

**Intergenerational inequality aversion, growth and the role of damages:  
Occam's rule for the global carbon tax**

**Armon Rezai<sup>1</sup> and Frederick van der Ploeg<sup>2,3</sup>**

**Abstract**

We derive a simple rule for a nearly optimal carbon tax that can be implemented and tested in a decentralized market economy. Our simple rule depends on the effect of the pure rate of time preference, growth and intergenerational inequality aversion and basic parameters of the carbon cycle, but also on any adverse effects of global warming on economic growth and mean reversion in climate damages. The performance of the simple rule is excellent and yields only tiny welfare losses compared with the true welfare optimum under a wide range of perturbations including some extreme runs designed to severely road-test the rule. Our IAM allows for scarce fossil fuel and endogenous energy transitions and generates cumulative carbon emissions and stranded assets which are also well predicted by our rule.

**Keywords:** simple rule, welfare losses, optimal carbon tax, intergenerational inequality aversion, growth, alternative climate damage specifications, stranded assets, transition times

**JEL codes:** H21, Q51, Q54

*Revised February 2016*

---

<sup>1</sup> Department of Socioeconomics, Vienna University of Economics and Business, Welthandelsplatz 1, 1020 Vienna, and IIASA, Schlossplatz 1, 2361 Laxenburg, Austria. Corresponding author: [armon.rezai@wu.ac.at](mailto:armon.rezai@wu.ac.at). Rezai is grateful for financial support from the OeNB Anniversary Fund (grant no. 15330) and the Schrödinger fellowship of Austrian Science Fund (FWF): J 3633.

<sup>2</sup> Department of Economics, Oxford University, Manor Road Building, Oxford OX1 3 UQ, U.K. Email: [rick.vanderploeg@economics.ox.ac.uk](mailto:rick.vanderploeg@economics.ox.ac.uk). Also affiliated with the St. Petersburg State University, 7/9 Universitetskaya nab., St. Petersburg, 199034 Russia and the VU University Amsterdam. Van der Ploeg is grateful for support from the ERC Advanced Grant 'Political Economy of Green Paradoxes' (FP7-IDEAS-ERC Grant No. 269788).

<sup>3</sup> We are very grateful to Cees Withagen of VU University Amsterdam for his involvement in the early stages of this project and for his detailed comments on earlier versions. We have benefited from his guidance in those stages of the project and from an earlier paper on additive versus multiplicative damages (Rezai et al., 2014). We are also grateful for the helpful and constructive comments and suggestions of Chuck Mason and a reviewer, which have led to improvements including an improved focus of the paper and a better road-testing of our simple rule.

## 1. Introduction

Integrated assessment models of climate change (IAMs) aim to integrate economics and climate science and to assess the impact the economy has on the climate and vice versa. Such IAMs are a crucial step in the design of optimal policies to fight the negative and potentially large effects of climate change on economic well-being. In quantifying the cost of the climate externality, many estimates of the optimal social cost of carbon have been obtained by maximizing welfare subject to the constraints of the IAM (for a summary, see Tol, 2015). Although a lot of valuable insights have been obtained from such studies, these projections for the carbon tax are necessarily conditioned on a set of specific parameter assumptions and it is difficult to comprehend the important drivers of the presented discretionary paths. A simple rule for the optimal carbon tax, like the much appreciated rules for fiscal policy and monetary policy (e.g., the Taylor rules for the nominal interest rate), would be easier to understand and communicate and can easily be adapted to changing economic circumstances.

Our main contribution is to show that such a simple policy rule for the optimal carbon tax that can be calculated on the back of an envelope yields negligible welfare losses and approximates the true numerical first-best optimal carbon tax of an IAM very well. To do this, we test for the first time a rule in the decentralized market economy mode of our IAM of growth, development and climate change as the first fundamental theorem of welfare economics does not hold if the simple rule rather than the first-best climate policy is implemented. Our simple rule allows for the effect on the optimal carbon tax of the pure rate of time preference, growth, intergenerational inequality aversion and basic parameters of the carbon cycle as in the classic paper of Nordhaus (1991), but also of the adverse effects of global warming on economic growth, mean reversion in climate damages and damages that do not have a unit elasticity of climate damages with respect to global GDP as recently discussed by Dietz and Stern (2015) and Moore and Diaz (2015).

Our simple rule is motivated by the observation that the Ramsey growth dynamics move much more quickly than the carbon cycles and temperature modules of most IAMs. Indeed, the so-called Iron Law refers to a convergence speed of 2 percent per year or a mean lag of 50 years (e.g., Barro, 2015) which is much faster than the carbon dynamics. Our approximation of the optimal carbon tax rests on the assumption that the Ramsey growth dynamics has converged and that the economy is moving along a trend growth path.<sup>1</sup> The resulting simple rule is then implemented in the decentralized market economy mode of our IAM to compare the welfare

---

<sup>1</sup> Our approach is akin to the analysis of fast-moving (economic) and slow-moving (carbon cycle) systems which is a technique often used in engineering analysis of dynamic systems (e.g., Majda, et al. 2001).

loss with the first-best optimal carbon tax. Anything that slows down the Ramsey growth dynamics (e.g., a much higher share of capital in value added) will worsen the performance of our rule. We use this insight to conduct some extreme road-testing of our rule which indicate that the performance does not worsen much, especially when compared to the welfare loss under business as usual.

Other papers that deal with simple rules for the global carbon tax are Golosov et al. (2014) and van den Bijgaart et al. (2013). The rule in these papers do not allow for the effects of effects of global warming on growth, mean reversion in damages, and non-multiplicative climate damages. It is only optimal in general equilibrium if the coefficient of intergenerational inequality aversion is unity, there is 100 percent depreciation of capital in each period, and the damage function is a negative exponential function of the atmospheric carbon stock (Golosov et al., 2014). Our simple rule generalizes this rule and shows that the welfare losses from using such a simple rule in a decentralized market economy are negligible compared with the first-best optimal carbon tax. These papers do not consider the welfare losses from simple rules and do not test them in a decentralized market economy mode of an IAM, which is crucial as the first fundamental theorem of welfare economics no longer holds if carbon is priced by the simple rules rather than the optimal tax. Furthermore, these papers do not highlight the effects of intergenerational inequality aversion on the optimal price of carbon.<sup>2</sup>

The generic micro-founded IAM of Ramsey growth and climate change we put forward to test our simple rule allows, in contrast to the cast-iron benchmark DICE (Nordhaus, 2014), for substitution between energy and a capital-labor composite, exhaustibility of fossil fuel and forward-looking dynamics of the scarcity rent, and extraction costs that rise as reserves are depleted. Due to rising extraction costs, a rising carbon tax and possibly a falling cost of renewable energy, fossil fuel is eventually priced out of the market and replaced by the carbon-free substitute. Our IAM also gives the optimal date to switch to the carbon-free era, the amount of fossil fuel reserves to leave untouched in the earth, and cumulative carbon emissions.<sup>3</sup>

---

<sup>2</sup> Van den Bijgaart et al. (2013) perform a ‘Monte Carlo’ simulation by fitting a truncated log-normal distribution function to 1000 realizations of the ‘optimal’ paths for the SCC corresponding to 1000 Monte Carlo simulations for 12 parameters and show that a less general version of our simple rule explains over 99% of within-model variations in these realizations (by fitting their model to the DICE model). The calibration of the probability density functions of their parameters is necessarily arbitrary and they do not calculate the correct stochastic social optimum of their IAM accounting for prudence as, e.g., Jensen and Traeger (2014) or Cai et al. (2015). Most crucially, they do not obtain and test how well simple rules perform in the market economy and are thus unable to quantify the welfare losses associated with the policy errors resulting from using the approximate rather than the actual rule.

<sup>3</sup> IPCC (2014) states that cumulative emissions cannot exceed 700-860 GtC if global warming is to remain below 2°C. By 2011 520 GtC have been emitted, giving a carbon budget of only 270 GtC. Proven

In section 2 we derive our simple rule for the optimal carbon tax. Section 3 sets out our IAM of climate change and Ramsey economic growth and shows how our simple rule can be implemented in a decentralized market economy. Section 4 uses a calibrated version of our IAM to highlight how well our simple rule performs compared with the first-best optimal carbon tax and the no-policy scenario. Section 5 illustrates the drivers of our simple rule and shows that the rule performs well with negligible welfare losses across a large number of variations in these drivers. Section 6 offers some extreme road-tests of our rule. Section 7 concludes.

## 2. Simple rules for the optimal carbon tax

The simple rule for the global carbon tax relies on the basics of the carbon cycle which are presented in section 2.1 and used in section 2.2 to heuristically derive an expression for the social cost of carbon. This rule is easy to understand and allows for general degrees of intergenerational inequality aversion, trend growth in living standards, population growth, and a general elasticity of climate damages with respect to GDP. Section 2.3 extends our simple rule to also allow for mean reversion in climate damages and for an adverse impact of global warming on productivity growth. Web appendix D further discusses the potential nature of damages.

### 2.1. Basics of the carbon cycle

Let a fraction  $\varphi_L$  of carbon emissions stay up permanently in the atmosphere and of the remainder the fraction  $(1-\varphi_0)$  be absorbed within the same period and the fraction  $\varphi_0$  decay at the rate  $\varphi$ , so that of an impulse of 1tC a fraction  $z_t \equiv \varphi_L + \varphi_0(1-\varphi_L)(1-\varphi)^{t-1}$  is still in the atmosphere after  $t$  periods. We employ an annual time grid and suppose that after 3 decades half of carbon has left the atmosphere ( $z_{30} = 0.5$ ), a fifth of carbon stays up in the atmosphere forever and the remainder has a half-time of 300 years ( $(1-\varphi)^{300-1} = 0.5$ ), hence we get  $\varphi_L = 0.2$ ,  $\varphi = 0.00231$  and  $\varphi_0 = 0.401$ .<sup>4</sup> With a distributed lag between temperature and

atmospheric carbon of  $\varphi_T$ ,  $\xi_t = \frac{1}{\varphi_T} \sum_{s=1}^t \left(1 - \frac{1}{\varphi_T}\right)^{t-s} z_s$  is the remainder relevant for the impact of temperature on marginal climate damages at time  $t$ , so the impact of 1 ton of carbon emitted at time 0 on temperature at time  $t$  is

---

fossil fuel resources exceed this budget by far. Here we optimize the optimal amount of stranded assets and degree of global warming.

<sup>4</sup> This way of modelling the climate system is based on Golosov et al. (2014) who employ a decadal instead of our annual time grid and derive  $\varphi = 0.0228$  and  $\varphi_0 = 0.393$ .

$$(1) \quad \xi_t = \varphi_L \left[ 1 - \left( 1 - \frac{1}{\varphi_T} \right)^t \right] + \left( \frac{\varphi_0(1-\varphi_L)}{1-\varphi_T\varphi} \right) \left[ (1-\varphi)^t - \left( 1 - \frac{1}{\varphi_T} \right)^t \right].$$

## 2.2. The social cost of carbon, growth and intergenerational inequality aversion

The SCC is the present value of all future marginal damages from emitting one ton of carbon:

$$SSC_0 \equiv \sum_{t=0}^{\infty} (1+r_t)^{-t} \omega_t \text{ with } \omega_t = \chi GDP_t^\varepsilon GDP_0^{1-\varepsilon} \xi_t \text{ the marginal damage in period } t \text{ and } r_t \text{ the}$$

average discount rate in period  $t$ . Golosov et al. (2014) have  $\chi = 0.02379$  giving marginal climate damages of 2.379% of GDP/tC. Since initial global GDP is \$70 trillion, initial marginal damages are \$1.67/tC.<sup>5</sup> Note that  $\varepsilon = 1$  gives *multiplicative* marginal climate damages as in Golosov et al. (2014), which corresponds closely to the specification used in DICE for the relevant range of stocks of atmospheric carbon and a climate sensitivity of 3 (Nordhaus, 2008). In contrast,  $\varepsilon = 0$  gives *additive* marginal climate damages which do not grow in line with GDP.<sup>6</sup> In general,  $\varepsilon$  is the elasticity of marginal damages with respect to GDP.

Let the trend exogenous growth rate of consumption be  $g$ , the rate of population growth be  $n$ , the rate of time impatience be  $\rho$  and the coefficient of relative intergenerational inequality aversion (IIA) be  $\Phi$ . The Euler equation then gives the long-run interest rate as  $r = \rho + \Phi(g - n)$  and we suppose  $r > g$  holds (see appendix A). Growth in living standards leads to wealthier future generations that require a higher interest rate, especially if IIA is large, since current generations are less prepared to sacrifice current consumption.

Derivations of Result 1 and all other labelled results in this paper are relegated to appendix A.

**Result 1:** A first-order approximation to the SCC along a trend growth path is:

$$(2) \quad SCC_t \equiv \left( \frac{\varphi_L}{\rho - \varepsilon n + (\Phi - \varepsilon)(g - n)} + \frac{\varphi_0(1-\varphi_L)}{\rho - \varepsilon n + (\Phi - \varepsilon)(g - n) + \varphi} \right) \times \\ \left( 1 + \varphi_T [\rho - \varepsilon n + (\Phi - \varepsilon)(g - n)] \right)^{-1} \chi GDP_t^\varepsilon GDP_0^{1-\varepsilon},$$

where we assume all denominators in (2) are strictly positive.  $\square$

<sup>5</sup> Golosov et al. (2014) derive this damage parameter as the ex-ante value to a high and low damage uncertainty. This value is significantly lower than their value  $\chi = 0.053$  which reproduces the damage function of Nordhaus (2008) best.

<sup>6</sup> For example, sea level rise may destroy part of production capacity, but at the cost of building dykes the loss can be mitigated. However, there is no reason why higher gross production requires higher dykes and, hence, higher net production losses in absolute terms. Although the damage costs of climate change are often expressed in percentages of gross product, sometimes the costs, such as for ecosystems, sea level rise, and water resource, are expressed in (billion) dollars (ToI, 2002a, b).

A special case of (2) prevails if marginal damages are proportional to GDP ( $\varepsilon = 1$ ):

$$(2') \quad SCC_t \equiv \left( \frac{\varphi_L}{\rho - n + (\Phi - 1)(g - n)} + \frac{\varphi_0(1 - \varphi_L)}{\rho - n + (\Phi - 1)(g - n) + \varphi} \right) \times \\ (1 + \varphi_T [\rho - n + (\Phi - 1)(g - n)])^{-1} \chi GDP_t.$$

If one further abstracts from a temperature lag ( $\varphi_T = 0$ ) and population growth ( $n = 0$ ), and restricts intergenerational inequality aversion to 1, we get

$$(2'') \quad SCC_t \equiv \left( \frac{\varphi_L}{\rho} + \frac{\varphi_0(1 - \varphi_L)}{\rho + \varphi} \right) \chi GDP_t \text{ or } SCC_t \equiv \beta \left( \frac{\varphi_L}{1 - \beta} + \frac{\varphi_0(1 - \varphi_L)}{1 - (1 - \varphi)\beta} \right) \chi GDP_t,$$

where  $\beta = 1/(1 + \rho)$  is the time impatience factor. Golosov et al. (2014) show that (2'') is the optimal carbon tax in a general equilibrium model with Cobb-Douglas production technology, full depreciation of capital each decade, and an energy production technology that only requires labour as input. Golosov et al. (2014, p.77) also give an approximate rule which corresponds to general IIA but without temperature lag, population growth or non-proportional climate damages.<sup>7</sup> In section 3.2 we discuss optimality of (2) further and present general optimality conditions.

With multiplicative damages the SCC (2') rises in line with GDP. The SCC is high for low time impatience and slow decay of atmospheric carbon. Expression (2'') over-estimates the true SCC since it ignores the temperature lag and the IIA typically exceeds one. If it is unethical to discount welfare of future generations (Stern, 2007; Gollier, 2013),  $\beta = 1$  and the SCC in (2'') is infinite. But with IIA  $> 1$  and positive growth in living standards, (2') indicates a finite SCC.

Equation (2) gives the following additional insights. First, given growth in living standards ( $g$ ), higher population growth (higher  $n$ ) boosts the SCC only in as far as climate damages increase with GDP ( $\varepsilon > 0$ ). Also, population growth boosts trend growth and thus boost the growth rate of the SCC. Second, with high degrees of IIA or climate damages that do not depend much on GDP ( $\Phi > \varepsilon$ ), growth in living standards (higher  $g$ ) boosts the interest rate and thus depresses the SCC. The lower temperature correction reinforces this effect. As future generations are better off, current generations are less prepared along a growth path to make sacrifices to combat global warming. However, with low enough IIA, growth in living standards boosts the SCC. Third, higher IIA (higher  $\Phi$ ) means that current generations are less prepared to temper future climate damages and thus the SCC is lower. Fourth, a longer lag between temperature

<sup>7</sup> Their approximate rule is  $SCC_t = \left( \varphi_L / (1 - \beta(1 + g)^{1 - \Phi}) + \varphi_0(1 - \varphi_L) / (1 - (1 - \varphi)\beta(1 + g)^{1 - \Phi}) \right) \chi GDP_t$ . This can be approximated by  $SCC_t \equiv \left( \varphi_L / (\rho + (\Phi - 1)g) + \varphi_0(1 - \varphi_L) / (\rho + (\Phi - 1)g + \varphi) \right) \chi GDP_t$ , a special case of (2').

and atmospheric carbon (higher  $\varphi_T$ ) depresses the SCC relative to GDP. Finally, if marginal damages react less than proportionally with GDP ( $\varepsilon < 1$ ), the initial SCC is lower than with multiplicative damages, especially if the economic growth rate is high. The effective discount rate has to correct less for growing climate damages, since growth of the SCC is only  $\varepsilon < 1$  times growth of GDP. Additive marginal damages ( $\varepsilon = 0$ ) give a flat time path for the SCC.

### 2.3. Mean reversion in climate damages

Pindyck (2013, p.867) claims that temperature does not affect the level but the growth rate of GDP.<sup>8</sup> Empirical evidence supports this hypothesis, especially for less developed economies (Dell et al., 2012; Bansal and Ochoa, 2011, 2012). We, therefore, introduce mean reversion in climate damages and see how this affects our rule for the SCC. To keep it simple, we set  $\varepsilon = 1$  and  $n = 0$ . Gross output is  $Y_t = \tilde{D}_t \bar{A}_t Z_t = A_t Z_t$  with  $\bar{A}_t = (1 + g_A)^t \bar{A}_0$  and  $\tilde{D}_t = \tilde{D}_{t-1}^{\varphi_\chi} e^{-\chi(E_t - E_0)}$  with  $g_A$  trend total factor productivity growth and  $\varphi_\chi$  the degree of mean reversion in damages.

**Result 2:** With mean reversion in climate damages,  $\varepsilon = 1$  and  $n = 0$ , a first-order approximation to the SCC along a trend growth path is

$$(3) \quad \begin{aligned} \frac{SSC_t}{\chi GDP_t} &= \frac{\varphi_L}{1 + \varphi_\chi [\rho - 1 + (\Phi - 1)g]} \frac{1}{\rho + (\Phi - 1)g} \\ &+ \frac{\varphi_0(1 - \varphi_L)}{1 - \varphi_T \varphi} \frac{1}{1 + \varphi_\chi [\rho - 1 + (\Phi - 1)g + \varphi]} \frac{1}{\rho + (\Phi - 1)g + \varphi} \\ &- \left\{ \varphi_L + \frac{\varphi_0(1 - \varphi_L)}{1 - \varphi_T \varphi} \right\} \frac{1}{1 - \varphi_\chi \left[ 1 - \rho - (\Phi - 1)g - \frac{1}{\varphi_T} \right]} \frac{1}{\rho + (\Phi - 1)g + \frac{1}{\varphi_T}}. \quad \square \end{aligned}$$

Mean reversion in climate damages give rise to extra terms in the first two lines of (3), which boost the SCC. The negative effects of the temperature lag on the SCC are given by the third line. Our formulation of mean reversion implies  $A_t = (1 + g_A)^{\varphi_\chi} \bar{A}_t^{1 - \varphi_\chi} A_{t-1}^{\varphi_\chi} e^{-\chi(E_t - E_0)}$ . Hence, if  $\varphi_\chi = 0$ ,  $\omega_t = \chi GDP_t \xi_t$  and a carbon impulse causes only damage to the *level* of GDP in which case it can be shown that (3) indeed boils down to (2') once  $n$  is set to zero). However, if a carbon impulse causes damage to the *trend growth rate* of GDP, we have  $\varphi_\chi = 1$ ,  $\tilde{D}_t = \tilde{D}_{t-1} e^{-\chi(E_t - E_0)}$ , and  $A_t = (1 + g_A) A_{t-1} e^{-\chi(E_t - E_0)}$ . In that case, it can be shown that a carbon impulse that hurts trend growth creates a much higher SCC than one that hurts the level of GDP as equation then (3) boils down (abstracting from a temperature lag) to

<sup>8</sup> The reason is that adaptation to climate change requires resources that could have been used for R&D.

$$(3') \quad \begin{aligned} SCC_t^{\text{growth damages}} &= \left( \frac{\varphi_L}{[\rho + (\Phi - 1)g]^2} + \frac{\varphi_0(1 - \varphi_L)}{[\rho + (\Phi - 1)g + \varphi]^2} \right) \chi GDP_t \\ &> SCC_t^{\text{temporary damages}} &= \left( \frac{\varphi_L}{\rho + (\Phi - 1)g} + \frac{\varphi_0(1 - \varphi_L)}{\rho + (\Phi - 1)g + \varphi} \right) \chi GDP_t. \end{aligned}$$

### 3. An integrated assessment model of Ramsey growth and energy transitions

Results 1 and 2 offer simple rules for the SCC which are heuristically derived and do not necessarily correspond to the optimal carbon taxes. Here we put forward a general equilibrium IAM and derive the first-best optimal SCC. The first fundamental theorem of welfare economics implies that the social optimum is sustained in a market economy if the carbon tax is set to the optimal SCC and the revenue rebated as lump-sum transfers. If the global carbon tax is set to the simple rules for the SCC given in Results 1 and 2, the first fundamental theorem of welfare economics no longer holds as the rule is only an approximation to the “correct” carbon price. Section 3.2 therefore shows how to implement the simple rules for the carbon tax in a decentralized market economy.

#### 3.1. The first-best optimal SCC and global carbon tax

Our IAM has the same utilitarian social welfare function as used in section 2:

$$(5) \quad \sum_{t=0}^{\infty} (1 + \rho)^{-t} L_t U(C_t / L_t) = \sum_{t=0}^{\infty} (1 + \rho)^{-t} L_t \frac{(C_t / L_t)^{1-\Phi}}{1-\Phi}, \quad \Phi \geq 0,$$

where  $C_t$  is aggregate consumption during period  $t$ ,  $L_t$  the population size at the outset of period  $t$ , and  $\Phi \equiv -\frac{U'(C_t/L_t)}{(C_t/L_t)U''(C_t/L_t)} > 0$  the coefficient of relative IIA. The ethics of climate

policy depend on how much the welfare of future generations is discounted (i.e.,  $\rho$ ) and on how small IIA ( $\Phi$ ) is or how easy it is to substitute current for future consumption per head. Optimal climate policy maximizes (5) subject to constraints (6)-(9) below.

First, gross output in period  $t$ ,  $Z(K_t, L_t, F_t + R_t)$ , is produced using capital at the outset of period  $t$ ,  $K_t$ , labor during period  $t$ ,  $L_t$ , and energy. Energy is either fossil fuels (oil, natural gas and coal),  $F_t$ , or renewables,  $R_t$  (e.g., wind or solar energy). The production function has imperfect factor substitution between capital, labor and energy, but has perfect substitution between fossil fuel and renewables. Renewable energy is infinitely elastically supplied at exogenously decreasing unit cost,  $b_t$ . Fossil fuel unit extraction cost in period  $t$  is  $G(S_t)$  with  $G'(S_t) < 0$ , where  $S_t$  denotes remaining reserves at the start of period  $t$ , hence extraction is more costly when there

are fewer reserves left. We assume sustained technical progress in aggregate and renewable energy production and an exogenous time profile for population growth. Damages can be multiplicative or additive. Production after depreciation at the rate  $\delta$ , cost of energy use and climate damages is used for consumption  $C_t$  and investment:

$$(6) \quad K_{t+1} = Z(K_t, L_t, F_t + R_t) - D(T_t)Z(K_t, L_t, F_t + R_t)^\varepsilon Z_0^{1-\varepsilon} - G(S_t)F_t - b_t R_t - C_t + (1 - \delta)K_t, \quad \forall t \geq 0,$$

where initial capital  $K_0$ , energy use  $F_0 + R_0$ , and gross output  $Z_0 = Z(K_0, L_0, F_0 + R_0)$  are given, and climate damages,  $D(T_t)$ , increase with global mean temperature  $T_t$ .

Second, given initial reserves  $S_0$ , finite fossil fuel reserves develop as follows:

$$(7) \quad S_{t+1} = S_t - F_t, \quad \forall t \geq 0, \quad \sum_{t=0}^{\infty} F_t \leq S_0.$$

Third, the two-box carbon cycle used in section 2 is based on Golosov et al. (2014):

$$(8) \quad E_{1,t} = E_{1,t-1} + \varphi_L F_t, \quad E_{2,t} = (1 - \varphi)E_{2,t-1} + \varphi_0(1 - \varphi_L)F_t, \quad \forall t \geq 1,$$

where  $E_{1,t}$  and  $E_{2,t}$  denote the permanent and transient end-of-period stocks of atmospheric carbon and  $E_{1,0}$  and  $E_{2,0}$  are given (Archer, 2005; Archer et al., 2009). Gerlagh and Liski (2012) allow for the delay between carbon and global warming. We capture this by assuming that the effective stock of atmospheric carbon and thus temperature follow from a distributed lag:

$$(9) \quad E_t = \frac{1}{\varphi_T} (E_{1t} + E_{2t}) + \left(1 - \frac{1}{\varphi_T}\right) E_{t-1}, \quad \forall t \geq 1, \quad T_t = \omega \frac{\ln(E_t / 596.4)}{\ln(2)}, \quad \forall t \geq 0,$$

where temperature is the deviation of global mean temperature from pre-industrial temperature in degrees Celcius, 596.4 GtC the IPCC figure for the pre-industrial stock of atmospheric carbon, and  $\omega$  is the climate sensitivity (the rise in temperature after a doubling of the total stock of atmospheric carbon, typically 3).<sup>9</sup>

**Result 3:** *The social optimum entails choosing  $\{C_t, F_t, R_t\}_{t=1}^{\infty}$  to maximize (5) and satisfies (6)-(9), the Euler equation for consumption growth*

$$(10) \quad \frac{C_{t+1} / L_{t+1}}{C_t / L_t} = \left[ (1 + r_{t+1}) / (1 + \rho) \right]^{1/\Phi}, \quad \forall t \geq 0,$$

---

<sup>9</sup> Alternatively, we can use  $T_t = \omega / \varphi_T \ln \left[ (E_{1t} + E_{2t}) / 596.4 \right] / \ln[2] + (1 - 1/\varphi_T) T_{t-1}$ , which gives a similar expression for the SCC as (13) below (Rezai and van der Ploeg, 2015).

with  $r_{t+1} \equiv \Gamma_{t+1} Z_{K_{t+1}} - \delta$  and  $\Gamma_t \equiv \left[ 1 - \varepsilon D(E_t) (Z(K_t, L_t, F_t + R_t) / Z_0)^{\varepsilon-1} \right]$ , and the efficiency conditions for energy use

$$(11a) \quad \Gamma_t Z_{F_t+R_t} (K_t, L_t, F_t + R_t) \leq G(S_t) + s_t + \theta_t, \quad F_t \geq 0, \quad \text{c.s.}, \quad \forall t \geq 1,$$

$$(11b) \quad \Gamma_t Z_{F_t+R_t} (K_t, L_t, F_t + R_t) \leq b_t, \quad R_t \geq 0, \quad \text{c.s.}, \quad \forall t \geq 1.$$

The scarcity rent and optimal SCC are

$$(12) \quad s_t = - \sum_{\zeta=1}^{\infty} \Delta_{t+\zeta} G'(S_{t+\zeta}) F_{t+\zeta}, \quad \forall t \geq 1,$$

$$(13) \quad \theta_t = \sum_{\zeta=0}^{\infty} \left[ \left\{ \varphi_L + \varphi_0 (1 - \varphi_L) (1 - \varphi)^{\zeta} \right\} \Delta_{t+\zeta} \theta_{Tt+\zeta} \right] / \varphi_T,$$

with  $\theta_{Tt} = \sum_{\zeta=0}^{\infty} \left[ (1 - 1/\varphi_T)^{\zeta} \Delta_{t+\zeta} D'(E_{t+\zeta}) Z_{t+\zeta}^{\varepsilon} Z_0^{1-\varepsilon} \right], \forall t \geq 1,$

the compound discount factors  $\Delta_{t+\zeta} \equiv \prod_{\zeta=1}^{\zeta} (1 + r_{t+\zeta})^{-1}, \zeta \geq 1$  and  $\Delta_t = 1, \forall t \geq 1$ , and initial stocks for  $K_0, S_0, E_{1,0}$ , and  $E_{2,0}$  given.  $\square$

In fact, our simulations in sections 4-6 extend the IAM presented above for the more complex and realistic carbon and temperature cycles and damage specification of Nordhaus (2014).<sup>10</sup> We also use an extension to allow for mean reversion in climate damages, but for sake of brevity we omit the analytical derivation of the optimal SCC for that case.

Equation (10) states that growth in consumption per capita increases with the social return on capital (net of depreciation) and decreases with the rate of time impatience. Equation (11a) implies that, if fossil fuel is used, its marginal product should equal its marginal extraction cost,  $G(S_t)$ , plus its scarcity rent plus the SCC. If the marginal product of fossil fuel is below its total marginal cost, it is not used. Equation (11b) states that, if renewable energy is used, its marginal product must equal its marginal cost,  $b_t$ . Although our IAM might display simultaneous use of fossil fuel and renewable energy, it rarely occurs and at most during one year.<sup>11</sup>

Equation (12) follows from the Hotelling rule and states that the scarcity rent of keeping an extra unit of fossil fuel unexploited equals the present discounted value of all future reductions

<sup>10</sup> This allows for separate dynamics of the stocks of carbon in the atmosphere and the upper and lower parts of the ocean as well as separate dynamics in the atmospheric and oceanic heat reservoirs and non-linear forcing relations between carbon and temperature.

<sup>11</sup> The period of simultaneous use increases in length with the share of energy in GDP. For historically observed values, this period only lasts for at most one period.

in fossil fuel extraction costs.<sup>12</sup> Finally, (13) states that the SCC is the present discounted value of all future marginal global warming damages from burning an additional unit of fossil fuel today. In contrast to section 2, it allows for the time-varying nature of the interest rate as well as the permanent and transient components of atmospheric carbon in the simple two-box carbon cycle and the temperature lag.

### 3.2. Implementing the simple rule in the decentralized market economy

The first fundamental theorem of welfare economics states that if markets are complete, a competitive market economy is Pareto efficient. Given that the government has enough instruments available, the theorem implies that the government can reproduce the social planner's first best by setting the carbon tax equal to the first-best SCC (13), i.e.,  $\tau_t = \theta_t \forall t \geq 1$ , and rebating revenue in lump-sum fashion.<sup>13</sup> If the government cannot set the global carbon tax to the first-best SCC according to (13) and, thereby, correct for the missing price signal in the energy market, it might set it to the simple rule for the social cost of carbon (2) or (3). In fact, it turns out that along a balanced growth the optimal carbon tax of the full IAM corresponds exactly to expression (2) derived heuristically in section 2.

**Result 4:** *If (i) the economy is on its balanced growth path and (ii) marginal damages are proportional to GDP, the SCC is given by (2). □*

To see how well the simple rule performs, we can no longer rely on the first fundamental theorem of welfare economics and have to see how such a simple rule can be implemented in a competitive market economy. In a competitive market economy firms choose inputs to maximize profits,  $Z_t - D(E_t)Z_t^\varepsilon Z_0^{1-\varepsilon} - w_t L_t - (r_{t+1} + \delta)K_t - (p_t + \tau_t)F_t - q_t R_t$ , taking the wage rate  $w_t$ , the market interest rate  $r_{t+1}$ , the market price for fossil fuel  $p_t$ , the specific tax  $\tau_t$  on carbon emissions, the market price for renewable energy  $q_t$ , and the atmospheric carbon concentration as given. Capital accumulation follows from  $K_{t+1} = (1 - \delta)K_t + I_t$ , where  $I_t$  is the investment rate at time  $t$ . Fossil fuel owners also operate under perfect competition and maximize profits,  $\sum_{t=0}^{\infty} \Delta_t [p_t F_t - G(S_t)F_t]$  with  $\Delta_t \equiv \prod_{s=0}^t (1 + r_{s+1})^{-1}$ ,  $t \geq 0$ , subject to the depletion equation (7), taking the market price of fossil fuel  $p_t$  as given and internalizing the

---

<sup>12</sup> This states that the return on extracting an extra unit of fossil, selling it and getting a return on it,  $r_{t+1}s_t$  minus the increase in future extraction cost  $-G'(S_{t+1})F_{t+1}$ , must equal the expected capital gain from keeping an extra unit of fossil fuel in the earth  $s_{t+1} - s_t$ .

<sup>13</sup> This can be verified by comparing conditions of Result 3 with those of Result 5 below. There are other ways of implementing the social optimum (e.g., an efficient market for tradable carbon permits).

effect of depletion on future extraction costs. Producers of renewable energy also operate under perfect competition and maximize their profits,  $\sum_{t=0}^{\infty} \Delta_t [q_t R_t - b_t R_t]$ , taking the market price of renewable energy  $q_t$  as given. Households maximize utility (5) subject to their budget constraint  $A_{t+1}^H = (1 + r_{t+1})A_t^H + w_t L_t + \tau_t F_t - C_t$ , where  $A_t^H$  denotes household assets and  $\tau_t F_t$  lump-sum carbon tax rebates at time  $t$ . Since Ricardian debt neutrality holds, there is no loss of generality in assuming that the government balances its books. Asset market and final good market equilibrium require  $A_t^H = K_t$  and  $GDP_t = Z_t - D(E_t)Z_t^\varepsilon Z_0^{1-\varepsilon} = C_t + I_t + G(S_t)F_t + b_t R_t$ . These optimality conditions give rise to the following result.

**Result 5:** *To implement a simple rule for the carbon tax (2) or (3) in a market economy, one needs to solve the dynamics of capital, in-situ fossil fuel reserves and carbon stocks (6)-(9), the Euler equation (10), the energy producers' optimality conditions*

$$(11a') \quad p_t \leq G(S_t) + s_t, \quad F_t \geq 0, \quad \text{c.s.}, \quad \forall t \geq 1,$$

$$(11b') \quad q_t \leq b_t, \quad R_t \geq 0, \quad \text{c.s.}, \quad \forall t \geq 1,$$

*the dynamics of the scarcity rent  $s_t$  (12) and the final good firms' optimality conditions*

$$(14) \quad \begin{aligned} \Gamma_t Z_L &= w_t, \quad \Gamma_t Z_K = r_t + \delta, \\ \Gamma_t Z_{F_t+R_t} &\leq p_t + \tau_t, \quad F_t \geq 0, \text{ c.s.}, \quad \Gamma_t Z_{F_t+R_t} \leq q_t, \quad R_t \geq 0, \text{ c.s.}, \quad \forall t \geq 1, \end{aligned}$$

*and the carbon tax rule (2) or (3) with the initial stocks for  $K_0$ ,  $S_0$ ,  $E_{1,0}$ , and  $E_{2,0}$  given.  $\square$*

This business as usual (BAU) scenario corresponds to zero carbon taxes,  $\tau_t = 0$ , and can be evaluated with the aid of Result 5 too. Numerically solving the market economy according to this Result 5 is more complicated than solving the command economy according to Result 3, because the decentralized economy poses a market equilibrium problem rather than an optimization problem and standard numerical optimization solvers such as CONOPT for GAMS cannot be readily applied.

#### 4. Baseline results: carbon taxes and inequality aversion in a growing economy

Here we present the first-best optimal carbon tax and the BAU outcomes from our IAM and compare them with the outcome under our simple rule for the carbon tax (2). Our IAM adopts the damage specification and more realistic carbon cycle with lags and separate temperature dynamics from Nordhaus (2014), but recalibrated to an annual time grid which is crucial to

describe the dynamics of the energy transitions accurately.<sup>14</sup> The functional forms and calibration are discussed in appendix B. In the baseline we let the global population increase from 7 billion to 11 billion over the next century and let total factor productivity grow at a rate of 2% per year. Our baseline parameters imply relatively low cost of fossil fuel extraction and high cost for renewable energy. Despite this bias toward initial fossil fuel use, both the first-best optimal carbon tax and the simple rule render the majority of fossil fuel reserves obsolete. Our baseline corresponds to the widely used intergenerational inequality aversion of two ( $\Phi = 2$ ), but to emphasize the importance of the IIA for the optimal carbon tax, we also report results for logarithmic utility ( $\Phi = 1$ ). Table 1 illustrates the six scenarios and the coding that is used to distinguish them in the simulation figures.

**Table 1**

**Policy Scenarios for the setting of the global carbon tax**

	<i>First best</i>	<i>BAU</i>	<i>Simple rule (2)</i>
<i>Baseline</i> ( $\Phi = 2$ )	—————	— — — — —	.....
<i>Logarithmic utility</i> ( $\Phi = 1$ )	—————	— — — — —	.....

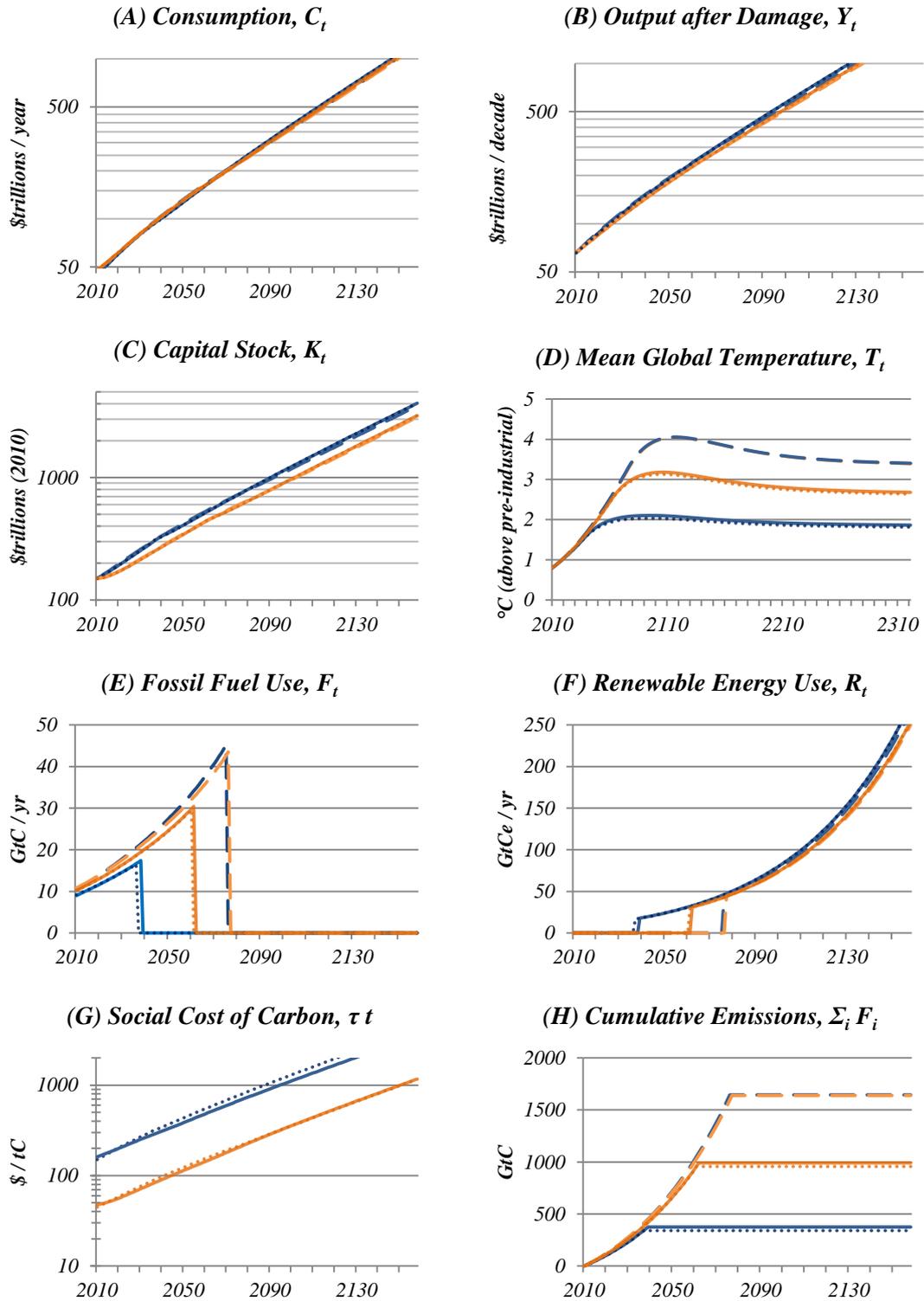
#### 4.1. Role of intergenerational inequality aversion

We start with figure 1 which shows the *first best* and how it depends on the degree of IIA; the solid line corresponds to  $\Phi = 2$  and the dark line to  $\Phi = 1$ .<sup>15</sup> Consumption and net output rapidly approach their steady growth path of 2 percent per year and are almost the same across the first-best, BAU and rule scenarios, but the aggregate capital stock is significantly lower under the first best and the rule than under BAU. Given the persistent trend in productivity growth, higher IIA lowers concern for future generations and makes climate policy less ambitious. In the baseline (with  $\Phi = 2$ ), table 2 indicates that fossil fuel is phased out in 2060 and temperature increases to slightly above 3°C. Under logarithmic utility, IIA is only half of that in the baseline so current generations can make more sacrifices by producing more costly as fossil fuels are phased out earlier, by 2038. As a result, temperature peaks slightly above 2°C. These sacrifices also mean that only 375 GtC are burnt under a unit IIA which is a third of the 990 GtC of the baseline with double the IIA. The choice of IIA has importance ramifications for climate policy in the presence of long-run growth. In the baseline higher emission levels imply higher marginal

<sup>14</sup> Nordhaus (2014) now uses a semi-decadal instead of decadal grid in his DICE-2013R model. Cai et al. (2012) show that a decadal grid introduces significant biases in the calculation of the SCC. An annual grid also helps to better to pinpoint the timing of the energy transitions and the level of stranded assets.

<sup>15</sup> In our numerical simulations and optimizations time runs from 2010 till 2610 and is measured in years,  $t=1,2,\dots, 600$ . The final time period is  $t=600$  or 2610, but we highlight the transitional dynamics in the first 150 years of the simulation.

Figure 1: Baseline simulations for baseline ( $\Phi=2$ ) and logarithmic utility ( $\Phi=1$ )



Note: Social optima: baseline,  $\Phi=2$  ( — ); logarithmic utility,  $\Phi=1$  ( — )  
 BAU: baseline,  $\Phi=2$  ( - - - ); logarithmic utility,  $\Phi=1$  ( - - - )  
 Simple rule for carbon tax: baseline,  $\Phi=2$  ( ..... ); logarithmic utility,  $\Phi=1$  ( ..... )

damages along the growth trajectory. The higher concern for intergenerational inequality, however, implies that the interest rate is much higher (by about 2 percentage points for most periods) and that these damages are discounted more heavily, thus leading to a lower SCC. In the baseline scenario the carbon tax starts at \$48/tC which is about a third of the \$162/tC under logarithmic utility. Both carbon taxes then rise at about 2 percent per year in line with the trend growth in productivity and living standards. Without a carbon tax the market price of fossil fuel is significantly lower and firms, not forced to internalize the deleterious effects of fossil fuel, use more of the cheaper input. Cumulative extraction and global temperature are higher under BAU as fossil fuel is used more and longer. Regardless of the degree of IIA, BAU leads to a maximal warming of 4°C and in total as much as 1,640 GtC are burnt which matches closely most recent baseline projections of the IPCC (2014).

**Table 2**

**Transition times and carbon budget**

	<i>Fossil fuel Only</i>	<i>Renewable Only</i>	<i>Carbon used</i>	<i>max. T</i>	<i>Welfare loss</i>
<i>Baseline (<math>\Phi=2</math>)</i>					
<i>First best</i>	2010-2061	2062 –	990 GtC	3.2 °C	0%
<i>BAU</i>	2010-2075	2076 –	1640 GtC	4.0 °C	- 7%
<i>Simple rule</i>	2010-2060	2061 –	960 GtC	3.1 °C	- 0.02%
<i>Log utility (<math>\Phi = 1</math>)</i>					
<i>First best</i>	2010-2038	2039 –	375 GtC	2.1 °C	0%
<i>BAU</i>	2010-2075	2076 –	1645 GtC	4.1 °C	- 158%
<i>Simple rule</i>	2010-2036	2037 –	340 GtC	2.0 °C	- 0.25%

*Note: Welfare losses are relative to the first best (which has welfare loss of 0%) and are reported as percentage losses of initial GDP.*

### 4.3. Performance of the simple rule for the global carbon tax

The simple rule (2'') put forward by Golosov et al. (2014) implies that the SCC rises in line with GDP and offers a useful benchmark. This rule give the first-best SCC exactly only under very restrictive assumptions (logarithmic utility, Cobb-Douglas production, decadal time steps, full depreciation each period, exponential damages, and zero capital intensity of fossil fuel extraction). Furthermore, this rule does not allow for the effects of trend growth and IIA bigger than unity on the SCC. Our simple rule (2) is more general, since it allows for sustained productivity and population growth and different levels of IIA as well as a temperature lag. The dotted light (for  $\Phi = 2$ ) and dark (for  $\Phi = 1$ ) lines in figures 2 and 3 represent the BAU outcome under our simple rule (2). We find that our rule predicts the level and growth of the first-best

carbon tax accurately and mimics the social optimum very closely. In the baseline the carbon tax according to our rule starts out \$3 too low at 45\$/tC, discourages fossil fuel use slightly too much, and induces the transition to renewable energy one year too early.<sup>19</sup> The welfare cost associated with this deviation is negligible, amounting to 0.02 percent of initial GDP (cf. table 2). For logarithmic utility our simple rule performs slightly worse, but it still reproduces the social optimum very well and leads to a welfare loss of only 0.25 percent of initial GDP thereby avoiding the huge welfare losses of 158% under BAU. These welfare losses under BAU are so high, because the lower IIA implies the need for a more ambitious climate policy and thus doing nothing as in BAU is extra costly. The rule (2) also predicts the amount of fossil fuel to be left abandoned and the optimal time of transition to the carbon-free era quite closely (cf. table 2).

The rule (2'') proposed by Golosov et al. (2014) is based on a simplified decadal two-box carbon cycle with no separate dynamics or lags in global mean temperature. In a calibrated model, it gives an initial SCC of \$75/tC which consequently rises at the same rate of growth as GDP. Our IAM is more realistic, since it allows for general levels of IIA, adopts the five-box carbon and non-linear temperature module of Nordhaus (2014), considers persistent long-run productivity growth, and is recalibrated to a finer annual time grid. These revisions increase the initial SCC significantly to 162\$/tC for IIA = 1 and curb it by half to \$48/tC for IIA = 2.

These differences in valuing intergenerational inequality also show up in the use of fossil fuel and the timing of transition to renewable energy. The relative wealth of future generations implies that current generations use more (less) fossil fuel and induce higher (lower) temperature rises, especially if IIA is high. Inequality aversion and economic growth are, therefore, just as important as the rate of time impatience (the pure rate of time discounting) in determining the social cost of carbon.

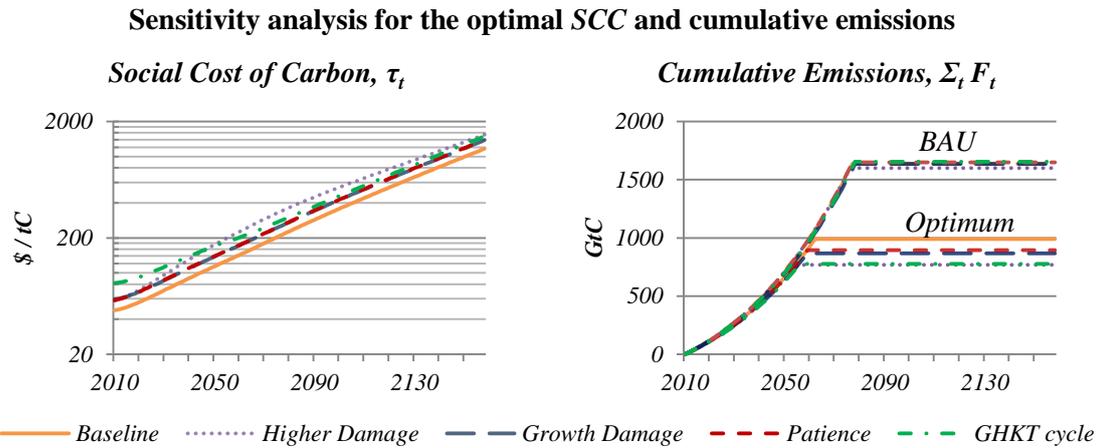
## **5. More ambitious climate policy: convex damages, growth effects and more patience**

In this section and in section 6 we explore the robustness of the performance of the rule (2) and, where applicable, its more general version (3). Figure 2 shows the sensitivity of the first-best, BAU and rule scenarios to more convex damages, adverse productivity growth effects of climate change, and risk to consumption growth. These effects make climate policy more ambitious by increasing the SCC and locking more fossil fuel in the ground. Table 3 gives the transition times and carbon budgets for the social optimum and the outcome under the rule for the carbon tax.

---

<sup>19</sup> The reason is that with  $IIA > 1$  the initial interest rate is lower than the long-run rate and thus the true optimal SCC is over-estimated. With  $IIA < 1$ , we get the opposite result.

Figure 2



*Note: Higher instantaneous and growth damages, prudence and more responsive climate system increase the SCC, curb fossil fuel use and lowering the carbon budget.*

The first sensitivity test increases the convexity of damages at high temperatures. Weitzman (2010) and Dietz and Stern (2015) argue that damages rise more rapidly at higher levels of temperature than suggested by Nordhaus (2014), but empirical studies on the costs of global warming at higher temperatures are not available due to lack of evidence (atmospheric carbon concentrations have not exceeded 400ppmv for 800,000 years). Greater convexity increases the SCC and optimal policy responds by limiting emissions, the carbon budget and global warming. Since the damage function does not increase much on the relevant temperature range, the outcome is not very different from the baseline.

Hence, even though the simple rule is based on the fairly flat marginal damages of Nordhaus (2014) and Golosov et al. (2014), it still mimics the social optimum quite closely with a welfare loss of 2 percent of initial GDP. Climate damages are commonly modelled as flow damages to the level of output. Recent empirical studies, however, found that climate change also has a negative impact on the growth rate of productivity, especially in developing countries (Dell et al., 2012). In the second sensitivity run, we replace the climate damages proposed by Nordhaus (2008) with these persistent damages to productivity.<sup>20</sup> This leads to less fossil fuel use, less short-run growth and an earlier transition to renewable energy. The rule (3) adjusts to the different damage specification with a welfare loss of 4 percent of initial output.

<sup>20</sup> We keep the same damage function but allow for persistent damages to productivity (see appendix B for details). We take these negative effects to be deviations from trend productivity growth and allow for mean reversion. In the sensitivity runs damages to productivity are assumed persistent with an AR(1) coefficient of 0.75.

**Table 3****Robustness of the carbon tax rule (2) or (3) to more ambitious policy**

<i>Scenario</i>	<i>Carbon used (First best)</i>	<i>max. T (First best)</i>	<i>Switch Time (First best)</i>	<i>Carbon used (simple rule)</i>	<i>max. T (simple rule)</i>	<i>Welfare loss (simple rule)</i>	<i>Welfare loss (BAU)</i>
<i>Baseline</i>	990 GtC	3.2 °C	2062	960 GtC	3.1 °C	- .02%	- 7%
<i>Higher damage</i>	770 GtC	2.8 °C	2054	960 GtC	3.1 °C	- 2.23%	- 35%
<i>Growth damage</i>	868 GtC	3.0 °C	2077	581 GtC	2.5 °C	- 4%	- 10%
<i>More patient</i>	895 GtC	3.0 °C	2058	865 GtC	3.0 °C	- .03%	- 12%
<i>GHKT cycle</i>	775 GtC	N/A	2055	814 GtC	N/A	- .09%	- 15%

*Note: Welfare losses are relative to the first best (which has welfare loss of 0%) and are reported as percentage losses of initial GDP.*

Next consider the effects of a lower discount rate from 1.0 to 0.6 percent per annum.<sup>21</sup> This gives greater weight to future marginal climate damages and thus increases the optimal carbon tax. The carbon budget reduces by more than 10 percent and global warming slightly falls to 3.0°C. The simple rule for the carbon tax aptly adjusts to the lower discount rate and maintains a welfare loss of 0.03 percent compared to the cost of inaction of 12% of initial GDP.

Table 3 also shows the effects of replacing the DICE carbon cycle with the simpler carbon cycle used in Golosov et al. (2014) and in the derivations of our simple rules (2) and (3). Somewhat surprisingly, the welfare loss resulting from the simple rule is somewhat bigger than with the DICE carbon cycle and temperature module. The reason is that the interest rate rises along the transition to the balanced growth path while the long-run rate has been applied in the derivation of the rules (2) and (3). Time lags in the DICE carbon cycle (absent in Golosov et al. (2014)) lessen this approximation error and improve the rule's performance. The immediacy of this carbon cycle variant also explains why the carbon tax is higher and more carbon is locked up relative to the baseline scenario.

Web appendix C discusses the effects of less ambitious climate policy resulting from additive damages, population stagnation and lower trend growth and shows that our simple rule still performs well when tested in the decentralized market economy.

<sup>21</sup> In a stochastic model this might arise from potential shocks to consumption growth (Kocherlokata, 1996; Gollier, 2013). Since global annual consumption growth had a standard deviation of 3.6 percent over much of the past century, the rate of time impatience is reduced by  $-0.5 \times \text{IIA} \times \text{CRP} \times \text{variance} = -0.4$  percent per annum with  $\text{IIA} = 2$ ,  $\text{CRP} = 1 + 2 = 3$  and the variance is 0.036 squared. This suggests that to allow for prudence the rate of time impatience should be lowered from 1.0 to 0.6 percent per annum. Stern (2007) discusses further important reasons for lower discounting.

## 6. Extreme road-tests of our simple rule

Here we road-test the rule (2) in a more severe context and show what happens if the speed of convergence of the Ramsey growth block of our IAM is much slower and thus closer to the speed of adjustment of the carbon cycle part of our IAM. This worsens performance of the simple rule relative to first best. In absolute terms, the loss remains, however, small. While the slow dynamics of the carbon cycle are governed by the geophysics of the climate system (about 0.25% per annum for a carbon cycle like the one of Golosov et al. (2014) or 0.04% per annum for the climate system of DICE-2013R),<sup>22</sup> the dynamics of the Ramsey growth block of our IAM (about 2% per annum) are influenced by economic variables. The speed of convergence of the Ramsey growth model is lower if productivity growth is lower ( $g$ ), the share of capital ( $\alpha$ ) is higher, the depreciation rate of capital ( $\delta$ ) is smaller, the coefficient of relative intergenerational inequality aversion ( $\text{IIA} = \Phi$ ) is higher (or  $\text{EIS} = 1/\Phi$  is lower), the rate of time preference ( $\rho$ ) is lower, and the value added share of energy ( $\beta$ ) is lower. We thus ‘load the dice’ against the simple rule by choosing parameters that slow down Ramsey convergence. In all scenarios we set productivity growth to zero. In addition we lower the depreciation and discount rates and increase the capital share. We find that the rule (2) performs well even in these adverse circumstances but less than in the baseline runs of sections 4-6.

Table 4 shows the time it takes for the capital stock to reach 90% of its steady-state level (in efficiency units) and confirms that convergence slows down as the rates of capital depreciation and discounting are lowered and especially if the share of capital is increased. The left panel of figure 3 shows the time paths of the carbon tax in the various scenarios. The right panel of figure 3 illustrates how these slower adjustment speeds translate into slower convergence of our rule for the carbon tax to the first-best carbon tax. As the time for convergences increases, the deviation between our rule for the carbon tax and the first-best carbon tax widens and persists for a longer period. The long-run deviation is due to different carbon cycles used in the derivation of our rule (2) (i.e., two boxes plus a temperature lag) and the simulations (carbon cycle module of DICE-2013R).

For the baseline scenario the welfare loss relative to the correct optimum is only 0.02% compared with 7% of initial GDP under BAU. In terms of welfare, table 5 shows that our rule converges less quickly and performs worse under these extreme conditions than the baseline of section 4 and the sensitivity tests discussed in section 5 (and web appendix C) but still quite well especially when compared with BAU. The highest welfare loss of only 1.8% of initial GDP comes from an increase in the share of capital from 30% to 65%, compared with a loss of 45.6%

---

<sup>22</sup> Web appendix D discusses speeds of convergence in different climate models in detail.

for the business as usual scenario. The welfare loss is the same as for the reference scenario where only productivity growth is set to zero, and again the welfare loss under the rule is much less than that under BAU (1.8% versus 54 % of initial GDP). Lower discounting and lower depreciation slow down convergence and worsen the performance of our rule by rather less. Relative to BAU, the rule yields at least 27% of initial GDP for lower depreciation and at most 265% under a lower discount rate. Our rule thus performs well even in more extreme scenarios.

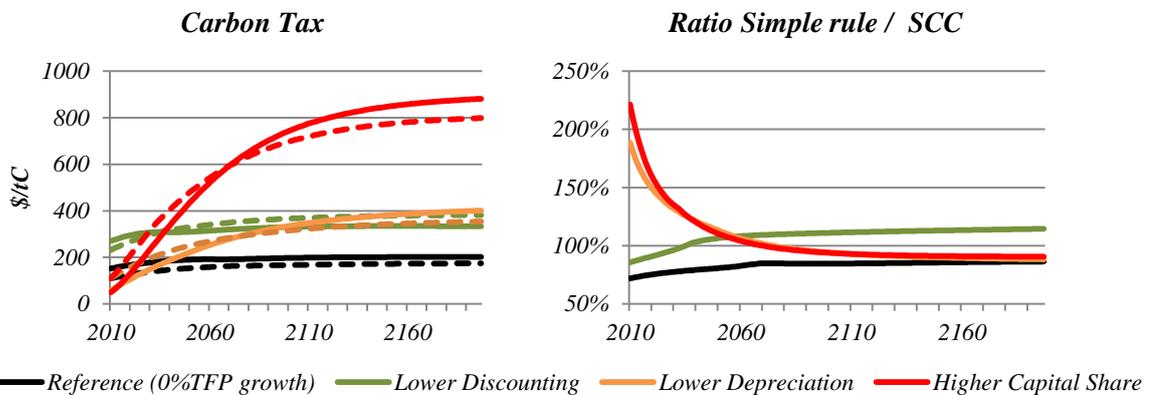
**Table 4**

**Convergence and welfare loss of simple rule in a slow convergence environment**

Scenario	Time to 90%	Welfare loss relative to first best (% of initial GDP)	Welfare gain relative to BAU (% of initial GDP)
Baseline of section 4 (2% TFP growth)	42 years	- 0.02%	7%
Reference (0% TFP growth) —	47 years	- 1.8%	54%
Lower Discounting (0.5%/year) —	65 years	- 0.2%	265%
Lower Depreciation (1%/year) —	123 years	- 0.7%	27%
Higher capital share (65%) —	177 years	- 1.8%	44%

**Figure 3**

**Robustness of simple rule in a slow convergence environment**



Note: Slower economic convergence lowers the SCC (solid) and worsens the performance of the simple rule (dashed).

**7. Conclusion**

The central questions of climate change economics have been what the level and what the time profile of the optimal price of carbon, whether achieved by levying a carbon tax or creating a carbon emission market, should be. We present a simple but robust rule for the optimal carbon tax which generalizes those put forward by Golosov et al. (2014) and van den Bijgaart et al. (2013). Our rule accounts thus not only for geophysical components, such as the dissipation speed of atmospheric carbon and temperature dynamics, but also for climate damages that are non-proportional to gross output and exhibit mean reversion and for socio-economic

characteristics, such as the growth rates of population and living standards and society's aversion to intergenerational inequality. The welfare losses from using our rule are tiny, especially when considering the business as usual alternative. Our rule predicts the first-best path for the carbon tax, the transition times to the carbon-free era and the amount of untapped fossil fuel reserves very well.

Our rule is easy to understand, calculate and implement. Furthermore, in contrast to the usual discretionary optimal time paths that are calculated from DICE and other IAMs, our rule allows for changing circumstances. It is a restricted feedback rule, since it reacts to GDP (net of climate damage) and does not react to levels of the capital stock, the stock of remaining fossil fuel reserves, atmospheric carbon stocks, or temperature. Due to the carbon cycle dynamics being much more sluggish than the Ramsey growth dynamics, implementing the simple rule in the decentralized economy mimics very close the first-best outcome even though our IAM which is used to capture the climate-economy interactions is much more complicated than the toy model used to calculate the simple rule and in contrast to DICE allows for forward-looking dynamics in scarcity rents and the effect of falling fossil fuel reserves on extraction costs. The rule also performs well in DICE as the supplementary material of Golosov et al. (2014) and van den Bijgaart et al. (2014) indicate. The rule is thus robust to model misspecification (highlighted by Stern, 2013) and may have more to command itself than the true optimal discretionary time path of carbon taxes generated by a particular IAM.

Our analysis highlights the importance of intergenerational inequality and economic growth rates for the carbon tax. Lower IIA in the presence of economic growth lead to a higher optimal carbon tax and a quicker phasing in of renewables and more fossil fuel left in the crust of the earth, just as a lower rate of time impatience. The lower concern for intergenerational inequality implies that, since society is more concerned with fighting global warming than with avoiding big differences in consumption of different generations, the carbon tax is borne much more by earlier generations than by later generations in the presence of sustained growth. This effect increases in the rate of economic growth. Our simple rule captures this effect correctly for a large range of intergenerational inequality aversion.

How global warming damages are modeled and calibrated matters for the SCC and climate policy. We rely on the common specification of marginal damages that are proportional to production, but also consider additive production damages and damages to productivity growth. We find that the climate policy is less ambitious, energy use higher, the stock of fossil fuel left in situ lower, global mean temperature higher and the optimal carbon tax lower with additive damages provided that the rate of economic growth is positive. Climate policy is more

ambitious and less fossil fuel is used if damages affect not only the level productivity but also its growth rate. Our extended rule for the carbon tax captures adjusts aptly to these different damage specifications.

Here we have considered the standard case of deterministic and continuous climate damages. To allow for risk and uncertainty, one might consider three extensions. First, to allow for the risk of catastrophic events one can revise the damage coefficients in our rule upwards as in Golosov et al. (2014). If the risk of catastrophic tipping points and regime switches rises with global warming, the rule for the optimal price of carbon should contain an extra term to curb the hazard of catastrophe. Second, the optimal price of carbon will increase in relative risk aversion and prudence, but decrease in intergenerational inequality aversion.<sup>23</sup> Third, if growth expectations are diminishing, the stochastic process for the growth rate of the economy displays mean reversion or the future rate of return on capital is governed by a gamma distribution, the precautionary effect is boosted and increases over time due to the persistency of shocks and thus the rate to discount climate damages falls with longer horizons (e.g., Gollier, 2013). The great benefit of our simple rule is that it can accommodate probability distributions of climate parameters and catastrophic shocks, but also modifications of the Euler equations due to ambiguity aversion if the distribution is unknown. Such aspects of stochastic control and decision theory as well as hyperbolic discounting and learning are left for future research.

## References

- Archer, David 2005. The fate of fossil fuel CO<sub>2</sub> in geologic time. *Journal of Geophysical Research* 110: C09S05.
- Archer, David, Michael Eby, Victor Brovkin, Andy Ridgwell, Long Cao, Uwe Mikolajewicz, Ken Caldeira, Katsumi Matsumoto, Guy Munhoven, Alvaro Montenegro, and Kathy Tokos. 2009. Atmospheric lifetime of fossil fuel carbon dioxide. *Annual Review of Earth and Planetary Sciences* 37: 117-134.
- Bansal, Ravi, and Marcelo Ochoa. 2011. Welfare costs of long term temperature shifts. NBER Working Paper 17574, National Bureau of Economic Research, Cambridge, MA.
- Bansal, Ravi, and Marcelo Ochoa. 2012. Temperature, aggregate risk, and expected returns. NBER Working Paper 17575, National Bureau of Economic Research, Cambridge, MA.
- Barro, Robert J. 2015. Convergence and modernisation. *Economic Journal* s125: 585, 911-942.
- Bijgaart, Inge van den, Reyer Gerlagh, Luuk Korsten, and Matti Liski. 2013. A simple formula for the social cost of carbon. Working Paper No. 83.2013, FEEM, Milan.

---

<sup>23</sup> One can separate risk aversion and intergeneration inequality aversion using Epstein-Zin (1989) preferences as discussed in Gollier (2013), Jensen and Traeger (2014) or Cai et al. (2015). This requires modifying our simple rule for the optimal price of carbon by lowering the trend growth rate with a term that is proportional to relative risk aversion and the variance of consumption.

- Cai, Yongyang, Kenneth L. Judd, and Thomas S. Lontzek. 2012. Open science is necessary. *Nature Climate Change* 2: 299.
- Cai, Yongyang, Kenneth L. Judd, Timothy M. Lenton, Thomas S. Lontzek, and Daiju Narita. 2015. Risk to ecosystem services could significantly affect the cost-benefit assessments of climate change policies, *The Proceedings of the National Academy of Sciences*. April, 14, 112: 15), 4606-4611.
- Dell, Melissa, Benjamin F. Jones, and Benjamin A. Olken. 2012. Temperature shocks and economic growth: evidence from the last half century. *American Economic Journal: Macroeconomics* 4: 66-95.
- Dietz, Simon and Nicholas H. Stern. 2015. Endogenous growth, convexity of damages and climate risk: how Nordhaus' framework supports deep cuts in carbon emissions. *Economic Journal* 125: 583, 574-602.
- Epstein, Larry G., and Stanley E. Zin. 1989. Substitution, risk aversion, and the temporal behavior of consumption and asset returns. *Econometrica* 57: 4, 937-969.
- Gerlagh, Reyer, and Matti Liski. 2012. Carbon prices for the next thousand years. Working Paper 3855, CESifo, Munich.
- Gollier, Christian. 2013 *Pricing the Planet's Future, The Economics of Discounting in an Uncertain World*. Princeton, New Jersey: Princeton University Press.
- Golosov, Michael, John Hassler, Per Krusell, and Aleh Tsyvinski (2014). Optimal taxes on fossil fuel in general equilibrium. *Econometrica* 82: 1, 41-88.
- IPCC. 2014. *Climate Change 2014, AR5 Synthesis Report*. Accessed 14 Nov 2014 at [http://www.ipcc.ch/pdf/assessment-report/ar5/syr/SYR\\_AR5\\_SPM.pdf](http://www.ipcc.ch/pdf/assessment-report/ar5/syr/SYR_AR5_SPM.pdf).
- Jensen, Sverre, and Christian P. Traeger. 2014. Optimal climate change mitigation under long-term growth uncertainty: stochastic integrated assessment and analytic findings, *European Economic Review*. 69: C, 104-125.
- Kocherlakota, Narayana. 1996. Implications of efficient risk sharing without commitment. *Review of Economic Studies* 63: 4, 595-609.
- Lutz, Wolfgang, William P. Butz, and Samir KC. 2014. *World Population and Global Human Capital in the Twenty-First Century*. New York: Oxford University Press.
- Majda, Andrew, Ilya Timofeyev, and Eric van den Eijnden. 2001. A mathematical framework for stochastic climate models. *Communications on Pure and Applied Mathematics* 54: 891-794.
- Moore, Frances C., and Delavane B. Diaz. 2015. Temperature impacts on economic growth warrant stringent mitigation policy. *Nature Climate Change* 5: 127-131.
- Nordhaus, W. (1991). To slow or not to slow: the economics of the greenhouse effect. *Economic Journal* 101: 407, 920-37.
- Nordhaus, William. 2008. *A Question of Balance. Weighing the Options on Global Warming Policies*. New Haven, Connecticut: Yale University Press.
- Nordhaus, William. 2014. Estimates of the social cost of carbon: concepts and results from the DICE-2013R model and alternative approaches. *Journal of the Association of Environmental and Resource Economists* 1: 273-312.
- Pindyck, Robert S. 2013. Climate change policy: what do the models tell us? *Journal of Economic Literature* 51: 860-872.
- Rezai, Armon, and Frederick van der Ploeg. 2015. Robustness of a simple rule for the social cost of carbon. *Economics Letters* 132: 48-55.

- Rezai, Armon, Frederick van der Ploeg, and Cees Withagen. 2014. Economic growth and the social cost of carbon: additive versus multiplicative damages. Research Paper 93, OxCarre, University of Oxford.
- Stern, Nicholas H. 2007. *The Economics of Climate Change: The Stern Review*. Cambridge, U.K.: Cambridge University Press.
- Stern, Nicholas H. 2013. The structure of economic modeling of the potential impacts of climate change: grafting gross underestimation of risk onto already narrow science models. *Journal of Economic Literature* 51: 3, 838-859.
- Tol, Richard. 2002a. Estimates of the damage costs of climate change, part I: benchmark estimates. *Environmental and Resource Economics* 21: 47-73.
- Tol, Richard. 2002b. Estimates of the damage costs of climate change, part II: dynamic estimates. *Environmental and Resource Economics* 21: 135-160.
- Tol, Richard. 2015. Economic impacts of climate change. Working Paper Series No. 75-2015, Department of Economics, University of Sussex.
- Weitzman, Martin L. 2009. Additive damages, fat tailed climate dynamics and uncertain discounting. *Economics: The Open-Access, Open-Assessment E-Journal* 3: 2009-39.
- Weitzman, Martin L. 2010. Risk-adjusted gamma discounting. *Journal of Environmental Economics and Management* 60: 1-13.

## Appendix A: Derivations

### Derivation of Result 1:

From (1) we get the following expression for the SCC:

$$(A1) \quad \frac{SCC_t}{\chi Y_t^\varepsilon Y_0^{1-\varepsilon}} = \left( \frac{\varphi_L}{1 - \frac{(1+g)^\varepsilon}{1+r}} + \frac{\varphi_0(1-\varphi_L)}{1 - \frac{(1+g)^\varepsilon}{1+r}(1-\varphi)} \right) \left( \frac{(1+g)^\varepsilon}{(1+g)^\varepsilon + \varphi_T [1+r - (1+g)^\varepsilon]} \right).$$

Per-capita consumption growth  $(1+g)/(1+n) = [(1+r)/\beta]^{1/\Phi}$  follows from the Euler equation, so the interest rate is  $1+r = [(1+g)/(1+n)]^\Phi / \beta$  and we get

$$(A2) \quad SCC_t = \left( \frac{\varphi_L}{1 - \beta(1+n)^\varepsilon \left(\frac{1+g}{1+n}\right)^{\varepsilon-\Phi}} + \frac{\varphi_0(1-\varphi_L)}{1 - \beta(1+n)^\varepsilon \left(\frac{1+g}{1+n}\right)^{\varepsilon-\Phi} (1-\varphi)} \right) \times \left( 1 + \varphi_T \left[ \left(\frac{1+g}{1+n}\right)^{\Phi-\varepsilon} \frac{(1+n)^{-\varepsilon}}{\beta} - 1 \right] \right)^{-1} \chi Y_t^\varepsilon Y_0^{1-\varepsilon},$$

For small  $\varphi$ ,  $r$ ,  $g$  and  $\rho \equiv \beta - 1$  (ignoring second-order terms), the Euler equation becomes  $g - n = (r - \rho) / \Phi$  and this expression for the SCC boils down to (2).  $\square$

**Derivation of Result 2:**

$E_t$  affects all future damages at time  $s \geq t$ , so we need  $\omega_t = \sum_{s=t}^{\infty} (1+r)^{-(s-t)} \frac{\partial D_s}{\partial E_t} =$

$-Y_t \sum_{s=t}^{\infty} \left( \frac{1+g}{1+r} \right)^{s-t} \frac{\partial A_s}{\partial E_t} \xi_s$ . We get  $A_s = \left( \prod_{t=0}^s e^{-\chi \varphi_\chi^{s-t} (E_t - E_0)} \left[ (1+g)^t \bar{A}_0 \right]^{(1-\varphi_\chi) \varphi_\chi^{s-t}} \right) \bar{A}_0 \varphi_\chi^{s-t}$  if

$\bar{A}_t = (1+g)^t \bar{A}_0$ , so that  $-\frac{\partial A_s}{\partial E_t} = \chi \varphi_\chi^{s-t} A_s$ ,  $s \geq t$ . Marginal damages of 1 tC emitted at time  $t$  is

$\omega_t = \chi Y_t \sum_{s=t}^{\infty} \left( \frac{1+g}{1+r} \varphi_\chi \right)^{s-t} \xi_s$ , so that

$$(A3) \quad \frac{\omega_t}{\chi Y_t} = \varphi_L \left[ \frac{1+r}{1+r-(1+g)\varphi_\chi} - \frac{1+r}{1+r-(1+g)\varphi_\chi(1-1/\varphi_T)} \left( 1 - \frac{1}{\varphi_T} \right)^t \right] \\ + \left( \frac{\varphi_0(1-\varphi_L)}{1-\varphi_T\varphi} \right) \left[ \frac{1+r}{1+r-(1+g)\varphi_\chi(1-\varphi)} (1-\varphi)^t - \frac{1+r}{1+r-(1+g)\varphi_\chi(1-1/\varphi_T)} \left( 1 - \frac{1}{\varphi_T} \right)^t \right].$$

Substituting this into  $SSC_t \equiv \sum_{s=0}^{\infty} (1+r)^{-(s-t)} \omega_s$ , we get

$$(A4) \quad \frac{SSC_t}{\chi Y_t} = \frac{\varphi_L}{1 - \left( \frac{1+g}{1+r} \right) \varphi_\chi} \frac{1}{1 - \left( \frac{1+g}{1+r} \right)} - \frac{\varphi_L}{1 - \left( \frac{1+g}{1+r} \right) \varphi_\chi \left( 1 - \frac{1}{\varphi_T} \right)} \frac{1}{1 - \left( \frac{1+g}{1+r} \right) \left( 1 - \frac{1}{\varphi_T} \right)} \\ + \frac{\varphi_0(1-\varphi_L)}{1-\varphi_T\varphi} \left[ \frac{1}{1 - \left( \frac{1+g}{1+r} \right) \varphi_\chi (1-\varphi)} \frac{1}{1 - \left( \frac{1+g}{1+r} \right) (1-\varphi)} - \frac{1}{1 - \left( \frac{1+g}{1+r} \right) \varphi_\chi \left( 1 - \frac{1}{\varphi_T} \right)} \frac{1}{1 - \left( \frac{1+g}{1+r} \right) \left( 1 - \frac{1}{\varphi_T} \right)} \right].$$

Upon substitution of the Euler equation  $\frac{1+g}{1+r} = (1+g)^{1-\Phi} / (1+\rho) \cong 1 - \rho - (\Phi-1)g$ , and ignoring other second-order terms, we get (3).  $\square$

**Derivation of Result 3:**

The Lagrangian for the problem of choosing  $\{C_t, F_t, R_t, \forall t \geq 1\}$  to maximize social welfare (5) subject to equations (6)-(9) is:

$$(A5) \quad L \equiv \sum_{t=0}^{\infty} (1+\rho)^{-t} \left[ L_t \frac{(C_t/L_t)^{1-\Phi}}{1-\Phi} - \mu_t (S_{t+1} - S_t + F_t) \right] \\ + \sum_{t=0}^{\infty} (1+\rho)^{-t} \left[ v_{1,t} (E_{1,t} - E_{1,t-1} - \varphi_L F_t) + v_{2,t} \{ E_{2,t} - (1-\varphi) E_{2,t-1} - \varphi_0 (1-\varphi_L) F_t \} \right]$$

$$\begin{aligned}
& + \sum_{t=0}^{\infty} (1+\rho)^{-t} \left[ v_t \left( E_t - \frac{1}{\varphi_T} (E_{1,t} + E_{2,t}) - \left( 1 - \frac{1}{\varphi_T} \right) E_{t-1} \right) \right] \\
& - \sum_{t=0}^{\infty} (1+\rho)^{-t} \lambda_t \left[ K_{t+1} - (1-\delta)K_t - Z(K_t, L_t, F_t + R_t) + D(E_t)Z(K_t, L_t, F_t + R_t)^\varepsilon Z_0^{1-\varepsilon} + G(S_t)F_t + b_t R_t + C_t \right],
\end{aligned}$$

where  $\mu_t$  denotes the shadow value of in-situ fossil fuel,  $v_{1t}$ ,  $v_{2t}$  and  $v_t$  the shadow disvalues of the permanent, transient and total delayed stocks of atmospheric carbon, and  $\lambda_t$  the shadow value of manmade capital. Consumption, fossil fuel use and renewable use in period zero,  $C_0$ ,  $F_0$  and  $R_0$ , the capital stock at the start of period zero,  $K_0$ , the stock of fossil fuel reserves at the start of period zero,  $S_0$ , and the carbon stocks at the end of period zero,  $E_{10}$ ,  $E_{20}$ , and  $E_0$ , are given, and thus not part of the optimization program. It follows that gross output,  $Z_0$ , fossil fuel extraction costs,  $G(S_0)$ , and GDP in period zero are given too. We use the same timing conventions as Golosov et al. (2014). Necessary optimality conditions for  $\{C_t, F_t, R_t, \forall t \geq 1\}$  are

$$(A.6) \quad U'(C_t / L_t) = (C_t / L_t)^{-\Phi} = \lambda_t, \quad \forall t \geq 1,$$

$$(A.7a) \quad \Gamma_t Z_{F_t+R_t}(K_t, L_t, F_t + R_t) \leq G(S_t) + [\mu_t + \varphi_L v_{1t} + \varphi_0(1-\varphi_L)v_{2t}] / \lambda_t, \\ F_t \geq 0, \quad \text{c.s.}, \quad \forall t \geq 1,$$

$$(A.7b) \quad \Gamma_t Z_{F_t+R_t}(K_t, L_t, F_t + R_t) \leq b_t, \quad R_t \geq 0, \quad \text{c.s.}, \quad \forall t \geq 1.$$

The co-state equations for this optimization problem are:

$$(A.8) \quad (1+\rho)\lambda_t = \left\{ 1 - \delta + \Gamma_{t+1} Z_{K_{t+1}}(K_{t+1}, L_{t+1}, F_{t+1} + R_{t+1}) \right\} \lambda_{t+1}, \quad \forall t \geq 1,$$

$$(A.9) \quad \mu_{t+1} = (1+\rho) \mu_t + G'(S_{t+1})F_{t+1} \lambda_{t+1}, \quad \forall t \geq 1,$$

$$(A.10) \quad (1+\rho)v_{1t} = v_{1t+1} + \frac{1}{\varphi_T} v_{t+1}, \quad \forall t \geq 1,$$

$$(A.11) \quad (1+\rho)v_{2t} = (1-\varphi)v_{2t+1} + \frac{1}{\varphi_T} v_{t+1}, \quad \forall t \geq 1,$$

$$(A.12) \quad (1+\rho)v_t = \left( 1 - \frac{1}{\varphi_T} \right) v_{t+1} + D'(E_t)Z_t(K_t, L_t, F_t + R_t)^\varepsilon Z_0^{1-\varepsilon} \lambda_t \quad \forall t \geq 1.$$

Equations (A.6) and (A.8) give  $U'(C_t / L_t) = (1+r_{t+1})(1+\rho)^{-1} U'(C_{t+1} / L_{t+1})$  with

$$r_{t+1} \equiv \Gamma_{t+1} Z_{K_{t+1}} - \delta \quad \text{or} \quad \frac{C_{t+1} / L_{t+1}}{C_t / L_t} = [(1+r_{t+1}) / (1+\rho)]^{1/\Phi} \quad \text{which is (10).}$$

The Kuhn-Tucker conditions (A.7a) and (A.7b) give (11a) and (11b) after defining  $s_t \equiv \mu_t / \lambda_t$  and  $\theta_t \equiv [\varphi_L v_{1t} + \varphi_0(1-\varphi_L)v_{2t}] / \lambda_t$  in final good units.

Equation (A.9) gives  $[s_{t+1} - G'(S_{t+1})F_{t+1}] \frac{\lambda_{t+1}}{\lambda_t} = (1 + \rho)s_t, \forall t \geq 1$ . Upon substitution of (A.8), we

get  $(1 + r_{t+1})s_t = s_{t+1} - G'(S_{t+1})F_{t+1}, \forall t \geq 1$ . Solving this difference equation forwards, we get

$s_t = -\sum_{\zeta=1}^{\infty} \Delta_{t+\zeta} G'(S_{t+\zeta})F_{t+\zeta}$  or (12). Finally, dividing (B.5)-(B.7) by  $\lambda_t$  and using (B.3) yields

$$\theta_{1t+1} = (1 + r_{t+1}) \left( \theta_{1t} - \frac{1}{\varphi_T} \theta_{Tt} \right), (1 - \varphi) \theta_{2t+1} = (1 + r_{t+1}) \left[ \theta_{2t} - \frac{1}{\varphi_T} \theta_{Tt} \right] \text{ and}$$

$$\left( 1 - \frac{1}{\varphi_T} \right) \theta_{Tt+1} = (1 + r_{t+1}) \left[ \theta_{Tt} - D'(E_t) Z_t^\varepsilon Z_0^{1-\varepsilon} \right]. \text{ Solving the first two difference equations}$$

forwards and using  $\Delta_{t+\zeta} \equiv \prod_{\zeta'=1}^{\zeta} (1 + r_{t+\zeta'})^{-1}, \zeta \geq 1$  and  $\Delta_t = 1$ , we get

$$\theta_t \equiv \varphi_L \theta_{1t} + \varphi_0 (1 - \varphi_L) \theta_{2t} = \sum_{\zeta=0}^{\infty} \left[ \left\{ \varphi_L + \varphi_0 (1 - \varphi_L) (1 - \varphi)^\zeta \right\} \Delta_{t+\zeta} \theta_{Tt+\zeta} \right] / \varphi_T. \text{ The third difference}$$

$$\text{equation gives } \theta_{Tt} = \sum_{\zeta=0}^{\infty} \left( 1 - \frac{1}{\varphi_T} \right)^\zeta \Delta_{t+\zeta} D'(E_{t+\zeta}) Z_{t+\zeta}^\varepsilon Z_0^{1-\varepsilon}. \quad \square$$

#### Derivation of Result 4:

Under assumption (i) that the economy is on its balanced growth path, (10) is replaced by

$$1 + g - n = \left[ (1 + r) / (1 + \rho) \right]^{1/\Phi}, \text{ which gives } 1 + r = (1 + g - n)^\Phi (1 + \rho) \text{ and the compound}$$

discount factors in Result 3 become  $\Delta_{t+\zeta} = (1 + r)^{-(\zeta-1)} = (1 + g - n)^{\Phi(\zeta-1)} (1 + \rho)^{-(\zeta-1)}$ . Together

with assumptions (ii) that marginal damages are proportional to GDP,  $\theta_{Tt}$  in expression (13) reduces to

$$\theta_{Tt} = \frac{1}{1 - \frac{(1+g)^\varepsilon}{(1+r)} \left( 1 - \frac{1}{\varphi_T} \right)} \chi GDP_t^\varepsilon GDP_0^{1-\varepsilon} \cong \frac{1}{\rho - \varepsilon n + (\Phi - \varepsilon)(g - n) + 1/\varphi_T} \chi GDP_t^\varepsilon GDP_0^{1-\varepsilon}$$

and the SCC becomes

$$\begin{aligned} \theta_t &= \left( \frac{\varphi_L}{1 - \frac{(1+g)^\varepsilon}{(1+r)}} + \frac{\varphi_0(1-\varphi_L)}{1 - (1-\varphi) \frac{(1+g)^\varepsilon}{(1+r)}} \right) \frac{1}{\varphi_T} \frac{1}{1 - \frac{(1+g)^\varepsilon}{(1+r)} \left( 1 - \frac{1}{\varphi_T} \right)} \chi GDP_t^\varepsilon GDP_0^{1-\varepsilon} \\ \text{(A13)} \quad &= \left( \frac{\varphi_L}{\rho - \varepsilon n + (\Phi - \varepsilon)(g - n)} + \frac{\varphi_0(1-\varphi_L)}{\rho - \varepsilon n + (\Phi - \varepsilon)(g - n) + \varphi} \right) \times \\ &\quad \left( 1 + \varphi_T [\rho - \varepsilon n + (\Phi - \varepsilon)(g - n)] \right)^{-1} \chi GDP_t^\varepsilon GDP_0^{1-\varepsilon}. \end{aligned}$$

This equals the simple rule in (2).  $\square$

## Appendix B: Functional forms and calibration

We have a CES utility function,  $\text{IIA} = \Phi = 2$  and  $\rho = 1\%$  per year. The initial capital stock is set to 150 (US\$ trillion) and  $\delta = 0.1$  per year. Output before damages is given by a CES production function with  $\theta$  the elasticity of substitution and  $0 < \beta < 1$  the share parameter for energy:

$$(C1) \quad Z_t = A_t \left[ (1-\beta) \left( A_0 K_t^\alpha (L_t)^{1-\alpha} \right)^{1-1/\vartheta} + \beta \left( \frac{F_t + R_t}{\sigma} \right)^{1-1/\vartheta} \right]^{\frac{1}{1-1/\vartheta}}, \quad \vartheta \geq 0, 0 < \alpha < 1.$$

The capital-labor composite is defined by a Cobb-Douglas function with  $\alpha$  the share of capital and  $A_t$  total factor productivity. The two energy types are perfect substitutes in production. Damages are calibrated so that they give the same level of global warming damages for the initial levels of output and mean temperature. We rewrite production before damages as

$$(C2) \quad Z_t = Z_0 A_t \left[ (1-\beta) \left( \frac{A_0 K_t^\alpha L_t^{1-\alpha}}{Z_0} \right)^{1-1/\vartheta} + \beta \left( \frac{F_t + R_t}{\sigma Z_0} \right)^{1-1/\vartheta} \right]^{\frac{1}{1-1/\vartheta}}.$$

We set the share of capital to  $\alpha = 0.35$  and the energy share parameter to  $\beta = 0.06$ . Golosov et al. (2014) use a Cobb-Douglas aggregate production function and thus have  $\vartheta = 1$ . Nordhaus (2014) assume a Leontief production function and thus  $\vartheta = 0$ . Hassler et al. (2012) give empirical evidence that suggests that  $\vartheta > 0$  but small. Hence, we have decided to simply use  $\vartheta = 0.5$ . Initial world GDP in 2010 is \$63 trillion. Given  $A_1 = 1$ , we calibrate  $A = 3.78$  to yield initial output under BAU. The energy intensity of output  $\sigma$  is calibrated to initial energy use of 9 GtC under BAU,  $\sigma = 0.15$  GtC/\$T.

### *Population growth and technical progress*

Population in 2010 ( $L_1$ ) is 7 billion people. In the baseline scenario we follow Nordhaus (2014) and UN projections where population growth is given by  $L_t = 11 - 4e^{-7 \cdot 0.01/2 \cdot t}$ . Population growth starts at around 1% per year and falls below 0.1% percent within six decades and flattens out at 11 billion people. In a sensitivity scenario we lower the upper limit to 9 billion. The exogenous trend in total factor productivity growth is 2% per year. In a sensitivity scenario we lower this trend growth rate to 1% per year.

### *Cost of energy*

We employ an extraction technology of the form  $G(S) = \gamma_1 (S_0 / S)^{\gamma_2}$ , where  $\gamma_1$  and  $\gamma_2$  are positive constants. This specification implies that reserves will not be fully be extracted; some fossil fuel remains untapped in the crust of the earth. Extraction costs are calibrated to give an initial share of energy in GDP between 5%-6% depending on the policy scenario. This translates to fossil production costs of \$300/tC (\$30/barrel of oil), where we take one barrel of oil to be

equivalent to 1/10 ton of carbon, giving  $G(S_0) = \gamma_1 = 0.3$ . The IEA (2008) long-term cost curve for oil extraction gives a doubling to quadrupling of the extraction cost of oil if another 1000 GtC are extracted. Since we are considering all carbon-based energy sources (not only oil) which are more abundant and cheaper to extract, we assume only a doubling of production costs if a total 2000 GtC is extracted. With  $S_0 = 4000$  GtC,<sup>24</sup> this gives  $\gamma_2 = 1$ .<sup>25</sup> This implies that we assume very low extraction costs and a high initial stock of reserves which biases our findings toward using more fossil fuel longer.

The unit cost of renewable energy is calibrated to the percentage of GDP necessary to generate all energy demand from renewables. It is customary to calibrate the cost of mitigation in IAMs as a share of output, e.g. Nordhaus (2008) assumes that it costs 5.6% of GDP to achieve full decarbonisation today with this share falling to 0.9% in 300 years and below 0.4% in 600 years. We want to study the energy market in more detail and explicitly model the production costs of fossil and renewable energy. To make our IAM comparable to the wider literature, we calibrate to the extreme but customary case of Leontief technology ( $\rho \rightarrow 0$ ) in which energy demand equals  $\sigma Z_t$  with  $Z_t = AK_t^\alpha (A_t^L L_t)^{1-\alpha}$  pre-damage capital-labour aggregate and  $\sigma$  the carbon intensity of output. Under Leontief technology, the cost of generating all energy carbon-free as a percentage of GDP reduces to  $\sigma Z_t b_t / Z_t = \sigma b_t$ . We are more conservative than Nordhaus and assume that it costs 6% to achieve this. Adding the existing cost of energy, we take  $\sigma b_1 = 0.12$  (i.e. we assume 12% of GDP would go to the energy sector if all carbon emissions were mitigated today) or, with  $\sigma = 0.15$ ,  $b_1 = 0.8$ . In the future this cost falls to current prices of fossil energy (with energy amounting to about 6% of GDP), that is  $b_t$  approaches 0.4). We assume that exogenous technical progress lowers the unit cost at a falling rate starting at a reduction of 1% per year. Specifically,  $b_t = 0.4 + 0.4e^{-0.8 \cdot 0.01/0.4 \cdot t}$ . This calibration assumes that renewable energy is initially very expensive and falls to current levels only in the very long run. Together with the assumption about fossil energy, this biases the model against rapid decarbonization. The calibration is done for a Leontief technology to make our calibration comparable to that of other IAMs. We assume that for a more general technology the same parameter values can be applied.

### *Damages*

Nordhaus (2008) has combined detailed micro estimates of the costs of global warming to get aggregate macro costs of global warming of 1.7% of world GDP when global warming is 2.5o

<sup>24</sup> Stocks of carbon-based energy are notoriously hard to estimate. IPCC (2007) assumes in its A2-scenario that 7000 GtCO<sub>2</sub> (with 3.66 tCO<sub>2</sub> per tC this equals 1912 GtC) will be burnt with a rising trend this century alone. We roughly double this number to get our estimate of 4000 GtC for initial fossil fuel reserves. Nordhaus (2008) assumes an upper limit for carbon-based fuel of 6000 GtC in the DICE-07.

<sup>25</sup> Since  $G(2000) / G(4000) = (4000 / 2000)^{\gamma_2} = 2^{\gamma_2}$  and  $2^{\gamma_2} = 2$ .

C. This figure is used to calibrate the fraction of production output that is left after damages from global warming:  $\tilde{Z}(T_t) = (1 + \zeta_1 T_t^{\zeta_2} + \zeta_3 T_t^{\zeta_4})^{-1}$  with  $\zeta_1 = 0.00284$ ,  $\zeta_2 = 2$ , and  $\zeta_3 = \zeta_4 = 0$ . This is our baseline damage calibration. Weitzman (2010) and Dietz and Stern (2015) argue that damages rise more rapidly at higher levels of temperature than suggested by (5), but empirical studies on the costs of global warming at higher temperatures are not available. With the heroic assumptions that damages are 50% of world GDP at 6o C and 99% at 12.5o C, Ackerman and Stanton (2012) recalibrate (5) with  $\zeta_1 = 0.00245$ ,  $\zeta_2 = 2$ ,  $\zeta_3 = 5.021 \times 10^{-6}$ , and  $\zeta_4 = 6.76$ . The extra term in the denominator captures potentially catastrophic losses at high temperatures. This revised calibration, which we utilize in our of sensitivity runs in section 7, is arguably more appropriate for higher temperatures and but matches baseline damages very closely with deviations of less than 0.5%-point up to 3°C of warming.

### Web appendix C: Simulations with less ambitious climate policy

Table C1 and figure C1 present the effects of allowing for additive production damages, population stagnation, and lower productivity growth. These variants give rise to less ambitious climate policy, leading to lower carbon taxes and less fossil fuel abandoned in situ. The rule (2) implies that with additive climate damages the carbon tax is lower and does not rise with trend GDP growth. Numerical simulations of our IAM confirm this. This implies that much more carbon is burnt in a prolonged fossil fuel era. The first-best carbon budget is 1600 GtC which is only 20 GtC less than the 1620 GtC burnt under BAU.<sup>26</sup> The reason is that with a growing economy additive marginal climate damages fall as a fraction of GDP over time.

**Table C1**

#### Robustness of the proportional tax to less ambitious policy

<i>Scenario</i>	<i>Carbon used (first best)</i>	<i>max. T (first best)</i>	<i>Switch Time (first best)</i>	<i>Carbon used (simple rule)</i>	<i>max. T (simple rule)</i>	<i>Welfare loss (simple rule)</i>	<i>Welfare loss (BAU)</i>
<i>Baseline</i>	990 GtC	3.2 °C	2062	960 GtC	3.1 °C	- .02%	- 7%
<i>Additive damage</i>	1590 GtC	4.0 °C	2075	1540 GtC	3.9 °C	- .03%	- .11%
<i>Population stagnation</i>	1010 GtC	3.2 °C	2065	970 GtC	3.1 °C	- .07%	- 4.9%
<i>Less prod. growth</i>	810 GtC	2.9 °C	2063	790 GtC	2.9 °C	- 0.04%	- 17%

*Note: Welfare losses are relative to the first best (which has welfare loss of 0%) and are reported as percentage losses of initial GDP.*

Our rule (2) also reproduces the first-best outcome also for additive damages remarkably well with the switch to renewable energy occurring in the socially optimal period and virtually the same amount of carbon used over the complete horizon. The welfare loss of using the rule instead of the first-best carbon tax path is thus very low (0.03 percent of initial GDP). As fossil fuel use is price-elastic, the slight difference in taxes yields differences in the carbon budget due to differences in time profile of fossil fuel use despite the synchronicity in transitions.

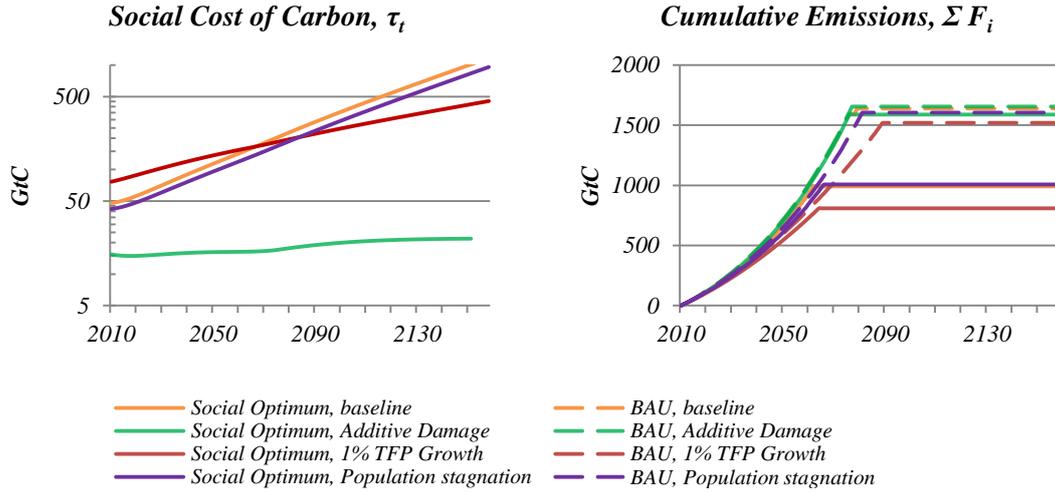
Since population growth is an important driver of climate policy and projections are prone to large revisions, we consider the effect of a stabilization of population at 9 rather than 11 billion people at the end of the century as predicted by Lutz et al. (2014). Fewer people in the future curb the benefit of climate policy so that current generations can consume fossil fuel for longer. The initial carbon tax is lower than under baseline. However, as the future population is larger,

<sup>26</sup> Weitzman (2009) finds that with additive global warming damages (in instantaneous welfare) the willingness to sacrifice current consumption to avoid future global warming is seven times higher than with multiplicative damages. In contrast, we find that additive damages lead to half the SCC at each point of time compared to multiplicative damages. However, this stark difference should serve as a reminder that the additive utility damages used in Weitzman and the additive production damages used in our model are not comparable. Weitzman's number applies to the willingness to pay to avoid any change in temperature along an exogenous consumption trajectory, whereas the SCC of our simulations reports the willingness to pay to avoid a marginal increase in atmospheric carbon along the optimal path. Our numerical results are in line with the analytical findings in section 2.3: the SCC is higher under multiplicative damages provided that the trend rate of economic growth is positive.

productive capacity is decreased and less fossil fuel is used during the longer fossil fuel era. Still, the overall effect of this drop in the future population is small. In total only 20 GtC more are burnt. Despite the fact that the simple rule (2) does not account for a temporary population boom, it still manages to capture the extent of the climate externality accurately.

**Figure C1**

**Sensitivity analysis for the optimal SCC and cumulative emissions**



*Note: Lower damages and higher population growth lower the SCC and boost fossil fuel use. Lower growth flattens the time profile and increases the carbon budget.*

Table C1 and figure C1 also show the effects of lower trend growth in living standards (1 instead of 2 percent per annum). Since future generations are worse off, climate policy is initially more ambitious but future carbon taxes are lower. Overall the initially higher carbon tax dominates climate policy and less carbon is burnt and a higher stock of carbon abandoned in situ.<sup>27</sup> Here, the welfare loss associated with the approximation (2) is less than 0.05 percent.

**Web appendix D: More general climate damages**

Theoretical models often have global warming affecting welfare negatively rather than productive activity. Although damages and consumption sometimes enter welfare in separable fashion (e.g., Weitzman, 2009), it is important to consider non-separable welfare functions

$$(B.1) \quad U(C_t, Q_t) = \left[ \frac{1}{1+\psi} C_t^{\frac{\zeta-1}{\zeta}} + \frac{\psi}{1+\psi} Q_t^{\frac{\zeta-1}{\zeta}} \right]^{\frac{\zeta}{\zeta-1(1-\Phi)}} / (1-\Phi)$$

with  $\zeta$  the elasticity of substitution between consumption of final goods  $C_t$  and environmental quality  $Q_t$  (a negative function of the stock of atmospheric carbon  $E_t$ ), and  $\Phi$  the coefficient of IIA as before (e.g., Guesnerie, 2004; Traeger, 2007; Hoel and Sterner, 2007; Sterner and Persson, 2008). This allows marginal utility of consumption to increase or decrease with

<sup>27</sup> We could also have discussed this scenario in section 5.

environmental quality.<sup>28</sup> The parameter  $\zeta$  is the percentage rate at which demand for  $C_t$  declines if the relative price of environmental quality increases by 1%. If  $\zeta \rightarrow 1$ , utility tends to  $C_t^{(1+\psi)^{-1}} Q_t^{\psi(1+\psi)^{-1}}$ .<sup>29</sup> Gollier (2010; 2013, Chapter 10, pp. 164-165) gives two results that follow from (4). First, the two goods are substitutes if  $\zeta\Phi > 1$  in which case a higher rate of economic growth and anticipated deterioration of environmental quality boost the ecological discount rate and thus depress the SCC. Second, there will be a term structure of interest rates as the Euler equation has an additional certainty-equivalent term on the right-hand side that falls with time  $t$  if  $(\zeta\Phi - 1)(1 - \zeta) > 0$ .<sup>30</sup> Future climate damages are then discounted less heavily.

Various IAMs specify damages as shock to the depreciation rate of physical capital (e.g., Fisher and Hanemann, 1993; Dietz and Stern, 2015; Fisher and Le, 2014; Hsiang and Jina, 2015). These damages are one-off as the capital stock typically recovers after some time. The resulting contribution to the SCC would thus not be very substantial (and not rise with GDP). As most IAMs employ a specification where climate change negatively affects net economy-wide production by reducing gross production by a percentage that depends on global mean temperature (see Kopp et al. (2011) for a survey), our focus is on production damages. We do not consider climate damages in social welfare or capital destruction shocks any further.

## References

- Barrage, Lint. 2014. Optimal dynamic carbon taxes in a climate-economy model with distortionary fiscal policy. Unpublished manuscript, Department of Economics, Brown University.
- Fisher, Anthony C., and Michael Hanemann. 1993. Assessing climate risks: valuation of effects, in *Assessing Surprises and Nonlinearities in Greenhouse Warming*, ed., Joel Darmstadter and Michael A. Toman. Washington, DC: Resources for the Future.
- Fisher, Anthony C., and Phu V. Le. 2014. Reflections - climate policy: science, and extremes. *Review of Environmental Economics and Policy* 8: 307-327.
- Gollier, Christian. 2010. Ecological discounting. *Journal of Economic Theory* 145: 812-829.
- Guesnerie, Roger. 2004. Calcul économique et développement durable. *Revue Economique* 55: 363-382.
- Hassler, John, Per Krusell, and Conny Olovsson. 2012. Energy-saving technical change. NBER Working Paper 18456, National Bureau of Economic Research, Cambridge, MA.
- Hsiang, Solomon, and Amir S. Jina. 2015. Geography, depreciation, and growth. *American Economic Review: Papers and Proceedings* 105: 5, 253-256.
- Hoel, Michael, and Thomas Sterner. 2007. Discounting and relative prices. *Climatic Change* 84: 265-280.
- Kopp, Robert, Alexander Golub, Nathaniel Keohane, and Chikara Onda. 2011. The influence of the specification of climate change damages on the social cost of carbon. *Economics-E-Journal* <http://dx.doi.org/10.5018/economics-ejournal.ja.2012-13>.

<sup>28</sup> Climate damages have been carefully split up into its multiplicative production and additive utility components (Barrage, 2014). In the model of Golosov et al. (2014) multiplicative production damages are equivalent to additive linear damages in utility.

<sup>29</sup> The model of Golosov et al. (2014) is equivalent to a model with a separable damage term in welfare that is linear in the carbon stock.

<sup>30</sup> Weitzman (2009) uses (4) with  $\zeta = 0.5$  and  $\Phi = 2$  in which case the term structure is flat.

- Stern, Thomas, and U. Martin Persson. 2008. An even Stern review: introducing relative prices into the discounting debate. *Review of Environmental Economics and Policy* 2: 61-76.
- Traeger, Christian P. 2007. Sustainability, limited substitutability and non-constant social discount rates. Discussion Paper 1045, Department of Agricultural and Resource Economics, University of California at Berkeley.