Optimal taxes on fossil fuel in general equilibrium

Michael Golosov, John Hassler, Per Krusell, and Aleh Tsyvinski*

Abstract

We embed a simple linear model of the carbon cycle in a standard neoclassical growth model where one input to the production function, oil, is non-renewable. The use of oil generates carbon emission, the key input in the carbon cycle. Changes in the amount of carbon in the atmosphere drive the greenhouse effect and thereby the climate. Climate change is modeled as a global damage to production and is a pure externality.

We solve the model for both the decentralized equilibrium with taxes on oil and for the optimal allocation. The model is then used to find optimal tax and subsidy polices. A robust model finding is that constant taxes on oil have no effect on the allocation: only time-varying taxes do. A key finding is that optimal ad valorem taxes on oil consumption should fall over time. In the simplified version of the model, optimal taxes per unit of oil should be indexed to GDP. A calibrated, less simplified model also generates declining, and initially rather substantial, taxes on oil.

*Golosov: MIT; Hassler: IIES; Krusell: IIES and Princeton; Tsyvinski: Yale. The paper is preliminary. Please do not distribute without author approval.

1 Introduction

In this paper we propose a global economy-climate model where taxes, or some other form of government policy, are called for in order to limit the negative impacts of the economy on our climate. The main goal of the paper is thus to use this model to characterize optimal global policy qualitatively and quantitatively. The background for the work and for our particular approach is that there now is widespread consensus that human activity is an important driver of climate change. First, when fossil fuel is burned, carbon (dioxide) is emitted, and through the carbon cycle this carbon leads to increasing atmospherical carbon concentrations. Second, these higher concentrations influence the global temperature, which in turn is a key determinant of our climate. Third, the direct and indirect damages to humans are largely caused not by higher average temperature but by extreme weather outcomes, such as droughts, floods, and storms, but these extreme outcomes are much more frequent at higher average global temperatures. Of course, some of these damages then in turn influence production and thus energy use: there is two-way economy-climate feedback. However, in typical climate projections like those issued by the IPCC, the two-way feedback is not taken into account; there, one takes a "scenario" for energy use as given without asking how it in turn would influence the economy. In the climate-economy model used in the present work, both energy use and climate outcomes are endogenous, and thus any energy projections coming out of the model are consistent with the model simulation of climate damages.

Any emission of carbon adds to a *global* stock of carbon in the atmosphere and it is the *global* concentration that determines *global* temperature. *Local* climates around the world, on the other hand, are a function of geophysical characteristics, i.e., primarily economy-independent factors, and of *global* temperature. This means that when someone burns oil in Uleåborg, to the extent there is an externality, it is global in nature. Thus, a study of the effect of the economy on the climate must involve a study of the global system with a pure externality. The global economy-climate model that we construct in this paper is a natural extension of non-renewable resource models along the lines of Dasgupta and Heal (1974) to include a climate externality and a carbon cycle. Quite importantly, our model is also an extension in that we study a global competitive equilibrium with an externality, allowing us to discuss explicitly, with standard welfare analysis, how economic policy could and should be used to correct this externality. The prime purpose of the paper is indeed to characterize optimal energy taxes in the global decentralized equilibrium economy.

We have two main, and related, analytical results. First, we show that an energy tax, no matter how high, that is constant over time is under quite general conditions ineffective for correcting the externality. The easiest way to understand the intuition for this result is to imagine a static economy. For simplicity, assume that there is a stock of oil R that will all be used up, i.e., that the extraction costs are zero for the resource. Then in a static economy, because all the oil will be used up, taxes cannot influence energy use, so all a tax does is alter the pre-tax price so that the post-tax price equals the marginal product of energy evaluated at R. Thus, taxes may redistribute resources between producers and consumers but will not influence energy use. More importantly, however, this result extends to a dynamic economy. If, namely, the total oil amount R is all used up over time in the dynamic economy as well, then the only issue is *when* it will be used up, and since a constant tax precisely will not influence the intertemporal incentives, oil use is not influenced at all by such a tax. This is, moreover, unfortunate, because the timing of when oil is used up turns out to be quite important for welfare.

The second result, then, more precisely characterizes what the optimal tax sequence must look like. Again under assumptions such that the resource will be used up entirely, we provide an optimal-tax formula that allows a world planner to attain the first best. Thus, the optimal tax sequence can be solved for given the first-best allocation. Our model has a dynastic household, so it is straightforward to first solve for the optimal allocation—numerically, if need be—and then simply generate the implied optimal tax sequence. The tax formula we provide applies under rather general assumptions for technology and for a somewhat special though, we think, flexible enough specification for the climate damage, and in order to obtain more specific characterization of the time path of the taxes on oil one needs to make more specific assumptions.

Thus, for particular technological assumptions that seem reasonable, we can establish that optimal tax rates on oil should be decreasing. Two main forces are behind this result. One is that the size of the tax should be in proportion to the size of the marginal externality, which is determined by carbon concentration in the atmosphere now and in the future (since energy use now generates increased carbon concentrations in the future). Since oil use must eventually be declining over time, as long as there is some mean reversion in carbon concentration—so that carbon concentration is not simply equal to the total of what was emitted in the past—the size of the marginal externality will (eventually, at least) be declining over time. Second, even if the externality cost were constant over time it would be beneficial to postpone the cost, as long as there is positive discounting. Thus, when the private value of oil is higher than its social value by a constant, it is beneficial to postpone extraction, thus calling for lower future tax rates on oil.

Aside from analytical characterization we also specify functional forms, calibrate the key parameters, and solve the model numerically for an optimal tax sequence. The results are quite striking. Optimal taxes should start out very high, in our benchmark at about 80% of the oil price, and decline slowly to about 50% in 200 years to reach close to zero in 500 years. The implied prescription is a drastic reduction in energy use now: by about 50%. An increase in global temperatures over the next 100 to 150 years is unavoidable given the current carbon concentration but the optimal policy will shave off more than a degree in temperature increase during the transition relative to the laissez-faire outcome, and this one degree is important. Since this would amount to a rather a gigantic immediate impact, what would the immediate costs be? We find them to be around 2% of GDP. However, output grows faster in the optimal allocation than in *laissez-faire* and overtakes the latter in around 50 years. In the long-run, output is around 5% higher in the optimal allocation due to *higher long*-run fossil fuel use.

The pioneering work in this area is due to William Nordhaus; for a nice recent description of his modeling, see Nordhaus and Boyer (2000). Nordhaus's main framework is a computational model called RICE—Regional dynamic Integrated model of Climate and the

Economy—and it is similar in spirit to the one we use here. Nordhaus's work is particularly pioneering in two areas. First, informed by modern climate and carbon-cycle modeling, he managed to summarize the key quantitative channels from the economy to the climate with a rather parsimonious, and mostly linear, dynamic system. This dynamic system is small enough that it can be embedded in a typical dynamic growth model. Second, Nordhaus did extensive work aimed at summarizing the damages from climate change. His modeling of these damages in RICE amounts to a multiplicative term on aggregate production which is a function of the average global temperature; thus, he lumps together damages of various sorts as a production loss measure. His formulation captures increasing marginal costs of temperature increase. In the present paper, we essentially use Nordhaus's formulations both for the carbon cycle and the climate system and for the damages. We make some simplifications, however. First, we simplify the carbon cycle somewhat, thus ignoring the distinction between different carbon deposits. Second, we abstract from the dynamics of the temperature in the oceans (which has a separate dynamic impact on global atmospheric temperature). However, overall we calibrate our simplified system to roughly match Nordhaus's system. Our treatment of damages uses Nordhaus's formulation directly. Nordhaus, moreover, has other model features that we ignore; his RICE model has eight regions, for example, and he includes trade in carbon permits. Our model, however, has the important advantage that it solves for a global equilibrium with externalities (with or without taxes); thus, using our setting it is straightforward to perform optimal tax analysis. Nordhaus also uses a finite horizon, different model solution techniques, and some other features that partly make our approaches difficult to compare directly. Our approach is firmly within the modern-macro tradition, thus relying on explicit microfoundations both in terms of consumers and firms; all prices are market-clearing, and markets work well, aside from not dealing properly with externalities. We thus study one region and make simplifications at this stage of the research project in order to draw out some central implications. Of course, it will be important at a later stage to consider more complex settings; work along those lines is already in progress (see Krusell and Smith, 2009, for multiregional modeling, Hassler, Krusell, and Olovsson, 2009, for some productivity accounting and an examination of endogenous technology, and Gars, Golosov, and Tsyvinski, 2009 for a model with a back-stop technology).

Section 2 describes the model and characterizes the solution to the planning problem. Section 3 then looks at a decentralized world economy and derives the optimal-tax formula. In Section 5 we then use particular functional forms and calibrate the model to obtain our main quantitative conclusions. We discuss some obvious limitations of our work in the concluding Section 6.

2 The economy and the climate: the planner's perspective

In this section, we describe the central planning problem. This will later be compared to the decentralized solution in order to establish the existence of a policy that replicates the solution to the planning problem as a decentralized equilibrium.

A rather general planning problem, that we momentarily specialize somewhat, is

$$\max_{\{C_t,K_{t+1},E_t,R_{t+1}\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t U(C_t) \tag{1}$$

$$\begin{aligned}
C_t + K_{t+1} &= \hat{F}(A_t, K_t, N_t, E_t, A_t^e, S_t) + (1-\delta) K_t - Q(R_{t+1}, R_t, A_t^r) \\
R_{t+1} &= R_t - E_t, \quad R_0 \text{ given}, \\
R_t &\geq 0 \forall t, \\
N_t &= 1 \forall t, \\
S_t &= L(E^t).
\end{aligned}$$

The function U is a standard concave utility function, C is consumption, and $\beta \in (0,1)$ is the discount factor. The second line of (1) is the aggregate resource constraint. The left-hand side is resource use—consumption and next period's capital stock. The first term on the right-hand side is output produced by an aggregate production function $\hat{F}$. The arguments of $\hat{F}$ include the standard inputs K_t and N_t (capital and labor) and A_t: an aggregate measure of technology. In addition, aggregate output depends on the energy input (fossil fuel) E_t, with an associated energy efficiency level A_t^e. We assume throughout that fossil fuel is essential in the sense that the production function satisfies the standard Inada conditions. Finally, we allow the climate variable S_t to affect output. This effect could in principle be both positive and negative, though here the focus is on various sorts of damages that are all captured in the production function. We will specify later how $\hat{F}$ depends on S, but note that we view climate to be well represented by one variable, which we take to be the global concentration of carbon in the atmosphere. This is reasonable given the medium-complexity climate models; these imply that the climate is quite well described by current carbon concentrations in the atmosphere (e.g., lags due to ocean heating are not so important).

Note also that we let the climate itself depend on previous use of fossil fuel through the history $E^t \equiv \{\ldots, E_{t-1}, E_t\}$ via the function L. Later, we will give $L(E_t)$ a simple linear lag structure. When we consider the decentralized equilibrium, the effect of emissions on climate damages will be assumed to be a pure externality, not taken into account by any private agent.

The parameter δ measures capital depreciation and Q represents total extraction costs for current extractions: it depends on the fossil fuel in the beginning of the period, R_t, and on the amount left at the end, R_{t+1}. This allows the costs of extracting one unit to increase over time as more easily available resources are used up; however, there can also be technological progress in the extraction technology through a changing A_t^r. Specifically, we assume

$$Q(R_{t+1}, R_t, A_t) = \frac{1}{A_t^r} \int_{R_{t+1}}^{R_t} q(R)\, dR$$

where $q(R)$ is a decreasing and differentiable function such that $q(0)$ is bounded. The interpretation of this is that given a technology level A_t^r, each unit of oil in the ground is associated with a given extraction cost, which falls as the technology (exogenously) improves. As the resource gets scarcer, more costly extraction sites are used. It is optimal to extract the resource in reverse order of extraction costs whenever marginal productivity of capital is above unity (whenever the market interest rate above zero): if a consumption unit can be saved, by extracting at a cheaper rate at one extraction source than at another, save it today rather than tomorrow since one consumption unit is worth more today than it is tomorrow.[1]

Specializing the setup somewhat, we let S_t be determined with in a simple mean-reverting manner:

$$S_{t+1} = (1 - \varphi) S_t + E_t,$$

where φ captures the rate at which carbon is absorbed by the deep oceans, thus no longer affecting the climate.

We also assume that the climate damage affects output proportionally:

$$Y_t = S(S_t) F(A_t, K_t, N_t, E_t, A_t^e) \equiv \hat{F}(.),$$

where the damage function satisfies

$$S(S_t) > 0, S'(S_t) < 0.$$

Thus, we summarize all damages, including direct utility damages or damages to the capital stock, as well as technical change that reduces the damages (adaptation), in the function S. This is a shortcut, the most important benefit of which is the connection with Nordhaus's work: Nordhaus uses this damage function and has elaborate estimates of it.

2.1 Solving the planning problem

The planner problem is then

$$\max_{\{K_{t+1},R_{t+1},C_t,S_t\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t U(C_t)$$

$$\begin{aligned} C_t &= S(S_t) F(A_t, K_t, N_t, R_t - R_{t+1}, A_t^e) \\ &\quad + (1-\delta) K_t - Q(R_{t+1}, R_t, A_t^r) - K_{t+1} \\ S_{t+1} &= (1-\varphi) S_t + R_t - R_{t+1} \end{aligned}$$

and R_t being a non-increasing non-negative sequence.

The first-order condition for K_{t+1} is the usual

$$\frac{U'(C_t)}{U'(C_{t+1})\beta} = S(S_{t+1}) F_{K,t+1} + 1 - \delta \equiv \rho_{t+1}, \tag{2}$$

where ρ denotes the gross return on saving (the real interest rate) and we use the notation

$$\frac{\partial F(K_t, N_t, E_t, A_t, A_t^e)}{\partial X} = F_{X,t+1}, X \in \{K_t, N_t, E_t\}$$

[1] This result goes back to HERFINDAHL, O. C.: "Depletion and Economic Theory," in Extractive Resources and Taxation, ed. by M. Gaffney. Madison: University of Wisconsin Press, 1967, pp. 63-90.

Let $\beta^t \lambda_t^s$ denote the shadow value on the constraint $S_{t+1} = (1-\varphi) S_t + E_t$. The interpretation is that λ_t^s is the damage cost, measured in current utils, of having one more unit of carbon in the atmosphere from next period and on, i.e., *the marginal damage cost* of one unit more of current emissions. The first-order condition for S_{t+1} can then be written

$$\lambda_t^s = -\beta U'(C_{t+1}) \frac{S'(S_{t+1})}{S(S_{t+1})} Y_{t+1} + \beta \lambda_{t+1}^s (1-\varphi).$$

This amounts to a law of motion for the the marginal damage cost. It states that this cost is equal to the (discounted) period damage cost in utils next period plus the (discounted) damage cost next period times one minus the depreciation rate for atmospheric carbon.

Defining the period climate damage in real (consumption) units to be $d_{t+1} = -\frac{S'(S_{t+1})}{S(S_{t+1})} Y_{t+1}$ and iterating forward on the first-order condition for S_{t+1}, we obtain

$$\lambda_t^s = \sum_{s=1}^{\infty} (1-\varphi)^{s-1} \beta^s U'(C_{t+s}) d_{t+s} + \lim_{s\to\infty} \beta^s (1-\varphi)^{s-1} U'(C_{t+s}) d_{t+s}.$$

Assuming that the limiting term is zero, using (2) and defining

$$\Gamma_t^s \equiv \prod_{j=0}^{s} \frac{1}{\rho_{t+j}}$$

and

$$\Lambda_t^s \equiv \frac{\lambda_t^s}{U'(C_t)},$$

we obtain that

$$\Lambda_t^s = \sum_{s=1}^{\infty} (1-\varphi)^{s-1} \Gamma_t^s d_{t+s}.$$

Λ_t^s measures the marginal cost of a unit of carbon in the atmosphere in terms of the consumption good. Thus, using the optimality condition, it can be written as the present discounted value of the production damages created by a marginal unit of extra carbon in the atmosphere. Note that discounting here involves both standard discounting (Γ_t^s is the value of a unit consumption at s in terms of consumption at t) and the depreciation of atmospheric carbon.

Finally, define

$$\varepsilon_t \equiv S(S_t) F_E(t) - q_t,$$
$$\varepsilon_t' \equiv S(S_t) F_E(t) - q_t'$$

where

$$q_t \equiv \frac{q(R_{t+1})}{A_t^r},$$
$$q_t' \equiv \frac{q(R_t)}{A_t^r}.$$

ε_t denotes the marginal product of oil net of extraction cost but excluding externalities, i.e., the private net marginal value of oil for the *last* unit extracted in period t. Correspondingly, ε_t' denotes the same value for the *first* unit extracted in period t. Note that

$$\frac{q_t}{q_{t+1}'} = \frac{A_t'}{A_{t+1}^r}$$

i.e., it measures technological progress in extraction technology.

The first-order condition for R_{t+1} can then be written

$$U'(C_t)\,\varepsilon_t - \lambda_t^s = \beta\left(U'(C_{t+1})\,\varepsilon_{t+1}' - \lambda_{t+1}^s\right) \tag{3}$$

Using (2) and assuming $\varepsilon_t - \Lambda_t^s > 0$, this condition becomes

$$\rho_{t+1} = \frac{\varepsilon_{t+1}' - \Lambda_{t+1}^s}{\varepsilon_t - \Lambda_t^s}. \tag{4}$$

This expression is a variant of the famous Hotelling rule[2], stating that the return on capital should be set equal to the return to postponing extraction of a marginal unit of oil to the next period. We should note that the last terms in both the numerator and the denominator are due to the climate externalities. We can think of this as a portfolio choice problem: how should the wealth we are accumulating for ourselves and for future generations be split into capital, on the one hand, and, on the other, oil resources left in the ground? They should be accumulated in such as way as to equalize returns.[3]

Already at this point, let us point to some important features of Hotelling's formula. First, abstract from the climate externality so that we can think of this formula immediately in terms of market outcomes. Then the formula says that the price of oil net of extraction costs, which through proper market pricing must equal ε, should rise over time at a rate equal to the real rate of interest. Second, since the extraction costs will not go to infinity by assumption, it must be that asymptotically, to the extent the real interest rate settles down to a constant above unity, the gross price of oil must also rise at the rate of the interest rate (since the production cost will be negligible at this point in relation to the price). Third, a special case of some interest is that where q_t is a positive constant and where the real interest rate is constant. In this case, the gross price of oil must grow at a declining rate over time (and then converge to a rate of the real rate of interest): postponing extraction now has the benefit of spending the extraction cost later, so the price increase does not have to be so large for the producer to be indifferent.

[2] The original Hotelling rule, derived in Hotelling (1931), applied to a monopolistic resource owner. Solow (1974) and Stiglitz (1974) derive an analogous condition for the case of perfect markets and no externalities, in which case the market implements the optimal extraction path. Finally, Sinn (2007) shows how to include an externality in the condition, arguing that this naturally leads to slower extraction than in *laissez-faire*.

[3] See Sinn (2008) for a derivation of the Hotelling rule above and for the portfolio-choice interpretation.

2.1.1 Backstop technology

Suppose now that we consider the case of a backstop technology such that as in Dasgupta and Heal (1974), an alternative non-exhaustable, energy source becomes available at time T. From this point in time, energy is produced with a clean technology. Specifically, we assume that energy is produced with a specific capital good good K_t^e. For simplicity, we assume that the introduction of the clean technolgy is drastic so that fossil fuel is no longer used. Even though $E_{T+s} = 0$ for all $s \geq 0$, S_{T+s} remains positive if $S_T > 0$ and $\varphi < 1$.

The necessary conditions above remain valid for $t < T$, but we now get an end-condition for R_T, namely

$$\left(\varepsilon_T - \Lambda_T^s\right) R_T = 0.$$

This condition says that either all remaining fossil fuel is used in period T,i.e., $R_T = 0$, or $\varepsilon_T - \Lambda_T^s = 0$. In the latter case, the marginal social value of fossil fuel should be set to zero, i.e., the private value (the marginal product of fuel minus the marginal extraction cost) should be set equal to the present discounted value of the damage caused by a marginal unit of fossil fuel burning.

An important implication of this comes from using $\left(\varepsilon_T - \Lambda_T^s\right) = 0 \Leftrightarrow U'\left(C_T\right) \varepsilon_T - \lambda_T^s = 0$ in (3);

$$U'\left(C_{T-1}\right) \varepsilon_{T-1} - \lambda_{T-1}^s = \beta\left(U'\left(C_T\right) \varepsilon_T' - \lambda_T^s\right) = 0.$$

By backward induction follows the following proposition;

Proposition 1 *Suppose that at som point in time T, fossil fuel becomes useless. If it is optimal to leave a strictly positive amount of fossil fuel in the ground at period T, the social value of fossil fuel at all dates before T is zero, i.e., $\varepsilon_t - \Lambda_t^s = 0 \forall t \leq T$.*

3 A decentralized economy and implementation of the optimum

We assume the government uses taxes on resource use in order to achieve the socially optimal allocation.

A representative individual solves

$$\max \sum_{t=0}^{\infty} \beta^t U\left(\rho_t K_t + \Pi_t^f + \Pi_t^e + T_t - K_{t+1}\right)$$
$$s.t.\ C_t + K_{t+1} = \rho_t K_t + \Pi_t^f + \Pi_t^e + T_t,$$

where Π_t^f and Π_t^e are profits from final goods production and resource extraction and T_t are government transfers that we assume are equal to the tax revenues in present value. Here, in equilibrium Π_t^f will be zero, due to perfect competition, but Π_t^e will be positive, essentially delivering the stock value of the oil in the ground.

The first-order condition of interest here, i.e., that for K_{t+1}, as usual delivers

$$U'\left(C_t\right) = \beta \rho_{t+1} U'\left(C_{t+1}\right). \tag{5}$$

Goods production takes place in perfect competition, implying that the price of the resource—the gross oil price, p_t^e—is given by its marginal product

$$p_t^e = \varepsilon_t + q_t, \tag{6}$$

where the marginal extraction cost q_t is added back because p^e is a gross price. Competitive goods production also implies that the competitive rental of capital satisfies

$$\rho_t = \frac{\partial Y_t}{\partial K_t} + 1 - \delta. \tag{7}$$

This implies that (5) coincides with the planner solution.

Now consider a representative atomistic resource extraction firm owning a share of fossil fuel resources of all remaining extraction-cost levels. Let us introduce an *profit-tax* tax τ_t and a *per-unit* tax θ_t. The *profit tax* taxes the oil-price net of extraction costs. Consequently, it is identical to a tax on the price of oil when extraction costs are zero. Two tax instruments are actually superflous in the sense that any allocation that can be achieved with both of them can also be achived with one only. We introduce them for pedagogical reasons.

The problem of a representative resource extraction firm is thus to maximize

$$\sum_{s=t}^{\infty} \Gamma_t^s \left(\left(p_s^e - \theta_s \right) \left(r_s - r_{s+1} \right) - Q \left(r_{s+1}, r_s; A_s \right) \right) \left(1 - \tau_s \right)$$
$$\text{s.t. } r_{t+1} \geq 0 \forall t.$$

The fact that we assume the oil extracting firms to be atomistic implies that they take all prices and the sequence of capital as exogenous.

Using (6), the first-order condition with respect to r_{t+1} can be written

$$\rho_{t+1} = \frac{\left(\varepsilon_{t+1} - \theta_{t+1} \right) \left(1 - \tau_{t+1}^e \right)}{\left(\varepsilon_t - \theta_t \right) \left(1 - \tau_t^e \right)},$$

provided $\left(\varepsilon_t - \theta_t \right) \left(1 - \tau_t^e \right) > 0$.

The intuition for this condition is that a unit of extraction today delivers a net benefit of $\left(\varepsilon_t - \theta \right) \left(1 - \tau_t \right)$ whereas if the unit is extracted next period it delivers $\left(\varepsilon_{t+1} - \theta_{t+1} \right) \left(1 - \tau_{t+1} \right)$. Since the present value of the latter has to equal the former for the firm to be indifferent between extracting today and next period, the equation follows. This is again a Hotelling-formula of sorts.

It should be noted that a constant profit-tax has no impact on equilibrium fossil fuel consumption provided that no finite amount of fossil fuel is left in the ground forever. The latter provision is satisfied since fossil fuel is assumed to be essential. Therefore, the first-order condition will hold at all points in time and this is sufficient for determining oil use.

3.1 The effects of taxes and the optimal-tax formula

To implement the planning solution using our taxes, we need to set the private return to keeping oil in the ground equal to the social return, i.e., setting

$$\frac{\left(\varepsilon'_{t+1}-\theta_{t+1}\right)\left(1-\tau^{e}_{t+1}\right)}{\left(\varepsilon_t-\theta_t\right)\left(1-\tau^{e}_{t}\right)}=\frac{\varepsilon_{t+1}-\Lambda^{s}_{t+1}}{\varepsilon_t-\Lambda^{s}_{t}}. \tag{8}$$

Thus, we have

Proposition 2 *The optimal allocation can be implemented by setting*

$$\begin{aligned}\tau_t &= \tau \text{ for any } \tau \text{ and}\\ \theta_t &= \Lambda^s_t \ \forall t,\end{aligned}$$

or, equivalently, by setting

$$\begin{aligned}\tau_t &= \frac{\Lambda^s_t}{\varepsilon_t} \text{ and}\\ \theta_t &= 0.\end{aligned}$$

Obviously, the proposition just states two kinds of possibilities; in practice time-varying ad-valorem and per-unit taxes can be used together in nontrivial ways, as long as (8) holds.

With the taxes given in the proposition, the tax per unit of oil is equal to the marginal externality cost, i.e., the tax makes firms internalize the externality.

3.2 The optimal tax path

In order to find the optimal tax path, operationally we can first solve the planning problem and then simply back out the tax rates from the optimal-tax formula (8), assuming that we choose among the ad-valorem and per-unit tax instruments. Let us focus on the former. Thus, for all t, we simply find $\frac{\Lambda^s_t}{\varepsilon_t}$ from the optimal allocation and set τ_t equal to it.

An important policy issue is to what extent optimal tax rates are falling or increasing over time. If tax rates are falling, their purpose would be to delay extraction by increasing the return to keeping oil in the ground.

A useful proposition, that goes further than the direct characterizations above, is as follows.

Proposition 3 *The optimal profit-tax rate falls if and only if*

$$d_{t+1} > \varphi\Lambda^s_{t+1} = \varphi\sum_{s=1}^{\infty}\left(1-\varphi\right)^{s-1}\Gamma^s_t d_{t+1+s}.$$

Although this proposition is in terms of endogenous variables it will turn out to be useful as an easy way to check whether taxes are falling. It also provides a useful intuition by comparing future and current damages.

Since $\varphi \sum_{s=1}^{\infty} (1-\varphi)^{s-1} = 1$, we can interpret this proposition as saying that taxes are falling when current direct production damages are higher than a weighted average of the future damages, each of these measured in present-value terms, and each discounted by the rate of depreciation of atmospheric carbon. The proposition implies that optimal taxes must be falling for sufficiently low depreciation rates of atmospheric carbon, provided the present discounted value of marginal damages $\sum_{s=1}^{\infty} \Gamma_t^s d_{t+1+s}$ is finite. An example of this case occurs under a the "pure stock externality", where S_t equals the sum of all energy emissions in the past, i.e., $S_0 + R_t - R_0$, implying that damages are a function of the current resource stock R_t (for exogenous initial conditions). Finally, we should note that a low depreciation rate is merely a sufficient condition for falling taxes; in the functional-form examples below, in fact, optimal taxes are falling also when $\varphi = 1$ (in which case emissions only have a one-period effect).

Consider now the case of a backstop technology and assume that the optimal allocation calls for some oil being left in the ground at period T. Clearly, this requires $(\varepsilon_T - \theta_T)(1 - \tau_T^e) = 0$, calling for a per-unit tax equal to the price minus extraction costs ($\theta_T = \varepsilon_T$) or a 100% percent *profit-tax.* Furthermore, the private optimality condition

$$\rho_{t+1} (\varepsilon_t - \theta_t) (1 - \tau_t^e) = \left(\varepsilon'_{t+1} - \theta_{t+1}\right) \left(1 - \tau_{t+1}^e\right)$$

we find;

Proposition 4 *Suppose that at some point in time T, fossil fuel becomes useless. If it is optimal to leave a strictly positive amount of fossil fuel in the ground at period T, the optimal allocation is implemented iff either $(\varepsilon_t - \theta_t)$ or $\tau_t^e = 1$ for all $t \leq T$.*

We can also find the *laissez-faire* allocation by noting that either some fossil fuel is left in the ground, in which case $\varepsilon_T = 0$, or all fossil fuel is used before it becomes obsolete. In the former case, the rents on fossil fuel are zero in all periods, i.e., $\varepsilon_t = 0 \ \forall t$. In the latter, the end condition that pins down the allcation is that all fuel is used.

4 Wedges and Taxes

In this section we characterize the relationship between wedges in the social planner's problem and profit, per unit and sales taxes implementing the optimum. We then provide sufficient conditions for the profit taxes to decrease. We also show that per-unit taxes increasing, but at a rate less than the interest. Throughout this section we assume that the extraction costs take the form of $Q(E_t)$.

4.1 Wedges in the social planner's problem

Let us now for convenience restate the planner's problem as follows:

$$\max_{\{K_{t+1},C_t,S_t\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t U\left(C_t\right)$$

$$[\lambda_t] : C_t = \hat{F}\left(A_t, K_t, N_t, E_t, A_t^e, S_t\right) + (1-\delta) K_t - K_{t+1} - Q\left(E_t\right)$$

$$[\eta_t] : S_{t+1} = (1-\varphi) S_t + E_t.$$

$$[\psi] : \sum_{t=0}^{\infty} E_t \leq R_0.$$

The first order conditions for this problem are as follows:

$$[C_t] : \beta^t U'(C_t) = \lambda_t,$$

$$[E_t] : \lambda_t \left[\hat{F}_{E,t} - Q'\left(E_t\right)\right] = \eta_t + \psi,$$

$$[S_t] : \lambda_t S'\left(S_t\right) = \eta_{t-1} - (1-\varphi)\,\eta_t.$$

A useful way to think about the characterization of the social planner's problem is in terms of the wedge:

$$wedge_t \equiv \frac{U'(c_t)\left[\hat{F}_{E,t} - Q'\left(E_t\right)\right]}{\beta U'(c_{t+1})\left[\hat{F}_{E,t+1} - Q'\left(E_{t+1}\right)\right]}.$$

Consider the first order condition for S_t and substitute for η_t:

$$\lambda_t S'\left(S_t\right) = \lambda_{t-1}\left[\hat{F}_{E,t-1} - Q'\left(E_{t-1}\right)\right] - \psi - (1-\phi)\left[\lambda_t\left(\hat{F}_{E,t} - Q'\left(E_t\right)\right) - \psi\right].$$

Note that $S'\left(S_t\right) > 0$:

$$(1-\varphi) < \frac{\lambda_{t-1}\left[\hat{F}_{E,t-1} - Q'\left(E_{t-1}\right)\right] - \psi}{\lambda_t\left(\hat{F}_{E,t} - Q'\left(E_t\right)\right) - \psi}.$$

This in turn implies that:

$$(1-\varphi) < \frac{U'(C_{t-1})\left[\hat{F}_{E,t-1} - Q'\left(E_{t-1}\right)\right]}{\beta U'(C_t)\left[\hat{F}_{E,t} - Q'\left(E_t\right)\right]}$$

We summarise the results in the following proposition.

Proposition 5 *For all t:*

$$wedge_t > 1 - \varphi$$

We continue our analysis by characterizing the profit taxes. Consider the first order conditions for the oil extracting firm faced with a sequence of profit taxes:

$$\frac{U'(C_t)}{\beta U'(C_{t+1})} = \frac{\left(1-\tau^e_{t+1}\right)\left[\hat{F}_{E,t+1} - Q'\left(E_{t+1}\right)\right]}{\left(1-\tau^e_t\right)\left[\hat{F}_{E,t} - Q'\left(E_t\right)\right]}$$

The above proposition then puts a bound on the ratio of profit taxes across periods.

$$\frac{1-\tau^e_{t+1}}{1-\tau^e_t} > 1-\varphi$$

An immediate corrollary describing taxes on profit follows.

Corollary 6 *If $\varphi = 0$, taxes on profit that implement the optimal allocation decrease.*

Let us now consider a per unit oil tax, θ_t. The first order conditions of the oil extracting firms in the competitive equilibrium are:

$$\frac{U'(C_t)}{\beta U'(C_{t+1})} = \frac{\hat{F}_{E,t+1} - Q'\left(E_{t+1}\right) - \theta_{t+1}}{\hat{F}_{E,t} - Q'\left(E_t\right) - \theta_t}.$$

or

$$\begin{aligned} & U'(C_t)\left[\hat{F}_{E,t} - Q'\left(E_t\right)\right] - \beta U'(C_{t+1})\left[\hat{F}_{E,t+1} - Q'\left(E_t\right)\right] \\ = & U'(C_t)\theta_t - \theta_{t+1}\beta U'(C_{t+1}). \end{aligned}$$

If $\varphi = 0$, then the above proposition implies that the wedge is greater than 1 and the above equation implies that

$$\begin{aligned} U'(C_t)\theta_t - \theta_{t+1}\beta U'(C_{t+1}) &> 0 \\ \frac{\theta_{t+1}}{\theta_t} &< \frac{U'(C_t)}{\beta U'(C_{t+1})} = \rho_{t+1}. \end{aligned}$$

We summarize the analysis in the corollary that follows.

Corollary 7 *If $\varphi = 0$, the rate of growth of per unit taxes that implement the optimal allocation is smaller than the interest rate.*

This logic implies that per-unit taxes θ_t may be increasing, but at a rate less than the interest. The fact that they may be increasing is not surprising. The after tax price of oil is exponentially increasing. If taxes were constant, a ratio of tax to the price of oil would quickly go to zero (a $1 tax on oil may affect some decisions when oil is $10 but not when it is $200), but because they do not raise as quickly as the profit per unit of oil, incentives to postpone extraction into the future are provided.

5 An analytical example and a Nordhaus calibration

We know that with log utility, full depreciation and Cobb-Douglas production, there is a closed form solution to the neoclassical growth model. Let us therefore use the same assumptions in the case of a non-renewable resource with externalities, since this model as well has a closed-form solution so long as extraction costs are zero.[4] Key in this analytical derivation is a proportionality result: the level of output drops out of all relevant first-order conditions. This also means that the model's implications for energy use, and for optimal energy taxes, are invariant to the key driver of output growth: improvements in total-factor productivity (TFP). Thus, we can shut down TFP growth here since it does not alter any of our results.[5]

More importantly, however, one can argue that these functional-form assumptions are not wildly at odds with what would seem to be quantitatively reasonable assumptions. First, logarithmic curvature for utility is in line with most applied macroeconomic studies. Second, full depreciation is not on short horizons, but with the 10-year periods we use here, it is not too far from a reasonable rate. Third, though one would have trouble over shorter time horizons with the assumption that energy enters like capital and labor in a Cobb-Douglas production function—since it seems reasonable to assume that installed equipment and structures have rather fixed energy requirements—but on a longer horizon, since the style of capital can be adjusted in response to energy prices, it is not so unreasonable with a Cobb-Douglas technology. In fact, it is also what Nordhaus uses in his RICE model, which is entirely quantitative in nature. Fourth, zero extraction costs is obviously an exaggeration but the "Hotelling rents", i.e., the price of oil minus the marginal extraction cost is quite high in percentage terms.

Production is thus assumed to be

$$Y_t = S\left(S_t\right) F\left(K_t, E_t, A_t\right) = S\left(S_t\right) A_t K_t^{\alpha} E_t^{\gamma}.$$

This together with logarithmic utility implies an Euler equation for physical capital investment that reads

$$\frac{\alpha Y_{t+1}}{K_{t+1}} = \frac{C_{t+1}}{\beta C_t}. \tag{9}$$

Because of full depreciation, $C_t + K_{t+1} = Y_t$. Now it is straightforward to see that $C_t = (1-\alpha\beta) Y_t$ solves (9). Furthermore,

$$\rho_{t+1} = \frac{U'\left(C_t\right)}{U'\left(C_{t+1}\right)\beta} = \frac{Y_{t+1}}{Y_t \beta} \tag{10}$$

[4]If extraction costs are positive and dependent on accumulated extraction, as assumed above, the analytical analysis is still tractable if extraction costs are proportional to output. This would be the case if "extraction services" are produced by a production function of the sam form as the one producing final output. In the appendix, such an analysis is executed.

[5]To be clear, higher TFP increases the demand for energy, but with Cobb-Douglas production it will simply increase the price of energy one-for-one, and the time path for energy will be unaffected.

Using this, we can compute the marginal damage (current utility) cost of the climate externality, λ_t^s, as

$$\lambda_t^s = -\sum_{s=1}^{\infty} (1-\varphi)^{s-1} \beta^s U'(C_{t+s}) \frac{S'(S_{t+s})}{S(S_{t+s})} Y_{t+s}.$$

Similarly, the same cost measured in current consumption units is

$$\Lambda_t^s = Y_t \sum_{j=1}^{\infty} (1-\varphi)^{j-1} \beta^j \frac{S'(S_{t+j})}{S(S_{t+j})}. \tag{11}$$

We notice now that there is proportionality to output also for Λ_t^s. Using the Hotelling equation, (4), and (10), we obtain

$$\frac{Y_{t+1}}{Y_t \beta} = \frac{\varepsilon_{t+1} - \Lambda_{t+1}^s}{\varepsilon_t - \Lambda_t^s} = \frac{\frac{\gamma Y_{t+1}}{E_{t+1}} - Y_{t+1} \sum_{j=1}^{\infty} (1-\varphi)^{j-1} \beta^j \frac{S'(S_{t+1+j})}{S(S_{t+1+j})}}{\frac{\gamma Y_t}{E_t} - Y_t \sum_{j=1}^{\infty} (1-\varphi)^{j-1} \beta^j \frac{S'(S_{t+j})}{S(S_{t+j})}}.$$

Noticing the proportionality to output everywhere, we can simplify to write

$$\frac{1}{\beta} = \frac{\frac{\gamma}{E_{t+1}} - \sum_{j=1}^{\infty} (1-\varphi)^{j-1} \beta^j \frac{S'(S_{t+1+j})}{S(S_{t+1+j})}}{\frac{\gamma}{E_t} - \sum_{j=1}^{\infty} (1-\varphi)^{j-1} \beta^j \frac{S'(S_{t+j})}{S(S_{t+j})}}.$$

This equation can be rewritten and simplified further; we will look at special cases below. One important implication of this equation, however, is that it, together with the conditions that $E_t \geq 0$ for all t and that $R_t \to 0$, fully determines the sequence $\{E_t\}_{t=0}^{\infty}$. This is true since each S_t only depends on past values for energy. The upshot is that our functional-form assumptions allowed us to simplify what would otherwise be a joint dynamic system for capital and energy into a system only involving energy. The equation also allows us to see that at time approaches infinity, since E_t has to approach zero, and since $S'(0)/S(0)$ is finite, energy use will have to shrink according to $E_{t+1} = \beta E_t$.[6]

Turning to the specific forms of the climate system and the damage function S, we have several cases of interest.

5.1 An exponential damage function

Suppose that

$$S(S_t) = e^{-\gamma_s S_t}.$$

Then it follows that

$$\frac{S'(S_t)}{S(S_t)} = -\gamma_s.$$

In other words, the marginal effect of an increase in S_t is constant in percentage terms in this case.

[6] This result does not just obtain under our functional-form assumptions.

It follows that

$$\lambda_t^s = \frac{\gamma_s \beta}{(1-\alpha\beta)(1-\beta(1-\varphi))}.$$

Thus, optimal per-unit taxes, θ^*, satisfy

$$\begin{aligned}\theta_t^* &= \lambda_t^s \\ &= \frac{\gamma_s \beta C_t}{(1-\alpha\beta)(1-\beta(1-\varphi))} \\ &= \frac{\gamma_s \beta}{(1-\beta(1-\varphi))} Y_t.\end{aligned}$$

Note that

$$\frac{\partial \theta_t^*}{\partial \beta} = \frac{\gamma_s}{(1-\beta(1-\varphi))^2} \geq 0,$$

i.e., you tax more the higher the degree of patience; and that

$$\frac{\partial \theta_t^*}{\partial \varphi} = -\frac{\beta^2 \gamma_s}{(1-\beta(1-\varphi))^2} \leq 0,$$

i.e., that a lower rate of atmospheric carbon depreciation also raises the tax rate. Finally, the tax formula shows that if output grows at a slower rate than does the gross price of fossil fuel, $\frac{\theta_t^*}{\varepsilon_t}$, and thus the ad-valorem tax, must fall over time. We shall confirm this supposition below.

Using the Euler equation (10) in the Hotelling rule (4) yields

$$\frac{Y_{t+1}}{Y_t \beta} = \frac{\varepsilon_{t+1} - \Lambda_{t+1}^s}{\varepsilon_t - \Lambda_t^s}. \tag{12}$$

This in turn gives

$$\frac{1}{\beta} = \frac{\frac{\gamma}{E_{t+1}} - \frac{\gamma_s \beta}{(1-\beta(1-\varphi))}}{\frac{\gamma}{E_t} - \frac{\gamma_s \beta}{(1-\beta(1-\varphi))}}.$$

Without externalities, this equation would yield $E_{t+1}/E_t = \beta$ and, given that the total resource stock must be exhausted exactly, $E_t = (1-\beta)R_t$. This result was derived by Dasgupta and Heal (1974), who (like the rest of the literature at that time) did not consider climate damages.

With externalities, we obtain

$$\frac{E_{t+1}}{E_t} = \frac{\beta}{1 - E_t \frac{\gamma_s}{\gamma} \beta \frac{1-\beta}{1-\beta(1-\varphi)}}. \tag{13}$$

Several things can be pointed out here. One is that $\frac{E_{t+1}}{E_t} > \beta$ if $\gamma_s > 0$. In words, the optimal allocation features postponed extraction relative to the case without externalities. This postponing is larger, the larger is γ_s and the smaller is φ. Moreover, $\frac{E_{t+1}}{E_t} < 1$, since if the right-hand side of (13) is larger than 1, $\varepsilon_t - \Lambda_t^s \leq 0$, which cannot be optimal. Further, as E_t falls over time, $\frac{E_{t+1}}{E_t}$ falls towards β. Note finally that as initial oil resources become very abundant, E_0 approaches $\gamma \frac{1-\beta(1-\varphi)}{\gamma_s \beta}$, implying that initially, oil use is close to constant over time.

Applying Lemma 3, we find that

$$\frac{d_{t+1}}{\varphi \Lambda_{t+1}^s} = \frac{\gamma_s Y_{t+1}}{\varphi Y_{t+1} \frac{\gamma_s \beta}{(1-\beta(1-\varphi))}} = 1 + \frac{1-\beta}{\beta \varphi},$$

implying that optimal ad-valorem tax rates are falling. This can also be seen by observing that ε_t, the marginal product of energy at t, equals $\gamma Y_t / E_t$, and since E_t falls over time, as seen above, ε_t must grow faster than output and thus than θ_t.

Together with

$$\begin{aligned} R_{t+1} &= R_t - E_t \\ K_{t+1} &= \alpha \beta A_t e^{-\gamma_s S_t} K_t^\alpha E_t^\gamma \\ S_{t+1} &= (1-\varphi) S_t + (R_t - R_{t+1}) \end{aligned}$$

the law of motion of the optimal allocation is determined. We have two initial conditions and the transversality condition implies that all oil will be used. This fully determines the optimal allocation.

5.2 A Nordhaus damage function: calibrating the model

Let us now use a description of damages in line with the literature: let us use Nordhaus's RICE-model. We let S_t denote the stock of carbon in the atmosphere above the pre-industrial level of 583 GTC (Giga tons of carbon), denoted $\bar{S}$. In line with Nordhaus's formulation, we assume the there is a log-linear relation between the atmospheric CO_2 concentration and the global mean temperature

$$T(S_t) = \lambda \ln \left(1 + \frac{S_t}{\bar{S}}\right) / \ln 2.$$

The parameter λ is set to 2.91, implying that a doubling of the CO_2 concentration increases the temperature by 2.91 degrees Celsius.

We also follow Nordhaus in assuming that the damage function is

$$\begin{aligned} S(S_t) &= \frac{1}{1 + \theta_1 T(S_t) + \theta_2 T(S_t)^2} \\ &= \frac{1}{1 + \theta_1 \lambda \ln\left(1 + \frac{S_t}{\bar{S}}\right) / \ln 2 + \theta_2 \left(\lambda \ln\left(1 + \frac{S_t}{\bar{S}}\right) / \ln 2\right)^2} \end{aligned}$$

with $\theta_1 = -4.5 * 10^{-3}$ and $\theta_2 = 3.5 * 10^{-3}$. We solve the model with 10-year time periods and use a discount factor of $\beta = 0.99^{10}$. We set the capital and fossil fuel shares (α and γ) to 0.3 and 0.03, respectively. The following figure plots $S(S_t)$, where we note that current S_t is a bit over 200 GTC, which is above the peak at $S_t = 96$ GTC.

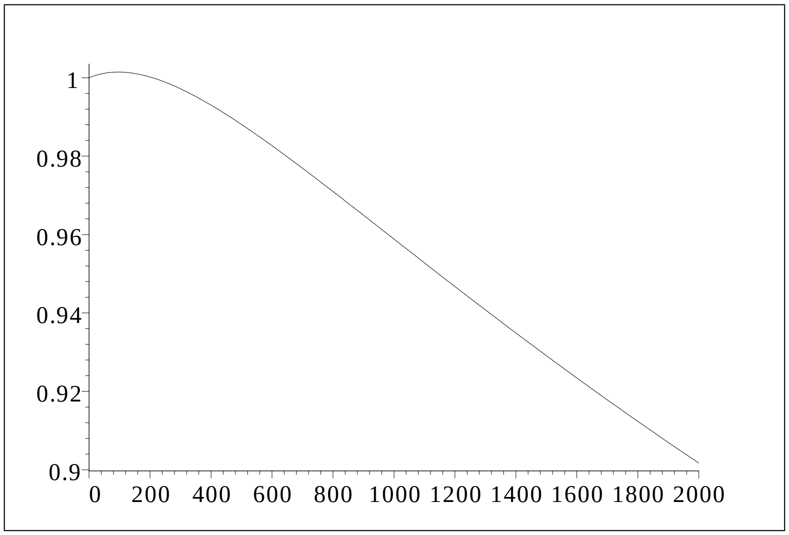

Figure 1. Damages as a function of CO2 stock.

Nordhaus uses a more elaborate law of motion for the stock of carbon in the atmosphere, with three sinks: the atmosphere, the biosphere with upper oceans, and the deep oceans. This system has the feature that emissions quickly mix between the two first sinks, but at a substantially slower rate between those two and the deep oceans. Here we target the long-run dynamics and set $\varphi = 1/11.7$ per decade, implying a a half-life of a deviation from the steady state of 117 years.

Using

$$\frac{S'(S_t)}{S(S_t)} = \frac{-\lambda\left(\theta_1 \frac{1}{(\bar{S}+S_t)} + \frac{2\theta_2\lambda}{\ln 2}\frac{\ln\left(1+\frac{S_t}{\bar{S}}\right)}{(\bar{S}+S_t)}\right)}{\ln 2 + \theta_1\lambda\ln\left(1+\frac{S_t}{\bar{S}}\right) + \frac{\theta_2}{\ln 2}\left(\lambda\ln\left(1+\frac{S_t}{\bar{S}}\right)\right)^2}$$

in (12) delivers

$$\frac{Y_{t+1}}{Y_t\beta} = \frac{\varepsilon_{t+1} - \Lambda^s_{t+1}}{\varepsilon_t - \Lambda^s_t}$$

$$\frac{1}{\beta} = \frac{\left(\frac{\gamma}{E_{t+1}} + \sum_{j=1}^{\infty}(1-\varphi)^{j-1}\beta^j\frac{S'(S_{t+j+1})}{S(S_{t+j+1})}\right)}{\left(\frac{\gamma}{E_t} + \sum_{j=1}^{\infty}(1-\varphi)^{j-1}\beta^j\frac{S'(S_{t+j})}{S(S_{t+j})}\right)},$$

which turns out to be easily solved numerically.[7] After having obtained the optimal path of E_{t+1}, all other variables follow directly.

[7] We solve this by guessing on an initial path of fossil fuel use (we use the laissez-faire path as an initial guess) from which we can calculate $\sum_{j=1}^{\infty}(1-\varphi)^{j-1}\beta^j\frac{S'(S_{t+j})}{S(S_{t+j})}$. Given this, there is a unique path of fossil fuel use that satisfies (12) and $\lim_{T\to\infty} R_T = 0$. We use this path to update $\sum_{j=1}^{\infty}(1-\varphi)^{j-1}\beta^j\frac{S'(S_{t+j})}{S(S_{t+j})}$ and iterate until convergence.

One remaining parameter is left: the initial stock of fossil fuel reserves R_0. In our benchmark simulation, we set this value to 2,000 GTC, which implies that the current reserves last around 250 years. In fact, this is likely to be a large underestimate, and we will also show results for $R_0 = 5,000$, which is more in line with independent estimates. However, the lower value delivers a value for E_0 that is less off the observed current level; with the higher value, the model predicts a level of E_0 that is much too small.

In Figure 2, the upper left panel shows the ratio of optimal to laissez-faire fossil fuel use for the coming 50 periods (500 years). As we see, the optimal fossil fuel use is substantially smaller than in the laissez-faire case early on, starting at about half of the laissez-faire use. This means that the fossil fuel reserves are exhausted faster under laissez-faire and that after a little over 100 years, fossil fuel use is higher in the optimal allocation forever after.

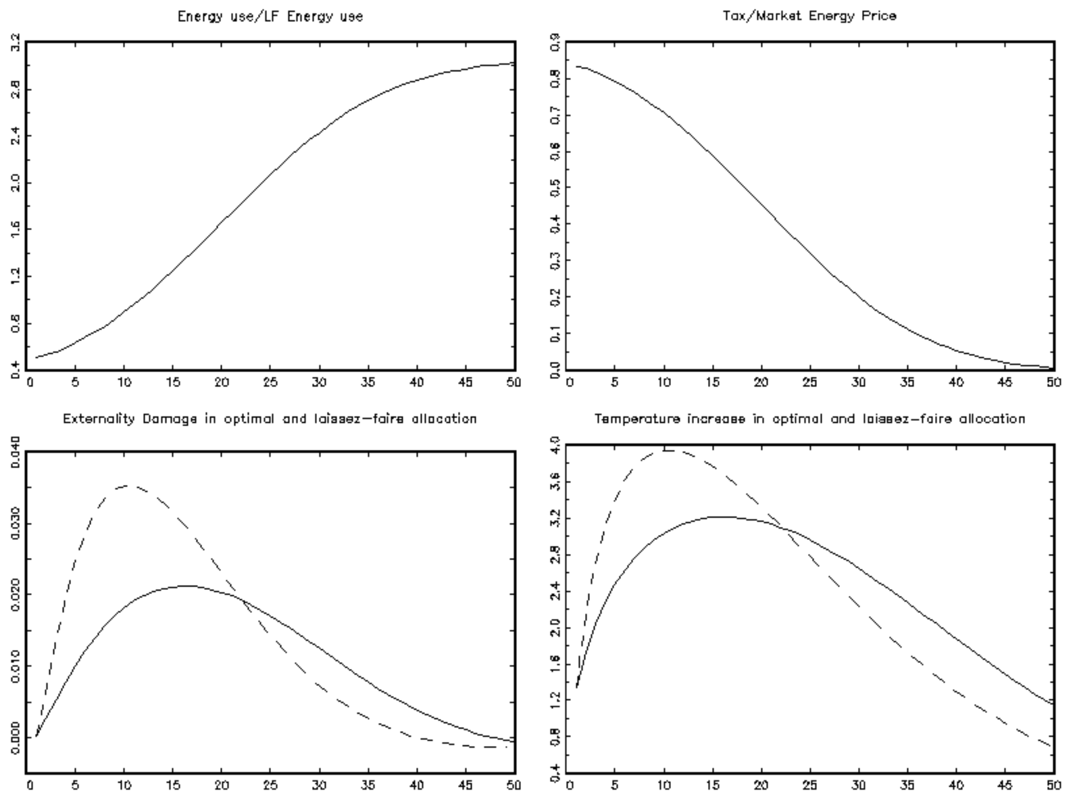

Figure 2. Baseline calibration.

The upper right panel shows optimal taxes as a share of the market price including taxes. As we see, optimal taxes are large, starting at over 80% of the market price. They then fall over time but remain high for a long period.

The lower left panel shows the externality damage, i.e., $1 - S(S_t)$, for the two scenarios. Climate damages are substantially larger in the laissez-faire allocation, reaching as much as 3.5% of GDP after a little more than 100 years. The maximum climate damage in the optimal allocation is 2.1% and occurs substantially later. The paths of the temperature increases resemble those of the climate damages, peaking at 3.9 and 3.2 degrees Celsius, respectively.

In the upper left panel of Figure3, the relative output in the two scenarios is shown. The introduction of a tax on fossil fuel use reduces output initially by 2% relative to the *laissez-faire.* However, optimal output then grows faster than in *laissez-faire* and is higher than in *laissez-faire* from period 6. In the very long run, climate externalities vanish but optimal output remains 4.9% higher than in *laissez-faire* due to the fact that fossil fuel consumption is pushed forward in time.

The upper right panel of figure 3 shows the oil price, inclusive of the tax, in the optimal allocation relative to that in the laissez-faire allocation.

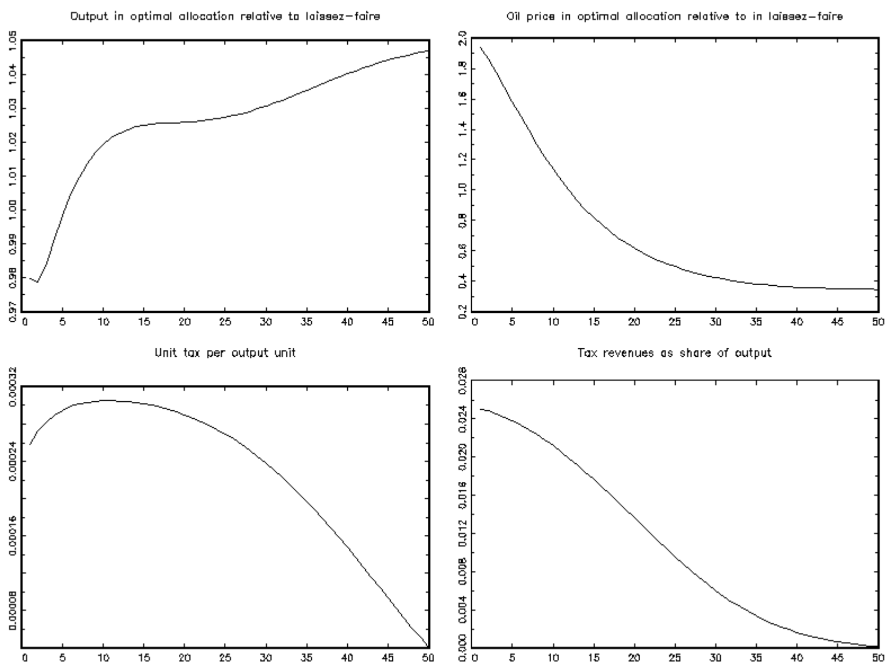

Figure 3. Baseline calibration.

The figure reveals that, although the tax is high, the postponement of energy use—the supply effect—makes the difference smaller: the optimal outcome only means twice as expensive an oil price, even though the tax rate is around 80%.

The lower left panel of figure 3 shows the unit tax as a fraction of output. We know that in the case with an exponential damage function (and otherwise identical preferences and technology) this amount is constant over time. Here, using Nordhaus's measured damage function, we obtain an outcome that does vary with time, but not in a major way. The lower right panel of figure 3 shows that not only the tax rate, but also tax revenues falls over time – starting at around 2.5% of GDP and falling towards zero.

In Figures 4 and 5, we redo the simulation using an initial stock of fossil fuel that is 2.5 times larger. This makes the differences between the scenarios quite dramatic. Almost the full fuel price must be taxed away for a long time. The output gain reaches a first peak at 9.8% above *laissez-faire* after 12 periods. Eventually, it settles at 17.3% above *laissez-faire* because of the postponement of fossil fuel use. In the laissez-faire allocation, the temperature increase is over 6 degrees Celsius.

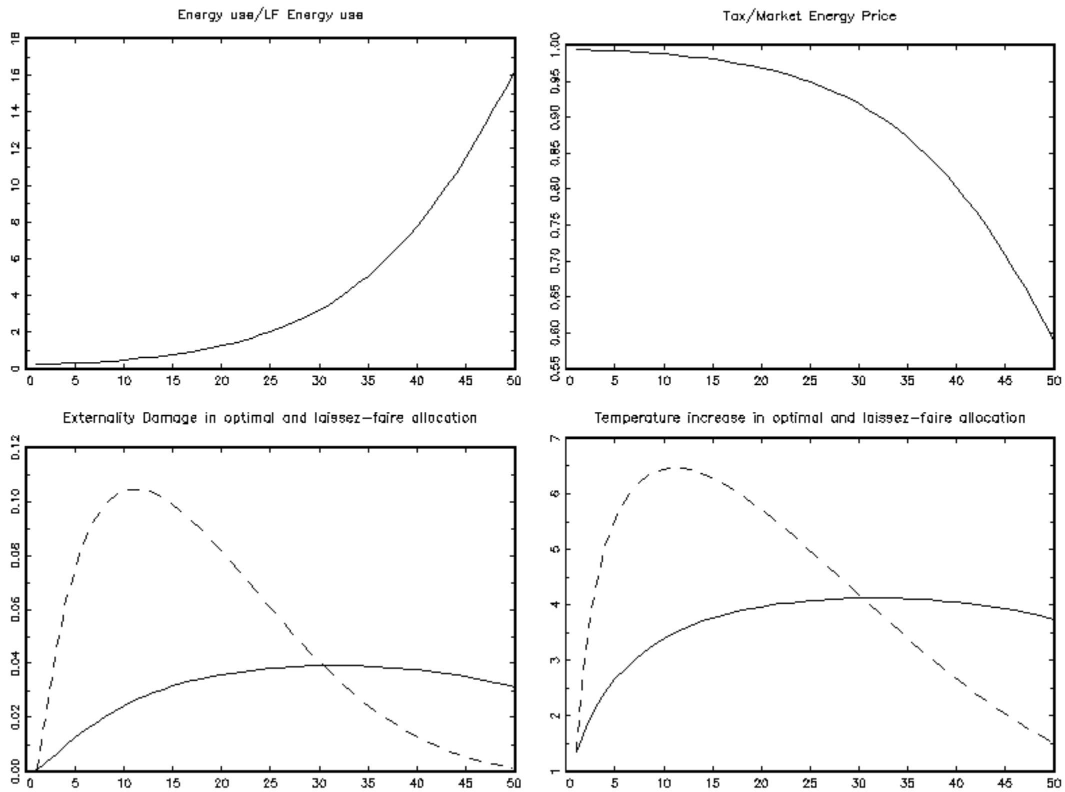

Figure 4. Larger fossil fuel reserves.

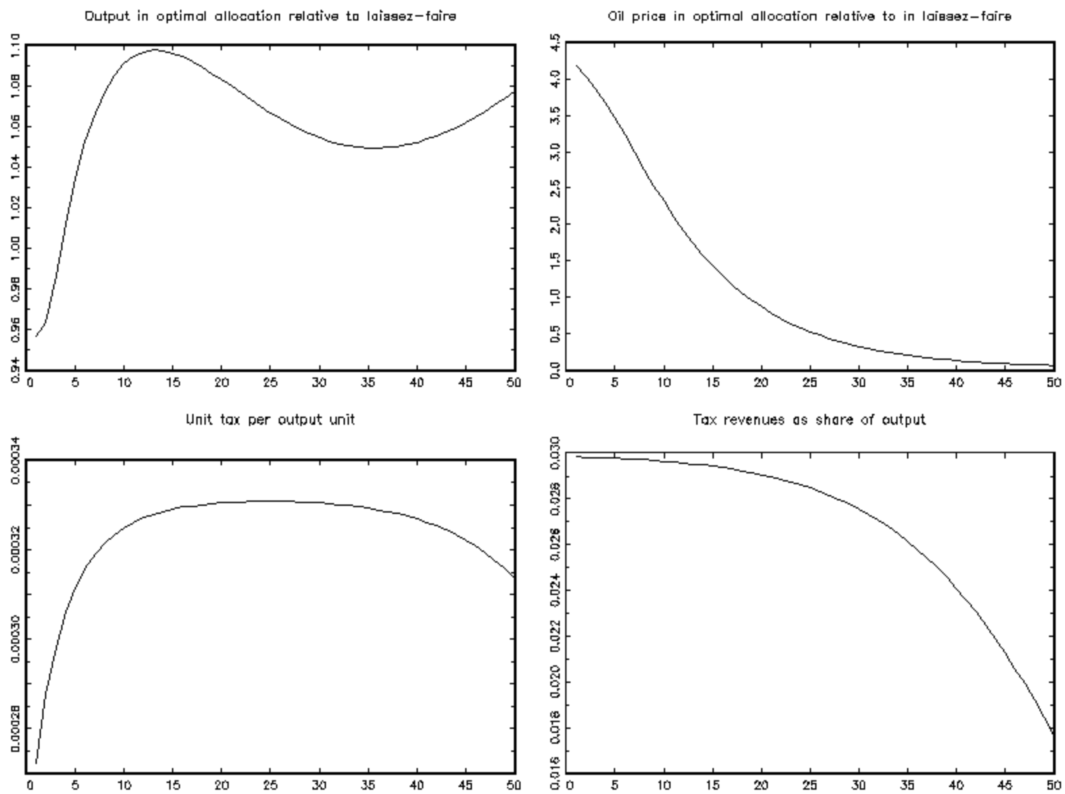

Figure 5. Larger fossil fuel reserves

A final issue we want to address is how much oil should maximally be extracted when there is a backstop technology and extraction costs are small. Of course, this depends on when the backstop technology is discovered. The following example assumes that fossil fuel becomes obsolete after $T = 20$ (200 years). Under the assumption that sufficiently large reserves exists, we know from proposition 1 that $\varepsilon_t = \Lambda_t$ for all $t < T$. In the case covered in this section, this implies that

$$E_t = \frac{\gamma}{-\sum_{j=1}^{\infty} (1-\varphi)^{j-1} \beta^j \frac{S'(S_{t+j})}{S(S_{t+j})}},$$

where $S_{t+1} = (1-\varphi) S_t + E_t$ and $E_t = 0 \forall t > T$.

Solving this equation gives an optimal path of E_t that is shown in upper left panel of Figure 6. As we see, fossil fuel consumption is fairly stable at levels close to today's. The fall slightly in the beginning of the period since carbon is accumulated in the atmosphere implying higher damages. At the end of the fossil fuel era, consumption optimally increases anticipating future lower damages. In total, 1947 GTC is used over the period, implying that if reserves today are at least this amount, all oil will be not be used if it becomes obsolete in 200 years. As already note, this is a small amount relative to expected existing reserves.

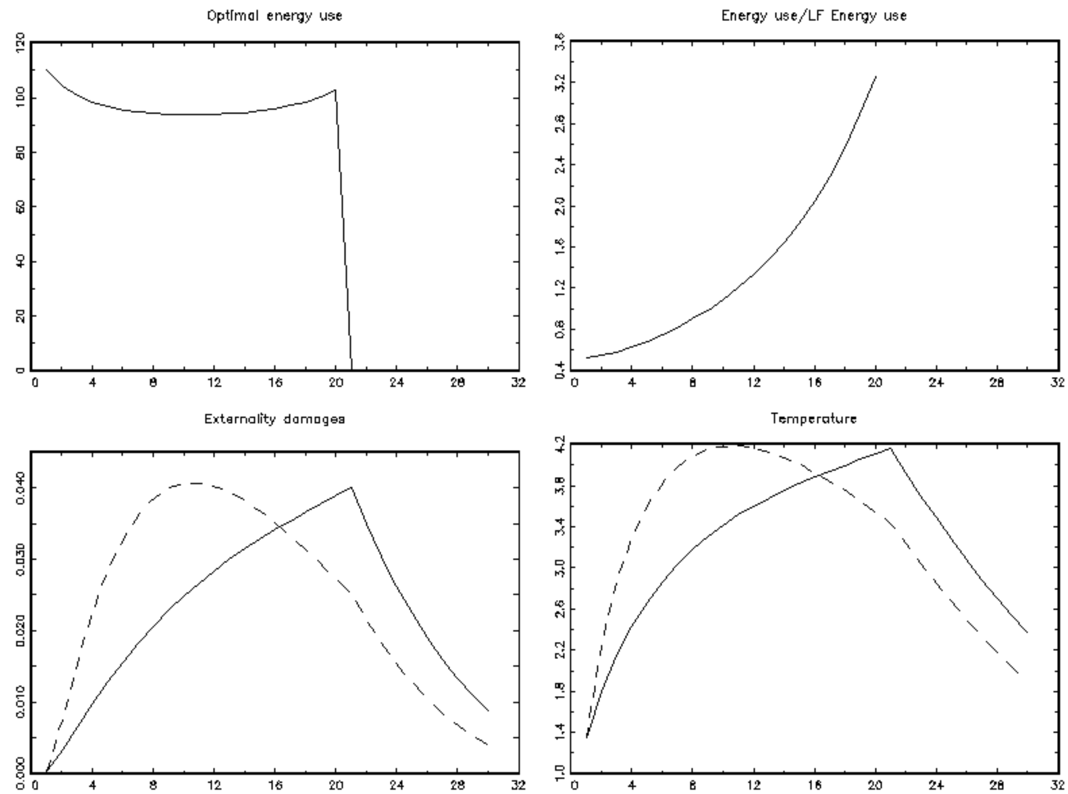

Figure 6. Fossil fuel consumption with a backstop technology.

In *laizzes-faire*, the allocation satisfies

$$E_{t+1} = E_t \beta,$$

which together with the condition

$$\sum_{t=0}^{T} E_t = E_0 \sum_{t=0}^{T} \beta^t = R_0$$

yields

$$E_t = R_0 \frac{1-\beta}{1-\beta^{T+1}} \beta^t.$$

In the remaing three panels of Figure 6, we compare the optimal allocation to the *laissez-faire* under the assumption that the amount of fossil fuel reserves is just large enough so that the planner wants to leave nothing, i.e., 1947 GTC. If fuel reserves are larger, the discrepancy between the two allocations if of course alrger.

The upper right panel shows the ratio of optimal to laizzes-faire allocations. The lower show the externality damage and the temperture increases respectively. As we see, the optimal allocation implies a substantial postponement of fuel use and of damages. We also note that substantially larger damages and temperature increases are allowed in the optimal allocation with a backstop technology than in the case above without such a technology.

6 Final comments

We have specified and characterized a global model with two-way climate-economy interactions. The tax prescriptions coming out of our setup call for large taxes on fossil fuel now and for these taxes to then decline slowly over time. Our work here is an obvious "first stab": this is, we believe, the first calibrated dynamic general-equilibrium model of the world economy with an externality due to a non-renewable resource, and it is important to proceed systematically and slowly in incorporating important features of the economy and of the climate. We will now briefly mention some of the assumptions used above that are key, and that should be important to relax in future work, as well as some challenges that we face with the present modeling.

As stated above, our result that time-invariant energy taxes are useless relies on the assumption that all oil in the grounds will be used up (or, more precisely, on the more primitive assumptions leading to that conclusion, such as the INADA condition making the marginal product of energy input infinity at zero). This may be an important assumption to relax. However, for practical purposes, the amount of cheap oil appears to be large enough that the assumption of zero extraction costs is a better approximation than one might first think, in particular when technological improvments in extraction technology are taken into account.

Perhaps more importantly, however, it appears reasonable that alternative energy sources—if not cold fusion so at least some radically better way of generating energy—will be developed, which might make oil that is difficult to extract look less cheap. Such a development would make the total amount of oil extraction "more endogenous", since now the important margin would involve this oil substitute and its associated costs. Would this speak against a declining tax on oil? As argued in Gars, Golosov, and Tsyvinski (2009), it would suggest that it is even more important to curb oil extraction and oil use today. The reason is that the announcement of an innovation of this sort would make the oil producers want to extract all the oil and sell it now before oil is superseded by a better technology, and given our arguments in this paper that the climate situation is particularly precarious now—and that it is particularly important to postpone, and not hasten, oil use—it would be all the more important to tax now. Thus, paradoxically, the appearance of a new, clean and cheap oil substitute calls for even firmer policy reactions in the short run. Having said all this, it is still obviously important to include technical change that involves the production of oil substitutes, or at least a back-stop technology. We do not yet know about future innovations, but economic theory can be used to capture the key features of how technology is endogenous and responsive to both environmental change and policy change.

In this paper, we represent the world as one region, so that any tax policy we consider should be interpreted as a global tax. This seems a reasonable starting point, since the fact that the emission of greenhouse gases is a global externality calls for equal taxes everywhere, assuming that there are no other frictions that differ across regions. In fact, if one large country, or the E.U. were to unilaterally impose a tax, the effect on the global externality would be small, since this would imply larger consumption of energy in the countries without the tax; in the absence of other frictions, there would be no effect at all of the unilateral tax increase. However, it is apparent from rounds of global negotiations that it is difficult to establish agreement on a global policy. One must therefore think about how the costs and benefits are unevenly distributed across the different countries/economies in the world. Nordhaus considers 8 regions in his model—defined both by economic and geographic characteristics—and this allows a better account of energy use, but more importantly it allows a better description of the differential damages that different regions will incur. Nordhaus, however, does not study the differential impact of taxes, and it would be important for our purposes to include heterogeneity for this reason. One obvious element of heterogeneity that involves conflict is that between oil producers and oil consumers; another is the free-riding problem, whereby one small oil-consuming country would like to use zero taxes, so long as all other countries make sure to combat the global externality. Thus, in more realistic descriptions of the world economy, various conflicts and commitment problems must be studied.

Our model also abstracts from the possibility of monopoly power in the oil-producing industry. It is, for example, possible that an oil monopoly would not want to extract all the oil, because that would lower their total profit. Relatedly, we make no efforts in the present paper to account for the historical data on oil prices, which seem difficult to do without taking market power into account. More generally, the present model does not deliver on some quantitative dimensions: it predicts current oil use that is much too low, it has difficulty in producing peak oil scenarios (either for past or future data), and it has difficulty in explaining that oil is simultaneously extracted at such different marginal extraction costs across the world (Norway, it seems, should not be extracting any oil yet, and nevertheless they extract their reserves at a much faster rate than do low-cost countries in the Middle East!). We need to extend our current setting to deal with these discrepancies.

References

Dasgupta, P. S. and Heal, G. M.: 1974, The optimal depletion of exhaustible resources, *Review of Economic Studies* **Symposium 1974**, 2–28.

Hotelling, H.: 1931, The economics of exhaustible resources, *Journal of Political Economy* **39**, 137–175.

Nordhaus, W. D. and Boyer, J.: 2000, *Warming the World – Economic Models of Global Warming*, The MIT Press.

Sinn, H.-W.: 2007, Pareto optimality in the extraction of fossil fuels and the greenhouse effect: A note. NBER WP 13453.

Solow, R. M.: 1974, Intergenerational equity and exhaustible resources, *Review of Econmic Studies* **41**, 29–45.

Stiglitz, J. E.: 1974, Growth with exhaustible natural resources: Efficient and optimal growth paths, *Review of Econmic Studies* **41**, 123–137.

7 Appendix

Deriving the first order condition for R_{t+1}.

$$\max_{\{K_{t+1},R_{t+1},C_t,S_t\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t U\left(S\left(S_t\right) F\left(A_t, K_t, N_t, R_t - R_{t+1}, A_t^e\right) + (1-\delta) K_t - Q\left(R_{t+1,} R_t, A_t^r\right) - K_{t+1}\right)$$
$$+\beta^t \lambda_t \left(S_{t+1} - (1-\varphi) S_t - R_t + R_{t+1}\right)$$

$$-U'\left(C_t\right)\left(F_E\left(t\right) - Q_1\left(R_{t+1,} R_t, A_t^r\right)\right) + \lambda_t + \beta U'\left(C_{t+1}\right)\left(F_E\left(t+1\right) - Q_2\left(R_{t+2,} R_{t+1}, A_{t+1}^r\right)\right) - \beta\lambda_{t+1} = 0$$

$$U'\left(C_t\right)\left(F_E\left(t\right) + Q_1\left(R_{t+1,} R_t, A_t^r\right)\right) - \lambda_t = \beta\left(U'\left(C_{t+1}\right) F_E\left(t+1\right) - Q_2\left(R_{t+2,} R_{t+1}, A_{t+1}^r\right)\right) - \beta\lambda_{t+1}$$
$$U'\left(C_t\right)\left(F_E\left(t\right) - q\left(R_{t+1}\right)/A_t^r\right) - \lambda_t = \beta U'\left(C_{t+1}\right)\left(F_E\left(t+1\right) - q\left(R_{t+1}\right)/A_{t+1}^r\right) - \beta\lambda_{t+1}$$

7.1 Proof of proposition 3

From the definition of Λ_t^s we have

$$\begin{aligned}
\frac{\Lambda_{t+1}^s}{\Lambda_t^s} &= \frac{U'\left(C_t\right)}{U'\left(C_{t+1}\right)} \frac{\sum_{s=1}^{\infty} (1-\varphi)^{s-1} \beta^s U'\left(C_{t+1+s}\right) d_{t+1+s}}{\sum_{s=1}^{\infty} (1-\varphi)^{s-1} \beta^s U'\left(C_{t+s}\right) d_{t+s}} \\
&= \frac{U'\left(C_t\right) \Lambda_{t+1}^s}{\beta U'\left(C_{t+1}\right) d_{t+1} + (1-\varphi) \beta U'\left(C_{t+1}\right) \Lambda_{t+1}^s} \\
&= \frac{U'\left(C_t\right)}{\beta U'\left(C_{t+1}\right)} \frac{\Lambda_{t+1}^s}{(1-\varphi) \Lambda_{t+1}^s + d_{t+1}} \\
&= \rho_{t+1} \frac{\Lambda_{t+1}^s}{(1-\varphi) \Lambda_{t+1}^s + d_{t+1}}.
\end{aligned}$$

Using the optimality condition $\rho_{t+1} = \frac{\varepsilon'_{t+1} - \Lambda_{t+1}^s}{\varepsilon_t - \Lambda_t^s}$, we then obtain

$$\begin{aligned}
\frac{\Lambda_{t+1}^s}{\Lambda_t^s} &= \frac{\varepsilon'_{t+1} - \Lambda_{t+1}^s}{\varepsilon_t - \Lambda_t^s} \frac{\Lambda_{t+1}^s}{(1-\varphi) \Lambda_{t+1}^s + d_{t+1}} \\
\frac{(1-\varphi) \Lambda_{t+1}^s + d_{t+1}}{\Lambda_t^s} &= \frac{\varepsilon'_{t+1} - \Lambda_{t+1}^s}{\varepsilon_t - \Lambda_t^s} \\
\frac{d_{t+1} - \varphi \Lambda_{t+1}^s}{\Lambda_t^s} &= \frac{\varepsilon_t}{\varepsilon_t - \Lambda_t^s} \left(\frac{\varepsilon'_{t+1}}{\varepsilon_t} - \frac{\Lambda_{t+1}^s}{\Lambda_t^s}\right).
\end{aligned}$$

From here, since $\varepsilon_t - \Lambda_t^s$ cannot be negative in an optimal allocation, the result follows. ∎

7.2 Extraction costs

In this section, we extend the simple calibrated example to allow for extraction costs. In order for the analysis to remain as simple as above, we assume that output (or equivalently, capital and labor with same production function as output) is used for extracting oil. Furthermore, we allow the extraction cost to depend on how much fossil fuel has been extracted before as well as on an exogeneous technology trend.

Specifically, we assume that

$$Q\left(R_{t+1}, R_t, A_t\right) = \tilde{Q}(R_{t+1}, R_{t+1}, A_t^r) S\left(S_t\right) A_t K_t^{\alpha} E_t^{\gamma},$$

where

$$\frac{1}{A_t^r} \int_{R_{t+1}}^{R_t} \tilde{q}\left(R\right) dR$$

implying that the aggregate budget constraint is

$$Y_t \equiv S\left(S_t\right) A_t K_t^{\alpha} E_t^{\gamma} \left(1 - \tilde{Q}(R_{r+1}, R_t, A_t^r)\right) = C_t + K_{t+1}$$

Using the formulas above, the Hotelling equation is

$$\rho_{t+1} = \frac{\varepsilon_{t+1}' - \Lambda_{t+1}^s}{\varepsilon_t - \Lambda_t^s}$$

where

$$\begin{aligned} \varepsilon_t &\equiv \frac{\gamma Y_t}{E_t} - \frac{Y_t}{1 - \tilde{Q}(R_{r+1}, R_t, A_t^r)} \frac{\tilde{q}\left(R_{t+1}\right)}{A_t^r}, \\ \varepsilon_{t+1}' &\equiv \frac{\gamma Y_{t+1}}{E_{t+1}} - \frac{Y_{t+1}}{1 - \tilde{Q}(R_{r+2}, R_{t+1}, A_{t+1}^r)} \frac{\tilde{q}\left(R_{t+1}\right)}{A_{t+1}^r}. \end{aligned}$$

ε_t measures the privat value (marginal product of fuel minus extraction costs) at the end of period t. ε_{t+1}' measures the same value at the beginning of period $t+1$.

The Euler equation remains

$$\rho_{t+1} = \frac{U'\left(C_t\right)}{U'\left(C_{t+1}\right) \beta} = \frac{Y_{t+1}}{Y_t \beta}. \tag{14}$$

Using this in the Hotelling equation, we get

$$\frac{1}{\beta} = \frac{\gamma \frac{1}{E_{t+1}} - \frac{1}{1 - \tilde{Q}(R_{r+2}, R_{t+1}, A_{t+1}^r)} \frac{\tilde{q}(R_{t+1})}{A_{t+1}^r} + \sum_{j=1}^{\infty} \left(1 - \varphi\right)^{j-1} \beta^j \frac{S'(S_{t+j+1})}{S(S_{t+j+1})}}{\gamma \frac{1}{E_t} - \frac{1}{1 - \tilde{Q}(R_{r+1}, R_t, A_t^r)} \frac{\tilde{q}(R_{t+1})}{A_t^r} + \sum_{j=1}^{\infty} \left(1 - \varphi\right)^{j-1} \beta^j \frac{S'(S_{t+j})}{S(S_{t+j})}} \tag{15}$$

Noting that $R_{t+1} = R_t - E_t$, we note given R_t, E_t, the technology trends and the externalities, this determines E_{t+1}. The solution is easily found by e.g., a shooting method as above.

We now specifying an increasing extaction cost function, for simplicity the affine function $\tilde{q}\left(SE\right) = \sigma_0 + \sigma_1\left(R_0 - R_t\right)$,implying $\tilde{Q}(SE_{t-1}E_t, A_t^r) = \frac{1}{A_t^r}E_t\left(\sigma_0 + \sigma_1\left(R_0 - R_t + \frac{E_t}{2}\right)\right)$.

We have set $\sigma_0 = 4 * 10^{-5}$ and $\sigma_1 = 8 * 10^{-6}$ and the growth rate of extraction technology is set to $1/\beta$. This generates extraction costs in *laissez-faire* starting at 24% of the price and peaking at 36%. Higher σ_1 implies that the marginal social value of fuel is zero.

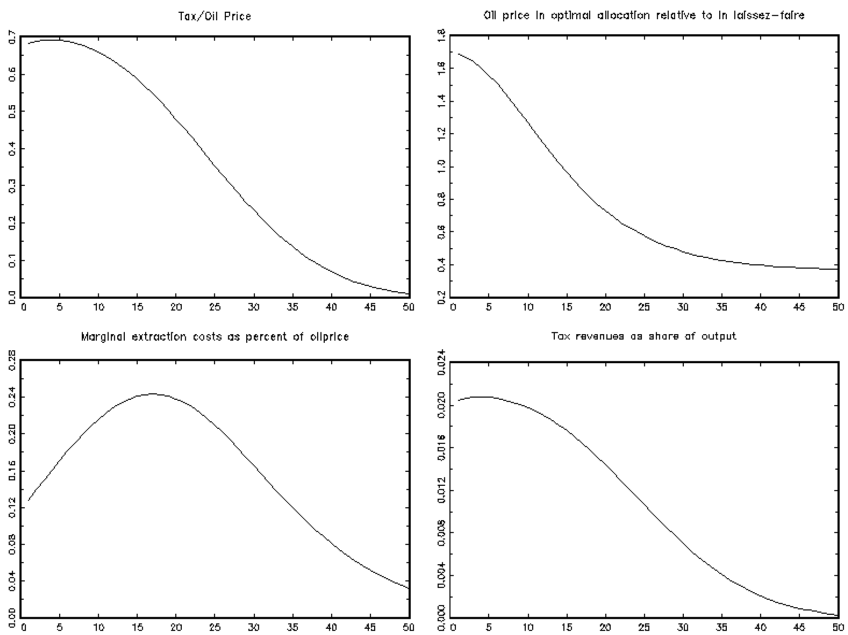

Figure A1.

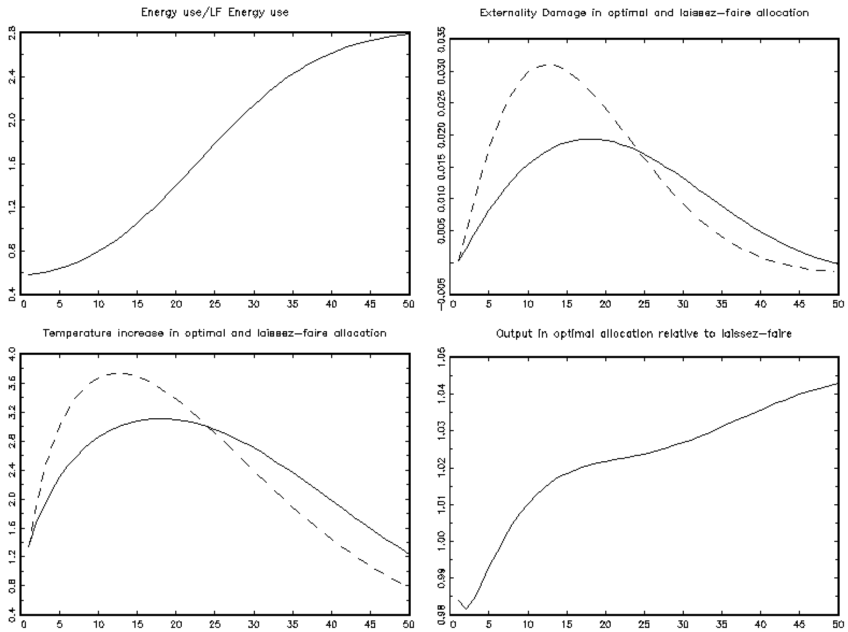

Figure A2.

Comparing the two cases with extraction costs with the base line case without, we find some differences. The oilprice relative to *laissez-faire* is slightly lower than in the case with extraction costs. The general tendency of falling taxes remains but the tax as a share of the price increases slightly the first few periods. The the tax rate starts at 68.2% of the market price, somewhat less than without extraction costs. It then increases and peaks in the fourth period when it is 69.2%. After that it montonically falls.

We can understand this by recalling that the externality damage, Λ_{t+1}^{s} increases over the first few periods, calling for an increasing tax per unit of fuel. Futhermore. the inclusion of extraction costs provides a market incentive to postpone extraction as long as marginal costs grows slower than at the rate of interest. This incentive seems strong enough for the initial periods not to need to be supported by falling tax rates. With falling extraction costs, the tendency to postpone extraction becomes stronger. To investigate this mechanism, we experimented with changing the growth rate of A_t^r. Increasing growth rate in extraction technology, lead to a more pronunced hump. However, also with a growth rate as high as 50% per period, the increase is still relatively modest, from 70.4% to 76.4% in the fifth period.