Grantham Research Institute on Climate Change and the Environment

Centre for Climate Change Economics and Policy

Cumulative carbon emissions and economic policy: in search of general principles

Simon Dietz and Frank Venmans

March 2018

Centre for Climate Change Economics and Policy Working Paper No. 317
ISSN 2515-5709 (Online)

Grantham Research Institute on Climate Change and the Environment
Working Paper No. 283
ISSN 2515-5717 (Online)

E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL

THE GRANTHAM FOUNDATION FOR THE PROTECTION OF THE ENVIRONMENT

UNIVERSITY OF LEEDS

LSE THE LONDON SCHOOL OF ECONOMICS AND POLITICAL SCIENCE

The Centre for Climate Change Economics and Policy (CCCEP) was established by the University of Leeds and the London School of Economics and Political Science in 2008 to advance public and private action on climate change through innovative, rigorous research. The Centre is funded by the UK Economic and Social Research Council. Its second phase started in 2013 and there are five integrated research themes:

1. Understanding green growth and climate-compatible development
2. Advancing climate finance and investment
3. Evaluating the performance of climate policies
4. Managing climate risks and uncertainties and strengthening climate services
5. Enabling rapid transitions in mitigation and adaptation

More information about the Centre for Climate Change Economics and Policy can be found at: www.cccep.ac.uk.

The Grantham Research Institute on Climate Change and the Environment was established by the London School of Economics and Political Science in 2008 to bring together international expertise on economics, finance, geography, the environment, international development and political economy to create a world-leading centre for policy-relevant research and training. The Institute is funded by the Grantham Foundation for the Protection of the Environment and the Global Green Growth Institute. It has six research themes:

1. Sustainable development
2. Finance, investment and insurance
3. Changing behaviours
4. Growth and innovation
5. Policy design and evaluation
6. Governance and legislation

More information about the Grantham Research Institute on Climate Change and the Environment can be found at: www.lse.ac.uk/grantham.

Cumulative carbon emissions and economic policy: in search of general principles

Simon Dietz[1,2] and Frank Venmans[3,1]

March 2018

Abstract

Recent advances in climate science make possible a physically consistent, yet surprisingly simple model of efficient climate policy. Optimal peak warming depends sensitively on several uncertain parameters. However, the optimal transition to peak warming is slow, because of the stock-flow nature of atmospheric CO_2. Consequently optimal warming in 2100 is much better constrained. A physical property of the climate system hitherto absent from economic models, the saturation of carbon sinks, underpins several important results that contradict previous work: (i) the optimal carbon price grows faster than output; (ii) the cost-effective carbon price follows the simple Hotelling rule after all.

Keywords: carbon price, climate change, cumulative emissions, peak warming, social cost of carbon

JEL codes: Q54

[1] ESRC Centre for Climate Change Economics and Policy, and Grantham Research Institute on Climate Change and the Environment, London School of Economics and Political Science.

[2] Department of Geography and Environment, London School of Economics and Political Science.

[3] Department of Micro-economics, Waroqué School of Economics and Management, University of Mons, Belgium.

Email for correspondence: s.dietz@lse.ac.uk. This paper has benefited from comments by Inge van den Bijgaart, Reyer Gerlagh, Andre Grimaud, Derek Lemoine, Rick van der Ploeg and Ivan Rudik, as well as seminar participants at EAERE 2017, FAERE 2017, FEEM Milan, LSE and Oxford. We would like to acknowledge the financial support of the ESRC, the Grantham Foundation and FNRS.

1 Introduction

In the last decade, climate science has delivered two important and related insights. First, global warming is approximately *linearly* proportional to cumulative emissions of carbon dioxide. Second, the temperature response to an emission of CO_2 is approximately instantaneous and then constant as a function of time. As Ricke and Caldeira (2014) write, "it is a widely held misconception that the main effects of a CO_2 emission will not be felt for several decades" (p1).

For the purposes of climate economics, these insights have two implications in turn. The first is that the climate system can be represented in a much simpler way than before, while staying true to the physics. We can simply write warming as a linear function of cumulative CO_2 emissions, with at most a very small delay. This is what we do in this paper, and we show how such a parsimonious representation of the climate enables a more general representation of the economy than some important previous work, and a wider variety of analytical results than previous work with similarly general foundations. But simplicitly is not the only implication. We also show that important economic models of climate change do not exhibit physically consistent behaviour. In particular, they ignore or give insufficient treatment to the saturation of carbon sinks, which leads them to give incorrect predictions about the dynamics of the optimal carbon price in both welfare-maximising and cost-minimising policy problems.

Our paper belongs to the literature assessing economically efficient pathways for global CO_2 emissions. For most of its history, this literature has rested on the use of numerical simulation models, including Integrated Assessment Models (IAMs) used to calculate welfare-maximising emissions and associated prices (e.g. the DICE, FUND and PAGE models), and energy systems models used to estimate the cost of meeting pre-determined climate goals like limiting global warming to 2°C (see Clarke et al., 2014).[1] What all of these models have in common is that, due to the complexity of the relationships they include (e.g. within the climate system), they are something of a 'black box'. This is discomfiting, especially given the policy traction these models have achieved, and has stimulated some economists to recently develop so-called 'analytical IAMs',[2] which yield closed-form solutions amenable to interpretation, for example simple rules for the optimal carbon price. Our model belongs in this class. Analytical IAMs are close in spirit to the theory of optimal stock pollution (see Xepapadeas, 2005). What distinguishes them is a more – but not too – detailed representation of the physics and economics of climate change. Generally, the

[1]IAMs, so defined, can also perform this latter task, though they typically have a much less detailed representation of the energy system.

[2]This term can be attributed to Traeger (2015).

balancing act of retaining physical and economic detail, while still obtaining (useful) closed-form analytical solutions, is hard to get right.

The original analytical IAM was developed by Golosov et al. (2014). They showed that a very simple closed-form solution for the optimal carbon price can be obtained from a quite general economic model, as long as a few assumptions are made. Arguably the most special assumptions they made were: (i) logarithmic utility; (ii) climate damages that are an exponential function of the current concentration of CO_2 in the atmosphere; and (iii) a constant savings rate.[3] The optimal carbon price in their model is very simple, because it only depends, as a proportion of output, on the utility discount rate, the damage intensity of atmospheric carbon and how fast atmospheric carbon depreciates. It does not depend on future output.

Subsequent contributions to the literature that are closely related to our paper include Rezai and van der Ploeg (2016) and van den Bijgaart et al. (2016). Both studies obtain closed-form solutions for the marginal damage cost of CO_2 emissions, a.k.a. the social cost of carbon. They make more general assumptions than Golosov et al. about the utility function, the damage function and the climate system's behaviour. Other key contributions that have a somewhat different focus to ours include Gerlagh and Liski (2017) and Traeger (2015).

Lemoine and Rudik (2017) seek to answer a different question; they find an analytical solution for the policy that minimises the discounted abatement cost of meeting a pre-determined climate goal, so they carry out 'cost-effectiveness' rather than cost-benefit analysis. They challenge the conventional wisdom that the cost-effective carbon price path follows an augmented Hotelling rule, increasing at the interest rate plus the depreciation rate of atmospheric CO_2. Instead, they argue that thermal inertia in the climate system allows abatement to be postponed, so that the cost-effective carbon price starts lower and initially grows more slowly than under the augmented Hotelling rule, before growing very fast to a very high level at the 'last minute' in order to respect the temperature constraint.

1.1 Our contribution

Our approach, based on cumulative carbon emissions, is particularly useful for evaluating optimal peak warming of the planet, and the circumstances in which the 1.5-2°C target range for peak warming that has been adopted in the Paris Agreement can be given support in a globally aggregated, welfarist framework. We show (Proposition 1 and Corollary 1) that optimal peak warming depends on: the utility discount rate; the elasticity of marginal

[3] The other special assumptions they made, which are arguably more representative of the literature, were (iv) damages proportional to output (so-called 'multiplicative'), and (v) a constant atmospheric CO_2 decay rate.

utility; population growth; output growth per capita; the marginal cost of abatement at zero emissions; the transient climate response to cumulative carbon emissions; and the damage function coefficient. Moreover optimal peak warming has a unit elasticity with respect to the last three of these parameters, and an elasticity of around one or more with respect to most of the others. Large uncertainty about some of these parameters therefore means there is large uncertainty about optimal peak warming. We suggest that if each parameter is calibrated on the breadth of relevant evidence and opinion, optimal peak warming is 3.4°C. However, we are also able to identify a wide range of circumstances in which peak warming of 2°C or less is optimal. We further show that the relatively short adjustment timescale of temperature to cumulative emissions can be ignored in calculating optimal peak warming and all that follows.

Our model is also simple enough to enable the characterisation of the optimal transition path to peak warming in closed-form. A key insight of this exercise is that the optimal transition is slow: it is optimal to put in significant effort early on, in order to slow the rate of increase of cumulative CO_2 emissions. Consequently the uncertainty about optimal transient warming a century from now is much lower than the uncertainty about optimal peak warming. We show that this is fundamentally due to the stock-flow nature of CO_2-induced warming, in the context of the structural assumptions made in our model about damages and abatement costs. Climate scientists have for some years been arguing that transient warming is a more policy-relevant variable than equilibrium warming (e.g. Allen et al., 2009) and our results therefore give this view an economic grounding.

We obtain a closed-form solution for the optimal carbon price (Proposition 2). It shows that the optimal carbon price does not just increase at the growth rate of the economy (Golosov et al., 2014), rather it increases faster. The fundamental reason why is the saturation of carbon sinks (Corollary 2), which is a physically realistic feature ignored or given insufficient treatment by economic models. The saturation of carbon sinks leads to a constant marginal effect of cumulative emissions on warming (except for a very short initial delay). Hence warming is proportional to cumulative emissions. We show that when damages are a convex function of warming, as is normally assumed, the optimal carbon price increases faster than aggregate output. Quantitatively, this effect adds around 0.5 percentage points to the initial growth rate of the optimal carbon price under central parameter values, falling to about zero in 100 years.

Having characterised what we might call the *unconstrained* optimal path, we consider the effect of a policy constraint to reflect the temperature limits set out in the Paris Agreement, namely "Holding the increase in the global average temperature to well below 2°C above pre-industrial levels and pursuing efforts to limit the temperature increase to 1.5°C above pre-industrial levels". In our model this can be represented by an inequality constraint – an upper limit – on cumulative CO_2 emissions. We show (Proposition 3) that the optimal carbon price under a binding temperature constraint comprises the social cost of carbon, plus a Hotelling premium to ensure inter-temporally efficient use of the cumulative emissions budget implied by the 1.5-2°C limit. Many studies have sought to derive optimal emissions and carbon prices under such a temperature constraint (see Lemoine and Rudik, 2017; Clarke et al., 2014, review numerical energy models). What distinguishes our approach is that the planner does not just minimise discounted abatement costs, rather the planner still values damages by minimising the discounted sum of abatement *and* damage costs.

We finish up by showing what difference this makes, by running the model ignoring damages. The optimal price path to minimise abatement costs just follows the simple Hotelling rule (Proposition 4), not the augmented Hotelling rule, and not with a large adjustment for thermal inertia either (Lemoine and Rudik, 2017). This result again comes from taking into account the feedback from the saturation of carbon sinks to the decay of atmospheric CO_2, as well as from not over-estimating thermal inertia. When we compare the cost-effective price path with the price path that maximises net benefits, we show that ignoring damages leads the planner to delay emissions cuts (Proposition 5). This effect is large: initial emissions are 31% lower when damages are included in its determination, under central parameter values.

1.2 Structure of the paper

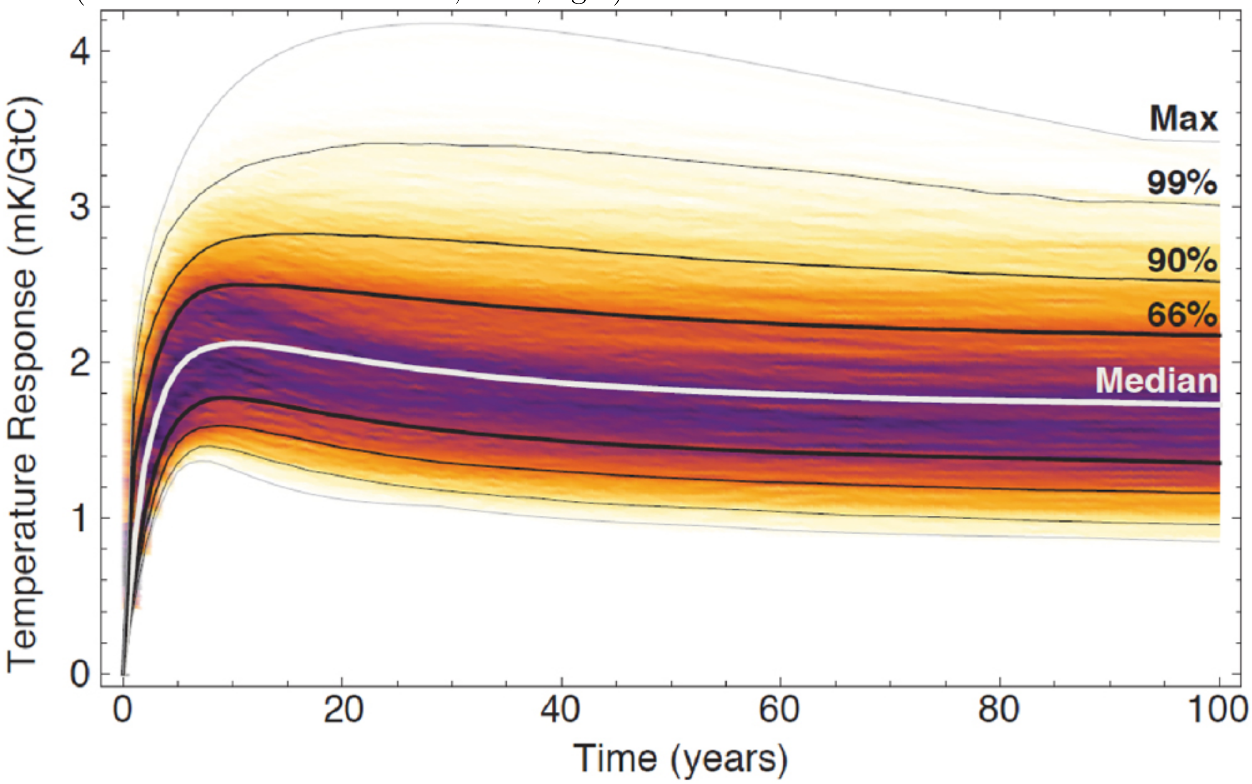
The rest of the paper is structured as follows. Section 2 lays out the building blocks of the model and provides a detailed justification of them, starting with the science alluded to above. Section 3 studies optimal emissions in the model, focusing on peak warming, the speed of transition to peak warming, and carbon prices. Section 4 introduces the constraint on warming made salient by the Paris Agreement. Section 5 concludes.

2 Elements of the model

2.1 A linear model of warming

Our climate model is based on two important physical properties. First, as mentioned above, the temperature response to a pulse emission of CO_2 is approximately constant as a function of time, except for an initial period of adjustment that is very short, i.e. five to ten years (Matthews and Caldeira, 2008; Shine et al., 2005; Solomon et al., 2009; Eby et al., 2009; Held et al., 2010; Joos et al., 2013; Ricke and Caldeira, 2014). Figure 1 reproduces a well-known example of this result, obtained by statistical fitting of the output of 15 carbon-cycle models

Figure 1: Temperature response to an instantaneous 100GtC pulse of CO_2, as a function of time (source: Ricke and Caldeira, 2014, fig 1)



and 20 atmosphere-ocean general circulation models (therefore the figure is representative). Second, the warming effect of an emission of CO_2 does not depend on the background concentration of CO_2 in the atmosphere (Matthews et al., 2009). As we now show, because the temperature response to CO_2 emissions is both time- and concentration-independent, warming is approximately linearly proportional to cumulative CO_2 emissions

The two stages of (i) CO_2 emissions raising the atmospheric CO_2 concentration and (ii) elevated atmospheric CO_2 causing global temperatures to rise can be collapsed into a single parametric relationship between cumulative emissions and warming. This has been defined by IPCC as the Transient Climate Response to Cumulative Carbon Emissions (TCRE: Collins et al., 2013). Formally, the TCRE ζ is (Matthews et al., 2009)

$$\zeta \equiv \frac{\Delta T}{\Delta S} = \frac{\Delta T}{\Delta M} \cdot \frac{\Delta M}{\Delta S}. \tag{1}$$

The TCRE is the product of temperature change per unit increase of atmospheric carbon, $\Delta T/\Delta M$, and the increase in atmospheric carbon per unit of cumulative emissions, $\Delta M/\Delta S$. This enables us to understand why warming from a pulse of emissions is constant over time, as in Figure 1 (also see Goodwin et al., 2015). $\Delta T/\Delta M$ is a concave increasing function of time, because of thermal inertia, i.e. it takes time before an energy imbalance will lead to a new equilibrium temperature, particularly given the large heat capacity of the oceans. Conversely $\Delta M/\Delta S$ is a convex decreasing function of time, because carbon is gradually absorbed by the biosphere and oceans. Except for the first five to ten years after a pulse of CO_2, the rate of increase of $\Delta T/\Delta M$ is cancelled out by the rate of decrease of $\Delta M/\Delta S$. It is perhaps worth belabouring this point: the fact that the TCRE is the product of $\Delta T/\Delta M$ and $\Delta M/\Delta S$ explains why the initial pulse-adjustment timescale of the climate system to an emission of CO_2 is so short, despite the fact that equilibrium $\Delta T/\Delta M$ is notoriously slow to reach.[4]

The TCRE is also independent of past emissions and in turn the atmospheric CO_2 concentration. As atmospheric CO_2 increases, it is well known that $\Delta T/\Delta M$ decreases, due to the saturation of wavelengths absorbed by CO_2. However, again this is cancelled out by an increase in $\Delta M/\Delta S$, due to saturation of carbon sinks. To the extent that the biosphere and oceans have already absorbed a lot of CO_2 in the past, their capacity to absorb further emissions decreases.

Together the time- and concentration-independence of the TCRE mean we can interpret it as a time-invariant parameter ζ. Global warming is approximately linearly proportional to cumulative CO_2 emissions (Matthews et al., 2009; Allen et al., 2009; Zickfeld et al., 2009, 2013; Gillett et al., 2013; Collins et al., 2013). Figure 2 reproduces an important chart from the *Fifth Assessment Report* of the Intergovernmental Panel on Climate Change (IPCC), which depicts evidence on this both from the observational record and projections by various kinds of physical climate model.

The resulting quasi-linearity between cumulative emissions and warming allows us to obtain an extremely simple, yet physically consistent, climate model. The global mean temperature at a point in time responds to cumulative emissions up until that point in time:

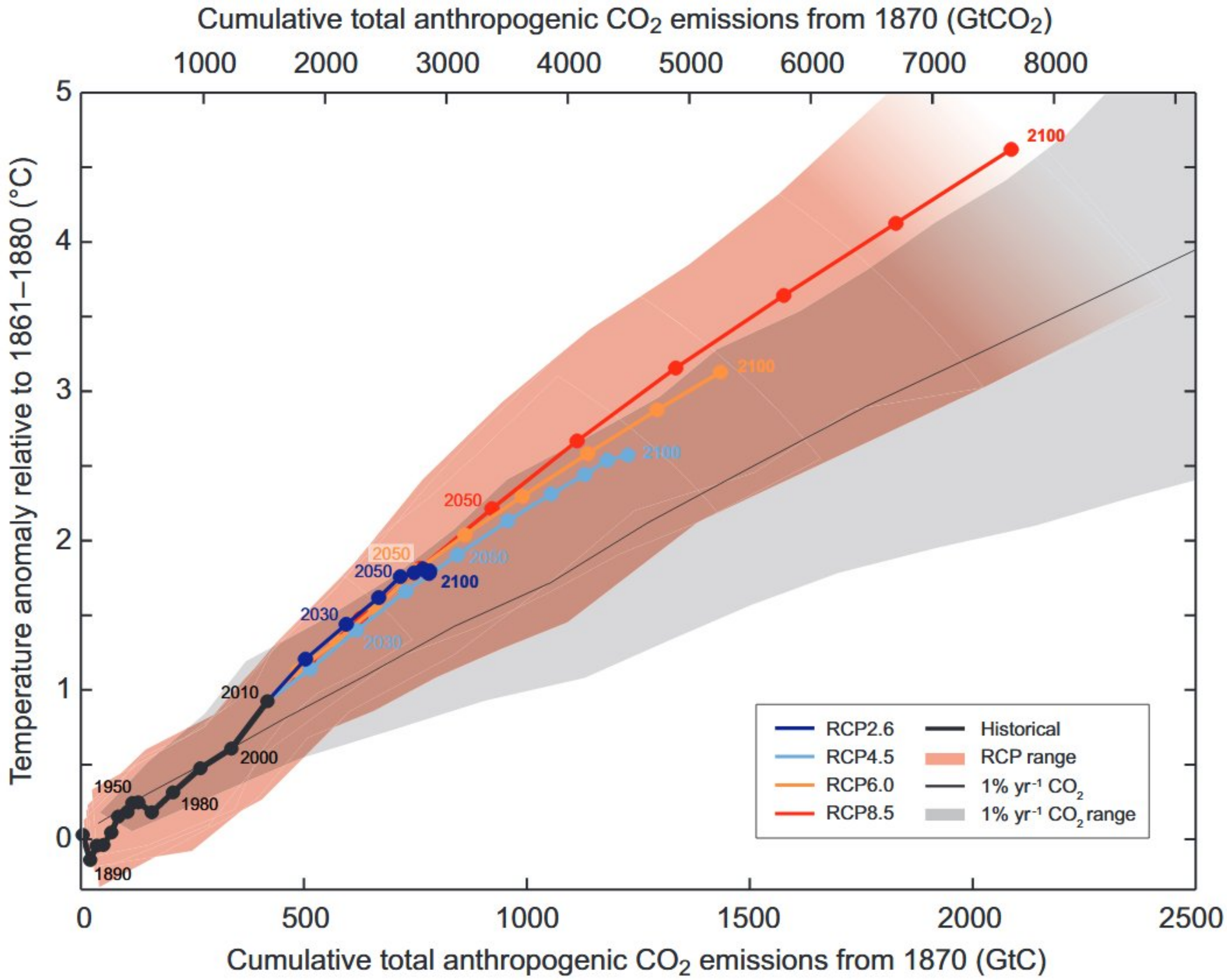
$$\dot{T} = \varepsilon \left(\zeta S - T \right), \tag{2}$$

where T is warming since pre-industrial times and ε parameterises the 'initial pulse-adjustment timescale' of the climate system (Allen, 2016), which is about 10 years as we have mentioned.[5]

[4]Indeed, Caldeira and Myhrvold (2013) show using the same models and data as Ricke and Caldeira (2014) that approximately one quarter of equilibrium $\Delta T/\Delta M$ occurs after more than one century.

[5]Strictly speaking, the physical basis of (2) only applies to CO_2, as opposed to other greenhouse gases (methane, for example, is much shorter-lived in the atmosphere than CO_2). However, it has been proposed that other greenhouse gases can be accommodated in models like this by simply assuming total anthropogenic warming remains a fixed fraction of warming induced by CO_2 alone, e.g. 10% higher (Allen, 2016). Figure 2 indicates this is a good characterisation of the past 150 years or so, while there is no clear case for assuming the ratio of CO_2-induced warming to total anthropogenic warming will be higher or lower in the future; it could go either way. Accordingly, in our numerical modelling we multiply ζ by a factor of 1.1 to account for

Figure 2: Transient warming as a function of cumulative global CO_2 emissions (source: Figure SPM.10 in IPCC, 2013). The coloured lines represent the mean of multiple physical models run under each of the IPCC's four scenarios for the atmospheric stock of greenhouse gases, i.e. the Representative Concentration Pathways (RCPs). The area shaded in colour represents 90 per cent of the spread between models.



S is cumulative emissions of CO_2, so

$$\dot{S} = E, \tag{3}$$

where E is the instantaneous flow of emissions.

The climate science set out here has significant implications for how IAMs and analytical IAMs are parameterised. In comparison with Figure 1, some IAMs like William Nordhaus' DICE assume significant thermal inertia in the climate system. Appendix A shows that in DICE it takes 40 years for the temperature response to a pulse of CO_2 to peak, not ten. Calel and Stainforth (2017) suggest that this is due to the effective heat capacity of the DICE climate being more than twice as high as a representative central estimate from physical climate models. Therefore analytical IAMs calibrated on the DICE climate (e.g. Lemoine and Rudik, 2017) will also warm up too slowly in response to emissions.

In addition, IAMs and analytical IAMs typically do not include the feedback created by saturating carbon sinks.[6] Millar et al. (2017) have shown that, without saturation of carbon sinks, the simple climate model used in the IPCC *Fifth Assessment Report* overestimates the historical accumulation of CO_2 in the atmosphere, underestimates atmospheric CO_2 at the end of this century, compared with the projections of physical climate models under a range of emissions scenarios, and is lastly unable to reproduce the linear response of warming to cumulative CO_2 emissions set out in Figure 2.[7]

2.2 Exponential-quadratic damages

We work with an aggregate damage function mapping the increase in global mean temperature since pre-industrial to a loss of output (first proposed by Nordhaus, 1991). Warming serves as an index of a set of climatic changes including, but not limited to, temperature change, while the accompanying loss of output is equivalent to a welfare loss. The appropriate form of the damage function is notoriously uncertain; there is little empirical evidence that is directly relevant (Pindyck, 2013) and it has been argued that the welfare cost of significant warming, of the order of 5°C or more, is underestimated (Stern, 2013; Weitzman, 2009, 2012). With this caveat front and centre, the existing data points on which the damage function might be fitted have been collected by Nordhaus and Moffat (2017). A quadratic

non-CO_2 greenhouse gases.

[6]Regarding the DICE model, which is often the benchmark for the climate system in analytical IAMs, as the ocean absorbs CO_2, it evolves towards a new equilibrium $\frac{[HCO_3^{2-}][H^+]}{[CO_2aq]} = 4,47 * 10^{-7}$mol/l. DICE has a feedback capturing the increase in $\left[HCO_3^{2-}\right]$, but it does not take into account the increase in $[H^+]$, i.e. acidification, a feedback that is of much greater importance.

[7]Underestimating the saturation of carbon sinks also appears to be why DICE is unable to replicate the approximately flat path of temperature in Figure 1, instead wrongly suggesting that, after peaking, the temperature response to a CO_2 emission recedes markedly (also see Appendix A).

damage function best fits these data and so we specify damages as

$$D(T) = \exp\left(-\frac{\gamma}{2}T^2\right), \tag{4}$$

where γ is the damage function coefficient.

2.3 Marginal abatement costs proportional to abatement and output

We capture the relationship between production and emissions by thinking of E as an input, the abatement of which reduces output, all else being equal (Brock, 1973). We assume that the marginal productivity of emissions is linear decreasing in emissions, when expressed as a proportion of GDP:

$$\frac{Q_E}{Q} = \phi - \varphi E. \tag{5}$$

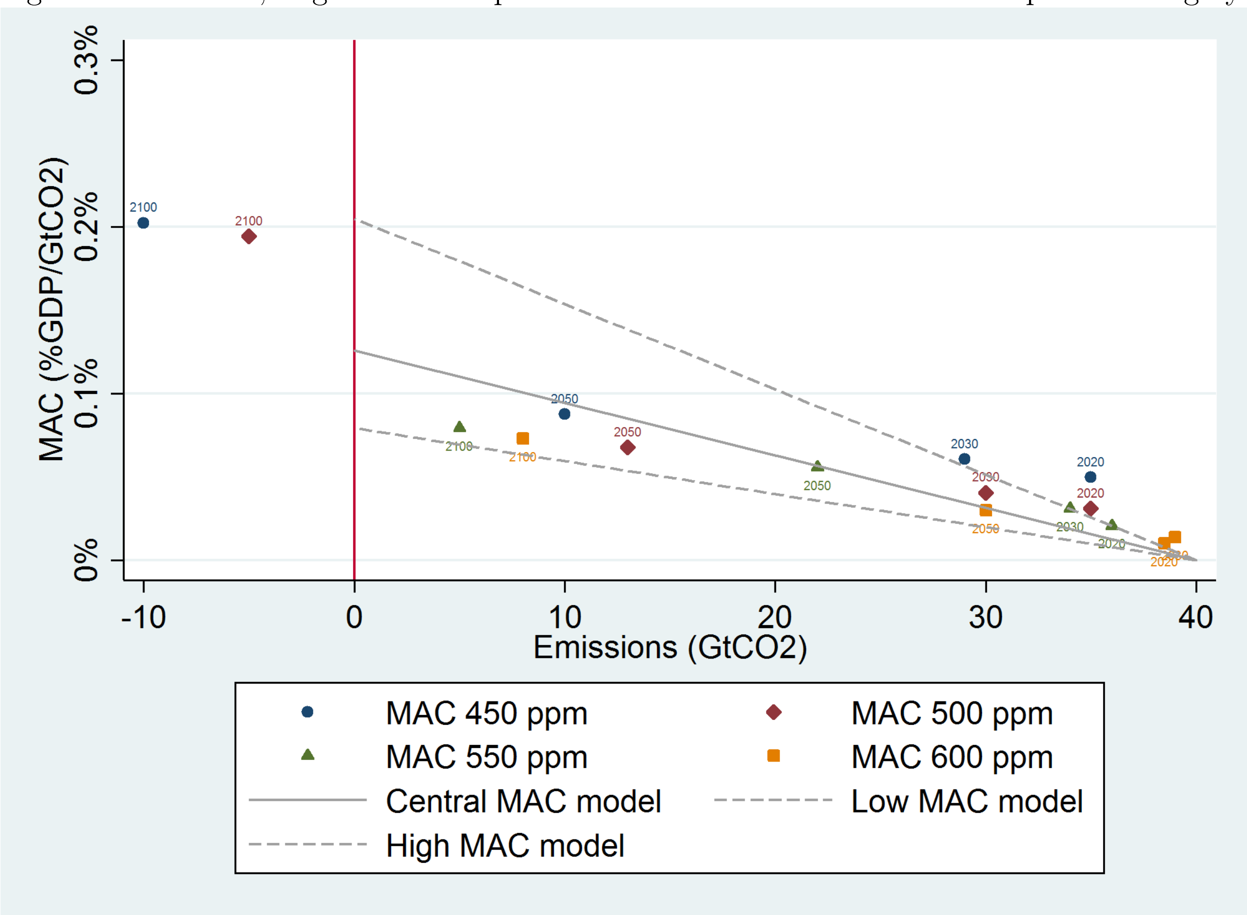
This also serves as the marginal abatement cost (MAC) function in our model, since abatement A can be defined as baseline or business-as-usual emissions ϕ/φ minus emissions, $A \equiv \phi/\varphi - E$, which allows us to rewrite (5) as

$$-Q_A = \varphi A Q. \tag{6}$$

This MAC function has two key properties. First, the MAC increases linearly as a function of abatement. Second, the MAC is proportional to output. Figure 3 looks at evidence from the IPCC *Fifth Assessment Report* on the shape of the MAC function, when expressed as a proportion of GDP. These are results derived from a variety of different energy models. It can be seen that a linear increasing function is a relatively good fit of the data.

It is often assumed that the MAC is convex increasing in abatement, at least with regard to *instantaneous* changes in abatement. The MAC rises, because increasingly expensive abatement technologies must be deployed (e.g. carbon capture and storage at the margin, instead of solar PV), and it is specifically assumed that the cost increase is more than proportional to the increase in the quantity of abatement. However, abatement can only change from one instant to another in our globally aggregated model, so we must also factor in – albeit in reduced-form – the empirical regularity that abatement technologies, such as renewable power, benefit from economies of scale over time, associated with learning-by-doing (Bramoullé and Olson, 2005; Neij, 2008). When the level of abatement increases from one instant to another, this means that learning-by-doing provides a countervailing effect

Figure 3: Global marginal abatement costs as a proportion of GDP under abatement scenario groups in the IPCC *Fifth Assessment Report* with peak atmospheric greenhouse gas concentrations of 430-480, 480-530, 530-580 and 580-630 ppm CO_2e. Median emissions for each scenario group are taken from Working Group III figure 6.7, median abatement costs are from Working Group III figure 6.21, and median growth rates are taken from figure 13 of the Synthesis Report. The MAC curves corresponding to the central, high and low parameter values in Section 3.1 are plotted in grey.



to any increase in the MAC that results from moving along the instantaneous MAC curve. This appears to be what is happening in the suite of IPCC energy models (Fig. 3).

The second assumption underpinning our MAC function is that the MAC is proportional to output. The main driver of an increasing MAC as a function of output is energy demand. Economic growth drives up energy demand, which will in turn drive up the MAC, because most low-carbon energy technologies have decreasing marginal productivity as the natural resources they require become scarcer (e.g. limited biomass, limited geological space for carbon storage, wind energy on less windy and/or more expensive locations, mineral resource constraints for batteries, etc.). For this reason, Rogelj et al. (2013), by way of a prominent example, find that abatement costs to limit warming to 2°C are ten times higher under the high energy demand scenario of the Global Energy Assessment, compared to the low energy demand scenario (refer to fig. 2 therein). Again, proportionality of the MAC with respect to output appears to be a feature of the IPCC energy models whose results are plotted in Figure 3.

2.4 Welfare and production

The last elements of our model are welfare, utility and production. Since we are interested in policy, we focus on the social planner's solution. Appendix B confirms this is equivalent to a decentralised competitive market equilibrium with a Pigouvian tax on CO_2 emissions. We assume the planner controls CO_2 emissions in order to maximise a discounted classical utilitarian social welfare functional:

$$\max_{E} W = \int_0^\infty e^{(n-\rho)t} u\left(c\right) dt, \tag{7}$$

where W is social welfare, n is the population growth rate (the initial population is normalised to unity), ρ is the utility discount rate and $u(c)$ is instantaneous utility as a function of per-capita consumption. Specifically

$$u(c) = \frac{c^{1-\eta}}{1-\eta}, \tag{8}$$

where η is the negative of the elasticity of marginal utility and consumption per capita is

$$c = e^{-nt} \left(1-s\right) Q,$$

where s is the savings rate and Q is aggregate output. It is best to think of s as an approximation of the savings rate that solves the planner's optimal investment problem. This is very unlikely to be exactly constant, however the assumption can be defended. Golosov et al. (2014) do so on three grounds: (i) a constant savings rate is optimal in a model with log utility, full depreciation of capital in one decade, and Cobb-Douglas production (with capital, labour and energy inputs); (ii) a constant savings rate is a good approximation of the optimal savings rate in a model where depreciation is less than full, the utility function is iso-elastic, or production is Leontief; (iii) observed savings rates worldwide do not exhibit much temporal variation. Our main reason for assuming constant savings is tractability: the model will not yield a closed-form solution otherwise (see below).

Aggregate output is given by

$$Q = F(t, K) \exp \left[nt - \frac{\gamma}{2} T^2 + \phi E - \frac{\varphi}{2} E^2 \right]. \tag{9}$$

Therefore production depends on capital and the evolution of the production technology $F(t, K)$,[8] as well as three terms contained within the exponent. The first term is climate damages according to (4). The second and third terms incorporate emissions as a factor of production and therefore abatement costs according to (5). Capital accumulation equals production minus consumption and depreciation:

$$\dot{K} = Q - ce^{nt} - \delta K. \tag{10}$$

3 The optimal path

The (unconstrained) optimal path is obtained when the planner maximises (7), subject to (2), (3), (10) and initial S, T and K. The current value Hamiltonian is

$$\mathcal{H} = \frac{1}{1-\eta} c^{1-\eta} - \lambda^S E - \lambda^T \varepsilon \left(\zeta S - T \right) + \lambda^K \left[Q \left(t, K, T, E \right) - e^{nt} c - \delta K \right],$$

where λ^S is the shadow price of cumulative emissions, λ^T is the shadow price of temperature and λ^K is the shadow price of capital. The Hamiltonian is defined such that all shadow prices are positive. Defining production per capita $q = Qe^{-nt}$ and substituting for λ^K, the

[8] We keep the functional form of $F(t, K)$ general. Note that if $F(t, K) = \left(\frac{K}{e^{nt}}\right)^{\alpha}$, production is Cobb-Douglas with respect to capital and labor.

necessary conditions for a maximum include

$$\lambda^S = c^{-\eta} q(\phi - \varphi E), \tag{11}$$

$$\dot{\lambda^S} = (\rho - n)\,\lambda^S - \varepsilon\zeta\lambda^T, \tag{12}$$

$$\dot{\lambda^T} = (\rho - n + \varepsilon)\,\lambda^T - c^{-\eta} q\gamma T, \tag{13}$$

$$Q_K - \delta = \eta\frac{\dot{c}}{c} + \rho \tag{14}$$

Equation (14) is the Ramsey rule and governs optimal capital accumulation. Equation (11) expresses the well-known optimality condition that the MAC must always equate to the social cost of carbon.

Integrating (13) gives

$$\lambda^T = \int_0^\infty e^{-(\rho-n+\varepsilon)t} c^{-\eta} q\gamma T dt\ . \tag{15}$$

In order that the climate system adjusts to a pulse of emissions quickly as per Figure 1, $\varepsilon \approx 0.5$ (and is no smaller than about 0.3), which makes it at least an order of magnitude larger than the population-adjusted discount rate $\rho - n$. Looking at (15), this means λ^T will be almost wholly dependent on the marginal disutility of warming $c^{-\eta}q\gamma T$ over the first few years. Over just a few years, we can safely assume that the marginal disutility of warming is constant: while marginal utility $c^{-\eta}$ decreases over the space of a few years in a growing economy, marginal damage $q\gamma T$ is increasing in a warming world, and neither will change much.

Assumption 1. *Because the climate system adjusts quickly to CO_2 emissions, $c^{-\eta}q\gamma T$ is constant over short periods and therefore*

$$\lambda^T \approx \frac{c^{-\eta} q\gamma T}{\rho - n + \varepsilon}. \tag{16}$$

This assumption allows us to rewrite (12) as

$$\dot{\lambda^S} = (\rho - n)\lambda^S + \frac{\varepsilon\zeta}{\rho - n + \varepsilon} c^{-\eta} q\gamma T.$$

Assumption 2. *The growth rate of output per capita net of climate damages and abatement costs $\hat{g}$ is approximately constant.*

Before justifying this assumption, note that it is only required if $\eta \neq 1$ (this is clear from (18) below). We assume the economy exhibits balanced growth at a constant rate g before climate damages and abatement costs. Constant growth of net output per capita then requires that both the rate of change of damages and the rate of change of abatement costs are much smaller than growth from capital accumulation and technological progress. This is also safe to assume. To demonstrate this, we can manipulate (9) into an expression for $\hat{g}$:

$$\hat{g} \equiv \frac{\dot{q}}{q} = g - \gamma T\dot{T} + \phi\dot{E} - \varphi E\dot{E}.$$

As Table 2 later shows, representative values of the damage function coefficient γ are smaller than a typical value of g. Moreover, on the optimal paths we find, $\dot{T}$ is also smaller than g, which means that $\gamma T\dot{T}$ will be much smaller than g. Similarly, even though $\dot{E}$ and $E\dot{E}$ can be significantly larger than g, Table 2 shows that the calibrated values of ϕ and φ are again very small relative to g, so that $\phi\dot{E} - \varphi E\dot{E}$ will amount to a small subtraction, overall.

Taking the time derivative of the first-order condition in (11) and substituting this into (12), we obtain

$$-\varphi\dot{E} = c\left[\left(\rho - n + \eta\frac{\dot{c}}{c} - \frac{\dot{q}}{q}\right)(\phi - \varphi E) - \frac{\varepsilon\zeta\gamma T}{\rho - n + \varepsilon}\right]. \tag{17}$$

Then applying Assumption 2 gives us an expression for the evolution of emissions:

$$\dot{E} = \left[\rho - n + (\eta - 1)\,\hat{g}\right](E - \phi/\varphi) + \frac{\varepsilon\zeta\gamma T}{(\rho - n + \varepsilon)\varphi}. \tag{18}$$

Integrating (2) gives

$$T_t = \int_{-\infty}^{t} e^{-\varepsilon(t-\tau)}\varepsilon\zeta S d\tau\ . \tag{19}$$

As was the case with the differential equation (15), the fact that $\varepsilon \approx 0.5$ means the value of the integral (19) is dominated by just a few years, in this case the most recent few years. Over such a short period, we can treat the growth rate of cumulative emissions as a constant, $\vartheta \equiv \dot{S}/S$. Then:

Assumption 3. *Because the climate system adjusts quickly to CO_2 emissions, ϑ is constant over short periods and*

$$T \approx \frac{\varepsilon}{\varepsilon + \vartheta}\zeta S. \tag{20}$$

We can then substitute (20) into (18) to obtain

$$\dot{E} = \left[\rho - n + (\eta - 1)\,\hat{g}\right](E - \phi/\varphi) + \frac{\varepsilon^2\zeta^2\gamma S}{(\rho - n + \varepsilon)\,(\varepsilon + \vartheta)\,\varphi}. \tag{21}$$

Rearranging (21) and then substituting $\dot{S}$ for E, we arrive at a linear differential equation for cumulative emissions:

$$\ddot{S} = \underbrace{\left[\rho - n + (\eta - 1)\,\hat{g}\right]}_{a}\dot{S} + \underbrace{\frac{\varepsilon^2\zeta^2\gamma}{(\rho - n + \varepsilon)\,(\varepsilon + \vartheta)\,\varphi}}_{b} S - \underbrace{\left[\rho - n + (\eta - 1)\,\hat{g}\right]\frac{\phi}{\varphi}}_{c}. \tag{22}$$

Clearly the linearity of (22), combined with constant coefficients and a constant term, is key to obtaining a closed-form solution for the optimal path.[9]

It is worth taking a moment to interpret the constants a, b and c, as they will often appear in the remainder of the analysis. The constant a is the standard 'Ramsey' discount rate minus the growth rate $\hat{g}$. As such it is the discount rate that is applied to the future flow of marginal damages from a tonne of CO_2 emitted at time t, when those damages are expressed as a proportion of output. This can be shown by integrating (17) with respect to time, dividing both sides by $c_t^{-\eta}q_t = c_\tau^{-\eta}q_\tau e^{\hat{g}(\eta-1)(\tau-t)}$ and defining

$$\mathrm{MAC}_{\%} \equiv \frac{Q_E}{Q} = -\frac{\epsilon}{\rho - n + \epsilon}\int_t^\infty e^{-(\rho-n+(\eta-1)\hat{g})(\tau-t)}\zeta\frac{Q_{T_\tau}}{Q_\tau}d\tau \equiv \mathrm{SCC}_{\%}, \tag{23}$$

where $\mathrm{SCC}_{\%}$ is the social cost of carbon as a proportion of GDP.

The reason that marginal damages as a proportion of output are discounted at the reduced rate $\rho - n + (\eta - 1)\,\hat{g}$ is that output growth has two countervailing effects on the social cost of carbon at any instant. On the one hand it reduces the present value of future damages, because it reduces marginal utility in the future. This is the conventional effect of discounting. On the other hand it increases the undiscounted value of future damages, because they are proportional to output in the model. This is an important feature of models where damages are multiplicative.

The constant b can be unpacked into

$$b = \frac{\varepsilon^2}{(\varepsilon + \vartheta)\,(\rho - n + \varepsilon)} \cdot \frac{\zeta^2\gamma}{\varphi}.$$

The first element is the delay factor, which can be further broken down into the physical effect of thermal inertia on marginal damages, $\varepsilon/\left(\varepsilon + \vartheta\right)$, and the discounting effect of thermal inertia, $\varepsilon/\left(\rho - n + \varepsilon\right)$. If temperature would adjust instantaneously to CO_2 emissions, then

[9] An extension to our model would be to have marginal damages and MACs that are not linearly proportional to consumption, i.e. $c_T = -c^\xi\gamma T$ and $c_E = c^\Phi\left(\phi - \varphi E\right)$, where ξ and Φ are the elasticities. This leads to an alternative differential equation,

$$\ddot{S} = \left[\rho - n + (\eta - \Phi)\,\hat{g}\right]\dot{S} + bSQ_0^{\xi-\Phi}e^{(\xi-\Phi)\hat{g}t} - \left[\rho - n + (\eta - \Phi)\,\hat{g}\right]\frac{\phi}{\varphi},$$

which is linear if $\xi = \Phi$.

the delay factor would be equal to one and $b = \zeta^2\gamma/\varphi$. This second element of b can also be written as $(Q_S/ - Q_A)(A/S)$. This can be interpreted the ratio of the slope of the marginal damage function with respect to S and the slope of the MAC function with respect to E, when both marginal damages and abatement costs are expressed as a proportion of output.[10] This ratio turns out to be central to interpreting our results for the optimal transition path. Lastly, the constant $c = a\phi/\varphi$, where ϕ/φ is baseline/business-as-usual emissions.

Returning to the task of solving the optimal path, the solution to the differential equation (22) is:

$$S_t = k_1 \exp \frac{1}{2}t\left(a - \sqrt{a^2+4b}\right) + k_2 \exp \frac{1}{2}t\left(a + \sqrt{a^2+4b}\right) + \frac{c}{b}. \tag{24}$$

The particular integral c/b is the inter-temporal equilibrium value of S. Bearing in mind that $\vartheta = 0$ at the inter-temporal equilibrium,

$$S^* = \frac{(\rho - n + \varepsilon)}{\varepsilon} \cdot \frac{[\rho - n + (\eta - 1)\hat{g}]\,\phi}{\zeta^2\gamma}. \tag{25}$$

Appendix C demonstrates that S^* is dynamically stable.

3.1 Peak warming

At S^*, the linear climate model dictates that the maximum increase in the global mean temperature relative to the pre-industrial level is simply $T^* = \zeta S^*$, so:

Proposition 1. *[Optimal peak warming] In the climate-economy system characterised by (2), (3) and (7)-(9), optimal peak warming is given by*

$$T^* = \zeta\frac{c}{b} = \underbrace{\frac{(\rho - n + \varepsilon)}{\varepsilon}}_{\textit{delay factor}} \cdot \frac{[\rho - n + (\eta - 1)\hat{g}]\,\phi}{\zeta\gamma}. \tag{26}$$

Proposition 1 tells us the maximum warming of the planet that is optimal from an economic point of view. The first element is the delay factor, but, not for the first time, the fact that ε is much larger than $\rho - n$ is significant. It means the delay factor will invariably be close to one. Take the central values of these three parameters as set out in Table 2; $\varepsilon = 0.5$ and $\rho - n = 0.006$. Then the delay factor is equal to 1.012. Even if we set $\rho - n = 0.03$, which we can take as about the maximum value that is plausible, the delay factor is equal to a still modest 1.06.

[10] In a version of the model without delay, $D(S) = \exp\left[-\gamma(\zeta S)^2\right]$ and so $Q_S = -2\gamma\zeta^2 SQ$. See Appendix D.

Table 1: Response of peak warming to changes in parameters

Parameter	Point elasticity of T^* with respect to parameter	Sign
ρ	$E_\rho = \frac{\rho}{\rho - n + (\eta - 1)\,\hat{g}} = \frac{\rho}{a}$	+
n	$E_n = -\frac{n}{[\rho - n + (\eta - 1)\,\hat{g}]} = -\frac{n}{a}$	-
η	$E_\eta = \frac{\eta\hat{g}}{\rho - n + (\eta - 1)\,\hat{g}} = \frac{\eta\hat{g}}{a}$	+
$\hat{g}$	$E_{\hat{g}} = \frac{(\eta - 1)\,\hat{g}}{\rho - n + (\eta - 1)\,\hat{g}} = \frac{(\eta - 1)\,\hat{g}}{a}$	+
ϕ	$E_\phi = 1$	+
ζ	$E_\zeta = -1$	-
γ	$E_\gamma = -1$	-

Corollary 1. *[The delay factor is insignificant to optimal peak warming] Because the climate system adjusts quickly to CO_2 emissions, optimal peak warming can be approximated by*

$$T^* \approx \frac{[\rho - n + (\eta - 1)\,\hat{g}]\,\phi}{\zeta\gamma}. \tag{27}$$

This is also naturally the exact solution of the model when warming is simply *assumed* to be an instantaneous function of cumulative emissions, as shown in Appendix D. Appendix E shows using numerical techniques that the versions of the model with and without a temperature delay give very similar optimal warming and are both very close approximations of the numerical solution to the maximisation problem, which takes the short delay into account, while not depending on Assumptions 1 or 3. Comforted by this, we henceforth work with the model without a temperature delay.

In Table 1 we compute the point elasticities of T^* with respect to the parameters that feature in (27). We find that optimal peak warming is an increasing function of the pure rate of time preference ρ, a new version of an old result. Since there is no delay between CO_2 emissions and warming from those emissions, this is fundamentally due to the long residence time of CO_2 in the atmosphere. Close inspection of the point elasticity of T^* with respect to ρ reveals that it is equal to the ratio of ρ to a, the discount rate on $SCC_{\%}$. Population growth n has the opposite effect on peak warming to ρ, because it reduces the population-adjusted discount rate.[11]

Increases in both η and the growth rate $\hat{g}$ result in an increase in optimal peak warming, provided that $\eta \geq 1$. Moreover, comparing the two elasticities, it is clear that the elasticity

[11]Notice that in the limit as $\eta \to 1$ (i.e. log utility), the elasticity of T^* with respect to $\rho - n$ is one: a doubling of $\rho - n$ leads to a doubling of optimal peak warming. Higher η tempers this, but given the magnitudes involved it does so only slightly.

of T^* with respect to η is larger by exactly $\hat{g}$, which reflects the fact that, whereas η only has an effect on the discount rate, $\hat{g}$ affects both the discount rate and the undiscounted value of marginal damages, as explained above.

Three of the model parameters have an especially simple relationship with optimal peak warming. There is a negative unit elasticity of T^* with respect to ζ, the TCRE parameter, and γ, the coefficient of the damage function. A one per cent increase in either of these parameters reduces optimal peak warming by one per cent. Conversely there is a unit elasticity of peak warming with respect to ϕ, the marginal cost of zero emissions. Notice that peak warming is independent of the parameter φ that governs the slope of the MAC function. Fundamentally this is because T^* is determined by comparing $\text{SCC}_{\%}$ at T^* with the abatement cost of zero emissions ϕ (see Eq. 23), which does not depend on φ.

Table 2 presents central values of the model's parameters, as well as ranges from the literature.[12] If we plug the parameters' central values into Eq. (27), we obtain optimal peak warming of 3.4°C, corresponding to stationary cumulative emissions of 7,014 gigatonnes of CO_2 since the beginning of the industrial revolution. With central values of ρ, n, η and $\hat{g}$, the consumption discount rate is about 3.1%, while the central value of γ implies that 2°C warming causes a loss of output of 2% and 4°C warming causes a loss of output of 8%. Therefore damages in the central case are relatively modest and they are discounted at a medium rate, which explains why optimal peak warming is well above 2°C.

Considering the ranges of parameter values in Table 2, it is clear that peak warming is highly sensitive to most of the model parameters. Take for instance the TCRE parameter ζ. A central estimate from climate science might be 0.00048°C per gigatonne of CO_2. But the range of uncertainty about ζ spans approximately +/-50%. Given that optimal peak warming has a unit elasticity with respect to ζ, optimal peak warming varies by +/-50% accordingly. Much the same is true of the other two parameters with a unit elasticity: the range of uncertainty either side of the central value of γ is -50% to +100%, while for ϕ it is -40% to +120%. The elasticities of T^* with respect to the other four parameters are non-constant, however in most cases they can also be expected to be large. Holding the other parameters to their central values, E_ρ will be close to one over the range of ρ, which according to Drupp et al. (2015) is -45% to +209%. E_η is particularly high, ranging from 1.3 for maximum η to 3.2 for minimum η, with $E_\eta = 2.1$ for the central value (again holding the other parameters to their central values). This makes clear the limitations of models that assume log utility when thinking about uncertainty governing optimal warming.

[12] We combine ρ and n in view of their diametrically opposing effects in the model (on the utility discount rate). The parameters ϕ and φ are jointly determined, so their respective minima, central values and maxima must be taken together, although φ does not feature at this juncture.

Table 2: Parameter values for numerical analysis

Parameter	Min.	Central value	Max.	Sources
$\rho - n$	0.006-0.005	0.011-0.005	0.034-0.003	Drupp et al. (2015); United Nations (2015)
η	1.01	1.35	3	Drupp et al. (2015)
$\hat{g}$	0.01	0.0195	0.03	g by assumption, $g - \hat{g}$ calibrated on numerical solution
ϕ	0.00079	0.00126	0.00205	Clarke et al. (2014)
φ	0.00002	0.00003	0.00005	
ζ	0.000240	0.000480	0.000749	Collins et al. (2013); Matthews et al. (2009)[a]
γ	0.005	0.01	0.02	Nordhaus and Moffat (2017); Weitzman (2012)

[a]Multiplied by 1.1 to adjust for non-CO_2 greenhouse gases (see Section 2.1).

Figure 4: Sensitivity of T^* to uncertain parameters

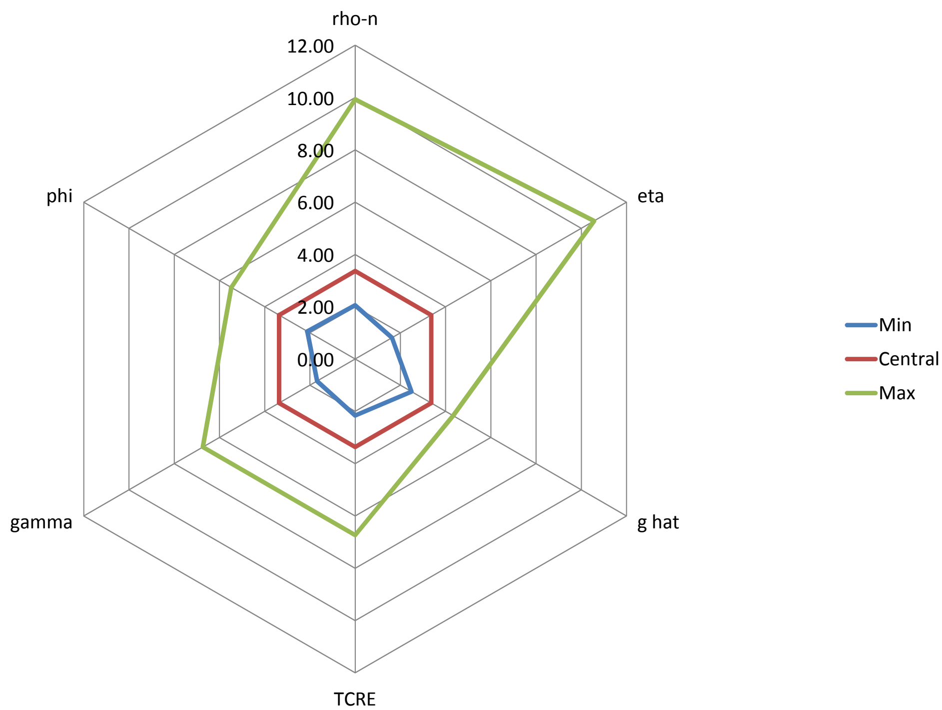


Figure 4 plots optimal peak warming as a function of variation in the model parameters. For this we impose a constraint that cumulative emissions may not exceed 'burnable carbon' embodied in the Earth's fossil fuel resources.[13] The constraint binds only with respect to η. When looking at sensitivity with respect to ζ, bear in mind that, not only does lower (higher) ζ result in higher (lower) optimal cumulative emissions, it also results in lower warming as a result of those emissions. Observe that when $\rho - n$ is set to its minimum value of 0.1%, optimal peak warming is 2.0°C. When η is set to its minimum value of roughly one, optimal peak warming is 1.6°C. When γ is set to its maximum value, such that 2°C warming causes a loss of output of 4% and 4°C warming causes a loss of output of 16%, optimal peak warming is 1.7°C. Many combinations of parameter values support optimal peak warming of 2°C or below.

[13] These are estimated to be in the region of 22,000 $GtCO_2$, including fossil fuels burned since the beginning of the industrial revolution (Nordhaus, 2008). When some parameters take extreme values, optimal cumulative emissions may exceed this. This constraint gives peak warming of 10.6°C for the central value of ζ.

3.2 The slow transition to equilibrium

While an analysis of optimal peak warming reveals useful information, it does not reveal how long it takes for warming to peak along the optimal path and therefore it is unlikely to reveal the key features of optimal emissions in the near future.

Appendix C demonstrates that the transition to S^* is governed by

$$S_t = \left(S_0 - \frac{c}{b}\right) \exp \frac{1}{2} t \left(a - \sqrt{a^2 + 4b}\right) + \frac{c}{b}. \tag{28}$$

Since $b > 0$, the exponent is negative and cumulative emissions approach their stationary value c/b asymptotically. Put another way, optimal emissions are strictly decreasing, at a decreasing rate. There is an intuitive explanation for this: the social cost of carbon as a proportion of output is an increasing function of S.[14] Since $E = \dot{S} > 0$, $\text{SCC}_{\%}$ increases all along the path. Since the MAC function is linear increasing as a proportion of output, the necessary condition for an optimum that $\text{SCC}_{\%} = \text{MAC}_{\%}$ means that emissions must decrease all along the path. It is not optimal for emissions to peak at $t > 0$, for instance.

But how fast do emissions approach zero? In other words, how long does it take for warming to approach its peak? It turns out that the answer is slowly, very slowly indeed. Figure 5 plots optimal paths of T over the next 250 years that correspond with our central parameter values, as well as with scenarios of high and low damages, which we choose as being illustrative of the transition path when optimal peak warming is low and high respectively. These optimal paths are obtained by plugging Eq. (28) into (2).

Although optimal peak warming corresponding with our central parameter values is 3.4°C, optimal (transient) warming a century from now is just 1.7°C; 250 years from now it is 2.5°C. When damages are high, optimal peak warming is 1.7°C, but optimal warming a century from now is just 1.3°C. When damages are low, optimal peak warming is 6.7°C, but optimal warming in a century's time is only 2.2°C. So, while peak warming is highly sensitive to the parameters that determine it, warming over the next couple of centuries is much less so.

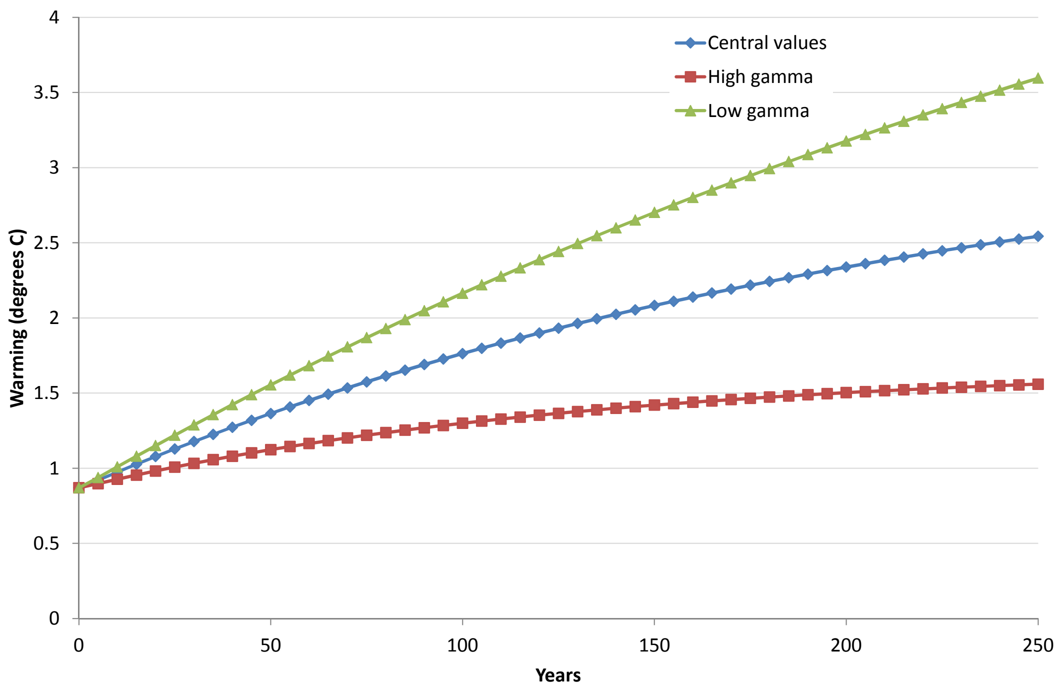
Why is the transition so long? The rate of change of emissions is

$$\frac{\dot{E}}{E} = \frac{1}{2}(a - \sqrt{a^2 + 4b}). \tag{29}$$

A slow transition to peak warming implies $\left|\dot{E}/E\right|$ is small. The reason for this is that b is very small. Recall that b is the ratio of the slope of the marginal damage function with respect to S and the slope of the MAC function with respect to E, when marginal damages

[14] $\text{SCC}_{\%} = \int_t^{\infty} e^{-(\rho - n + (\eta+1)\hat{g})(\tau - t)} \zeta^2 \gamma S_\tau d\tau$.

Figure 5: The optimal transition path of T for central parameter values, and low and high damages (γ)



and abatement costs are expressed as a proportion of output:

$$b=\frac{\zeta^2\gamma}{\varphi}=\frac{Q_S/S}{Q_E/\left(E_{\mathrm{BAU}}-E\right)}=\frac{Q_S}{-Q_A}\frac{A}{S}.$$

Equation (23) shows that Q_A is much larger than Q_S, because the latter is a perpetual stream of damages from a non-decaying stock of CO_2. The second factor of b, A/S, is also small, because abatement A is a flow and S is a non-decaying stock. Therefore this result bears the imprint of the flow-stock nature of CO_2-induced warming, which means that there is a certain affinity between it and the comparison between carbon price and quantity instruments under uncertainty about marginal abatement costs (Weitzman, 1974; Pizer, 2002).[15] It is this flow-stock property that leads to the result illustrated by Figure 5, where optimal emissions in the near term are much less sensitive to parameter variations that lead to large differences in optimal peak warming. The short delay between emissions and warming is a driving force behind this result. Since damages occur almost immediately, it is worth avoiding them from the start. The flow-stock dynamic also stems from the fact that warming does not decay in our climate model.

However, a weakness of the model in characterising the transition to peak warming is that it ignores 'locked-in' emissions from the capital stock existing at $t = 0$, which will in reality constrain near-term emissions reductions, presumably leading to a transition path where emissions are higher in the near term and lower in the long term, and where warming thereby approaches its peak faster. A simple way to account for this and therefore to test the robustness of our stylised finding of a slow transition is to increase initial S by the cumulative emissions embodied in the global capital stock today, assuming it is operated to the end of its economic lifetime. Davis and Socolow (2014) have estimated that future cumulative CO_2 emissions embodied in global power plants in 2012 was 307 $GtCO_2$. [16] Adding this to initial S, the transition to peak warming is faster, but only marginally so. For central parameter values, optimal warming a century from now rises from 1.7°C to 1.9°C. When damages are high, it rises from 1.3°C to 1.4°C.

3.3 Carbon prices

As well as peak warming, we can characterise the optimal carbon price by differentiating (28) with respect to time, substituting the resulting expression into (5) and rearranging:

[15] The presumption in favour of prices rests on the slope of the marginal damage function being presumably much gentler than the slope of the MAC function over a timescale relevant for deciding between prices and quantities.

[16] This is of course most likely to be an underestimate of so-called 'committed emissions', because it only covers the power sector.

Proposition 2. *[The optimal carbon price] In the climate-economy system characterised by (2), (3) and (7)-(9), the optimal carbon price is*

$$p^* = MAC = \underbrace{Q_0 e^{(\hat{g}+n)t}}_{growth\ effect} \cdot \underbrace{(\phi - \varphi E)}_{emissions\ effect}, \tag{30}$$

where $E = \left(S_0 - \frac{c}{b}\right)\frac{1}{2}\left(a - \sqrt{a^2+4b}\right)\exp\left[\frac{1}{2}t\left(a - \sqrt{a^2+4b}\right)\right]$.

Proposition 2 shows that the evolution of the carbon price depends on two factors. On the one hand, the carbon price is proportional to output, so as output grows at the rate $\hat{g}+n$ the carbon price does likewise, all else being equal. We call this the growth effect. On the other hand, the carbon price depends on emissions, which means that the evolution of the carbon price is also subject to the emissions dynamics set out above. In particular, what we call the emissions effect increases, but it does so at a decreasing rate, since emissions converge to zero in the long run. The overall effect is that p^* grows at a rate that is initially faster than aggregate output grows, but converges to $\hat{g}+n$ asymptotically, with the transition governed via E by a and b:

$$\begin{aligned} \ln p^* &= \ln Q_0 + (\hat{g}+n)t + \ln(\phi - \varphi E), \\ \tfrac{\dot{p^*}}{p^*} &= \hat{g} + n + \frac{\dot{A}}{A}. \end{aligned}$$

In the steady state, the optimal carbon price expressed as a percentage of GDP is ϕ.

As a corollary to Proposition 2, we can show that in our model the optimal carbon price grows at the same rate as aggregate output if damages are an exponential-linear rather than exponential-quadratic function of warming. If damages are exponential-linear in warming, then marginal damage is constant in cumulative emissions.

Corollary 2. *[The optimal carbon price under exponential-linear damages] In a climate-economy system where* $D(T) = \exp(-\gamma T)$*, the optimal carbon price grows at the rate* $\hat{g}+n$.

Proof. If $D(T) = \exp\left(-\frac{\gamma}{2}T\right) = \exp(-\frac{\gamma}{2}\zeta S)$, marginal damage as a function of cumulative emissions is $Q_S = \zeta\gamma Q$, assuming away the temperature delay. Instead of Eq. (38), we have

$$\frac{Q_{E_t}}{Q_t} = -\int_t^\infty e^{-(\rho-n+(\eta-1)\hat{g})(\tau-t)}\zeta\gamma d\tau = \frac{\zeta\gamma}{\rho - n + (\eta-1)\hat{g}}.$$

Hence the carbon price is a fixed proportion of aggregate output,

$$Q_E = Q\frac{\zeta\gamma}{\rho - n + (\eta-1)\hat{g}},$$

and increases at $\hat{g} + n$. $\square$

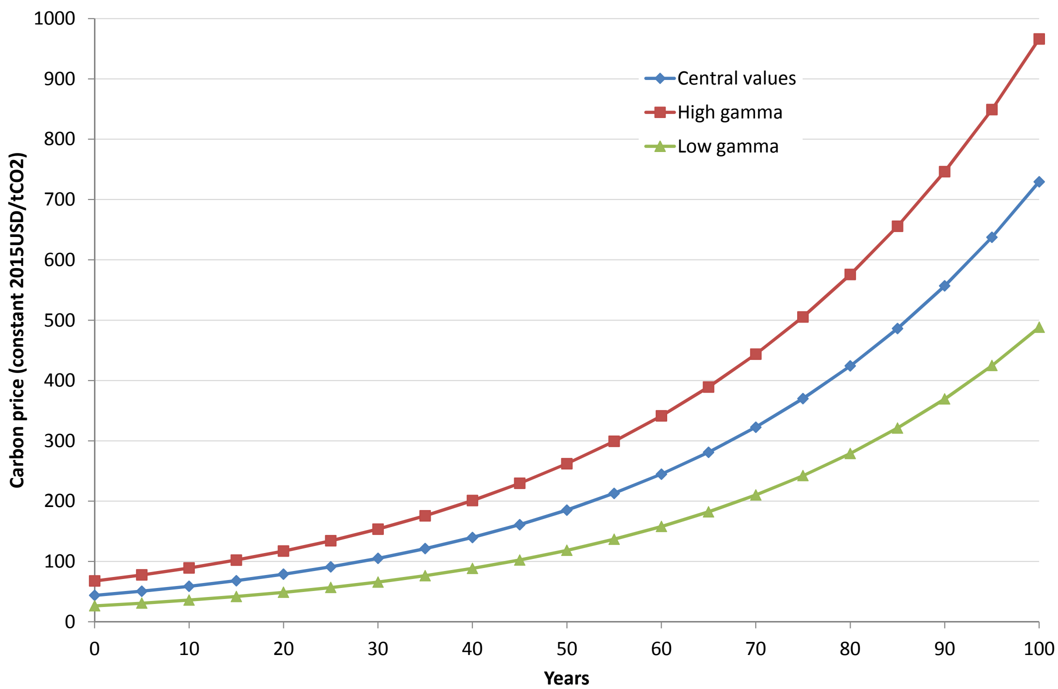
Golosov et al. (2014) also found that the optimal carbon price grows at the same rate as the economy, although they assumed damages are an exponential-linear function of atmospheric CO_2, not of temperature, i.e. $Q = Q_0 \exp(-\gamma M)$. The relationship between the two approaches can be better understood if we decompose marginal damage as a function of cumulative emissions,

$$\frac{d \ln Q}{dS} = \frac{d \ln Q}{dT} \frac{dT}{dM} \frac{dM}{dS}.$$

Relating this back to Eq. (1), the right-hand side is marginal damage as a function of warming, multiplied by the TCRE in the limit as $\Delta \to 0$. In Golosov et al. (2014), $d \ln Q/dS$ is constant, because increasing marginal damages with respect to temperature $(d^2 \ln Q/dT^2 > 0)$ are exactly offset by decreasing marginal climate sensitivity $(d^2T/dM^2 < 0)$, and marginal carbon sensitivity is constant $(d^2M/dS^2 = 0)$. By contrast, in our model $d \ln Q/dS$ is constant if and only if marginal damages are constant with respect to temperature, because decreasing marginal climate sensitivity is exactly compensated by increasing marginal carbon sensitivity. That is, the TCRE is constant. So the optimal carbon price grows faster than the economy in our standard model (Proposition 2), because marginal damages are an increasing function of cumulative emissions, and the saturation of carbon sinks means that marginal carbon sensitivity is increasing.

Figure 6 plots optimal carbon prices under our central parameter values, and in scenarios of low and high damages. The optimal carbon price corresponding with our central parameter values starts at \$44/tCO_2 today and increases to \$59 in 10 years' time, \$185 at $t = 50$ and \$729 at $t = 100$. The rate of increase of the optimal price falls from 3.0% (real) initially to 2.7% after 100 years, which is close to the growth rate of aggregate output, assumed to be just under 2.5%. The optimal price in the low damages scenario starts at \$26/tCO_2 and increases to \$36 after 10 years, \$118 at $t = 50$ and \$488 at $t = 100$. This reinforces the message of the previous passage that, even if optimal peak warming is high, optimal transient warming over the coming centuries is low. Achieving this requires a significant and significantly increasing carbon price. Again the rate of increase of the optimal price in this scenario falls over time, but at 3.2% it is initially higher than the central case, falling to 2.8% after 100 years. The optimal price in the high damages scenario starts at \$68/tCO_2 and rises to \$966 after a century. The price grows in this scenario at a rate of 2.8% initially, falling to 2.6% after a century.

Figure 6: Optimal carbon prices for central parameter values and low and high damages (γ)



4 The optimal path under a temperature constraint

Important as it is to examine the unconstrained optimum of the model, the majority of countries in the world have ratified the Paris Agreement, the central aim of which is "Holding the increase in the global average temperature to well below 2°C above pre-industrial levels and pursuing efforts to limit the temperature increase to 1.5°C above pre-industrial levels". This indicates that, as a description of the real world, the maximisation problem in Section 3 could be under-specified. Rather, we might say that the Paris Agreement leaves us with the objective of maximising (7), subject to (2) and (3), initial S and T, *and* the inequality constraint that $S \leq \overline{S}$, where $\overline{S} = \zeta\overline{T}$ and $\overline{T}$ is 2°C (or even 1.5°C).[17]

4.1 Maximising welfare subject to the temperature constraint

Technical details are relegated to Appendix F. Solving the constrained maximisation problem, we find:

Proposition 3. *[The optimal carbon price under a binding temperature constraint] When cumulative CO_2 emissions are constrained such that $S \leq \overline{S}$, where $\overline{S} = \zeta\overline{T}$, the optimal carbon price is*

$$\mathrm{MAC} = \mathrm{SCC} + \left(\phi - \frac{\gamma\zeta^2\bar{S}}{\rho - n + (\eta - 1)\hat{g}}\right) Q_{\bar{t}} e^{-(\rho - n + \eta g)(\bar{t} - t)}, \tag{31}$$

where $\bar{t}$ is the time when the cumulative emissions constraint binds. Therefore the optimal carbon price under a temperature constraint equals the social cost of carbon, plus a premium, which is a function of the cumulative emissions constraint and which increases at the discount rate (van der Ploeg, 2017, obtains a structurally similar result). The premium therefore follows Hotelling's rule, ensuring that the cumulative emissions budget implied by $\overline{S} < S^*$ is allocated in an inter-temporally efficient manner.

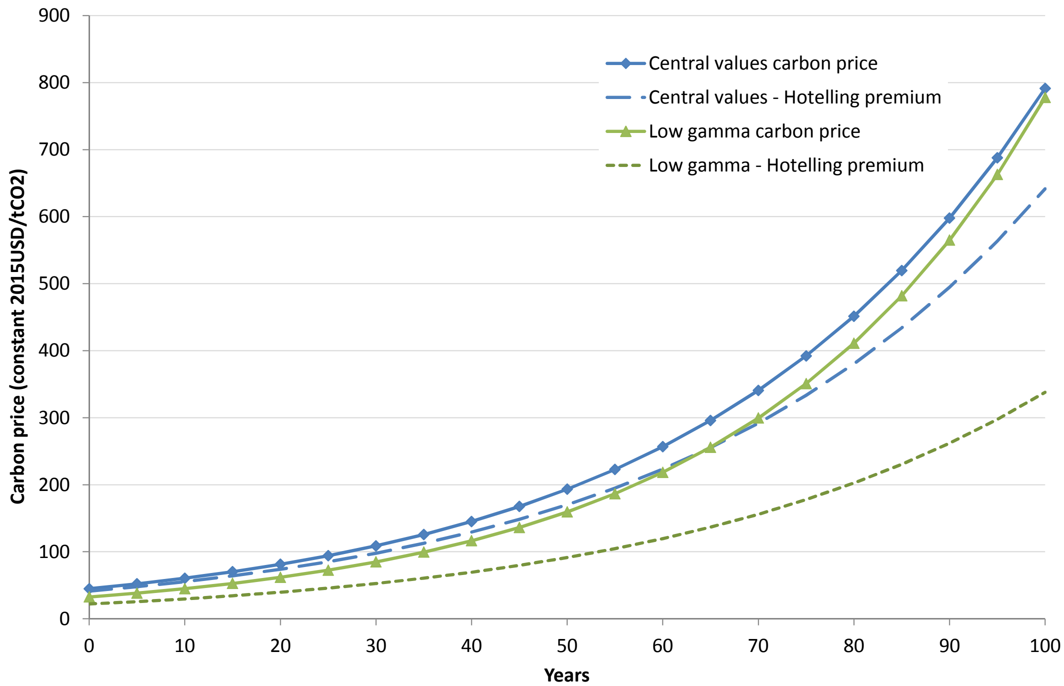
If the temperature constraint binds, we have

$$\bar{S} < S^* \Leftrightarrow \bar{S} < \frac{(\rho - n + (\eta - 1)g)\,\phi}{\zeta^2\gamma} < \frac{(\rho - n + \eta g)\,\phi}{\zeta^2\gamma} \Leftrightarrow \phi - \frac{\gamma\zeta^2\bar{S}}{\rho - n + \eta\hat{g}} > 0.$$

Therefore the premium is strictly positive, which further implies that emissions will be lower everywhere on the constrained path compared with the unconstrained path.

[17] If the constraint binds, then it is obviously at variance with the planner's optimal policy, based on the parameter values that the planner believes in: more than 1.5-2°C peak warming would result from these. Some might then regard the existence of a constraint on the planner's problem as a logical inconsistency. However, in our view there is no logical inconsistency, once one recognises that the Paris Agreement is a political constraint, which is partly motivated by non-welfarist principles.

Figure 7: Optimal carbon prices under a temperature constraint of 2°C for central parameter values and low γ



The Hotelling price premium required to stay within the temperature constraint is significant, even in the relatively near term. Figure 7 shows that the Hotelling premium begins at \$4/tCO_2 today, rising to \$5 in 10 years, \$23 in 50 years and \$150 in 100 years, under central parameter values. This is on top of a social cost of carbon of \$41/tCO_2 today, \$55 in 10 years, \$170 in 50 years and \$641 in 100 years. Notice that the social cost of carbon is lower than in the corresponding unconstrained optimisation (Figure 6), because cumulative emissions and therefore warming are lower. When the Hotelling premium is added on, however, the overall carbon price is higher than its equivalent in the unconstrained optimisation. Figure 7 also shows that when $\gamma = 0.005$ the Hotelling premium is a larger share of the carbon price, both because the social cost of carbon is lower and because, with higher optimal unconstrained warming, the constraint binds earlier.

4.2 Minimising abatement costs to meet the temperature constraint

Most studies on the costs of emissions abatement solve a different problem to the preceding section. In particular, they ignore climate damages and determine the emissions path that meets the constraint $\overline{S}$ at minimum total discounted abatement cost (Clarke et al., 2014). This is often referred to as cost-effectiveness analysis, as opposed to cost-benefit analysis. In our set-up, the cost-effective policy is the solution to maximising (7), subject to (2) and (3), initial S and T, and $S \leq \overline{S}$, but where the marginal disutility of warming is zero. The optimal carbon price path follows straightforwardly from Eq. (37):

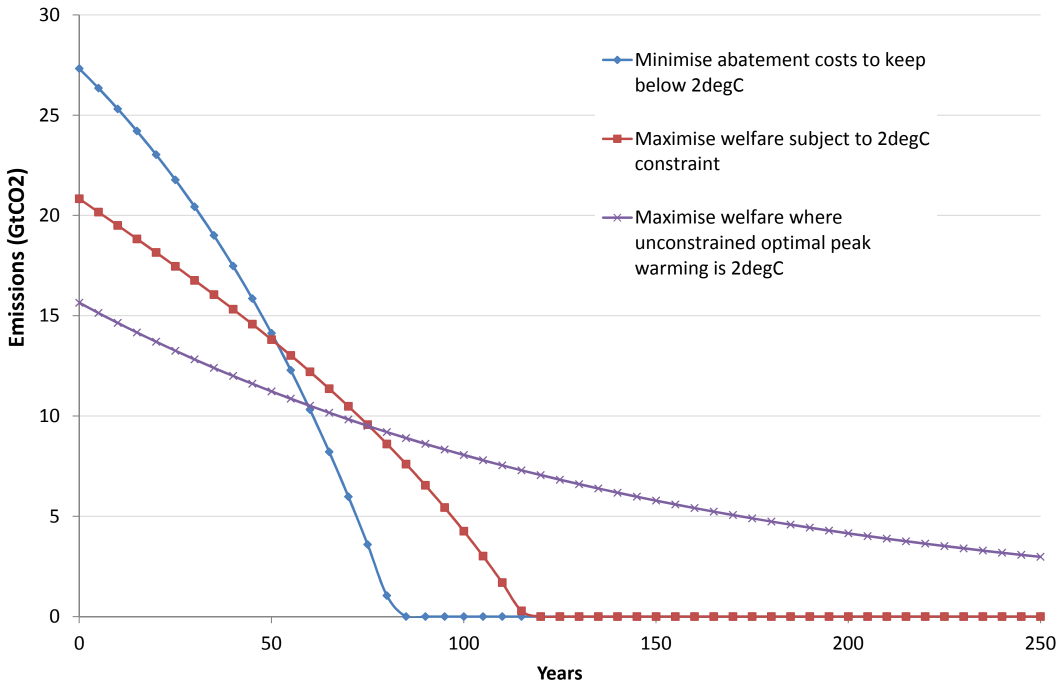
Proposition 4. *[The cost-effective carbon price] When cumulative CO_2 emissions are constrained such that $S \leq \overline{S}$, where $\overline{S} = \zeta\overline{T}$, and damages are ignored, the optimal carbon price is*

$$Q_E = Q_0 \left(\phi - \varphi E_0\right) e^{(\rho - n + \eta\hat{g})t}. \tag{32}$$

That is, inter-temporal efficiency is ensured by letting the carbon price follow the *simple* Hotelling rule. This is different to the standard assumption that the cost-effective carbon price increases at the 'augmented' Hotelling rate, i.e. at the consumption discount rate *plus* the decay rate of CO_2 in the atmosphere. This assumption rests on atmospheric decay creating a reason to postpone abatement, since CO_2 emitted earlier has the chance to decay more. Decay also enlarges the carbon budget for given $\overline{T}$. However, while this is true in and of itself, the saturation of carbon sinks, which our model accounts for, has the opposite effect; additional emissions today saturate the carbon sinks earlier. Saturation of carbon sinks reduces the carbon budget for given $\overline{T}$. Lemoine and Rudik (2017) have argued for a different kind of augmented Hotelling rule, because of thermal inertia, which again enlarges the carbon budget for given $\overline{T}$. But we have shown that modest thermal inertia together with saturation of carbon sinks more-or-less exactly offset the effect of decay of atmospheric CO_2. Lemoine and Rudik's result depends on significant thermal inertia, which they calibrated on DICE. Section 2 showed that DICE is arguably too slow to respond to CO_2 emissions. Consequently the simple Hotelling rule is in fact appropriate.

Appendix F shows that the rate of emissions reduction must be faster on the cost-effective path than on the cost-benefit path. Because both paths must result in the same cumulative emissions, the cost-effective emissions path must therefore begin with higher emissions, but eventually cross the constrained cost-benefit path and reach zero emissions faster.

Figure 8: Optimal emissions under a temperature constraint of 2°C when the discounted sum of total abatement and damage costs are minimised, compared with when only abatement costs are minimised, and when temperature is unconstrained but optimal peak warming is 2°C (high γ).



Proposition 5. *[Cost-effective emissions abatement is lower initially, but higher eventually] Compared with the emissions path that maximises net benefits, subject to the emissions constraint, the cost-effective emissions path has higher emissions initially, but emissions fall to zero earlier.*

Figure 8 shows the difference in the cost-benefit and cost-effective emissions paths, for central parameter values. We also include for illustration an unconstrained, welfare-maximising emissions path, where γ is solved backwards so that optimal peak warming is 2°C. Initial emissions on the cost-effective path are about 44% higher than on the constrained cost-benefit path, but the rate of emissions reduction is always higher and the two paths cross after about 50 years. Finally, observe how low and flat the emissions path is when optimal peak warming is 2°C; initial emissions are about 31% lower than on the constrained cost-benefit path.

5 Conclusions

In this paper we have built a model of optimal CO_2 emissions by exploiting recent advances in climate science, which have identified a near-instantaneous and quasi-linear warming response to cumulative CO_2 emissions, and combining them with reduced-form representations of climate damages and the costs of CO_2 emissions abatement, which are capable of capturing the stylised facts of the large applied literatures on each topic.

The model is surprisingly simple and yields closed-form solutions for optimal peak warming, optimal emissions along the transition to peak warming and optimal carbon prices, including under a temperature constraint that is consistent with the Paris Agreement. We draw five conclusions:

1. Optimal peak warming has an elasticity of one or more with respect to several parameters that are highly uncertain. This implies optimal peak warming is itself highly uncertain.

2. Even if optimal peak warming is high, optimal transient warming over the coming centuries is not. The transition is slow, because of the stock-flow nature of CO_2-induced warming. If optimal peak warming is 3.4°C, optimal transient warming one century from now is only 1.7°C.

3. The optimal carbon price initially grows faster than output per capita, but in the long run it grows at the same rate, because there is an emissions effect that converges to a constant factor. The underlying reason, however, is that damages are a convex function of cumulative emissions, which is amplified by the saturation of carbon sinks. For central parameter values, we calculate that the optimal carbon price grows 0.5 percentage points faster than the economy initially.

4. The optimal carbon price under a binding temperature constraint comprises the social cost of carbon, plus a Hotelling premium. If we take account of damages, then we should abate emissions more quickly than if we simply meet the temperature constraint at the lowest discounted abatement cost. This effect is quantitatively large.

5. When the objective is to minimise abatement costs alone, the optimal carbon price follows the simple Hotelling rule, not various kinds of augmented Hotelling rule, as in previous work. This is because modest thermal inertia and the saturation of carbon sinks more-or-less exactly offset the effect of decay of atmospheric CO_2.

Finally, our paper has generated many points of comparison with the literature, particularly other analytical IAMs. We synthesise these points of comparison in Table 3, with a focus on rules for optimal carbon price growth and the cumulative emissions budget. The rate of decay of atmospheric CO_2 is denoted δ. The results are independent of the shape of the MAC curve, and the damage functions in the cost-benefit models are all virtually equivalent (assuming a unit elasticity of marginal damages with respect to income), so the differences between the pricing rules and cumulative emissions budgets come down to features of the climate system. The Table highlights the crucial role of feedback from the saturation of carbon sinks to the decay of atmospheric CO_2, which is not present in other models and is a key driver of warming being linearly proportional to cumulative emissions.

References

Allen, Myles R, "Drivers of peak warming in a consumption-maximizing world," *Nature Climate Change*, 2016.

_ , David J Frame, Chris Huntingford, Chris D Jones, Jason A Lowe, Malte Meinshausen, and Nicolai Meinshausen, "Warming caused by cumulative carbon emissions towards the trillionth tonne," *Nature*, 2009, *458* (7242), 1163–1166.

Bramoullé, Yann and Lars J Olson, "Allocation of pollution abatement under learning by doing," *Journal of Public Economics*, 2005, *89* (9), 1935–1960.

Brock, William A, "A polluted golden age," in V L Smith, ed., *Economics of Natural and Environmental Resources*, New York: Gordon and Breach, 1973, pp. 441–461.

Caldeira, K and NP Myhrvold, "Projections of the pace of warming following an abrupt increase in atmospheric carbon dioxide concentration," *Environmental Research Letters*, 2013, *8* (3), 034039.

Calel, Raphael and David A Stainforth, "On the physics of three Integrated Assessment Models," *Bulletin of the American Meteorological Society*, 2017, (2016).

Clarke, L., K. Jiang, K. Akimoto et al., "Asessing transformation pathways," in O. Edenhofer, R. Pichs-Madruga, Y. Sokona et al., eds., *Climate Change 2014: Mitigation of Climate Change. Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, Cambridge, UK and New York, NY, USA: Cambridge University Press, 2014.

Collins, M., R. Knutti, J. Arblaster, J.-L. Dufresne, T. Fichefet, P. Friedlingstein, X. Gao, W.J. Gutowski, T. Johns, G. Krinner, M. Shongwe, C. Tebaldi, Weaver A.J., and M. Wehner, "Long-term climate change: projections, commitments and irreversibility," in T.F. Stocker, C. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia,

Table 3: Overview of the effect of different climate features on carbon prices

Model / pricing rule	CO_2 decay	Saturating IR wavelengths	Thermal inertia	Saturating CO_2 sinks	Growth rate of optimal carbon price	Cumulative CO_2 budget (#1 is smallest)
Cost-benefit models						
Golosov et al.	x	x			$\hat{g}+n$	2
Rezai and van der Ploeg van den Bijgaart et al.	x	x	x		$\hat{g}+n$ with unit elasticity of marginal damages wrt. income[a]	3
Our model	x	x	x	x	$\delta+\eta\hat{g}-\frac{\gamma T}{\phi-\varphi E}$, converging to $\hat{g}+n$	1
Cost-effectiveness models						
Augmented Hotelling	x				$\delta+\eta\hat{g}+\delta$	2
Lemoine and Rudik	x	x	x		$\delta+\eta\hat{g}+\delta$ $-\frac{\lambda^T}{u_c}\frac{\varepsilon z}{M_{pre}+M}$	3
Our model	x	x	x	x	$\delta+\eta\hat{g}$	1

[a]The growth rate of the carbon price is $\xi(g+n)$, if the elasticity of marginal damage with respect to output$\xi \neq 1$. This is a feature that could be applied to any model here in the cost-benefit class.

V. Bex, and P.M. Midgeley, eds., *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, Cambridge, UK: Cambridge University Press, 2013.

Davis, Steven J and Robert H Socolow, "Commitment accounting of CO2 emissions," *Environmental Research Letters*, 2014, *9* (8), 084018.

Drupp, Moritz, Mark Freeman, Ben Groom, Frikk Nesje et al., "Discounting disentangled: an expert survey on the determinants of the long-term social discount rate," *Centre for Climate Change Economics and Policy Working Paper*, 2015, *195.*

Eby, M, K Zickfeld, A Montenegro, D Archer, KJ Meissner, and AJ Weaver, "Lifetime of anthropogenic climate change: millennial time scales of potential CO2 and surface temperature perturbations," *Journal of Climate*, 2009, *22* (10), 2501–2511.

Gerlagh, Reyer and Matti Liski, "Consistent climate policies," *Journal of the European Economic Association*, 2017.

Gillett, Nathan P, Vivek K Arora, Damon Matthews, and Myles R Allen, "Constraining the ratio of global warming to cumulative CO2 emissions using CMIP5 simulations," *Journal of Climate*, 2013, *26* (18), 6844–6858.

Golosov, Mikhail, John Hassler, Per Krusell, and Aleh Tsyvinski, "Optimal taxes on fossil fuel in general equilibrium," *Econometrica*, 2014, *82* (1), 41–88.

Goodwin, Philip, Richard G Williams, and Andy Ridgwell, "Sensitivity of climate to cumulative carbon emissions due to compensation of ocean heat and carbon uptake," *Nature Geoscience*, 2015, *8* (1), 29–34.

Held, Isaac M, Michael Winton, Ken Takahashi, Thomas Delworth, Fanrong Zeng, and Geoffrey K Vallis, "Probing the fast and slow components of global warming by returning abruptly to preindustrial forcing," *Journal of Climate*, 2010, *23* (9), 2418–2427.

IPCC, "Working Group I Contribution to the IPCC Fifth Assessment Report: Summary for Policymakers," in "Climate Change 2013: The Physical Science Basis," IPCC, 2013.

Joos, Fortunat, Raphael Roth, JS Fuglestvedt, GP Peters, IG Enting, W von Bloh, V Brovkin, EJ Burke, M Eby, NR Edwards et al., "Carbon dioxide and climate impulse response functions for the computation of greenhouse gas metrics: a multi-model analysis," *Atmospheric Chemistry and Physics*, 2013, *13* (5), 2793–2825.

Lemoine, Derek and Ivan Rudik, "Steering the climate system: using inertia to lower the cost of policy," *American Economic Review*, 2017, *107* (10), 2947–57.

Matthews, H Damon and Ken Caldeira, "Stabilizing climate requires near-zero emissions," *Geophysical Research Letters*, 2008, *35* (4), L04705.

_ , Nathan P Gillett, Peter A Stott, and Kirsten Zickfeld, "The proportionality of global warming to cumulative carbon emissions," *Nature*, 2009, *459* (7248), 829–832.

Millar, Richard J, Zebedee R Nicholls, Pierre Friedlingstein, and Myles R Allen, "A modified impulse-response representation of the global near-surface air temperature and atmospheric concentration response to carbon dioxide emissions," *Atmospheric Chemistry and Physics*, 2017, *17* (11), 7213–7228.

Neij, Lena, "Cost development of future technologies for power generation: a study based on experience curves and complementary bottom-up assessments," *Energy Policy*, 2008, *36* (6), 2200–2211.

Nordhaus, William D, "To slow or not to slow: the economics of the greenhouse effect," *Economic Journal*, 1991, *101* (407), 920–937.

_ , *A Question of Balance: Weighing the Options on Global Warming Policies*, Yale University Press, 2008.

_ and Andrew Moffat, "A survey of global impacts of climate change: replication, survey methods, and a statistical analysis," Technical Report, National Bureau of Economic Research Working Paper 23646 2017.

Pindyck, R. S., "Climate change policy: What do the models tell us?," *Journal of Economic Literature*, 2013, *51* (3), 860–872.

Pizer, William A, "Combining price and quantity controls to mitigate global climate change," *Journal of Public Economics*, 2002, *85* (3), 409–434.

Rezai, Armon and Frederick van der Ploeg, "Intergenerational inequality aversion, growth, and the role of damages: Occam's rule for the global carbon tax," *Journal of the Association of Environmental and Resource Economists*, 2016, *3* (2), 493–522.

Ricke, Katharine L and Ken Caldeira, "Maximum warming occurs about one decade after a carbon dioxide emission," *Environmental Research Letters*, 2014, *9* (12), 124002.

Rogelj, Joeri, David L McCollum, Andy Reisinger, Malte Meinshausen, and Keywan Riahi, "Probabilistic cost estimates for climate change mitigation," *Nature*, 2013, *493* (7430), 79–83.

Shine, Keith P, Jan S Fuglestvedt, Kinfe Hailemariam, and Nicola Stuber, "Alternatives to the global warming potential for comparing climate impacts of emissions of greenhouse gases," *Climatic Change*, 2005, *68* (3), 281–302.

Solomon, Susan, Gian-Kasper Plattner, Reto Knutti, and Pierre Friedlingstein, "Irreversible climate change due to carbon dioxide emissions," *Proceedings of the National Academy of Sciences*, 2009, *106* (6), 1704–1709.

Stern, Nicholas, "The structure of economic modeling of the potential impacts of climate change: grafting gross underestimation of risk onto already narrow science models," *Journal of Economic Literature*, 2013, *51* (3), 838–859.

Traeger, Christian, "Analytic integrated assessment and uncertainty," 2015. UC Berkeley.

United Nations, *World Population Prospects: The 2015 Revision, DVD Edition.*, United Nations, Department of Economic and Social Affairs, Population Division, 2015.

van den Bijgaart, Inge, Reyer Gerlagh, and Matti Liski, "A simple formula for the social cost of carbon," *Journal of Environmental Economics and Management*, 2016, *77*, 75–94.

van der Ploeg, Frederick, "The safe carbon budget," 2017.

Weitzman, Martin L, "Prices vs. quantities," *Review of Economic Studies*, 1974, *41* (4), 477–491.

—, "On modeling and interpreting the economics of catastrophic climate change," *Review of Economics and Statistics*, 2009, *91* (1), 1–19.

—, "GHG targets as insurance against catastrophic climate damages," *Journal of Public Economic Theory*, 2012, *14* (2), 221–244.

Xepapadeas, Anastasios, "Economic growth and the environment," in Karl-Göran Mäler and Jeffrey R. Vincent, eds., *Handbook of Environmental Economics*, Vol. 3, North Holland, 2005, pp. 1219–1271.

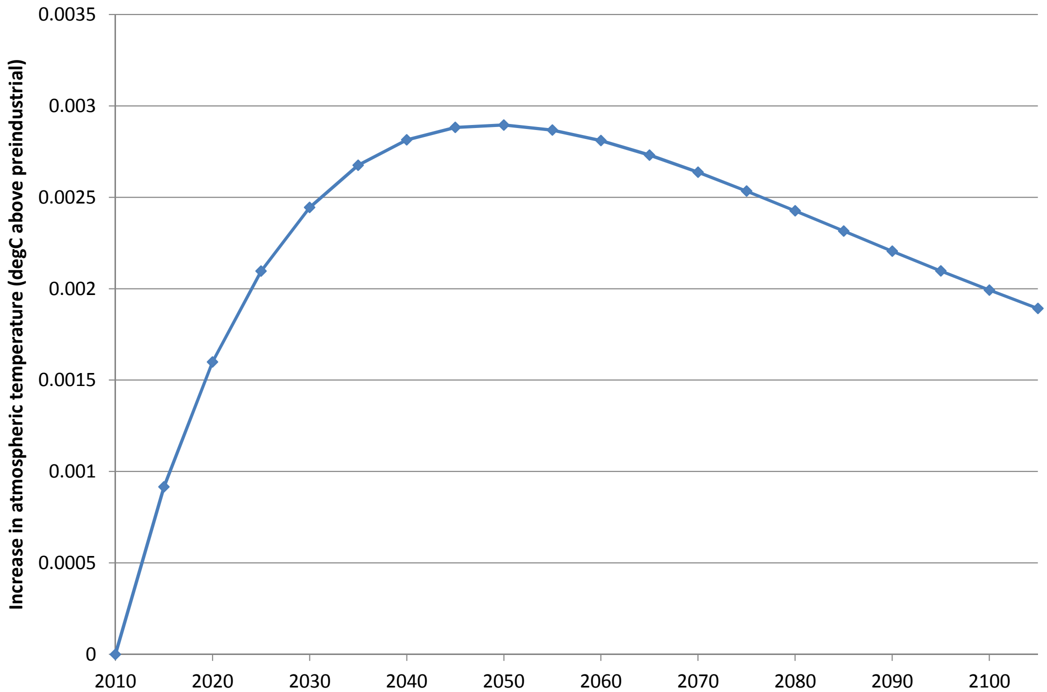
Zickfeld, Kirsten, Michael Eby, H Damon Matthews, and Andrew J Weaver, "Setting cumulative emissions targets to reduce the risk of dangerous climate change," *Proceedings of the National Academy of Sciences*, 2009, *106* (38), 16129–16134.

—, —, Kaitlin Alexander, Andrew J Weaver, Elisabeth Crespin, Thierry Fichefet, Hugues Goosse, Gwenaëlle Philippon-Berthier, Neil R Edwards, Philip B Holden et al., "Long-term climate change commitment and reversibility: An EMIC intercomparison," *Journal of Climate*, 2013, *26* (6), 5782–5809.

A The temperature response to a pulse of CO_2 in DICE

Figure 9 plots warming from an emission of one gigatonne of CO_2 in DICE-2013R in the model base year, 2010. Peak warming from the additional gigatonne occurs 40 years later.

Figure 9: The temperature response to a pulse of CO_2 in DICE



B Equilibrium in a decentralised economy

Competitive firms maximise profit

$$\Pi = F(t, K)e^{(n+g)t+\phi E-\frac{\varphi}{2}E^2-\frac{\gamma}{2}T^2} - wL - \tau E - iK - \delta K$$

taking T and wage payments wL as given. τE are emission tax payments, iK are interest payments on household savings[18] and δK is depreciation of capital. The representative

[18] iK can be thought of as both 'normal' interest and dividend payments, while Π represents extra-ordinary profits, such as resource rents or oligopoly rents, unrelated to the marginal productivity of capital.

household maximises

$$\int_0^\infty e^{-\rho t} u(c) dt,$$

subject to an aggregate budget constraint

$$\dot{K} = iK + wL + \tau E + \Pi - cL = Q - cL - \delta K$$

With $L = e^{nt}$ and $k = K/L$ this boils down to the following household budget constraint:

$$\dot{k} = q - c - \delta k + nk.$$

The government hands back the income from an emissions tax as a lump-sum transfer. Household utility maximisation yields the Ramsey rule $Q_K - \delta = \rho - n + \eta \dot{c}/c$. Profit maximisation ensures that net marginal productivity equals the yield paid on capital $Q_K - \delta = i$. Firms choose emissions that maximise profits:

$$\frac{\partial \Pi}{\partial E} = 0 \Leftrightarrow Q_E = \tau. \tag{33}$$

If the government sets a Pigouvian emissions tax at $\tau = \lambda^S e^{nt} c^\eta$, with λ^S satisfying (11), (12) and (13), the decentralised economy will follow the same emissions path as the social planner's solution.

C The transition to stationary cumulative emissions

Convergence to S^* is dictated by the complementary function

$$y_c \equiv k_1 \exp \frac{1}{2} t \left(a - \sqrt{a^2 + 4b} \right) + k_2 \exp \frac{1}{2} t \left(a + \sqrt{a^2 + 4b} \right).$$

We may assume $b > 0$ and hence the characteristic roots are real.

In order to satisfy the transversality condition on cumulative emissions,

$$\lim_{t \to \infty} e^{(n-\rho)t} \lambda^S = 0,$$

λ^S may not increase at a rate larger than $\rho - n$:

$$\lim_{t \to \infty} e^{(n-\rho)t} \lambda^S = 0 \Longleftrightarrow \lim_{t \to \infty} \frac{-(\rho - n)t + \ln \lambda^S}{t} < 0.$$

Applying l'Hôpital's rule gives

$$\lim_{t\to\infty} -(\rho - n) + \dot{\lambda^S}/\lambda^S < 0.$$

Substituting this with the state equation (12) yields

$$\lim_{t\to\infty} \left[-\epsilon\zeta\lambda^T / \left(Q\left(\phi - \varphi E\right)c^{-\eta}\right)\right] < 0.$$

Since λ^T is always positive, the transversality condition requires the denominator to be positive. Hence the transversality condition is violated if $E > \frac{\phi}{\varphi}$. If $k_2 > 0$, cumulative emissions would be on an explosive increasing path, leading to negative marginal productivity of emissions and violating the transversality condition. Consequently $k_2 = 0$. The initial condition on cumulative emissions S_0 implies $k_1 = S_0 - \frac{c}{b}$, so the transition to S^* is described by

$$S_t = \left(S_0 - \frac{c}{b}\right)\exp\frac{1}{2}t\left(a - \sqrt{a^2+4b}\right) + \frac{c}{b}.$$

D The optimal path in a model without delay

The model without delay has Eqs. (5)-(9) in common, but the climate model and its relationship with damages are now different. Because warming is an instantaneous function of cumulative emissions, it is simply the case that

$$T = \zeta S.$$

Hence we can write damages as a direct function of cumulative emissions,

$$D(S) = \exp\left[-\frac{\gamma}{2}\left(\zeta S\right)^2\right], \tag{34}$$

and dispense with a state variable in the Hamiltonian, which is now just

$$\mathcal{H} = u(c) - \lambda^S E + \lambda^K\left((q-c)e^{nt} - \delta K\right). \tag{35}$$

From the first-order conditions we obtain

$$\dot{\lambda^S} = \dot{u_c}q_E + u_c\dot{q_E} = (\rho - n)u_c q_E + u_c q_S. \tag{36}$$

Since we have an iso-elastic utility function whereby $\dot{u_c}/u_c = -\eta\frac{\dot{c}}{c} \cong -\eta\hat{g}$,

$$\dot{q_E} = (\rho - n + \eta\hat{g})\, q_E + q_S. \tag{37}$$

By integrating (37) and substituting in marginal damages from (34), we see that the MAC must equal the discounted sum of all future damages:

$$q_E = -\int_t^\infty e^{-(\rho-n+\eta\hat{g})(\tau-t)} q_\tau \zeta^2 \gamma S_\tau d\tau,$$

which can be scaled to the size of the economy using $c_\tau = e^{\hat{g}(\tau-t)}c_t$:

$$\frac{q_{E_t}}{q_t} = -\int_t^\infty e^{-(\rho-n+(\eta+1)\hat{g})(\tau-t)} \zeta^2 \gamma S_\tau d\tau. \tag{38}$$

Substituting in the marginal productivity of emissions from (9) we obtain

$$\dot{q}(\phi - \phi E) - q\varphi\dot{E} = (\rho - n + \eta\hat{g})\, q(\phi - \varphi E) - q\zeta^2\gamma S. \tag{39}$$

Rearranging gives:

$$\dot{E} = (\rho - n + (\eta - 1)\hat{g}) \left(E - \frac{\phi}{\varphi}\right) + \frac{\zeta^2\gamma}{\varphi} S, \tag{40}$$

which, after following the same steps as in Section 3, eventually delivers

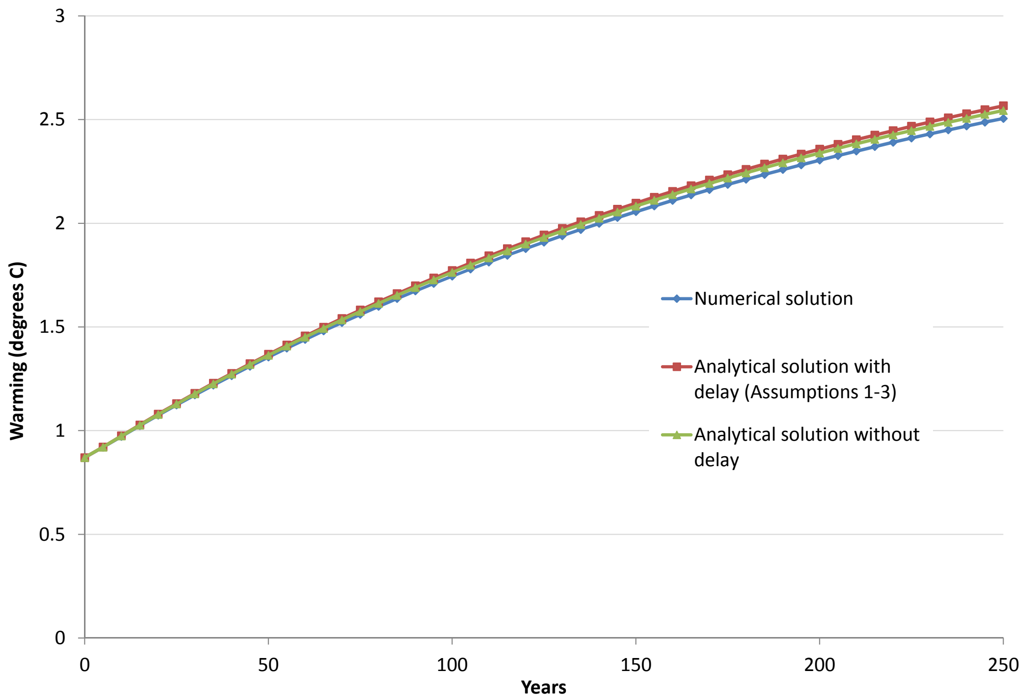
$$T^* = \frac{[\rho - n + (\eta - 1)\,\hat{g}]\,\phi}{\zeta\gamma}. \tag{41}$$

E Model comparison

In this paper we have shown that exact solutions can be obtained for the optimal path of CO_2 emissions and warming in a quite general framework, albeit we have to take one of two shortcuts. Either we take into account the short delay between cumulative emissions and associated warming of the atmosphere, which on the other hand requires making Assumptions 1 and 3, or we ignore the short delay.

Here we compare the performance of these two simplified analytical models with the numerical solution of the 'full' model. The full model comprises discrete-time equivalents of Eqs. (2), (3) and (7)-(9), a five-year time step in the interests of rapid computation, and a finite model horizon, where the terminal period is chosen to be far enough in the future (1000 years) that it does not exert a discernible effect on the optimal path on a decision-relevant timescale (which we take to be 250 years). Optimisation proceeds by choosing $\{E_t\}_{t=0}^{1000}$ so as to maximise $W = \sum_0^\infty u(c_t)(1+\rho-n)^{-t}$, taking $F(t, K) = e^{1.98t}$, i.e. assuming that in the absence of both climate damages and abatement costs, the economy would be on a balanced growth path. As Figure 10 shows, the solutions of the three models are very close. After 50 years, the difference between the solutions is at most 0.01°C (or 1%), while in 100 years' time it is 0.02°C (or 1.6%).

Figure 10: The optimal path of T in the simplified model with an analytical solution and in the full model with a numerical solution.



F Maximising welfare subject to the temperature constraint

We add the inequality constraint that $S \leq \overline{S}$, where $\overline{S} = \zeta\overline{T}$, to the model that has an instantaneous temperature response to emissions. The current value Lagrangian is

$$\mathcal{L} = \frac{1}{1-\eta}c^{1-\eta} - \lambda^S E + \lambda^K \left[Q\left(t, K, T, E\right) - e^{nt}c - \delta K\right] - \theta E. \tag{42}$$

The necessary conditions for a maximum include

$$c^{-\eta} q\left(\phi-\varphi E\right)=\lambda^{S}+\theta, \tag{43}$$

$$\dot{S}=E, \tag{44}$$

$$\dot{\lambda^{S}}=(\rho-n)\lambda^{S}-\gamma\zeta^{2}Sqc^{-\eta}, \tag{45}$$

$$E\geq 0;\theta\geq 0;\theta E=0, \tag{46}$$

$$S\leq\overline{S};\theta(S-\overline{S})=0, \tag{47}$$

$$\dot{\theta}\leq 0\ (=0\ when\ S<\overline{S}). \tag{48}$$

The constrained problem results in a modified differential equation for cumulative emissions:

$$\ddot{S}=a\dot{S}+bS-c+\frac{\rho-n}{\varphi qc^{-\eta}}\theta-\frac{1}{\varphi qc^{-\eta}}\dot{\theta}. \tag{49}$$

The constraint binds if $\overline{S}<c/b$. We define $\bar{t}$ as the time when the constraint binds so that

$$\begin{aligned} t=[0,\bar{t}) &\iff S<\overline{S}\ \&\ \theta=0, \\ t=[\bar{t},\infty] &\iff S=\overline{S};E=0;\theta>0;\dot{\theta}\leq 0. \end{aligned}$$

Note that $E=0$ at $\bar{t}$, because the costate variable θ is required to be continuous. This prevents a discontinuous fall in emissions from taking place at $t=\bar{t}$. Until $t=\bar{t}$, $\theta=0$ and the state equation of the Lagrangian (45) results in the same general solution as (23), but with the added boundary condition that $E_{\bar{t}}=0$, resulting in Proposition 3.

The optimal path of cumulative emissions again derives from the general solution (24) to the differential equation (22). To find k_1, k_2 and $\bar{t}$ we have a system of three boundary conditions. The system has an analytical solution using the following approximation:

$$\begin{aligned} S_t &= (S_0-\frac{c}{b}-k_2)\exp\frac{1}{2}t\left(a-\sqrt{a^2+4b}\right)+k_2\exp\frac{1}{2}t\left(a+\sqrt{a^2+4b}\right)+\frac{c}{b} \\ &\cong (S_0-\frac{c}{b})\exp\frac{1}{2}t\left(a-\sqrt{a^2+4b}\right)+k_2\exp\frac{1}{2}t\left(a+\sqrt{a^2+4b}\right)+\frac{c}{b} \end{aligned}$$

at $t=\bar{t}$. The approximation is based on the insight that, at $t=\bar{t}$, the exponent of the first term is much smaller than unity, while the exponent of the second term is much larger than unity:

$$\begin{aligned}
S_{\bar{t}} = \overline{S} \quad &\Leftrightarrow \quad k_2 = \frac{\overline{S} - \frac{c}{b} + \left(\frac{c}{b} - S_0\right) \exp\left[\frac{1}{2}\bar{t}\left(a - \sqrt{a^2+4b}\right)\right]}{\exp\left[\frac{1}{2}\bar{t}\left(a + \sqrt{a^2+4b}\right)\right]}, \\
E_{\bar{t}} = 0 \quad &\Leftrightarrow \quad k_2 = \frac{\left(\frac{c}{b} - S_0\right)\left(a - \sqrt{a^2+4b}\right) \exp\left[\frac{1}{2}\bar{t}\left(a - \sqrt{a^2+4b}\right)\right]}{\left(a + \sqrt{a^2+4b}\right)\exp\left[\frac{1}{2}\bar{t}\left(a + \sqrt{a^2+4b}\right)\right]}, \\
S(0) = S_0 \quad &\Leftrightarrow \quad S_0 = k_1 + k_2 + \frac{c}{b}.
\end{aligned}$$

Solving this system of equations gives:

$$\bar{t} = \frac{2}{a - \sqrt{a^2+4b}} \ln \frac{\frac{c}{b} - \overline{S}}{\left(\frac{c}{b} - S_0\right)\left(1 - \frac{a-\sqrt{a^2+4b}}{a+\sqrt{a^2+4b}}\right)}, \tag{50}$$

$$k_2 = \left(\frac{c}{b} - S_0\right)\left(\frac{a - \sqrt{a^2+4b}}{a + \sqrt{a^2+4b}}\right)\left(\frac{\frac{c}{b} - \overline{S}}{\left(\frac{c}{b} - S_0\right)\left(1 - \frac{a-\sqrt{a^2+4b}}{a+\sqrt{a^2+4b}}\right)}\right)^{-\frac{2\sqrt{a^2+4b}}{a-\sqrt{a^2-4b}}} \tag{51}$$

When damages are ignored and the problem is to meet the constraint $\overline{S}$ at minimum total discounted abatement cost, Equation (49) becomes

$$\ddot{S} = (\rho - n + (\eta - 1)\hat{g})\left(E - \frac{\phi}{\varphi}\right) = aE - c, \tag{52}$$

integration of which allows us to obtain a general solution for cost-effective emissions:

$$E = \frac{\phi}{\varphi} - \left(\frac{\phi}{\varphi} - E_0\right) e^{(\rho - n + (\eta-1)\hat{g})t} = \frac{c}{a} - \left(\frac{c}{a} - E_0\right) e^{at}. \tag{53}$$

On the cost-effective emissions path $\dot{E_{\text{ce}}} = aE - c$, whereas on the constrained cost-benefit path $\dot{E_{\text{cb}}} = aE + bS - c$. Since bS is positive, the rate of emissions reduction is faster on the cost-effective path. Because both paths must result in the same cumulative emissions, the cost-effective emissions path must begin with higher emissions, but eventually cross the constrained cost-benefit path and reach zero emissions faster (Proposition 5). Note that for a general damage function, the differential equation is $\dot{E_{cb}} = aE + (Q_S/Q)/\varphi - c$. Therefore Proposition 5 holds for any damage function that has positive damages over the whole path.