

When the long run matters

The joint effect of carbon decay and discounting

Terrence Iverson · Scott Denning · Sammy Zahran

Received: 15 May 2014 / Accepted: 29 December 2014 / Published online: 27 January 2015
© Springer Science+Business Media Dordrecht 2015

Abstract Roughly 20 percent of current CO₂ emissions will likely remain in the atmosphere for thousands of years (Solomon et al. 2008). Despite this, climate damages attributable to current emissions that occur beyond 150 years or so have almost no effect on the current optimal carbon tax in typical integrated assessment models. The source of this strong result is conventional economic discounting. The current paper builds on recent work by Gerlagh and Liski (2013) and Iverson (J Environ Econ Manag 66:598–608, 2013a, b) to demonstrate this fact in a simple way and to show that it is not robust to plausible changes in the calibration approach for discounting parameters. Specifically, when time preference rates decline, a possibility supported by a wide variety of studies from psychology and economics, long run consumption impacts are potentially very important and so are long run features of the carbon cycle. The paper follows (Gerlagh and Liski 2013) in showing that this remains true even when the discounting parameters are calibrated to match historical interests rates, thus avoiding the main economic critique of the Stern Review (Stern 2007; Nordhaus 2008; Weitzman Rev Econ Stat 91:1–19 2009). The effects are quantified using a formula for the optimal carbon tax from Iverson (J Environ Econ Manag 66:598–608, 2013a, b), which we use to decompose the current optimal tax into the cumulative contribution from consumption impacts at different horizons.

1 Introduction

From a scientific perspective, climate change is a problem in which the very long run is clearly important. A significant portion of carbon dioxide emissions remain in the atmo-

T. Iverson (✉) · S. Zahran
Department of Economics, Colorado State University, Fort Collins, CO, USA
e-mail: terry.iverson@colostate.edu

S. Denning
Department of Atmospheric Science, Colorado State University, Fort Collins, CO, USA

sphere for “many thousands of years” (Solomon et al. 2008), and damages attributable to these emissions include irreversible changes, such as the potential collapse of the Greenland or West Antarctic ice sheets. But viewed through the lens of cost benefit analysis, typical integrated assessment models suggest that climate outcomes in the long run have little or no bearing on current policy decisions. In particular, the optimal carbon tax, a summary measure of the economic value of current emission reductions, is largely insensitive to climate outcomes beyond a hundred years or so.

The source of this strong result is economic discounting.¹ Discounting converts costs and benefits at different times into a common unit of account that reflects foregone investment returns when a payoff stream is received in the future rather than today. In typical economic models, the rate at which future payoffs are discounted is governed by two parameters: the rate of time preference—the rate at which future welfare (or “utility”) is discounted—and the consumption elasticity, which reflects society’s aversion to consumption inequality across generations. In economics, these parameters are most commonly calibrated by (i) assuming that they are constant and (ii) choosing them so savings decisions are consistent with data on financial transactions in real economies. This implies a discount rate (on real consumption goods) of about 5.5 percent (Nordhaus 2008). Thus, a billion dollars in 200 years has the value today of an economy car (\$ 17K), while a billion dollars in 400 years has the value today of a gumball (28 ¢).

While insensitivity to long run outcomes is potentially unsettling, the argument for calibrating the discounting parameters to match historical interest rates cannot be readily dismissed. Suppose, as in the widely-publicized Stern Review on the Economics of Climate Change (Stern 2007), that we instead adopt a lower discount rate. Motivated by ethical considerations, Stern (2007) adopts (constant) discounting parameters that imply a discount rate (on real consumption goods) of about 1.4 percent. The trouble is that the assumed parameters imply savings rates twice as high as observed in real economies (Nordhaus 2008), so the high value attributed to climate policy reflects an unrealistically high willingness to transfer wealth to future generations.

The accompanying assumption, a constant rate of time preference, is not nearly as compelling. Since it was suggested by Samuelson (1937), a constant rate of time preference has been part of the dominant model in economics for evaluating welfare outcomes over time. But Samuelson himself stated that “any connection between utility as discussed here and any welfare concept is disavowed” (Samuelson 1937; Frederick et al. 2002). Meanwhile, the assumption stands in conflict with experimental evidence from psychology and economics (Thaler 1981; Loewenstein 1987; Cropper et al. 1994). For example, Thaler (1981) asks subjects to indicate the dollar amount, to be received in a month, a year, or ten years, that would make them indifferent to receiving \$ 15 today. The mean responses imply discount rates of 345 %, 120 %, and 19 % over the respective horizons, consistent with declining rates of time preference (DRTP). At long horizons, DRTP are consistent with the observation that people can more readily imagine differences between the current generation and the generation in 20 years than they can differences between the generation in 300 years and that in 320 years (Rubinstein 2003; Karp and Tsur 2011; Layton and Brown 2000).

A variety of normative rationale for DRTP have also been suggested. These are arguably more important for motivating policy since they speak to how policy ideally *should* be set.

¹Gerlagh and Liski (2013) (hereafter “GL”) refer to this insensitivity to long-term outcomes under conventional discounting assumptions as a climate-policy “puzzle”. Their paper is discussed below.

One of the strongest arguments applies when agents in society disagree about the appropriate rate of time preference. In this case, a utilitarian social planner who seeks to aggregate preferences will act “as if” using DRTP. The argument is conceptually related to seminal work by Weitzman (1998), and it has been developed in a variety of settings by Li and Löfgren (2000), Gollier and Zeckhauser (2005), Heal and Millner (2013), and Heal and Millner (2014). We build on a version of this argument in the application.

Incorporating DRTP into policy analysis introduces complexities that do not arise when time preference rates are constant. Consider a decision maker (or a sequence of decision makers) with DRTP who plans a series of actions over time. Today, when evaluating outcomes between ten and fifteen years out, she uses a relatively-low time preference rate. But when the time comes to implement the year ten action, she (or her successor) would reevaluate the decision by applying a relatively-high short-term rate over the same, now proximate, dates. This shift in preferences can lead to “time inconsistent” investment decisions. In principle, the difficulty is overcome by requiring agents each period to correctly anticipate how future agents are expected to respond (Strotz 1955).² But solving the problem in this way can be difficult in practice. Indeed, in most quantitatively interesting integrated assessment models, solving for the time consistent solution is numerically infeasible.

To make progress, the climate policy literature that considers DRTP has either employed relatively simple climate-economy models (Karp 2007; Fujii and Karp 2008) or restricted time preference rates to differ over two periods only (Karp 2005; Gerlagh and Liski 2013). The latter assumption is referred to as quasi-hyperbolic discounting (Phelps and Pollak 1968; Laibson 1997). It simplifies the analysis, but is not quantitatively plausible for a long term problem like climate change. Iverson (2013a), which builds on Gerlagh and Liski (2013), is the first paper to solve for (time consistent) optimal carbon taxes in a quantitatively plausible integrated assessment model while allowing for arbitrary non-constant time preference rates. It employs a recent integrated assessment model developed in Golosov et al. (2014).

With DRTP, it is possible for the model to oblige historical evidence on savings rates and for the current carbon tax to respond in a sensitive way to climate outcomes in the distant future. This point is emphasized in Gerlagh and Liski (2013). We evaluate the same question while building on the model in Iverson (2013a). The paper goes beyond prior work by providing a way to explicitly decompose the contribution to the optimal tax from consumption impacts at different times.³ This makes it possible to study interactions between discounting assumptions and long run features of the carbon cycle in determining the importance of long run outcomes for current policy. Among other things, our results have implications for assessing the potential for scientific evidence regarding long run carbon decay to influence the current case for climate action.

2 Model

The paper employs the integrated assessment model of Golosov et al. (2014) (the “GHKT model”). The model structure emulates the widely-used DICE model (Nordhaus 2013).

²This implies a subgame perfect equilibrium.

³The employed decomposition is not possible in typical, quantitatively-interesting integrated assessment models (e.g. Nordhaus 2008; Tol 1997; Hope 2008) because the optimal tax in those models is inferred numerically from the shadow value on the carbon stock constraint at the numerical optimum.

It combines a neoclassical economic growth model where carbon emissions arise as a byproduct of production activities with a carbon cycle that describes the manner in which carbon emissions accumulate in the atmosphere. Accumulated pollution stocks cause damages that feed back on economic growth. The model equations are presented in the [Appendix](#).

The GHKT model differs from DICE in three important ways. First, the climate module abstracts from temperature inertia. As a result, the dynamic link between emissions and damages is somewhat less realistic than in DICE. Fortunately, GL show that temperature inertia can be recovered within the GHKT model by modifying the climate module in a tractable way. We build on GL to include temperature inertia in our analysis. Second, the model contains a full “general equilibrium” specification of clean and dirty energy sectors. Due to this, the model does a better job simulating the effect of a carbon tax in the equilibrium of a competitive market economy. Finally, the GHKT model allows for stochastic uncertainty about future damages.

Analytic tractability is obtained in Golosov et al. (2014)—and in Gerlagh and Liski (2013) and Iverson (2013a) and —by imposing restrictive functional form assumptions. These are presented in the [Appendix](#). As argued in Golosov et al. (2014), the assumptions are close to those in most of the prior climate economics literature. Nevertheless, they have some strong consequences. First, they effectively shut down strategic interactions among generations (Iverson 2013a; Karp 2014). This feature arises because the equilibrium of the model is linear in the stock of carbon. Second, they have restrictive implications when modeling the time path of fossil fuel extraction (Rezai and van der Ploeg 2014).

The employed model simplifies the “full” GHKT model by assuming that fossil fuel reserves are not used up along the optimal path. Provided coal extraction costs remain low for the foreseeable future, this assumption is unlikely to matter much for the analysis. This is the case because coal is the most important source of carbon emissions in long run analyses (van der Ploeg and Withagen 2012), and coal deposits are not used up along the optimal path under plausible parameter values in typical integrated assessment models. This simple interpretation breaks down, however, if the marginal extraction cost of coal increases significantly as more coal is extracted.

2.1 Optimal carbon taxes

As shown in Gerlagh and Liski (2013) and Iverson (2013a), the GHKT model provides a highly convenient setting in which to consider the implications of DRTP for climate policy. Gerlagh and Liski (2013) use the model to derive an explicit formula for the optimal carbon tax when time preference rates differ over two periods (so-called “quasi-hyperbolic” discounting), and Iverson (2013a) derives an analogous formula when time preference rates vary arbitrarily over time. We employ Iverson’s formula. It reduces to the optimal tax formula in Golosov et al. (2014) when discounting is constant, and it reduces to the tax formula in Gerlagh and Liski (2013) when discounting is quasi-hyperbolic.⁴

Without corrective policy, consumers and firms in a market economy largely ignore the environmental burden associated with emission-generating activities because these costs are external—they fall primarily on people and societies distant in space and time. The optimal carbon tax is the price on the carbon content of fossil fuels that would induce

⁴The formula is presented in the [Appendix](#).

consumers and firms to “internalize” these external costs (Pigou 1924), thus guiding the market system to an abatement path that is socially optimal in the broad sense of maximizing the present value of benefits minus costs. The optimal tax equals the present value of all current and future social damages associated with an incremental ton of carbon emissions today, and it can be viewed as a summary measure of the current case for climate action.

The optimal tax depends on three sets of parameters: discounting parameters, carbon decay and climate damage parameters, and a parameter that determines the relationship between capital and output in final-goods production. Calibration assumptions for these parameters are discussed below.

2.2 Calibration

2.2.1 Discounting

Time preference rates (TPR) are calibrated to ensure the real return on capital in the model is consistent with historical interest rates. A similar approach is considered in Gerlagh and Liski (2013). To construct alternative TPR paths, we use a functional form for declining rates implied by the preference aggregation argument of Li and Löfgren (2000). Agents in society disagree about the correct rate of time preference, and a social planner aggregates preferences by taking a population-weighted average of welfare functions.⁵

To keep the analysis simple, we assume that agents adopt one of two discounting perspectives, and we loosely model these perspectives after the popular discounting debate between economists William Nordhaus and Nicholas Stern. Nordhaus (2008) calibrates a constant TPR to ensure the real return on capital is consistent with historical interest rates, while Stern (2007) adopts a rate near zero on ethical grounds. If r_S is the Stern rate, r_N the Nordhaus rate, and p_S the population weight on the Stern perspective, the preference aggregation argument implies an aggregate period t discount factor of

$$D(t) = p_S \cdot e^{-r_S t} + (1 - p_S) \cdot e^{-r_N t}. \tag{1}$$

It is easy to show that the corresponding discount rate, defined as the instantaneous decay rate of the discount factor, begins in time zero at the mean rate, declines continuously in time, and converges asymptotically to the lower rate, r_S .⁶

As described, the approach would not satisfy the calibration commitment. It would with a Stern weight of zero—this is true by construction since the Nordhaus rate is calibrated to match observed returns. But with higher Stern weights, agents in the model are compelled to save more, which pushes the real return on capital downwards. To maintain the calibration commitment, we modify the procedure by allowing the higher of the two rates

⁵This approach is consistent with Utilitarian ethics when the consumption good is public, so the same amount of the aggregate good is consumed by all agents. Gollier and Zeckhauser (2005) consider the problem in which consumption shares are endogenous while income is exogenous, and Heal and Millner (2013) consider the problem in which consumption shares and income are both endogenous.

⁶This is seen by computing $r(t) = -\frac{\dot{D}(t)}{D(t)} = \frac{p_S \cdot e^{-r_S t}}{p_S \cdot e^{-r_S t} + (1-p_S) \cdot e^{-r_N t}} r_S + \frac{(1-p_S) \cdot e^{-r_N t}}{p_S \cdot e^{-r_S t} + (1-p_S) \cdot e^{-r_N t}} r_N$, where $\dot{D}(t)$ is the time derivative of $D(t)$. Weitzman (1998) proves the same limiting result for an arbitrary finite probability distribution. Iverson (2013b) shows that an analogous result holds when decision makers are unable to assign probabilities to potential discount rates and employ the non-Bayesian decision criterion Minimax Regret.

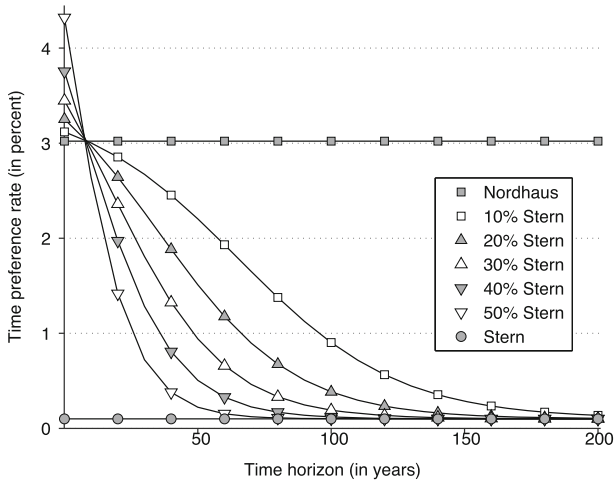


Fig. 1 Comparison of alternative discounting paths

to adjust to ensure that the path of real returns remains fixed at what it is under the Nordhaus calibration.⁷ Specifically, we fix the Stern rate at 0.1 % (as in (Stern 2007)), then for a given population weight on Stern (p_S), we consider an “adjusted Nordhaus rate” (r_{AN}), which is adjusted endogenously so the implied declining path of time preference rates gives rise to an average real return on capital (over the first 5 simulated decades) of 5.25 percent. r_{AN} increases with the Stern weight. The corresponding path of discount factors is given by Eq. 1 with r_N replaced by r_{AN} . TPR paths obtained in this way are shown in Fig. 1.

2.2.2 Carbon cycle and climate damages

Carbon dioxide is chemically inert in the atmosphere, so its removal is limited to net uptake by the land and oceans. The lifetime of anthropogenic CO₂ has been widely and persistently misunderstood in the popular culture and even in some of the natural science literature. To first order, the decay of the anthropogenic carbon pulse in the atmosphere can be approximated by assuming that 80 percent of it has a half life of about 300 years, while the remaining 20 percent remains in the atmosphere permanently (Archer and Brovkin 2008). Over many centuries, a more accurate representation of CO₂ removal is a sum of three exponential decay terms operating on the excess atmospheric CO₂: fast decay of a small carbon pulse through CO₂ fertilization (Hyvönen et al. 2007; Norby and Zak 2011) and dissolution into the surface ocean (Sabine et al. 2004); slow decay of a larger pulse by chemical reaction with calcium carbonate in the deep ocean (Tans 1998); and extremely slow decay of the remainder by reaction with rocks over geologic time (Archer and Brovkin 2008).

Larger pulses saturate the faster reservoirs so the long-lived fraction increases with total historical emissions. Archer et al. (2009) compare calculations of carbon pulse response in eight Earth System models capable of representing the millennial timescales of these geochemical processes. They found that 15 to 30 percent of a 1000 GtC pulse remained in the

⁷The qualitative results of the paper are unchanged if we instead adopt the more conventional calibration procedure above.

atmosphere after 1000 years. For a larger pulse of 5000 GtC, 25 % to 65 % remains airborne after 1000 years and 10 % to 30 % remains airborne after 10,000 years. All emissions scenarios summarized by the IPCC (Solomon et al. 2008; Van Vuuren et al. 2011) are intermediate between these extremes, so the approximation outlined above (80 % with a 300 year half life and 20 % permanent) is plausible.

To account for these features of the carbon cycle, Golosov et al. (2014) model decay of atmospheric carbon with the following three-parameter decay function:

$$1 - d_s = \phi_L + (1 - \phi_L)\phi_0(1 - \phi)^s. \quad (2)$$

$1 - d_s$ is the fraction of a unit of emissions that remains in the atmosphere s periods ahead. ϕ_L is the long-run emission fraction. ϕ is the geometric decay rate of the remaining fraction, and $1 - \phi_0$ is the fraction (of the non-permanent portion of emissions) that falls out of the atmosphere within the first decade. GHKT further link atmospheric carbon to temperature using a relationship adopted in Nordhaus (2008). This models steady-state mean temperature as a logarithmic function of the atmospheric carbon stock.⁸

A feature of the problem omitted from this formulation is temperature inertia. Temperature inertia increases the lag between emissions and damages, and has a potentially important impact on the optimal carbon tax. Fortunately, (Gerlagh and Liski 2013) show that temperature inertia can be accommodated within the GHKT economic model without sacrificing analytic tractability. We build on Gerlagh and Liski (2013) to incorporate temperature inertia while also preserving the carbon cycle specification from GHKT. We do this because the GHKT specification allows us to specify and calibrate the long run emission fraction directly, which is important for our analysis.

To calibrate the carbon decay parameters in Eq. 2, we view ϕ_L as a free parameter that falls somewhere between 0 and 0.3.⁹ For each possible value, we follow (Golosov et al. 2014) in calibrating ϕ_0 and ϕ to match two calibration targets from the literature. First, Archer (2005) argues that excess carbon that does not enter the permanent reservoir has a half-life of about 300 years. Second, according to Solomon et al. (2008), roughly half of emissions leave the atmosphere within thirty years. We further assume that “permanent carbon” remains in the atmosphere for 2000 years. To calibrate the rate of temperature adjustment, we target a six-decade delay between emissions and peak damages, as in Gerlagh and Liski (2013) and Nordhaus (2008).

Damages are driven by a parameter that indicates the percent loss in Gross World Product (GWP) associated with an extra ton of atmospheric CO₂. The parameter is stochastic, reflecting uncertainty about future harm. Only the expected value of the damage parameter in each future period enters the optimal tax formula. Our baseline calibration follows (Golosov et al. 2014). They assume the (current) expected value of the damage elasticity parameter is the same for all future periods, and they calibrate it to match two data points from a meta analysis in Nordhaus and Boyer (2000): a 2.5 degree Celsius temperature increase leads to a 0.48 % loss of GWP, while a catastrophic 6 degree rise causes a 30 % loss in GWP. Combining these assumptions implies an expected damage elasticity in each future period of 2.379×10^{-5} .

⁸Climate sensitivity in this equation is set to 3.0 degrees Centigrade.

⁹In contrast, Golosov et al. (2014) fix it at 0.2 based on summary evidence from Solomon et al. (2008).

Table 1 Optimal taxes for discounting scenarios in Fig. 1

	Optimal tax, 2015 <i>no</i> temp. inertia (\$ per ton carbon)	Optimal tax, 2015 <i>with</i> temp. inertia (\$ per ton carbon)	Average real return, first 50 years (in percent)
Nordhaus (2008)	31	17	5.25
10 % Stern	73	61	5.25
20 % Stern	116	104	5.25
30 % Stern	160	149	5.25
40 % Stern	205	194	5.25
50 % Stern	250	239	5.25
Stern (2007)	440	429	5.25

The “no temperature inertia” scenario adopts the calibrated climate module from Golosov et al. (2014). The “with temperature inertia” scenario augments that specification to include temperature inertia

3 Quantifying long-run importance

3.1 Carbon taxes

The optimal carbon taxes are shown for each discounting path in Table 1.¹⁰ The scenario without temperature inertia follows GHKT. This is augmented to include temperature inertia using the analysis of GL.¹¹ The average real return is reported for the GHKT calibration, but it is essentially identical with temperature inertia included.

Under Nordhaus discounting, the optimal tax with temperature inertia is just over half what it is without.¹² In this case, temperature inertia has a very significant impact on the optimal tax. But as more weight is put on the distant future, the relative importance of temperature inertia declines, becoming almost insignificant when the Stern weight is high. The source of the latter result is intuitive: temperature inertia creates a delay in the onset of damages without changing the long run path, and this has the largest relative impact for high discounting scenarios where near term damages comprise a larger fraction of the total tax.

The results can be compared to a first period carbon tax in Nordhaus (2008) of about \$ 30 per ton carbon and a first period carbon tax in Stern (2007) of about \$ 250. Our significantly higher tax under Stern discounting is due to the model allowing for “permanent carbon”, something not considered in Stern’s original analysis.

The results clearly show that fixing the calibration procedure to oblige historical interest rates does not uniquely pin down climate policy when TPR decline—nor does it necessitate climate policy that is insensitive to long run outcomes. The same point is made in GL.

¹⁰The remaining calibration assumptions follow (Iverson 2013a), which with slight differences to account for model differences, are the same as in Golosov et al. (2014). Total factor productivity evolves as in Nordhaus (2013).

¹¹In the notation of GL, we take $a = (\phi_L, (1 - \phi_L)\phi_0)$, $\eta = (0, \phi)$, and $\pi = \gamma_{\text{GHKT}} = 2.379 \times 10^{-5}$.

¹²Our calibration is hybrid between GHKT and GL. With the full climate module calibration from GL, including their higher damage parameter, the tax under Nordhaus discounting is \$ 27 per ton carbon.

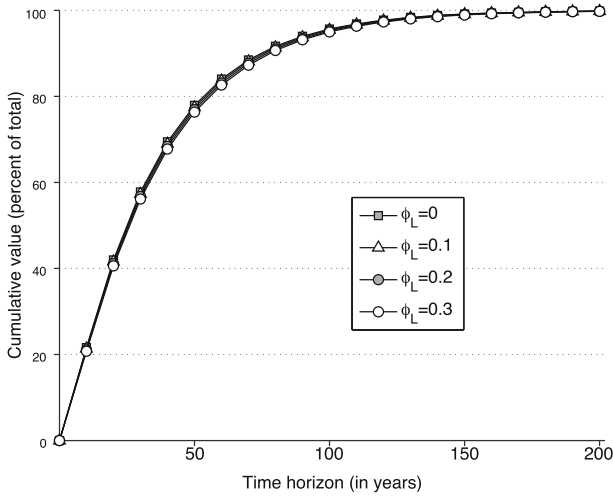


Fig. 2 Percent of total value accrued over time. Nordhaus discounting. ϕ_L is the long run emission fraction

3.2 Cumulative value

The optimal tax aggregates the present value of future consumption impacts from climate damages attributable to an extra ton of carbon emissions today. Only damages associated with the portion of current emissions that remain in the atmosphere in a given future period enter. These damages impact consumption and investment the period in which they occur, and the drop in investment in turn impacts consumption in subsequent periods. The formula (presented in the [Appendix](#)) allows us to decompose the contribution to the current optimal tax from consumption impacts at different times.

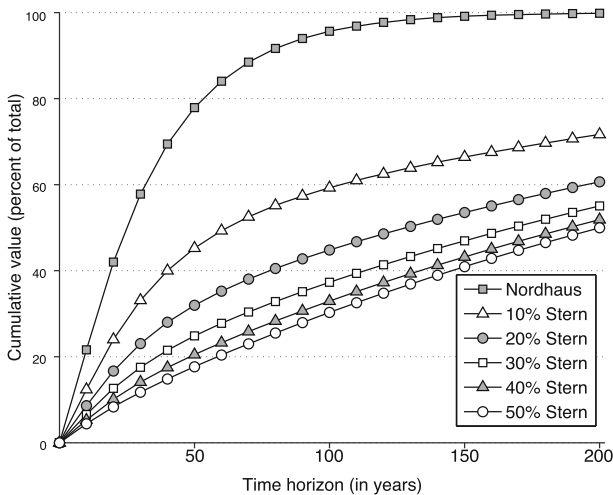


Fig. 3 Percent of total value accrued over time. $\phi_L = 0$

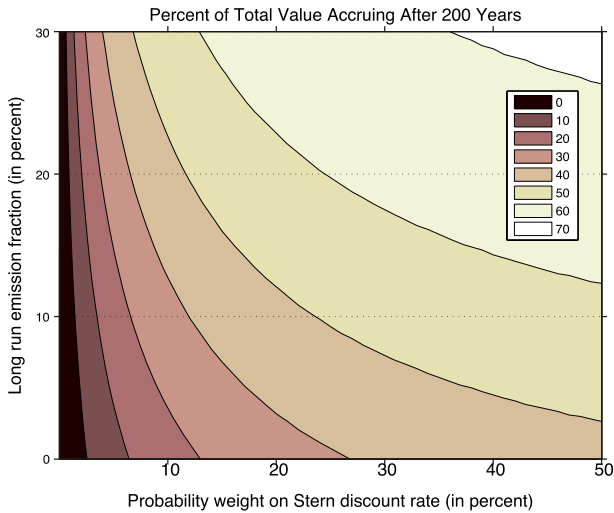


Fig. 4 Percent of current optimal carbon tax attributable to consumption impacts beyond 200 years (in percent). Horizontal axis depicts alternative discounting assumptions. Vertical axis depicts the fraction of emissions that remain in the very long run (ϕ_L)

To present this information, we construct “cumulative value functions” (CVFs) in the spirit of cumulative distribution functions from probability theory. These plot the cumulative fraction of the optimal tax attributable to consumption impacts at or before a given horizon. When presenting results in this section, we have chosen to adopt the carbon cycle model without temperature inertia. This is the scenario for which our conclusions regarding the impact of declining time preference on long run valuation are smallest, and we do this in the interest of being conservative. Results generated in the model *with* temperature inertia are reported in the [Appendix](#).¹³

Figure 2 superimposes four CVFs to demonstrate the earlier claim that conventional discounting overwhelms the long run emission fraction (ϕ_L) in determining the importance of long run outcomes for current policy. For each CVF, discounting is fixed at the Nordhaus (2008) calibration, and the long run emission fraction varies from 0 to 0.3. The optimal tax—thus the denominator used in computing cumulative value as a fraction of total—varies over the range from \$ 38 per ton when $\phi_L = 0$ to \$ 30 per ton when $\phi_L = 0.3$. That the carbon tax decreases in ϕ_L under this discounting scenario is a consequence of the calibration procedure which implies that near-term carbon decay is higher for lower ϕ_L . Thus, these movements in the carbon tax are due to changes in the path of carbon decay over the first hundred years or so and not to differences in the long run path. Despite these small differences, the temporal distribution of cumulative value is nearly identical across scenarios. Moreover, the importance of the long-run, viewed as the fraction of total

¹³Specifically, we report a revised version of Fig. 4, which includes the information presented in Fig. 3.

value attributable to consumption impacts beyond 200 years, is trivial in all cases—less than 0.2 %.

The story changes dramatically when time preference rates decline. Figure 3 fixes the long run emission fraction at zero—the assumption for which the long run matters least—then plots CVFs for the discounting scenarios in Fig. 2.¹⁴ The temporal distribution of cumulative value varies substantially across discounting scenarios. The effect of declining rates on the shape of the CVF is most dramatic at low Stern weights, and it attenuates with further increases. This is driven by Weitzman (1998) limiting result—at even small weights on the Stern view, the path of time preference rates still declines in the long run to the Stern rate. With just 10 percent weight on the Stern view, consumption impacts beyond 200 years account for almost 30 percent of the total tax today; the long run clearly matters.

Increasing the long run emission fraction above zero makes the differences in Fig. 3 even more pronounced. Figure 4 plots long-run importance—fraction of value from impacts beyond 200 years—as a function of the discounting assumptions and the long run emission fraction. Reading from left to right, we see that while the long run does not matter for the Nordhaus calibration examined above, its importance increases immediately when rates decline. Indeed, with just 1 percent weight on Stern and $\phi_L = 0.2$, long run impacts make up over 10 percent of the current tax. Moreover, as the Stern weight increases, the long run emission fraction becomes increasingly important. With 20 percent weight on Stern, long run importance increases from 36 percent to 65 percent as ϕ_L increases from 0 to 0.3.

4 Conclusions

The main critique of economic analyses of climate change that simply adopt a low discount rate is that the implied social cost of carbon does not accurately reflect society's willingness to transfer wealth to future generations. A key source of information concerning people's willingness to transfer resources over times are transactions in financial markets. Because of this, a standard consistency requirement imposed on climate-economy models is that the assumed preferences should give rise to equilibrium returns in the capital market of the model that are consistent with historical interest rates.

Accepting this restriction, a great deal hinges on how policy-makers view the temporal structure of time preference. When evaluating investment decisions that play out within a decade or so the difference is small. But when the consequences of current actions span centuries, as for climate change, the difference becomes crucially important. If we accept the conventional assumption of a constant rate of time preference, long run outcomes are essentially irrelevant for current policy. In contrast, when time preference rates decline, this stark feature of typical integrated assessment models goes away. For example, in our analysis with the long run decay fraction set to the IPCC suggested value of 0.2, consumption impacts beyond 200 years account for half the total carbon tax when the calibrated weight on the Stern rate of time preference is a mere 12 %.

¹⁴The corresponding carbon taxes are different from those reported in Table 1, where the long run emission fraction is 0.2. With $\phi_L = 0$, the corresponding carbon taxes for the discounting scenarios in the table are, respectively, (\$ 38, \$ 67, \$ 96, \$ 126, \$ 157, \$ 189, and \$ 327).

Mirroring this result, the potential for long run features of the carbon cycle to affect the current case for action also hinges on the structure of time preference. With constant time preference, the effect of discounting overwhelms the effect of carbon decay, and there is no room for long run features of the carbon cycle to influence current policy. This probably explains why long-run features of the carbon cycle are rarely emphasized in economic analyses of climate change. But with DRTP, long run features of the carbon cycle have the potential to matter a great deal.

Finally, our findings have implications for interpreting a recent contribution to this journal. Roe and Bauman (2013) argue that fat-tailed uncertainty about climate sensitivity is unlikely to have a large impact on current climate policy due to the long transition time needed for the earth system to adjust to long run equilibrium if climate sensitivity turns out to be very high. They use a model to show that high temperature outcomes associated with high climate sensitivity scenarios do not come about for several centuries. Under conventional economic discounting, even very large impacts at this horizon do not significantly effect current decisions. Our paper does not explicitly consider this issue, but the general insights carry over. Because outcomes several hundred years and beyond have much greater weight in current decision-making under declining time preference, allowing for this presents a challenge to the Roe and Bauman conclusions.

Acknowledgments The authors thank William Brock, Reyer Gerlagh, and three anonymous referees for helpful comments.

Appendix: Technical appendix

A.1 Model

The model is a multi-sector neoclassical growth model that includes a final-goods sector, one fossil energy sector (“coal”) and one clean energy sector (“wind”).¹⁵ Net output in the final-goods sector is determined by

$$Y_t = \omega(S_t)A_{0,t}K_t^\alpha N_{0,t}^{1-\alpha-v} E_t^v,$$

where $\omega(S_t)$ is a multiplicative damage function described below and

$$E_t = (\kappa_1 E_{1,t}^\rho + \kappa_2 E_{2,t}^\rho)^{1/\rho}.$$

Here, $E_{1,t}$ is energy output from the “coal” (or composite fossil) sector and $E_{2,t}$ is energy output from the “wind” (or composite low carbon energy) sector. Both are measured in units of carbon. The parameter $\rho < 1$ pins down the elasticity of substitution between coal and wind, while the κ parameters determine the relative energy-efficiency of each energy source. Output in each energy sector i is linear in labor:

$$E_{i,t} = A_{i,t}N_{i,t}.$$

Labor is mobile across sectors with

$$N_{0,t} + N_{1,t} + N_{2,t} = N_t,$$

¹⁵The analytic formula for the optimal tax is unchanged when there are an arbitrary number of energy sectors of each type.

where N_t is the total workforce in the economy at date t . Both N_t and $A_{i,t}$ are exogenous. The aggregate resource constraint for final consumption goods assumes 100 % depreciation of physical capital each period.¹⁶ Thus,

$$Y_t = C_t + K_{t+1}.$$

In addition, climate change arises because CO2 is produced as a byproduct of burning fossil fuels. The stock of atmosphere carbon accumulates according to

$$S_t - \bar{S} = \sum_{j=0}^{t+H} (1 - d_j) E_{\tau-j}^f,$$

where

$$E_s^f = \sum_{i=1}^n E_{i,s}$$

is total carbon emissions (from fossil sectors) in period s . In the summation above, H is the number of periods between period 0 in the model and the start of the industrial revolution. Thus, the concentration of carbon in the atmosphere is an arbitrary linear function of all prior anthropogenic emissions. $1 - d_k$ is the fraction of a unit of emissions that remains in the atmosphere k periods after it is emitted, and it is given in the paper by the decay formula in Eq. 2.

Cumulative emissions lead to climate damages that impact economic output through a multiplicative “exponential–linear” damage function:

$$\omega(S_t) = \exp(-\gamma_\tau(S_t - \bar{S})). \tag{3}$$

γ_τ , an elasticity, is the percent output loss associated with an extra unit of atmospheric carbon in τ . It is stochastic.

The key innovation in the model is to allow for non-constant time preference. Specifically, utility is discounted with a sequence of potentially non-constant discount factors, $\{\beta_j\}$. For each j , $0 < \beta_j < 1$. The typical case in which the TPR declines implies $\{\beta_j\}$ increasing. For each time horizon k periods ahead, $R_k = \prod_{j=1}^k \beta_j$ is the cumulative discount factor. Using this, the representative household in generation t evaluates future consumption streams using the following “social welfare function”:

$$\mathbb{E}_t \sum_{\tau=t}^T R_{\tau-t} \ln(C_\tau).$$

Utility is logarithmic.

Analytical tractability is a consequence of five key assumptions: log utility, Cobb Douglas production (in capital), “exponential-linear” damages, 100 percent depreciation of physical capital each period, and the exclusion of capital as a factor of production in the various energy sectors.

A.2 Equilibrium concept

Non-constant time preference introduces the potential for time inconsistency. This is true in particular if the problem is solved in a way that (explicitly or implicitly) assumes a

¹⁶This strong assumption is partly offset by the fact that the period length in the calibrated model is a decade. It is relaxed later on.

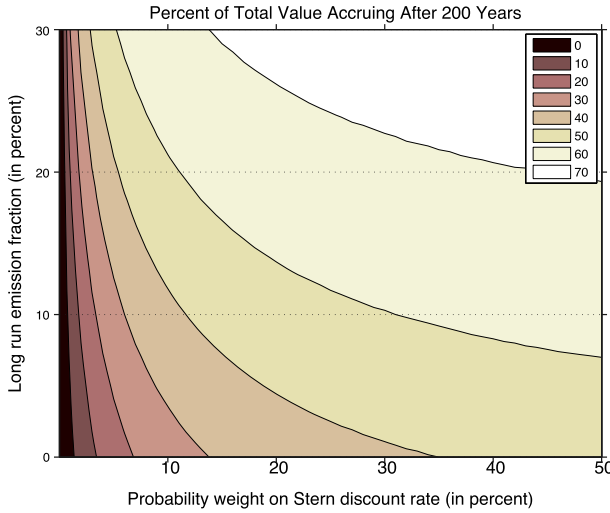


Fig. 5 Percent of current optimal carbon tax attributable to consumption impacts beyond 200 years (in percent). Horizontal axis depicts alternative discounting assumptions. Vertical axis depicts the fraction of emissions that remain in the very long run (ϕ_L)

commitment device. Unfortunately, standard optimal control methods do precisely this. To avoid this problem, we follow Iverson (2013a) in solving for the subgame perfect equilibrium in a finite horizon version of the model. This means that we assume that the policy variables will be chosen each period by the living generation who controls policy for one period only.¹⁷ At the same time, the planner in each period correctly anticipates the way in which future policy makers will set policy (in response to the stock of physical capital and the stock of atmospheric carbon that they inherit from the past). This assumption implies an equilibrium problem that must be solved using backward induction. The corresponding equilibrium is unique. For more details, see Iverson (2013a).

A.3 Optimal carbon taxes

The optimal carbon tax is the tax on the carbon content of fossil fuels in the model economy that would induce firms to choose the abatement path that corresponds to the subgame perfect equilibrium to the planner’s problem without commitment. This is the best one can do on cost-benefit grounds in a world in which commitment is infeasible.

The adopted notation follows Iverson (2013a). He defines $R_{l,m}^{(t)}$ as the price at $t + l$ of a unit of utility received at $t + m$, viewed by an agent in t . For $m \geq l$,

$$R_{l,m}^{(t)} = \begin{cases} \prod_{j=l+1}^m \beta_j^{(t)} & m > l \\ 1 & m = l. \end{cases} \tag{4}$$

With this notation, the optimal tax in fossil sector i can be written

$$\tau_{i,t} = \mathbb{E}_t \left[\sum_{k=0}^{\infty} R_k^{(t)} \gamma_{t+k} (1 - d_k) \cdot \Gamma_k^{(t)} \right] Y_t, \tag{5}$$

¹⁷The period length in the model is a decade.

where

$$\Gamma_k^{(t)} = \frac{\sum_{m=0}^{\infty} \alpha^m R_{k,k+m}^{(t)}}{\sum_{n=0}^{\infty} \alpha^n R_n^{(t)}}.$$

The optimal tax on clean-energy firms is zero. Here $\mathbb{E}_t \gamma_{t+k}$ is the expected damage elasticity parameter in future period $t+k$ given that we are standing today in period t . α is the elasticity of final goods production with respect to physical capital. Y_t is Gross World Product in t .

A. 4 Results with temperature inertia

Temperature inertia is introduced by considering a two-box version of the GL climate module. Maintaining the carbon cycle and climate damage calibration assumptions from GHKT implies $a = (0.2, 0.3144)$, $\eta = (0, 0.0228)$, and $\mathbb{E}_t \gamma_{t+k} = 2.379e - 5$. Meanwhile, temperature inertia is calibrated so peak temperature occurs 70 years after an emission pulse, as in GL and (Nordhaus 2008). This implies $\epsilon = 0.5$.

Comparing Fig. 5 to Fig. 4 shows the sense in which the results in the text are conservative. For any given value of ϕ_L , the impact of increasing the weight on Stern is more pronounced with temperature inertia (Fig. 5) than it is without temperature inertia (Fig. 4). The difference is especially large when the weight on Stern is small.

References

- Archer D (2005) Fate of fossil fuel CO₂ in geologic time. *J Geophys Res Oceans* (1978–2012):110
- Archer D, Brovkin V (2008) The millennial atmospheric lifetime of anthropogenic CO₂. *Clim Chang* 90:283–297
- Archer D, Eby M, Brovkin V, Ridgwell A, Cao L, Mikolajewicz U, Caldeira K, Matsumoto K, Munhoven G, Montenegro A et al (2009) Atmospheric lifetime of fossil fuel carbon dioxide. *Ann Rev Earth Planet Sci* 37:117
- Cropper M, Aydede S, Portney P (1994) Preferences for life saving programs: how the public discounts time and age. *J Risk Uncertain* 8:243–265
- Frederick S, Loewenstein G, O'Donoghue T (2002) Time discounting and time preference: a critical review. *J Econ Lit* 40:351–401
- Fujii T, Karp L (2008) Numerical analysis of non-constant pure rate of time preference: a model of climate policy. *J Environ Econ Manag* 56:83–101
- Gerlagh R, Liski M (2013) Carbon prices for the next thousand years. Working paper
- Gollier C, Zeckhauser R (2005) Aggregation of heterogeneous time preferences. *J Polit Econ* 113:878–896
- Golosov M, Hassler J, Krusell P, Tsyvinski A (2014) Optimal taxes on fossil fuel in general equilibrium. *Econometrica* 82:41–88
- Heal GM, Millner A (2013) Discounting under disagreement, No. w18999
- Heal GM, Millner A (2014) Agreeing to disagree on climate policy. *Proc Natl Acad Sci* 111:3695–3698
- Hope C (2008) Discount rates, equity weights and the social cost of carbon. *Energy Econ* 30:1011–1019
- Hyyönen R, Ågren G, Linder S, Persson T, Cotrufo MF, Ekblad A, Freeman M, Grelle A, Janssens IA, Jarvis PG et al (2007) The likely impact of elevated [CO₂], nitrogen deposition, increased temperature and management on carbon sequestration in temperate and boreal forest ecosystems: a literature review. *New Phytol* 173:463–480
- Iverson T (2013a) Optimal carbon taxes with non-constant time preference. MPRA Working Paper No. 49588
- Iverson T (2013b) Minimax regret discounting. *J Environ Econ Manag* 66:598–608
- Karp L (2005) Global warming and hyperbolic discounting. *J Public Econ* 89:261–282
- Karp L (2007) Non-constant discounting in continuous time. *J Econ Theory* 132:557–568
- Karp L (2014) Provision of a public good with altruistic overlapping generations and many tribes. Working Paper
- Karp L, Tsur Y (2011) Time perspective and climate change policy. *J Environ Econ Manag* 62(1):1–14
- Laibson D (1997) Golden eggs and hyperbolic discounting. *Q J Econ* 112:443–478

- Layton DF, Brown G (2000) Heterogeneous preferences regarding global climate change. *Rev Econ Stat* 82:616–624
- Li C-Z, Löfgren K-G (2000) Renewable resources and economic sustainability: a dynamic analysis with heterogeneous time preferences. *J Environ Econ Manag* 40:236–250
- Loewenstein G (1987) Anticipation and the valuation of delayed consumption. *Econ J* 40:236–250
- Norby RJ, Zak DR (2011) Ecological lessons from free-air CO₂ enrichment (FACE) experiments. *Ann Rev Ecol Evol Syst* 42:181
- Nordhaus W (2008) *A question of balance*. Yale University Press
- Nordhaus W (2013) *The climate casino: risk, uncertainty, and economics for a warming world*. Yale University Press
- Nordhaus W, Boyer J (2000) *Warming the world: economic models of global warming*. MIT Press, Cambridge
- Phelps E, Pollak R (1968) On second-best national saving and game-equilibrium growth. *Rev Econ Stud* 35:185–199
- Pigou AC (1924) *The economics of welfare*. Transaction Publishers
- Rezai A, van der Ploeg F (2014) Robustness of a simple rule for the social cost of carbon. CESIFO Working Paper No. 4703
- Roe GH, Bauman Y (2013) Climate sensitivity: should the climate tail wag the policy dog? *Clim Chang* 117:647–662
- Rubinstein A (2003) Economics and psychology? The case of hyperbolic discounting. *Int Econ Rev* 44(4):1207–1216
- Sabine CL, Feely RA, Gruber N, Key RM, Lee K, Bullister JL, Wanninkhof R, Wong C, Wallace DW, Tilbrook B et al (2004) The oceanic sink for anthropogenic CO₂. *Science* 305:367–371
- Samuelson P (1937) A note on measurement of utility. *Rev Econ Stud* 4:155–161
- Solomon S, Qin D, Manning M, Chen Z, Marquis M, Averyt KB, Tignor M, Miller HL (eds) (2008) *Climate change 2007: the physical science basis*. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge
- Stern N (2007) *The economics of climate change: the stern review*. Cambridge University Press, Cambridge
- Strotz R (1955) Myopia and inconsistency in dynamic utility maximization. *Rev Econ Stud* 23:165–180
- Tans P (1998) *Why carbon dioxide from fossil fuel burning wont go away*. Oxford University Press, New York
- Thaler R (1981) Some empirical evidence on dynamic inconsistency. *Econ Lett* 8:201–207
- Tol R (1997) On the optimal control of carbon dioxide emissions: an application of FUND. *Environ Model Assess* 2:151–163
- van der Ploeg F, Withagen C (2012) Too much coal, too little oil. *J Public Econ* 96:62–77
- Van Vuuren DP, Edmonds J, Kainuma M, Riahi K, Thomson A, Hibbard K, Hurtt GC, Kram T, Krey V, Lamarque J-F et al (2011) The representative concentration pathways: an overview. *Clim Chang* 109:5–31
- Weitzman W (1998) Why the far-distant future should be discounted at its lowest possible rate. *J Environ Econ Manag* 36:201–208
- Weitzman W (2009) On modeling and interpreting the economics of catastrophic climate change. *Rev Econ Stat* 91:1–19