CESifo Working Papers

www.cesifo.org/wp

Carbon Prices for the Next Hundred Years

Reyer Gerlagh
Matti Liski

CESIFO WORKING PAPER NO. 4671
CATEGORY 10: ENERGY AND CLIMATE ECONOMICS
FEBRUARY 2014

CESifo *Working Paper No. 4671*

Carbon Prices for the Next Hundred Years

Abstract

World income grows fast without verifiable climate-change impacts on the economy. The growth spell can end if climate impacts turn real but this can take decades to learn. We develop a tractable stochastic climate-economy model with a hidden-state impact process to evaluate the contributions of the expanding economy and changing impact beliefs to the social cost of carbon. Taking a dataset of estimates for the social cost as a representation of beliefs, we assess how robust climate policies are to the delays of hard information. The carbon price should rise with income to the next century, even without observed impacts. The carbon price should grow faster than the economy as long as climate warming is not enough for generating impacts that are informative about the true social cost.

JEL-Code: H430, H410, D610, D910, Q540, E210.

Keywords: carbon price, learning, climate change.

Reyer Gerlagh
Economics Department
Tilburg University
Tilburg / The Netherlands
r.gerlagh@uvt.nl

Matti Liski
Economics Department
Aalto University
Helsinki / Finland
matti.liski@aalto.fi

Jan 28, 2014
We have received helpful comments on an early version from Derek Lemoine, and audiences at the CESifo area conference “Energy and Climate Economics”, Paris School of Economics, Berlin TU, EAERE 2013, and Helsinki. We have also benefitted from discussions with Lint Barrage, David Hemous, Terry Iverson, Anthony Millner, and Svenn Jensen. We thank James Corbishley for research assistance.

"Estimating impacts has been the most difficult part of all climate science"
—William W.D. Nordhaus, EAERE lecture 2012

1 Introduction

A price for carbon measures the social cost of releasing a unit of carbon dioxide to the atmosphere, based on expected climate-change impacts.[1] However, there is little or no quantitative information on the impacts of persistent climate change on our economies, although there is extensive research on what such impacts might be.[2] The social cost of carbon is based on *beliefs* about impacts that will be updated when the "climate experiment" generates actual impacts. But this can take a long period of time; the past century of carbon emissions has not yet led to precise estimates, and another 50-100 years may pass without additional hard evidence on the ultimate consequences of current emissions. In view of such time delays for evidence, the carbon price as a concept may appear elusive and difficult to defend.

Roe and Baker establish (2007) that, because of positive feed-back mechanisms of the climate system, it is unlikely that we will better understand the temperature sensitivity to emissions in the near future. The economic literature modeling the learning of climate impacts has almost exclusively focused on the structural uncertainties of the climate system, including those related to the climate sensitivity (Kelly and Kolstad 1999; Leach 2007; Kelly and Tan 2013) and to unknown thresholds leading to tipping points (Lemoine and Traeger 2014; Cai, Judd, and Lontzek, 2013). For many economists, such climate uncertainties and the implied low-probability but high-consequence events, which cannot be ruled out by new information any time soon, have become the prime argument for having a price for carbon (Weitzman, 2009, 2011, 2013; Pindyck 2013).[3]

[1] Most evaluations of the social cost of carbon build on a set of middle-of-the-road assumptions on climate change impacts, commonly expressed in terms of GDP losses, and then use climate-economy models such as DICE, FUND, or PAGE (see Greenstone, Kopits, and Wolverton, 2011, for a succinct description and references) that combine the impact assumptions with background scenarios to obtain a monetized value for the social cost. There is a pressing demand for such a number as it is required, for example, in the cost-benefit analyses to assess regulations across wide domains; however, see Pindyck (2013) for a critical review of the integrated-assessment models used in producing the numbers.

[2] See Tol (2009) for a survey on methods and results. There is a growing empirical literature on how climate impacts various sectors of the economy (e.g., Deschenes and Greenstone, 2007, and Schlenker and Roberts, 2009, Dell, Jones, and Olken, 2012).

[3] See also Heal and Millner (2013) for a survey on uncertainties in climate-change economics.

Accepting climate-change "unknowns" as reasons for beliefs supporting a current price for carbon, in this paper we argue that the reason for sustaining and increasing the price over long periods without hard evidence is fundamentally different and a well-understood economic variable: the growth of global income. For the coming century, the global income is expected to grow by multiple factors, in part due to the rise of the middle class in major emerging economies.[4] The US government has recently developed estimates for the carbon price, for regulatory purposes, assuming that the global GDP increases by a factor that varies between five and seven in this time-span (see Greenstone et al., 2011).

An economy grown five to seven times bigger values information on the social cost of climate change differently — but how exactly? For the contributions of income and beliefs to the social cost, we develop a tractable climate-economy model where impacts are initially neither observed nor experienced — they may arrive through a hidden-state impact response characterized by long delays and dependence on the past emissions history.

We develop policy rules that separate sharply the carbon price determinants that are clearly understood, such as the size of the economy, from those that involve beliefs about future impacts. In contrast with the general tone of the previous literature, "threats" originating from the natural science uncertainties of the climate problem are not the source of time-increasing carbon prices.[5] The well-understood part, that is, the expansion of the economic stake through the growth of the global economy is enough. We find conditions when carbon prices grow faster or slower than the economy. The price should grow faster than the economy if the current level of climate change cannot generate impacts substantial enough for learning the true social cost.

The model is set up to answer the question how the currently perceived social cost should develop if income growth progresses without verifiable economic climate-change impacts over periods such as the next 50-100 years. When no impacts occur over time, it becomes more likely that impacts from a given climate change will never become

[4]See, for example, the IPCC Special Report on Emissions Scenarios (2000), U.S. Climate Change Science Program (2007), Stanford Energy Modeling Forum (for example, in Weyant et al. 2006).

[5]By the nature of our quantitative exercise below, we rule out "tail events". The supporting potential high-damage climate event that justifies the estimated initial carbon price is equivalent to a GDP-loss of about 10 per cent at temperatures that are 3 degrees Celsius above the pre-industrial level. Such an event is economically significant but not a "tail event" in the sense of Weitzman (2009) where policies become undefined since, effectively, it is not possible to transfer wealth to the high consequence events; see, for example, Nordhaus (2010) and Millner (2013).

very large. Considering carbon prices for such a "no news is good news" scenario, we intentionally devise a conservative test for a climate policy ramp, that is, a gradual tightening of policies as advocated by previous studies (Nordhaus, 2007).[6] Taking a dataset of estimates for the social cost as a representation of current beliefs, put together by Tol (2009) in a study of existing estimates, we quantify how these initial social cost perceptions would have to change to overturn the contribution of rising income to the carbon price. To reverse the upward trend of the carbon price, the climate experts would have to become more optimistic, and rule out severe impacts on the economy, by orders of magnitude faster than what is implied by the scenario in our explorative calibration.

Arguably, the distribution of existing estimates arises from fundamental differences between the individual studies, introducing also a strong subjective component to the estimates (see Pindyck, 2013). But the subjective dispersion of views is exactly the reason why, in our framework, beliefs are introduced for interpreting the existing carbon price distribution. The approach to the calibration of beliefs is explorative; however, the general conclusions seem quantitatively robust. We are unaware of previous attempts to use current estimates of the social cost for assessing its dynamic development.

Interestingly, through growth, the economy may become fully decarbonized without experiencing economic climate-change impacts. The key by-product of income growth for climate policies is the increased willingness to pay for emissions reductions; Chichilnisky, Heal, and Starrett (1993) discuss a similar effect in a static context, and Stokey (1998) considers the effect in a dynamic setting. In our model with long-delayed learning, the pollution impact on the economy may not occur but the carbon emissions still experience a rise and decline much in the vein of the Environmental Kuznets curve –literature (Grossman and Krueger, 1993, 1995; Selden and Song 1994; Holz-Eakin and Selden, 1992).

We build on the Brock-Mirman model (1972) for the climate-economy interactions, following Golosov, Hassler, Krusell, and Tsyvinski (2011); however, we introduce climate change differently through a hidden state that determines whether a negative productivity shock can hit the economy in the future. Beliefs on the hidden state allows including heterogeneity of views, and the structural interpretation of the social cost data.

The paper is structured as follows. In Section 2, we first explain the basic planning problem, and then the learning dynamics as well as the emissions-temperature response

[6] It should be noted that there are other well-received arguments such as green technological change for not following gradualism but rather a jump-start in emissions pricing (van der Zwaan et al. 2002; Gerlagh, Kverndokk and Rosendahl 2009; and Acemoglu, Aghion, Bursztyn, and Hemous, 2012).

that follows from the description of the global carbon cycle; this description is in full detail in Gerlagh and Liski (2013).[7] In Section 3, we first introduce the optimal policies when the true state of nature is known, and then consider the policy before learning the impacts. Section 4 introduces the calibration and the quantitative assessment. The online supplementary file contains a program for reproducing the graphs in the text.[8]

2 The climate-economy model

2.1 The basic setting

We consider a climate-economy planning problem where production possibilities at time t depend on capital k_t inherited, and potentially also on the full history of carbon input use,

$$s_t = (z_0, ..., z_{t-1}).$$

Given k_t and history s_t at time period t, consumption, c_t, and carbon inputs, z_t, are chosen to maximize the expected discounted utility

$$\max \mathbb{E}_t \sum\nolimits_{\tau=0}^{\infty} \delta^{\tau} u_{t+\tau} \tag{1}$$

where $0 < \delta < 1$ is the discount factor and $u_{t+\tau}$ is the periodic utility, specified below. The chosen allocations must satisfy

$$c_t + k_{t+1} = y_t, \tag{2}$$

with $y_t = f_t(k_t, s_t, z_t)$ denoting the output at time t. Losses due to climate change arise as reduced output, and depend on the history of emissions s_t through variable D_t that is a measure of the global mean temperature increase above the pre-industrial levels at time t. We assume that this measure is a function of history s_t,

$$D_t = \sum\nolimits_{\tau=1}^{t} \mathcal{R}(\tau) z_{t-\tau} \tag{3}$$

where the weights $\mathcal{R}(\tau)$ define the "emissions-temperature response". That is, current emissions z_t affect temperatures at some later time $t+\tau$ according to a known response function $\mathcal{R}(\tau)$:

[7] A longer working paper version Gerlagh and Liski (2012) contains a detailed description of the energy sector that is needed in the quantitative analysis of the current paper. That paper focuses on the valuation of far-distant climate impacts, without uncertainty.

[8] Follow the link `https://www.dropbox.com/sh/7meos655j14jh5p/_dlr8X_FHI`

$$\frac{dD_{t+\tau}}{dz_t} = \mathcal{R}(\tau) > 0. \tag{4}$$

The key characteristic of the calibrated $\mathcal{R}(\tau)$, explained below, is the considerable delay of the response following an impulse of emissions; it has a non-linear shape peaking several decades after the date of the emissions, and a fat tail of almost permanent impacts. Simplistically, there is no uncertainty about $\mathcal{R}(\tau)$; the response serves the purpose of introducing delays to the potential impacts on the economy that, in turn, will be uncertain.

Output is given by a production function where capital contribution takes the Cobb-Douglas form, with $0 < \alpha < 1$,

$$y_t = k_t^\alpha A_t(z_t) \exp(-\Delta_{y,t} D_t), \tag{5}$$

where the contribution of carbon inputs z_t enter through the function $A_t(z_t)$ that captures the energy sector of the economy. The current policy will be free of details of the energy sector; we merely assume that carbon input z_t has a positive but diminishing marginal product.[9] Losses from climate change arise as reduced output, as in most applied climate-economy models (e.g., Nordhaus, 2008); moreover, they depend on the history of emissions through variable D_t and damage coefficient $\Delta_{y,t} \geqslant 0$, as in Golosov et al. (2011).

There are two climate-economy states, $I_t \in \{0, 1\}$. If $I_t = 0$, no damages have been experienced by t. If $I_t = 1$, damages have appeared, and once $I_t = 1$, then $I_{t+\tau} = 1$ for all $\tau \geq 0$. The damage coefficient at time t is $\Delta_{y,t} = \Delta_y I_t$, where $\Delta_y > 0$ is a constant, independent of time. Thus, there is a dichotomy between climate change, captured by D_t, and impacts, $\Delta_{y,t}$, where only the latter will be unknown.[10] The economy starts with $\Delta_{y,t} = 0$; below, we specify the learning process for the future values of $\Delta_{y,t} > 0$.

Periodic utility is

$$u_t = u(c_t) - \Delta_{u,t} D_t, \tag{6}$$

where $u(c_t) = \ln(c_t)$, $\Delta_{u,t} = \Delta_u I_t$, and $\Delta_u \geqslant 0$. We thus allow for intangible damages that can appear together with the production losses, once $I_t = 1$.

[9]However, the details of the energy sector will affect the future development of the economy and thus the future states of the economy and future policies. For this reason, we introduce a structure for the energy sector in detail in Section 4.2 and in the Appendix

[10]The dichotomy can be broken by assuming a smooth arrival process for impacts; the extension in the Appendix can be interpreted this way. The approach in the main text allows sharper analytics, and the substance-related differences between the two approaches are small.

The economic problem defined through (1)-(6) has a well-explored structure, apart from the climate-economy interactions. The state vector is (k_t, s_t, I_t). Because of the log-utility for consumption, full capital depreciation in one period, and Cobb-Douglas capital contribution, the consumption choice model follows Brock and Mirman (1972) so that share

$$g = \alpha\delta$$

of gross output will be saved; the dynamic programming arguments leading to this policy are well known in analytical macro-economics (Sargent, 1987). Moreover, given the exponential form for the potential output loss, the contribution of k_t and s_t to value of the program in (1) will be separable in these variables. Thus, the climate policy analysis can be conducted by taking savings g as given and by tracking the direct utility impacts of the potential loss from climate change.[11] It proves useful to aggregate both the potential output and direct utility losses into one measure:

Remark 1 *For $I_t = 1$, the present-value loss of utils from marginal climate change at time t is*

$$\Delta \equiv -\sum\nolimits_{\tau=0}^{\infty} \delta^{\tau} \frac{du_{t+\tau}}{dD_t} = \Delta_u + \frac{\Delta_y}{1-g}. \tag{7}$$

Thus, in equilibrium, output and direct utility losses can be made interchangeable in terms of utility; for convenience, we will use Δ as an aggregate measure of both losses. For the proof, consider the effect of temperature $D_{t+\tau}$ on utility in period $t+\tau$ when $I_t = 1$ (climate impacts have arrived). Recall that the consumption utility is $\ln(c_{t+\tau}) = \ln((1-g)y_{t+\tau}) = \ln(1-g) + \ln(y_{t+\tau})$ so that, through the exponential output loss, the consumption utility loss is given by $\partial \ln(c_{t+\tau})/\partial D_{t+\tau} = -\Delta_y$. As there is also the direct utility loss, captured by Δ_u in (6), the full loss in utils at $t+\tau$ is

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

But, part g of the output loss at $t+\tau$ also propagates through savings to period $t+\tau+1$ and further to periods $t+\tau+n$ with $n > 0$, so that the full loss of utils, discounted to time t and denoted by Δ, is given by (7)

2.2 Beliefs

The hazard rate for damages, denoted as p, is the probability that damages start and $I_t = 0$ moves to $I_{t+1} = 1$. The hazard rate is a given constant for each period, but

[11] See Golosov et al. (2011) or Gerlagh and Liski (2013).

unknown to the policy maker. We assume that p has a discrete prior distribution: it can either take value $p = 0$ or $p = \lambda$. The hazard rate can depend on the degree of climate change as measured by D_t, for example, so that only for periods where $D_t > \overline{D} \geqslant 0$ the state can switch. We postpone this extension in Section 3.3, and assume now learning in all periods by setting $\overline{D} = 0$.[12]

There is no prior climate experiment; we do not know the value of p, but we assume a subjective prior probability $\mu_0 > 0$ for a positive hazard rate, $p = \lambda$. The probability for eventual climate impacts satisfy:

$$\begin{aligned} 1 - \mu_0 &= \Pr(\lim_{t\to\infty} I_t = 0) = \Pr(p = 0) \\ \mu_0 &= \Pr(\lim_{t\to\infty} I_t = 1) = \Pr(p = \lambda > 0). \end{aligned}$$

Let μ_t denote the posterior probability that $p = \lambda$, at time t, conditional on no learning by time t, $I_t = 0$. Each period where $D_t > \overline{D} = 0$, but where no damages have appeared so far, $I_t = 0$, climate change runs an experiment. If the outcome is $I_{t+1} = 1$, which happens with probability $\mu_t\lambda > 0$, we have learned that $p = \lambda$, so $\mu_{t+1} = 1$. If the outcome is $I_{t+1} = 0$, we have not learned the state of nature with certainty, but the beliefs are updated to μ_{t+1}. We can write the Bayesian updating rule as[13]

$$\begin{aligned} \mu_t &= \Pr(p = \lambda \,|\, I_t = 0) \\ &= \frac{\mu_0(1-\lambda)^t}{\mu_0(1-\lambda)^t + 1 - \mu_0} \end{aligned} \tag{8}$$

which is the probability that climate change impacts will ultimately arrive even though such damages have not been experienced by time t. Note that μ_t declines over time: "no news is good news"; the assessment of the distribution for damages becomes more optimistic over time.[14] The triple (μ_t, λ, Δ) describes the current beliefs, the underlying stochastic process for damages, and the size of damages, respectively.

[12]For example, $\overline{D}$ can correspond to 2-degrees Celsius warming, but since we have little information about the learning thresholds, we will set $\overline{D} = 0$ in the calibration. The solution of the model can be easily extended to the case of different temperature brackets, all having different hazard rates.

[13]Note that $\Pr(p = \lambda \,|\, I_t = 0) \times \Pr(I_t = 0) = \Pr(p = \lambda \cap I_t = 0)$. The probability that there has been no news by time t is $\Pr(I_t = 0) = \mu_0(1-\lambda)^t + 1 - \mu_0$. The probability that there has been no news by time t and that $p = \lambda$ is $\Pr(p = \lambda \cap I_t = 0) = \mu_0(1-\lambda)^t$. Combining gives the equation.

[14]One could argue that impacts must ultimately arrive for a sufficiently severe climate change. While the model can be extended to include temperature brackets where impacts arrive almost surely, it is also reasonable to think that, for example, a long period of 2-degrees warming without impacts is evidence for not having impacts at such temperatures. Even if one considers "no news is good news" learning to be biased, this bias is consistent with the idea of having a conservative test against the climate policy ramp, as explained in the Introduction.

Variants of the learning dynamics considered here are common in other fields of economics but some features of the setting deserve attention. Malueg and Tsutsui (1997) were among the first to consider learning of unknown Poisson rates in an R&D race; see also, for example, Keller, Rady, and Cripps (2005), and Bonatti and Hörner (2011). In this literature, new information is generated by periodic effort; no current effort means no new information. In climate change, the arrival of new information depends on persistent temperature increases that follow from past actions with a considerable delay. The literature on catastrophic environmental events assumes that the hazard rates for the high-consequence events depend on past actions by making them dependent on variables such as pollution stocks (Clarke and Reed 1994; Tsur and Zemel 1996; and, for example, Polaski, de Zeeuw, and Wagener 2011). However, this literature has not considered uncertainty in the sense that the parameters of the primitive distributions are not known at the outset, as in our case. We connect to this literature in Section 3.3, where belief updating depends on the temperature level.

2.3 Climate dynamics

The temperature response to emissions is a key determinant of the expected present-value utility impacts of the current emissions, that is, the social cost of carbon emissions. For tractable policies, we build on a closed-form for $\mathcal{R}(\tau)$ that is derived in Gerlagh and Liski (2013); see Theorem 1. For exposition, we outline here the two main determinants of the response: the carbon cycle and the relationship between carbon concentrations and temperatures.

The carbon cycle refers to a diffusion process of carbon between reservoirs of carbon, such as those in the atmosphere, oceans and biosphere. Obviously, the atmospheric reservoir is the one relevant for climate warming but the other reservoirs are relevant for the delays and persistencies of changes in the atmospheric stock. Assuming a linear diffusion, the system can be de-coupled by eliminating interactions between the reservoirs, leading to an isomorphic system of separable impulse-responses for carbon stocks (Maier-Reimer and Hasselman 1987). Let $\mathcal{I}$ denote the set of impulse-responses, with fraction $0 < a_i < 0$ of emissions having decay rate $0 \leqslant \eta_i < 1$, $i \in \mathcal{I}$. The shares and decay rates have intuitive meanings, discussed below, and they follow from the physical description of the system of carbon reservoirs.[15]

[15]The true diffusion process is non-linear (Joos et al. 2013). The linear representation is an approximation.

The carbon cycle is relatively well understood in natural sciences but the relationship between temperatures and carbon concentrations is fundamentally uncertain (see, for example, Roe and Baker, 2007). Acknowledging these complications, we note that economic impacts introduce yet another layer of fundamental uncertainty; we focus on this uncertainty and make the following simplistic assumptions on the determinants of the climate equilibrium. Emissions z_t increase the atmospheric CO_2 stock, through the carbon cycle, and there is a linear relationship between the steady state atmospheric CO_2 stock and the steady state level of D_t. This relationship is captured by parameter π: a one-unit increase in the steady-state atmospheric CO_2 stock leads to a π-unit increase in the steady-state level of D_t. Outside steady state, there is a delay in the effect from concentrations to temperatures, and this delay is captured by parameter $0 < \varepsilon < 1$: a one-unit increase in emissions increases the next period CO_2 stocks one-to-one but the direct temperature increase is only $\varepsilon\pi$ -units

Remark 2 *Consider a carbon diffusion process, described by shares* $0 < a_i < 0$ *for depreciation rates* $0 \leqslant \eta_i < 1$, $i \in \mathcal{I}$. *For temperature sensitivity* π *and adjustment speed* ε, *the impact of emissions at time* t *on temperatures at time* $t + \tau$ *is*

$$\frac{dD_{t+\tau}}{dz_t} = \mathcal{R}(\tau) = \sum\nolimits_{i \in \mathcal{I}} a_i \pi \varepsilon \frac{(1-\eta_i)^\tau - (1-\varepsilon)^\tau}{\varepsilon - \eta_i} > 0. \tag{9}$$

The result follows from Gerlagh and Liski (2013, Theorem 1). Parameter η_i captures, for example, the carbon uptake from the atmosphere by forests and other biomass, and oceans. The term $(1-\eta_i)^\tau$ measures how much of carbon z_t under decay i still lives after τ periods, and the term $-(1-\varepsilon)^\tau$ captures the slow temperature adjustment. The limiting cases can be helpful. Consider one CO_2 reservoir. If atmospheric carbon-dioxide does not depreciate at all, $\eta = 0$, then the temperature slowly converges at speed ε to the long-run equilibrium climate sensitivity π, giving $\mathcal{R}(\tau) = \pi[1-(1-\varepsilon)^\tau]$. If atmospheric carbon-dioxide depreciates fully, $\eta = 1$, the temperature immediately adjusts to $\pi\varepsilon$, and then slowly converges to zero, $\mathcal{R}(\tau) = \pi\varepsilon(1-\varepsilon)^{\tau-1}$. If temperature adjustment is immediate, $\varepsilon = 1$, then the temperature response function directly follows the carbon-dioxide depreciation $\mathcal{R}(\tau) = \pi(1-\eta)^{\tau-1}$. If temperature adjustment is absent, $\varepsilon = 0$, there is no response, $\mathcal{R}(\tau) = 0$.

When multiplying temperature measure D_t by given output-loss coefficient $\Delta_y > 0$, we can interpret the emissions-temperature response as an emissions-damage response. Fig. 1 shows the life path of damages (percentage of total output) caused by an impulse of one Teraton of Carbon [TtCO_2] in the first period.[16] The output loss is thus measured

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

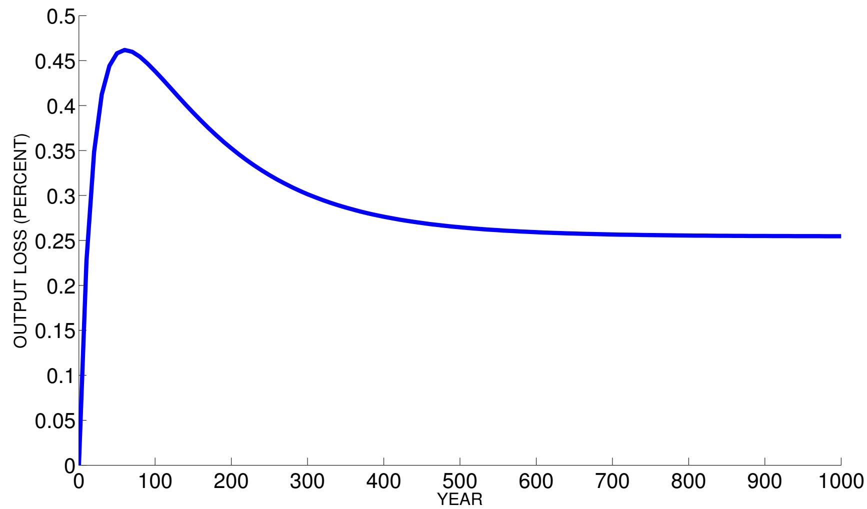

Figure 1: Emissions-damage response. The path depicts the output loss associated with $1TtCO_2$ impulse of carbon at time $t = 0$ for $\Delta_y = 1$ and $\pi = .0156$.

per TtCO_2, and it equals $1 - \exp(-\Delta_y \mathcal{R}(\tau))$, τ periods after the impulse. The non-monotonicity of the response, as depicted in Fig. 1, captures well the climate impact dynamics, for example, in DICE-2007 (Nordhaus, 2008).

The physical data on carbon emissions, stocks in various reservoirs, and the observed concentration developments can be used to calibrate a three-reservoir carbon cycle representation; we choose to the following emission shares and depreciation factors per decade:[17]

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

Thus, about 16 per cent of carbon emissions does not depreciate while about 45 per cent has a half-time of one decade. We assume $\varepsilon = .183$ per decade, implying a global temperature adjustment speed of 2 per cent per year. Normalizing the output loss parameter at unity, $\Delta_y = 1$, and setting $\pi = .0156$ [per $TtC0_2$, see Gerlagh and Liski (2013)] is consistent with the Nordhaus (2008) baseline where a temperature rise of 3 degrees Celsius leads to about 2.7 per cent loss of output.[18] These quantitative choices parametrize the

[17]Some fraction of emissions depreciates within one decade from the atmosphere, and therefore the shares a_i do not sum to unity. The choices here are based on Gerlagh and Liski (2013) but similar representative numbers can be found in the scientific literature; see, e.g., Maier-Reimer and Hasselman (1987).

[18]To clarify the units, the damages are measured per Teraton of CO2 [TtonCO2], and the 3 degrees

emissions-temperature response that is depicted in Figure 1. In the calibration below, we allow Δ_y to be determined by the distribution for the carbon prices obtained from previous studies; throughout, $\Delta_y = 1$ refers to the Nordhaus' baseline.

3 General-equilibrium policies

3.1 After learning, $I = 1$

To obtain the carbon price, that is, the social cost of current carbon emissions z_t, consider the effect of emissions at t on a stream of future utilities. The full loss of utils per increase of temperatures as measured by $D_{t+\tau}$, caused by z_t at time t, when discounted to t with factor $0 < \delta < 1$, is denoted by h. It follows with the aid of (7) and (9):

$$\begin{aligned} h &\equiv -\sum_{\tau=1}^{\infty} \delta^{\tau} \frac{du_{t+\tau}}{dz_t} \\ &= \Delta \sum_{\tau=1}^{\infty} \delta^{\tau} \frac{dD_{t+\tau}}{dz_t} = \Delta \sum_{\tau=1}^{\infty} \delta^{\tau} \mathcal{R}(\tau) \qquad (10) \\ &= \Delta \sum_{i \in \mathcal{I}} \frac{a_i \pi \varepsilon}{\varepsilon - \eta_i} \sum_{\tau=1}^{\infty} \delta^{\tau} (1-\eta_i)^{\tau} - \delta^{\tau} (1-\varepsilon_j)^{\tau} \\ &= \delta \Delta \pi \frac{\varepsilon}{1-\delta(1-\varepsilon)} \sum_{i \in \mathcal{I}} \frac{a_i}{1-\delta(1-\eta_i)}. \qquad (11) \end{aligned}$$

The present-value utility costs of current emissions can thus be compressed to a number, h, that will be an input to the determination of the currently optimal carbon price. The first term, $\delta\Delta\pi$, describes the utility loss associated with one emission unit when steady state damages would happen immediately at the next period. The second term discounts damages because of the time-delay associated with temperature adjustment. The third term with the summation describes the persistence of damages as the atmospheric CO_2 stock decays slowly.

Proposition 1 *Conditional on $I_t = 1$, the optimal carbon price is*

$$\tau_t = \frac{\partial y_t}{\partial z_t} = (1-g) y_t \delta \Delta \pi \frac{\varepsilon}{1-\delta(1-\varepsilon)} \sum_{i \in \mathcal{I}} \frac{a_i}{1-\delta(1-\eta_i)}. \qquad (12)$$

Thus, the optimal carbon price in (12) is proportional to income, with proportionality depending only on δ, Δ, and the carbon cycle parameters in (9). Given loss parameter Δ,

Celsius rise follows from doubling the CO_2 stock. We have chosen the value of π such that the normalization $\Delta_y = 1$ gives the Nordhaus case. For this reason, the interpretation of π is "climate damage sensitivity" rather than "climate sensitivity".

the same tax is optimal for any division between utility and production losses satisfying (7).

For the proof, given the Brock-Mirman structure (1972), the payoff implications of temperature changes are separable from capital capital wealth. The climate policy can be found by balancing the present-value of future utility costs of emissions (11) with the current utility-weighted marginal product of carbon: $\frac{\partial y_t}{\partial z_t}\frac{\partial u}{\partial c} = h$. Since $\frac{\partial u_t}{\partial c_t} = 1/c_t = 1/(1-g)y_t$, we can express the optimal carbon price as

$$\tau_t = \frac{\partial y_t}{\partial z_t} = h(1-g)y_t$$

which gives the result.

Let us comment on the property that the optimal tax is proportional to income. This follows since, effectively, through the Brock-Mirman structure we assume a unit elasticity of losses with respect to income, which represents an intermediate position in the literature. Some economic climate-change losses, such as decreased agricultural yields in tropical areas, are likely to increase less than one-to-one with income, as the share of the agricultural sector tends to decrease when income grows. At the same time, as these agricultural impacts are expected to be more severe in the currently warm-climate and less-developed countries, the share of damages in world-wide income will increase when those economies grow at rates larger than the world-wide average growth rate. Also, the monetary evaluation of economically intangible impacts such as ecological losses are expected to increase more than proportionally with income (Mendelsohn, Dinar and Williams 2006; Mendelsohn et al 2012).

3.2 Carbon price distribution before learning, $I = 0$

Once damages appear, the policies can be determined exactly as in Proposition 1. Prior to their appearance, the model generates a parametric distribution for the time when damages occur. Let Z be the stochastic variable, measuring the full future utility cost from increasing current emissions z_t marginally. Let $h_t = \mathbb{E}_t Z$ be the expected present value of future utility losses associated with one unit of current emissions. Z can take the values $Z_1, Z_2, ...,$ where Z_τ is the current social cost of carbon if damages appear for the first time, precisely at period $t + \tau$. Thus, Z_τ characterizes the present-value marginal utility losses from current emissions z_t, assuming that the damage indicator I_t remains at zero for all periods prior to $t + \tau$ but then turns positive. Proceeding as in Section 3.1, and using the emissions-temperature response from Section 2.3, we can obtain the present-value of such delayed utility losses in closed-form:

$$
\begin{aligned}
Z_\tau &= \Delta \sum\nolimits_{s=\tau}^{\infty} \delta^s \mathcal{R}(s) \\
&= \Delta \sum\nolimits_{i\in\mathcal{I}} \frac{\pi a_i \varepsilon}{\varepsilon - \eta_i} \delta^\tau \left(\frac{(1-\eta_i)^\tau}{[1-\delta(1-\eta_i)]} - \frac{(1-\varepsilon)^\tau}{[1-\delta(1-\varepsilon)]} \right).
\end{aligned}
$$

Given our model of learning, we find for the distribution of Z that

$$
\Pr(Z = Z_\tau | I_t = 0) = \Pr(I_\tau = 1 \cap I_{\tau-1} = 0 | I_t = 0)
$$

which gives the probability that damages turn positive exactly after τ periods when the current time t subjective belief for the climate problem is μ_t. To find the corresponding cumulative distribution function for the utility losses, denoted by $F_t(Z)$, we first establish the probability that the damage has revealed itself at period t, irrespective of if t is the first time:

$$
\begin{aligned}
\Pr(I_t &= 1) = (1-\mu_0)\Pr(I_t = 1|p=0) + \mu_0 \Pr(I_t = 1|p=\lambda) \\
&= \mu_0[1 - \Pr(I_t = 0|p=\lambda)] \\
&= \mu_0[1 - \Pr(I_1 = ... = I_t = 0|p=\lambda)] \\
&= \mu_0[1-(1-\lambda)^t].
\end{aligned}
$$

We can generalize this to expectations at period t,

$$
\Pr(I_{t+\tau} = 1 | I_t = 0) = \mu_t[1-(1-\lambda)^\tau]
$$

so that the distribution for Z is then given by

$$
\begin{aligned}
F_t(Z_\tau) &= \Pr(Z \le Z_\tau | I_t = 0) = \Pr(I_{t+\tau-1} = 0 | I_t = 0) \\
&= 1 - \mu_t + \mu_t(1-\lambda)^{\tau-1}.
\end{aligned} \tag{13}
$$

We can use this distribution to determine the social cost of carbon at time t as dependent on beliefs μ_t.

Theorem 1 *Conditional on no experience of impacts by time t ($I_t = 0$), the previous-period distribution of the social cost of carbon $F_{t-1}(Z)$ stochastically dominates the current distribution $F_t(Z)$. The social cost of carbon as measured by $h_t = \mathbb{E}_t Z$ declines over time conditional on $I_t = 0$. Moreover,*

$$
\begin{aligned}
h_t &\equiv \mathbb{E}_t Z = \sum\nolimits_{\tau=1}^{\infty} \delta^\tau \mathbb{E}_t \frac{du_{t+\tau}}{dz_t} = \mu_t h^l \\
h^l &\equiv \delta\Delta\pi \frac{\varepsilon}{1-\delta(1-\varepsilon)} \sum\nolimits_{i\in\mathcal{I}} \frac{a_i}{1-\delta(1-\eta_i)} \\
&\quad -\delta(1-\lambda)\Delta\pi \frac{\varepsilon}{1-\delta(1-\lambda)(1-\varepsilon)} \sum\nolimits_{i\in\mathcal{I}} \frac{a_i}{1-\delta(1-\lambda)(1-\eta_i)}.
\end{aligned}
$$

Proof. The expected utility losses from current emissions are equal to

$$\begin{aligned} h_t &= \mathbb{E}_t \Delta \textstyle\sum_{\tau=1}^{\infty} \delta^{\tau} I_{t+\tau} \dfrac{dD_{t+\tau}}{dz_t} \\ &= \Delta \textstyle\sum_{\tau=1}^{\infty} \delta^{\tau} \Pr(I_{t+\tau} = 1 | I_t = 0) \mathcal{R}(\tau) \\ &= \mu_t \Delta [\textstyle\sum_{\tau=1}^{\infty} \delta^{\tau} \mathcal{R}(\tau) - \sum_{\tau=1}^{\infty} (1-\lambda)^{\tau} \delta^{\tau} \mathcal{R}(\tau)]. \end{aligned}$$

Using our temperature-response function leads to the expression for h_t. Decreasing carbon prices measured in utils and stochastic dominance follow from (13) and μ_t decreasing over time. ■

The result gives a closed-form expression for the optimal carbon price policy depending both on the climate system parameters and on the current belief of the damage distribution characterized by (μ_t, λ, Δ). The first term that defines h^l equals h, the full information policy variable (defined in (11). The second term subtracts the present value of damages that in expectations do not occur, substituting $\delta(1-\lambda)$ for the discount factor.

Recall that the optimal general-equilibrium carbon price is the income-weighted future utility-cost of current actions, analogous to (12), giving:

Proposition 2 *The optimal learning-adjusted carbon price is*

$$\tau_t = \mu_t h^l (1-g) y_t. \tag{14}$$

The results follows from the same arguments as for the full information case, using Theorem 1 for the expected future utility-costs.

The "climate policy ramp", that is, the gradually tightening carbon-price policy over time, can follow even with increasing climate optimism over time: despite the declining μ_t, sufficient growth of income growth y_t, implies that the economy becomes, in expected terms, more exposed to losses from climate change.

Limiting cases reveal the mechanisms at work. Consider time $t = 0$, where the subjective belief of damages is given by $\mu_0 < 1$. If damages are almost surely observable, $\lambda \nearrow 1$, the optimal initial policy prior to experimentation is the full information policy, weighted with the subjective probability for damages, $h^l \to h$. However, if damages do not appear the next period, $I_1 = 0$, then the subjective assessment μ_1 drops to zero by the updating rule (8) as beliefs become very optimist, and the carbon price drops to zero, $\mu_1 \searrow 0, h_1 = \mu_1 h^l \searrow 0$. In this case, no news reveals the true climate-economy state precisely. On the other hand, if climate change damages are not easily observable, $\lambda \searrow 0$, climate change is a problem with a non-significant rate of appearances in all cases and carbon prices are low, $h^l \searrow 0$. But this case also implies that climate experiments are not very informative; there will be no learning, and the subjective assessment μ_t in (8) remains almost unchanged over time.

The carbon price formula developed here differs from that in Golosov et al. (2011), who also build on the Brock-Mirman consumption choice framework for climate change impacts, in two main ways. First, our formula incorporates a delayed response of temperatures to atmospheric CO_2, without losing tractability. Golosov et al. assume that the temperature and associated potential impact of emissions reaches its maximum immediately after the date of emissions, which is hard to reconcile with the carbon cycle representations of the applied models typically used for carbon pricing.[19] Second, we introduce a structure for beliefs and their tractable updating (8) so that the carbon price has a closed form and the contributions of beliefs and income become explicit in (14). Moreover, we will exploit in a following section the closed-form distribution of Z in (13) to connect the quantitative assessment to carbon price estimates in the literature.

3.3 Learning thresholds

Before moving to the quantitative assessment, consider the case the degree of climate change determines the intensity of experimentation. Suppose learning takes place only above a temperature threshold, $D_t \geq \overline{D}$, corresponding, for example, to 1 or 2 degrees Celsius above the pre-industrial temperature levels.

Proposition 3 *Assume that temperatures generate information on impacts only if $D_t \geq \overline{D}$. Let $D_0 < \overline{D}$ and $t' < \infty$ be the first period such that $D_{t'} \geqslant \overline{D}$. Then, prior to t', the expected present-value utility impact of emissions increases over time: $h_t < h_{t+1}$ for $0 < t < t'$.*

[19] Gerlagh and Liski (2013) compare the emissions-damage responses of DICE-2007, Golosov et al. 2011, and the one presented here. Moreover, the supplementary material of that paper contains a note that illustrates the importance of the non-monotonicity of the response in replicating the carbon price predictions of the applied climate-economy models.

Proof. Let T be the set of periods τ such that $D_\tau \geqslant \overline{D}$. The expected utility losses for $0 < t < t'$ satisfy

$$\begin{aligned}
h_t &= \mathbb{E}_t \Delta \textstyle\sum_{s=1}^{\infty} \delta^s I_{t+s} \dfrac{dD_{t+s}}{dz_t} \\
&= \Delta \sum_{\tau \in T} \left\{ \Pr(I_\tau = 1 \cap I_{\tau-1} = 0 | I_t = 0) \textstyle\sum_{s=\tau-t}^{\infty} \delta^s \dfrac{dD_{t+s}}{dz_t} \right\} \\
&= \Delta \sum_{\tau \in T} \left\{ \Pr(I_\tau = 1 \cap I_{\tau-1} = 0 | I_t = 0) \textstyle\sum_{s=\tau-t}^{\infty} \delta^s \mathcal{R}(s) \right\} \\
&< \Delta \sum_{\tau \in T} \left\{ \Pr(I_\tau = 1 \cap I_{\tau-1} = 0 | I_t = 0) \textstyle\sum_{s=\tau-t-1}^{\infty} \delta^s \mathcal{R}(s) \right\} \\
&= \Delta \sum_{\tau \in T} \left\{ \Pr(I_\tau = 1 \cap I_{\tau-1} = 0 | I_{t+1} = 0) \textstyle\sum_{s=\tau-(t+1)}^{\infty} \delta^s \mathcal{R}(s) \right\} \\
&= \Delta \sum_{\tau \in T} \left\{ \Pr(I_\tau = 1 \cap I_{\tau-1} = 0 | I_{t+1} = 0) \textstyle\sum_{s=\tau-(t+1)}^{\infty} \delta^s \dfrac{dD_{t+1+s}}{dz_{t+1}} \right\} \\
&= \mathbb{E}_{t+1} \Delta \textstyle\sum_{s=1}^{\infty} \delta^s I_{t+1+s} \dfrac{dD_{t+1+s}}{dz_{t+1}} \\
&= h_{t+1}
\end{aligned}$$

The second line follows because $I_t = 0$ with certainty for $0 < t < t'$. The inequality follows as we subtract one period to take one period of delay away. The fifth line follows as beliefs do not change between t and $t+1$. ■

As long as no information can be obtained, no damages will occur but policy h_t becomes more strict over time as the expected first appearance of damages comes closer. The tightening of policies continues until the temperatures start generating information.

Proposition 4 *For $0 < t < t'$, defined in Proposition 3, the optimal carbon tax grows faster than the economy.*

Since the actual carbon tax is a multiple of income, the tax implied by h_t for $D_t < \overline{D}$ will be growing over time at a rate exceeding the growth of the economy, by Proposition 3. Further, recall that our emissions-temperature response implies that the temperature peak for a given emissions impulse lags 60-70 years behind the date of emissions: the learning of effects described here can start several decades *after* the emissions that caused climate change to break through the threshold. Meanwhile, optimal policies are characterized by constant beliefs, but by potentially sharply increasing carbon prices.

The shape of the emissions-temperature response, $\mathcal{R}(\tau)$ is thus not only important as a measure of the development over time for the potential shock on the economy; it also dictates how quickly the climate experiment can become informative. The result above can be extended to a more sophisticated representation of arrival rates, depending on the temperature level. However, in the interest of designing a conservative test for the climate policy ramp, we assume that any level of temperature increase allows learning of the climate impacts on the economy in the quantitative analysis below. Moreover, our approach ignores the possibility of learning the climate impacts from the shorter-term temperature volatility (see, for example, Kelly and Tan 2013); in the Appendix, we extend the tractable carbon price formula to this case.

4 Quantitative assessment

Throughout the quantitative analysis, we assume 10-year periods; the first year is '2010' corresponding to period 2006-2015. We assume only (potential) output losses from climate change so that $\Delta_y = \Delta$ and $\Delta_u = 0$, to maintain an easy comparison with earlier studies. We take the Gross Global Product as 600 Trillion Euro [$Teuro$] for the decade, 2006-2015 (World Bank, using PPP). Throughout we assume a capital share of $\alpha = .3$ and one per cent pure rate of annual time preference, implying $\delta = .90$ for decadal periods and resulting in savings $g = .27$.

Normalizing the output loss parameter at unity, $\Delta = 1$, and assuming that these losses exist at the outset gives us results comparable to Nordhaus' (2007) baseline. Together with our carbon cycle, such damages result in a carbon price of 22 EUR/tCO_2, equivalent to about 105 USD/tC, for 2010. This estimate is higher than the Nordhaus baseline (2007) because of our lower pure rate of time preference that facilitates the calibration presented in Section 4.1.[20]

4.1 Matching carbon price distributions

For an informed approach to quantifying the belief component in the model, we use now a distribution of existing carbon price estimates as external data. The underlying idea in this, admittedly unorthodox, calibration is that each number in the data presents a point estimate of the social cost. Our model gives a structural interpretation for the dispersion of the estimates, allowing calibration of the parameters that quantify the initial beliefs

[20]Note that 1 tCO2 = 3.67 tC, and 1 Euro is about 1.3 USD. Our number 105 USD/tC is almost precisely equal to the DICE-2007 carbon price when in that model the elasticity of substitution parameter is set to one and the pure rate of time preference is set to 1 per cent per year, as in our analytical model. The number appearing in Nordhaus (2007), that is $35USD/tC$, can be matched by setting 2.7 per cent pure rate of time preference. However, Tol's data, used in the calibration below, does not exist for this value of time preference.

in the model. The policy-maker, that is, the decision-maker in the model, then forms one initial point estimate for the social cost, and evaluates its evolution over time given the learning dynamics assumed.[21]

Tol (2009) conducted a comprehensive survey of the existing estimates for the social cost of carbon. From a sample of 232 estimates he derived a distribution for the carbon price measured in 1995 USD/tC, controlling for the time discount rates used in the studies. We focus on Tol's sample corresponding to 1 percent pure rate of time preference.[22] Tol's mean value for the carbon price is 32.7 for 2010 EUR/tCO_2 (his Table 2, 2009). We calibrate the climate system parameters as reported in Section 2.3 and choose economic parameters as above, and then fit our cumulative damage distribution function $F(Z)$ by choosing the initial prior μ_0, the hazard rate λ, and the damage parameter Δ. Note that in this interpretation of the data, the heterogeneity in the point estimates comes from different possible outcomes for the arrival date of the damage.

Fig. 2 depicts a spline connecting the 33, 50, 67, 90 and 95 percentiles of the carbon price distribution, expressed in 2010 EUR/tCO_2, as reported by Tol, jointly with the distribution that follows from our calibration, depicted as a smooth line. We can match the two cumulative distributions either by minimizing the errors at the reported percentile points, or, more directly, by matching the means and the end-points of the distributions. The approaches are almost outcome-equivalent. We followed the latter approach to allow for the interpretation set out below.

There is a mass point at zero, corresponding to a 20 per cent assessment that insignificant or positive climate change impacts will occur.[23] For interpretation, we may think that $1 - \mu_0$ represents the share of climate experts having the assessment that climate-change impacts will be negligible or even positive; to match the lower end of the distribution, we set $\mu_0 = .8$.

In the other extreme, there are experts who have strong views that income losses are

[21]The social cost of carbon is an elusive concept in the applied work that has generated the data discussed below. Many of the studies do not optimize to find the optimal shadow value of the current carbon constraint; rather, the cost of carbon is the evaluated cost from a marginal increase of emissions given a background scenario for the economy; see, for example, the model descriptions in the Stanford Energy Modeling Forum (in Weyant et al, 2006). Our planner optimizes the social cost which, obviously, differs from the non-optimized estimates but is not necessarily inconsistent with them.

[22]Tol reports distributions for 0, 1 and 3 per cent discount rates, respectively. Our analysis of the 3-percent case produced very similar qualitative results; the levels of the policy variables are systematically lower.

[23]This number we inferred from Tol (2008).

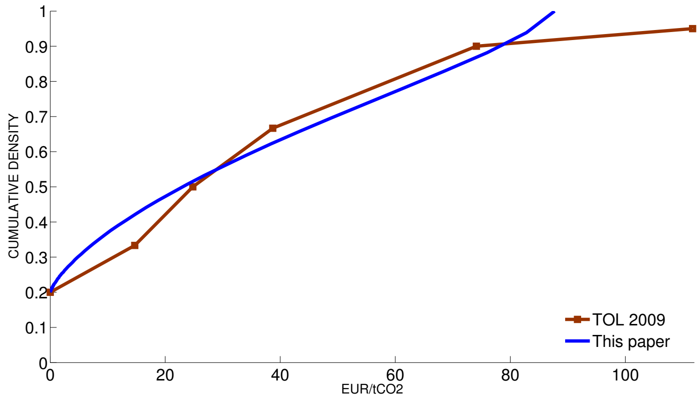

Figure 2: Fitting cumulative distribution with Tol's (2009) distribution.

high, arrive almost surely and soon: the high end of the carbon prices pins down the damages conditional on bad news, as captured by the value Δ_y. To avoid giving too much weight to a few extreme cost estimates in the sample, we truncated the fitted distribution at 87.7 EUR/tCO_2 by setting $\Delta_y = 4$. That is, maximum damages are by factor four higher than the middle-of-the-road damages assumed in Nordhaus (2007) — the implied output loss is then about 10.7 per cent from doubling the CO_2 stock, if climate impacts materialize.

The continuum of views between the extremes are described through the third parameter, λ. We obtain the value $\lambda = .077$ such that the initial carbon price implied by our model exactly matches Tol's mean value of 32.7. Choice $\lambda = .077$ means that information is generated very slowly – there is about 8 per cent probability of learning per decade. A geometric distribution with this arrival rate per decade means that the expected arrival time for a severe climate change damage event is about 130 years. After 100 years without damages, the posterior for the eventual impact arrival μ_t is still 64 per cent.

The matching of our carbon price distribution with Tol's distribution ensures consistency between our quantitative assessment of the first-period social cost of carbon and the views as held by the profession; updating of the distribution depends on our structural interpretation of the learning process.

4.2 The climate-policy ramp

The above calibration sets the optimal initial carbon price at the mean in Tol's survey: 32.7 EUR/tCO_2 in 2010. Consider now the development of the optimal carbon price over time. We set $\overline{D} = 0$, assuming that any level of temperature increase produces information.

Given the closed-form formula for the carbon price in (14), one approach is to conjecture future output or income levels, say, in 2050 and, conditional on no observed impacts by that time, obtain the future carbon price for that state of the world. However, future states of the world result partly from past policy decisions; carbon pricing decisions shape current, and through investments, also future income levels. To obtain consistent policy scenarios, in the Appendix we introduce a structure for $A_t(z_t)$ in the production function to describe the two main mechanisms through which the economy adjusts to carbon policies: energy savings that typically feature the early decades of the adjustment, and then decarbonization of energy that is needed to meet the long-run climate targets.

The benchmark for our assessment is the "Climate policy ramp" (dotted line in Fig. 3), based on Nordhaus' DICE (2007) middle-of-the-road damage estimate, corresponding to $\Delta_y = 1$ sure-loss damages; that is, damages are immediately observed with no uncertainty. For 2010, with 1 per cent annual pure time-discounting and log-utility, the benchmark sure-loss policy path gives 22 EUR/tCO_2 as the optimal price which is almost identical to what DICE produces under this choice for discounting and preferences.[24] This middle-of-the-road sure-loss path involves a tightening of the policies over the coming century, typical for most no-uncertainty climate-policy assessments.

We now look at the optimal time path for the carbon price for *high potential damages, but conditional on not observing these damages*; that is, we consider the evolution of the policy when future impacts are potentially severe, $\Delta_y = 4$, as determined by the calibration procedure above, but when no news on climate impacts arrive. Then, we compare this policy path to the baseline. Without impacts, the economy is unaffected by climate change but, since the carbon policies are in place, emissions and output will be reduced below the business-as-usual path. The optimal carbon price is depicted as a solid line in Fig. 3 over the coming century and beyond. The two climate policies —one with immediate damages based on the central estimate, and the other with high but only potential damages and gradual updating of beliefs to the no-news situation— have the same shape for the first century.[25] The main result of the quantitative assessment follows: policies should become tighter over time even if climate optimism increases. Strikingly, for this particular learning scenario, it takes close to 200 years without observed climate

[24]Illustration available on request.

[25]The difference in levels follows since the baseline estimate by Nordhaus is close to the median, but lower than the mean in Tol's distribution.

damages for beliefs to become optimistic enough for the carbon price to decline – the social cost of carbon declines very slowly.

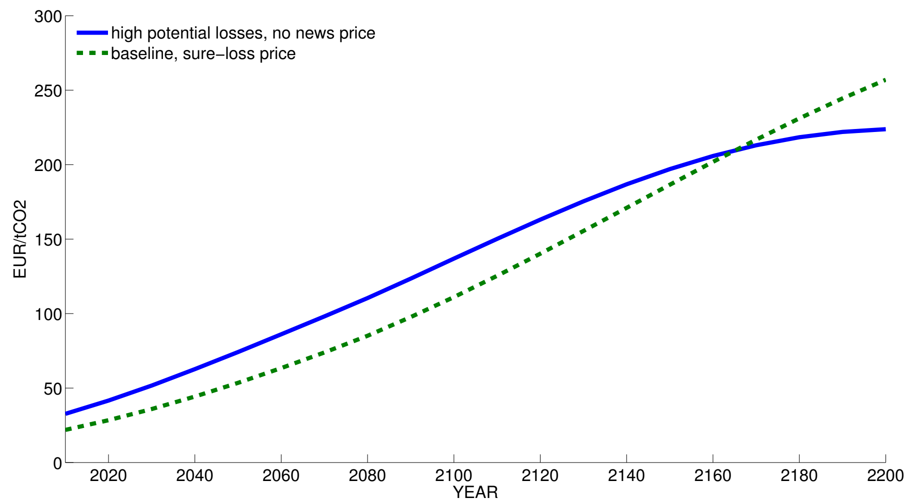

Figure 3: The carbon price for a sure income loss of 2.7 per cent from doubling the carbon concentration ($\mu = 1; \Delta_y = 1$), and for uncertain damages, conditional on no news on damages ($\mu_0 = 0.8; \lambda = 0.077, \Delta_y = 4$).

To assess the shape of the carbon price path, we decompose its level into its two main components. Recall that the optimal carbon price is proportional to h_t capturing the expected utility losses from current emissions, and to income y_t; $\tau_t = h_t(1-g)y_t$, $h_t = \mu_t h^l$. See Table 1, for the contribution of income (y_t) and learning (μ_t) to the carbon price.[26] Expected income growth is prodigious; in our evaluation, based on the IPCC scenarios (see Appendix), income rises five-fold during the coming century. Such an estimate is not unheard of, and is driven by an increasing population and the rise of the middle class in emerging economies. The development of beliefs is captured through μ_t in the Table. Observing no major climate damages over the coming century, leads to substantial increase in optimism, but, as is evident from the Table, it is the changing scale of the global economy that matters for the development of carbon pricing. The stake affected by the potential inverse income shocks from climate change increase so

[26] It is illuminating to consider the units of measurement for the utility loss measure $h_t = \mu_t h^l$, which has the same unit as the constant in the legend of the table: years per emissions. The variable h_t measures the life-time equivalent of welfare that is lost per unit of emissions. For the year 2010, annual emissions are about .04 TtonCO_2, implying $.75 \times .04 = .03$ years of expected life-time destroyed by these emissions.

much that stabilizing carbon prices at the initial level — thus ruling out a climate policy ramp completely — would require that the climate experiment is by orders of magnitude more informative than considered here. The assessment of the probability of major utility losses, as captured by μ_t, would need to decline by 50 per cent by 2050.

	income [T€/yr]	beliefs, μ_t [.]	carbon price [€/tCO$_2$]
2010	60	.80	33
2050	146	.74	74
2100	304	.66	137
2150	510	.57	197
2200	703	.47	224

Table 1: Decomposing the contribution of income and learning to the carbon price. Multiplying the first column and the second column, with a constant $h^l(1-g) = 0.68$ $[yr/TtCO_2]$, gives the last column.

Such carbon prices imply substantial value. Current CO_2 emissions exceed 30 Gton annually, while the annual world output is about 60 trillion euro. A carbon price of 100 EUR/tCO_2, worldwide, then represents about 5% of the value of the output, but such a value share is unlikely to be reached along the path as emissions decline in response to higher carbon prices; even without carbon prices, emissions tend to increase less than proportionally to output. But, then, can carbon prices continue to rise as emissions go down and the climate returns to its natural state? To answer this question consider the persistence of atmospheric carbon, as shown in Figure 1. Current atmospheric CO_2 concentrations are about 400 particles per million (ppm), 125 ppm above the pre-industrial levels of 275 ppm. Even when CO_2 emissions fall to zero before 2100, it is expected that atmospheric CO_2 concentrations will not drop below 400 ppm before the end of the century, and stay above 350 ppm for centuries to come. In that sense, the climate is not expected to return to its natural state for a long period, and carbon prices continue to rise with income. For this reason, in our calibration of the energy sector (in the Appendix), the economy becomes decarbonized during the coming century; the combination of rising incomes and the persistence of carbon concentrations leads to an Environmental Kuznets Curve for the observables of the economy.

4.3 Bad news

It may seem surprising that carbon prices under learning, as depicted in Fig. 3, reach such high levels, despite no actual damage taking place. Obviously, in addition to the income growth development, the persistent tightness of the climate policy is supported by the possibility of real damages that may arrive at any time period, and, if damages arrive, the historical emissions have persistent real impacts on future utilities. The bad news carbon price captures the economic meaning of the threat; it is the carbon price at time t that would be socially optimal if bad news arrived at time t. Fig. 4 depicts both the no news and bad news carbon price path for the near and longer terms. Note that the bad news price path is "virtual" because it is drawn against the economy that does not, but could, experience the damage, that is, output has not dynamically adjusted. The starting level is given by our calibration at 87.7 EUR/tCO_2, as this is the highest price estimate that we pulled from Tol's survey and applied to the immediate arrival of impacts. The virtual price increases for a long period of time reflecting the expanding world economy.

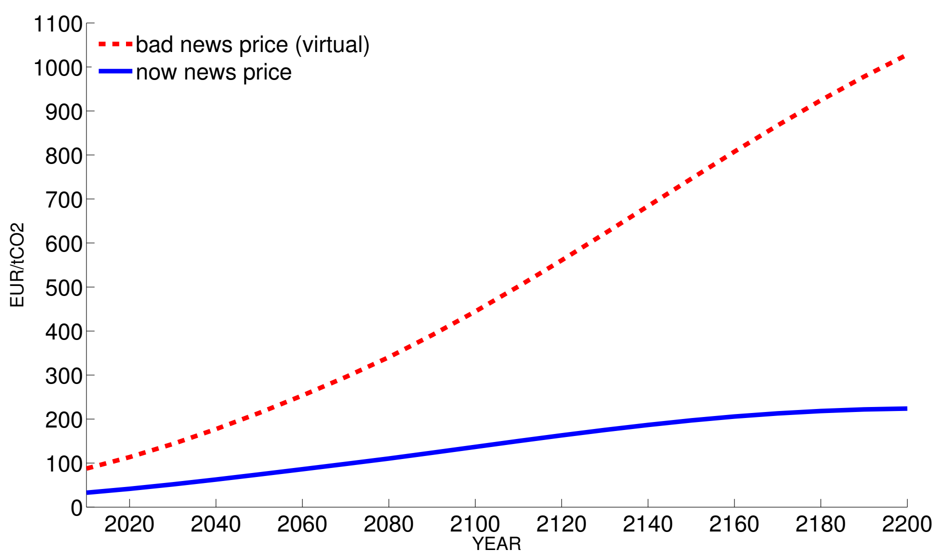

Figure 4: Bad news and the no news carbon price

5 Concluding Remarks

We developed a tractable climate-economy model that allows a stylized but transparent and self-contained quantitative assessment of the optimal carbon price when the impacts of climate change can only be learned gradually over time. Rather than producing another estimate for the carbon price, we took the distribution of the existing estimates as given and provided a structural interpretation for it, stemming from the strong subjective components in the estimates. The current paper is the first attempt to use the current estimates in a quantitative assessment of the sensitivity of policies to the fact the dispersion of views may converge slowly over time. The optimal carbon pricing policies building on the current estimates is robust to significant delays in obtaining hard evidence on the socio-economic impacts of climate change; it is the size of the economy that drives the carbon price.

Since there is a pressing policy need for a meaningful estimate of the carbon price, it is important that the framework is detailed enough for replicating the more comprehensive applied climate-economy models that, despite their shortcomings, are currently used for regulatory purposes such as those reported in Greenstone et al. (2011). Other analytical approaches are often more stylized (e.g., Weitzman 2009), and as such provide valuable insights but cannot directly contribute to the quantitative determination of the optimal policies. Our model, while still very stylized, has a tractable emissions-damage response building on the insights from the natural science literature that, when combined with the macro-economic approach of Golosov et al. (2011), enables us to construct a transparent quantitative policy tool, with explicit component for beliefs.[27]

The closed-form approach to policies and their calibration to the distribution of estimates is very different from the recent numerical approaches to uncertainty and learning in climate change (Lemoine and Traeger, 2014; Cai et al. 2013). The tractability can be maintained while extending the current simplistic approach to learning that allows a gradual arrival of information, say, learning of economic losses from extreme weather events. In the Appendix, we show how the analysis can be extended to this direction while keeping the main results: a tractable carbon price distribution that allows calibration and that features similar qualitative implications for belief updating over time. More precisely, in the Appendix we introduce random productivity shocks where we cannot, before experimenting, tell apart persistent climate impacts from temporary shocks. Using a standard normal learning approach about the underlying climate-economy fun-

[27]The supplementary material of Gerlagh and Liski (2013) includes a note that evaluates numerically the deviation of our carbon price prediction (in the absence of uncertainty and learning) from the prediction produced by Nordhaus-DICE (2007). We generate data for the DICE carbon price by sampling the key model parameters. Our carbon price formula, that uses only a subset of the sampled parameters, can explain 99 per cent of the variation in the DICE carbon price.

damental, we can reproduce the policies of the simple framework in the text for a case that allows a richer set of observables. We find that, when observed damages accord with the prior median damage estimate, learning tends to lower the carbon price; but the rise of income increases the carbon price much in the same way as here.

The main observation that follows from the the carbon price formulas and their quantitative assessment is novel and likely to hold in more general settings: the trend in the optimal carbon price path is mostly driven by the expansion of the global economy and the resulting growth of the potential expected economic losses. Changes in impact assessments will likely have smaller impacts on optimal carbon policies.

References

[1] Acemoglu D., Aghion P., Bursztyn L., Hemous D (2012). The environment and directed technical change, American Economic Review 102, 131–166.

[2] Bonatti, A., and J. Hörner(2011). Collaborating, American Economic Review, 101(2): 632-63.

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

[4] Cai, Y, K. L. Judd, and T. S. Lontzek, (2013) The social cost of stochastic and irreversible climate change, NBER working paper 18704.

[5] Chichilnisky, G., G.M. Heal, and D. Starrett (1993), International Emission Permits: Equity and Efficiency, Columbia University, Discussion Paper No. 686.

[6] 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.

[7] Deschenes, O., and M. Greenstone. 2007. "The Economic Impacts of Climate Change: Evidence from Agricultural Output and Random Fluctuations in Weather." American Economic Review, 97(1): 354-385.

[8] De Groot, M. Optimal Statistical Decisions. New York: McGraw-Hill, 1970.

[9] Dell, M., B. F. Jones, and B. A. Olken. 2012. "Temperature Shocks and Economic Growth: Evidence from the Last Half Century." American Economic Journal: Macroeconomics, 4(3): 66-95.

[10] Fischer C. and R. Newell, 2008, Environmetnal and technology policies for climate mitigation, Journal of Environmental and Economic Management 55:142-162.

[11] Gerlagh, R., and M. Liski,Carbon Prices for the Next Thousand Years (2013). Available at http://hse-econ.fi/liski/ Longer working paper: Carbon Prices for the Next Thousand Years (June 28, 2012). CESifo Working Paper Series No. 3855.

[12] Gerlagh, R., S. Kverndokk, and K. E. Rosendahl. 2009. Optimal Timing of Climate Change Policy: Interaction between Carbon Taxes and Innovation Externalities. Environmental and Resource Economics 43(3): 369-390.

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

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

[15] Grossman G,.M. and A.B. Krueger, "Environmental Impacts of a North American Free Trade Agreement," in P. Garber, ed., The U.S.-MexicoFree TradeAgreement(Cambridge, MA: MIT Press, 1993), 13-56.

[16] Grossman G,.M. and A.B. Krueger (1995). "Economic Growth and the Environment," The Quarterly Journal of Economics, MIT Press, vol. 110(2), pages 353-77, May.

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

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

[19] Heal, G., A. Millner (2013), uncertainty and decision in climate-change economics, NBER WP 18929.

[20] Holtz-Eakin, D. and T. M. Selden (1995), Stoking the Fires? Co2 Emissions and Economic Growth, Journal of Public Economics, vol. 57, pp. 85-101.

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

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

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

[24] IPCC (2007), Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, Solomon, S., D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M. Tignor and H.L. Miller (eds.) Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

[25] Joos, F. and Roth, R. and Fuglestvedt, J. S. and Peters, G. P. and Enting, I. G. and von Bloh, W. and Brovkin, V. and Burke, E. J. and Eby, M. and Edwards, N. R. and Friedrich, T. and Frölicher, T. L. and Halloran, P. R. and Holden, P. B. and Jones, C. and Kleinen, T. and Mackenzie, F. and Matsumoto, K. and Meinshausen, M. and Plattner, G.-K. and Reisinger, A. and Segschneider, J. and Shaffer, G. and Steinacher, M. and Strassmann, K. and Tanaka, K. and Timmermann, A. and Weaver, A. J. (2013), Carbon dioxide and climate impulse response functions for the computation of greenhouse gas metrics: a multi-model analysis, Atmospheric Chemistry and Physics Discussions, 13, 2793—2825.

[26] Keller, G., and S. Rady, and M. Cripps (2005), Strategic experimentation with exponential bandits, Econometrica, Vol. 73, No. 1, 39-68.

[27] Kelly, D.L., and C.D. Kolstad (1999), Bayesian learning, growth, and pollution, Journal of economic dynamics and control 23, 491-518.

[28] Kelly, D.L., and Z. Tan (2013), Learning and Climate Feedbacks: Optimal Climate Insurance and Fat Tails, working paper.

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

[30] Leach, A. J. (2007), The climate change learning curve, Journal of Economic Dynamics and Control 31, 1728 -1752.

[31] Lemoine, D. M. and Traeger, C. P. (2014), Watch Your Step: Optimal Policy in a Tipping Climate, American Economic Journal: Economic Policy 6(1): forthcoming.

[32] Malueg, D., and S. O. Tsutsui (1997), Dynamic R&D Competition with Learning, The RAND Journal of Economics , Vol. 28, No. 4, pp. 751-772.

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

[34] Mendelsohn R., A. Dinar and L. Williams (2006), The distributional impact of climate change on rich and poor countries, Environment and Development Economics 11: 159-178.

[35] Mendelsohn R. K. Emanuel, S. Chonabayashi, and L. Bakkensen (2012), The impact of climate change on global tropical cyclone damage, Nature Climate Change 2: 205-209.

[36] Millner, A. (2013), On welfare frameworks and catastrophic climate risks, Journal of Environmental Economics and Management, Volume 65, Issue 2, Pages 310325.

[37] National Research Council. Evaluating Progress of the U.S. Climate Change Science Program: Methods and Preliminary Results. Washington, DC: The National Academies Press, 2007.

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

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

[40] Nordhaus, W. D.(2010). Economic Policy in the Face of Severe Tail Events, Journal of Public Economic Theory, 14 (2), 2012, pp. 197- 219.

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

[42] 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.

[43] Roe, G. H., and M.B. Baker (2007). Why Is Climate Sensitivity So Unpredictable? Science 26 October 2007: Vol. 318 no. 5850 pp. 629-632.

[44] Sargent, T. J. (1987). Dynamic Macroeconomic Theory. Harvard University Press.

[45] Schlenker, W., and M. J. Roberts (2009), Nonlinear temperature effects indicate severe damages to U.S. crop yields under climate change, Proc Natl Acad Sci U S A. 2009 September 15; 106(37): 1559415598.

[46] Selden T.M, and D. Song (1995), Environmental Quality and Development: Is There a Kuznets Curve for Air Pollution Emissions?, Volume 27, Issue 2, Pages 147162.

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

[48] Stokey, N. (1998), Are there limits to growth?, International Economic Review , Vol. 39, No. 1 , pp. 1-31

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

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

[51] 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.

[52] van der Zwaan, B. C. C., R. Gerlagh, G. Klaassen, and L. Schrattenholzer. 2002. Endogenous Technological Change in Climate Change Modelling. Energy Economics 24(1): 1-19.

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

[54] Weitzman, M. (2009), On modeling and interpreting the economics of catastrophic climate change, The Review of Economics and Statistics 91(1), 1-19.

[55] Weitzman, M. (2011). Additive Damages, Fat-Tailed Climate Dynamics, and Uncertain Discounting, Chapter in NBER book The Economics of Climate Change: Adaptations Past and Present (2011), Gary D. Libecap and Richard H. Steckel, editors (p. 23 - 46).

[56] Weitzman, M. (2013). Tail-Hedge Discounting and the Social Cost of Carbon. Journal of Economic Literature, 51(3): 873-82.

[57] Weyant, J.P, F. C. de la Chesnaye, and G. J. Blanford (2006). Overview of EMF-21: Multigas Mitigation and Climate Policy, The Energy Journal, Multi-Greenhouse Gas Mitigation and Climate Policy Special Issue.

Appendix: Extension to smooth learning

We consider the same model as in the main text but introduce smooth learning about the true damage parameters Δ_y and Δ_u. Assume that priors are normally distributed with mean $\mu_{\Delta,y}$ ($\mu_{\Delta,u}$ respectively) and variance $\sigma^2_{\Delta,y}$ ($\sigma^2_{\Delta,u}$, resp.). Signals are the realizations of damages that come from the true distributions, but initially we cannot tell apart damages from weather volatility and those from more persistent climate impacts. For illustration, we consider the two cases separately: first, a log-normal distribution for output losses that lead to a normal distribution for utility impacts; and second, log-normal direct utility losses. For each case, we denote the posterior mean for Δ_y (Δ_u) based on cumulative experience at time t by μ_t. There is 'no news' at time t when the posterior mean equals the prior mean, $\mu_t = \mu_{\Delta,y}$ ($\mu_{\Delta,u}$ respectively). We consider the carbon policy h_t as defined in (12), and study how h_t develops; specifically, we test whether h_t remains constant or declines when 'no news' appears. This allows us to see which formulation is consistent with the model in the main text.

Output losses: normally distributed utility impacts

Consider first output losses only (i.e., $\Delta_u = 0$), and assume experienced (relative) losses given by

$$1 - \exp(-\Lambda_{y,t} D_t)$$

where the state of the climate captured by D_t is observed, and output losses contain a stochastic signal for the persistent damage sensitivity

$$\Lambda_{y,t} = \Delta_y + \varepsilon_{y,t}$$

with $\varepsilon_{y,t} \sim N(0, \sigma_\varepsilon^2)$ and i.i.d. across periods and also independent of Δ_y. We observe $\Lambda_{y,t}$ but cannot tell apart the contribution of the noise and the true damage that has an initial prior $\Delta_y \sim N(\mu_{\Delta,y}, \sigma^2_{\Delta,y})$ with $\mu_{\Delta,y} > 0$. Thus, in expectations, temperature causes output losses but there can also be temporary positive productivity shocks, $\varepsilon_{y,t} < 0$. Then, in this setting, we can apply the normal learning rule (De Groot, 1970) to see that after t observations, the posterior distribution for $\Lambda_{y,\tau}$, $\tau > t$ is given by

$$\Lambda_{y,\tau} \sim N(\mu_t, \sigma_t^2) \tag{15}$$

$$\mu_t \equiv \mathbb{E}_t[\Lambda_{y,\tau}] = \mu_{\Delta,y} + \frac{t}{t + \sigma^{-2}_{\Delta,y}\sigma^2_{\varepsilon,y}}(\overline{\Lambda}_{y,t} - \mu_{\Delta,y}) \tag{16}$$

$$\sigma_t^2 \equiv Var_t[\Lambda_{y,\tau}] = \frac{\sigma^2_{\Delta,y}}{1 + t\sigma^2_{\Delta,y}\sigma^{-2}_{\varepsilon,y}} + \sigma^2_{\varepsilon,u}. \tag{17}$$

where $\overline{\Lambda}_{y,t}$ is the average observation after t observations. Thus, the future utility losses due to output reduction at $\tau > t$ can be obtained using the posterior at t:

$$\mathbb{E}_t[\Lambda_{y,\tau} D_\tau] = \mathbb{E}_t[\Lambda_{y,\tau}]D_\tau = \mu_t D_\tau. \tag{18}$$

We define the sure-loss policy as in the main text through h which gives the present-value utility losses for initial mean expectation $\mu_{\Delta,y}$ and $\sigma^2 = 0$.

Proposition 5 *(output losses) The optimal policy at time t is proportional to the sure-loss policy h,revised upwards or downwards only because of more pessimistic or optimistic beliefs on expected damages:*

$$h_t = \frac{\mu_t}{\mu_{\Delta,y}} h,$$

with $h_0 = h$.

Proof. Follows directly from the independence between the variation in future impacts and the delay structure of impacts:

$$
\begin{aligned}
h_t &\equiv \mathbb{E}_t[-\textstyle\sum_{\tau=1}^{\infty}\delta^{\tau}\dfrac{du_{t+\tau}}{dz_t}] \\
&\equiv \mathbb{E}_t[-\textstyle\sum_{\tau=1}^{\infty}\delta^{\tau}\Lambda_{y,\tau}\dfrac{dD_{t+\tau}}{dz_t}] \\
&= \mu_t\textstyle\sum_{\tau=1}^{\infty}\delta^{\tau}\dfrac{dD_{t+\tau}}{dz_t} = \mu_t\textstyle\sum_{\tau=1}^{\infty}\delta^{\tau}\mathcal{R}(\tau) \\
&= \dfrac{\mu_t}{\mu_{\Delta,y}}h.
\end{aligned}
$$

■

The proposition reveals no trend in carbon policies when no news arrives: $h_t = h_0$ if $\mu_t = \mu_{y,\Delta}$. From the perspective at $t = 0$, future policies h_t have a normal distribution determined by the distribution of μ_t:

$$
\begin{aligned}
\mathbb{E}_0[h_t/h] &= \mathbb{E}_0[\mu_t/\mu_{\Delta,y}] = 1 &\text{(19)}\\
Var_0[h_t/h] &= Var_0[\mu_t/\mu_{\Delta,y}] = \left(\frac{t}{t+\sigma_{\Delta,y}^{-2}\sigma_{\varepsilon,y}^{2}}\right)^2 \frac{\sigma_{\Delta,y}^2 + t^{-1}\sigma_{\varepsilon,y}^2}{\mu_{\Delta,y}^2} &\text{(20)}\\
&= \frac{t}{t+\sigma_{\Delta,y}^{-2}\sigma_{\varepsilon,y}^{2}}\frac{\sigma_{\Delta,y}^2}{\mu_{\Delta,y}^2} &\text{(21)}
\end{aligned}
$$

From the ex-ante perspective, the expected future policies show a slow divergence towards the prior distribution for Δ_y as information and better observations enter; more precise obserations as captured by smaller $\sigma_{\varepsilon,y}$ result in faster adjustment.

Direct utility losses: right-skewed utility impacts

Consider then utility losses only (i.e., $\Delta_y = 0$), and assume that the damage parameter is log-normally distributed. As was the case in the main text, here, the level of policies will adjust in a "no news" scenario. Consider intangible damages in the periodic utility, having a log-normal distribution,

$$u_t = \ln(c_t) - \exp(\Lambda_{u,t})D_t,$$

where

$$\Lambda_{u,t} = \Delta_u + \varepsilon_{u,t},$$

with zero-mean normal realizations $\varepsilon_{u,t}$ that are i.i.d. across periods and also independent of Δ_u. Here, too, the initial prior is normal, $\Delta_u \sim N(\mu_{\Delta,u}, \sigma_{\Delta,u}^2)$. Thus, again, the realized (experienced) damage depends on the unknown damage-generating process and on the noise term. As above, we obtain the expected intangible damage, after a given history of observed damages at time t; for notational convenience, we will now use for $\tau > t$: $\Lambda_{u,\tau} \sim N(\mu_t, \sigma_t^2)$.

The expected damages then have a right-skewed distribution with potentially a fat tail for large damages, so that the expected future utility loss at $\tau > t$ is given by

$$\mathbb{E}_t[\exp(\Lambda_{u,\tau})D_\tau] = \exp(\mu_t + \frac{1}{2}\sigma_t^2)D_\tau \tag{22}$$

where $\mu_t = \mathbb{E}[\Lambda_{u,\tau}]$ and $\sigma_t^2 = Var[\Lambda_{u,\tau}]$ are given by the same equations as above for $\mathbb{E}[\Lambda_{y,t}]$ and $Var[\Lambda_{y,t}]$.

Proposition 6 *(utility losses) The optimal policy at time t is revised upwards or downwards consistently with the updated damage estimate, but it also has a markup because of uncertainty, which declines through learning:*

$$h_t = \exp(\mu_t - \mu_{\Delta,u} + \frac{1}{2}\sigma_t^2)h, \tag{23}$$

with

$$h_0 = \exp(\frac{\sigma_{\Delta,u}^2 + \sigma_{\varepsilon,u}^2}{2\mu_{\Delta,u}^2})h. \tag{24}$$

Proof. The proof follows from independence between variation in damages and the time structure of damages, see the previous proposition, with the damage coefficient defined through (22). ■

The initial distribution for the carbon price ($\mu_{\Delta,u}$ and $\sigma_{\Delta,u}$) can be calibrated to Tol's (2009) data as in the main text. Importantly, the policy at $t = 0$ immediately deviates from the sure-loss policy h, as the optimal carbon price at $t = 0$ now demands a markup due the probability of high-damage events. When no news arrives, $\mu_t = \mu_{\Delta,u}$, there is a systematic downwards drift in future policies, as $\frac{1}{2}\sigma_t^2$ goes down: (17) implies that the fat-tail markup effect on damages declines, through learning, $\sigma_t^2 \searrow \sigma_{\varepsilon,u}^2$. No news is good news in this setting, as in the main text.

We can also consider more generally the distribution of expected carbon policies at future time t: they are distributed log-normally, based on the underlying normal distribution for $\mu_t - \mu_{\Delta,u} + \frac{1}{2}\sigma_t^2$. There is a systematic downwards drift in future policies, as $\frac{1}{2}\sigma_t^2$ goes down (17), while the expected value of $\mu_t - \mu_{\Delta,u}$ is zero. The variation in future policies comes from μ_t (16), while the expected value of $\mu_t - \mu_{\Delta,u}$ is zero. Consider the dynamics for the underlying normal distribution:

$$\begin{aligned}\mathbb{E}_0[\ln(h_t/h_0)] &= \frac{1}{2}\sigma_t^2 - \frac{1}{2}\sigma_0^2 = -\frac{1}{2}\frac{t}{t+\sigma_{\Delta,y}^{-2}\sigma_{\varepsilon,y}^2}\sigma_{\Delta,y}^2 \\ Var_0[\ln(h_t/h_0)] &= Var_0[\mu_t] = \frac{t}{t+\sigma_{\Delta,y}^{-2}\sigma_{\varepsilon,y}^2}\sigma_{\Delta,y}^2\end{aligned}$$

We can then derive the expected mean and variation of future carbon policies through:

$$\begin{aligned}\mathbb{E}_0[h_t/h_0] &= \exp(\mathbb{E}_0[\ln(h_t/h)] + \frac{1}{2}Var_0[\ln(h_t/h)]) = 1 & (25)\\ Var_0[h_t/h_0] &= \exp(Var_0[\ln(h_t/h)] - 1)(\mathbb{E}_0[h_t/h])^2 & (26)\\ &= \exp(\frac{t}{t+\sigma_{\Delta,y}^{-2}\sigma_{\varepsilon,y}^2}\sigma_{\Delta,y}^2) - 1 & (27)\end{aligned}$$

Here we see that the expected future carbon policy is constant: $\mathbb{E}_0[h_t] = h_0$. There are two opposing forces. In a majority of scenarios, the policy variable h_t will decrease as the impact estimate becomes more precise and not more pessimistic; so that the probability of future high-impacts decreases. But when a high impact is observed, the size of expected impacts is revised upwards and carbon policies h_t go up relatively sharply, as in the main text.

Appendix: detailed energy-sector model

We specify the economy's production function (5) as follows:[28]

$$\begin{aligned}y_t &= k_t^{\alpha}[A_t(l_{y,t}, e_t)]^{1-\alpha}\exp(-\Delta_{y,t}D_t) & (28)\\ A_t(l_{y,t}, e_t) &= \min\{A_{y,t}l_{y,t}, A_{e,t}e_t\} & (29)\\ e_t &= e_{f,t} + e_{n,t} & (30)\\ e_{f,t} &= \min\{A_{f,t}l_{f,t}, B_t z_t\} & (31)\\ e_{n,t} &= \frac{\varphi+1}{\varphi}(A_{n,t}l_{n,t})^{\frac{\varphi}{\varphi+1}} & (32)\\ l_t &= l_{f,t} + l_{n,t} + l_{y,t}. & (33)\end{aligned}$$

Before explaining the structure in detail, we note that there will be time-trends for total labor l_t, and for productivities (A_y, A_e, A_f, A_n); solution to the energy-sector allocation problem allows us to express the energy-labor composite $A_t(l_{y,t}, e_t)$ as depending only on time and carbon inputs, $A_t(z_t)$, as in the main text.

[28]This specification builds on Gerlagh and Liski (2012).

The labor-energy composite $A_t(l_{y,t}, e_t)$ takes a Leonfief form capturing an extremely low elasticity of substitution between labor in the final-good sector $l_{y,t}$ and energy e_t. The assumption avoids unrealistically deep early reductions of emissions through labor reallocation; see also Hassler, Krusell and Olovsson (2011). Final-good and energy productivities $A_{y,t}$ and $A_{e,t}$ are calibrated so that the model matches the business-as-usual (BAU) quantities for fossil-fuel and non-carbon energy with the A1F1 SRES scenario from the IPCC (2007). We introduce production for total energy e_t that depends, effectively, only on labor allocation at t: the core allocation problem in the energy sector is how to allocate a given total labor l_t at time t between final output $l_{y,t}$, fossil-fuel energy, $l_{f,t}$, and non-carbon energy, $l_{n,t}$. Thus, the energy and climate policy steers the labor allocation $(l_{y,t}, l_{f,t}, l_{n,t})_{t\geq 0}$ and thereby the quantities of fossil-fuel, $e_{f,t}$, and non-carbon energy, $e_{n,t}$. Both energy sources are intermediates, summing up to the total energy input, $e_t = e_{f,t} + e_{n,t}$.

In (31), we assume that $e_{f,t}$ can be produced with a constant-returns to scale technology using labor $l_{f,t}$ and the fossil-fuel z_t, where $A_{f,t}$ and B_t describe productivities. The fuel resource is not a fixed factor and commands no resource rent; by this assumption, our focus is on the "coal phase", as in Golosov et al. (2011), where the fossil-fuel resource is in principle unlimited. In contrast, in equation (32), where $\varphi > 0$ describes the elasticity of supply from the non-carbon sector; the non-fossil fuel energy production is land-intensive and subject to diminishing returns and land rents (as in Fischer and Newell, 2008).

The model structure described here can reasonably well capture the two main adjustment channels to carbon policies: energy savings that typically feature the early decades of the adjustment, and then decarbonization that is needed to meet the long-run climate targets; that is, a transition to non-carbon energy is a long-run rather than short-run option. The calibration of the energy sector is detailed Gerlagh and Liski (2012), where it progresses as follows. Without carbon policy, $h = 0$, the labor market allocation can be solved in closed form; thus, we can invert the model to map from quantities $(l, y, e_f, e_n)_{t\geq 0}$ to productivities $(A_y, A_e, A_f, A_n)_{t\geq 0}$. We express all energy in carbon units; to obtain this, we set $B_t = 1$ and then obtain three distinct energy productivities (A_e, A_f, A_n). We match the business-as-usual (BAU) quantities $(y, e_f, e_n)_{t\geq 0}$ with the A1F1 SRES scenario from the IPCC (2007). Population follows a logistic growth curve based on World Bank forecasts. Population in 2010 is set at 6.9 [billion], while the maximum population growth rate is chosen such that in 2010 the effective population growth rate per decade equals 0.12 [/decade]. The maximum expected population (reached at about 2200) is set at 11 [billion].

The online supplementary file contains a program for reproducing the graphs in the text, `https://www.dropbox.com/sh/7meos655j14jh5p/_dlr8X_FHI`. The labor allocation is numerically obtained as follows. The allocation can be solved period-by-period taking the (i) productivity parameters, (ii) total labor, (iii) savings g, and (iv) carbon policies h_t as given. We drop the time subscript in the variables:

1. We normalize prices for the final good to equalize marginal utility, so that factor prices can be interpreted as marginal welfare per factor endowment:

$$p = \frac{1}{c} = \frac{1}{(1-g)y}.$$

2. Final-good producers of y take capital k, wages w, and prices of energy q and output p as given. Since $y = k^{\alpha}[\min\{A_y l_y, A_e e\}]^{1-\alpha}\exp(-\Delta_{y,t}D_t)$, factor compensation for labour and energy together receives a share $(1-\alpha)$ of the value of output py:

$$wl_y + qe = (1-\alpha)py$$

where $e = e_f + e_n$.

3. Fossil-fuel energy production combines labor and fuels, with technology $e_{f,t} = \min\{A_{f,t}l_{f,t}, B_t z_t\}$. Fossil fuel use and labour employed, $z, l_f \geq 0$, are strictly positive if q covers the factor payments, including the carbon price τ

$$\left[q - \left(\frac{w}{A_f} + \frac{\tau}{B}\right)\right] \times l_f \leq 0.$$

The zero profit condition for fossil fuel energy allocates the value of fossil fuel energy to labour and emission payments; using the production identity we can express it in terms of labour employed,

$$qe_f = wl_f + \tau z = (w + \frac{\tau A_f}{B})l_f.$$

4. Carbon-free energy inverse supply is given by the first-order condition

$$q = w\frac{\partial l_n}{\partial e_n} = \frac{w_t}{(A_n)^{\frac{\varphi}{\varphi+1}}}(l_n)^{\frac{1}{\varphi+1}}.$$

The value share of labour employed in the carbon-free energy sector equals $\varphi/(1+\varphi)$, so that the rent value is expressed in labour employed:

$$qe_n = (1 + \frac{1}{\varphi})wl_n$$

We obtain four equations in four unknowns l_y, l_f, l_n, w:

$$\begin{aligned} A_y l_y &= A_e(A_f l_f + \frac{\varphi+1}{\varphi}(A_n l_n)^{\frac{\varphi}{\varphi+1}}) & (34) \\ wl + \frac{\tau A_f}{B} l_f + \frac{1}{\varphi} w l_n &= \frac{1-\alpha}{1-g} & (35) \\ \frac{w}{A_f} + \frac{\tau}{B} &\geq \frac{w}{(A_n)^{\frac{\varphi}{\varphi+1}}}(l_n)^{\frac{1}{\varphi+1}} \perp l_f \geq 0 & (36) \\ l_y + l_f + l_n &= l & (37) \end{aligned}$$

For (34) note that, for strictly positive input prices, $A_t(\cdot) = \min\{A_y l_y, A_e e\} \Rightarrow A_y l_y = A_e e$. In equation (35) we allocate the value of output that is not attributed to capital (the right-hand side) to the labour, carbon emissions, and land rent for the non-carbon energy (where we latter two terms are expressed in labour units). Equation (36) compares the production costs for fossil fuel energy with non-carbon energy, and the last equation is the labor market clearing equation. Note that the solution depends on the state of the economy only through total labor l and productivities A_y, A_e, A_f, A_n.

In the absence of a carbon policy, $\tau = 0$, we can solve the allocation in closed-form:

$$\begin{aligned} l_{n,t} &= \frac{A_{n,t}^{\varphi}}{A_{f,t}^{\varphi+1}} & (38) \\ w_t &= \frac{1-\alpha}{1-g} \frac{\varphi}{\varphi l_t + l_{n,t}} & (39) \\ l_{y,t} &= \frac{A_{e,t}}{A_{y,t} + A_{e,t} A_{f,t}} [A_{f,t}(l_t - l_{n,t}) + \frac{\varphi+1}{\varphi}(A_{n,t} l_{n,t})^{\frac{\varphi}{\varphi+1}}] & (40) \\ l_{f,t} &= l_t - l_{y,t} - l_{n,t} & (41) \end{aligned}$$

Here we include the time subscripts to emphasize the drivers of the solution. This business-as-usual allocation is used to calibrate the productivities. When $\tau > 0$, the solution is numerical, and available in the supplementary file.