)(f_t)^2 < 0$, so that the carbon price measured in units $h_t$ and the carbon price measured in units $\tau_t$ 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:

$$\frac{1}{y_t}\frac{\partial y_t}{\partial e_t}\frac{\partial E_t}{\partial z_t} = h_t(1-g_t), \tag{36}$$

$$\frac{\partial A_t}{\partial l_{y,t}} = \frac{\partial A_t}{\partial e_t}\frac{\partial E_t}{\partial l_{e,t}} \tag{37}$$

Equation (37) 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 (3), labour market clearance (4), and the two first-order conditions (36)-(37), 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

$$k_{\tau+1} = g_\tau y_\tau, \tag{38}$$

$$\frac{\partial y_\tau}{\partial z_\tau} = h_\tau(1-g_\tau)y_\tau, \tag{39}$$

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 (38), $k_{t+2} = g_{t+1}y_{t+1}$. 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 5. 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 $g_\tau$ and $z_\tau$ for $\tau > t$, and that do not depend on the next-period state variables $k_{t+1}$ and $s_{t+1}$, we get the constant part of $V_{t+1}(k_{t+1})$:

$$\widetilde{A}_{t+1} = \ln(1-g_{t+1}) + \delta\xi\ln(g_{t+1}) + (1+\delta\xi)\ln(A_{t+1}) - \delta\zeta_1 z_{t+1} + \delta\widetilde{A}_{t+2}. \tag{40}$$

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 $\xi$,

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

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 $\zeta_\tau$ 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 $\tau$, 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 $\zeta_1$, 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 $\eta_i = \varepsilon$, $\zeta_1$ still has a closed-form solution; this derivation is available on request) Q.E.D.

## Proof of Remark 1

In text.

## Proof of Proposition 1

In text.

## Proof of Lemma 2

Competitive factor markets (18)-(20) and constant returns to scale with respect to these inputs, ensure the value identity $r_t k_t + \tau_t^z z_t + q_t l_t = y_t$. Lump-sum tax transfers, $T_t = \tau^k k_{t+1} + \tau_t^z z_t$, combined with the individual's consumers budget (21) ensure that aggregate budget balance holds, $c_t + k_{t+1} = y_t$. Using the consumer's budget, the properties of the production function, and the assumed savings function $\mathcal{G}^i = g y_t$, we can write consumption, as given by the rule:

$$c_t^i = [1 - \alpha + g\tau^k + (\alpha - g(1+\tau^k))\frac{k_t^i}{k_t}]y_t.$$

Consumer's utility maximization requires that $k_{t+1}^i = \mathcal{G}^i(k_t^i; k_t, s_t)$ is a solution to the consumption choice that maximizes $u_t^i + \beta\delta w_{t+1}^i$ in (23), with budget $c_t^i + (1+\tau^k)k_{t+1}^i = q_t l_t^i + r_t k_t^i + T_t$ holding, giving

$$(1+\tau^k)\frac{\partial u_t^i}{\partial c_t^i} = \beta\delta\frac{\partial w_{t+1}^i}{\partial k_{t+1}^i} \tag{41}$$

The consumer pays a tax on capital investments, so that effective costs of capital relative to consumption is distorted by $\tau^k$. Using the assumed form $W_t^i(k_t^i; k_t, s_t) = a_t + b\ln(k_t) + c\ln(k_t + \varphi k_t^i)$ in (41):

$$\begin{aligned}
\frac{1+\tau^k}{c_t^i} &= \beta\delta\frac{\partial w_{t+1}^i}{\partial k_{t+1}^i} \Rightarrow \\
\frac{1+\tau^k}{[1-\alpha+g\tau^k+(\alpha-g(1+\tau^k))\frac{k_t^i}{k_t}]y_t} &= \beta\delta c\frac{\varphi}{g y_t + \varphi g\frac{k_t^i}{k_t}y_t} \Rightarrow \\
(g + \varphi g\frac{k_t^i}{k_t})(1+\tau^k) &= \beta\delta c\varphi[(1-\alpha+g\tau^k) + (\alpha - g(1+\tau^k))\frac{k_t^i}{k_t}].
\end{aligned}$$

Since the equation must be valid for any $k_t^i$, it results in two conditions. Condition (42) is for the constant term (independent of $k_t^i$), and (43) is for the linear term in $\frac{k_t^i}{k_t}$:

$$g(1+\tau^k) = \beta\delta c\varphi(1-\alpha+g\tau^k) \tag{42}$$

$$g(1+\tau^k) = \beta\delta c(\alpha - g(1+\tau^k)) \tag{43}$$

We now verify (24), $W_t^i = u_t^i + \delta W_{t+1}^i$, by using the consumers choice rule and the assumed functional form:

$$a_t + b\ln(k_t) + c\ln(k_t + \varphi k_t^i) = \ln([1-\alpha+g\tau^k + (\alpha - g(1+\tau^k))\frac{k_t^i}{k_t}]y_t) +$$
$$\delta[a + b\ln(gy_t) + c\ln(gy_t + \varphi g y_t \frac{k_t^i}{k_t})] \Rightarrow$$
$$a_t + (b+c)\ln(k_t) + c\ln(1+\varphi k_t^i/k_t) = \alpha\ln(k_t) + \ln(1 + \frac{\alpha - g(1+\tau^k)}{1-\alpha+g\tau^k}\frac{k_t^i}{k_t}) +$$
$$\alpha\delta(b+c)\ln(k_t) + \delta c\ln(1+\varphi\frac{k_t^i}{k_t}) + ...$$

where we left out constant terms (independent of $k_t^i$ and $k_t$) associated with $a_t$. We find three more conditions:

$$b + c = \alpha + \alpha\delta(b+c) = \frac{\alpha}{1-\alpha\delta} \tag{44}$$

$$c = 1 + \delta c = \frac{1}{1-\delta} \tag{45}$$

$$\varphi = \frac{\alpha - g(1+\tau^k)}{1-\alpha+g\tau^k} \tag{46}$$

Condition (46) is implied by (42)-(43). Thus, there is one redundant condition. We have 4 conditions to determine the four parameters $b$, $c$, $\varphi$, $g$. The parameters $b$, $c$ and $\varphi$ are directly derived above. Substitution of $c$ in (43) gives (25): savings $g$ as dependent on the technology-preference parameters and policy $\tau^k$.

## Proof of Theorem 3

The tax setting game is defined as follows. Each planner $t = 1, 2, 3, ..$ has instruments $(\tau_t^k, \tau_t^z, T_t)$. Markov strategy for planner $t$ consists of state-dependent triple $(\tau_t^k(k_t, s_t), \tau_t^z(k_t, s_t), T_t(k_t, s_t))$. For $\tau > t$, the set of tax rules generates allocation $\{c_\tau, z_\tau, k_\tau, s_\tau\}_{\tau > t}^{\infty}$, and through Theorem 2, the current planner's continuation value $W_{t+1}(k_{t+1}, s_{t+1})$. If all future planners $\tau > t$, use taxes

$$\tau^{k*} = \frac{\delta(1-\alpha)(1-\beta)}{1-\delta(1-\beta)},$$
$$\tau_t^{z*} = h^*(1-g^*)y_t,$$

then, by Lemma 2, future policies are $(g,h)_{\tau \geqslant t} = (g^*,h^*)_{\tau \geqslant t}$. By Theorem 2, the planner's continuation value coincides with the Markov planning continuation value. The best response for the planner at $t$ is to implement $(g,h)_{\tau=t} = (g^*,h^*)_{\tau=t}$, which, by the description of the competitive equilibrium, is obtained by setting $(\tau_t^k,\tau_t^z) = (\tau_t^{k*},\tau_t^{z*})$, and returning the tax receipts to the consumer.

## Proof of Lemma 3

Consider a given policy path $(g_\tau, z_\tau)_{\tau \geq t}$. We look at variations of policies at time $\tau$, 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 $\xi$ and $\zeta$ do not depend on future polices $(g_\tau, z_\tau)$, but term $\widetilde{A}_t$ does. Technically, we need to show that, for some given $\tau > t$, $\widetilde{A}_t$ increases in $g_\tau$ for $g_\tau < \alpha\delta$. In the proof of Theorem 2, consider (40). Term $\widetilde{A}_t$ increases with $\widetilde{A}_\tau$ for some $\tau > t$. Moreover, $\widetilde{A}_\tau$ is strictly concave in $g_\tau$, and maximal when $g_\tau$ maximizes $\ln(1-g_\tau)+\delta\xi\ln(g_\tau)$, that is, for $g_\tau = \frac{\delta\xi}{1+\delta\xi} = \alpha\delta$. We have now shown the "if" part of the Lemma. The "only if" follows from the strict concavity of $\widetilde{A}_\tau$ with respect to $g_\tau$. Q.E.D.

## Proof of Proposition 2

We start by formally defining the subgame-perfect equilibrium. As for the Markov equilibrium, we confine attention to policies defined through a sequence of constants $(g_t,h_t)_{t\geq 1}$ satisfying (13) and (14) but now allow strategies to depend on the history of policies, defined for $t > 1$ as

$$\mathbf{H}_{t-1} = ((g_1,h_1),...,(g_{t-1},h_{t-1})).$$

Let $\mathcal{H}_\infty$ be the set of all histories. Strategy is a function that maps from the history of policies to current actions, $\mathbf{s}(\mathbf{H}_{t-1}) : \mathcal{H}_\infty \to \mathcal{R}_+^2$.

**Definition 5** *A subgame-perfect equilibrium is a sequence of savings and carbon price rules* $(g_t,h_t)_{t\geq 1}$ *satisfying (13) and (14) such that* $\mathbf{s}(\mathbf{H}_{t-1}) = (g_t,h_t)$ *maximizes welfare at each* $t$ *and all* $\mathbf{H}_{t-1}$*, given* $\mathbf{s}(\mathbf{H}_{\tau-1}) = (g_\tau,h_\tau)_{\tau>t}$.

In the Proposition we consider subgame-perfect coordination of savings, that is, policy $\widehat{g}$ differing from Markov policy $g^*$. The carbon price rule remains at the Markov level $h^*$. Formally, for all $t$, the coordination strategy takes the form

$$\mathbf{s}(\mathbf{H}_{t-1}) = \begin{cases} (\widehat{g}, h^*) & \text{if } \mathbf{H}_{t-1} = ((\widehat{g}, h^*), ..., (\widehat{g}, h^*)) \\ (g^*, h^*) & \text{otherwise.} \end{cases}$$

Now we construct $\widehat{g}$ consistent with capital tax $\widehat{\tau}^k$ defined in the Proposition. Consider the constant saving fraction of output, to be followed at each future date, that maximizes welfare at $t$. Such $\widehat{g}_t$ 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

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

$$(\forall \tau > t, g_\tau = g) \Rightarrow$$

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

$$\Rightarrow$$

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

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

$$\Rightarrow$$

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

Since $\widehat{g}$ it is independent of $t$, this same proposal is optimal for any agent at $\tau > t$. The associated capital tax follows from (25) in Lemma 2.

Note that $\widetilde{A}_{t+1}$ is concave in $g$, so that welfare is monotonic in $g$ between Markov $g^*$ and $\widehat{g}$. This implies $\widehat{\tau}^k < 0$, and for $\tau^k = 0$, welfare still exceeds the level reached for the Markov policy $\tau^{k*}$.

The proof that both the policies $\widehat{\tau}^k$ and $\widehat{\tau}^k = 0$ are self-enforcing (subgame perfect) is straightforward. Anticipating that any deviation from $g_t = \widehat{g}$ triggers $(g_\tau = g^*)_{\tau > t}$, it follows from Theorem 2 that any profitable deviation must be the Markov policy, $g_t = g^*$. But as we have shown above, the deviation that leads to the Markov policy leads to a strict loss compared to both alternative capital tax policies. Q.E.D.

## Proof of Proposition 3

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

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

Since the planner sets

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

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

## Proof of Proposition 4

Because of Lemma 5 (ii), the proposition, stated as $\frac{\tau^{z(\beta,\delta)}}{\tau_t^{z(\gamma)}} > 1$, 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, $\beta < 1$:

$$\frac{h^{\beta\delta}}{h_t^{\gamma}} = \frac{(1-\gamma)^2}{(1-\delta)^2} \frac{\beta\delta}{\gamma}$$

The equality follows from substitution of $\eta_i = \varepsilon = 0$ in the equation for the equilibrium carbon price and efficient carbon price. We note that the ratio decreases in $\gamma$. It thus suffices to check the ratio for highest $\gamma : \widehat{\gamma} = \frac{\beta\delta}{1-\delta(1-\beta)(1+\alpha(1-\delta))}$:

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

Q.E.D.

## Proof of Proposition 5

From Proposition 1, $\tau^{z(\beta,\delta)} \to \infty$ as $\delta \to 1$. From Proposition, 4, we see that $\tau^{z(\gamma)}$ remains bounded. Q.E.D.

### Proof of Lemma 4

The proof runs parallel to the proof for Lemma 3. Consider a given policy path $(g_\tau, z_\tau)_{\tau\geq t}$. We look at variations of policies at time $\tau$, 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 $\xi$ and $\zeta$ do not depend on future polices $(g_\tau, z_\tau)$, but term $\widetilde{A}_t$ does. Technically, we need to show that, for some given $\tau > t$, $\widetilde{A}_t$ decreases in $z_\tau$ for $z_\tau > z^\delta_\tau$, where $z^\delta_\tau$ is the emission level that is consistent with the policy variable $h^\delta$ and $z_\tau$ is the emission level consistent with some $h < h^\delta$. In the proof of Theorem 2, consider (40). Term $\widetilde{A}_t$ increases with $\widetilde{A}_\tau$ for some $\tau > t$. Moreover, $\widetilde{A}_\tau$ is strictly concave in $z_\tau$ and maximal when $z_\tau$ 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 $z_\tau$ consistent with $h^\delta$. We have now shown the "if" part of Lemma 4. The "only if" follows from the strict concavity of $\widetilde{A}_\tau$ with respect to $(g_\tau, z_\tau)$. Q.E.D.

### Proof of Proposition 6

Capital tax is a given constant so policy policy $g$ remains unaffected; thus, we can focus on the change in current welfare $w_t$ due to changes in carbon taxes. Also, by Lemma 5, a higher policy $h$ implies a higher carbon price $\tau$. Let $\beta < 1$ so that $\beta\delta < \gamma < \delta$, and let climate change be a slow process such that $\tau^{z(\delta)} > \tau^{z(\beta,\delta)} > \tau^{z(\widehat{\gamma})} > \tau^{z(\gamma^*)}$; see Proposition 4. Imposing the capital-returns based carbon tax will then decrease the future carbon price, taking it further away from $\tau_t^{z(\delta)}$, decreasing current welfare as shown in Lemma 4. The same mechanism applies for $\beta > 1$, when we have $\tau_t^{(\delta)} < \tau_t^{z(\beta,\delta)} < \tau_t^{z(\gamma)}$. Moreover, imposing the capital-returns based 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.

## 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.[53] We calibrate the model parameters $\mathbf{M}$, $b$, $\mu$, 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

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

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 $\mu$. 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 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 the 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.[54] 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. (2014). The GAMS source

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

code for the DICE2007 model provides a precise description of the carbon cycle through a three-reservoir model. We use the linear algebra from Appendix "Calibrating carbon cycle" 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 $\pi$ and $\varepsilon$ 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 a counterfactual scenario with equal emissions, apart from a the first period when we decreased emissions by 1Gt$CO_2$ (Gigaton rather than Teraton used in the text to keep the impulse marginal for the purposes here). Comparison of the damages, relative of output, then defines the response function $\theta_\tau$ 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 values of $\pi$ and $\varepsilon$ that approximate the average response $\theta_\tau$ 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.

# APPENDIX FOR ONLINE PUBLICATION

# Appendix: A three-period extension to general functional forms

## Technologies and preferences

Consider three generations, living in periods $t = 1, 2, 3$. In each period, consumers are represented by an aggregate agent having a concern also for future consumers' utilities and welfare. Generations care about current and future utilities as follows

$$w_1 = u_1(c_1) + \beta[\delta u_2(c_2) + \delta^2 u_3(c_3)] \tag{47}$$

$$w_2 = u_2(c_2) + \beta[\delta u_3(c_3)] \tag{48}$$

$$w_3 = u_3(c_3), \tag{49}$$

where all utility functions $u_t$ are assumed to be continuous and, in addition, strictly concave, differentiable, and satisfying $\lim_{c \to 0} u_t' = \infty$. The condition $\beta < 1$ is equivalent to pure altruism towards future decision makers (Saez-Marti and Weibull 2005):

$$\begin{aligned} w_1 &= u_1(c_1) + a_2 w_2 + a_3 w_3 \\ a_2 &= \beta\delta > 0, a_3 = \beta(1-\beta)\delta^2 > 0, \end{aligned} \tag{50}$$

where $a_2, a_3$ can be interpreted as welfare weights given by the first generation, implied by increasing patience over time. When $\beta = 1$, there is one-period pure altruism, and the typical recursive-dynastic representation of welfare follows.

In the first period, the consumption possibilities are determined by a strictly concave neoclassical production function $f_1(k_1, z)$, where $k_1$ is the capital stock, and $z$ is the use of fossil fuels, or emissions of carbon dioxide, both having positive marginal products, $\frac{\partial f_1}{\partial k} = f_{1,k}, \frac{\partial f_1}{\partial z} = f_{1,z} > 0$. The first generation starts with a capital stock $k_1$, and produces output using $z$, which can be used to consume $c_1$, or to invest in capital for the immediate next period $k_2$:

$$c_1 + k_2 = f_1(k_1, z). \tag{51}$$

We abstract from fossil-fuel use in the second and third period, but the first-period fossil-fuel use impacts production negatively in the third period: this captures the delay of climate-change impacts. The second agent starts with the capital stock $k_2$, produces output using a strictly concave neoclassical production function $f_2(k_2)$, and can use its income to consume $c_2$, or to invest in capital for the third period $k_3$:

$$c_2 + k_3 = f_2(k_2). \tag{52}$$

The third consumer derives utility from its consumption, which equals production. Past emissions now enter negatively, as damages, in the production function, $f_{3,k} > 0, f_{3,z} < 0$:

$$c_3 = f_3(k_3, z). \tag{53}$$

We assume that also this production function is strictly concave.

An allocation $(\mathbf{c}, \mathbf{k}, z) = (c_1, c_2, c_3, k_2, k_3, z) \in A \subseteq R_+^6$ (convex set) constitutes a consumption level for each generation $c_t$, the first-period use of fossil fuels $z$, which we thus also consider a proxy for the emissions of carbon dioxide emissions, and capital stocks $k_2$ and $k_3$ left for future agents ($k_1$ is given).

## Equilibrium carbon price

In the subgame-perfect equilibrium generations choose consumptions and emissions in the order of their appearance in the time line, given the preference structure (47)-(49) and choice sets defined through (51)-(53).

The third agent consumes all capital received and cannot influence past emissions. The second agent decides on the capital $k_3$ transferred to the third agent, given the capital inherited $k_2$ and the emissions $z$ chosen by the first agent. We thus have a policy function $k_3 = g(k_2, z)$, defined by

$$\max_{k_3} u_2(c_2) + \beta\delta u_3(f_3(k_3)), \tag{54}$$

leading to equilibrium condition

$$u_2' = \beta\delta u_3' f_{3,k} \Rightarrow 1 = \frac{R_{2,3}}{MRS_{2,3}^{t=2}}, \tag{55}$$

where we introduce the notation $R_{i,j}$ for the rate of return on capital from period $i$ to $j$, and $MRS_{i,j}^t$ for the absolute value of the marginal rate of substitution between consumptions in periods $i$ and $j$ for generation $t$.

The strict concavity of utility implies consumption smoothing, and thus if the second agent inherits marginally more capital $k_2$, the resulting increase in output is not saved fully but rather split between the second and third generation:

**Lemma 6** *Policy function $g$ satisfies* $0 < g_k < R_{1,2}$.

**Proof.** Substitute the policy function $k_3 = g(k_2, z)$ in (55),

$$\beta\delta u_3'(f_3(g(k_2, z), z))f_{3,k}(g(k_2, z), z) = u_2'(f_2(k_2) - g(k_2, z)). \tag{56}$$

Full derivatives with respect to $k_2$ lead to

$$\beta\delta g_k(u_3'' f_{3,k} f_{3,k} + u_3' f_{3,kk}) = u_2''(f_2' - g_k)$$

$$\Rightarrow g_k = \frac{f_2' u_2''}{\beta\delta u_3'' f_{3,k} f_{3,k} + \beta\delta u_3' f_{3,kk} + u_2''} < f_2' = R_{1,2}. \quad (57)$$

as $u_t'', f_{3,kk} < 0$ and $f_{3,k}, u_3' > 0$. ■

Understanding the second agent's policy, the first agent decides on consumption and fossil-fuel use to maximize its welfare

$$w_1 = u_1 + \beta\delta[u_2(f_2(k_2) - g(k_2, z)) + \delta u_3(f_3(g(k_2, z), z)].$$

The choice for leaving capital $k_2$ satisfies

$$\begin{aligned} u_1' &= \beta\delta(f_{2,k} - g_k)u_2' + \beta\delta^2 f_{3,k} g_k u_3' \\ &\Rightarrow MRS_{1,2}^{t=1} = R_{1,2} + (\frac{1}{\beta} - 1)g_k. \end{aligned} \quad (58)$$

where we use (55). When $\beta = 1$, preferences are consistent, and the term in brackets vanishes as in standard envelope arguments for single decision makers; capital $k$ is then valued according to the usual consumption-based asset pricing equation $MRS_{1,2}^{t=1} = R_{1,2}$. For $\beta < 1$, the second agent has a steeper indifference curve between consumptions in periods 2 and 3: the first-order effect in the bracketed term remains positive, leading to capital returns that no longer reflect the first generation's consumption trade-offs. Letting $MRS_{1,3}^{t=1} = MRS_{1,2}^{t=1} \times MRS_{2,3}^{t=1}$, we have

**Lemma 7** *The compound capital return satisfies* $MRS_{1,3}^{t=1} < R_{1,3}$ *if and only if* $\beta < 1$ *.*

**Proof.** Using (58), $MRS_{2,3}^{t=1} = \beta MRS_{2,3}^{t=2} = \beta R_{2,3}$, and Lemma 6:

$$\begin{aligned} MRS_{1,3}^{t=1} &= ... = \left[R_{1,2} + (\frac{1}{\beta} - 1)g_k\right] \times MRS_{2,3}^{t=2} \\ &\Rightarrow MRS_{1,3}^{t=1} = \left[R_{1,2} + (\frac{1}{\beta} - 1)g_k\right] \beta R_{2,3} \\ &< \left[R_{1,2} + (\frac{1}{\beta} - 1)R_{1,2}\right] \beta R_{2,3} = R_{1,3}, \end{aligned}$$

where the inequality holds iff $\beta < 1$. ■

Capital returns are generally excessive from the first agent's point of view when $\beta < 1$, that is, the result holds without any restrictions on how emissions alter savings.

But, for the implications of the excessive capital returns on carbon pricing, we must make assumptions on the effect of first-period emissions on the second-period policy, $g_z$. Taking the full derivatives of (56) with respect to $z$, we get

$$g_z = -\frac{\beta(u_3'' f_{3k} f_{3,z} + u_3' f_{3,kz})}{u_2'' + \beta u_3'' f_{3,k} f_{3,k} + \beta u_3' f_{3,kk}}. \tag{59}$$

Assuming $f_{3,kz} \leqslant 0$, all terms in the denominator are negative, so that with the overall negative sign in front, the signs of the numerator's elements inform us about the mechanisms in play. The first term in the numerator captures the income effect of emissions and is positive. If the first generation emits more, the third generation has lower utility levels and the second generation will tend to save more, as the marginal utility of the third generation increases. The second term in the numerator captures the productivity effect and is negative. If the first generation emits more, productivity of capital in the third period will fall, and the return to investments in the second period will fall alongside. The relative strength of both mechanisms depends on the elasticity of marginal utility versus the elasticity of marginal damages:

$$g_z > 0 \text{ iff } EMU > EMD \tag{60}$$

where $EMU = -c_3 u_3''/u_3'$ is the elasticity of marginal utility, and $EMD = f_3 f_{3,kz}/f_{3k} f_{3z}$ is the elasticity of marginal damages, and we use $c_3 = f_3$. If utility is more concave, then the left-hand side of the last inequality will increase, and the second generation will tend to save more with higher past emissions. If marginal damages increase more than proportionally with income, the right-hand side will increase and the second generation will tend to save less with higher emissions.

Assuming log utility, and that the production damage is multiplicative:

$$u_t(c_t) = \ln(c_t) \tag{61}$$

$$f_3(k_3, z) = f_3(k_3)\omega(z), \tag{62}$$

where $\omega(z)$ is a strictly decreasing damage function, sets both sides of the inequality to unity, and implies that the direct effect of emissions on savings vanishes, $g_z = 0$, as can be easily verified from (59).

**Lemma 8** *The second generation does not adjust its savings to past emissions if utility is logarithmic and damages are proportional to output. More elastic marginal utility (or damages that increase less than proportionally with output) imply that second generation's savings increase with past emissions.*

Consider then the first generation's equilibrium carbon policy $z$:

$$u_1' f_{1,z} = \beta\delta g_z u_2' - \beta\delta^2(f_{3,k}g_z + f_{3,z})u_3'. \tag{63}$$

which after substitution of (55) can be rewritten as

$$u_1' f_{1,z} = (1-(1-\beta)g_z\frac{f_{3,k}}{-f_{3,z}})\beta\delta^2(-f_{3,z})u_3' \tag{64}$$

or

$$MCP = (1-(1-\beta)g_z\frac{f_{3,k}}{-f_{3,z}})\frac{MCD}{MRS_{1,3}^{t=1}} \tag{65}$$

where we let $MCP = f_{1,z}$ denote the marginal carbon product, and $MCD = -f_{3,z}$ denote the marginal carbon damages. If $\beta = 1$, then capital returns reflect consumption trade-offs, $MRS_{1,3}^{t=1} = R_{1,3}$, so that from (65) the carbon price becomes just equal to the damage, discounted with capital return:

$$MCP = \frac{MCD}{R_{1,3}}. \tag{66}$$

This is the general-equilibrium Pigouvian carbon price, under consistent preferences $\beta = 1$. If we impose (61)-(62) and thus $g_z = 0$, the first term in the carbon policy implied by (63) is unity. Yet, if $\beta \neq 1$, in equilibrium, while (65) continues to hold as an internal cost-benefit rule for $t = 1$, Lemma 7 implies that the discounted damage no longer equals the carbon price (if $g_z = 0$):

$$MCP > \frac{MCD}{R_{1,3}} \text{ if and only if } \beta < 1. \tag{67}$$

In equilibrium, the first agent establishes a higher carbon price, compared to the Pigouvian level, if and only if $\beta < 1$, i.e., when the first agent gives a higher weight to the long-term utility than the second agent. The result has a very simple intuition. The first consumer would like to transfer more wealth to the third consumer, compared with the preferred wealth transfer of the second consumer: the high capital returns reflect this distortion (Lemma 7). The higher capital returns depress the present-value damages below the true valuation by the first consumer. The opposite deviation — carbon price below the Pigouvian price — occurs if $\beta > 1$.

**Proposition 7** *Assume (61)-(62). If $g_z = 0$ but $\beta \neq 1$, the first-period carbon price does not satisfy the Pigouvian pricing rule, i.e., $MCP \neq \frac{MCD}{R_{1,3}}$. The carbon price exceeds the Pigouvian level if and only if $\beta < 1$. Furthermore, for $g_z \neq 0$ and $\beta < 1$, we find that a larger elasticity of marginal utility with respect to consumption tends to lower* *carbon prices while a larger elasticity of marginal damages with respect to income tends to increase carbon prices.*

***Proof.*** *Above.* ■