Optimal Carbon and Income Taxation

Alex Schmitt*

January 31, 2014

Abstract

At what rate should a government tax carbon emissions? This paper analyzes optimal carbon taxation while taking into account that taxes are set by national policy makers. I add three real-world features to a standard climate-economy model, namely distortionary income taxation, lack of commitment to future policies, and no policy cooperation between regions. I find that in a calibrated two-region model, the optimal carbon tax decreases by more than half relative to the global first-best rate as a result of these features, assuming that a regional government only maximizes domestic welfare. The intuition for this result is twofold: first, under unilateral policy making, the government does not take into account climate damages to the rest of the world and therefore faces a lower welfare cost of emitting carbon. Second, resorting to distortionary income taxes reduces the optimal carbon tax further, partly because of a tax-interaction effect. This effect, which is caused by the presence of second-best costs and benefits of reducing emissions, drives a wedge between the optimal tax and the Pigouvian rate.

Key words: climate-economy modeling, carbon tax, optimal income taxation, unilateral climate change policy

JEL Classification: E61, E62, H21, H23, Q54

*IIES, Stockholm University, SE-106 91 Stockholm, Sweden. Email: alex.schmitt@iies.su.se. Phone: +46 (0)8 16 30 72. I am very grateful to John Hassler and Per Krusell for their supervision and support. I thank Philippe Aghion, Lint Barrage, David von Below, Inge van den Bijgaart, Timo Boppart, Tobias Broer, Pamela Campa, Tom Cunningham, Reyer Gerlagh, Sebastian Koehne, Andreas Mueller, Torsten Persson, Rick van der Ploeg, Nick Sheard and seminar participants at the IIES brown bag seminar, the Oxcarre seminar at Oxford University, the Stockholm-Uppsala Doctoral Workshop in Economics, the AERE summer conference, the European Macroeconomics Workshop at the LSE, the Conference on Computing in Economics and Finance, the Nordic Summer Symposium in Macroeconomics, the GSS seminar at University of Tilburg, the European Winter Meeting of the Econometric Society, and the brown bag seminar at the Department of Economics at Stockholm University for their useful comments. All errors are my own.

1 Introduction

At what rate should a policy maker tax emissions of greenhouse gases such as CO_2 in order to internalize climate change? In this paper, I address this question while taking into account that climate policy is chosen by national or regional governments. Specifically, I consider three real-world features that are relevant from the perspective of national policy makers, and show how optimal carbon taxes are affected when adding these features to an otherwise standard climate-economy model. These features are i) distortionary income taxation, ii) an inability of the government to commit to future policies, and iii) unilateral policy making, that is, a lack of full policy cooperation between countries or regions. Previous climate-economy models focusing on globally optimal carbon taxation have typically abstracted from these features. Hence, this paper sheds some light on the question of how useful these "global first-best" models are in guiding policy makers: if their results are very sensitive to the introduction of arguably realistic features, then this limits their policy relevance.

The main findings of this paper are the following. First, resorting to distortionary income taxes rather than first-best lump-sum taxation to finance public goods gives rise to a "tax-interaction effect", which drives a wedge between the level of marginal climate damage - i.e., the Pigouvian rate - and the optimal carbon tax. This effect is caused by costs and benefits of reducing emissions that only materialize in "second best". Quantitatively, I find that the magnitude and the direction of the net tax-interaction effect depend on the size of the carbon tax and hence on the cost of climate damages. If the effect of climate change on social welfare is sufficiently small, then the tax-interaction effect results in an optimal carbon tax equal to or even above the Pigouvian rate. Second, when climate policy is implemented unilaterally, the size of the carbon tax that is optimal from the perspective of a regional government depends on whether the government only maximizes the welfare of domestic households, rather than global welfare. In this case, in a model with two equally-sized regions, the optimal carbon tax is less than half of the global first-best Pigouvian rate. Part of this gap is due to the fact that without policy cooperation, a domestic government does not take into account climate damages in the rest of the world, while the remainder can be attributed to the distortionary impact of income taxes and to the tax-interaction effect. In contrast, if the national government maximizes global welfare, unilateral policy making and distortionary taxation partially offset each other, resulting in an optimal carbon tax which is about 20% below the global first-best rate.

This paper builds on "integrated assessment models" (IAM), such as, for example, the DICE model (Nordhaus, 2008) or Golosov et al. (2013), that are widely used to inform policy makers about the size of the social cost of carbon and hence the optimal carbon fee.[1,2] These models typically prescribe a Pigouvian tax that equals the marginal global climate damage, which is defined as the present value of the damage caused globally by emitting an additional unit of carbon. As in these studies, I analyze optimal climate policy in a standard deterministic neoclassical growth model. However, I relax two assumptions that are crucial for obtaining a global Pigouvian tax and that are not realistic in the context of national policy making: lump-sum taxation and full policy cooperation between regions.

Specifically, I take into account that a government's role is not limited to implementing climate policy, but it must also raise revenue in order to finance expenditures on public goods, by taxing labor and capital income. This introduces distortions into the economy that make the first-best allocation unfeasible, even if the government is assumed to be benevolent. When modeling distortionary taxation, I assume that all tax rates are chosen optimally.[3] In addition, it is well-known that in models with distortionary income taxes, one has to take a stand on whether the government is able to credibly commit to future policies (Klein et al., 2008). I assume that there is no such commitment device. This assumption is arguably more realistic and plausible than allowing a government to commit to all future tax rates.

Moreover, previous climate-economy models often feature a global planner or government implementing climate policy, either in one region – the world – or in a multiple-region setting. In other words, they analyze carbon taxation assuming that all countries participate in pricing carbon optimally. This is a restrictive assumption: neither is there a governmental body that can impose and enforce regulation on a global level, nor do we observe cooperation between countries on a global scale when setting carbon prices. Instead, existing pricing schemes, for example the European Union's Emission Trading System (EU ETS) or the carbon tax in Australia, were implemented unilaterally, while many major polluting countries, such as the US or China, do not tax carbon emissions at all.[4]

I start by considering each feature separately. In the first part of the paper, for simplicity, I only focus on distortionary fiscal policy. Specifically, I analyze optimal climate policy both qualitatively and quantitatively in a global planner economy with distortionary income taxes and lack of commitment.[5] I provide an analytical characterization of the social cost of carbon

[1] In this paper, a carbon fee could be either a direct carbon tax or the price of tradable emission permits. I will use the terms "tax" or "fee" as synonyms. Note that in order to optimally correct a pollution externality, the emission fee must be set equal to the marginal social cost of pollution evaluated at the efficient emission level (Kolstad, 2000). The question of what carbon tax is optimal is therefore equivalent to asking what is (a good estimate for) the marginal social cost of carbon.

[2] For example, an Interagency Working Group of the US government recently published a report determining the social cost of carbon to be used in cost-benefit analysis (IWG, 2010, 2013). They have used DICE as well as the PAGE model (Hope, 2006, 2008) and the FUND model (Tol, 2002a,b; Anthoff et al., 2009).

[3] There is a large literature in environmental economics that instead analyzes partial tax reforms, where non-environmental tax rates, and possibly the pollution tax, are exogenously given. An example is Glomm et al. (2008) for a dynamic growth model. Compare also the quantitative results in Barrage (2013).

[4] Figure B in the appendix gives an overview of current carbon pricing schemes.

[5] Note that throughout this paper, I focus on Markov-perfect equilibria when solving the model with distor-

and show that the optimal second-best carbon tax is in general not equal to the marginal climate damage and hence is not at the Pigouvian rate. This is due to additional effects of carbon taxation that only materialize in a setting with distortionary income taxes, and that are caused by the interaction of carbon emissions with current and future "wedges", that is, distortions of a first-best margin. For example, under certain conditions, an emission tax leads to a decrease in the labor supply, which results in a welfare loss if the intratemporal labor-leisure margin is distorted by a labor income tax. Similarly, taxing carbon has a negative effect on household savings, thereby exacerbating the intertemporal distortion caused by the capital income tax. These effects represent second-best costs of emission reductions and thus decrease the optimal carbon tax. In addition, I show that there are future wedges that are affected by contemporaneous carbon emissions.

The deviation of the optimal emission tax from the Pigouvian rate was a prominent finding in earlier studies, which considered second-best environmental taxation in a static model with a labor income tax (Bovenberg and de Mooij, 1994; Parry, 1995; Bovenberg and Goulder, 1996) and hence focused on the labor-leisure wedge.[6] In a dynamic model, however, taking into account only the second-best effect of carbon emissions on the current labor supply gives an incomplete characterization of the social cost of carbon.

A quantitative assessment shows that the second-best effects go in different directions. In other words, some constitute costs of emission reduction, while others represent benefits. The net effect depends on the size of the carbon tax and hence on the welfare effect of climate damages. In the baseline calibration of my global climate-economy model, where preferences are such that temperature stabilizes at a level of $3°C$ above the pre-industrial level in the long run, I find that the optimal carbon tax is about 14% below the Pigouvian fee. Thus, the second-best costs of reducing carbon emissions outweigh the benefits. For higher target temperatures, and hence lower tax rates, these costs become smaller. If the disutility of climate damages is sufficiently low, then they are more than offset by the benefits, implying that the optimal carbon tax is above the Pigouvian level.

The second part of the paper analyzes optimal climate policy in a stylized two-region model with unilateral policy. I divide the global economy into two equally-sized regions and assume that the government in the "domestic" region chooses climate and fiscal policies optimally. In the other region, there is no role for the government and, in particular, there is no carbon taxation. This precludes strategic interaction between the two regions: the domestic government takes into account the behavior of agents in the rest of the world when choosing its policies, but not vice versa. While this is an extreme assumption, it is arguably a good approximation of reality.

In this setting, I compute the time path of the optimal carbon tax for the domestic region, and compare it to the Pigouvian tax schedule from the first-best global planner problem. I make

tionary taxation.

[6]Those studies usually found that the tax-interaction effect dominates "revenue recycling", that is, using the revenue raised with the emission tax to lower distortionary income taxes, which results in an efficiency gain.

the key assumption that capital is perfectly mobile between regions. I distinguish two cases. If the domestic government only maximizes the welfare of its own households, rather than global welfare, I find that under distortionary taxation and unilateral climate policy, the carbon tax in 2011 amounts to 96 \$/tC, as compared to 211 \$/tC in global first best. In comparison, with lump-sum taxation, going from a global planner economy to a two-region model with unilateral policy results in an optimal carbon tax of 126 \$/tC in 2011. Hence, the absence of policy cooperation between regions accounts for the larger share of the total decrease. Intuitively, this is due to the fact that the domestic government only takes climate damages in its own region into account, while disregarding the rest of the world. The distortionary impact of the income taxes and the tax-interaction effect further increase the gap between the global Pigouvian tax and the optimal second-best carbon tax in the domestic region.

In the second case, assuming that the domestic government maximizes global welfare and internalizes all climate damages, the optimal carbon tax under lump-sum taxation is above the global first-best rate. By reducing its carbon emissions, the domestic government partly offsets higher emissions in the rest of world. Adding distortionary taxes once more leads to a lower carbon tax rate. Hence, the two distortions move in opposite directions, and therefore partially offset each other. The net effect is an optimal carbon tax which is only about one fifth below the global first-best rate.

1.1 Related Literature

This paper is related to different strands of the literature, both in climate-economy modeling and public finance. In the context of climate change, numerous studies have used IAMs to compute the social cost of carbon and the optimal carbon tax, often in a global planner model without distortionary taxation.[7] One of the earliest and most influential IAMs is the DICE model (Nordhaus, 2008), which features a neoclassical growth model similar to the one used below.[8,9] The estimates for the SCC in DICE have increased over the past versions. The newest version, DICE-2013R, finds an optimal carbon price of 66\$/tC in 2015 (Nordhaus, 2013). A multi-region version of this framework is the RICE model (Nordhaus and Yang, 1996; Nordhaus, 2010). Here, in a cooperative regime, a planner sets carbon prices optimally in all regions to maximize global welfare for a given set of welfare weights.[10] Nordhaus (2010) finds an optimal

[7]One example of a "second-best" model is Gerlagh and Liski (2012) They compute Markov-perfect optimal carbon fees in a setting without distortionary taxes, but where the government is unable to commit to future policies due to time-inconsistent preferences, that is, hyperbolic discounting.

[8]A notable difference is that in DICE, fuel use is not explicitly modeled. Instead, carbon emissions are linked proportionally to output. The planner can invest in abatement, which reduces the amount of pollution for a given output level (Nordhaus, 2008).

[9]Other examples of frequently used IAMs are the PAGE model (Hope, 2006, 2008) and the FUND model (Tol, 2002a,b; Anthoff et al., 2009). Note that these are not optimal growth models, but instead take output scenarios as given.

[10]Hence, in the cooperative regime, the RICE model does not explicitly model climate agreements. See Barrett (2005) for an overview of the literature on environmental agreements and Harstad (2012, 2013) for

carbon tax of 29$/tC in 2010.

In addition, RICE can be used to analyze a setting without cooperation among regions. Specifically, Nordhaus and Yang (1996) compute a full-information "open loop" Nash equilibrium. An IAM related to RICE is the WITCH model (Bosetti et al., 2006), which also consists of multiple regions that play a noncooperative game. However, WITCH features a more detailed representation of energy production. Moreover, technical change is endogenous.

Golosov et al. (2013) consider a climate-economy model similar to DICE, although with a different formulation of the carbon cycle and explicit modeling of fossil fuel use. Notably, they derive a closed-form expression for the expected social cost of carbon under certain conditions - in particular, logarithmic utility and full depreciation - and show that the optimal tax as a share of contemporaneous output is constant. Quantitatively, they find an optimal carbon tax 57$/tC in 2010.

As indicated above, the present model is a generalization of the literature on static models initiated by Bovenberg and de Mooij (1994). While Bovenberg and de Mooij (1994) analyze a model with a pollution-intensive consumption commodity, Parry (1995) and Bovenberg and Goulder (1996) examine the case of a "dirty" input in the production process, similar to the setting in this paper.[11]

In addition, this paper is related to more recent work by Barrage (2013). She also considers the interaction between optimal income and carbon taxation in a neoclassical climate-economy model with distortionary income taxes. However, her analysis differs from this paper in three key dimensions. First, throughout her paper, she studies a setting in which the government is assumed to be able to commit to future tax rates. In her baseline model, this results in a zero capital tax.[12] Second, her model comprises only one region and hence focuses on global climate policy. Finally, with regard to the research question, Barrage (2013) is interested in the qualitative difference between climate damages to utility and to productivity, and under which conditions they are not fully internalized. By contrast, my main qualitative result does not make a distinction between utility and production damages, but instead relates the tax-interaction effect to different wedges.

Methodologically, this paper is related to Klein et al. (2008), Azzimonti et al. (2009), and Martin (2010) who analyze time-consistent Markov-perfect equilibria in a standard neoclassical growth model without environmental quality.[13] Analogous to Klein et al. (2008), I derive the current government's generalized Euler equations, which are weighted sums of intertemporal and

recent work on agreements in a dynamic climate model.

[11]Compare also Bovenberg and van der Ploeg (1994), Goulder (1995), Goulder et al. (1997) and Bovenberg and Goulder (2002).

[12]She also considers extensions to her model where capital taxes can either be temporarily non-zero, due to an upper bound that binds for a finite number of periods, or where a permanently positive capital tax is exogenously given.

[13]See Fischer (1980) and Lucas and Stokey (1983) for earlier work on the time inconsistency of optimal policy in the presence of distortionary incomes taxes. Kehoe (1989) extended the model in Fischer (1980) to a two-country setting.

intratemporal wedges that the government trades off against each other. Similar to Azzimonti et al. (2009), this paper analyzes a *stock* of a public good, rather than a pure flow.[14] Note that while the application here is with respect to an environmental public good, the analysis would be similar to the case of the stock of a non-environmental public good. For example, one could think of infrastructure such as public roads and buildings as a persistent public good, i.e., expenditures today matter for the stock tomorrow.

The remainder of this paper is structured as follows. Section 2 presents the framework. In section 3, I analyze a global climate-economy model with distortionary taxation. Section 4 considers the two-region setting with unilateral policy making, first with lump-sum taxes, and then with both distortions. Section 5 concludes the paper.

2 The Framework

In this section, I introduce a simple dynamic framework in which I analyze second-best carbon taxation. Consider the standard neoclassical growth model, extended by "fossil fuel" or "energy" and "environmental quality", that is, the state of the climate. Fossil fuel is used as a factor of production, in addition to capital and labor. Burning fuel causes emissions of a pollutant, here carbon. Hence, the amount of fuel used in production is a determinant of climate change, which affects both the utility function of the representative household - as in the static second-best literature following Bovenberg and de Mooij (1994) - and the productivity of the representative firm, as, for example, in Golosov et al. (2013). Producers do not take into account how their decisions affect the climate; hence, carbon emissions represent an externality.

Note that throughout this paper, I employ deterministic models and abstract from all uncertainty related to the climate or economic development. Moreover, technological growth is exogenous.

2.1 Utility and Household's Problem

The representative household's per-period utility is given by $u(C, 1-h, G, T)$, where C denotes private consumption of a final good, h hours worked, G public consumption and T is an indicator of climate change. Specifically, T represents the change in mean global surface temperature relative to the preindustrial period. The latter two variables are not chosen by the household; hence, they represent public goods. u is increasing in its first three arguments and decreasing in T. In other words, a higher T corresponds to a "worse" state of the climate. There are several channels through which permanently warmer temperatures cause disutility (Barrage, 2013), for example, by affecting health and general well-being.

[14] Battaglini and Coate (2007, 2008) also consider an environment with distortionary income taxation and public good provision, but their focus is on the political economy of fiscal policy and in particular on legislative decision making.

Note that in contrast to many papers in public finance, I have assumed here that the public consumption good is valued by the household; therefore, the amount provided is a choice variable of the government or planner. This assumption is important for the infinite-horizon version of the model in sections 3 and 4 for technical reasons.[15] For the two-period model in 3.1, it is not essential, and therefore I will simplify the analysis by assuming exogenous government expenditures.

The population size is constant. Let A_t denote the level of labor-enhancing productivity, which grows exogenously at the rate ζ.[16] From here on, I express variables in efficiency levels. That is, $c_t \equiv C_t/A_t$ and $g_t \equiv G_t/A_t$.

Households maximize lifetime utility, subject to their budget constraint, taking price and tax sequences as given:

$$\max_{\{c_t,h_t,k_{t+1}\}_{t=0}^{\bar{t}}} \sum_{t=0}^{\bar{t}} \beta^t u(c_t, 1-h_t, g_t, T_t), \tag{2.1}$$

subject to

$$c_t + k_{t+1} \leq [1 + (1-\tau_t^k)(r_t - \delta)]k_t + (1-\tau_t^h)w_t h_t, \tag{2.2}$$

where $\bar{t} \leq \infty$. $k_t \equiv K_t/A_t$ denotes the capital stock in period t. r_t and $w_t = W_t/A_t$ are the factor prices of capital and labor, respectively, while τ^k and τ^h are the corresponding linear tax rates. δ denotes the rate of capital depreciation. Solving this problem yields two standard optimality conditions, one intertemporal (consumption-savings) and one intratemporal (consumption-leisure).

When solving the model numerically in sections 3 and 4, I need to choose a functional form for the per-period utility function. In the baseline model, following Klein et al. (2008), u is additively separable and logarithmic in the first three arguments.[17] Moreover, I assume a quadratic utility damage due to the externality:

$$u(c,l,g,T) = [(1-\alpha_g)\alpha_c \ln c + (1-\alpha_g)(1-\alpha_c)\ln l + \alpha_g \ln g] - \frac{\alpha_T}{2}T^2. \tag{2.3}$$

As is standard in the environmental economics literature, I let preferences between private consumption and environmental quality be additively separable (for example Cremer and Gahvari, 2001) and assume that disutility is convex in temperature change (Weitzman, 2010). The assumption of additivity is convenient, since it facilitates the analytical and numerical analysis. Moreover, relaxing this assumption is not straightforward. While there is some evidence that higher temperatures have an effect on the marginal utility of leisure (for example Zivin and Neidell, 2010), it is unclear whether, on aggregate, leisure and climate are substitutes or complements. In addition, how to specify a non-separable utility function with temperature change in a macroeconomic model is an open research question.

[15] Compare the discussion in the appendix, section A.3, for details.

[16] That is, $A_{t+1} = \zeta A_t$. For the theoretical analysis, to save notation, I will assume that $\zeta = 1$.

[17] This specification is convenient since it allows for the existence of a balanced growth path (King et al., 2002).

2.2 Production, Fuel Use and Firm's Problem

The consumption good is produced with a technology represented by a function F, which uses as inputs capital, labor and fuel, denoted by M. Moreover, the temperature change T does not only affect utility, but also has an impact on the production process. "Net" output - taking damages from climate change into account - is given by $Y_t = F(K_t, A_t h_t, B_t M_t, T_t)$. B_t denotes energy-augmenting productivity or "energy efficiency", which grows exogenously. Hence, there is a second source of technological progress, apart from the labor-enhancing productivity A_t. I assume that A_t and B_t grow at the same exogenous rate ζ, which is necessary for a balanced growth path, as further discussed in section 3.3.1 below.

Following Nordhaus (2008) and Golosov et al. (2013), I assume that T enters the production function multiplicatively:

$$Y_t = F(K_t, A_t h_t, B_t M_t, T_t) = [1 - d(T_t)] f(K_t, A_t h_t, B_t M_t), \tag{2.4}$$

where the "damage function" d captures damages to productivity, with $0 \leq d(T) \leq 1$. It is assumed to increase in temperature ($d_T > 0$). Moreover, f exhibits constant returns to scale. Hence, dividing both sides of (2.4) by A_t gives the output in efficiency units:

$$y_t = \frac{Y_t}{A_t} = [1 - d(T_t)] f(k_t, h_t, m_t), \tag{2.5}$$

where $m_t \equiv B_t M_t / A_t$.

Regaring fuel use, to simplify the exposition, I assume that there is no scarcity problem and thus, m is available at "infinite" capacity. Hence, I abstract from the Hotelling problem of how to optimally extract a finite resource. Fuel here should be interpreted as coal rather than (conventional) oil or gas.

Moreover, let the *private* marginal cost of fuel use be given by an exogenous process κB_t. This could, for example, be interpreted as a per-unit extraction cost paid in terms of the final good, which is assumed to grow at the same rate as energy-augmenting productivity. In efficiency units, using one unit of m is then associated with the constant cost κ.[18]

A representative firm then solves the following problem:

$$\max_{k_t, h_t, m_t} [1 - d(T_t)] f(k_t, h_t, m_t) - r_t k_t - w_t h_t - (\theta_t + \kappa) m_t, \tag{2.6}$$

where θ_t denotes a tax on fuel use (per efficiency unit) and hence on carbon emissions. From the first-order condition, it follows that the carbon tax is given by $\theta_t = f_m(t) - \kappa$.[19] The economy's resource constraint in efficieny units is given by:

$$c_t + g_t + k_{t+1} + \kappa m_t = F(k_t, h_t, m_t, T_t) + (1 - \delta) k_t. \tag{2.7}$$

[18]Alternatively, one could let the cost be a function of the resources left in the ground, or model energy production as a separate production sector that uses labor and possibly capital as in Golosov et al. (2013) or Barrage (2013).

[19]Note that $B_t \theta_t$ gives the carbon tax in $\$/GtC$.

For the quantitative analysis, I assume a nested CES for the gross production function:

$$f(k, h, m) = \left[(1 - \xi)\left(k^{\nu} h^{1-\nu}\right)^{\frac{\sigma-1}{\sigma}} + \xi m^{\frac{\sigma-1}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}}. \tag{2.8}$$

In the baseline model, I let $\sigma \to 1$, which results in a Cobb-Douglas function:

$$f(k, h, m) = k^{\rho} m^{\xi} h^{1-\rho-\xi}, \tag{2.9}$$

with $\rho = \nu(1 - \xi)$. Finally, a common functional form for the damage function is:

$$d(T) = (1 + \gamma T^2)^{-1} \gamma T^2. \tag{2.10}$$

This is, for example, used in the DICE model (Nordhaus, 2008).

2.3 Climate Change

In addition to the private extraction cost, using fuel has a social cost: it causes carbon emissions, which negatively affect the state of the climate by increasing the mean global surface temperature. Many IAMs model this mechanism in two steps (Nordhaus, 2008; Golosov et al., 2013): first, past (and possibly current) carbon emissions, plus a vector of initial carbon concentrations in the atmosphere and other reservoirs like the upper and lower oceans, translate into current carbon concentrations. Second, the current vector of carbon stocks, $\mathbf{s}_t$, maps into the mean global temperature change in period t: $T_t = \mathcal{F}(\mathbf{s}_t)$.[20] One consequence of this modeling strategy is that many IAMs typically feature multiple variables summarizing the state of the climate (Nordhaus, 2008; Cai et al., 2013).

In this paper, to keep the number of state variables low, I use a more reduced-form approach. Specifically, I assume a direct mapping $T_t = q(\mathbf{m}^t)$, where $\mathbf{m}^t = \{m_t, m_{t-1}, ..., m_{t_0}\}$ denotes the history of past global CO_2 emissions back to period t_0, and $\partial q / \partial m_j > 0 \; \forall j$. In words, the current *flow* of carbon has an impact on the state of the climate in future periods. This is a reasonable assumption, given that carbon stays in the atmosphere for a very long time horizon.

The functional form for q that is chosen below is based on Matthews et al. (2009). They define the "climate-carbon response" (CCR) as the ratio of global mean temperature change and total cumulative carbon emissions over some period of time. Using both historical emission data and an ensemble of climate models, they show that this variable is almost constant over time, and in particular, it is independent of the atmospheric carbon concentration. From this observation, it follows that the increase in global mean surface temperature can be written recursively as:

$$T_{t+1} - T_t = \text{CCR} \cdot m_t. \tag{2.11}$$

[20] Carbon is sometimes referred to as a "stock pollutant", since the stock in the atmosphere, rather than the flow, matters for climate change.

For the quantitative analysis, this specification is convenient, since it summarizes the climate side of the model in one state variable, T_t, while abstracting from the carbon concentrations in the different reservoirs. Note that in contrast to other climate-economy models (Nordhaus, 2008; Golosov et al., 2013; Gerlagh and Liski, 2012), this specification implies that an increase in the global mean temperature due to carbon emissions is irreversible. This has quantitative implications on the size of the optimal carbon tax, as discussed in more detail in section 3.3.3.

2.4 Government

The government has to finance the public good by taxing labor and capital income. By assumption, lump-sum taxes are not available. Its budget constraint reads:

$$g_t \leq \tau_t^k (r_t - \delta) k_t + \tau_t^h w_t h_t + \theta_t m_t. \tag{2.12}$$

Note that throughout this paper, I assume that the government has to balance its budget in every period. In other words, it can neither borrow from nor lend to households. The latter assumption is crucial, as pointed out by Azzimonti et al. (2006). They show that if the government were allowed to accumulate assets, it would be able to dispense with distortionary taxation after a finite number of periods. Hence, even in the absence of commitment, a government would set a zero tax rate on capital income in the long run. The intuition for this result is that the government could confiscate all income in the first period, and then lend to households every period and accumulate assets over time. After a sufficiently large number of periods, the government's wealth would be large enough to finance the public good without resorting to distortionary taxation. Therefore, since I want income tax rates to be non-zero in every period, I abstract from government assets.

3 A Global Model With Distortionary Taxation

3.1 Finite Horizon

In this section, I consider two finite-horizon versions of the framework introduced in the previous section, a static model where $\bar{t} = 0$ and a two-period model with $\bar{t} = 1$. The purpose of this section is twofold. First, it allows me to review the main result in the literature that started with Bovenberg and de Moij (1994) and illustrate its key mechanism that will be generalized in the infinite-horizon setting in the next section. Second, it introduces some notation and provides some intuition for the results found later on.

Throughout the section, for simplicity, assume that the public good is not valued by the household, and hence its per-period utility $u(c, l, T)$ depends on private consumption, leisure and the state of the climate. Instead, government expenditures are exogenously given.

3.1.1 The One-Period Model Revisited

In the one-period model, investment is zero and the capital stock is given. Environmental quality is a flow variable. That is, assume that temperature change is given by $T = q(m)$. Note that such a static model is not really appropriate to analyze long-term effects on the climate: as mentioned above, greenhouse gases are stock rather than flow pollutants.

Start by defining two wedges, i.e. distortions of the first-best margins. Let ω_{LL} denote the "labor-leisure wedge" and ω_{Env} the "environmental wedge", respectively. They are given by:

$$\omega_{LL} \equiv u_c F_h - u_l \tag{3.1}$$

$$\omega_{Env} \equiv u_c(F_m - \kappa) + q_m[u_s + u_c F_T]. \tag{3.2}$$

The first-best equilibrium, when lump-sum taxes are available, is characterized by both wedges being zero. For the environmental wedge, this just implies that the marginal benefit of increasing emissions net of private cost - the first term on the right-hand side of (3.2) - in terms of utility equals the marginal climate damage, captured by the second term.

Turning to the second-best setting with distortionary taxation, I use these wedges to express the government's optimality condition, which characterizes optimal policy. This approach follows Klein et al. (2008) and will be at the core of the analysis in the remainder of the paper.

The government's problem in this economy is given by

$$\max_{m,h} u(F(k,h,m) - g - \kappa m, 1-h, q(m))$$

s.t. an "implementability constraint" which is derived from the household's optimality condition and its budget constraint:[21]

$$\begin{aligned}\chi(h,m) \equiv u_c\big(F(k,h,m) - g - \kappa m, 1-h, q(m)\big)[F(k,h,m) - g - \kappa m]\frac{F_h h}{F_h h + F_k k} \\ -u_l\big(F(k,h,m) - g - \kappa m, 1-h, q(m)\big)h = 0.\end{aligned}$$

Note that I have assumed a "total income tax", that is, both capital and labor income are taxed at the same rate $\tau = \tau^k = \tau^h$.[22]

Define the function $\mathcal{H}$ implicitly by

$$\chi(\mathcal{H}(m), m) = 0. \tag{3.3}$$

In words, for given fossil fuel use m, $\mathcal{H}(m)$ gives the household's optimal labor supply, i.e. the labor choice that satisfies its optimality condition. In a sense, it is the household's "best

[21] That is, from consolidating $u_c F_h(1-\tau) - u_l = 0$ and $c = (1-\tau)(F_k k + F_h h)$. Note that if both the household's budget constraint and its resource constraint, $c = F(k,h,m) - g - \kappa m$, are satisfied, so is the government's budget constraint, $g \leq \tau(F_h h + F_k k) + (F_m - \kappa)m = \tau(F_h h + F_k k) + \tau_m m$.

[22] Note that since the tax on capital income works as a lump-sum tax in this setting, assuming separate tax rates would eliminate the distortion, that is, would result in a zero labor tax.

response" to the fossil fuel use announced by the government, assuming within-period commitment. Using $\mathcal{H}$, the government's problem can be more compactly written as:

$$\max_{m} u(F(\mathcal{H}(m), m) - g - \kappa m, 1 - \mathcal{H}(m), q(m)). \tag{3.4}$$

Taking the derivative of (3.4) w.r.t. m gives the following optimality condition:

$$\underbrace{u_c(F_m - \kappa) + q_m[u_s + u_c F_s]}_{\omega_{Env}} + \mathcal{H}_m \underbrace{(u_c F_h - u_l)}_{\omega_{LL}} = 0. \tag{3.5}$$

This shows that, in equilibrium, the government trades off wedges. In first best, since both wedges are zero, (3.5) holds. In second best where $\tau > 0$, I have that $\omega_{LL} > 0$, and hence the environmental wedge cannot be zero unless $\mathcal{H}_m = 0$. More precisely, if $\mathcal{H}_m > 0$ ($\mathcal{H}_m < 0$), ω_{Env} must be negative (positive) for (3.5) to be satisfied.

Expression (3.5) illustrates that there is an interaction between climate policy and the non-environmental tax, in the sense that in the presence of a distortionary tax on labor income, environmental quality is not provided at the first-best margin. Moreover, whether or not the climate externality is less than fully internalized, i.e. whether or not $\omega_{Env} < 0$, depends on the sign of $\mathcal{H}_m$.

What is the intuition behind (3.5)? For the sake of the argument, assume that $\mathcal{H}_m > 0$. That is, a marginal increase in fuel use and thus, carbon emissions, raises the labor supply. In first best, where $\omega_{LL} = 0$, a marginal change in hours worked does not affect welfare. In contrast, in second best it leads to a first-order welfare gain since in equilibrium, the benefit from a marginal increase in hours worked, $u_c f_h$, is greater than the marginal disutility of working more (u_l). This is only if $\omega_{LL} > 0$, i.e. as long as the income tax is positive. In this sense, fuel use mitigates the intratemporal distortion caused by the income tax. The positive effect on labor supply is an additional benefit of fuel use, apart from the usual benefit of increasing output and consumption. Thus, it reduces the social cost of carbon below the value of the marginal damage - or, in terms of tax rates, the optimal second-best tax is lower than the corresponding Pigouvian fee.[23] In the remainder of this paper, I will refer to this effect on the labor supply, represented by $\mathcal{H}_m \omega_{LL}$, as the "static labor effect".

In order to make a statement on in which direction this effect goes, one must determine the sign of $\mathcal{H}_m$. In the appendix, I show that the general equilibrium effect is given by:

$$\mathcal{H}_m = A^{-1}(\epsilon_{h,w} + \epsilon_{h,y}) \left(\frac{\upsilon_h}{\upsilon_k + \upsilon_h} (F_m - \kappa + F_s q_m) \right),$$

where $A > 0$. The second term is the sum of the uncompensated wage elasticity of labor supply, $\epsilon_{h,w}$, and the elasticity of labor with respect to non-labor income, $\epsilon_{h,y}$. Note that increasing fuel use has both a substitution effect - by affecting the net wage - and an income effect, through

[23]Naturally, the same logic applies if $\mathcal{H}_m < 0$. In that case, there is an additional second-best cost of using fuel, since it decreases labor supply and hence exacerbates the intratemporal distortion.

both the net wage and non-labor income. If $\epsilon_{h,w}+\epsilon_{h,y} > 0$ ($\epsilon_{h,w}+\epsilon_{h,y} < 0$), the former dominates (is dominated by) the latter.

Finally, υ_h and υ_k denote the income share of labor and capital, respectively. Then, as shown in the appendix, the third term captures the change both in the net wage and the non-labor income when increasing fuel use, i.e. $\partial(1-\tau)F_h/\partial m$ and $\partial(1-\tau)F_k k/\partial m$, assuming that $\upsilon_h/(\upsilon_k+\upsilon_h)$ is constant.

Note that in the special case when there are no direct utility damages in the model, I have

$$\omega_{Env} = F_m - \kappa + F_s q_m, \tag{3.6}$$

and hence a zero environmental wedge implies that $\mathcal{H}_m = 0$, which is consistent with (3.5). In other words, as noted by Bovenberg and Goulder (2002), the climate externality is in general not fully internalized only if pollution has a negative effect on utility.[24] Moreover, this term captures the interplay between the tax-interaction effect and the revenue-recycling effect discussed above. If it is positive, a decrease in the carbon tax – and hence a decrease in green revenue – leads to a higher net wage, implying that the tax-interaction effect dominates the revenue-recycling effect.

3.1.2 A Two-Period Model With Capital

Let $\bar{t} = 1$, so the representative household lives for two periods and can postpone consumption by saving into an asset, referred to as capital. The government can impose a tax on capital income.

Crucially, I assume that the government in period 0 does not have access to a commitment device which would allow it to commit to future tax rates. From this, it follows that even in the case of separate tax rates on labor and capital income in period 1, the capital tax will be non-zero.[25] Hence, there is an intertemporal distortion, implying that the "consumption-savings wedge", i.e. the distortion of the consumption-savings margin, is positive:[26]

$$\omega_{CS} \equiv -u_c + \beta u_c'[F_k' + 1 - \delta] > 0. \tag{3.7}$$

Under lack of commitment, the government's problem is solved using backwards induction. In the appendix, I show that this results in a linear combination of wedges, similar to (3.5) in the static setting with labor taxation:

$$\omega_{Env} + \mathcal{H}_m \omega_{LL} + \mathcal{K}_m \omega_{CS} = 0. \tag{3.8}$$

[24] Intuitively, in the case of contemporaneous productivity damages only, a marginal increase in fuel use from the first-best level leaves output – and hence the wage – unchanged. This is not the case if there are direct damages to utility.

[25] In the case of separate tax rates, the capital tax acts as a de-facto lump-sum tax since the tax base is inelastic. Hence, the labor tax is zero, while government expenditures are solely financed through capital taxation. If I instead allowed for commitment, the capital tax rate may be zero.

[26] Throughout this paper, primes denote future variables. For example, $c' = c_{t+1}$, $u_c' = u_c(t+1)$, $c'' = c_{t+2}$, $u_c'' = u_c(t+2)$ etc.

Here, the function $\mathcal{K}(m)$ is defined analogous to $\mathcal{H}(m)$, but with respect to current savings. That is, it gives the best-response function for households' savings in period 0 given an emission level m. Note that (3.8) contains the derivatives of the policy functions for savings and labor, and hence, following Klein et al. (2008), I refer to this as a *generalized* Euler equation. This equation shows that, as in the one-period model, optimal environmental policy interacts with fiscal policy. This implies that as long as current savings and labor supply are affected by current fuel use, $\mathcal{K}_m \neq 0$ and $\mathcal{H}_m \neq 0$, emissions are not in general at the Pigouvian level.

The last term on the left-hand side of (3.8), $\mathcal{K}_m \omega_{CS}$, reflects what I will call the "savings effect". To understand the mechanism, abstract from labor taxation, and hence let $\omega_{LL} = 0$. Then, the sign of the environmental wedge depends on the sign of $\mathcal{K}_m$: if current savings increase (decrease) in current fuel use, the wedge is negative (positive), i.e., the carbon tax is below (above) its Pigouvian level. The intuition is similar to the static case above. First, note that since households understand that capital will be taxed in the following period and thus their return to savings will be lower, they consume more and save less than in first best.

Then, if $\mathcal{K}_m > 0$, by increasing fossil fuel use and thus "underproviding" the environmental public good, i.e. by not fully internalizing climate damages, the first-period government can increase current savings. This has a first-order welfare gain in second best since $\omega_{CS} > 0$, and hence the discounted marginal increase in utility due to more consumption in the subsequent period is higher than the marginal utility loss due to less consumption today. It follows that using fossil fuel has an additional benefit which is not present in first best, and hence the social cost of carbon does not fully internalize the utility damage caused by the climate externality. In other words, the margin between private consumption and environmental quality is distorted compared to the first best.

In the case with labor taxation, fossil fuel use affected labor supply through a change in the return to labor. Here, the mechanism is slightly different. The return to current savings is determined by next period's fuel use, which is not affected directly by the current government. Instead, more fuel use today affects the amount of resources available to the household by increasing today's capital income, i.e. the return to past savings. Note that this result is analogous to Klein et al. (2008) who only consider not-environmental public consumption. In general, underproviding a public good today dampens underinvestment and thus mitigates the intertemporal distortion caused by the positive tax on capital income.

Returning to the more general case with both $\omega_{LL} > 0$ and $\omega_{CS} > 0$, whether the optimal carbon tax is below or above the Pigouvian rate depends on the sign of $\mathcal{H}_m \omega_{LL} + \mathcal{K}_m \omega_{CS}$. This is in general ambiguous. Note that under particular assumptions - logarithmic utility, a Cobb-Douglas production structure, endogenous government expenditures and a flow pollutant - one can solve for ω_{Env} analytically. In this case, $\mathcal{H}_m = 0$, while $\mathcal{K}_m > 0$, and hence $\omega_{Env} < 0$. Under more general assumptions, numerical simulations indicate that this results still holds for reasonable parameter values.

To summarize, this section has illustrated that in finite-horizon settings without commitment, carbon taxation interacts with fiscal policy, in the sense that a distortion of the non-environmental margins affects the environmental margin. This possibly leads to an optimal pollution tax which is below the Pigouvian level, and hence the climate externality is not fully internalized.

In the following sections, I will turn to infinite-horizon models. Increasing the number of periods gives rise to additional dynamic effects, i.e. adds more terms to the equation (3.8). Each of them is linked to a distorted margin in the future, and will reflect how a change in current fuel use affects these future wedges in addition to the contemporaneous wedges that are at the root of the savings and the static labor effect.

3.2 Infinite Horizon

I now consider an infinite-horizon version of the simple global climate-economy model outlined above.[27] I derive the government's generalized Euler equations and show that the result obtained in the previous section carries over to the longer horizon: in dynamic models without commitment, climate policy interacts with distortionary taxation and the optimal carbon tax is in general not at the Pigouvian level. In addition to the second-best effects of emissions present in the two-period model above, I identify further benefits and costs from current fuel use.

3.2.1 First Best

In the presence of lump-sum taxation, the first-best equlibrium is characterized by the following set of equations:

$$\omega_{CS} \equiv u_c - \beta u_c'[F_k' + 1 - \delta] = 0 \tag{3.9}$$

$$\omega_{LL} \equiv u_l - u_c F_h = 0 \tag{3.10}$$

$$\omega_{PG} \equiv u_g - u_c = 0 \tag{3.11}$$

$$\omega_{Env} \equiv u_c(F_m - \kappa) + \beta q_m \left[u_T' + u_c' F_T' - \frac{q_T'}{q_m'} u_c'(F_m' - \kappa) \right] = 0. \tag{3.12}$$

As before, I define wedges for the consumption-savings margin (ω_{CS}), the labor-leisure margin (ω_{LL}), and the environmental margin (ω_{Env}). In addition, since public consumption is endogenous, there is a wedge for the public-private good margin, denoted by ω_{PG}. The first-best equilibrium is characterized by all these wedges being simultaneously zero.

Note that (3.12) is a compact way of writing the first-best environmental margin. Since

$$u_c'(F_m' - \kappa) = -\beta q_m' \left[u_T'' + u_c'' F_T'' - \frac{q_T''}{q_m''} u_c''(F_m'' - \kappa) \right],$$

[27] For simplicity, I derive the analytical results in an infinite-horizon setting where fossil fuel never stops. In the calibrated model below, fossil fuel stops after a finite number of periods.

(3.12) can be rewritten as

$$\begin{aligned}
\omega_{Env} &= u_c(F_m - \kappa) + \beta q_m \left[u'_T + u'_c F'_T + \frac{q'_T}{q'_m} \left(\beta q'_m \left[u''_T + u''_c F''_T - \frac{q''_T}{q''_m} u''_c (F''_m - \kappa) \right] \right) \right] \\
&= u_c(F_m - \kappa) + \beta q_m \left[u'_T + u'_c F'_T + q'_T \beta \left(u''_T + u''_c F''_T - \frac{q''_T}{q''_m} u''_c (F''_m - \kappa) \right) \right] \\
&= u_c(t)(F_m(t) - \kappa) + q_m(t) \sum_{i=t+1}^{\infty} \beta^{i-t} \left(\prod_{j=t+1}^{i-1} q_T(j) \right) [u_T(i) + u_c(i) F_T(i)].
\end{aligned}$$

The first term captures the marginal benefit of carbon emissions (in utils), the second term the discounted marginal utility and productivity damages in the future. From this, it follows that the Pigouvian tax is defined as:

$$\theta_t^p = -\frac{q_m(t)}{u_c(t)} \sum_{i=t+1}^{\infty} \beta^{i-t} \left(\prod_{j=t+1}^{i-1} q_T(j) \right) (u_T(i) + u_c(i) F_T(i)). \tag{3.13}$$

3.2.2 Second-best without Commitment

Next, I relax the assumption of lump-sum taxation being feasible. Instead, the government has to resort to a distortionary tax on labor and capital income to finance a public consumption good. Once more, I assume that it does not have access to a technology that allows it to commit to all future tax rates. When solving for the outcome under lack of commitment, I look for the time-consistent differentiable Markov-perfect equilibrium in this economy.[28] The basic idea of this equilibrium concept is that only current payoff-relevant states, but not the history of states and actions, matter for a player's action choice.

Note that in this setting, the current goverment plays a game with its successor.[29] This implies that the current government takes into account the optimal behavior of next period's government when solving its problem. While the current government cannot directly choose policies in the following period, it can affect them indirectly by choosing the economy's state variables.

I define the Markov-perfect equilibrium in a setting where the government has access to a total income tax. That is, it is restricted to impose the same tax rate on labor and capital income. This assumption is necessary since I want to consider a setting in which both the intertemporal consumption-savings margin and the intratemporal labor-leisure margin are distorted simultaneously, i.e. where I have positive tax rates on both labor and capital income. As shown by Martin (2010), in a setting where the government has ex-ante access to two separate

[28]An alternative approach would be to look for all sustainable equilibria, along the lines of Phelan and Stacchetti (2001) or Reis (2011).

[29]If governments in different periods are identical, the current government actually plays a game "against itself". That is, even though I have the same government making the decisions in both periods, it must be treated as different players, due to the lack of commitment. Equivalently, announcements in the current period about how the government will behave in the following period are not credible.

tax rates, the equilibrium features a zero tax rate on labor income, assuming that labor taxes are bounded to be non-negative and that the capital income tax is not bounded from above. This is intuitive: recall that taxes on capital are ex post non-distortionary and thus, can be considered as de facto lump-sum taxes. Hence, in the presence of such a tax, assuming that it is unbounded, it cannot be optimal to have a positive distortionary tax on labor income.[30] Empirically, a setting with a total income tax appears more realistic than an equilibrium in which the government finances its expenditures using only a tax on capital or on labor income.[31]

The analysis is a straightforward extension of Klein et al. (2008) and Azzimonti et al. (2009), adding a second public good, environmental quality, which in contrast to the other good does not only affect utility, but also the production process. Hence, as in Azzimonti et al. (2009), there is a second state variable in addition to capital, here the current state of the climate T.

Moreover, by comparing the Markov-perfect carbon tax to the outcome under commitment, I will show that in the presence of distortionary taxes, the optimal pollution price is in general not time-consistent. That is, the tax schedule set by a government which had access to a commitment device is different than the one chosen under lack of commitment. Note that this time inconsistency is due to the interaction between environmental and non-environmental taxes: as seen above, the optimal pollution price depends on the optimal tax structure. If other taxes are time-inconsistent - for example a positive tax on labor income in a scenario where separate tax rates on labor and capital are feasible - so is the environmental tax.

3.2.3 Equilibrium Definition

A stationary Markov-perfect equlibrium is defined as a value function v, differentiable policy functions ψ and ϕ, a savings function n^k and a labor-supply function n^h such that for all k and T, $\psi(k,T)$, $\phi(k,T)$, $n^k(k,T)$ and $n^h(k,T)$ solve

$$\max_{k',T',h,g,m} u(\mathcal{C}(k,k',h,g,m,T), 1-h, g, T) + \beta v(k',T'), \tag{3.14}$$

subject to

$$\begin{aligned} & u_c\big(\mathcal{C}(k,k',h,g,m,T), 1-h, g, T\big) \\ & - \beta u_c' \left(\mathcal{C}\left(\begin{array}{c} k', n^k(k',T'), n^h(k',T'), \\ \psi(k',T'), \phi(k',T'), T' \end{array} \right), 1 - n^h(k',T'), \psi(k',T'), T' \right) \cdot \\ & \cdot \left[1 + \left[1 - \mathcal{T}\left(\begin{array}{c} k', n^h(k',T'), \psi(k',T'), \\ \phi(k',T'), T' \end{array} \right) \right] \left[F_k(k', n^h(k',T'), \phi(k',T'), T') - \delta \right] \right] = 0, \end{aligned} \tag{3.15}$$

[30] In other words, the government is not allowed to subsidize labor. Martin (2010) shows that subsidizing labor is optimal in a setting with unrestricted separate tax rates on labor and capital. However, this appears to be less relevant empirically than zero labor taxes.

[31] There are other ways of modifying the model such that one would get positive tax rates on both production factors in equilibrium. Martin (2010) considers an exogenous upper bound on the capital tax, as well as making the utilization rate of capital endogenous. The former is somewhat unsatisfying since it leaves the origin of the bound unmodeled. The latter slightly changes the logic of the mechanism.

$$\frac{u_l\big(\mathcal{C}(k,k',h,g,m,T),1-h,g,T\big)}{u_c\big(\mathcal{C}(k,k',h,g,m,T),1-h,g,T\big)} - F_h(k,h,m,T)[1-\mathcal{T}(k,h,g,m,T)] = 0, \tag{3.16}$$

and $T' = q(T,m)$. $\mathcal{C}$ yields private consumption from the resource constraint, while $\mathcal{T}$ gives the income tax rate that balances the government's budget:

$$\mathcal{C}(k,k',h,g,m,T) = F(k,h,m,T) + (1-\delta)k - g - \kappa m - k', \tag{3.17}$$

$$\mathcal{T}(k,h,g,m,T) = \frac{g-(F_m(k,h,m,T)-\kappa)m}{(F_k(k,h,m,T)-\delta)k + F_h(k,h,m,T)h}. \tag{3.18}$$

Moreover, for all k and T,

$$\begin{aligned} v(k,T) =& u[\mathcal{C}(k,n^k(k,T),n^h(k,T),\psi(k,T),\phi(k,T),T),1-n^h(k,T),\psi(k,T),T)] \\ &+ \beta v(n^k(k,T),q(T,\phi(k,T))). \end{aligned} \tag{3.19}$$

As outlined above, the current government plays a game against its successor, possibly itself. Following the one-stage deviation principle, the current government's strategy constitutes an equilibrium if it maximizes its objective function, subject to all relevant constraints, taking the strategies of the other player - the future government - as given. In other words, assuming that the future government chooses policies according to the equilibrium decision rules, it must be optimal for the current government to follow the same policy functions.

3.2.4 Solution

When solving the model, I follow Klein et al. (2008). Denote the left-hand side of (3.15) as $\eta(k,T,g,m,k',h)$ and the left-hand side of (3.16) as $\epsilon(k,T,g,m,k',h)$, respectively. Define the functions $\mathcal{K}(k,T,g,m)$ and $\mathcal{H}(k,T,g,m)$ implicitly as

$$\eta(k,T,g,m,\mathcal{K}(k,T,g,m),\mathcal{H}(k,T,g,m)) = 0 \tag{3.20}$$

$$\epsilon(k,T,g,m,\mathcal{K}(k,T,g,m),\mathcal{H}(k,T,g,m)) = 0. \tag{3.21}$$

As in the finite-horizon models above, $\mathcal{K}$ ($\mathcal{H}$) can be interpreted as the household's response function for savings (hours worked) to the current government's policy choice, assuming that future governments follow the equilibrium policies: it gives the household's optimal savings level if the current governments set expenditures g and a carbon tax that results in emission level m. In equilibrium, $\mathcal{K}(k,T,\psi(k,T),\phi(k,T)) = n^k(k,T)$ and $\mathcal{H}(k,T,\psi(k,T),\phi(k,T)) = n^h(k,T)$.

Solving the government's problem, taking $\mathcal{K}$ and $\mathcal{H}$ as given, results in the following system of optimality conditions that characterize the stationary Markov-perfect equilibrium:[32]

$$u_c - \beta u_c'[1 + (1 - \mathcal{T}(k', T', h', g', m'))(F_k' - \delta)] = 0 \tag{3.22}$$

$$u_l - u_c(1 - \mathcal{T}(k, T, h, g, m))F_h = 0 \tag{3.23}$$

$$\begin{aligned} &u_g - u_c + \mathcal{H}_g(u_c F_h - u_l) + \mathcal{K}_g[-u_c + \beta u_c'(F_k' + 1 - \delta)] + \\ &+\beta \mathcal{K}_g \left\{ \mathcal{H}_k'(F_h' u_c' - u_l') - \frac{\mathcal{K}_k'}{\mathcal{K}_g'} \left[\mathcal{H}_g'(F_h' u_c' - u_l') + u_g' - u_c' \right] \right\} = 0. \end{aligned} \tag{3.24}$$

$$\begin{aligned} &u_c(F_m - \kappa) + \beta q_m \left[u_T' + u_c' F_T' - \frac{q_T'}{q_m'} u_c'(F_m' - \kappa) \right] + \\ &+\mathcal{K}_m[-u_c + \beta u_c'(F_k' + 1 - \delta)] + \beta(u_g' - u_c') \left[-\mathcal{K}_m \frac{\mathcal{K}_k'}{\mathcal{K}_g'} - q_m \frac{\mathcal{K}_T'}{\mathcal{K}_g'} + q_m \frac{q_T'}{q_m'} \frac{\mathcal{K}_m'}{\mathcal{K}_g'} \right] + \\ &+\mathcal{H}_m(u_c F_h - u_l) + \beta \mathcal{K}_m(u_c' F_h' - u_l') \left[\mathcal{H}_k' - \frac{\mathcal{K}_k'}{\mathcal{K}_g'} \mathcal{H}_g' \right] + \\ &+\beta(u_c' F_h' - u_l') q_m \left[\mathcal{H}_T' - \frac{\mathcal{K}_T'}{\mathcal{K}_g'} \mathcal{H}_g' - \frac{q_T'}{q_m'} \mathcal{H}_m' + \frac{q_T'}{q_m'} \frac{\mathcal{K}_m'}{\mathcal{K}_g'} \mathcal{H}_g' \right] = 0 \end{aligned} \tag{3.25}$$

where the latter two equations are once more generalized Euler equations. Using the wedges defined in (3.9)-(3.12) above, I can write (3.24) and (3.25) as a linear combination of wedges:

$$\omega_{PG} + \mathcal{K}_g \omega_{CS} - \beta \mathcal{K}_g \frac{\mathcal{K}_k'}{\mathcal{K}_g'} \omega_{PG}' + \mathcal{H}_g \omega_{LL} + \beta \omega_{LL}' \mathcal{K}_g \left[\mathcal{H}_k' - \frac{\mathcal{K}_k'}{\mathcal{K}_g'} \mathcal{H}_g' \right] = 0, \tag{3.26}$$

$$\begin{aligned} &\omega_{Env} + \mathcal{K}_m \omega_{CS} - \beta \omega_{PG}' \mathcal{K}_m \frac{\mathcal{K}_k'}{\mathcal{K}_g'} - \beta \omega_{PG}' \left[q_m \frac{\mathcal{K}_T'}{\mathcal{K}_g'} - q_m \frac{q_T'}{q_m'} \frac{\mathcal{K}_m'}{\mathcal{K}_g'} \right] \\ &+ \mathcal{H}_m \omega_{LL} + \beta \omega_{LL}' \mathcal{K}_m \left[\mathcal{H}_k' - \frac{\mathcal{K}_k'}{\mathcal{K}_g'} \mathcal{H}_g' \right] \\ &+ \beta \omega_{LL}' q_m \left[\mathcal{H}_T' - \frac{\mathcal{K}_T'}{\mathcal{K}_g'} \mathcal{H}_g' - \frac{q_T'}{q_m'} \mathcal{H}_m' + \frac{q_T'}{q_m'} \frac{\mathcal{K}_m'}{\mathcal{K}_g'} \mathcal{H}_g' \right] = 0. \end{aligned} \tag{3.27}$$

As before, the government trades off wedges in equilibrium. In first best, as $\omega_{CS} = \omega_{LL} = \omega_{PG} = \omega_{env} = 0$, (3.26) and (3.27) are satisfied. If the government has to resort to distortionary taxation, the household's optimality conditions (3.22) and (3.23) imply that if $\mathcal{T}(\cdot) > 0$, the consumption-savings and labor-leisure wedges are positive. Then, assuming that the derivatives of the best-response functions are non-zero, it follows from (3.26) and (3.27) that ω_{PG} and ω_{env} cannot be zero at the same time: in optimum, neither public good is provided at the first-best margin. Recall that for the environmental public good, the result that $\omega_{env} \neq 0$ just says that the social cost of carbon is not equal to the marginal climate damage.

Equations (3.26) and (3.27) are useful in computing stationary Markov-perfect equilibria.

[32] Compare the appendix, section A.3, for details. Note that deriving the generalized Euler equations for the less general case with a capital income tax only is completely analogous, but with $\omega_{LL} = \omega_{LL}' = 0$.

To facilitate interpretation, I show in the appendix that they can be rewritten as:[33]

$$
\begin{aligned}
u_c(t)(F_m(t)-\kappa) - \bar{q}_m \sum_{i=t+1}^{\infty} \beta^{i-t} \left(\prod_{j=t+2}^{i} q_T(j) \right) [-u_T(i) - u_c(i)F_T(i)] + \\
+\mathcal{K}_m(t)\omega_{CS}(t) - \beta\omega_{PG}(t+1)\mathcal{K}_m(t)\frac{\mathcal{K}_k(t+1)}{\mathcal{K}_g(t+1)} + \\
+\mathcal{H}_m(t)\omega_{LL}(t) + \beta\omega_{LL}(t+1)\mathcal{K}_m(t)\left[\mathcal{H}_k(t+1) - \frac{\mathcal{K}_k(t+1)}{\mathcal{K}_g(t+1)}\mathcal{H}_g(t+1)\right] + \\
+\bar{q}_m \sum_{i=t+1}^{\infty} \beta^{i-t} \left(\prod_{j=t+2}^{i} q_T(j) \right) \left[\omega_{LL}(i)\left(\mathcal{H}_T(i) - \frac{\mathcal{K}_T(i)}{\mathcal{K}_g(i)}\mathcal{H}_g(i)\right) - \omega_{PG}(i)\frac{\mathcal{K}_T(i)}{\mathcal{K}_g(i)}\right] = 0,
\end{aligned}
\tag{3.28}
$$

where I have assumed that $q_m(t) = q_m(t+1) = q_m(t+2) = ... = \bar{q}_m$.[34]

This expression, together with (3.27), represents the main theoretical result in this paper. Note that the first term in (3.28) is the marginal benefit of pollution (in utils), while the second term is the present value of the sum of future marginal damages when increasing current emissions, or more concisely, the current marginal pollution damage. Denote the difference between the latter and the former, that is, the tax-interaction effect, by $Q(t)$:

$$
-\bar{q}_m \sum_{i=t+1}^{\infty} \beta^{i-t} \left(\prod_{j=t+2}^{i} q_T(j) \right) [u_T(i) + u_c(i)F_T(i)] - u_c(t)(F_m(t)-\kappa) \equiv Q(t).
$$

For further interpretation, divide both sides of this expression by the current marginal utility of consumption, $u_c(t)$. On the left-hand side, this gives the difference between the Pigouvian carbon fee, θ^p, and the social cost of carbon (in dollar terms) and hence the optimal carbon tax. Denote this by Δ:

$$
\Delta_t = \theta_t - \theta_t^p = -\frac{Q(t)}{u_c(t)}. \tag{3.29}
$$

In first best, $Q(t) = 0$, and hence $\theta = \theta^p$, as seen above. In second best, (3.28) shows that

[33]Note that in the special case of a flow pollutant (and contemporaneous damages), (3.27) directly characterizes the tax-interaction effect. With $q_m = 0$ in every period, (3.27) can be rewritten as

$$
\omega_{Env} + \mathcal{K}_m\omega_{CS} - \beta\omega'_{PG}\mathcal{K}_m\frac{\mathcal{K}'_k}{\mathcal{K}'_g} + \mathcal{H}_m\omega_{LL} + \beta\omega'_{LL}\mathcal{K}_m\left[\mathcal{H}'_k - \frac{\mathcal{K}'_k}{\mathcal{K}'_g}\mathcal{H}'_g\right] = 0,
$$

where ω_{Env} is the difference between the social cost of emissions and the marginal damage.

[34]In words, an additional unit of emissions affects the pollutant stock in the same way, independently of when it occurs. This is a standard assumption in many climate models (Nordhaus, 2008; Golosov et al., 2013).

the tax-interaction effect is given by

$$
\begin{aligned}
Q(t) = &\underbrace{\mathcal{H}_m(t)\omega_{LL}(t)}_{Q_1} + \underbrace{\mathcal{K}_m(t)\omega_{CS}(t)}_{Q_2} \\
&+ \underbrace{\beta\omega_{LL}(t+1)\mathcal{K}_m(t)\left[\mathcal{H}_k(t+1) - \frac{\mathcal{K}_k(t+1)}{\mathcal{K}_g(t+1)}\mathcal{H}_g(t+1)\right]}_{Q_3} \\
&+ \underbrace{\left(-\beta\omega_{PG}(t+1)\mathcal{K}_m(t)\frac{\mathcal{K}_k(t+1)}{\mathcal{K}_g(t+1)}\right)}_{Q_4} \\
&+ \underbrace{\phi\sum_{i=t+1}^{\infty}\beta^{i-t}\left(\prod_{j=t+1}^{i} q_T(j)\right)\omega_{LL}(i)\left[\mathcal{H}_T(i) - \frac{\mathcal{K}_T(i)}{\mathcal{K}_g(i)}\mathcal{H}_g(i)\right]}_{Q_5} \\
&+ \underbrace{\phi\sum_{i=t+1}^{\infty}\beta^{i-t}\left(\prod_{j=t+1}^{i} q_T(j)\right)\left[-\omega_{PG}(i)\frac{\mathcal{K}_T(i)}{\mathcal{K}_g(i)}\right]}_{Q_6} \\
&= \sum_{n=1}^{6} Q_n(t).
\end{aligned} \tag{3.30}
$$

The questions of whether there is a tax-interaction effect and in what direction it goes can then be asked in terms of $Q(t)$: is it different from zero, and if so, what is the sign? Moreover, to what extent do the different components of $Q(t)$ move in the same direction or cancel each other out?

At this point, note that the wedges and derivatives that show up in (3.30) are endogenous and solved for simultaneously when computing the equilibrium. Moreover, for most of these terms, it is not possible to determine the sign analytically. Hence, the questions just asked are ultimately quantitative questions and hence require a numerical analysis. This is done in the next subsection. Before solving for the terms Q_1 through Q_6, I proceed with a qualitative interpretation of the different terms.

3.2.5 Interpretation

The first component of $Q(t)$, $\mathcal{H}_m\omega_{LL}$, captures the static labor effect of current fuel use on labor supply, described in section 3.1.1. Recall from above that if $\mathcal{H}_m > 0$, an increase in current fuel use has a positive impact on current labor supply, which mitigates the intratemporal distortion, and hence has non-zero impacts on utility in second best.

Similarly, the second component, $\mathcal{K}_m\omega_{CS}$, is the savings effect that was present in the two-period model. For $\mathcal{K}_m > 0$, a higher fuel use today increases the resources available for the household and allows it to move more resources to the next period, thereby mitigating the intertemporal distortion. With a slight abuse of terminology, call Q_1 and Q_2 the "contemporaneous second-best effects" of carbon emissions.

The remaining components of Q did not show up in the one- and two-period models considered above. Q_3 reflects the second-best effect of an increase in current fuel use on tomorrow's labor-leisure wedge through a change in the future capital stock. A higher k' affects future labor supply in two ways. First, there is a direct effect, captured by the term $\mathcal{H}'_k$: a higher capital stock increases tomorrow's real wage, which in turn affects the amount of hours worked.

To understand the second channel, note that a change in current savings affects future savings. For example, assuming that $\mathcal{K}'_k > 0$, a higher capital stock tomorrow leads to more savings. This, in turn, dampens the intertemporal distortion *in the future* and hence reduces the incentive for the future government to "underprovide" public goods in order to increase savings. This may imply a higher income tax (to finance the non-environmental good) and a higher emission fee (to increase the provision of the environmental public good), which will affect labor supply. This channel is captured by the term $\frac{\mathcal{K}'_k}{\mathcal{K}'_g}\mathcal{H}'_g$

The same logic applies to Q_4. This component reflects the direct welfare effect of a change in future public goods provision due to a change in the capital stock. For example, assume that the public consumption good tomorrow is underprovided, and that a change in current fuel use increases tomorrow's capital stock, which has a positive effect on tomorrow's savings and public goods provision. Formally, this implies that $-\mathcal{K}_m\frac{\mathcal{K}'_k}{\mathcal{K}'_g} > 0$. Then, by underproviding public goods today, the current government is able to dampen the desire of the successive government to underprovide public goods tomorrow, by increasing savings in the future and thereby reducing the need to provide additional resources in the subsequent period. This would be an additional benefit of current fuel use.

Note that these channels were also present in the model by Klein et al. (2008), and hence they do not depend on the public good being persistent. Therefore, I will refer to Q_3 and Q_4 as the "second-best flow effects".

The two remaining terms in (3.30) reflect the effect of a change in current fuel use on the future pollutant stock and thus, on the labor-leisure and public-good margins in all future periods, due to the persistence of the pollutant. More specifically, Q_5 captures the impact of current fuel use on future labor supply decisions. Once more, this component can be disentangled in a direct effect of a change in s on hours worked ($\mathcal{H}'_T$), and its effect on future savings and hence public goods provision and taxation ($-\frac{\mathcal{K}'_T}{\mathcal{K}'_g}\mathcal{H}'_g$), analogous to a change in the capital stock. For example, through its impact on productivity, a higher pollutant stock has a negative impact on output and thus wages and household income and hence on future labor supply and savings.

The latter channel is also reflected in Q_6, which captures the effect of current fuel use on all future distortions to the public-good margins. Both Q_5 and Q_6 depend on the decay rate of the pollutant, given by $1 - \phi$, being less than one. Hence, these terms will be referred to as "second-best stock effects" of current emissions.

3.3 Quantitative Analysis

Above, I have characterized the tax-interaction effect in a dynamic infinite-horizon model without commitment. I have identified the second-best benefits and costs of emissions that affect the social cost of carbon. However, this section has given no indication about how important the tax-interaction effect is. As argued above, this is a quantitative question. Therefore, I now compute the optimal Markov-perfect carbon tax in a calibrated climate-economy model and compare it to both the tax in first best and to the Pigouvian rate. Moreover, for comparison, I also compute the outcome if the government had access to a commitment device, and hence could credibly commit to all future tax rates.[35]

3.3.1 Calibration

I calibrate the model using 2011 as the base year. That is, I assume that in 2011, there is no climate policy, and the world economy is on a balanced growth path (BGP), featuring a total income tax. I solve the global model in second best without any climate damages, and calibrate the labor-augmenting productivity level A_{2011} such that annual output equals world GDP in 2011, which amounted to 70 trillion US\$. Moreover, I calibrate the energy efficiency level B_{2011} such that emissions equal around 9 GtC (Olivier et al., 2013). Recall that labor- and energy-augmenting productivity are assumed to grow at the same rate. I set the annual growth rate to be 1.8% ($\zeta = 1.018$). Note that one period in the model equals two years.

When using a Cobb-Douglas production function, a standard choice for the income share of capital, θ, is between 0.3 and 0.36. I set $\rho = 0.35$, to get a capital-output ratio of around 3 along the business-as-usual BGP. As for the income share of fossil fuel, I follow Golosov et al. (2013) and set $\xi = 0.03$.[36] Note that without climate policy, the marginal product of fuel is equal to the private extraction cost, and hence:

$$\kappa_{2011} = \frac{\xi y_{2011}}{m_{2011}} = 0.03\frac{70 \cdot 10^{12}}{8.5 \cdot 10^{9}} = 247\$/tC. \tag{3.31}$$

This value is between what is found by Golosov et al. (2013) for the private cost of using oil and coal, respectively.

With respect to the utility function, I choose the parameter values for α_c and α_g to target hours worked and the ratio of government expenditures to output. Following Klein et al. (2008), the latter should be around 0.2. The labor supply is typically assumed to be between 0.22 and 0.25 (Klein et al., 2008; Barrage, 2013).

Regarding the calibration of parameters γ and α_s, Barrage (2013) finds that production damages account for about 74% of total output damages if $T = 2.5°C$. From this, she calibrates $\gamma = 0.00172$. I fix this value and then choose α_s to target a first-best temperature increase of $3°C$ in the baseline setting. In other words, in the long run, the global mean temperature

[35] Compare the appendix, section A.5, for how to compute the second-best equilibrium with commitment.

[36] Other studies use a higher income share. For example, Fischer and Springborn (2011) choose $\xi = 0.09$.

stabilizes at $3°C$ above the pre-industrial level. I conduct some sensitivity analysis by varying the target value for temperature and check for the robustness of the results.

The following table lists the parameter settings for the baseline calibration.

Parameter	Description	Value	Source
Preferences			
α_c	utility weight consumption	.255	Calibration
α_g	utility weight public good	.12	Calibration
α_s	utility weight climate	.0063	Calibration, sensitivity
Production			
ρ	income share capital	.35	
ξ	income share fuel	.03	Hassler et al. (2012)
Other			
δ	depreciation rate	.08	Klein et al. (2008)
β	time discount rate	.985	Nordhaus (2010)
γ	damage function	.00172	Barrage (2013)

Table 1: Baseline Calibration

3.3.2 Solution Method

In the analytical analysis above, for simplicity, I have characterized optimal income and carbon taxation in a stationary economy. When quantitatively analyzing climate change, however, stationarity is a problematic assumption if it implies that fossil fuel is used forever. Therefore, in this section, I turn to a non-stationary framework. Nevertheless, the key equation derived above, (3.30), can still be used to quantify the tax-interaction effect and its individual components.

I follow a common modeling strategy (Golosov et al., 2013; Barrage, 2013; Cai et al., 2012) and assume that in the long run, energy use does no longer cause carbon emissions. More specifically, I impose a deterministic number of J periods, after which the one-to-one mapping between the amount of fuel used and the amount of carbon emitted is no longer valid.[37] In the baseline model, I let $J = 120$. From this period onwards, T is constant, and the economy converges to a balanced growth path (BGP). Let $\bar{v}$ denote the terminal value function (Cai et al., 2012):

$$\bar{v}(k_{J+1}, T_{J+1}) = \sum_{t=J+1}^{\infty} \beta^{t-J+1} u(c_t, 1-h_t, g_t, T_t). \tag{3.32}$$

This function gives the continuation value for the stationary phase of the economy.

Numerically, both the first-best and the second-best equilibrium with commitment are straightforward to solve for, by maximizing over all variables for J periods, given the terminal value function. This is the approach employed by Golosov et al. (2013) or Barrage

[37] This could be interpreted in different ways. For example, at a given point in time, "clean" energy, i.e. a perfect substitute for fossil fuel whose production and use does not cause carbon emissions, becomes available. Alternatively, one could assume the employment of technologies like Carbon Capture and Storage (CCS).

(2013), among others. Under lack of commitment, this method is not feasible. Instead, I have to resort to dynamic programming, following Cai et al. (2012, 2013). The problem is formally described in the appendix in section A.6. As discussed above, the current government, when choosing its optimal policy, takes as given the next period's value function, as well as its successor's time-consistent decision rules. When solving the problem, I approximate these functions using Chebyshev polynomials. Computations are performed in AMPL, using the KNITRO optimization solver (Ziena, 2013) for the maximization problems.

3.3.3 Results

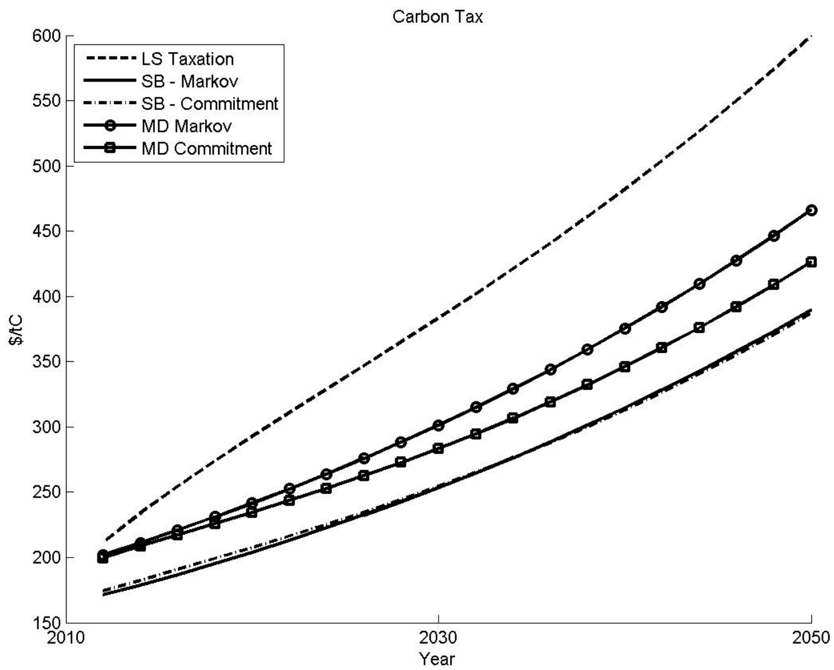

Figure 1: Optimal carbon tax - baseline

Figure 1 shows the time path of the optimal carbon price under a total income tax for the first best, the equilibrium with commitment and the Markov-perfect equilibrium, as well the corresponding levels of marginal climate damage in the latter two settings.

The first three rows of table 2 contain the numerical results for the baseline model. In first best, the tax schedule starts at 211\$/tC in 2011 and rises to about 600\$/tC in 2050. In the presence of distortionary taxation, by contrast, the carbon tax increases from 174\$/tC (2011) to 387\$/tC (2050) under commitment, and from 171\$/tC (2011) to almost 390\$/tC (2050) in the Markov-perfect equilibrium.

Several results are noteworthy. In both second-best settings, the optimal carbon tax is lower than in first best, by about 18% in 2011. The reason for this is twofold: first, due to

Year	$\$/tC$ 2010	2050	θ/θ^{FB} 2010	2050	θ/θ^{p} 2010	2050	τ 2010	2050
First Best - 3°	211.4	599.4	-	-	1.00	1.00	-	-
Markov - 3°	171.1	389.5	0.81	0.65	0.85	0.84	0.247	0.246
Commitment - 3°	174.1	387.0	0.82	0.65	0.87	0.91	0.247	0.310
First Best - 4°	79.0	237.3	-	-	1.00	1.00	-	-
Markov - 4°	70.7	162.5	0.89	0.68	1.05	1.00	0.250	0.250
Commitment - 4°	71.6	135.7	0.91	0.57	1.03	0.87	0.250	0.316
First Best - 4.5°	55.2	165.9	-	-	1.00	1.00	-	-
Markov - 4.5°	54.7	125.2	0.99	0.75	1.20	1.13	0.251	0.251
Commitment - 4.5°	54.6	91.5	0.99	0.55	1.13	0.83	0.251	0.317

Table 2: Results

the distortions in the consumption-savings and the consumption-labor margin, the second-best equilibria both with and without commitment feature lower output and lower consumption than the first-best outcome. This, in turn, reduces the damages from climate change and hence the absolute level of the marginal damages, i.e. the Pigouvian carbon tax. Moreover, the social cost of carbon does not equal the marginal damage. Instead, the optimal carbon tax is about 15% below the corresponding Pigouvian fee in 2011 due to the tax-interaction effect.

Comparing the Markov-perfect outcome with the equilibrium under commitment, note that the difference between the carbon tax rates is very small throughout the time period considered here. This is due to the fact that under lack of commitment, there is less provision of either public good, which leads to two counteracting effects. On the one hand, the commitment solution features more public consumption and hence a higher income tax rate than in the Markov-perfect equilibrium.[38] This results in a lower output and consumption level, and hence a smaller Pigouvian rate, as seen in figure 1. On the other hand, the optimal emission level in the Markov-perfect equilibrium is higher than under commitment, which would translate into lower carbon taxes, all else equal. In the baseline calibration, these effects cancel each out almost completely.

The reminder of table 2 shows the analogous results when changing α_s, the weight on temperature change in the utility function. When calibrating it such that global mean temperature stabilizes at $4°C$ in first best, instead of $3°C$, intuitively the level of the carbon tax decreases, to about 72\$/tC in the Markov-perfect equilibrium. Interestingly, the tax-interaction effect is smaller and goes in a different direction than in the baseline calibration. In 2011, the optimal carbon tax exceeds the Pigouvian fee by 5%, while in 2050, the tax-interaction effect is zero. When decreasing α_s even further, such that the target temperature in first best is $4.5°C$, the optimal carbon tax is substantially greater than the marginal damage level throughout the whole period.

[38]Compare Klein et al. (2008).

To shed some light on these findings, recall the expression for the absolute deviation Δ of the social cost of carbon from the marginal climate damage, derived above:

$$\Delta_t = \theta_t - \theta_t^p = -\frac{Q(t)}{u_c(t)} = \sum_{j=1}^{6} \tilde{Q}_j(t).$$

T	θ	Δ	$\tilde{Q}_1$	$\tilde{Q}_2$	$\tilde{Q}_3$	$\tilde{Q}_4$	$\tilde{Q}_5$	$\tilde{Q}_6$
3°	171.10	-29.59	-30.02	-4.99	38.01	-37.04	4.11	0.32
4°	70.65	2.91	-1.25	-1.34	10.01	-9.81	4.80	0.51
4.5°	54.65	9.26	3.41	-0.73	5.46	-5.36	4.70	1.77

Table 3: Results

Table 3 shows the values of the components of Δ in the three scenarios. In the baseline model, the overall effect is negative. Looking at the different components $\tilde{Q}_1$ though $\tilde{Q}_6$, one can see that what drives this result is mainly the first term. Recall that $\tilde{Q}_1$ quantifies the static labor effect: disregarding all other effects, the government would have an incentive to decrease the optimal carbon tax in order to account for the positive effect of fuel use on the labor supply in second best. By itself, the static labor effect would prescribe an optimal carbon tax more than 30\$/tC below the Pigouvian level. Similarly, $\tilde{Q}_2$ captures the second-best benefit of stimulating savings. It also has a negative effect on the social cost of carbon, but by considerably less than the static labor effect.

In terms of absolute value, $\tilde{Q}_3$ and $\tilde{Q}_4$ are the largest effect, albeit going in different directions and almost completely offsetting each other. The dynamic labor effect, $\tilde{Q}_3$, is positive, thus representing a second-best cost of current fuel use. That is, higher savings today due to more fuel use would lead to a net reduction in future labor supply, thereby exacerbating the intratemporal distortion tomorrow. In contrast, the dynamic public-good effect is positive: by emitting more carbon today – and hence underproviding environmental quality– the current government could induce its successor to provide more of the public consumption good.

Finally, the dynamic stock effects $\tilde{Q}_5$ and $\tilde{Q}_6$ are both positive, reflecting the negative impact of a higher pollutant stock on labor supply and savings in the future. Overall, note that the combined sign of the dynamic terms $\tilde{Q}_3$ through $\tilde{Q}_6$ is positive. All these terms capture the response of future governments to the policy chosen in the current period. In other words, disregarding contemporaneous second-best benefits of emissions - reflected by $\tilde{Q}_1$ and $\tilde{Q}_2$ - the current government would have an incentive to increase the optimal carbon tax above the Pigouvian level to avoid suboptimal policies chosen in the future.

With a first-best temperature target of $4°C$, the optimal carbon tax is slightly above the Pigouvian rate, by $\Delta = 2.91\$/\text{tC}$. In this case, the contemporaneous second-best benefits of emitting carbon, $\tilde{Q}_1$ and $\tilde{Q}_2$, are considerably smaller than before, due to the fact that the tax rate is much lower. At the same time, they are more than offset by the dynamic stock effects,

$\tilde{Q}_5$ and $\tilde{Q}_6$, which are slightly higher than in the baseline model, thus resulting in a positive Δ. In other words, the small tax-interaction is due to the fact that the different channels described above almost cancel each other out. Recall that the stock terms capture the effect of current carbon emissions on labor supply and savings decisions in all future period, through their impact on future productivity and hence on output and households' income. Thus, the stock effects are not affected by changes in α_s.

Finally, when targeting $4.5°C$ in first best, $\tilde{Q}_1$ is positive, implying that the static labor effect now drives up the optimal carbon tax and thus captures a second-best cost of burning carbon. The intuition for this change in the sign is that at sufficiently low carbon tax rates, while increasing fossil fuel still increases the net wage, the income effect now dominates the substitution effect, and hence the labor supply is reduced. Once more, this exacerbates the intratemporal distortion. Together with the positive stock effects, this results in a larger positive tax-interaction effect than in the previous setting.

Going back to the baseline model, note that the global first-best carbon tax in this paper is much higher than in other studies. In particular, Golosov et al. (2013) find an optimal carbon tax of 57\$/tC in 2010, while the newest version of the DICE model (Nordhaus, 2013) reports a tax of 66\$/tC in 2015 (18\$/tCO_2). The revised estimate of the Interagency Working Group on Social Cost of Carbon of the US government is higher, amouting to 121\$/tC in 2010 (33\$/tCO_2) for a comparable discount rate, but still well below my findings (IWG, 2013). The main difference to these studies, as explained above, is the law of motion for global mean temperature change: they explicitly model an atmospheric carbon stock that determines temperature change, and assume that some carbon leaves the atmosphere over time. By contrast, my model, based on Matthews et al. (2009), assumes that the effect of emitting carbon on temperature is permanent. Hence, when calibrating the model such that temperature change stabilizes at $3°C$ above the pre-industrial level in the long-run, this temperature is reached much later (2250) than in other models. For example, in Nordhaus (2013), global mean temperature change is at $3°C$ already around the year 2100. After that, temperature increases for three more decades, before it starts falling, and by 2200, temperature change is well below $3°C$. In my model, this decrease is not possible. As a consequence, in order to stabilize temperature change at $3°C$, climate policy must be much more stringent, which results in higher optimal carbon taxes. By contrast, when instead targeting a long-run increase of $4°C$, temperature change is at $3°C$ already in 2110, implying higher carbon emissions and much smaller carbon taxes.

To summarize, in this section I have shown that introducing distortionary income taxation in a global climate-economy model gives rise to a tax-interaction effect, which results in a wedge between the optimal carbon tax and the Pigouvian fee. This wedge is caused by the presence of second-best costs and benefits of burning fossil fuel. Quantitatively, the size and direction of this effect depends on the level of the carbon tax and hence on the socially desired temperature change.

4 A Two-region Model with Unilateral Climate Policy and Distortionary Taxation

In the previous section, I have considered a one-region global model of fiscal and climate policy. In reality, there is no global government, and both income and carbon taxes are set on the national level. Moreover, in the context of climate policy, cooperation between countries or regions has so far been restricted to initiatives like the Kyoto protocol, in which some countries committed to reducing carbon emissions. With respect to carbon pricing schemes, there is no cooperation and existing policies, such as the European Union Emission Trading Scheme, have been implemented unilaterally. As illustrated by figure B in the appendix, many major polluters, most notably the US and China, do not currently impose a price on carbon emissions.

Hence, unilateral climate policy and the absence of a carbon tax in many countries appears to be empirically much more relevant than the assumption of full participation across countries or regions made by one-region global models as the one in the previous section, as well as multi-region models that feature a cooperative regime (Nordhaus and Yang, 1996; Nordhaus, 2010).[39] Therefore, in this section, I modify the above model to a multiregion setting without full policy cooperation. For convenience, I divide the global region from the previous section into two equally-sized subregions.[40] For most of the analysis below, the only difference between the two regions will be the way policies are chosen.

In the baseline model, I follow Kehoe (1989) and Nordhaus and Yang (1996) and define a *noncooperative* regime as a setting in which a government maximizes its region's welfare. In contrast, in a *cooperative* regime, a single policy maker maximizes global welfare. To keep the analysis simple, however, I abstract from strategic interaction between regions. Specifically, I assume that fiscal and climate policy are set endogenously in only one of the two regions, while there is no carbon taxation regulation and no distortionary income taxation in the other region.

4.1 The Model

Without loss of generality, let the government in the "domestic" region D choose its policy optimally. In contrast, income and carbon tax rates in the rest of the world, denoted by R, are exogenously zero. Both regions are assumed to have identical preferences and production technologies, which are represented by the same functions u and F as in the global economy in section 2.

There are two channels through which the regions are linked with each other. First, climate change is a global externality. That is, the state of the climate, represented by the global mean

[39] In other words, these models compute the optimal carbon tax for a particular country, given that all other countries also price carbon emissions optimally.

[40] However, note that the main industrialized countries that have so far shown little willingness to engage in climate policy cooperation, specifically the US, China and Russia, account for about half of the global carbon emissions.

surface temperature change, is a function of aggregate global carbon emissions:

$$T_{t+1} = q(T_t, m_t^D + m_t^R). \tag{4.1}$$

Since utility and productivity in either region depend on T, emitting carbon in one region affects welfare in the other and vice versa. This link is present even if the two regions are otherwise closed economies.

The second channel arises from assuming that capital is perfectly mobile between the two regions. The motivation for this assumption is twofold. First, capital mobility is empirically relevant and important in the context of climate change. In particular, it allows for "carbon leakage", i.e. carbon emissions may rise in one region as a result of emission reductions in another. Here, if productive capital moves to the rest of the world due to the carbon tax in the domestic region, this may increase output and hence fuel use in R. Moreover, by allowing for capital mobility, the equilibrium outcome can feature distinct, positive tax rates on labor and capital income. Recall that in the global model, I had to restrict the analysis to a total income tax, since for separate tax rates, the labor tax would optimally be zero. In the two-region model, if capital can be reallocated between regions after the tax rates have been (credibly) announced, this results in an *endogenous* upper bound on the capital income tax.[41]

Assuming capital mobility requires that a distinction is made between the capital stock of a region, i.e. the capital owned by the household in this region, and the amount of capital used in production. Let z_t^D and z_t^R denote the levels of utilized capital in regions D and R, respectively.[42] Then, region j's capital export in period t is given by the difference between owned and utilized capital, $x_t^j = k_t^j - z_t^j$. If $x_t^j > 0$ ($x_t^j < 0$), region j owns more (less) capital than it uses in production. The amount of capital exported in one region must equal the amout imported in the other, and hence:

$$x_t^D + x_t^R = k_t^D - z_t^D + k_t^R - z_t^R = 0. \tag{4.2}$$

This is a market-clearing condition for the world market for capital.

The resource constraint in region D is then given by:

$$c_t^D + k_{t+1}^D + g_t^D + \kappa m_t^D = F(z_t^D, h_t^D, m_t^D, T_t) + (1-\delta)k_t^D + r_t^*(k_t^D - z_t^D), \tag{4.3}$$

where r_t^* denotes the world market rental rate for capital. Note that if region D is a capital exporter (importer), the last term on the right-hand side is positive (negative); hence, there is an inflow (outflow) of resources. An analogous constraint holds for region R:

$$c_t^R + k_{t+1}^R + g_t^R + \kappa m_t^R = F(z_t^R, h_t^R, m_t^R, T_t) + (1-\delta)k_t^R + r_t^*(k_t^R - z_t^R), \tag{4.4}$$

[41]Martin (2010) shows that assuming an exogenous upper bound on the capital tax results in positive labor tax rates.

[42]Output is then a function of z rather than k: $y^j = F(z^j, h^j, m^j, T$.

In equilibrium, with perfectly mobile capital, the net return to utilized capital in either region must equal the world rental rate r^*:

$$r_t^* = (1 - \tau_t^{k,D})(F_q^D(t) - \delta) = (1 - \tau_t^{k,R})(F_z^R(t) - \delta), \tag{4.5}$$

where the tax on capital income in R, $\tau^{R,k}$, is zero. That is, the price of the mobile factor is equalized across regions.

Taking the behavior by the policymaker and the households in D as given, a competitive equilibrium in R is characterized by the following set of equations:

$$-u_c^R(t) + \beta u_c^R(t+1)[1 - (1 - \tau_t^{k,R})(F_z^R(t+1) - \delta)] = 0, \tag{4.6}$$

$$u_l^R(t) - u_c^R(t)(1 - \tau_t^{h,R})F_h^R(t) = 0, \tag{4.7}$$

$$u_g^R(t) - u_g^R(t) = 0, \tag{4.8}$$

$$F_m^R(t) - \kappa = 0, \tag{4.9}$$

representing the foreign household's Euler equation, its labor-leisure optimality condition, the first-order condition for public good provision, and the foreign firm's first-order condition for fuel use. Since there is no climate policy, (4.9) states that the price of fossil fuels must equal the private extraction cost. For simplicity, I assume that the public consumption good in R is financed with a lump-sum tax.

4.1.1 Cooperative Regime

Under policy cooperation, a single policy maker maximizes global welfare (Kehoe, 1989), that is, the sum of welfare over regions. It is straightforward to show that if the two regions are identical, the aggregate outcome in this setting is the same as in the global planner model in section 3 above.[43] To see this, first note that apart from temperature change T, all other variables in the global model were expressed in efficiency units. Reducing the scale of the economy, say, by half, is equivalent to decreasing the productivity level to $0.5A_t$. However, this does not affect the allocation in efficiency units if the time path of T (which is expressed in degrees Celsius) does not change.

Moreover, since both regions are identical, there are no capital flows in equilibrium. Without the global externality, this would imply that the problem is that of two closed economies (Kehoe, 1989). With climate change, assume that in every period, both regions emit half of the total carbon emissions in the global planner economy. Then, the time path global mean temperature change is the same as in the one-region model. This implies that the allocations, in efficiency units, in each region are also identical to the global setting.

[43] If the regions are not identical and capital is mobile, the equilibrium features capital movements between the regions, which makes the computation of the cooperative outcome more involved. One approach is applying the Negishi algorithm (Nordhaus and Yang, 1996; Ljungqvist and Sargent, 2004).

4.1.2 The Unilateral Problem with Lump-sum Taxation

Due to both the global climate externality and capital mobility, the government in D has to take into account the behavior of households in R when choosing its optimal policies. Hence, compared to a global economy, the domestic government faces a number of additional choice variables and constraints.

As in the global model, assume that carbon is emitted until an exogenously given period J. Then, if lump-sum taxes are available, the domestic government's sequential problem in period $t = 0$ can be written as:

$$\max_{\substack{c_t^D, k_{t+1}^D, z_t^D, h_t^D, g_t^D, m_t^D \\ c_t^R, k_{t+1}^R, z_t^R, h_t^R, m_t^R, T_{t+1}}} \sum_{t=0}^{J} \beta^t \big(u(c_t^D, 1-h_t^D, g_t^D, T_t) + \varphi u(c_t^F, 1-h_t^F, g_t^F, T_t) \big)$$
$$+ \beta \bar{v}(k_{J+1}^D, k_{J+1}^R, T_{J+1}) + \beta \varphi \bar{v}(k_{J+1}^R, k_{J+1}^D, T_{J+1}),$$

subject to the law of motion for climate change (4.1), the two resource constraints (4.3) and (4.4), the equilibrium conditions in R, (4.6) - (4.9), the world market clearing condition (4.2) and the factor price equalization constraint (4.5). Moreover, $\bar{v}$ denotes the terminal value function for the stationary phase after period J, which for simplicity I assume to be the same in both regions.

Note that with respect to the domestic government's objective function, φ denotes the weight on the welfare of the households in the rest of the world. In the quantitative analysis below, I consider two extreme cases. For $\varphi = 0$, the domestic government only cares about the welfare of the domestic household, but not about the agents in the other region. In contrast, if $\varphi = 1$, there are equal weights on both regions, and hence the domestic government is benevolent in the sense that it maximizes global welfare. Thus, the results for the optimal carbon tax computed below constitute an upper and a lower bound.[44]

4.1.3 The Unilateral Problem with Distortionary Taxation

I now turn to the "third-best" setting with unilateral policy and distortionary income taxes. As in the global model above, I can solve for the non-stationary equilibrium using backward induction.

Consider the case that the domestic government maximizes welfare in D, i.e. $\varphi = 0$. Let ς_{t+1}^D denote the policy function for the capital income tax that the future domestic government follows in period $t+1$. For period $t < J$, given next period's policy functions for both domestic ($n_{t+1}^{c,D}$, $n_{t+1}^{z,D}$, $n_{t+1}^{h,D}$, ψ_{t+1}^D, ϕ_{t+1}^D, ς_{t+1}^D) and foreign variables ($n_{t+1}^{c,R}$, $n_{t+1}^{z,R}$, $n_{t+1}^{h,R}$, ϕ_{t+1}^R), and the value

[44] In its report, the US government interagency working group argued that a proper estimate of the social cost of carbon should include all damages, domestic and foreign (IWG, 2010), arguing that climate policy should be done by international agreements. However, in the absence of such cooperation, the question of which value for φ is more relevant is ambiguous, and eventually comes down to moral judgement.

function v_{t+1}, the domestic government in t solves the following problem:[45]

$$\max_{\substack{c_t^D, k_{t+1}^D, z_t^D, h_t^D, g_t^D, m_t^D, \tau_t^{k,D} \\ c_t^R, k_{t+1}^R, z_t^R, h_t^R, m_t^R, T_{t+1}}} u(c_t^D, 1 - h_t^D, g_t^D, T_t) + \beta v_{t+1}(k_{t+1}^D, k_{t+1}^R, T_{t+1}),$$

subject to

1. the law of motion for temperature change:

$$T_{t+1} = q(T_t, m_t^D + m_t^R) \tag{4.10}$$

2. the resource constraint in D:

$$\begin{aligned} F(z_t^D, h_t^D, m_t^D, T_t) + (1-\delta)k_t^D - c_t^D - g_t^D - \kappa m_t^D - k_{t+1}^D \\ + \left(1 - \tau_t^{k,R}\right)\left[F_z\left(\begin{array}{c} z_t^R, h_t^R, \\ m_t^R, T_t \end{array}\right) - \delta\right](k_t^D - z_t^D) = 0 \end{aligned} \tag{4.11}$$

3. the household's intratemporal optimality condition in D:

$$\frac{u_l(c_t^D, 1 - h_t^D, g_t^D, T_t)}{u_c(c_t^D, 1 - h_t^D, g_t^D, T_t)} - F_h^D(t)[1 - \mathcal{T}(z_t^D, h_t^D, g_t^D, m_t^D, T_t)] = 0, \tag{4.12}$$

4. the household's Euler equation in D:

$$\begin{aligned} u_c(c_t^D, 1 - h_t^D, g_t^D, T_t) - \beta u_c'\left(n_{t+1}^{c,D}, 1 - n_{t+1}^{h,D}(t+1), \psi_{t+1}^D(t+1), T_{t+1}\right) \cdot \\ \cdot \left[1 + [1 - \varsigma_{t+1}^D(t+1)]\left[F_z\left(\begin{array}{c} n_{t+1}^{z,D}(t+1), n_{t+1}^{h,D}(t+1), \\ \phi_{t+1}^D(t+1), T_{t+1} \end{array}\right) - \delta\right]\right] = 0, \end{aligned} \tag{4.13}$$

5. the resource constraint in R:

$$\begin{aligned} F(z_t^R, h_t^R, m_t^R, T_t) + (1-\delta)k_t^R - c_t^R - g_t^R - \kappa m_t^R - k_{t+1}^R \\ + \left(1 - \tau_t^{k,R}\right)\left[F_z\left(\begin{array}{c} z_t^R, h_t^R, \\ m_t^R, T_t \end{array}\right) - \delta\right](k_t^R - z_t^R) = 0 \end{aligned} \tag{4.14}$$

6. the household's intratemporal optimality condition in R:

$$\frac{u_l(c_t^R, 1 - h_t^R, g_t^R, T_t)}{u_c(c_t^R, 1 - h_t^R, g_t^R, T_t)} - F_h^R(t)[1 - \tau_t^{h,R}] = 0, \tag{4.15}$$

[45] In period J, the last period with carbon emissions, the government's problem reads

$$\max_{\substack{c_J^D, k_{J+1}^D, z_J^D, h_J^D, g_J^D, m_J^D, \tau_J^{k,D} \\ c_J^R, k_{J+1}^R, z_J^R, h_J^R, m_J^R, T_{J+1}}} u(c_J^D, 1 - h_J^D, g_J^D, T_J) + \beta \bar{v}(k_{J+1}^D, k_{J+1}^R, T_{J+1}),$$

subject to constraints (4.10)-(4.19).

7. the household's Euler equation in R:

$$\begin{aligned} & u_c(c_t^R, 1-h_t^R, g_t^R, T_t) - \beta u_c' \left(n_{t+1}^{c,R}, 1 - n_{t+1}^{h,R}(t+1), \psi_{t+1}^R(t+1), T_{t+1} \right) \cdot \\ & \cdot \left[1 + \left(1 - \tau_{t+1}^{k,R} \right) \left[F_z \left(\begin{array}{c} n_{t+1}^{z,R}(t+1), n_{t+1}^{h,R}(t+1), \\ \phi_{t+1}^R(t+1), T_{t+1} \end{array} \right) - \delta \right] \right] = 0, \end{aligned} \tag{4.16}$$

8. the firm's optimality condition for fuel use in R:

$$F(z_t^R, h_t^R, m_t^R, T_t) - \kappa = 0, \tag{4.17}$$

9. the factor price equalization constraint:

$$\left(1 - \tau_t^{k,D}\right) \left[F_z \left(\begin{array}{c} z_t^D, h_t^D, \\ m_t^D, T_t \end{array} \right) - \delta \right] - \left(1 - \tau_t^{k,R}\right) \left[F_z \left(\begin{array}{c} z_t^R, h_t^R, \\ m_t^R, T_t \end{array} \right) - \delta \right] = 0, \tag{4.18}$$

10. and the clearing condition for the world capital market:

$$k_t^D - z_t^D + k_t^R - z_t^R = 0, \tag{4.19}$$

Note that $n_{t+1}^{c,D}(t+1) = n_{t+1}^c(k_{t+1}^D, k_{t+1}^R, T_{t+1})$ etc. and

$$\mathcal{T}(z_t^D, h_t^D, g_t^D, m_t^D, \tau_t^{k,D}, T_t) = \frac{g - (F_m^D(t) - \kappa) m_t^D - \tau_t^{k,D}(F_z^D(t) - \delta) z_t^D}{F_h^D(t) h_t^D}.$$

denotes the labor tax function, where $F_m^D(t) = F_m(z_t^D, h_t^D, m_t^D, T_t)$, etc.

The solution to this problem then gives the contemporaneous policy functions $n_t^{c,D}$, $n_t^{k,D}$, $n_t^{z,D}$, $n_t^{h,D}$, ψ_t^D, ϕ_t^D, ς_t^D, $n_t^{c,R}$, $n_t^{k,R}$, $n_t^{z,R}$, $n_t^{h,R}$, ϕ_t^R, as well as the value function v_t:

$$\begin{aligned} v_t(k_t^D, k_t^R, T_t) =& u \left(n_t^{c,D}(k_t^D, k_t^R, T_t), 1 - n_t^{h,D}(k_t^D, k_t^R, T_t), \psi_t^D(k_t^D, k_t^R, T_t), T_t \right) \\ & + \beta v_{t+1} \left(n_t^{k,D}(k_t^D, k_t^R, T_t), n_t^{k,R}(k_t^D, k_t^R, T_t), q(T_t, \phi_t^D(k_t^D, k_t^R, T_t) + \phi_t^R(k_t^D, k_t^R, T_t)) \right). \end{aligned}$$

4.2 Quantitative Analysis

As in section 3, I use finite-horizon dynamic programming to solve this problem, and approximate future policy and value functions using Chebyshev polynomials. Compared to the global model above, I have more choice variables and constraints in the maximization problem and more policy functions to approximate.

Table 4 shows the estimates for the optimal carbon tax for the baseline calibration. For comparison, the first line displays the result for the first-best global tax, θ^{FB}, from section 3. Figures 2 and 3 display the optimal carbon tax schedules under both lump-sum and distortionary taxation, for $\varphi = 0$ and $\varphi = 1$, respectively, and compare them to the global first-best outcome.

Year	$\$/tC$ 2011	2050	θ^m/θ^{FB} 2011	2050	θ^m/θ^p 2011	2050	τ^k 2011	τ^h 2011
Global First-Best	211.4	599.4	-	-	1.00	1.00	-	-
Two-region, LST, $\varphi = 0$	126.2	390.2	0.60	0.65	0.96	0.99	-	-
Two-region, DT, $\varphi = 0$	96.1	247.1	0.45	0.41	0.88	0.81	0.14	0.26
Two-region, LST, $\varphi = 1$	244.7	737.6	1.16	1.23	0.95	0.97	-	-
Two-region, DT, $\varphi = 1$	193.9	495.0	0.92	0.83	0.91	0.88	0.29	0.28

Table 4: Results

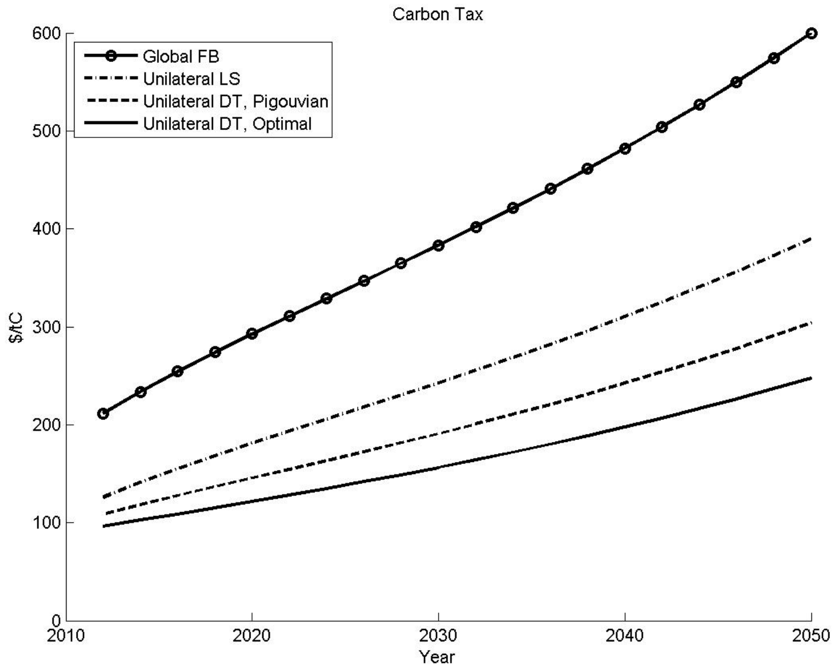

Figure 2: Optimal unilateral carbon tax - $\varphi = 0$

4.2.1 Lump-Sum Taxation

I start by only considering a noncooperative regime, while abstracting from distortionary taxation. In the two-region setting, if lump-sum taxes are available and if $\varphi = 0$, the optimal carbon tax amounts to 126\$/tC in 2011 and 390\$/tC in 2050. Intuitively, these values are considerably lower, by about 40%, than in the global model, since the government maximizes the welfare of its own households, rather than global welfare. Hence, in contrast to the global planner in the previous section, it does not take into account the climate damages suffered by the other region.

Note also that the optimal carbon tax under lump-sum taxation in 2011 exceeds half the global first-best rate, despite the assumption that each of the two regions initially comprises half of the world economy. This is because the rest of world emits more carbon than in the

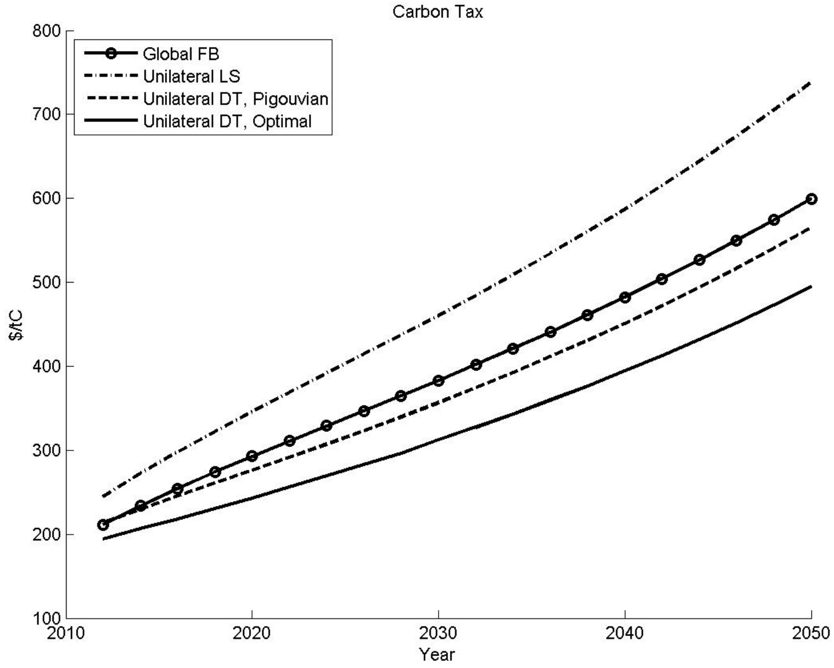

Figure 3: Optimal unilateral carbon tax - $\varphi = 1$

cooperative regime, due to the absence of climate policy. To illustrate this formally, first note that for a general φ, the Pigouvian tax is given by:[46]

$$\theta_t^{p,D} = -\frac{q_m(t)}{u_c(t)} \sum_{j=t+1}^{\infty} \beta^{j-t} \left[(u_T^D(j) + u_c^D(j) F_T^D(j)) + \varphi (u_T^R(j) + u_c^R(j) F_T^R(j)) \right]. \tag{4.20}$$

In the cooperative regime, $\varphi = 1$ and the allocations in both regions are identical. It follows that $\theta_t^{p,D} = \theta_t^p$, i.e. the global Pigouvian tax rate. Without policy cooperation, set $\varphi = 0$, but assume that both regions still implement the allocations from the cooperative regime. Then, (4.20) implies that $\theta_t^{p,D} = 0.5 \cdot \theta_t^p$, as the domestic government does not put any weight on the welfare cost from climate damages in R. Finally, consider the noncooperative regime, i.e. abstract from climate policy in R and let the government in D choose its optimal policy. Relative to the previous scenario, carbon emissions in R increase, which implies a higher temperature change. Due to the convexity of both utility and productivity damages, this implies a higher marginal damage in D:

$$\frac{\partial}{\partial T}(u_T^D + u_c^D F_T^D) > 0.$$

Hence, the Pigouvian tax level increases: $\theta_t^{p,D} > 0.5 \cdot \theta_t^p$.

Moreover, note from table 4 that under lump-sum taxation, the optimal carbon tax is close to but not exactly at the Pigouvian level. To understand the intuition behind this result,

[46]Note that $q_T(t) = 1$ for all periods t.

consider the first-order conditions from the sequential government problem with lump-sum taxation described in section 4.1.2. Let λ_t^R denote the Lagrange multiplier associated with the resource constraint (4.4), respectively, and χ_t the Lagrange multiplier associated with (4.5), the factor price equalization condition. The first-order conditon with respect to m_t^D then reads:

$$\begin{aligned} u_c(t)(F_m^D(t) - \kappa) + q_m(t) \sum_{j=t+1}^{\infty} \beta^{j-t} \left[u_T^D(j) + u_c^D(j)F_T^D(j)\right] \\ + F_{zm}^D(t)\chi_t - \sum_{j=t+1}^{\infty} \lambda_j^R F_T^R(j) - F_{zm}^D(t)\lambda_t^R(k_t^R - z_t^R) = 0 \end{aligned} \tag{4.21}$$

The first and the second term are the marginal benefit and the marginal damage from fossil fuel use in utils, respectively. Hence, the remaining term introduces a wedge between the optimal carbon tax and the Pigouvian fee. The third term captures the increase in the return to capital use in D if more fossil fuel is burnt, due to the complementarity of capital and energy in the production function ($F_{zm}^D > 0$). This is a benefit of emitting carbon, since it relaxes the factor price equalization constraint.

The fourth and fifth term reflect a change in the resource constraint in R due to higher fossil fuel use in D. On the one hand, emitting more carbon leads to a higher temperature change and hence reduces productivity in all future periods. On the other hand, it increases the return to capital use and hence R's income from capital export. The sum of the last two terms is very small. The fourth row in table 4 gives the optimal unilateral carbon tax under lump-sum taxation for $\varphi = 1$, i.e. assuming that the domestic government maximizes global welfare. Now, the optimal carbon tax exceeds the global first-best rate, by 16% in 2011 and 23% in 2050. The intuition is similar to the argument made in the previous setting: since the rest of the world emits too much carbon, compared to the global optimum, the domestic region optimally reduces emissions relative to the first-best, to partially counteract this increase. Figure 4 illustrates this by plotting carbon emissions in D and R with and without policy cooperation. It also shows that the emission reductions in D do not completely make up for the increase in R: total emissions in the unilateral setting - the sum of the top and the bottom curve - rise as compared to the first best.

4.2.2 Full Model

I now consider the model with both unilateral policy making and distortionary income taxation. The third row in table 4 shows the outcome in the unilateral setting with income taxes and lack of commitment. With $\varphi = 0$, this "third-best" optimal carbon fee is 96\$/tC in 2011 and 247\$/tC in 2050 and hence about 55% below the global first-best global tax. Hence, compared to the outcome with lump-sum taxation, distortionary taxation increases the gap between the first-best and the unilateral carbon tax by about 50%.

The intuition for this decrease in the optimal carbon tax is the same as in the global planner model. It is partly due to the negative impact of distortionary taxes on economic activity, in

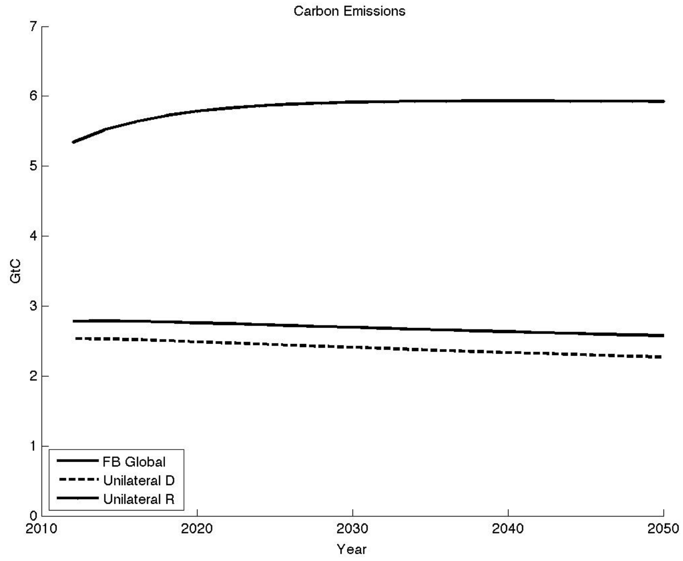

Figure 4: Carbon Emissions under lump-sum taxation and $\varphi = 1$

particular output and consumption, which results in a lower value of the climate damages than in a setting with lump-sum taxes. An additional reduction in the carbon tax is caused by the tax-interaction effect, resulting in a wedge between the optimal carbon tax and the Pigouvian rate. Note that in the two-region model here, the optimal tax is about 12% lower than the marginal climate damage.

For $\varphi = 0$, both distortions have a negative effect on the social cost of carbon. For the other extreme, $\varphi = 1$, distortionary income taxation and unilateral policy making go in different directions. While the absence of policy cooperation increases optimal carbon tax, the lack of lump-sum taxation reduces the social cost of carbon. In the baseline calibration, the overall effect is negative: the optimal carbon tax is initially about 8% below the global first-best rate. This gap increases over time to 17% in 2050.

To summarize, at least for the case $\varphi = 0$, the deviation of the optimal carbon tax from the global first-best rate is considerable. In this sense, a first-best integrated assessment model is not robust to the addition of distortions. In contrast, when $\varphi = 1$, the optimal carbon tax is much closer to the global first best since the effects of distortionary taxation and unilateral carbon pricing partially offset each other.

5 Conclusion

This paper has analyzed optimal carbon taxation in a world where governments have to resort to distortionary taxation to finance public expenditures and where regions do not cooperate when setting climate policy. I have added these features to an otherwise standard dynamic climate-economy model and computed optimal carbon tax schedules for different settings.

The main findings of this paper are the following. First, I characterize optimal policy analytically in a global planner model. I illustrate that the optimal second-best carbon tax is in general not at the Pigouvian level, due to the presence of additional costs and benefits of carbon emissions that only materialize under distortionary income taxation. In contrast to previous studies, I show that it is not only the current labor-leisure margin that is affected by climate policy, but there are other current and future wedges that interact with carbon taxes. Second, a quantitative analysis shows that the size and direction of the net tax-interaction effect depends on the level of the carbon tax and hence on the cost of climate change. In the baseline model, the overall effect is negative, resulting in an optimal second-best carbon tax which is about 20% below the Pigouvian rate.

Third, in a model with two equally-sized regions and in the absence of policy cooperation, the optimal carbon tax decreases considerably if a government does not take into account climate damages in the rest of the world. I find that the optimal carbon tax decreases by about 55%, relative to the global first-best outcome. In contrast, if the government is benevolent in the sense that it maximizes global welfare, this decrease is much smaller, and the optimal carbon tax is only about one fifth below the first-best Pigouvian rate.

In order to facilitate the analysis, I have made a number of simplifying assumptions. Relaxing those could result in potentially interesting extensions of the above framework. Most notably, by restricting the analysis to optimal climate policy making in only one of the two regions, I have abstracted from strategic interaction between governments. Considering a richer framework, where climate and fiscal policy are chosen optimally in both regions, may change the quantitative results. In particular, if the rest of the world has an incentive to curb carbon emissions, this may reduce the social cost of carbon in the domestic country, and hence lead to a lower optimal carbon tax under both lump-sum and distortionary taxation.

Moreover, the framework in this paper is deterministic and has abstracted from uncertainty, both with respect to climate change and to long- and short-run economic growth. Regarding the latter, it is straight forward to add exogenous productivity or taste shocks to the model, in the tradition of the real business cycle literature. Previous work by Heutel (2012) has shown that in a first-best setting, climate policy is procyclical in the sense that a carbon tax optimally increases during expansion, while it must be reduced in a recession. The framework used in this paper would allow to analyze how robust this finding is to the introduction of distortionary income taxes into the economy.

Given its long term nature and its complexity, climate change gives rise to many types of uncertainty, both related to the science and to the economic damages.[47] Integrated assessment models are usually either deterministic, or consider parametric uncertainty (Nordhaus, 2008; Golosov et al., 2013). Some recent papers instead focus on "intrinsic" uncertainty, that

[47] Compare Stern (2013) for a recent summary.

is, uncertainty caused by the random occurrence of exogenous events (Cai et al., 2012). In particular, studies by Lemoine and Traeger (2013) and Cai et al. (2013) incorporate so-called "tipping points", defined as irreversible and abrupt shifts in the climate system, in stochastic versions of the DICE model. Jensen and Traeger (2013) analyze optimal carbon mitigation under long-term growth uncertainty. While such questions are beyond the scope of this paper, it is important to keep in mind that these channels have a potentially large quantitative impact on optimal carbon taxes.

Integrated assessment models such as the one used in this paper have other limitations. Two important areas of ongoing research are the modeling of economic growth and the representation of climate damages. The above framework has built upon the neoclassical growth model, assuming exogenously given progress of both general-purpose technology and energy efficiency. A recent strand of the literature has instead considered optimal environmental policy in endogenous-growth models. Acemoglu et al. (2012), using a model of directed technical change, endogenize productivity growth for both a clean and a dirty production input. They show that optimal climate policy in such a setting consists of both a carbon tax and a research subsidy to the clean type of energy. Hemous (2013) embeds this framework in a setting with two regions and analyzes unilateral policy. He also finds that research subsidies are an important component of optimal policy. These results suggest that incorporating endogenous growth along similar lines would affect the results in this paper qualitatively and quantitatively.

Moreover, Stern (2013) notes that integrated assessment models usually assume that economic growth is not affected by climate damages. Instead, the multiplicative damage function that is used in this paper, as well as in many others studies, relates contemporaneous damages to the current flow of output, but not to, for example, the stock of capital or other factors determining the growth potential of the economy.[48] This shortcoming is exacerbated by the fact that the damage function as used in the DICE model yields quantitatively small damages.[49] These issues are summarized by Stern (2013), arguing that the "exogeneity of a key driver of growth, combined with weak damages" is one of the key weaknesses of current integrated assessment modeling. While this paper had a different focus and did not attempt to contribute to advances in climate-economy modeling in these areas, one should keep these limitations in mind when interpreting its results.

[48] A recent paper by Moyer et al. (2013) analyzes a model in which climate change has direct effects on productivity. They find that this leads to a considerable increase in the social cost of carbon.

[49] As shown by Ackerman et al. (2010), the damage function in Nordhaus (2008) would imply a decrease in output by only 50% when temperature increases by $19°C$ relative to the preindustrial level. However, it should be noted that William Nordhaus warns that there is not sufficient evidence to reliably use this damage function for temperature changes above $3°C$ (Stern, 2013).

References

Acemoglu, D., P. Aghion, L. Bursztyn, and D. Hemous (2012). The environment and directed technical change. *The American Economic Review 102*(1), 131–66.

Ackerman, F., E. A. Stanton, and R. Bueno (2010). Fat tails, exponents, extreme uncertainty: Simulating catastrophe in dice. *Ecological Economics 69*(8), 1657–1665.

Anthoff, D., C. Hepburn, and R. S. Tol (2009). Equity weighting and the marginal damage costs of climate change. *Ecological Economics 68*(3), 836–849.

Azzimonti, M., P. Sarte, and J. Soares (2006). Optimal policy and (the lack of) time inconsistency: Insights from simple models.

Azzimonti, M., P.-D. Sarte, and J. Soares (2009). Distortionary taxes and public investment when government promises are not enforceable. *Journal of Economic Dynamics and Control 33*(9), 1662–1681.

Barrage, L. (2013). Optimal dynamic carbon taxes in a climate-economy model with distortionary fiscal policy. Mimeo, University of Maryland.

Barrett, S. (2005). The theory of international environmental agreements. *Handbook of environmental economics 3*, 1457–1516.

Battaglini, M. and S. Coate (2007). Inefficiency in legislative policymaking: A dynamic analysis. *The American Economic Review 97*(1), 118–149.

Battaglini, M. and S. Coate (2008). A dynamic theory of public spending, taxation, and debt. *The American Economic Review 98*(1), 201–36.

Bosetti, V., C. Carraro, M. Galeotti, E. Massetti, and M. Tavoni (2006). A world induced technical change hybrid model. *The Energy Journal* (Special Issue), 13–38.

Bovenberg, A. and R. de Mooij (1994). Environmental levies and distortionary taxation. *The American Economic Review 84*(4), 1085–1089.

Bovenberg, A. and L. Goulder (1996). Optimal environmental taxation in the presence of other taxes: general-equilibrium analyses. *The American Economic Review 86*(4), 985–1000.

Bovenberg, A. and L. Goulder (2002). Environmental taxation and regulation. *Handbook of Public Economics 3*, 1471–1545.

Bovenberg, A. L. and F. van der Ploeg (1994). Environmental policy, public finance and the labour market in a second-best world. *Journal of Public Economics 55*(3), 349–390.

Cai, Y., K. Judd, and T. Lontzek (2012). DSICE: a dynamic stochastic integrated model of climate and economy.

Cai, Y., K. L. Judd, and T. S. Lontzek (2013). The social cost of stochastic and irreversible climate change. *NBER Working Paper 18704*.

Chamley, C. (1986). Optimal taxation of capital income in general equilibrium with infinite lives. *Econometrica: Journal of the Econometric Society*, 607–622.

Cremer, H. and F. Gahvari (2001). Second-best taxation of emissions and polluting goods. *Journal of Public Economics 80*(2), 169–197.

EESI (2012). Fact Sheet: Carbon Pricing around the World. `http://files.eesi.org/FactSheet_Carbon_Pricing_101712.pdf`. Accessed November 29, 2013.

Fischer, C. and M. Springborn (2011). Emissions targets and the real business cycle: Intensity targets versus caps or taxes. *Journal of Environmental Economics and Management 62*(3), 352–366.

Fischer, S. (1980). Dynamic inconsistency, cooperation and the benevolent dissembling government. *Journal of Economic Dynamics and Control 2*, 93–107.

Gerlagh, R. and M. Liski (2012). Carbon prices for the next thousand years. *CESifo Working Paper Series*.

Glomm, G., D. Kawaguchi, and F. Sepulveda (2008). Green taxes and double dividends in a dynamic economy. *Journal of Policy Modeling 30*(1), 19–32.

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

Goulder, L. H. (1995). Effects of carbon taxes in an economy with prior tax distortions: an intertemporal general equilibrium analysis. *Journal of Environmental economics and Management 29*(3), 271–297.

Goulder, L. H., I. W. Parry, and D. Burtraw (1997). Revenue-raising versus other approaches to environmental protection: The critical significance of preexisting tax distortions. *The Rand Journal of Economics*, 708–731.

Harstad, B. (2012). Climate contracts: A game of emissions, investments, negotiations, and renegotiations. *The Review of Economic Studies 79*(4), 1527–1557.

Harstad, B. (2013). The dynamics of climate agreements.

Hemous, D. (2013). Environmental policy and directed technical change in a global economy: the dynamic impact of unilateral environmental policies.

Heutel, G. (2012). How should environmental policy respond to business cycles? optimal policy under persistent productivity shocks. *Review of Economic Dynamics 15*(2), 244–264.

Hope, C. (2006). The marginal impact of co2 from page2002: An integrated assessment model incorporating the ipcc's five reasons for concern. *Integrated assessment 6*(1).

Hope, C. (2008). Optimal carbon emissions and the social cost of carbon over time under uncertainty. *Integrated Assessment 8*(1).

IWG (2010). Social cost of carbon for regulatory impact analysis under executive order 12866. Interagency Working Group on Social Cost of Carbon (IWG), United States Government. `http://www.whitehouse.gov/sites/default/files/omb/inforeg/for-agencies/Social-Cost-of-Carbon-for-RIA.pdf`.

IWG (2013). Technical update of the social cost of carbon for regulatory impact analysis under executive order 12866. Interagency Working Group on Social Cost of Carbon (IWG), United States Government. `http://www.whitehouse.gov/sites/default/files/omb/inforeg/social_cost_of_carbon_for_ria_2013_update.pdf`.

Jensen, S. and C. Traeger (2013). Mitigation under long-term growth uncertainty: Growing emissions but outgrowing its consequences–sure?

Judd, K. L. (1985). Redistributive taxation in a simple perfect foresight model. *Journal of public Economics 28*(1), 59–83.

Judd, K. L. (1998). *Numerical methods in economics.* The MIT press.

Kehoe, P. J. (1989). Policy cooperation among benevolent governments may be undesirable. *The Review of Economic Studies 56*(2), 289–296.

King, R. G., C. I. Plosser, and S. T. Rebelo (2002). Production, growth and business cycles: Technical appendix. *Computational Economics 20*(1), 87–116.

Klein, P., P. Krusell, and J. Rios-Rull (2008). Time-consistent public policy. *The Review of Economic Studies 75*(3), 789–808.

Kolstad, C. D. (2000). *Environmental economics*, Volume 15. Oxford university press Oxford.

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

Ljungqvist, L. and T. Sargent (2004). *Recursive macroeconomic theory.* MIT press.

Lucas, R. E. and N. L. Stokey (1983). Optimal fiscal and monetary policy in an economy without capital. *Journal of monetary Economics 12*(1), 55–93.

Martin, F. (2010). Markov-perfect capital and labor taxes. *Journal of Economic Dynamics and Control 34*(3), 503–521.

Matthews, H. D., N. P. Gillett, P. A. Stott, and K. Zickfeld (2009). The proportionality of global warming to cumulative carbon emissions. *Nature 459*(7248), 829–832.

Moyer, E., M. Woolley, M. Glotter, and D. A. Weisbach (2013). Climate impacts on economic growth as drivers of uncertainty in the social cost of carbon. *Center for Robust Decision Making on Climate and Energy Policy Working Paper*, 13–02.

Nordhaus, W. (2008). *A question of balance: Weighing the options on global warming policies*. Yale University Press.

Nordhaus, W. D. (2010). Economic aspects of global warming in a post-copenhagen environment. *Proceedings of the National Academy of Sciences 107*(26), 11721–11726.

Nordhaus, W. D. (2013). DICE 2013R: Introduction and User's Manual. `http://www.econ.yale.edu/~nordhaus/homepage/documents/DICE2013R_Manual_2ndEd_000.pdf`. Accessed October 29, 2013.

Nordhaus, W. D. and Z. Yang (1996). A regional dynamic general-equilibrium model of alternative climate-change strategies. *The American Economic Review*, 741–765.

Olivier, J. G., G. Janssens-Maenhout, M. Muntean, and J. A. Peters (2013). *Trends in global co2 emissions: 2013 report*. PBL Netherlands Environmental Assessment Agency.

Parry, I. W. (1995). Pollution taxes and revenue recycling. *Journal of Environmental Economics and management 29*(3), S64–S77.

Phelan, C. and E. Stacchetti (2001). Sequential equilibria in a ramsey tax model. *Econometrica 69*(6), 1491–1518.

Reis, C. (2011). Taxation without commitment. *Economic Theory*, 1–24.

Stern, N. (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, R. S. (2002a). Estimates of the damage costs of climate change. part 1: Benchmark estimates. *Environmental and Resource Economics 21*(1), 47–73.

Tol, R. S. (2002b). Estimates of the damage costs of climate change, part ii. dynamic estimates. *Environmental and Resource Economics 21* (2), 135–160.

Weitzman, M. L. (2010). What is the "damage function" for global warming - and what difference does it make? *Climate Change Economics 1* (01), 57–69.

Ziena (2013). KNITRO Documentation, Release 8.1. Ziena Optimization LLC. `http://www.ziena.com/docs/Knitro81_UserManual.pdf`. Accessed October 29, 2013.

Zivin, J. G. and M. J. Neidell (2010). Temperature and the allocation of time: Implications for climate change. *NBER Working Paper 15717*.

A Appendix A

A.1 Static Model

The government's problem is given by

$$\max_{m,h} u(F(k,h,m) - g - \kappa m, 1-h, q(m))$$

s.t.

$$u_c F_h(1-\tau) - u_l = 0 \tag{A.1}$$

$$c = (1-\tau)(F_k k + F_h h) \tag{A.2}$$

Using $c = F(k,h,m) - g - \kappa m$, these two equations can be collapsed into

$$\begin{aligned}\chi(h,m) \equiv\ & u_c\big(F(k,h,m) - g, 1-h, q(m)\big)[F(k,h,m) - g - \kappa m]\frac{F_h h}{F_h h + F_k k} \\ & - u_l\big(F(k,h,m) - g, 1-h, q(m)\big)h = 0.\end{aligned}$$

Let $h = \sigma(\tilde{w}, y)$ be the Marshallian (uncompensated) labor supply, as a function of the net wage $\tilde{w}$ and non-labor income y. Note that these are themselves functions of h and m. Then, totally differentiating

$$h = \sigma(\tilde{w}(h,m), y(h,m))$$

yields

$$dh = \frac{\partial \sigma}{\partial \tilde{w}}\left[\frac{\partial \tilde{w}}{\partial h}dh + \frac{\partial \tilde{w}}{\partial m}dm\right] + \frac{\partial \sigma}{\partial y}\left[\frac{\partial y}{\partial h}dh + \frac{\partial y}{\partial m}dm\right].$$

Rearranging gives

$$\mathcal{H}_m = \frac{dh}{dm} = \frac{\sigma_{\tilde{w}}\frac{\partial \tilde{w}}{\partial m} + \sigma_y \frac{\partial y}{\partial m}}{1 - \sigma_{\tilde{w}}\frac{\partial \tilde{w}}{\partial h} - \sigma_y \frac{\partial y}{\partial h}} = \frac{\epsilon_{h,\tilde{w}}\frac{h}{\tilde{w}}\frac{\partial \tilde{w}}{\partial m} + \epsilon_{h,y}\frac{h}{y}\frac{\partial y}{\partial m}}{1 - \epsilon_{h,\tilde{w}}\frac{h}{\tilde{w}}\frac{\partial \tilde{w}}{\partial h} - \epsilon_{h,y}\frac{h}{y}\frac{\partial y}{\partial h}}, \tag{A.3}$$

where the denominator is positive.

Next, note that with a total income tax,

$$(1-\tau)F_h = \left(1 - \frac{g - (F_m - \kappa)m}{F_k k + F_h h}\right)F_h = \left(\frac{F - g - \kappa m}{F_k k + F_h h}\right)F_h = \frac{\gamma_h h^{-1}(F - g - \kappa m)}{\gamma_k + \gamma_h},$$

where $F_k k = \gamma_k F$, $F_k k = \gamma_k F$, $F_k k = \gamma_k F$ and $\gamma_k + \gamma_h + \gamma_m = 1$ by Euler's theorem. Similarly,

$$y = (1-\tau)F_k k = \frac{\gamma_k (F - g - \kappa m)}{\gamma_k + \gamma_h}.$$

Assume that $\frac{\gamma_k}{\gamma_k + \gamma_h}$ is constant, which is the case for a many CES functions. Then,

$$\frac{\partial \tilde{w}}{\partial m} = \frac{\partial (1-\tau)F_h}{\partial m} = \frac{\gamma_h h^{-1}}{\gamma_k + \gamma_h}(F_m + F_s q_m - \kappa m) \tag{A.4}$$

and

$$\frac{\partial y}{\partial m} = \frac{\partial(1-\tau)F_k k}{\partial m} = \frac{\gamma_k}{\gamma_k+\gamma_h}(F_m + F_s q_m - \kappa m) \tag{A.5}$$

Substituting (A.4) and (A.5) in (A.3) yields:

$$\mathcal{H}_m = \frac{\epsilon_{h,\tilde{w}}\frac{\gamma_h}{\gamma_k+\gamma_h} + \epsilon_{h,y}\frac{\tilde{w}h}{y}\frac{\gamma_k}{\gamma_k+\gamma_h}}{\tilde{w} - \epsilon_{h,\tilde{w}}h\frac{\partial\tilde{w}}{\partial h} - \epsilon_{h,y}\frac{\tilde{w}h}{y}\frac{\partial y}{\partial h}}(F_m + F_s q_m - \kappa m) = \frac{\epsilon_{h,\tilde{w}} + \epsilon_{h,y}}{A}\frac{\gamma_h}{\gamma_k+\gamma_h}(F_m + F_s q_m - \kappa m)$$

Note that in the case without capital and hence exogenous non-labor income, this becomes

$$\mathcal{H}_m = \frac{\epsilon_{h,\tilde{w}}}{\tilde{w} - \epsilon_{h,\tilde{w}}\tau F_h}(F_m + F_s q_m - \kappa m).$$

A.2 Two-period Model

Assume that the stock of the pollutant s evolves in the following way:

$$s_0 = q_0(m_0), \quad s_1 = q(s_0, m_1). \tag{A.6}$$

That is, s_t denotes the pollutant stock *at the end* of period t.

Similarly in the static setting above, the environmental wedge in period 0 is defined as the sum of the direct marginal benefit and cost of pollution, now taking into account stock effects:

$$\omega_{Env} \equiv u_c(F_m - \kappa) + q_m[u_s + u_c F_s] + \beta q_m q_s' \left[u_s' + u_c' F_s'\right] = 0. \tag{A.7}$$

I solve the model using backwards induction. For simplicity, assume that there are separate tax rates feasible on labor and capital income in the final period. The government's problem in period 1 can then be written compactly as:

$$\max_{h_1,m_1} u[F(k_1, h_1, m_1, q(s_0, m_1)) - g_1, 1 - h_1, q(s_0, m_1)]. \tag{A.8}$$

Note that this is identical to the problem of a social planner that takes the capital stock and the stock of the pollutant as given. Although the government does raise a tax on capital income, this tax is not distortionary, but equivalent to imposing a lump-sum tax: in period 1, the investment decision has been made and the capital stock is sunk. Due to the presence of this lump-sum tax, the labor tax is optimally zero. Thus, for a given capital stock, the optimal pollution tax in $t = 1$ is determined as in first-best such that $\omega_{Env} = 0$.

Let $M(k_1, s_0)$ and $H(k_1, s_0)$ denote the policy rules that solves the government's problem in period 1. Since government expenditures are financed by the revenue generated from the tax on capital income and from the pollution tax, hence the tax rate on capital in $t = 1$ is given by

$$T(k_1, s_0) = \frac{g_1 - F_m[k_1, M(k_1, s_0), H(k_1, s_0), q(s_0, M(k_1, s_0))]}{F_k[k_1, M(k_1, s_0), H(k_1, s_0), q(s_0, M(k_1, s_0))]k_1}. \tag{A.9}$$

I assume that this tax is always positive, i.e. government expenditures are large enough so that they cannot be financed by the revenue from the emission tax alone. To simplify notation, let $M(k_1, s_0) = M_1$, $H(k_1, s_0) = H_1$ and $T(k_1, s_0) = T_1$.

The government in period 0 takes these function as given when solving its problem[50]It then solves the following problem:

$$\max_{k_1,m_0,h_0} u[F(k_0, h_0, m_0, q_0(m_0)) - \kappa m_0 - g_0 - k_1, 1 - h_0, q_0(m_0)] \\ + \beta u\left[F(k_1, H_1, M_1, q(q_0(m_0), M_1)) - \kappa M_1 - g_1, 1 - H_1, q(q_0(m_0), M_1)\right]$$

subject to implementability constraints. Assuming that the government imposes a tax on total income in period 0, and hence $\omega_{LL} > 0$, there are two such constraints[51]. The first one comes from the intratemporal optimality condition and is analogous to the one-period model, but with savings k_1 as an additional argument. Hence, it can be compactly written as $\chi(k_1, h_0, m_0) = 0$. The second one combines the household's budget constraint with the household's intertemporal optimality condition:

$$0 \geq \beta u_c\big(F_h(1)H_1 + (1 - T_1)F_k(1)k_1, 1 - H_1, q(q_0(m_0), M_1)\big) \\ \cdot \big(F_k(k_1, H_1, M_1, q(q_0(m_0), M_1))(1 - T_1\big) - u_c\big(F_h(0)h_0 + (1 - \tau_0)F_k(0)k_0, 1 - h_0, q_0(m_0)\big).$$

Using the definition of T_1 above, and a similar definition for the tax in period 0, one can rewrite this as:

$$0 \geq \beta u_c\big(F(k_1, H_1, M_1, q(q_0(m_0), M_1)) - g_1 - \kappa M_1, 1 - H_1, q(q_0(m_0), M_1)\big) \\ \cdot \big(F_k(1)(1 - T_1\big) - u_c\big(F(0) - g_0 - \kappa m_0 - k_1, 1 - h_0, q_0(m_0)\big) \equiv \eta(k_1, h_0, m_0), \tag{A.10}$$

where I have omitted the arguments of the production function.

Let $\mathcal{H}(m)$ and $\mathcal{K}(m)$ denote the decision rules implicitly defined by the solutions to the two constraints:

$$\epsilon(\mathcal{K}(m), \mathcal{H}(m), m) = 0 \tag{A.11}$$

$$\eta(\mathcal{K}(m), \mathcal{H}(m), m) = 0. \tag{A.12}$$

Similar to $\mathcal{H}$, $\mathcal{K}$ can be interpreted as the household's response function to the current fossil fuel choice. Then, the government's problem can be rewritten as

$$\max_{m_0} u[F(k_0, m_0, \mathcal{H}(m_0), q_0(m_0)) - \kappa m_0 - g_0 - \mathcal{K}(m_0), 1 - \mathcal{H}(m_0), q_0(m_0)] \\ + \beta u\left[F(\mathcal{K}(m_0), M_1, H_1, q(q_0(m_0), M_1) - \kappa M_1 - g_1, q(q_0(m_0), 1 - H_1, M_1\right]. \tag{A.13}$$

where $M_1 = M(\mathcal{K}(m_0), s_0)$ etc. Taking derivatives with respect to m and using the envelope theorem gives the generalized Euler equation (3.8).

[50]In game-theoretic terms, these are the response functions of the follower (the future government) in this sequential game, which the leader (the current government) must take into account when optimizing.

[51]If instead there was a separate tax on labor and capital income, it would follow that $\omega_{LL} = 0$, and there would be only one distortion and hence implementability constraint.

A.3 Derivation of the GEE in the Infinite-horizon Model

Let the equilibrium policy functions for saving, hours worked, fuel use and government expenditures be denoted by $n^k(k,s)$, $n^h(k,s)$, $\psi(k,s)$ and $\phi(k,s)$, respectively.

Define the residual functions that capture the household's optimality conditions when set to zero:

$$\begin{aligned}&\eta(k,s,g,m,k',h) = u_c[\mathcal{C}(k,k',h,g,m,s),1-h,g,s]\\&-\beta u_c'[\mathcal{C}(k',n^k(k',s'),p(k',s'),\psi(k',s'),\phi(k',s'),s'),1-n^h(k',s'),\psi(k',s'),s']\\&\cdot\left\{1+[1-T(k',n^h(k',s'),\psi(k',s'),\phi(k',s'),s')][F_k(k',n^h(k',s'),\phi(k',s'),s')-\delta]\right\},\end{aligned} \tag{A.14}$$

and

$$\epsilon(k,s,g,m,k',h) = \frac{u_l[\mathcal{C}(k,k',h,g,m,s),1-h,g,s]}{u_c[\mathcal{C}(k,k',h,g,m,s),1-h,g,s]} - F_h(k,h)[1-T(k,h,g,m,s)]. \tag{A.15}$$

That is, $\eta(k,s,g,m,k',h)=0$ gives the household's Euler equation and $\epsilon(k,s,g,m,k',h)=0$ the household's labor-leisure condition.

Define functions $\mathcal{K}(k,s,g,m)$ and $\mathcal{H}(k,s,g,m)$ implicitly by

$$\eta(k,s,g,m,\mathcal{K}(k,s,g,m),\mathcal{H}(k,s,g,m))=0 \tag{A.16}$$
$$\epsilon(k,s,g,m,\mathcal{K}(k,s,g,m),\mathcal{H}(k,s,g,m))=0. \tag{A.17}$$

Using those functions and

$$\mathcal{C}(k,s,k',h,g,m) = F(k,h,m,s)+(1-\delta)k-g-k',$$

write the government's problem in t compactly as:

$$\max_{k',s',h,g,m} u(\mathcal{C}(k,s,k',h,g,m),1-h,g,s)+\beta v(k',s'),$$

s.t.

$$\begin{aligned}s' &= q(s,m)\\ k' &= \mathcal{K}(k,s,g,m)\\ h &= \mathcal{H}(k,s,g,m).\end{aligned}$$

or

$$\max_{g,m} u[\mathcal{C}(k,s,\mathcal{K}(k,s,g,m),\mathcal{H}(k,s,g,m),g,m),1-\mathcal{H}(k,s,g,m),g,s]+\beta v[\mathcal{K}(k,s,g,m),q(s,m)],$$

where

$$\begin{aligned}v(k,s) =& u[\mathcal{C}(k,h(k,s),n^h(k,s),\psi(k,s),\phi(k,s),s),1-n^h(k,s),\psi(k,s),s]\\&+\beta v(n^k(k,s),q(s,\phi(k,s))).\end{aligned} \tag{A.18}$$

Taking f.o.c. yields:

$$-u_c(\mathcal{K}_g - F_h\mathcal{H}_g + 1) - u_l\mathcal{H}_g + u_g + \beta v_k'\mathcal{K}_g = 0 \tag{A.19}$$

$$-u_c[\mathcal{K}_m - F_h\mathcal{H}_m - (F_m - \kappa)] - u_l\mathcal{H}_m + u_m + [\beta v_k'\mathcal{K}_m + v_s' q_m] = 0. \tag{A.20}$$

.

The derivatives of the value function $v(k,s)$ with respect to k and s read:

$$v_k = u_c[F_k + 1 - \delta - \mathcal{K}_k + F_h\mathcal{H}_k] - u_l\mathcal{H}_k + \beta v_k'\mathcal{K}_k, \tag{A.21}$$

and

$$v_s = u_c[F_s - \mathcal{K}_s + F_h\mathcal{H}_s] - u_l\mathcal{H}_s + u_s + \beta[v_k'\mathcal{K}_s + v_s' q_s]. \tag{A.22}$$

Note that at this point, one can see why I need a second *endogenous* public good in addition to environmental quality. If government expenditures were exogenous, I would have only one first-order condition above, namely (A.20). However, in order to substitute for v_k' and v_s' in (A.21) and (A.22), I need two equations; (A.20) alone would not be sufficient. From (A.19) and (A.20):

$$\beta v_k' = \frac{1}{\mathcal{K}_g}\left[u_c(\mathcal{K}_g - F_h\mathcal{H}_g + 1) + u_l\mathcal{H}_g - u_g\right] \tag{A.23}$$

$$\begin{aligned}\beta v_s' &= \frac{1}{q_m}\left[u_c(\mathcal{K}_m - F_h\mathcal{H}_m - (F_m - \kappa)) + u_l\mathcal{H}_m - \frac{\mathcal{K}_m}{\mathcal{K}_g}\left[u_c(\mathcal{K}_g - F_h\mathcal{H}_g + 1) + u_l\mathcal{H}_g - u_g\right]\right] \\ &= -\frac{1}{q_m}\left[u_c(F_h\mathcal{H}_m + (F_m - \kappa)) - u_l\mathcal{H}_m + \frac{\mathcal{K}_m}{\mathcal{K}_g}\left[u_c(1 - F_h\mathcal{H}_g) + u_l\mathcal{H}_g - u_g\right]\right]\end{aligned} \tag{A.24}$$

Inserting (A.23) in (A.21) and updating gives:

$$v_k' = u_c'(F_k' + 1 - \delta) + \mathcal{H}_k'(u_c'F_h' - u_l') - \frac{\mathcal{K}_k'}{\mathcal{K}_g'}\left[u_g' - u_c' + \mathcal{H}_g'(u_c'F_h' - u_l')\right] \tag{A.25}$$

Similarly, (A.24) in (A.22) yields:

$$\begin{aligned}v_s' =& u_s' + u_c'F_s' - \frac{q_s'}{q_m'}u_c'(F_m' - \kappa) + (u_g' - u_c')\left[-\frac{\mathcal{H}_s'}{\mathcal{H}_g'} + \frac{q_s'}{q_m'}\frac{\mathcal{H}_m'}{\mathcal{H}_g'}\right] \\ &+ (u_c'F_h' - u_l')\left[\mathcal{H}_s' - \frac{\mathcal{K}_s'}{\mathcal{K}_g'}\mathcal{H}_g' - \frac{q_s'}{q_m'}\mathcal{H}_m' + \frac{q_s'}{q_m'}\frac{\mathcal{K}_m'}{\mathcal{K}_g'}\mathcal{H}_g'\right]\end{aligned} \tag{A.26}$$

Finally, substituting (A.25) for v_k' in (A.19) gives:

$$\begin{aligned}&u_g - u_c + \mathcal{H}_g(u_cF_h - u_l) + \mathcal{K}_g[-u_c + \beta u_c'(F_k' + 1 - \delta)] \\ &+ \beta\mathcal{K}_g\left\{\mathcal{H}_k'(F_h'u_c' - u_l') - \frac{\mathcal{K}_k'}{\mathcal{K}_g'}\left[\mathcal{H}_g'(F_h'u_c' - u_l') + u_g' - u_c'\right]\right\} = 0.\end{aligned}$$

Moreover, substituting (A.26) for v_s' in (A.20) yields:

$$\begin{aligned}
&u_c(F_m-\kappa)+\beta q_m\left[u_s'+u_c'F_s'-\frac{q_s'}{q_m'}u_c'(F_m'-\kappa)\right]+\mathcal{H}_m(u_cF_h-u_l)\\
&+\mathcal{K}_m[-u_c+\beta u_c'(F_k'+1-\delta)]+\beta(u_g'-u_c')\left[-\mathcal{K}_m\frac{\mathcal{K}_k'}{\mathcal{K}_g'}-q_m\frac{\mathcal{K}_s'}{\mathcal{K}_g'}+q_m\frac{q_s'}{q_m'}\frac{\mathcal{K}_m'}{\mathcal{K}_g'}\right]\\
&+\beta\mathcal{K}_m(u_c'F_h'-u_l')\left[\mathcal{H}_k'-\frac{\mathcal{K}_k'}{\mathcal{K}_g'}\mathcal{H}_g'\right]+\beta(u_c'F_h'-u_l')q_m\left[\mathcal{H}_s'-\frac{\mathcal{K}_s'}{\mathcal{K}_g'}\mathcal{H}_g'-\frac{q_s'}{q_m'}\mathcal{H}_m'+\frac{q_s'}{q_m'}\frac{\mathcal{K}_m'}{\mathcal{K}_g'}\mathcal{H}_g'\right]=0
\end{aligned}$$

Using the wedges defined above, these two equation can be written as

$$\omega_{PG}+\mathcal{K}_g\omega_{CS}-\beta\mathcal{K}_g\frac{\mathcal{K}_k'}{\mathcal{K}_g'}\omega_{PG}'+\mathcal{H}_g\omega_{LL}+\beta\omega_{LL}'\mathcal{K}_g\left[\mathcal{H}_k'-\frac{\mathcal{K}_k'}{\mathcal{K}_g'}\mathcal{H}_g'\right]=0, \tag{A.27}$$

and

$$\begin{aligned}
&\omega_{Env}+\mathcal{K}_m\omega_{CS}-\beta\omega_{PG}'\mathcal{K}_m\frac{\mathcal{K}_k'}{\mathcal{K}_g'}-\beta\omega_{PG}'\left[q_m\frac{\mathcal{K}_s'}{\mathcal{K}_g'}-q_m\frac{q_s'}{q_m'}\frac{\mathcal{K}_m'}{\mathcal{K}_g'}\right]\\
&\quad+\mathcal{H}_m\omega_{LL}+\beta\omega_{LL}'\mathcal{K}_m\left[\mathcal{H}_k'-\frac{\mathcal{K}_k'}{\mathcal{K}_g'}\mathcal{H}_g'\right]\\
&\quad+\beta\omega_{LL}'q_m\left[\mathcal{H}_s'-\frac{\mathcal{K}_s'}{\mathcal{K}_g'}\mathcal{H}_g'-\frac{q_s'}{q_m'}\mathcal{H}_m'+\frac{q_s'}{q_m'}\frac{\mathcal{K}_m'}{\mathcal{K}_g'}\mathcal{H}_g'\right]=0.
\end{aligned} \tag{A.28}$$

A.4 Derivation of the tax-interaction effect

In this section, I derive the explicit version of the environmental GEE, given by (3.30), from the implicit version (3.27). To keep the notation simple, start by defining $X(t)$ as the sum of the current second-best benefit and cost of pollution; that is,

$$\begin{aligned}
X(t)=&\mathcal{K}_m\omega_{CS}-\beta\omega_{PG}'\mathcal{K}_m\frac{\mathcal{K}_k'}{\mathcal{K}_g'}-\beta\omega_{PG}'\left[q_m\frac{\mathcal{K}_s'}{\mathcal{K}_g'}-q_m\frac{q_s'}{q_m'}\frac{\mathcal{K}_m'}{\mathcal{K}_g'}\right]\\
&+\mathcal{H}_m\omega_{LL}+\beta\omega_{LL}'\mathcal{K}_m\left[\mathcal{H}_k'-\frac{\mathcal{K}_k'}{\mathcal{K}_g'}\mathcal{H}_g'\right]\\
&+\beta\omega_{LL}'q_m\left[\mathcal{H}_s'-\frac{\mathcal{K}_s'}{\mathcal{K}_g'}\mathcal{H}_g'-\frac{q_s'}{q_m'}\mathcal{H}_m'+\frac{q_s'}{q_m'}\frac{\mathcal{K}_m'}{\mathcal{K}_g'}\mathcal{H}_g'\right].
\end{aligned}$$

Then, (3.27) can be written as $\omega_{Env}(t)+X(t)=0$. Using the definition of ω_{Env}, this implies that in period $t+1$,

$$\begin{aligned}
u_c(t+1)(F_m(t+1)-\kappa)=&-X(t+1)-\beta q_m(t+1)\cdot\\
&\left[u_s(t+2)+u_c(t+2)F_s(t+2)-\frac{q_s(t+2)}{q_m(t+2)}u_c(t+2)(F_m(t+2)-\kappa)\right].
\end{aligned}$$

Inserting this repeatedly in the GEE (3.27) and assuming that $q_m(t+1) = q_m(t+2) = ... = q_m(t)$ gives

$$\begin{aligned}\omega_{Env}(t) + X(t) &= u_c(t)(F_m(t) - \kappa) + \sum_{i=t+1}^{\infty} \beta^{i-t-1} \left(\prod_{j=t+2}^{i} q_s(j) \right) [q_m(t)\beta(u_s(i) + u_c(i)F_s(i)) + X(i)] \\ &= u_c(t)(F_m(t) - \kappa) + q_m(t) \sum_{i=t+1}^{\infty} \beta^{i-t} \left(\prod_{j=t+2}^{i} q_s(j) \right) (u_s(i) + u_c(i)F_s(i)) \\ &\quad + \sum_{i=t+1}^{\infty} \beta^{i-t-1} \left(\prod_{j=t+2}^{i} q_s(j) \right) X(i) = 0.\end{aligned}$$

A.5 Second-best with commitment

Once more, I assume that the government is restricted to impose a total income tax. sIn the commitment case, if it could instead impose separate tax rates on income from labor and capital, the seminal work by Judd (1985) and Chamley (1986) has shown that capital taxes are zero in the long run (steady state). Moreover, if per-period utility is separable in consumption and leisure, and has a constant intertemporal elasticity of substitution with respect to consumption, capital taxes are zero as of the second period.

The household takes the sequences of before-tax factor prices and taxes as given. As before, solving its problem gives rise to the usual optimality conditions:

$$u_c(t) - \beta u_c(t+1)[1 + (1 - \tau_{t+1})(r_{t+1} - \delta)] = 0 \tag{A.29}$$

$$u_l(t) - u_c(t)(1 - \tau_t)w_t = 0.. \tag{A.30}$$

The tax rate is given by the government's budget constraint:

$$\tau_t = \frac{g_t - (F_m(k_t, h_t, m_t, T_t) - \kappa)m_t}{(F_k(k_t, h_t, m_t, T_t) - \delta)k_t + F_h(k_t, h_t, m_t, T_t)h_t}$$

With $c_t = \mathcal{C}(k_t, k_{t+1}, h_t, g_t, m_t, T_t)$ from the economy's resource constraint in (3.17), (A.29) and (A.30) can then be written as:

$$\tilde{\eta}(k_t, k_{t+1}, k_{t+2}, T_t, T_{t+1}, g_t, g_{t+1}, m_t, m_{t+1}, h_t, h_{t+1}) = 0 \tag{A.31}$$

$$\tilde{\epsilon}(k_t, k_{t+1}, T_t, g_t, m_t, h_t) = 0 \tag{A.32}$$

The government's problem can be written as

$$\max_{c_t, k_{t+1}, T_{t+1}, h_t, g_t, m_t, \tau_t} \sum_{t=0}^{\infty} \beta^t u(c_t, 1 - h_t, g_t, T_t)$$

s.t.

$$F(k_t, h_t, m_t, T_t) + (1 - \delta)k_t - c_t - g_t - \kappa m_t - k_{t+1} = 0 \tag{A.33}$$

$$T_{t+1} \geq q(T_t, m_t) \tag{A.34}$$

as well as (A.31) and (A.32). That is, the government maximizes its objective function, lifetime utility, subject to the resource constraint, its budget constraint and the household's optimality conditions.

A.6 The Government's Problem in a Global Finite-Horizon Economy

A.6.1 Problem

Let J denote the final period with carbon emissions. For period $t < J$, given next period's policy functions ψ_{t+1}, ϕ_{t+1}, n^k_{t+1} and n^h_{t+1}, and the value function v_{t+1}, the government in t solves the following problem:

$$\max_{k_{t+1},T_{t+1},h_t,g_t,m_t} u(\mathcal{C}(k_t,k_{t+1},h_t,g_t,m_t,T_t),1-h_t,g_t,T)+\beta v_{t+1}(k_{t+1},T_{t+1}), \tag{A.35}$$

subject to

$$T_{t+1}=q(T_t,m_t)$$

$$\frac{u_l\big(\mathcal{C}(k_t,k_{t+1},h_t,g_t,m_t,T_t),1-h_t,g_t,T_t\big)}{u_c\big(\mathcal{C}(k_t,k_{t+1},h_t,g_t,m_t,T_t),1-h_t,g_t,T_t\big)}-F_h(t)[1-\mathcal{T}(k_t,h_t,g_t,m_t,T_t)]=0,$$

and

$$\begin{aligned}
&u_c\big(\mathcal{C}(k_t,k_{t+1},h_t,g_t,m_t,T_t),1-h_t,g_t,T_t\big)\\
&-\beta u_c'\left(\mathcal{C}\left(\begin{array}{c} k',n^k_{t+1}(t+1),n^h_{t+1}(t+1),\\ \psi_{t+1}(t+1),\phi_{t+1}(t+1),T_{t+1},\end{array}\right),1-n^h_{t+1}(t+1),\psi_{t+1}(t+1),T_{t+1}\right)\cdot\\
&\cdot\left[1+\left[1-\mathcal{T}\left(\begin{array}{c} k_{t+1},n^h_{t+1}(t+1),\psi_{t+1}(t+1),\\ \phi_{t+1}(t+1),T_{t+1}\end{array}\right)\right][F_k(k_{t+1},n^h_{t+1}(t+1),\phi_{t+1}(t+1),T_{t+1})-\delta]\right]=0,
\end{aligned}$$

where $n^k_{t+1}(t+1)=n^k_{t+1}(k_{t+1},T_{t+1})$ etc. Moreover,

$$\mathcal{C}(k_t,k_{t+1},h_t,g_t,m_t,T_t)=F(k_t,h_t,m_t,T_t)+(1-\delta)k_t-g_t-\kappa m_t-k_{t+1},$$

and

$$\mathcal{T}(k_t,h_t,g_t,m_t,T_t)=\frac{g-(F_m(k_t,h_t,m_t,T_t)-\kappa)m_t}{(F_k(k_t,h_t,m_t,T_t)-\delta)k_t+F_h(k_t,h_t,m_t,T_t)h_t}.$$

The solution to this problem gives the contemporaneous policy functions ψ_t, ϕ_t, n^k_t and n^h_t, as well as the value function v_t:

$$\begin{aligned}
v_t(k_t,T_t)=&u\left(\begin{array}{c}\mathcal{C}\big(k_t,n^k_t(k_t,T_t),n^h_t(k_t,T_t),\psi_t(k_t,T_t),\phi_t(k_t,T_t),T_t\big),\\ 1-n^h_t(k_t,T_t),\psi_t(k_t,T_t),T_t\end{array}\right)\\
&+\beta v_{t+1}\big(n^k_t(k_t,T_t),q(T_t,\phi_t(k_t,T_t))\big).
\end{aligned}$$

A.6.2 Algorithm

This algorithm for finite-horizon value function iteration follows Cai et al. (2012).

Initialization

- For each period $t \leq J+1$, choose m_t^k approximation nodes for the capital stock and m_t^T nodes for temperature change:

$$X_t^k = \{x_{it}^k : 1 \leq i \leq m_t^k\},\ \ X_t^T = \{x_{it}^T : 1 \leq i \leq m_t^T\}.$$

  Here, I solve the first-best problem sequentially using KNITRO, and then choose Chebyshev approximation nodes from an interval around the first-best outcome for k and T, i.e. $x_{1t}^k = a_{min}^k k_t$ and $x_{mt}^k = a_{max}^k k_t$, with $a_{min}^k < 1$ and $a_{max}^k > 1$.

- Choose a functional form for approximating value and policy functions. Here, I apply Chebyshev polynomial approximation (Judd, 1998; Cai et al., 2012), using complete Chebyshev polynomials as basis functions. That is, for a function $v(k,T)$, the degree-n complete Chebyshev approximation is given by $P_n(k,T;\mathbf{b})$, where $\mathbf{b}$ is a vector of coefficients.

Step 1 - Continuation Value

a. Maximization step. For any $(x_{i,J+1}^k, x_{j,J+1}^T)$, $1 \leq i \leq m_{J+1}^k$, $1 \leq j \leq m_{J+1}^T$ solve the stationary problem in period $J+1$:

$$\max_{\{c_t^{i,j}, h_t^{i,j}, k_{t+1}^{i,j}, g_t^{i,j}, m_t^{i,j}\}_{t=J+1}^{\infty}} \sum_{t=J+1}^{\infty} \beta^{t-J+1} u(c_t^{i,j}, 1-h_t^{i,j}, g_t^{i,j}, x_{j,J+1}^T)$$

s.t.

$$c_t^{i,j} + g_t^{i,j} + k_{t+1}^{i,j} + \kappa m_t^{i,j} = F(k_t^{i,j}, h_t^{i,j}, m_t^{i,j}, x_{j,J+1}^T) + (1-\delta)k_t^{i,j},$$

and $k_{J+1}^{i,j} = x_{i,J+1}^k$. The continuation value is given by:

$$v_{J+1}^{i,j} = \sum_{t=J+1}^{\infty} \beta^{t-J+1} u(c_t^{i,j}, 1-h_t^{i,j}, g_t^{i,j}, x_{j,J+1}^T). \tag{A.36}$$

b. Fitting step. Approximate $v_{J+1}(k,T)$ with $P_n(k,T;\mathbf{b}_{J+1})$. That is, compute $\mathbf{b}_{J+1}$ such that $P_n(x_{i,J+1}^k, x_{j,J+1}^T; \mathbf{b}_{J+1})$ approximates $v_{J+1}^{i,j}$.

Step 2 - Backwards Induction For $t = J, J-1, ..., 0$, iterate through the following steps.

a. Maximization step. For any $(x_{i,t}^k, x_{j,t}^T)$, $1 \leq i \leq m_t^k$, $1 \leq j \leq m_t^T$, solve the maximization problem in (A.35), given the approximated policy functions and the value function in $t+1$.

b. Fitting step. Choose coefficients to approximate

$$
\begin{aligned}
v_t(k,T) &\approx P_n(k,T;b_t^v) \\
\psi_t(k,T) &\approx P_n(k,T;b_t^\psi) \\
\phi_t(k,T) &\approx P_n(k,T;b_t^\phi) \\
n_t^k(k,T) &\approx P_n(k,T;b_t^k) \\
n_t^h(k,T) &\approx P_n(k,T;b_t^h)
\end{aligned}
$$

B Appendix B

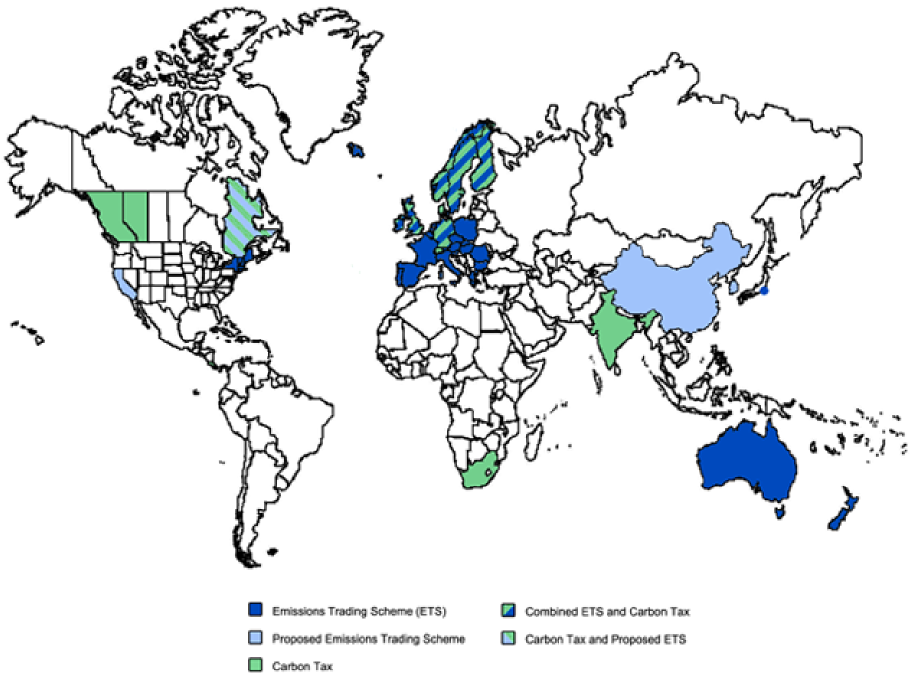

Figure 5: Source: EESI (2012)