DEPARTMENT OF ECONOMICS
OxCarre
Oxford Centre for the Analysis of Resource Rich Economies

Manor Road Building, Manor Road, Oxford OX1 3UQ
Tel: +44(0)1865 281281 Fax: +44(0)1865 271094
oxcarre@economics.ox.ac.uk www.oxcarre.ox.ac.uk

UNIVERSITY OF OXFORD

OxCarre Research Paper 122

Abrupt Positive Feedback and the Social Cost of Carbon

Frederick van der Ploeg

OxCarre

Direct tel: +44(0) 1865 281281 E-mail: oxcarre@economics.ox.ac.uk

ABRUPT POSITIVE FEEDBACK AND THE SOCIAL COST OF CARBON

Frederick van der Ploeg*

Abstract

Optimal climate policy should act in a precautionary fashion to deal with tipping points that occur at some future random moment. The optimal carbon tax should include an additional component on top of the conventional present discounted value of marginal global warming damages. This component increases with the sensitivity of the hazard to temperature or the stock of atmospheric carbon. If the hazard of a catastrophe is constant, no correction is needed of the usual Pigouvian tax. The results are applied to a tipping point resulting from an abrupt and irreversible release of greenhouse gases from the ocean floors and surface of the earth, which set in motion a positive feedback loop. Convex enough hazard functions cause overshooting of the carbon tax, but a linear hazard function gives rise to undershooting. A more convex hazard function and a high discount rate speed up adjustment.

Keywords: social cost of carbon, tipping point, positive feedback, climate

JEL codes: D81, H20, Q31, Q38

Revised January 2014

---

* Department of Economics, University of Oxford, Manor Road, Oxford OX1 3UQ, U.K. Also affiliated with VU University Amsterdam. Support from the ERC Advanced Grant 'Political Economy of Green Paradoxes' is gratefully acknowledged.

Email: rick.vanderploeg@economics.ox.ac.uk . Phone: +44-1865-281285.

1. Introduction

The idea that the prime role of climate policy is to deal with the small risk of abrupt and often irreversible climate disasters and tipping points at high temperatures rather than to internalize smooth global warming damages at low and moderate temperatures is gaining traction (e.g., Lenton and Ciscar, 2013; Kopits et al., 2013; Pindyck, 2013). Climate policy must deal with catastrophic events such as destroying a large chunk of productive capacity or unleashing positive feedback loops at higher temperatures. A well-known example is the ice-albedo effect. Global warming may be accelerated with sudden melting of ice sheets (e.g., Greenland), since water and earth reflect less solar radiation than ice and absorb more heat. The warming up causes more ice to melt and sets in motion even more global warming. The positive feedback acts more quickly over the oceans than over land, because sea ice can melt faster than continental ice sheets. Positive feedbacks can also occur with the death of rain forests as plants have a lower reflectivity than bare soil and there will be less transpiration. A final example is the Clathrate gun hypothesis, which states that a rise in sea temperatures and/or a rise in sea levels can trigger the sudden release of methane from methane clathrate compounds buried in sea-beds and permafrost (e.g., from the tundra in the Arctic, mostly Eastern Siberia). Since methane is itself a powerful (albeit shorter lived) greenhouse gas, this methane release will increase global warming and set in motion further methane clathrate destabilization. This type of positive feedback might trigger a runaway process, which in the long run is stabilized via the natural decay of the stock of atmospheric carbon. There is a lot of debate about whether these positive feedback effects will occur at higher temperatures and also about the magnitude of such effects. We have nothing to contribute to this debate. Our objective is to study a regime switch caused by a temperature-dependent risk of an irreversible, sudden release of greenhouse gases and show how such a risk of a regime switch resulting in a disastrous change in the climate dynamics affects the social cost of carbon and the optimal carbon tax. We analyze what to do when burning more fossil fuel leads to a higher stock of carbon in the atmosphere and higher global temperature and thus to a higher risk of a climate calamity.

Our contribution is to show that with temperature-dependent risks of a climate calamity the optimal carbon tax and the social cost of carbon have to be higher than indicated by the normal Pigouvian formula for the social cost of carbon (i.e., the present discounted value of marginal global warming damages resulting from burning an additional unit of fossil fuel). Our contribution is also to do this in a tractable partial equilibrium model of climate policy with tipping points generating regime switches and changes in the carbon cycle and the system dynamics at some random future moment of time. We thus investigate how the carbon tax should respond to a sudden unleashing of positive feedback loops at higher global mean temperatures and change in system dynamics (cf., Naevdal, 2006)[1]. The expected time it takes to unleash such positive feedback loops decreases with temperature and the accumulated carbon stock, which results from a hazard function which increases in the carbon stock.

We show how the optimal social cost of carbon associated with burning fossil fuel can be decomposed into three components. The first component arises from the usual need to correct for marginal global warming damages which internalizes the adverse effects on all future global warming damages arising from burning an additional unit of fossil fuel. This requires that this component of the social cost of carbon is set to the present value of all future marginal global warming damages where the relevant discount rate is the social rate of discount augmented by the hazard of a tipping point as well as the rate of atmospheric decay. The hazard makes society more impatient, so that the marginal global warming damages are discounted more heavily which depresses this conventional expression for the Pigouvian social cost of carbon. The second component of the social cost of carbon arises because burning an additional unit of fossil fuel increases the stock of atmospheric carbon and thus curbs the welfare after the tipping point. This raises the cost of a tipping point and thus requires a boost to the social cost of carbon before the disaster strikes. This is called the 'raising the stakes' effect. The third component of the social cost of carbon arises because burning an additional unit of fossil fuel increases global mean temperature and thus increases the risk of a tipping point and a discrete catastrophic loss in value. This component is also positive. The resulting boost to the social cost of carbon curbs fossil fuel use and the risk of a tipping point. This is referred to as the 'averting risk' effect. We illustrate our results with some calibrated simulations. This permits us to also investigate how the optimal carbon tax and social cost of carbon depend on the shape and especially the convexity of the hazard function, since it is plausible that the risk of a tipping point increases relatively more at higher temperatures.

Our analysis of the effects of the catastrophic unleashing of positive feedback loops and what to do about differs from earlier literature in five respects. First, we differ from Nordhaus (2008) and Golosov et al. (2013) who deal with a moderate cost of global warming at low temperatures and a catastrophic cost of global warming at higher temperatures but combine the two in an expected damage parameter without dealing with the analysis of regime switches and tipping points. Second, our analysis deals with the risk of extreme global warming increasing the risk of climate calamity and pushing up the carbon tax but differs from Weitzman (1998) who argues the case for an ambitious climate policy using fat-tailed risk of global warming. Our analysis also differs from Gollier (2008, section 3.2; 2012, Chapter 5) who uses a

[1] This is related to recent work on discrete thresholds which result in a regime of much less assimilative capacity of carbon once the stock of carbon in the atmosphere crosses a given threshold (Amigues and Moreaux, 2012) and on a an economy with exhaustible resources and a regime switch entailing a total destruction of assimilative capacity with zero decay of atmospheric carbon (Prieur et al, 2013).

Markov 2-regime switching model to show that an exogenous small risk of a very large drop in the growth rate of GDP depresses the efficient discount rate for long maturities. Since this boosts the social cost of carbon, this also makes a case for a more aggressive climate policy. Third, we differ from work on discrete thresholds for the stock of carbon in the atmosphere or global mean temperature which once passed result in a regime of much less assimilative capacity of carbon (Amigues and Moreaux, 2012; Prieur et al., 2013).[2] Although this work deals with positive feedback and changes in capacity to assimilate carbon, it takes the threshold for the catastrophe as given. In contrast, we have an uncertain threshold as we have the risk of tipping increasing with the carbon stock or global mean temperature. Fourth, our catastrophe directly affects the intrinsic dynamics of the carbon cycle whilst most earlier work deals with a catastrophic shock to global warming damages[3]. We also differ from Lemoine and Traeger (2013) who consider a catastrophic shock to the equilibrium climate sensitivity.[4] Finally, as van der Ploeg and de Zeeuw (2013), we offer a decomposition of the optimal carbon tax in the face of impending catastrophe[5] and highlight how the shape of the hazard function affects the optimal carbon tax and the before-catastrophe dynamics of fossil fuel use and the stock of atmospheric carbon.

A limitation of our framework is that it is partial equilibrium. Recently, general equilibrium studies of growth and climate change have analyzed climate tipping in extended versions of the DICE model and show that the threat of a tipping point which increases with global mean temperature induces significant and immediate increases in the social cost of carbon and thus the optimal carbon tax (Lemoine and Traeger, 2012; Cai et al., 2012; van der Ploeg and de Zeeuw, 2013).[6] These studies do not consider catastrophic shocks to the intrinsic dynamics of the carbon cycle and deal with catastrophic shocks to total factor productivity and/or the equilibrium climate sensitivity instead. These studies do allow for general equilibrium with an endogenous interest rate and endogenous growth and development. Furthermore, as

---

[2] Prieur et al. (2013) deal with an economy with exhaustible resources and a regime switch entailing a total destruction of assimilative capacity with zero decay of atmospheric carbon.

[3] For example, the loss of the Greenland Ice Sheet or the Antarctic Ice Sheet might lead to sea levels rising by 7 and 3 meters, respectively, and the resulting damages can take millennia and may already be occurring. A more imminent example is the collapse of the Atlantic thermohaline circulation and is likely to already occur at relatively moderate degrees of global warming. This will have different effects on damages from global warming (viz. the stopping or reversal of the Gulf Stream will mainly hurt northern Europe).

[4] They model the release of methane from melting permafrost as an instantaneous doubling of the equilibrium climate sensitivity, but it is not clear that this makes sense from a climate science point of view. We prefer to model this as a sudden unleashing of a positive feedback loop in the carbon cycle.

[5] However, van der Ploeg and de Zeeuw (2013) do not focus on the 'raising the stakes' term in the decomposition of the optimal carbon tax, because their catastrophic shock affects total factor productivity and the stock of carbon or global mean temperature does not affect the post-calamity value function.

[6] These follow earlier partial equilibrium analytical studies of climate catastrophes (e.g., Tsur and Zemel, 1996, 2008, 2009; Naevdal, 2006; Karp and Tsur, 2011; Tsur and Withagen, 2013) and on sudden collapses of the resource stock and changes in the system dynamics (regime switches) in resource management and pollution control (e.g., Cropper, 1976; Heal, 1984; Clarke and Reed, 1994; Polasky et al., 2011; de Zeeuw and Zemel, 2012).

an earlier paper has argued, this leads to a novel argument for precautionary capital accumulation over and above the carbon tax (van der Ploeg and de Zeeuw, 2013).

Section 2 sets up our model of a regime switch corresponding to a change in the dynamic properties of the carbon cycle resulting from a sudden, irreversible unleashing of positive feedback which is more likely to occur at higher temperatures. Section 3 derives the optimal climate policies in the presence of abrupt and irreversible unlocking of positive feedback loops in the carbon cycle which result from a tipping point at some future moment in time. Section 4 offers an illustrative calibration of our model paying attention to both linear and convex hazard functions. Section 5 presents indicative calculations of how the optimal carbon should be adjusted upwards given an endogenous risk for the sudden release of greenhouse gases and investigates the sensitivity of the social cost of carbon to the shape of the hazard function for such a climate catastrophe. Section 6 concludes.

2. Regime switches and the economics of climate change

Climate policy is usually derived in the face of smooth global warming damages. To study how the optimal climate policy is affected by non-smooth climate catastrophes, we will contrast and analyze both the usual smooth damages and a climate catastrophe leading to a reduction in the rate of decay of atmospheric carbon. We thus focus at the best way to deal with a shock corresponding to the sudden release of greenhouse gases resulting from the emergence of positive feedback which is more likely to occur at higher temperatures or stocks of carbon in the atmosphere.[7]

To keep matters simple, to focus on the economics of regime switches and tipping points and to highlight the difference between these two approaches to the design of optimal climate policy, we will adopt a partial equilibrium perspective so we abstract from capital accumulation, the interest rate, and growth and development. We will also assume athat fossil fuel reserves are abundant with an infinitely elastic supply.

Private utility is quasi-linear, so consists of the consumption of all goods other than fossil fuel, i.e., aggregate global income Y minus the cost of fossil fuel extraction $d\,E$ with d the constant cost of fossil fuel extraction, plus the utility of fossil fuel consumption E (measured in tons of carbon). Being a partial equilibrium model aggregate global income is exogenous. The utility function $U(E)$ is assumed to be concave. Utility is the area under the demand curve for fossil fuel, so that the price of fossil fuel must

[7] Temperature measured as the difference from pre-industrial temperature is often modeled as $\Xi \ln(P / P_I) / \ln(2)$, where P and P_I are the prevailing and pre-industrial stocks of atmospheric carbon, respectively, and Ξ is the equilibrium climate sensitivity (with a typical estimate 3) measuring the temperature increase caused by a doubling of the carbon stock. We thus use the effect of temperature and the carbon stock on the hazard interchangeably.

equal $p = U'(E, \tilde{\Omega})$. The quasi-linearity of preferences implies that all these terms (in the integrand of (1) below) are measured in units of the other consumption goods (e.g., 2010 US dollars).

Social welfare is private utility minus the damage of global warming $D(P)$, where P is the stock of carbon in the atmosphere. The damage function is supposed to be weakly convex. The stock of atmospheric carbon is the net effect of the accumulation of all past carbon emissions which result from burning fossil fuel and the natural rate of decay of atmospheric carbon, where the parameter $\delta > 0$ indicates the natural decay rate which prevails before the catastrophe (about 1/250). The social planner maximizes the expected present discounted value of social welfare subject to the dynamics of carbon accumulation in the atmosphere and the hazard of the climate catastrophe occurring, where the social rate of discount is denoted by $\rho > 0$. The social planner thus solves the following problem:

$$\text{(1)} \qquad \underset{E}{\text{Max}}\ \mathrm{E}\left[\int_0^\infty \left\{U\left((E(t)\right) + Y - dE(t) - D\left(P(t)\right)\right\} e^{-\rho t} dt\right]$$

subject to the dynamics describing the stock of atmospheric carbon,

$$\text{(2a)} \qquad \dot{P} = E + \tilde{\Theta}P - \delta P, \quad P(0) = P_0,$$

and the specification of the climate catastrophe,

$$\text{(2b)} \qquad \tilde{\Theta}(0) = 0, \quad \dot{\tilde{\Theta}} = 0, \quad \tilde{\Theta}(T^+) - \tilde{\Theta}(T^-) = \Theta,$$

$$\text{(2c)} \qquad \Pr\left[T < t\right] = 1 - \exp\left(-\int_0^t H\left(P(s)\right) ds\right), \quad \forall t \geq 0,$$

where Θ is the size of the climate catastrophe and T the date at which the climate catastrophe strikes (a stochastic variable). Equation (2b) indicates what happens if the catastrophe strikes at time T. The climate catastrophe corresponds to the sudden unleashing of a positive feedback loop which causes ongoing extra release of carbon emissions from the ocean floors into the atmosphere, $\Theta\, P$, where $\Theta > 0$ is a constant. The amount of carbon released once the positive feedback loop is unleashed thus increases with temperature or the total atmospheric carbon stock, P. This unleashed feedback loop amounts to a regime switch which causes a sudden and irreversible reduction in the rate of atmospheric decay from δ to $\delta - \Theta$.

Equation (2c) gives the probability of the catastrophes occurring before time t. We model the probability of the catastrophe occurring in an infinitesimally small interval of time starting at time t with the hazard rate $h(t) = H\left(P(t)\right) = \lim_{\Delta t \to 0} \Pr\left[T \in (t, t+\Delta t) \mid T \notin (0,t)\right] / \Delta t$. This hazard rates corresponds to the conditional density in the sense that $h(t)\Delta t$ is the probability of a disaster given that no climate disaster has taken place in the interval before time t. The unconditional probability that the event occurs between t and $t+\Delta t$

can be approximated by $f(t)\Delta t$, where $f(t)$ is the probability density function. A constant hazard rate h implies that the probability density function for the date T of climate change is given by $f(t) = he^{-ht}$ and the corresponding cumulative density function is given by $\Pr[T < t] = 1 - e^{-ht}$ and gives the (unconditional) probability of the catastrophe occurring before time t. The expected date for the climate calamity to occur is $1/h$, so that the probability of "survival" is given by e^{-ht}. If the hazard rate h is not constant, ht in the expressions above has to be replaced by the term $\int_0^t h(s)ds$. Equation (2c) implies that the passing of time t gradually raises the probability that a climate disaster has occurred at some time T before that, especially if the hazard rates are high and rising with the stock of atmospheric carbon and temperature.

Although we will consider the implications of a constant hazard rate of a climate catastrophe, our main focus is on a stock-dependent hazard rate $h(t) = H(P(t))$ to capture the effect that a higher stock of carbon in the atmosphere increases global temperature and thus increases the probability of a regime switch resulting in climate change (i.e., $H'(P) > 0$). Hence, if the stock of atmospheric carbon increases over time, the expected duration before the catastrophe occurs, $1/H(P)$, decreases over time. This implies that a failing climate policy makes catastrophe more imminent. The stock-dependent hazard function $H(P)$ is increases in the carbon stock at a non-decreasing rate and is thus weakly convex. Furthermore, we suppose that positive feedback will lead to 'runaway' global warming in the short and medium run but will eventually be checked by the process of natural decay of the stock of atmospheric carbon.

The sufficient conditions for this problem to have a solution are summarized in the following assumption.[8]

Assumption 1: $U' > 0, U'' < 0, D' > 0, D'' \geq 0, H > 0, H' \geq 0, H'' \geq 0,\ \delta > 0,\ 0 < \Theta < \delta.$

We are solving (1)-(2) as a social planner problem but the social optimum can be realized in a decentralized market economy provided the climate externalities resulting from smooth global warming damages and from impending catastrophes are internalized via an appropriate carbon tax. Hence, we ignore imperfect competition in fossil fuel markets and other externalities. We also suppose that the carbon tax revenues are rebated in lump-sum fashion and abstract from other distorting taxes. In that case, the social optimum is replicated if the carbon tax is set to the optimal social cost of carbon.

One can extend the analysis to allow for both a transient and a permanent component of the stock of atmospheric carbon, as has been put forward by Golosov et al. (2013) in their reduced-form model of the carbon cycle used by Nordhaus (2008), for the stock of carbon which is absorbed by the upper and lower

---

[8] Strictly speaking, one might need $H'(P_0) > 0$ to avoid that the optimum features a corner solution. However, for the calibration and functional forms considered below this is not an issue.

layers of the oceans and the surface of the earth and for a finite stock of fossil fuel reserves. The sum of the stocks of carbon in situ, the atmosphere and the oceans must be constant. So far, we have allowed for positive feedback leading to the release of greenhouse gases from the oceans and surface of the earth into the atmosphere. In addition, one could also allow for different types of climate catastrophes where the hazard rate functions for each of these catastrophes may differ. For example, one could allow for an abrupt and irreversible shock to the utility obtained from using fossil fuel, which may arise when a disaster (e.g., a hurricane) hits productive capacity, pushes up the price of fossil fuel and thus reduces utility. Alternatively, one might consider the effects of an abrupt and irreversible shock to global warming damages, which might arise with a climate disaster which destroys natural habits (e.g., a sudden acidification of oceans destroying coral reefs or the sudden loss of rainforests).[9] However, we focus on the simpler problem (1)-(2) with only one type of climate catastrophe and a simplified carbon cycle, since it highlights how climate policy is affected by regime switches without making the analysis to cumbersome.

3. Optimal climate policy with positive feedback loops

To find the optimal climate policy, we solve the problem (1)-(2) by backward recursion, so we first solve for the post-catastrophe regime and then for the pre-catastrophe regime.

3.1. Post-catastrophe regime

The Hamiltonian function for the post-catastrophe outcome is defined as follows:

(3) $$H \equiv U(E) + Y - dE - D(P) - \lambda[E + \Theta P - P^T], \quad t \geq T,$$

where λ denotes the shadow cost of the stock of atmospheric carbon, respectively. We thus have the following optimality conditions for $t \geq T$:

(4a) $$U'(E) = d + \lambda,$$

(4b) $$\dot{\lambda} = \left[\rho + \delta - \Theta\right]\lambda - D'(P), \quad \lim_{t\to\infty} \lambda(t)P(t)e^{-\rho t} = 0.$$

To replicate the social optimum in a decentralized market economy a price has to be charged for carbon which has to correspond to the social cost of carbon. This can be a specific carbon tax or the price of an emission permit. For ease of the discussion, we will refer to it as the carbon tax and will use it

[9] These first two types of climate calamities are discussed in van der Ploeg and de Zeeuw (2013) within the context of a general equilibrium Ramsey growth model with energy as a factor of production.

interchangeably with the social cost of carbon. The carbon tax has to be set to $\tau = \lambda$. Revenues from the carbon tax are handed back in lump-sum fashion.

Optimality condition (4a) can be used to give energy demand as decreasing function of the carbon tax:

$$(5) \qquad U'(E) = d + \tau \quad \Rightarrow \quad E = E(\tau), \quad E' = 1/U'' < 0.$$

We thus have the following after-calamity saddle-point system of differential equations:

$$(6a) \qquad \dot{P} = \left[E(\tau) + \Theta P)\right] - \delta P, \quad P(T) \text{ given,}$$

$$(6b) \qquad \dot{\tau} = (\rho + \delta - \Theta)\tau - D'(P),$$

where equation (6b) follows from the optimality condition (4b).

Immediately after the climate catastrophe has taken place, the carbon tax must jump to place the system on its stable after-calamity manifold:

$$(7) \qquad \tau = \tau^A(P), \quad t \geq T.$$

The Bellman equation for the after-catastrophe problem is:

$$(8) \qquad \rho V^A(P) = \underset{E}{\text{Max}} \left[U(E) + Y - dE - D(P) + V^{A'}(P)\left(E + \Theta P - \delta P\right) \right],$$

where $V^A(P)$ is the after-catastrophe value function and $\lambda = -V^{A'}(P)$. It follows from (8) and equations (5) and (6) that the after-catastrophe value function can be calculated from:

$$(9) \qquad V^A(P,\Theta) = \frac{\bar{U}(\tau) - D(P) + \tau^A(P)(\delta - \Theta)P}{\rho}, \qquad V_P^A = -\tau^A(P) < 0, \quad V_\Theta^A = -P\tau^A(P)/\rho < 0,$$

where utility of both fossil fuel and other consumption (excluding the carbon tax rebates) is defined as:

$$(10) \qquad \bar{U}(\tau) \equiv U\left(E(\tau)\right) + Y - (d + \tau)E(\tau), \quad \bar{U}' = -E < 0.$$

From (5) we see that marginal effect of an increase in the carbon tax on total utility is negative, namely the negative of fossil fuel consumption. The after-catastrophe value function (9) decreases in the stock of atmospheric carbon and the size of the disaster. Of course, the marginal cost of the carbon stock is the after-catastrophe carbon tax. The after-disaster value is also negatively affected by the catastrophe Θ resulting in extra release of carbon emissions from the ocean floors caused by positive feedback.

To make it easier to highlight the effects of regime switches, we let global warming damages be linear which is what is effectively assumed in the approximation to the Nordhaus (2008) damages used by Golosov et al. (2013) in their calibration of the optimal carbon tax.

Assumption 2: $D(P) = \psi P, \psi > 0.$

In that case, the after-catastrophe optimal carbon tax is constant (from equation (6b)) and thus energy use is constant too in the post-catastrophe regime:

$$(11) \qquad \tau(t) = \frac{\psi}{\rho + \delta - \Theta} \equiv \tau^A > \tau^N = \frac{\psi}{\rho + \delta}, \quad E(t) = E(\tau^A), \quad t \geq T.$$

Hence, with linear global warming damages the optimal carbon tax is independent of the prevailing stock of carbon and temperature. This after-catastrophe carbon tax exceeds the naïve no-shock Pigouvian carbon tax which does not take account of the unleashed positive feedback and associated change in system dynamics, denoted by τ^N. The optimal carbon tax thus corrects for the positive feedback in the carbon cycle. It equals the ratio of the marginal global warming damage divided by the discount rate *plus* the decay rate *minus* the rate of carbon accumulation resulting from positive feedback. Positive feedback thus pushes up the after-catastrophe Pigouvian carbon tax.

3.2. Pre-catastrophe regime

The Hamilton-Jacobi--Bellman equation for the pre-catastrophe regime given assumptions 1 and 2 is:

$$(12) \qquad \rho V^B(P,\Theta) = \underset{E}{\text{Max}} \left[U(E) + Y - dE - \psi P + V_P^B(P,\Theta)(E - \delta P) - H(P)\left\{ V^B(P,\Theta) - V^A(P,\Theta) \right\} \right].$$

The final term on the right-hand side of (12) is the expected loss in value at the time that the catastrophe strikes. Through this term the before-catastrophe value function depends on the size of the catastrophe, Θ.

The optimality condition for (12) is, as before, equation (5) with now the carbon tax given by $\tau = -V_P^B(P,\Theta)$ and energy use by $E = E(\tau)$. Using this, we get the before-catastrophe value function:

$$(13) \qquad V^B(P,\Theta) = \frac{\bar{U}(\tau) - \psi P + \tau\delta P + H(P)V^A(P,\Theta)}{\rho + H(P)}.$$

Differentiating equation (12) or (13) with respect to time and using equations (9) and (10) yields:

$$(14) \qquad \left[(\rho + H)V_P^B + H'V^B \right]\dot{P} = -E\dot{\tau} + (\delta\tau - \psi)\dot{P} + \delta P\dot{\tau} + (H'V^A - H\tau^A)\dot{P}.$$

Simplifying equation (14) and using $\dot{P} = E - \delta P$, we get:

$$(14') \qquad \left[\dot{\tau} + \psi + H\tau^A + H'(V^B - V^A) - (\rho + \delta + H)\tau \right]\dot{P} = 0.$$

Since this equation must hold for any $\dot{P} \neq 0$, we get the following dynamics of the social cost of carbon:

$$(15) \qquad \dot{\tau} = \left[\rho + \delta + H(P)\right]\tau - \psi - H(P)\tau^A - H'(P)\left[V^B(P,\Theta) - V^A(P+\Theta/\rho)\right], \quad 0 \le t < T.$$

Integration of (15) yields the expression for the pre-disaster optimal social cost of carbon:[10]

$$(16) \qquad \tau(t) = \underbrace{\psi\int_t^\infty e^{-\int_t^s [\rho+\delta+H(P(s'))]ds} ds}_{\text{modified conventional no-shock social cost of carbon}} + \underbrace{\int_t^\infty H\left(P(s)\right)\tau^A e^{-\int_t^s [\rho+\delta+H(P(s'))]ds} ds}_{\text{'raising the stakes' effect}}$$
$$+ \underbrace{\int_t^\infty H'\left(P(s)\right)\left\{V^B\left(P(s),\Theta\right) - V^A\left(P(s),\Theta\right)\right\} e^{-\int_t^s [\rho+\delta+H(P(s'))]ds} ds}_{\text{'risk averting' effect}}, \quad 0 \le t < T.$$

Expression (16) shows that our expression for the before-disaster carbon tax reduces to the conventional Pigouvian carbon tax, $\tau(t) = \psi/(\rho+\delta), 0 \le t < T,$ if there is no hazard of a climate catastrophe ($H(P) = 0$).

With a constant hazard rate, say $H(P) = h_0 > 0$, the third 'risk-averting' term in (16) drops out and it is easy to demonstrate that the carbon tax and fossil fuel use are constant[11]:

$$(16') \qquad \tau^N < \tau(t) = \left(\frac{\rho+\delta+h_0-\Theta}{\rho+\delta+h_0}\right)\tau^A < \tau^A, \quad E(t) = E(d+\tau^A), \forall t \ge 0.$$

With a constant hazard rate the before-calamity carbon tax is thus bigger than the no-shock Pigouvian carbon tax, so that the no-shock carbon tax term is dominated by the 'raising the stakes' term in (16). However, the before-catastrophe carbon tax is set below the after-catastrophe carbon tax. Hence, fossil fuel use with a constant hazard is higher before than after the climate calamity.

In general, the hazard rate increases with the carbon stock and global mean temperature in which case the following components of optimal climate policy can be highlighted:

1. The first term in (16) is the conventional present value of marginal global warming damages, but is reduced by the '*making hay while the sun shines*' effect as the hazard rate has to be added to the sum of the discount rate and the atmospheric decay rate. The hazard of a catastrophe and the belief that good times eventually come to an end makes society more impatient, so marginal global warming damages are discounted more heavily which depresses the social cost of carbon below the expression given in (11′). In fact, this 'making hay' effect also depresses the 'raising the stakes' and 'averting risk' effects in (16), and is stronger if the risk of catastrophe is higher.

[10] A similar decomposition of the social cost of carbon for a general model of growth and climate change can be found in an earlier study (van der Ploeg and de Zeeuw, 2013).

[11] With a constant hazard rate (16) gives $\tau(t) = \psi(1+h_0\tau^A)/(\rho+\delta+h_0), 0 \le t < T.$ Upon substitution of τ^A from (11′), we get expression for $\tau(t)$ in (16′) which is clearly less than τ^A given that $\Theta > 0$. It is also easy to establish that $\tau(t) > \tau^N$ provided h_0 is strictly positive.

2. The '*raising the stakes*' effect in (16) is a consequence of burning more fossil fuel increasing the stock of atmospheric carbon and thereby increasing marginal global warming damages. Since a marginal increase in the carbon stock curbs the value after the climate calamity (with the size of the effect equal to the after-calamity carbon tax), this raises the stakes (the drop in welfare following a catastrophe) and thus boosts the social cost of carbon before the calamity strikes.
3. The '*risk averting*' effect in (16) arises from taking account of the fact that burning fossil fuel leads to a higher stock of atmospheric carbon and higher global mean temperature. This increases the risk of climate catastrophe and the chance of a discrete *non-marginal* catastrophic loss in value resulting from the eventual regime switch.

The expression for the optimal social cost of carbon is, in general, difficult to evaluate, since one needs to know the whole time trajectory of the stock of atmospheric carbon. To do this, one needs to solve equation (15) together with the pre-catastrophe dynamics of the atmospheric carbon stock,

$$\text{(17)} \qquad \dot{P} = E(\tau) - \delta P, \quad P(0) = P_0, \quad 0 \le t < T,$$

as a saddle-path system to give the stable manifold:

$$\text{(18)} \qquad \tau = \tau^B(P,\Theta), \quad 0 \le t < T.$$

The condition for saddle-path stability of the system of ordinary equations (15) and (17) is that the determinant of the Jacobian matrix J, say $\det(J)$, has to be negative. This is the case if:

$$\text{(19)} \qquad \det(J) \equiv -\left[\rho + \delta + H(P)\right]\delta - 2E'(\tau)H'(P)(\tau - \tau^A) + E'(\tau)H''(P)\left[V^B(P,\Theta) - V^A(P,\Theta)\right] < 0.$$

The first term in (19) is always negative, the second term is negative given that the pre-catastrophe carbon tax is less than the after-catastrophe carbon tax, and the third term is positive for convex hazard functions given that welfare is higher before the catastrophe strikes than afterwards. Hence, we assume that the saddle-point stability condition (19) is satisfied which will be the case if the hazard function is linear or not too convex.

Upon substitution of (18) into (17), we can simulate forwards in time. By substituting (18) into (13) we get the value function for the pre-catastrophe system. The following proposition gives the steady state before-catastrophe social cost of carbon (denoted by an asterisk) with a stock-dependent hazard rate. The local dynamic properties are given in proposition 2.

Proposition 1: The steady-state pre-catastrophe carbon tax and carbon stock follow from:

(20a) $$\tau^{B*} = \left[\frac{\rho+\delta+H(P^{B*})-\Theta}{\rho+\delta+H(P^{B*})}\right]\tau^A + \frac{H'(P^{B*})\left[V^B(P^{B*},\Theta)-V^A(P^{B*},\Theta)\right]}{\rho+\delta+H(P^{B*})} > \tau^N,$$

(20b) $$P^{B*} = E(\tau^{B*})/\delta.$$

A bigger positive feedback (higher Θ) curbs the carbon tax with a constant hazard (he first term in (20a)), but increases the risk-averting component of the carbon tax (the second term in (20a)).

Proof: see appendix.

The sensitivity of the hazard rate to the carbon stock or global mean temperature thus implies that the steady-state carbon tax must be set even higher than the no-shock carbon tax, τ^N (given in (11′)), and the carbon tax corresponding to a constant hazard rate is set to the steady-state hazard rate. If the hazard increases in the carbon stock, $H(P)$ with $H' > 0$, the climate tipping point becomes more imminent as the atmosphere continues to accumulate carbon and global mean temperature rises. By setting a higher carbon tax, the economy steers away from the risk of a disaster by reducing the hazard of a sudden release of carbon from the ocean floors occurring setting in motion a process of positive feedback. Hence, the steady-state optimal carbon tax is pushed up by the 'risk averting' effect above the carbon tax suggested by the conventional Pigouvian carbon to deal with smooth damages and the 'making hay' effect, corrected for the 'making hay' effect (i.e., the first term in equation (20a)).

We can solve equations (15) and (17) for the *pre-catastrophe* transient time paths of the carbon tax and the stock of atmospheric carbon that are relevant so long as the catastrophe has not struck yet. Two types of pre-catastrophe dynamic response are feasible depending on whether the parameter Γ, which is defined by $\Gamma \equiv 2H'(P^{B*})(\tau^{B*}-\tau^P) - H''(P^{B*})(V^B - V^A)$, is positive or negative.[12] The first type of dynamic behaviour corresponds to undershooting of the social cost of carbon and occurs if $\Gamma > 0$. It is illustrated in the phase diagram portrayed in fig. 1(a), where the red line is the locus of points for which $\dot{P} = 0$ holds and the green line is the locus of points for which $\dot{\tau} = 0$ holds. We then have that, upon the realization of an impending catastrophe, the carbon tax jumps up and *undershoots* its pre-catastrophe steady-state value and subsequently rises over time as it travels along the saddle-path in north-westerly direction.

Fig. 1(b) shows the phase diagram for the case $\Gamma < 0$ in which case the carbon tax *overshoots* its pre-catastrophe steady-state value before falling over time as it travels along the saddle-path in south-westerly direction. The undershooting case occurs for linear (in which case $\Gamma \equiv 2H'(P^{B*})(\tau^{B*}-\tau^P)$ is always

[12] The saddle-path stability condition (19) implies that Γ cannot be too large, i.e., $\Gamma < -0.5\left[\rho+\delta+H(P)\right]\delta/E'(\tau)$.

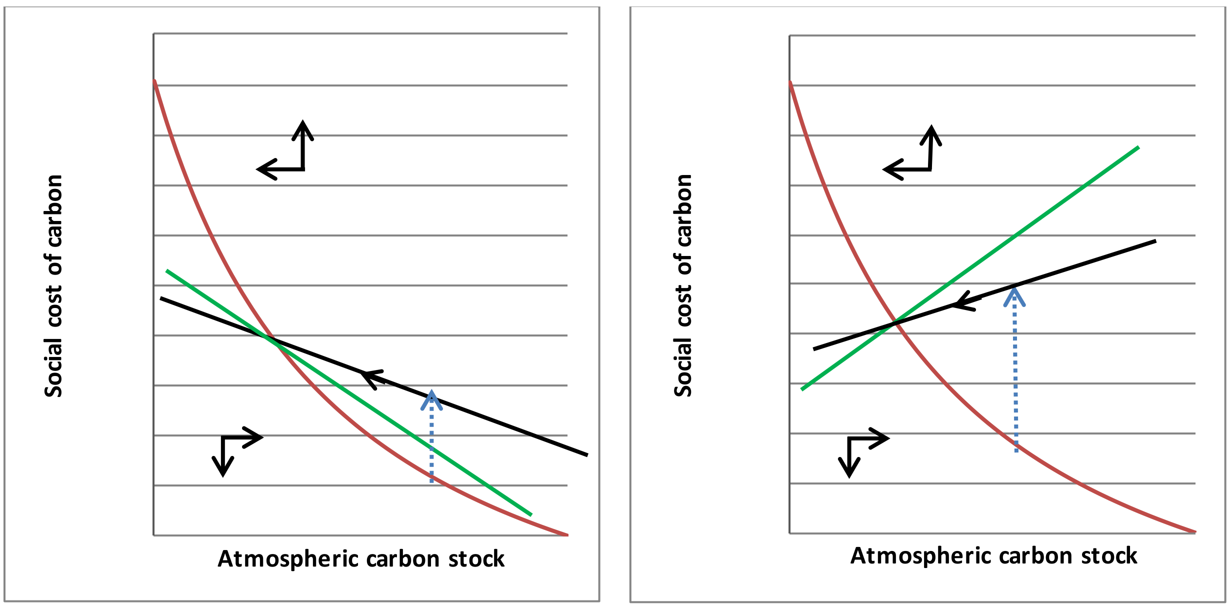
positive) and not too convex hazard functions, but overshooting of the social cost of carbon occurs for convex enough hazard functions.

As soon as the catastrophe strikes, the social cost of carbon jumps from wherever the social cost of carbon is to its post-catastrophe value $\tau^A > \tau^N$ as given in (11) (see the illustration reported in fig. 3 below).

We offer an approximation of these transient time paths based on log-linearization of the saddle-path system defined by (15) and (17) given the pre-catastrophe steady state characterized in proposition 2.

Figure 1: Pre-catastrophe response of the social cost of carbon to an impending climate calamity

(a) Undershooting, Γ > 0 **(b) Overshooting, Γ < 0**



Proposition 2: Suppose condition (19) is satisfied. First-order approximations of the before-catastrophe time paths for the optimal carbon tax and the stock of carbon in the atmosphere are then given by:

$$\text{(21)} \qquad \tau(t) \cong \tau^{B*}\left(\frac{P_0}{P^{B*}}\right)^{\omega e^{-\mu t}}, \quad \omega \equiv \frac{(\delta-\mu)\tau^{B*}}{E'(\tau^{B^*})P^{B*}},$$

$$\text{(22)} \qquad P(t) \cong P^{B*^{1-e^{-\mu t}}} P_0^{e^{-\mu t}}, \quad 0 \le t < T,$$

where $\mu \equiv \frac{1}{2}\sqrt{\left[\rho + H(P^{B*})\right]^2 + 4\det(J)} - \frac{1}{2}\left[\rho + H(P^{B*})\right] > 0$ is the speed of adjustment of the pre-catastrophe system with det(J) defined in (20) and ω is the elasticity of the carbon tax manifold (17), $\omega \equiv \tau_P^B(P,\Delta)P/\tau$. Undershooting of the carbon tax occurs if $\omega < 0$ (or equivalently $\Gamma < 0$), overshooting occurs if $\omega > 0$ (or $\Gamma > 0$), and immediate adjustment occurs if the hazard rate is constant and $\omega = \Gamma = 0$.

Proof: see appendix.

These paths are local approximations around the pre-catastrophe steady state and cannot be used to characterize optimal trajectories that are far from the steady state. Local convergence to the before-catastrophe steady state given in proposition 1 is guaranteed if the dynamics of the pre-catastrophe system defined by (15) and (17) is indeed saddle-path stable if condition (19) is satisfied.[13] Examining the definition of $\det(J)$ in (19), we see that $\det(J)$ increases in absolute value if the hazard function is more convex and thus the speed of adjustment of the pre-catastrophe system will be higher. We will now apply propositions 1 and 2 to an illustrative calibration of our model to inform us of how the carbon tax, carbon stock and temperature are affected by the hazard of a sudden and irreversible release of greenhouse gases from the ocean floors when the hazard itself increases with the carbon stock and global mean temperature.

4. Illustrative calibration

We first calibrate energy demand, global mean temperature, the marginal global warming damage, the discount rate, and the implied conventional social cost of carbon. We also discuss the calibration of the size of the climate hazard and the functional form of the hazard function. We then offer some calculations on how the steady-state and transient paths of the social cost of carbon are affected by impending climate catastrophes resulting in abrupt and irreversible release of greenhouse gases.

4.1. Energy demand and utility

From the BP Statistical Review fossil fuel consumption in 2010 is 8.3 Giga tons of carbon. We take the cost fossil fuel to be 3.7 US \$ per million BTU or equivalently 504 US \$ per ton of carbon, which is a fairly low figure to reflect the abundance of coal. We suppose a constant elasticity of fossil fuel demand of $\varepsilon = 0.85$, which implies that utility of fossil fuel consumption is $U(E) = AE^{1-1/\varepsilon} / (1-1/\varepsilon)$ with A a positive constant to be determined. Calibrating our model to the global economy without a carbon tax based on 2010 figures for fossil energy use and using $E = (A/d)^{\varepsilon}$ we back out the constant $A = (8.3\times 10^{9})^{-1/0.85} \times 504 = 2.35\times 10^{14}$. With a carbon tax, we have the fossil fuel demand function $E = \left[A/(d+\tau)\right]^{\varepsilon}$. Utility of consumption of fossil fuel and other consumption goods is given by $\bar{U}(\tau) = Y + A^{\varepsilon}(d+\tau)^{1-\varepsilon} / (\varepsilon - 1)$, where Y is set to the 2010 value of world GDP, 63 trillion US dollars.

[13] Generally, the dynamic system for the social cost of carbon and the carbon stock will be non-autonomous due to the occurrence of a probability that enters via the expectation and depends explicitly on time (unless a new state is defined for the cumulative hazard). However, the special nature of the exponential hazard function adopted in this paper makes the system autonomous and local convergence is therefore ensured provided (19) holds.

4.2. Global mean temperature and the initial carbon stock in the atmosphere

To calculate the global mean temperature *Temp*, we follow IPPC (2007) and assume an equilibrium climate sensitivity of 3. This implies that doubling the stock of atmospheric carbon yields an increase in global mean temperature of 3 degrees Celsius. Hence, $Temp = 3\ln(P/581)/\ln(2)$ with 581 GtC equal to the pre-industrial stock of atmospheric carbon. The 2010 stock of carbon equals P_0 = 826 GtC.

4.3. Marginal global warming damage, the discount rate and carbon decay

Empirical integrated assessment models of climate change yield estimates of the social cost of carbon ranging from 5 to 35 US $ per ton of carbon in 2010 and rising to US $16 to $50 per ton in 2050 (e.g., the DICE, PAGE and FUND models described in Nordhaus (2008), Hope (2006) and Tol (2002), respectively). We will calibrate our model to a ballpark social cost of carbon of 25 US $ per ton of carbon when only taking account of smooth global warming damages and not of impending climate disasters. We set the discount rate to 1.5% per annum, $\rho = 0.015$, and the rate of atmospheric decay of the stock of atmospheric carbon to 0.4% per annum, $\delta = 0.004$, which gives a marginal global warming damage parameter of $\psi = (\rho + \delta)\ \tau^N = 0.019 \text{ x } 25 = 0.475$. The Stern Review obtains much higher estimates of the social cost of carbon with a much lower discount rate of only 0.1% per annum (e.g., Stern (2007)). This yields a no-shock social cost of carbon of 0.475/(0.001+0.004) = 95 US $ per ton of carbon.

4.4. The size and hazard of a climate catastrophe

Unfortunately, there is almost no information on the functional specification and calibration of the hazard function. To make some headway, we make bold assumptions about the size and hazard of the disaster. As far as the size of the impending climate disaster is concerned, we set the positive feedback parameter $\Theta = 0.028$ which corresponds to a massive release of carbon from the ocean floors.

To calibrate the climate calamity hazard function, we suppose that the expected duration for the climate catastrophe to hit is 14.7 years if the carbon concentration stays put at 2324 GtC and global mean temperature remains unchanged at 6 degrees Celsius (cf. Nordhaus, 2008; Golosov et al., 2012). This gives a hazard rate of 0.068. We suppose that the expected duration for disaster to strike at the current carbon stock of 826 GtC is hundred years, so that H(826 GtC) = 0.01. We use these two points to calibrate the following linear hazard function:

$$(23) \qquad H(P) = 0.01 + 3.872 \times 10^{-5} \times (P - 826), \quad P \geq 826 \text{ GtC}.$$

Although this function implies a hazard rate of 0.0005 of this climate calamity occurring in pre-industrial times, this hazard function does not apply in this range. It may be more reasonable to suppose that the hazard function is convex, so that the hazard rate increases more rapidly once the carbon stock and temperature are high enough. Hence, we also calibrate a quadratic hazard function to the same two points and in the same relevant range for the atmospheric carbon stock:

$$(23') \qquad H(P) = 0.01 + 2.585 \times 10^{-8} \times (P - 826)^2, \quad P \geq 826 \text{ GtC}.$$

The quadratic hazard function does not apply to atmospheric carbon stocks lower than the present one.[14]

Figure 2: Calibration of the hazard function

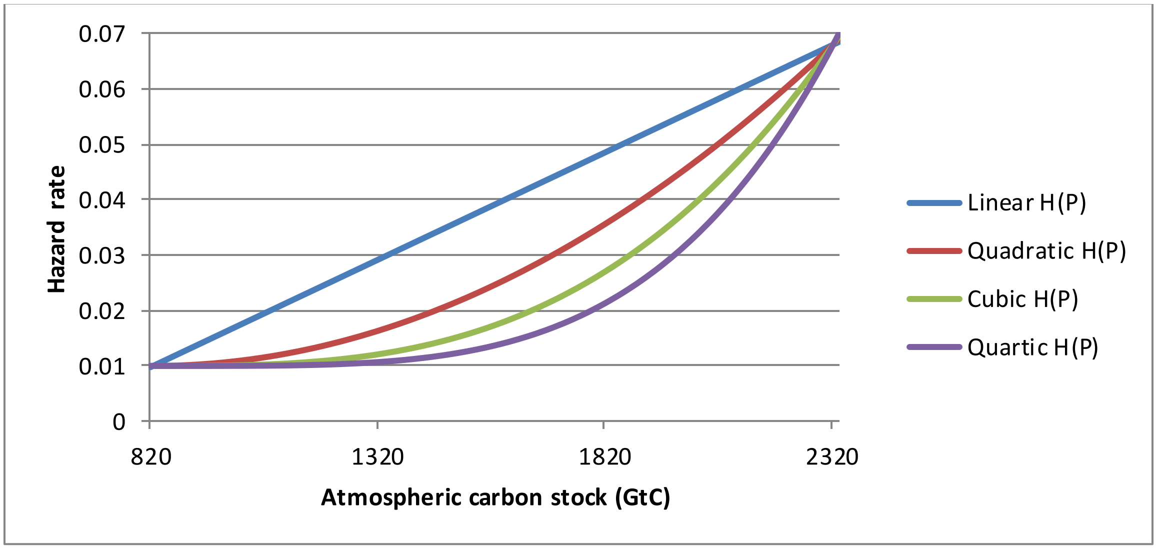


Fig. 2 indicates that the quadratic hazard function lies over the relevant range below the linear hazard function, but it rises more steeply at higher levels of the carbons stock. Fig. 2 also plots a cubic and a quartic hazard function calibrated again to the same two points which imply that the hazard rate rises even more steeply at higher temperatures.

5. Correcting the social cost of carbon for an impending climate catastrophe

5.1. After-catastrophe social cost of carbon

We have calibrated our model such that the social cost of carbon with a discount rate of 1.5% per year equals 25 US \$/tC if no account is taken of the impending abrupt release of carbon and positive feedback.

---

[14] The quadratic hazard function (23′) would imply hazard rates bigger than 0.01 for $P < 826$ GtC which is clearly not realistic and therefore the relevant range is restricted to $P > 826$ GtC. For this range the quadratic hazard function (23′) gives a useful convex alternative to the linear function (23) provided the optimal path for the stock of atmospheric carbon does not go outside the relevant range. This is the case for our calibration as can be witnessed from the second panel of the simulation reported in fig. 3 but in general this need not be the case. One might resolve this problem (and gain more realism) by calibrating more complex hazard functions to the points $H(581) = 0$, $H(826) = 0.01$ and $H(2324) = 0.068$ whilst ensuring $H' > 0$ over the entire range, but this would not fundamentally alter our insights from the simulations reported in section 5 about how to best respond to climate catastrophes.

This curbs fossil use from 8.3 to 7.5 GtC per year. After the disaster has struck, the social cost of carbon and the global carbon tax have to be increased to curb fossil fuel use:

$$(24)\qquad \tau^A = \frac{\psi}{\rho+\delta-\Theta} = 29 > \tau^N = 25 \text{ US \$/tC} \quad \text{and} \quad E(t) = E(\tau^A) = 7.4 < 7.5 \text{ GtC}, \quad t \geq T.$$

The steady-state carbon stock is thus curbed from 1882 to 1852 GtC, which leads to a modest cut in ultimate global warming from 5.09 to 5.02 degrees Celsius. The increase in the social cost of carbon as a result of the climate catastrophe is much higher with a very low discount rate of 0.1% per year which thus necessitates a much more ambitious climate policy (cf., Stern, 2007):

$$(24')\qquad \tau^A = \frac{\psi}{\rho+\delta-\Theta} = 216 > \tau^N = 95 \text{ US \$/tC} \quad \text{and} \quad E(t) = E(\tau^A) = 4.5 < 6.0 \text{ GtC}, \quad t \geq T.$$

As a result, the steady-state carbon stock is substantially reduced from 1501 GtC in the absence of disaster to 1126 GtC with disaster which makes a much bigger dent in global warming from 4.11 to 2.86 degrees Celsius. The speed of adjustment of the after-calamity system is much slower than of the no-calamity system, i.e., $\delta - \Theta$ = 0.0012 < δ = 0.005, which is a dire consequence of the positive feedback in the carbon cycle.

5.2. Before-catastrophe social cost of carbon: constant hazard rate

With a constant hazard rate set to h_0 = 0.03 and a discount rate of 1.5% per annum we obtain a carbon tax in between the no-shock and after-catastrophe carbon tax:

$$(25)\qquad \tau^N = 25 < \tau(t) = \left(\frac{\rho+\delta+h_0-\Theta}{\rho+\delta+h_0}\right)\tau^A = 27.6 < \tau^A = 29.3, \forall t \geq 0 \quad \text{if} \quad \rho = 0.015.$$

With a lower discount rate of ρ = 0.001 (as in Stern (2007)), the effects are much more pronounced:

$$(25')\qquad \tau^N = 95 < \tau(t) = \left(\frac{\rho+\delta+h_0-\Theta}{\rho+\delta+h_0}\right)\tau^A = 198.6 < \tau^A = 215.9, \forall t \geq 0 \quad \text{if} \quad \rho = 0.001.$$

5.3. Before catastrophe social cost of carbon: temperature-dependent hazard rate

These results and further results with temperature-dependent hazard functions are presented in table 1. There are various striking observations. First, the carbon tax undershoots with linear hazard functions but overshoots its steady-state value on impact of the realization of the impending catastrophe with all of the three convex hazard functions shown in fig. 2. Second, the overshooting is more severe for more convex hazard functions as can be seen from the increasing modulus of the manifold elasticity, ω, and also more severe for lower discount rates. Third, adjustment of the before-catastrophe carbon tax and the carbon stock is faster for more convex hazard functions but slower for lower discount rates. Fourth, the magnitude of the optimal carbon tax and the mitigating impact of the catastrophe on carbon accumulation and temperature are much higher for a low than for a high discount rate.

Table 1: Catastrophe and the social cost of carbon

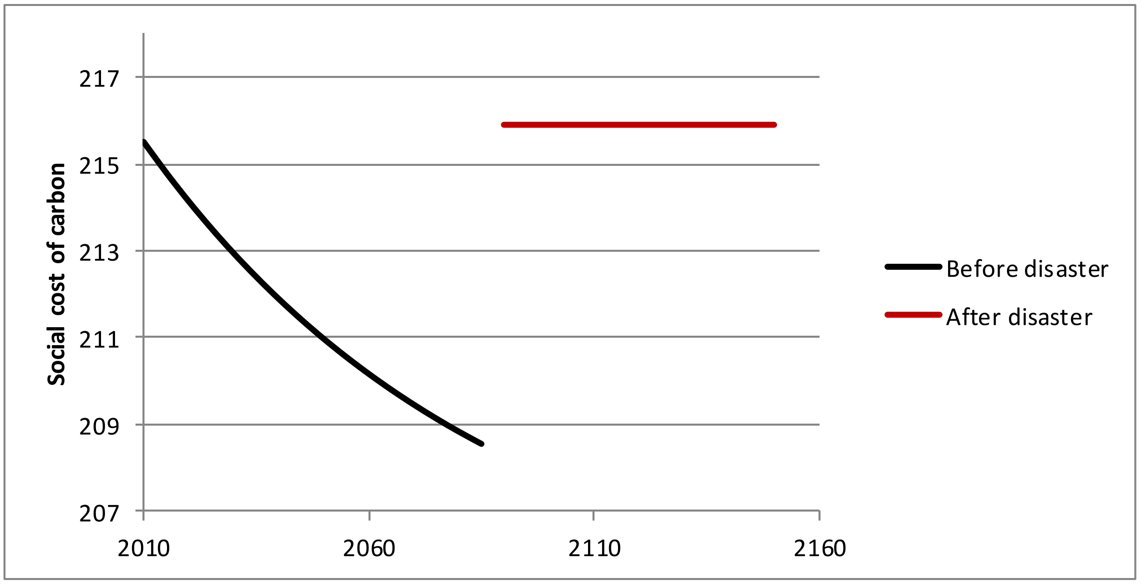
	$\tau(0)$ (\$/tC)	τ^{B*} (\$/tC)	E^* (GtC/year)	P^* (GtC)	*Temp* (^{o} Celsius)	Manifold elasticity	Adjustment speed
Business as usual: no disaster	0	0	8.3	2075	5.51	n.a.	0.004
Business as usual: disaster	0	0	8.3	6917	10.72	n.a.	0.0012
Optimum with no catastrophe:							
ρ = 0.015	25.0	25.0	7.5	1882	5.09	0	0.004
ρ = 0.001	95.0	95.0	6	1501	4.11	0	0.004
Optimum after catastrophe:							
ρ = 0.015	29.3	29.3	7.4	6174	10.23	0	0.0012
ρ = 0.001	215.9	215.9	4.5	3753	8.07	0	0.0012
Optimum before catastrophe:							
ρ = 0.015, $H(P)$ = 0.03	27.6	27.6	7.5	1864	5.04	0	0.004
ρ = 0.015, linear hazard	27.7	28.1	7.4	1860	5.04	0.0177	0.0357
ρ = 0.015, quadratic hazard	27.8	27.4	7.5	1865	5.05	-0.0190	0.0393
ρ = 0.015, cubic hazard	27.9	27.4	7.5	1865	5.05	-0.0206	0.0423
ρ = 0.015, quartic hazard	27.9	27.4	7.5	1865	5.05	-0.0218	0.0044
ρ = 0.001, $H(P)$ = 0.03	198.6	198.6	4.7	1167	3.02	0	0.004
ρ = 0.001, linear hazard	184.8	191.9	4.7	1183	3.08	0.1043	0.0105
ρ = 0.001, quadratic hazard	209.2	200.4	4.6	1162	3.00	-0.1252	0.0111
ρ = 0.001, cubic hazard	213.4	202.8	4.6	1156	2.98	-0.1505	0.0123
ρ = 0.001, quartic hazard	215.5	204.3	4.6	1153	2.97	-0.1601	0.0127

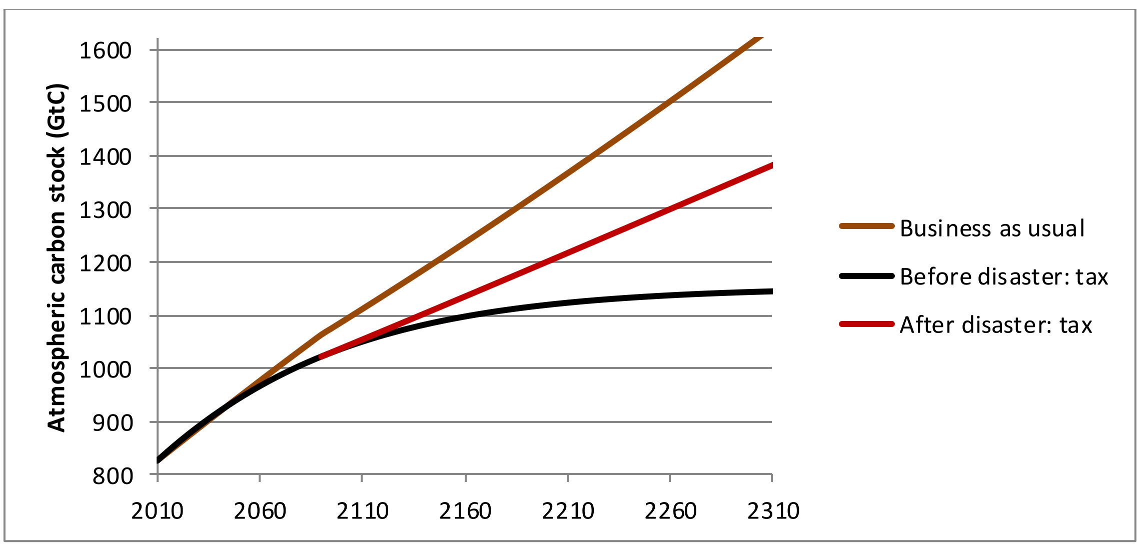
Illustrative time paths for when the catastrophe strikes after in 2090 for ρ = 0.1% per annum and a quartic hazard function are presented in fig. 3. This is the case of a convex hazard function and short-run overshooting of the social cost of carbon. Although the stock of atmospheric carbon rises to an alarming 3753 GtC in the long run, this occurs at a very slow rate of only 0.12% per annum. This is why the carbon stock has only increased to 1135 GtC whilst it would have been 1097 GtC if the disaster would not have struck in 2090. This suggests that the adverse effects on global warming of unleashed positive feedback in the carbon cycle can take many decades, even centuries before all the cumulated harm has been done.

5.4. Elasticity of fossil fuel demand and the pre-catastrophe social cost of carbon

Given linear smooth global warming damages (assumption 2), the after-catastrophe social cost of carbon is unaffected by the price elasticity of fossil fuel demand, ε, or by the cost of supplying fossil fuel to the market, d. However, this is not the case for the pre-catastrophe social cost of carbon as these factors affect fossil fuel emissions and thus eventually affect temperature and the risk of climate disaster as well. For example, if the price elasticity of fossil fuel demand is only 0.4 instead of 0.85, the business-as-usual scenario is unaffected (as it is the base for the calibration). The post-catastrophe carbon tax is (215.9 US

Figure 3: Transient effects of impending tipping point on social cost of carbon and carbon stock





\$/tC if $\rho = 0.1\%$ as in Stern (2007)) also unaffected, but fossil fuel demand is higher after the disaster and thus the carbon stock is higher also (5187 instead of 3753 GtC) and global warming is more severe (9.47 versus 8.07 degrees Celsius). The pre-catastrophe carbon tax is, however, reduced from 204.3 to 197.1 US \$/tC. The initial carbon tax is only somewhat lower (214.7 instead of 215.5 US \$/tC). The optimal social cost of carbon is thus not much affected by the price elasticity of fossil fuel demand. We conclude that, if better substitutes for fossil fuel that arrive on the market push up the elasticity, the carbon tax does not have to be raised by much.

6. Conclusion

We have analyzed how the first-best optimal carbon tax should be adjusted in the face of an impending climate disaster consisting of a sudden, irreversible release of greenhouse gases from the ocean floors when such a disaster is more likely to occur if the stock of atmospheric carbon and temperature are high. We have shown how the optimal global carbon tax should be increased over and above the conventional Pigouvian carbon tax to curb the risk of various types of climate catastrophe occurring. This extra component of the carbon tax increases in the expected gap between welfare before and after the catastrophe arising from burning an additional unit of fossil fuel. This captures that an increase in the stock of atmospheric carbon or temperature pushes up the risk of a climate catastrophe and the potential loss in welfare that this may cause. This term disappears if the hazard of a climate calamity is exogenous. There is also a second term pushing up the carbon tax, which captures that burning an additional unit of fossil fuel raises temperature and thus lowers after-calamity welfare. This means that the stakes are raised, since the fall in welfare following a climate catastrophe is bigger.

We have also shown that the initial social cost of carbon or carbon tax undershoots its steady-state value for a linear hazard function, but overshoots for convex enough hazard functions. We also find that the before-catastrophe carbon tax and the carbon stock adjust more quickly for more convex hazard functions but more slowly for lower discount rates. The optimal carbon tax and the mitigating impact of the catastrophe on carbon accumulation and temperature are much higher for a low than a high discount rate.

Calibrating our model to a social cost of carbon of 25 US $/tC, a rate of atmospheric decay of 0.4% of the carbon stock per annum and a discount rate of 1.5% per annum, we consider the hazard of a sudden, irreversible release of carbon corresponding to a positive feedback of 0.12% of the carbon stock per annum. Pinning down the hazard rate at 6 degrees Celsius to 0.068, we find that for a linear hazard function the social cost of carbon jumps up to 27.7 US $/tC and then rises gradually to 28.1 US $/tC until the catastrophe strikes from which point onwards the social cost of carbon is raised to 29.3 US $/tC. Hence, the effects of the impending catastrophe are 'small potatoes'.[15] However, with a discount rate of 0.1% per annum as used in Stern (2007) the effects of the tipping point on the social cost of carbon are much more pronounced. The carbon tax jumps up to 184.8 US $/tC and then rises to 191.9 US $/tC.

---

[15] Our drops in global warming are small even for relatively high carbon taxes, since we abstract from the potential of the carbon tax to curb global warming via other means such as bringing on a carbon-free era based on renewable energy, stimulating R&D into energy-saving and renewable technology (e.g., Rezai and van der Ploeg, 2013).

When the catastrophe strikes the carbon tax is increased to 215.9 US $/tC. The effects are stronger for convex (such as the quadratic, cubic or quartic) hazard functions.

In a general equilibrium framework with Ramsey growth and an impending shock to total factor productivity one needs, in addition to a higher carbon tax to curb the risk of climate calamities, also precautionary capital accumulation to be better prepared for when the catastrophe strikes (e.g., van der Ploeg and de Zeeuw, 2013). This precautionary capital accumulation must take place even if the hazard rate is exogenous. If the hazard does depend on temperature, the carbon tax is pushed up which curbs fossil fuel demand and capital demand thereby offsetting precautionary capital accumulation. The risk of climate disaster will, in general, also make it attractive to shift capital that is used as engine of growth from the productive process to precautionary adaptation capital with the aim to mitigate the adverse effects of potential climate disasters (water defences etc.). Hence, as a result of the hazard of a climate catastrophe, adaptation capital increases and capital used in the production process falls, especially if the hazard increases strongly with global warming.

Future work should allow for multiple, possibly reversible and gradually emerging catastrophes with different temperature-dependent hazard rates and other types of climate uncertainty (cf., Lemoine and Traeger, 2013; Cai et al., 2012). The challenge for future empirical research is to pin down the effect of raising the carbon tax and curbing temperature on the probability of climate disaster as well as on the baseline probability of climate disaster. With more knowledge of the shape of the hazard function we can improve our understanding of how much the carbon tax should be raised to cut the risk of climate disaster. Future work should also allow for a world in which there is a small risk that a climate disaster occurs but also a small probability that such a disaster is never going to materialize in which case Bayesian updating is needed to revise those probabilities. This implies that on top of the 'risk averting' component of the social cost of carbon there should be a 'probing' component to find out more about the properties of the climate system and the risk of tipping points.

Acknowledgements

Support from an Advanced Instigator Grant is gratefully acknowledged. Useful comments from Cees Withagen and Amos Zemel on an earlier version of this paper and helpful discussions with Aart de Zeeuw on the topic of climate catastrophes and regime switches are gratefully acknowledged.

References

Amigues, J.-P. and M. Moreaux (2012). Potentially irreversible catastrophic shifts of the assimilative capacity of the environment, TSE Working Paper 12-276, Toulouse School of Economics.

Cai, Y., K.L. Judd and T. Lontzek (2012). The social cost of abrupt climate change, mimeo., Hoover Institution, Stanford University.

Clarke, H.R. and W.J. Reed (1994). Consumption/pollution tradeoffs in an environment vulnerable to pollution-related catastrophic collapse, Journal of Economic Dynamics and Control, 18, 991-1010.

Cropper, M.L. (1976). Regulating activities with catastrophic environmental effects, Journal of Environmental Economics and Management, 3, 1-15.

Gollier, C. (2008). Discounting with fat-tailed growth, Journal of Risk and Uncertainty, 37, 171-186.

Gollier, C. (2012). Pricing the Planet's Future: The Economics of Discounting in an Uncertain World, Princeton University Press, Princeton, New Jersey.

Golosov, M., J. Hassler, P. Krusell and A. Tsyvinski (2013). Optimal taxes on fossil fuel in general equilibrium, Econometrica, forthcoming.

Heal, G.M. (1984). Interactions between economy and climate: a framework for policy design uncertainty, 151-168 in V. Kerry Smith and A. White (eds.), Advances in Applied Micro-Economics, Vol. 3, JAI Press, Greenwich, Connecticut.

Hope, C.W. (2006). The marginal impact of CO2 from PAGE 2002: an integrated assessment model incorporating the IPCC's five reasons for concern, Integrated Assessment, 6, 1-16.

IPCC (2007): Climate Change 2007, the Fourth IPCC Assessment Report. http://www.ipcc.ch/ipccreports/ar4-syr.htm

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

Kimball, M. (1990). Precautionary saving in the small and in the large, Econometrica, 58, 1, 53-73.

Kopits, E., A.L. Martin and A. Wolverton (2013). Moving forward with integrating "catastrophic" climate change into policy analysis, Working Paper #13-01, National Center for Environmental Economics, U.S. Environmental Protection Agency, Washington, D.C.

Lemoine, D.M. and C.P. Traeger (2013). Watch your step: optimal policy in a tipping climate, American Economic Journal: Economic Policy, in print.

Lenton, T.M. and J.-C. Ciscar (2013). Integrating tipping points into climate impact assessments, Climatic Change, 117, 585-597.

Naevdal, E. (2006). Dynamic optimization in the presence of threshold effects when the location of the threshold is uncertain – with an application to a possible disintegration of the Western Antarctic Ice Sheet, Journal of Economic Dynamics and Control, 30, 1131-1158.

Nordhaus, W. (2008). A Question of Balance: Economic Models of Climate Change, Yale University Press, New Haven, Connecticut.

Pindyck, R.S. (2013). The climate policy dilemma, Review of Environmental Economics and Policy, 7, 2, 219-237.

Ploeg, F. van der and A.J. de Zeeuw (2013). Climate policy and catastrophic change: be prepared and avert risk, Research Paper 118, OxCarre, University of Oxford.

Polaski, S., A.J. de Zeeuw and F. Wagener (2011). Optimal management with potential regime shifts, Journal of Environmental Economics and Management, 62, 2, 229-240.

Prieur, F., M. Tidball and C. Withagen (2013). Optimal emission-extraction policy in a world of scarcity and irreversibility, Resource and Energy Economics, 35, 4, 637-658.

Rezai, A. and F. van der Ploeg (2013). Abandoning fossil fuel: the role of carbon taxes and renewable subsidies, Research Paper 123, OxCarre, University of Oxford.

Stern, N.H. (2007). The Economics of Climate Change: The Stern Review, Oxford University Press, Oxford.

Tol, R. (2002). Estimates of the damage costs of climate change, Environmental and Resource Economics, 21, 135-160.

Tsur, Y. and A. Zemel (1996). Accounting for global warming risks: resource management under event uncertainty, Journal of Economic Dynamics and Control, 20, 1289-1305.

Tsur, Y. and A. Zemel (2008). Regulating environmental threats, Environmental and Resource Economics, 39, 297-310.

Tsur, Y. and A. Zemel (2009). Endogenous discounting and climate policy, Environmental and Resource Economics, 44, 507-520.

Tsur, Y. and C. Withagen (2013). Preparing for catastrophic climate change, Journal of Economics, in print.

Weitzman, M.L. (1998). Why the far-distant future should be discounted at its lowest possible rate?, Journal of Environmental Economics and Management, 36, 201-208.

Zeeuw, A.J. de and A. Zemel (2012). Regime shifts and uncertainty in pollution control, Journal of Economic Dynamics and Control, 36, 7, 939-950.

Appendix

Proof of Proposition 1: The stationarity conditions for (15) and (17) give rise to the expressions given in (20a) and (20b). To perform the comparative statics, we substitute (20b) into (20a) and rearrange to get:

$$\text{(A1)}\qquad \left[\rho+\delta+H\left(E(\tau^{B*})/\delta\right)\right]\left\{\tau^{B*}-\left[\frac{\rho+\delta+H(P^{B*})-\Theta}{\rho+\delta+H(P^{B*})}\right]\tau^{A}\right\}= H'\left(E(\tau^{B*})/\delta\right)\left[V^{B}\left(E(\tau^{B*})/\delta,\Theta\right)-V^{A}\left(E(\tau^{B*})/\delta,\Theta\right)\right].$$

Totally differentiating equation (20′) and making use of (9′), (11′) and (13) yields:

$$\left[\rho+\delta+H+\left\{\tau^{B*}-\left[\frac{\rho+\delta+H(P^{B*})-\Theta}{\rho+\delta+H(P^{B*})}\right]\tau^{P}\right\}\frac{E'H'}{\delta}+(\tau^{B*}-\tau^{P})\frac{E'H'}{\delta}-(V^{B}-V^{A})\frac{E'H''}{\delta}\right]d\tau^{B*}$$

(A2) $$-\left[\frac{\Theta\tau^{P}}{\rho+\delta+H(P^{B*})}\right]\frac{E'H'}{\delta}d\tau^{B*}=\left[\rho+\delta+H+2(\tau^{B*}-\tau^{P})\frac{E'H'}{\delta}-(V^{B}-V^{A})\frac{E'H''}{\delta}\right]d\tau^{B*}$$

$$=\Theta\delta d\tau^{B*}=(V_{\Theta}^{B}+P^{B*}\tau^{A}/\rho)H'd\Theta=\left(1-\frac{H}{\rho+H}\right)\frac{\tau^{A}P^{B*}H'}{\rho}d\Theta=\left(\frac{\tau^{A}P^{B*}H'}{\rho+H}\right)d\Theta.$$

Invoking the correspondence principle we can use condition (19) to establish that the term in the square brackets on the left-hand side of (A2) must be positive. It follows from (A2) that $d\tau^{B*}/d\Theta>0$ if $H'>0$ and $d\tau^{B*}/d\Theta=0$ if $H'=0$. Since with $\Theta=0$, we have $V^{B}=V^{A}$ and thus $\tau^{B}(t)=\tau^{P},\forall t\geq 0$. We thus establish that for $\Theta>0$ we must have $\tau^{B*}>\tau^{P}$. □

Proof of Proposition 2: We define $\tilde{P}(t)\equiv\ln\left(P(t)/P^{B*}\right)$ and $\tilde{\tau}(t)\equiv\ln\left(\tau(t)/\tau^{B*}\right)$ and loglinearize (15) and (17) around the pre-catastrophe steady state:

$$\begin{pmatrix}\dot{\tilde{P}}\\ \dot{\tilde{\tau}}\end{pmatrix}=J\begin{pmatrix}\tilde{P}\\ \tilde{\tau}\end{pmatrix},\quad J=\begin{pmatrix}-\delta & E'(\tau^{B*})\tau^{B*}/P^{B*}\\ \Gamma P^{B*}/\tau^{B*} & \rho+\delta+H(P^{B*})\end{pmatrix}.$$

Det[J] = $-\Theta<0$, since (19) holds and the system displays saddle-path stability. Solving the characteristic polynomial of the Jacobian matrix J, $\lambda^{2}-\left[\rho+H(P^{B*})\right]\lambda-\det(J)=0$, yields $\lambda_{1}=-\mu<0$ as the negative eigenvalue and λ_2 as the positive eigenvalue. Spectral decomposition yields

$J=N\begin{pmatrix}\lambda_{1} & 0\\ 0 & \lambda_{2}\end{pmatrix}N^{-1}=M^{-1}\begin{pmatrix}\lambda_{1} & 0\\ 0 & \lambda_{2}\end{pmatrix}M$, where M and N are the matrices of column and row eigenvectors.

Defining $\begin{pmatrix}y_{1}\\ y_{2}\end{pmatrix}\equiv M\begin{pmatrix}\tilde{P}\\ \tilde{\tau}\end{pmatrix}$, we get $\begin{pmatrix}\dot{y}_{1}\\ \dot{y}_{2}\end{pmatrix}\equiv\begin{pmatrix}\lambda_{1} & 0\\ 0 & \lambda_{2}\end{pmatrix}\begin{pmatrix}y_{1}\\ y_{2}\end{pmatrix}$. Ruling out explosive paths gives

$y_{2}(t)=0,\forall t\geq 0$. We also have $y_{1}(t)=y_{10}e^{\lambda_{1}t}$, where the constant y_{10} needs to be determined. It follows that $\tilde{P}=N_{11}y_{1}$, hence $y_{10}=\tilde{P}_{0}/N_{11}$. We also have $\tilde{\tau}=N_{21}y_{1}$. We can normalize $N_{11}=1$ and calculate $N_{21}=(\lambda_1-A_{11})/A_{12}=\omega$ with ω as defined in (21). Alternatively, we can write $\omega=A_{21}/(\lambda_1-A_{22})$ which has the same sign as A_{21} or Γ. This gives $\tilde{P}(t)=\tilde{P}_{0}e^{\lambda_{1}t}$ and $\tilde{\tau}(t)=\omega\tilde{P}_{0}e^{\lambda_{1}t}$. These can be readily transformed to equations (22) and (21), respectively. □