CESifo Working Papers

www.cesifo.org/wp

Optimal Carbon Tax with a Dirty Backstop
Oil, Coal, or Renewables?

Frederick van der Ploeg
Cees Withagen

CESIFO WORKING PAPER NO. 3334
CATEGORY 9: RESOURCE AND ENVIRONMENT ECONOMICS
JANUARY 2011

CESifo *Working Paper No. 3334*

Optimal Carbon Tax with a Dirty Backstop
Oil, Coal, or Renewables?

Abstract

Optimal climate policy is studied. Coal, the abundant resource, contributes more CO2 per unit of energy than the exhaustible resource, oil. We characterize the optimal sequencing oil and coal and departures from the Herfindahl rule. “Preference reversal” can take place. If coal is very dirty compared to oil, there is no simultaneous use. Else, the optimal outcome starts with oil, before using oil and coal together, and finally coal on its own. The “laissez-faire” outcome uses coal forever or starts with oil until it is no longer profitable to do so and then switches to coal. The optimum requires a steeply rising CO2 tax during the oil-only phase and a less steeply rising CO2 tax during the subsequent oil-coal and coal-only phases to avoid the abrupt switch from oil to coal thus leaving a lot of oil in situ. Finally, we analyze the effects on the optimal transition times and carbon tax of a carbon-free, albeit expensive backstop (solar or wind). Without a carbon tax, a prohibitive coal tax leads to less oil in situ, substantially delays introduction of renewable, and thus curbs global warming substantially. Subsidizing renewables to just below the cost of coal does not affect the oil-only phase. The gain in green welfare dominates the welfare cost of the subsidy if the subsidy gap is small and the global warming challenge is acute.

JEL-Code: Q300, Q420, Q540.

Keywords: Herfindahl rule, Hotelling rule, non-renewable resource, dirty backstop, coal, global warming, carbon tax, renewables, tax on coal, subsidy on renewables.

Frederick van der Ploeg
Department of Economics
University of Oxford
Manor Road Building
UK – Oxford, OX1 3UQ
rick.vanderploeg@economics.ox.ac.uk

Cees Withagen
Department of Economics
VU University Amsterdam
De Boelelaan 1105
NL – 1081 HV Amsterdam
cwithagen@feweb.vu.nl

12 January 2011
Support from the BP financed Oxford Centre for the Analysis of Resource Rich Economies is gratefully acknowledged. We are grateful to Michel Moreaux and Tony Venables for helpful suggestions.

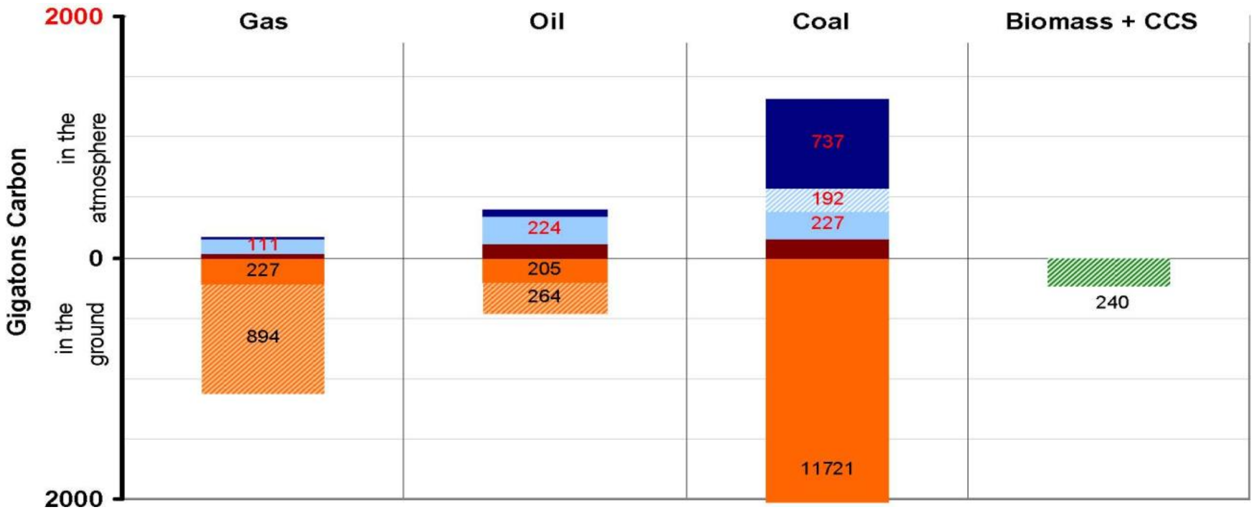
1. Introduction

Oil[3] will last at most another half a century after which society has to switch to alternative sources of energy. If no breakthrough on inventing viable and cost-effective carbon-free energy has been realized by then, the world has to make do with coal which could last for another three or four centuries. Oil is relatively expensive to extract compared with coal (see fig. 1), but it has the advantage of less CO2 emissions per unit of energy and thus less global warming damages. Wind and solar energy are clean, albeit expensive backstops, but account for at most a few percentage points of global energy supplies. Furthermore, they suffer from intermittence and are difficult to roll out on a large scale. Our objective is to analyze the optimal depletion of oil in the face of having to switch to coal as oil reserves become less accessible, taking due account of the social costs of global warming. The optimal policy of extracting oil and combating climate change must take into account the optimal order in which oil and coal have to be extracted. In doing so, differences in extraction costs for the various sources of energy as well as differences in the contributions the resources make to climate change play a role. We explicitly consider exhaustibility of oil with extraction costs rising as reserves diminish, and suppose that coal is available in limitless supply and is a perfect substitute for oil. The extraction cost advantage of coal thus increases over time as more oil is depleted. As the atmospheric concentration of CO2 is increased, marginal global warming damages rise so that the social cost of extracting oil increases more strongly than the private cost of extracting oil as oil fields are depleted. We abstract from R&D into developing new, clean forms of energy (e.g., Bosetti et al., 2009; Aghion et al., 2009; Acemoglu et al., 2010), so our analysis offers a conservative guide to what should be done if there is no hope of a technological fix for global warming.

In contrast to what is optimal if there is no dirty backstop, there may be a phase where it is socially optimal to use oil and the backstop, coal, simultaneously (cf., Tahvonen, 1979). However, if coal is very dirty compared to oil, we show that there is no simultaneous use of oil and coal. The socially optimal outcome is likely to start with oil, before moving on to using oil and coal together, and finally coal. We show that the "laissez-faire" market outcome never has simultaneous use of oil and coal. It either uses coal forever if it is cheap enough, or in the more realistic case starts with oil until it is no longer profitable to do so and then switches to using only coal. We show that the optimum policy requires a steeply rising CO2 tax during the oil-only phase followed by a less steeply rising CO2 tax during the subsequent oil-coal and coal-only phases. The reason for the steep initial rises in the carbon tax is to make sure that the market does not quickly and abruptly stop using oil, leaving a lot of oil in situ, and moving to use the relatively dirty coal as its energy source.

[3] Throughout this paper we refer to oil when we mean oil or gas.

Figure 1: Coal Reserves dominate Gas and Oil Reserves



Key: The CO2 that has been emitted and will in the future be emitted under the scenario that CO2-equivalent levels are stabilized at 400 parts per million compatible with the objective of limiting the increase in global temperature to a maximum of 2 degrees centigrade over pre-industrial levels is: estimated energy consumption since 1990 has already led to 227, 224 and 111 Giga tons of carbon for coal, oil and gas, respectively; estimated additional consumption to hit targets 737 Gigaton for coal, but almost none for gas and oil and coal plus CCS much less, namely 192 Giga ton of carbon. The remaining bit immediately above the line is Giga tons of carbon emitted in the atmosphere due to cumulative historical energy consumption. Biomass and CCS has and will not lead to CO2 emissions.
Under-the-ground reserves imply the following further future CO2 emissions: probable and proven resources and reserves for gas, respectively, 894 and 227 Giga ton; oil 264 and 205 Giga ton of carbon; reserves for coal 11,721 Giga ton of carbon; biomass plus carbon, capture and storage (CCS) 240 Giga ton of carbon unless it is fully sequestrated in which case there will be zero CO2 emissions.
Source: Edenhofer and Kalkuhl (2009)

Finally, we consider a carbon-free, albeit expensive backstop (e.g, solar or wind energy) that kicks in eventually and takes over from coal. We show that this reduces the optimal carbon tax. The optimal carbon tax steeply rises during the oil-only phase, less steeply during the oil-coal and coal-only phases, and finally remains a positive constant during the renewable phase. Interestingly, initially more oil is pumped and, although coal is phased in earlier, much less coal is used for a shorter period of time. Hence, the CO2 concentration is less with renewables. A lower cost of renewables brings forward and shortens the only-coal period and leads to slightly more aggressive pumping of oil in the earlier periods. Although the optimal carbon tax falls, the stock of CO2 and global warming are curbed. In the absence of a carbon tax, we show that a tax on coal just high enough to ensure that the market never uses coal (or a moratorium on coal) leads to less and possibly no oil left in situ and substantially delays the introduction of renewables. This curbs global warming damages substantially and consequently this second-best policy can do surprisingly well. Interestingly, the first-best policy does a little better by using coal alongside oil and for a while on its own. The alternative of subsidizing renewables to such an extent that they are a little cheaper than coal neither affects oil left in situ nor the time of transition to the backstop in the absence of a carbon tax, but does stabilize the stock of CO2 once renewables are introduced. This gain in green welfare outweighs the cost of the lump-sum taxes needed to finance the subsidy if the required subsidy is small and the global warming challenge is not too acute. These two policies suggest that a green paradox[4] need not occur if coal rather than fossil fuel is the alternative source of energy.

Our results shed light on the Herfindahl rule which says that reserves with lowest cost to extract should be used first (Herfindahl, 1967).[5] It has recently been shown for an economy where all resources are non-renewable and have the same (zero) cost of extraction that, if an economy starts below its CO2 cap, it may be socially optimal to start using coal, postpone using clean natural gas, and finally use coal again (Chakravorty et al., 2008). In contrast to what the Herfindahl rule seems to suggest, it is not optimal to start with resources with the lowest environmental cost but to start with coal in order to benefit from natural decay of the concentration of CO2 in the atmosphere. Only if gas is abundant will it be used first. We extend these results to allow for heterogeneity in the cost of extraction as well as differences in pollution intensities, and also obtain “preference reversals” in the sense of starting with coal and ending with coal after an intermediate phase where oil is used (cf., Chakravorty et al., 2008) and the possibility that coal and oil are used at the same time (cf., Tahvonen, 1979; Kemp and Long, 1980; Chakravorty and Krulce, 1994). We thus show that it may be optimal to use the dirty resource first, then a period where the cleaner resource is used (possibly alongside the dirty resource), and then finally the dirty resource again.

Section 2 analyzes the socially optimal transition from conventional oil to coal with climate externalities and stock-dependent extraction costs. Damages from CO2 emissions can be modelled through a negative externality in production (cf., Heal, 1985; Sinn, 2008ab), but we suppose that they directly damage social welfare. We abstract from capital accumulation.[6] Section 3 studies the outcome in a market economy and shows how the social optimum can be sustained with an appropriately designed carbon tax. Section 4 analyzes what happens if carbon-free renewables kick in eventually and investigates the effect of lowering the cost of renewables and a coal tax. Section 5 concludes and discusses policy implications.

[4] This says that subsidizing a carbon-free backstop if a carbon tax is politically or otherwise infeasible can have adverse climate effects as the anticipation of oil being made obsolete more quickly by such renewables encourages oil companies to pump faster (Sinn, 2008ab; Hoel, 2008; Gerlagh, 2009; Grafton et al., 2010). However, the Green Paradox does not occur if the backstop subsidy ensures that more oil is kept in situ and thus that less CO2 is emitted (van der Ploeg and Withagen, 2010a).

[5] This result has been studied in a general equilibrium framework (Kemp and Long, 1980; Lewis, 1982), in a situation with setup costs (Gaudet et al., 2001), and under heterogeneous demands (Chakravorty and Krulce, 1994).

[6] Golosov et al. (2010) and van der Ploeg and Withagen (2010b) also study a general equilibrium growth model of fossil fuels and a backstop fuel with capital accumulation, but do not consider dirty backstops such as coal.

2. Social optimum

We study optimal extraction of oil with coal being phased in when energy prices become high enough. The backstop coal is a perfect substitute for oil and its supply is infinitely elastic. We add a convex function in past CO_2 emissions to the felicity function to capture the damage done by accumulated CO_2 emissions into the atmosphere from burning oil or coal. We abstract from natural decay of the stock of CO_2 in the atmosphere and thus suppose that this stock evolves according to:

(1) $$\dot{E}(t) = q(t) + \psi x(t), \quad E(0) = E_0, \quad \psi > 1,$$

where *E*, *q* and *x* denote the CO_2 concentration in the atmosphere, oil use, and coal use, respectively. The emission coefficient of oil is normalized to one, so that the emission coefficient of relatively dirty coal ψ is bigger than one.[7] With quasi-linear preferences, the social planner's problem then reads:

(2) $$\text{Max} \int_0^\infty \exp(-\rho t)[U(q(t) + x(t)) - G(S(t))q(t) - bx(t) - D(E(t))]\text{d}t$$

subject to (1), the non-negativity condition $x(t) \geq 0$ and the depletion equation for oil,

(3) $$\dot{S}(t) = -q(t),\ q(t) \geq 0,\ S(t) \geq 0,\ S(0) = S_0, \text{ given},$$

where ρ denotes the rate of time preference, *G* per unit extraction cost of oil, and *b* cost of supplying coal. Equation (3) implies that total current and future oil depletion cannot exceed oil reserves, $\int_0^\infty q(t)\text{d}t \leq S_0$.

Assumption 1: Instantaneous utility is concave, $U' > 0$ and $U'' < 0$, global warming damages are convex, $D' > 0$ and $D'' > 0$, unit oil extraction costs increase as less reserves remain, $G' < 0$.

So we suppose that marginal global warming damages are high when the atmospheric CO_2 concentration is already high and that it becomes more expensive to extract oil as more of the more accessible oil fields have been mined. We will consider both the case that, ignoring global warming damages, coal is always cheaper than oil, i.e., $b < G(S_0)$, and the case that oil is currently cheaper than coal, $b > G(S_0)$. The latter case is not unrealistic; e.g., the cost of extracting oil in the Gulf is less than the cost of coal (including transport cost) from South Africa.

[7] Our model extends the model analysed by Hoel and Kverndokk (1996) to allow for a dirty backstop, but we restrict attention to the case where there is no natural decay.

2.1. First-order conditions

The current-value Hamiltonian function for the social planner is defined by:

(4) $$H(q,x,S,\lambda,\mu)\equiv U(q+x)-G(S)q-bx-D(E)-\lambda q-\mu\left(q+\psi x\right),$$

where λ is the social value of oil and μ (> 0) is the social cost of the stock of CO2 in the atmosphere. The necessary conditions for a social optimum are:

(5a) $U'(q+x)-G(S)\le\lambda+\mu,\ q\ge 0,$ c.s.,

(5b) $U'(q+x)-b\le\psi\mu,\ x\ge 0,$ c.s.,

(5c) $\dot{\lambda}=\rho\lambda+G'(S)q,$

(5d) $\dot{\mu}=\rho\mu-D'(E),$

(5e) $\lim_{t\to\infty}\exp(-\rho t)\left[\lambda(t)S(t)-\mu(t)E(t)\right]=0,$

where c.s. stands for complimentary slackness. We define the cheapness of coal relative to oil, both in terms of unit extraction costs and the present value of marginal global warming damages as:

(6) $$\Omega(S(t),Y(t))\equiv\rho\left[G(S(t))-b-(\psi-1)\frac{D'(E_0+S_0-S(t)+Y(t))}{\rho}\right],$$

where Y denotes the accumulated CO2 emissions from using coal in the past,

(7) $$Y(t)=\psi\int_0^t x(s)ds,\quad \dot{Y}=\psi x.$$

Since the damage function is convex and coal is dirtier per unit of energy than oil, we have $\Omega_Y=(1-\psi)D''(E)<0.$ Since oil is more expensive to extract if oil reserves are low, the private component of the cost advantage of coal rises as oil reserves are depleted (i.e., $\Omega_S=\rho G'(S)<0$ if $D'=0$). This is the case for the market outcome, which does not internalize global warming externalities. However, taking account of the social cost of warming, we see that the cost advantage of using coal rather than oil might fall as oil reserves are depleted and the CO2 concentration rises (i.e., $\Omega_S>\rho G'(S)$ if $D'>0$). This highlights a key dilemma for climate change policy: the market might want to switch to coal as oil reserves are depleted whilst duly taking account of global warming externalities one might want to carry on to use oil rather than switching to dirtier coal. The sign of $\Omega_S=\rho G'(S)+(\psi-1)D''(E)$ is thus negative if the positive effect of a falling stock of remaining oil reserves on oil extraction costs dominates the mitigating effect on global warming damages of using oil rather than coal; and is positive, otherwise. For example, with the functional forms $G(S)=\gamma-\delta S$ and $D(E)=\frac{1}{2}\kappa E^2$, we have $\Omega_S<0$ iff $\delta>(\psi-1)\kappa/\rho$. Furthermore, we suppose that Ω is monotonically increasing, or monotonically decreasing in S (which is the case for the functional forms that we in the analysis of section 2.3 and thereafter in our policy simulations).

2.2. Optimal sequence of using oil, coal or both

The next proposition helps to characterize the optimal sequencing of oil and coal use on the optimal path.

Proposition 1: When to use oil, coal or both?

Consider an interval of time $V=[t_1,t_2]$ with $0\le t_1<t_2$.

i. If both oil and coal are used at the same time, $q(t)>0$ and $x(t)>0,\forall t\in V$, then $\Omega(S(t),Y(t))=0$ and $\partial\Omega(S(t),Y(t))/\partial S(t)<0,\forall t\in V$.

ii. If $\Omega(S(t),Y(t))=0,\forall t\in V$, then *either* oil and coal are both used, $q(t)>0$, $x(t)>0$, and $\partial\Omega(S(t),Y(t))/\partial S(t)<0,\forall t\in V$, or *neither* oil nor coal is used, $q(t)=x(t)=0,\forall t\in V$.

iii. If $\Omega(S(t),Y(t))>0,\forall t\in V$, oil is not used, $q(t)=0,\forall t\in V$.

iv. If $\Omega(S(t),Y(t))<0$ and $S(t)>0,\forall t\in V$, coal is not used, $x(t)=0,\forall t\in V$.

Proof:

i. Making use of the c.s. conditions in (5a) and (5b), we have $U'(q+x)=G(S)+\lambda+\mu=b+\psi\mu$ or $G(S)-b+\lambda+(1-\psi)\mu=0$. $\Omega(S(t),Y(t))=0$ then follows immediately from (5a)-(5d) and differentiation of $G(S)-b+\lambda+(1-\psi)\mu$ with respect to time: $0=G'(S)\dot{S}+\dot{\lambda}+(1-\psi)\dot{\mu}=\Omega$. Then $\Omega_S\dot{S}(t)+\Omega_Y\dot{Y}(t)=0$. We have $\dot{S}(t)<0,\dot{Y}(t)>0,\Omega_Y<0$, so $\Omega_S<0$ must hold.

ii. If $\Omega(S(t),Y(t))=0$ along V, it cannot be excluded that there is no supply of fuel at all. But if only fossil fuel is supplied, Y is constant and S is decreasing. This contradicts constancy of Ω. And if only backstop energy is supplied, we get a contradiction as well.

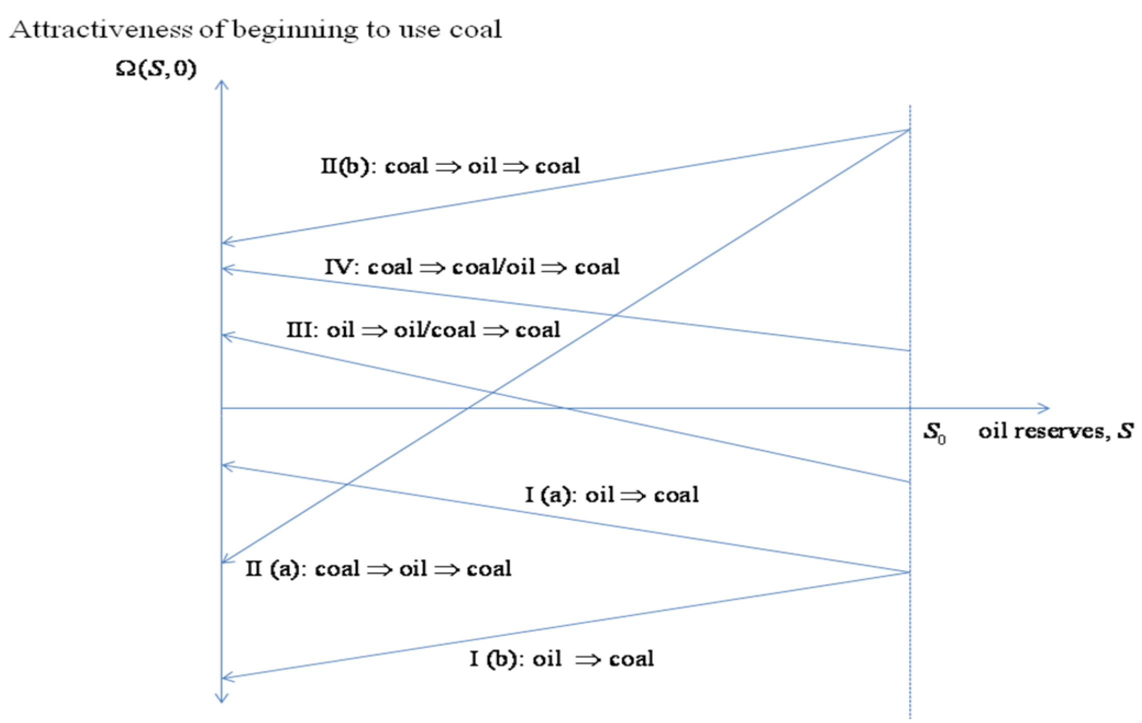
iii. Suppose $\Omega(S(t),Y(t))>0$ for all $t\in V$ and $q(t)>0$ along some sub-interval $M\in V$. Hence, along M we have $U'(q(t)+x(t))=G(S(t))+\lambda(t)+\mu(t)\le b+\psi\mu(t)$. So, $G(S(t))-b+\lambda(t)+(1-\psi)\mu(t)\le 0$. The derivative of the left-hand side with respect to time is $G'(S)\dot{S}+\dot{\lambda}+(1-\psi)\dot{\mu}(t)=-G'(S)q+\rho\lambda+G'(S)q+(1-\psi)(\rho\mu-D'(E))=$

$\rho(\lambda+(1-\psi)\mu-(1-\psi)D'(E)/\rho)\leq\rho(b-G(S)+(\psi-1)D'(E)/\rho)\leq 0.$ The latter inequality is strict in the interior of M. Along any programme $G(S)$ is non-decreasing and $D'(E)$ is non-decreasing. Therefore $G(S(t))-b+\lambda(t)+(1-\psi)\mu(t)\to-\infty$ as $t\to\infty$. But $\lambda(t)\geq 0$ and $\mu(t)$ is bounded from above because otherwise E will become a constant eventually and the transversality condition is violated. Therefore, $q(t)=0$ for all $t\in V$.

iv. Suppose $H(S(t),Y(t))<0$ for all $t\in V$ and $x(t)>0$ along some subinterval $M\in V$. Hence, along M we have $U'(q(t)+x(t))=b+\psi\mu(t)\leq G(S(t))+\lambda(t)+\mu(t)$. So, $G(S(t))-b+\lambda(t)+(1-\psi)\mu(t)\geq 0$. The derivative of the left hand side with respect to time is
$G'(S)\dot{S}+\dot{\lambda}+(1-\psi)\dot{\mu}(t)=-G'(S)q+\rho\lambda+G'(S)q+(1-\psi)(\rho\mu-D'(E))=$
$\rho(\lambda+(1-\psi)\mu-(1-\psi)D'(E)/\rho)\geq\rho(b-G(S)+(\psi-1)D'(E)/\rho)\geq 0.$
Along this part of the programme $G(S)$ is constant and $D'(E)$ is increasing. No transition to simultaneous use or only fossil fuel use can take place. Therefore
$G(S(t))-b+\lambda(t)+(1-\psi)\mu(t)\to\infty$ as $t\to\infty$. But $\mu(t)\geq 0$ for all $t\geq 0$. Hence $\dot{\lambda}(t)/\lambda(t)=\rho$. But this contradicts that $S(t)>0$ for all $t\geq 0$. Q.E.D.

We will use proposition 1 to distinguish various regimes for the optimal path, where we use the coal price b as one of the pivotal parameters to delineate the various regimes. These regimes are portrayed in fig. 2.

Figure 2: Regimes of oil and coal use and stock of remaining oil reserves



The regimes are determined by the relative cheapness of coal, Ω, being increasing or decreasing in the stock of oil which matters for the occurrence of simultaneous use, by $\Omega(0,0)$ being positive or negative which matters for what happens after a phase with only oil use, and by $\Omega(S_0,0)$ being positive or negative which matters for what happens initially. For a given CO2 concentration Y due to emissions from coal, one can draw $\Omega(S,Y)$ below $\Omega(S,0)$. If we start on the curve $\Omega(S,0)$ and coal is not used, we move along this curve to the left as the stock of oil decreases. If we start on the curve $\Omega(S,0)$ and only coal is used, we move vertically downward in the plane as the stock of oil remains constant and the pollution stock due to coal increases. Along intervals of time with simultaneous use, we have $\Omega(S,Y)=0$ and we are on a downward-sloping curve (as the stock of oil decreases, we move to the left, but Y increases, so we move downward which requires the curve to slope downwards).

Regimes I(a) and (b): Use oil, then switch to coal

Regime I (a) corresponds to the situation where it is optimal to first use oil and to switch to coal when oil has run out. It occurs if the coal price *b* is high enough to ensure that it is never advantageous to start using coal, i.e., $\Omega(S, 0) < 0$, for all $0 \leq S \leq S_0$. Since $\Omega_Y < 0$, we have $\Omega(S,Y) < 0$ for all $0 \leq S \leq S_0$ and $Y \geq 0$ as well. It follows from proposition 1 that there will never be simultaneous use of oil and coal. Initially, coal is not used because once the cost of global warming is taken into account it is expensive relative to oil. Coal kicks in, but only after oil has been fully exhausted (see proposition 1 (iv)). A necessary and sufficient condition to start using only oil is that the marginal utility of extracting the first drop of oil exceeds the marginal cost of extracting the first drop of oil including the discounted value of all future global warming damages associated with this first drop of oil, i.e., $U'(0) > G(S_0) + D'(E_0)/\rho$. We suppose that this condition is satisfied, so that there is an initial interval of time with only use of oil. The social cost of using oil during this phase, $G(S) + D'(E_0 + S_0 - S)/\rho$, increases over time as oil reserves are depleted.

We know from proposition 1 (iv) that coal will never be used unless oil is completely exhausted. So, we have to look for a condition for the full exhaustion of oil. Suppose there exists T_1 with $S(T_1) > 0$. Suppose we are at T_1 and consider what to do. First, it follows from proposition 1 (iv) that it is suboptimal to start using coal. Hence, we should either stop using energy altogether or continue to use oil. Now suppose that $U'(0) = G(S(T_1)) + D'(E_0 + S_0 - S(T_1))/\rho$. Extracting a marginal unit of oil yields utility $U'(0)$. The marginal cost consist of the marginal extraction cost and the discounted total marginal costs of global warming damages, $G(S(T_1)) + D'(E_0 + S_0 - S(T_1)) / \rho$. Hence, we will stop extracting oil (and using energy altogether) if $U'(0) = G(S(T_1)) + D'(E_0 + S_0 - S(T_1)) / \rho$. Then, if the marginal utility of the last drop of oil exceeds its social cost, i.e., $U'(0) > G(0) + D'(E_0 + S_0) / \rho,$ oil reserves will be fully exhausted, say at time T_2. Two possibilities arise in that case: either it does not pay to use coal, $U'(0) < b + \psi D'(E_0 + S_0)/\rho,$ so coal is never used; or it pays to use coal, $U'(0) \geq b + \psi D'(E_0 + S_0)/\rho,$ so coal takes over until some instant of time T_3 where it becomes unprofitable to use coal (where T_3 follows from the condition that $U'(0) = b + \psi D'(E_0 + S_0 + Y(T_2)) / \rho$ as reserves are fully exhausted). It may be that coal is never phased out, so that the switch time T_2 is infinite (see the example given below).

Regime I (b) is similar to regime I (a) except that the social disadvantage of using coal rather than oil becomes bigger rather than smaller as oil reserves are depleted ($\Omega_S(S, 0) > 0$), which occurs if the effect via marginal global warming damages dominates the effect via rising oil extraction costs. For the optimal sequence it does not matter. It is still not optimal to oil and coal alongside each other, because we have $\Omega(S,Y) < 0$ for all $S_0 > S \geq 0$ and all $Y \geq 0.$ Moreover, oil is still always cheaper than coal and it is thus optimal to start using oil and to switch to coal when oil has run out and it is profitable to use coal. Hence, in this regime it is optimal to start with oil and then either one exhausts oil fully and then switches to coal or one leaves oil in situ and never uses coal.

Regimes II(a) and (b): Use coal, then oil, and finally switch back to coal

Regime II (a) applies if oil is initially relatively more expensive than coal from a social perspective, and occurs for a relatively low coal price, *b*. In this regime there will never be simultaneous use (as Ω increases in S). From proposition 1(iii) we have that there is no oil use initially. If, in spite of the fact that the coal price is low compared to the previous two cases, we still have that the marginal utility of the coal falls short of the economic cost of oil plus the induced future marginal global warming damages caused by using coal, $U'(0) < b + \psi D'(E_0) / \rho,$ then neither coal nor oil is used. Oil is then more expensive than coal and coal is unprofitable from a social welfare perspective. However, in the more likely case that coal is profitable from a social perspective, at least initially, i.e., $U'(0) > b + \psi D'(E_0) / \rho$, we only use coal initially.

It could be that at some instant of time it is optimal to stop using coal ($U'(0) = b + \psi D'(E_0 + Y(t)) / \rho$), since the marginal global warming damages of accumulated coal use have become prohibitive. However, it may not be optimal at that point for oil to take over if the accumulated CO2 in the atmosphere resulting from past coal use is high enough (i.e., $U'(0) < G(S_0) + D'(E_0 + Y(t)) / \rho$ despite assumption 1). Otherwise, oil does take over from coal, and from that moment onwards the economy is in case I (b) again so that the oil phase will eventually be taken over by coal once again.

In regime II (b) the social cost of coal is initially less than that of oil and this relative attractiveness diminishes as oil reserves are depleted, but now coal stays more attractive than oil all the time. Now the optimal sequence is again to start with coal, then shift to using oil once coal has become prohibitive to use from a global warming perspective, and finally switch back to using coal again. There will never be simultaneous use of oil and coal, because Ω decreases as oil reserves are depleted and S falls. Moreover, it is not optimal to start using only oil from the outset. This regime of using coal, then oil and finally coal again is an illustration of what has been coined “policy reversal” (Chakravorty et al., 2008). Therefore, if $U'(0)$ is large enough, we start using only coal until we reach $\Omega(S_0, Y(t)) = 0$. From then on we are back in regime II (a). Whether oil reserves are fully exhausted or not thus depends on whether the social cost of extracting the last drop of oil is less or more than the social cost of coal.

Regimes III and IV: Use oil or coal, then oil and coal alongside each other, and finally coal only

Regime III applies if the social cost of oil is initially less than that of coal and if the social cost of using the last drop of oil exceeds that of coal. Once oil has run out, the social cost of coal is thus less than that of oil. In this case, Ω increases as oil is depleted and S falls. Hence, we could in principle have simultaneous use, but not initially. The optimal sequence has an initial phase with only oil, because coal is too expensive. Of course, $U'(0)$ needs to be large enough. Once we arrive at $\Omega(S, 0) = 0$ there will be an interval of time with simultaneous use of oil and coal. This phase may last forever, or just for a finite period of time, after which coal takes over indefinitely. Which possibility occurs depends on the parameters, as explained below. In regime IV coal is cheaper from a social perspective than oil and becomes even cheaper as oil reserves are depleted. Hence, we could in principle again have simultaneous use, but not initially. Initially we should use only coal provided that the initial price of energy exceeds the initial social cost of using the first unit of coal, i.e., $U'(0) > b + \psi D'(E_0) / \rho$. Once we arrive at $\Omega(S_0, Y(t)) = 0$ there will be a phase where oil and coal are used simultaneously. This phase may last forever, or just for a finite period of time until it is no longer profitable to use oil or until oil reserves are fully exhausted, after which coal takes over indefinitely.

The above discussion on the optimal sequences of using only oil, coal and oil alongside each other, and only coal are summarized in table 1. With the additional assumption that oil becomes infinitely expensive to extract as reserves go to zero, $\lim_{S\to 0} G(S)=\infty$, one always has Ω(0, 0) > 0 and thus regimes I (a) and (b), and case II (a) can be ruled out. For regimes III and IV we will show later (see proposition 3) that after the phase where oil and coal are used alongside each other there will only be a switch to using coal forever if oil reserves become fully exhausted.

Table 1: Regimes for optimal sequences of oil extraction and coal use

High coal price; last drop of oil cheap I. (a) $\Omega(S_0,0)<\Omega(0,0)<0$ (b) $\Omega(0,0)<0$ and $\Omega(S_0,0)<\Omega(S,0)<0$	First use oil; then switch to coal as long as it is profitable to do so. Fully exhaust oil reserves if the marginal utility of the last drop of oil exceeds its social cost.
Low coal price; last drop of oil cheap II. (a) $\Omega(S_0,0)>0>\Omega(0,0)$ (b) $\Omega(S_0,0)>\Omega(0,0)>0$	First use coal; then use oil if marginal global warming damages become prohibitive; finally switch back to using coal if there is no other option.
High coal price; last drop of oil expensive III. $\Omega(S_0,0)<0<\Omega(0,0)$	First use oil; then use oil and coal together. There will be a switch to using only coal forever if oil becomes fully exhausted.
Low coal price; last drop of oil expensive IV. $0<\Omega(S_0,0)<\Omega(0,0)$	First use coal; then use coal and oil together. There will be a switch back to using only coal if oil becomes fully exhausted.

Note that from (6) the advantage of starting to use coal is $\Omega(S,0)\equiv\rho\left[G(S)-b\right]-(\psi-1)D'(E_0+S_0-S)$ which for the functional forms $G(S)=\gamma-\delta S$ and $D(E)=\frac{1}{2}\kappa E^2$ gives

$$\Omega(S,0)/\rho=\gamma-b-\frac{(\psi-1)\kappa}{\rho}(E_0+S_0-S)-\delta S=\gamma-b-\frac{(\psi-1)\kappa}{\rho}(E_0+S_0)-\left[\delta-\frac{(\psi-1)\kappa}{\rho}\right]S$$

. From the first equality we see that a large emission-ratio of coal (ψ), a large marginal damage parameter (κ) and a small rate of time preference (ρ) make oil attractive compared to coal. However, from the second equality we see that also $\delta-(\psi-1)\kappa/\rho$ matters. If oil extraction costs rise rapidly with declining reserves (high δ), this expression is positive so with a falling stock of oil the cost advantage of coal is increasing over time. So, then there is a case for coal taking over if we start with oil only (as we suppose in section 3).

2.3. Characterization of the various phases

With our characterization of the optimal sequencing of coal and oil, propositions 2, 3, and 4 below describe the solution for the phases where, respectively, only coal, coal and oil, and only oil are used. For this purpose, we use the functional forms $U(x)=\alpha x-\frac{1}{2}\beta x^2$, $G(S)=\gamma-\delta S$ and $D(E)=\frac{1}{2}\kappa E^2$.

Proposition 2: Coal-only phase

Suppose there exists an interval of time $V=[T_1,T_2]$ with $0\le T_1\le T_2$ such that $q(t)=0$ and $x(t)>0$ along V. Then the optimal use of coal and accumulated CO2 emissions resulting from coal use are given by:

(8) $$x(t)=K_1e^{\lambda_1 t}+K_2e^{\lambda_2 t} \text{ and } Y(t)=\frac{\rho}{\psi\kappa}(\alpha-b)-\left[E_0+S_0-S(T_1)\right]+\hat{K}_1e^{\lambda_1 t}+\hat{K}_2e^{\lambda_2 t},\quad t\in V,$$

where $\lambda_1=\tfrac{1}{2}\rho+\tfrac{1}{2}\sqrt{\rho^2+4\psi^2\kappa/\beta}>\rho>0$ and $\lambda_2=\rho-\lambda_1<0,$ and K_1 and K_2 are to be determined by the boundary conditions and $\lambda_1\hat{K}_1=\psi K_1$ and $\lambda_2\hat{K}_2=\psi K_2$.

If $T_2=\infty,$ then $\alpha>b+\psi\kappa\left[E_0+S_0-S(T_1)\right]/\rho$ and coal use is given by

(8′) $$x(t)=\left(\frac{\psi\kappa}{\beta\lambda_1}\right)[Y^*-Y(T_1)]e^{\lambda_2(t-T_1)} \text{ and } Y(t)=Y^*+[Y(T_1)-Y^*]e^{\lambda_2(t-T_1)},\quad \forall t\ge T_1$$

with the long-run level of accumulated CO2 emissions resulting from coal use given by $Y^*\equiv\left(\frac{\rho}{\psi\kappa}\right)(\alpha-b)-\left[E_0+S_0-S(T_1)\right].$ Coal use vanishes asymptotically.

Proof: Along V we have $\alpha-\beta x-b=\psi\mu$ and $\dot{\mu}(t)=\rho\mu(t)-\kappa(E_0+S_0-S(T_1)+Y(t))$. This yields $\dot{x}=-\frac{\rho}{\beta}(\alpha-\beta x-b)+\frac{\psi\kappa}{\beta}\left(E_0+S_0-S(T)+Y\right)$ and $\dot{Y}=\psi x.$ This system displays saddlepoint stability with one positive characteristic root, λ_1, and one negative characteristic root, λ_2. We obtain $-\beta\ddot{x}+\rho\beta\dot{x}+\psi^2\kappa x=0$, which readily gives the first part of (8). For Y we get $\beta\ddot{Y}-\rho\beta\dot{Y}-\kappa\psi^2Y=-\rho\psi(\alpha-b)+\kappa\psi^2(E_0+S_0-S(T_1))$ yielding the second part of (8). If $T_2=\infty$, it should be the case that Y is finite. Since $\lambda_1>0$, we thus have $K_1=0.$ Making use of $\lambda_1\lambda_2=-\psi^2\kappa/\beta,$ we obtain (8′). To ensure that $Y^*>0,$ we require that $\alpha>b+\psi\kappa\left[E_0+S_0-S(T_1)\right]/\rho$. Hence, $x(t)\to 0$ and $Y(t)\to Y^*$ as $t\to\infty.$ The rest is straightforward. Q.E.D.

If coal is used forever, it starts with $x(T_1)\equiv\left(\rho/\beta\lambda_1\right)(\alpha-b)-\left(\psi\kappa/\beta\lambda_1\right)\left[E_0+S_0-S(T_1)\right]<(\alpha-b)/\beta$ and must then vanish asymptotically as the marginal global warming damages increase steadily as accumulated CO2 emissions resulting from coal increase over time. The rate at which use of the backstop falls increases with the rate at which marginal global warming damages increase (κ), the emission intensity of the backstop (ψ), and the sensitivity of demand for fossil fuels with respect to the price ($1/\beta$). Total use of the backstop (Y^*) is less if past use of fossil fuels has already led to a high concentration of CO2 emissions in the atmosphere and if the unit cost of the backstop (b) is high, but more if autonomous demand for energy (α) is high. Coal use starts especially low if marginal global warming damages are high (high ψ, κ, E_0, S_0, low $S(T_1)$). If global warming externalities are not internalized, as would be the case in the "laissez-faire" market economy, coal use would be higher.

Proposition 3: Oil-coal phase

Suppose there exists an interval of time $V=[T_1,T_2]$ with $0\le T_1\le T_2$ such that $q(t)>0$ and $x(t)>0$ along V. Then the time paths for the stock of oil reserves, oil use and coal use are given by:

$$S(t)=\frac{(\alpha-\gamma)(1-\psi)-b+\gamma}{\psi\delta}+L_1e^{\pi_1 t}+L_2e^{\pi_2 t},\quad q(t)=-\pi_1L_1e^{\pi_1 t}-\pi_2L_2e^{\pi_2 t},$$

(9)

$$x(t)=\frac{\alpha-\left[b+\psi\int_t^\infty D'(E_0+S_0-S(t'))e^{-\rho(t'-t)}dt'\right]}{\beta}+\pi_1L_1e^{\pi_1 t}+\pi_2L_2e^{\pi_2 t},\quad t\in V,$$

where $\pi_1=\tfrac{1}{2}\rho+\tfrac{1}{2}\sqrt{\rho^2+4\psi^2\kappa/\beta[1+\kappa(1-\psi)^2/\rho\delta]}>0$, $\pi_2=\rho-\pi_1<0$ and L_1 and L_2 are to be determined from the boundary conditions on fossil fuel stocks or extraction rates. If $(\alpha-\gamma)(1-\psi)-b+\gamma<0$ then the oil-coal phase lasts for a finite time and oil becomes fully exhausted. In the optimal program there will be a final phase where only coal is used. Coal use and accumulated CO2 emissions from coal use can be calculated from:

(9′)
$$x(t)=\left(\frac{\rho\delta-(\psi-1)\kappa}{\psi(\psi-1)\kappa}\right)q(t),\quad Y(t)=\frac{\left[(\psi-1)\kappa-\rho\delta\right]S(t)+\rho(\gamma-b)}{(\psi-1)\kappa}-E_0-S_0,\quad t\in V.$$

Proof: Using the necessary conditions and the fact that $\Omega(S,Y)=0$ along V gives the differential equation $\beta\ddot{S}-\rho\beta\dot{S}+\dfrac{\psi^2\delta\kappa}{-\delta-\kappa(1-\psi)^2/\rho}S=\dfrac{\psi\kappa\left[(\alpha-\gamma)(1-\psi)-b+\gamma\right]}{-\delta-\kappa(1-\psi)^2/\rho}$, which yields the expression for $S(t)$ in (9). Differentiating this expression yields the expression for $q(t)$ in (9). The expression for $x(t)$ in (9) then comes from the optimality conditions (5b) and (5d). Suppose exhaustion occurs at some instant of time T. Suppose $(\alpha-\gamma)(1-\psi)-b+\gamma<0$. Then we cannot have $T_2=\infty$. Oil will be fully depleted at some instant of time T and $\gamma-b+\kappa(1-\psi)(E_0+S_0+Y(T))/\rho=0.$ This implies

$\alpha > b + \kappa\psi(E_0 + S_0 + Y(T)) / \rho,$ so that after the coal-oil phase there will be a final phase where only coal is used. The expression for Y in (9′) comes from solving $\Omega(S,Y) = 0$ (i.e., using proposition 1 (i)) and the expression for *x* in (9′) comes from $x = \dot{Y} / \psi$ and using $\dot{S} = -q.$ Q.E.D.

We thus see from the first part of (9′) that during the oil-coal phase more coal than oil is used if society is impatient and cares little about global warming (high ρ, low κ), oil extraction costs rise rapidly as reserves are depleted (high δ), and coal generates little more CO2 per unit of energy than oil (low ψ).

Proposition 4: Oil-only phase

Suppose there exists an interval of time $V = [T_1, T_2]$ with $0 \le T_1 \le T_2$ such that $q(t) > 0$ and $x(t) = 0$ along V. Then the time paths of the stock of oil reserves and oil use are given by:

(10) $$S(t) = M_1 e^{\omega_1 t} + M_2 e^{\omega_2 t} - \frac{\alpha - \gamma - \kappa(E_0 + S_0 + Y(T_1)) / \rho}{\delta + \kappa / \rho} \text{ and } q(t) = -\omega_1 M_1 e^{\omega_1 t} - \omega_2 M_2 e^{\omega_2 t},$$

where $\omega_1 = \tfrac{1}{2}\rho + \tfrac{1}{2}\sqrt{\rho^2 + 4(\rho\delta + \kappa) / \beta}$, $\omega_2 = \rho - \omega_1$ and the constants M_1 and M_2 are determined from the boundary conditions given in (11) below.

Proof: Along V $\beta\ddot{S} - \rho\beta\dot{S} - (\rho\delta + \kappa)S = \rho(\alpha - \gamma) - \kappa(E_0 + S_0 + Y(T_1)) / \rho,$ which yields (10). Q.E.D.

2.4. Solving for the boundary conditions and transition times

With the classification of regimes given in table 1 and propositions 2-4, we can fully characterize the various regimes. We will illustrate this for regime III in which the economy starts with using only oil until time T_1, then has a phase from time T_1 to T_2 where both oil and coal are used alongside each other until oil is fully exhausted, and finally a phase from time T_2 onwards where only coal is used. The various constants of integration and the transition times T_1 and T_2 are determined as follows.

We have no accumulated CO2 from coal at the end of the oil phase, so $Y(T_1) = 0.$ From $\Omega(S(T_1), 0) = 0$ we find $S(T_1) = \dfrac{\rho(\gamma - b) - (\psi - 1)\kappa(E_0 + S_0)}{\rho\delta - (\psi - 1)\kappa}$; if this turns out to be negative, we set it equal to zero. At the end of the combined oil-coal phase, oil reserves are fully exhausted, $S(T_2) = 0.$ From $\Omega(0, Y(T_2)) = 0$

we find $Y(T_2) = \dfrac{\rho(\gamma - b)}{(\psi - 1)\kappa} - E_0 - S_0).$ We use the first part of (8′) in proposition 2 to obtain

$x(T_2+) = (\psi\kappa / \beta\lambda_1)\left[Y^* - Y(T_2)\right]$, where $Y^* = \rho(\alpha - b)/\psi\kappa - E_0 - S_0.$ Continuity requires

$q(T_2-) + x(T_2-) = x(T_2+)$ and from (9′) we have $q(T_2-) = \left(\dfrac{\psi(\psi-1)\kappa}{\rho\delta - (\psi-1)\kappa}\right)x(T_2-)$, so we can solve for

$x(T_2-) = \left[\dfrac{\rho\delta - (\psi-1)\kappa}{\rho\delta + (\psi-1)^2\kappa}\right]x(T_2+) < x(T_2+)$ and $q(T_2-) = \left[\dfrac{\psi(\psi-1)\kappa}{\rho\delta + (\psi-1)^2\kappa}\right]x(T_2+) > 0.$ We have from (9)

that $S(T_2) = \dfrac{(\alpha-\gamma)(1-\psi) - b + \gamma}{\psi\delta} + L_1 e^{\pi_1 T_2} + L_2 e^{\pi_2 T_2} = 0,\;\; S(T_1) = \dfrac{(\alpha-\gamma)(1-\psi) - b + \gamma}{\psi\delta} + L_1 e^{\pi_1 T_1} + L_2 e^{\pi_2 T_1}$

and $-\pi_1 L_1 e^{\pi_1 T_2} - \pi_2 L_2 e^{\pi_2 T_2} = q(T_2-)$, where $S(T_1)$ and $q(T_2-)$ (via $x(T_2+)$) are known. Hence, for a given T_1, we can solve these three equations for the constants L_1, L_2 and T_2. We can also solve for

$q(T_1+) = -\pi_1 L_1 e^{\pi_1 T_1} - \pi_2 L_2 e^{\pi_2 T_1}$ and $x(T_1+)$ from $x(T_1+) = \left(\dfrac{\rho\delta - (\psi-1)\kappa}{\psi(\psi-1)\kappa}\right)q(T_1+)$ (using (9′)) and

$q(T_1-) = q(T_1+) + x(T_1+)$ (using continuity). Now consider equation (10) with $Y(T_1) = 0.$ We then obtain the boundary conditions for the oil phase:

(11)
$$S_0 = -\frac{\alpha - \gamma - \kappa(E_0 + S_0)/\rho}{\delta + \kappa/\rho} + M_1 + M_2, \quad S(T_1) = -\frac{\alpha - \gamma - \kappa(E_0 + S_0)/\rho}{\delta + \kappa/\rho} + M_1 e^{\omega_1 T_1} + M_2 e^{\omega_2 T_1},$$
$$q(T_1-) = -\omega_1 M_1 e^{\omega_1 T_1} - \omega_2 M_2 e^{\omega_2 T_1} = q(T_1+) + x(T_1+).$$

Hence, we can solve these three equations for the constants M_1, M_2 and T_1 and all the constants of integration and transition times. Coal use and accumulated CO_2 from past coal use in the final phase, $x(t), Y(t), \forall t \geq T_2$, follow from (8′).

Similarly, the four integration constants and two transition times for regime IV can be found from the boundary conditions $S(T_1) = S_0, \Omega(S_0, Y(T_1)) = 0, S(T_2) = 0, \Omega(0, Y(T_2)) = 0,$ $x(T_1-) = x(T_1+) + q(T_1+),$ and $q(T_2-) + x(T_2-) = x(T_2+) = (\alpha - b)/\beta.$

In general there may be full or partial exhaustion of oil reserves. Full exhaustion occurs if the marginal utility of zero energy consumption is large enough and oil extraction is not too expensive compared to the use of coal, taking due account of the climate damages in addition to the direct user cost.

3. Which carbon tax sustains the socially optimal outcome?

3.1. Comparing the "laissez-faire" and the socially optimal outcome

To focus our attention on some relevant cases, we make the following additional assumption.

Assumption 2: Initially, it is profitable from a social perspective to use oil, $U'(0) > G(S_0) + D'(E_0)/\rho$ and the social cost of extracting coal exceeds that of the first drop of oil, $b + \psi D'(E_0)/\rho > G(S_0) + D'(E_0)/\rho.$ Coal is eventually profitable to use from a utility perspective as well, $U'(0) > G(0) + D'(E_0 + S_0)/\rho.$ Coal has a lower social cost than the final drop of oil, $b + \psi D'(E_0 + S_0)/\rho > G(0) + D'(E_0 + S_0)/\rho.$ Finally, $G(S_0) < b < G(0).$

In the "laissez-faire" market outcome private agents do not internalize global warming externalities ($D(.) = 0$). If the pure extraction cost of oil is higher than that of coal, $G(S_0) > b$, then $G(S) > b$ for all $S < S_0$. In that case, in the "laissez-faire" outcome only coal is ever used. Hence, we assume the more interesting case $G(S_0) < b < G(0)$. This says that initially oil is cheaper for the market to extract than coal, but as less accessible oil fields have to be depleted coal becomes cheaper to extract than oil from a private perspective. Since oil in the Gulf is easier to extract and cheaper to transport than, say, coal from South-Africa, this case seems relevant. Note that the assumption that the social cost of coal is more than the first drop of oil, $b + \psi D'(E_0) > G(S_0) + D'(E_0)$, allows for both $b > G(S_0)$ and $b < G(S_0)$. After all oil is depleted, it is still attractive to use coal, at least for some interval of time, $U'(0) > G(0) + D'(E_0 + S_0)/\rho.$

We assume that the last drop of oil is always more expensive to extract than coal, both from a private and a social perspective. A degenerate form of regime III prevails for the "laissez-faire" outcome as initially the private cost of extracting oil is less than that of coal, $G(S_0) < b$, but the final drop of oil is more expensive than coal, $G(0) > b$. Regime III also prevails for the socially optimal outcome for two reasons. First, initially it is more attractive in view of the global warming damages to burn oil rather than coal, $\Omega(S_0, 0) < 0$. Second, once oil reserves have run out, the social cost of using oil is more than that of coal $\Omega(0, 0) > 0$. Both are ensured by assumption 2. In the oil-only phase the relative advantage of coal must wear off as oil reserves are depleted, $\Omega_S < 0.$ It is thus optimal to start burning oil, then a phase where both coal and oil are used alongside each other, and finally switch to using only coal.

The socially optimal and the "laissez-faire" outcomes will leave oil in situ at the end of the phase where only oil is used. We will show below that the intermediate phase of using coal alongside oil for regime III of the "laissez-faire" outcome will be degenerate, since it is never optimal to use oil and coal simultaneously. Once oil reserves are fully exhausted in the socially optimal outcome, the economy switches to using coal forever.

In the sequel we use the functional specifications introduced earlier, but many qualitative results can be shown to hold more generally.

Proposition 5: Given assumption 2, the "laissez-faire" outcome leads to a regime where first oil is used until instant *T*; from then on, it is no longer profitable to do so and coal is used. The "laissez-faire" outcome never has simultaneous use of oil and coal and oil use falls monotonically during the oil-only phase. Oil use and coal use are given by:

(12a) $$S(t)=\left(\frac{\frac{\alpha-b}{\delta}-\left(S_0+\frac{\alpha-\gamma}{\delta}\right)e^{\omega_2 T}}{e^{\omega_1 T}-e^{\omega_2 T}}\right)e^{\omega_1 t}+\left(\frac{\left(S_0+\frac{\alpha-\gamma}{\delta}\right)e^{\omega_1 T}-\frac{\alpha-b}{\delta}}{e^{\omega_1 T}-e^{\omega_2 T}}\right)e^{\omega_2 t}-\left(\frac{\alpha-\gamma}{\delta}\right),$$

(12b) $$q(t)=-\omega_1\left(\frac{\frac{\alpha-b}{\delta}-\left(S_0+\frac{\alpha-\gamma}{\delta}\right)e^{\omega_2 T}}{\left[e^{\omega_1 T}-e^{\omega_2 T}\right]}\right)e^{\omega_1 t}-\omega_2\left(\frac{\left(S_0+\frac{\alpha-\gamma}{\delta}\right)e^{\omega_1 T}-\frac{\alpha-b}{\delta}}{\left[e^{\omega_1 T}-e^{\omega_2 T}\right]}\right)e^{\omega_2 t},\quad 0\le t\le T,$$

(12c) $$x(t)=0,\quad 0\le t\le T,\qquad S(t)=\frac{\gamma-b}{\delta},\quad q(t)=0,\quad x(t)=\frac{\alpha-b}{\beta},\quad t\ge T,$$

where the transition time *T* is such that $q(T)=(\alpha-b)/\beta$ and increases in *b*.

Proof: If there is simultaneous use of oil and coal, then $\Omega(S,0)\equiv\rho\left[G(S)-b\right]=0$ during this phase. This implies a constant value of *S* and thus $q = 0$, which contradicts the assumption of simultaneous use. The optimal regime thus corresponds to a degenerate version of regime III. To establish that *q* falls during the oil-only phase, note that for the market μ plays no role so that $U'(q)=G(S)+\lambda,$ which implies $U''(q)\dot{q}=G'(S)\dot{S}+\rho\lambda+G'(S)q=\rho\lambda$ and thus (as $\lambda>0$) we have $\dot{q}<0.$ From proposition 4 the initial oil phase is described by $S(t)=M_1e^{\omega_1 t}+M_2e^{\omega_2 t}-\frac{\alpha-\gamma}{\delta}$ and $q(t)=-\omega_1M_1e^{\omega_1 t}-\omega_2M_2e^{\omega_2 t}$, where $\omega_1=\frac{1}{2}\rho+\frac{1}{2}\sqrt{\rho^2+4\rho\delta/\beta}$, $\omega_2=\rho-\omega_1$. From proposition 2 the final coal phase implies that coal use is $x(t) = (\alpha - b)/\beta$, $t \ge T$ as $\lambda_1 = \rho$ and $\lambda_2 = 0$ if $\kappa = 0$ and $Y(T) = 0$. Solving for M_1 and M_2 and the transition time *T* for the case $\gamma > b$ from $S(0) = S_0$ and $S(T)=G^{-1}(b)=(\gamma-b)/\delta,$ we get

$M_1 = \dfrac{\alpha - b - (\delta S_0 + \alpha - \gamma)e^{\omega_2 T}}{\delta\left[e^{\omega_1 T} - e^{\omega_2 T}\right]}$ and $M_2 = \dfrac{(\delta S_0 + \alpha - \gamma)e^{\omega_1 T} - (\alpha - b)}{\delta\left[e^{\omega_1 T} - e^{\omega_2 T}\right]}$, and thus the expressions for $S(t)$ and $q(t)$ given in (12). We get the transition time by requiring continuity:

$q(T) = \dfrac{(\omega_1 - \omega_2)(\delta S_0 + \alpha - \gamma) - (\gamma - b)\left(\omega_1 e^{-\omega_2 T} - \omega_2 e^{-\omega_1 T}\right)}{\delta\left(e^{-\omega_2 T} - e^{-\omega_1 T}\right)} = \dfrac{\alpha - b}{\beta} = x(T).$ The denominator increases in T and, as $\gamma > b$, the numerator decreases in T, hence $q'(T) < 0$. Since ω_1 and ω_2 do not depend on α or b, we see that T increases with b. Q.E.D.

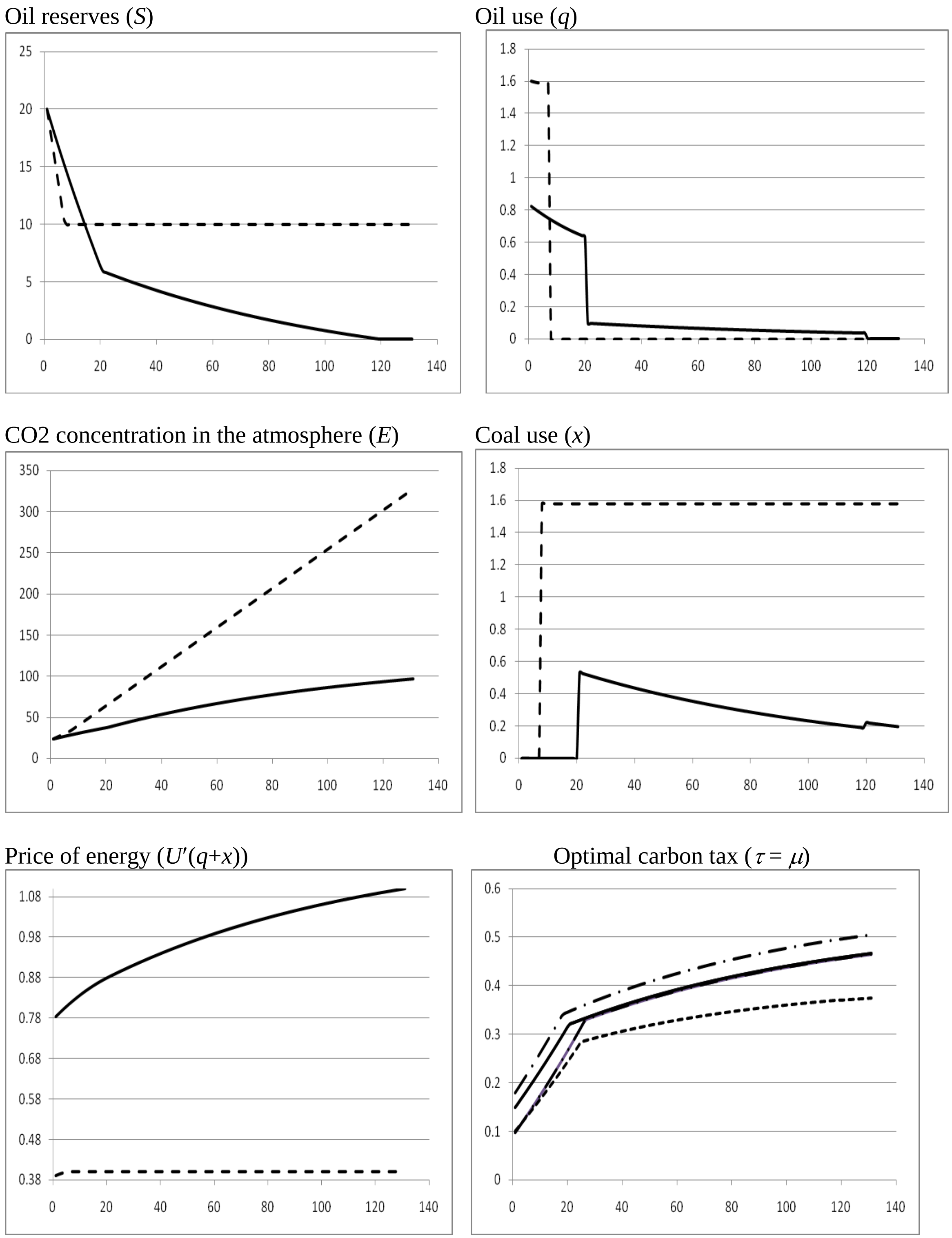
Once oil is phased out, the "laissez-faire" decentralized market outcome leaves a positive amount of oil in situ and more so if coal is cheap (low b) and the extraction cost of oil is high and does not decrease rapidly as oil reserves are depleted (high γ, low δ). The market wants to make the transition to coal too soon, and too abrupt. In contrast, the socially optimal outcome fully exhausts oil reserves before having to rely on coal as the sole source of energy. The reason is, of course, that coal emits more CO2 per unit of energy and oil is relatively clean so it pays from the point of view of combating climate change to use up all oil. Also, note that coal use is high and remains high in the final coal phase of the "laissez-faire" outcome but coal use is lower and vanishes asymptotically in the social optimum.

The above theoretical insights are confirmed by the simulations presented in fig. 3.[8] The "laissez-faire" economy phases out relatively clean oil and switches to coal much earlier, at instant $T = 6.3$, than the time coal is phased in and oil is phased out in the optimal economy. The optimal economy has depleted oil reserves from 20 at instant zero to 5.92 at instant $T_1 = 19.7$ before coal is phased in alongside oil, and then moves to using coal only once oil reserves are fully exhausted at the much later instant $T_2 = 118.4$. Furthermore, the "laissez-faire" economy uses much more substantial and non-vanishing amounts of dirty coal than the social optimum. Since coal contributes more to global warming than oil per unit of energy, both effects exacerbate global warming and curb social welfare. This is also reflected in fig. 3, which shows a much steeper rise in the concentration of CO2 in the atmosphere resulting from a much lower price of energy in the "laissez-faire" outcome than in the socially optimal outcome. Note that with our

---

[8] The following parameter values have been used: $S_0 = 20, E_0 = 24, b = 0.4, \gamma = 0.6, \delta = 0.02, \psi = 1.5, \rho = 0.014,$ $\kappa = 0.00006, \alpha = 1.2, \beta = 0.506.$ These values satisfy assumptions 1 and 2. γ and δ have been calibrated so that initial oil extraction costs are half those of coal, $G(S_0) = b/2$, and in the "laissez-faire" outcome half of oil reserves is left in situ, $(\gamma - b)/\delta = S_0/2$. The parameters $\alpha = 1.2$ and $\beta = 0.51$ are obtained by taking a second-order Taylor series approximation around $x = (\alpha - b)/\beta$ of the utility function $U(x) = \dfrac{x^{1-1/\sigma}}{1-1/\sigma}$, where the intertemporal elasticity of substitution has been set to $\sigma = 0.5$ and thus the elasticity of intertemporal inequality aversion to $1/\sigma = 2$.

Figure 3: Simulations of "laissez-faire" and optimum economy



Key: Solid lines = social optimum; dashed lines = "laissez-faire" outcome. Long dashed-dotted-dotted, short dashed and long dashed-dotted lines: optimal carbon tax with $\Delta S_0 = 5$, $\Delta b = -0.1$ & $\Delta(1/\sigma) = -0.5$.

chosen parameter values $\Omega_S(S,Y)=(\psi-1)\kappa-\rho\delta=-0.00025<0$ for the optimum economy and $\Omega_S(S,Y)=-\rho\delta=-0.00028<0$ for the "laissez-faire" economy, so that the effect of oil extraction costs rising dominates the effect of marginal global warming damages rising and therefore the advantage of using oil rather than coal falls in both cases as oil reserves are depleted. There are discrete jumps in both oil use and coal use. When in the optimal economy coal is phased in at instant T_1 = 19.7, oil use jumps down by an amount 0.53 and coal use jumps up by the same amount so that there is no jump in the price of energy. Also, oil use jumps from 0.034 to zero at instant T_2 = 118.4 as coal jumps up by the same amount, again to ensure that there is no jump in the price of energy.

In the "laissez-faire" outcome the rent on oil, i.e., the difference between the price of oil and the extraction cost of oil, vanishes at the time the economy switches to the backstop as at that point the extraction cost of oil must equal that of coal (cf., Heal, 1976). In the socially optimal outcome all oil reserves are depleted. Oil still has a positive shadow price as it has lower CO2 emissions per unit of energy than coal.

3.2. The optimal carbon tax

The "laissez-faire" outcome relies in the long run too much on coal instead of oil, and thus leads to too much CO2 emissions compared to the socially optimal outcome. In the short run it uses up oil quickly, although it leaves more oil in situ. The key challenge for policy makers is to design an optimal carbon tax in such a way that the market outcome replicates the socially optimal outcome. Such a tax should persuade private agents to use up all oil reserves, even the more expensive fields, rather than using highly CO2-intensive coal, and to give proper incentives for private agents to use no coal at all in the long run. Such an optimal carbon tax inevitably rises over time and faces the difficult task of encouraging private agents to use both oil and coal alongside each other for a period of time.

Proposition 6: The optimal carbon tax rises during all three phases. In the final phase where coal is used indefinitely, the socially optimal outcome is attained with the following carbon tax:

(13) $$\tau(t)=\frac{\alpha-\beta x(t)-b}{\psi} \text{ and } x(t)=\left(\frac{\psi\kappa}{\beta\lambda_1}\right)\left[\left(\frac{\rho}{\psi\kappa}\right)(\alpha-b)-\left(E_0+S_0\right)\right]e^{\lambda_2(t-T_2)}, \quad t\geq T_2.$$

In the phase where oil and coal are used alongside each other, $q(t)$ follows from (9) and the optimal outcome is attained with the following carbon tax:

(13′) $$\tau(t) = \frac{\alpha - \beta\left[q(t) + x(t)\right] - b}{\psi}, \quad q(t) + x(t) = \left(\frac{\rho\delta + (\psi - 1)^2\kappa}{\psi(\psi - 1)\kappa}\right) q(t), \quad T_1 \le t \le T_2.$$

For the initial only-oil phase the optimal carbon tax is given by: [9]

(13″) $$\tau(t) = e^{-\rho(T_1 - t)}\left[\frac{U'\left(q(T_1) + x(T_1)\right) - b}{\psi}\right] - \int_t^{T_1} D'(E_0 + S_0 - S(s))e^{-\rho(s-t)}ds, \quad \forall 0 \le t \le T_1.$$

Proof: It follows from (5d) that $\ddot{\mu} = \rho\dot{\mu} - D''(E)\dot{E}$. CO2 emissions are always non-decreasing and the damage function is convex. Hence, once the social cost of carbon and thus the optimal carbon tax $\tau = \mu$ starts decreasing, it will decrease forever and tends towards minus infinity which violates the optimality conditions. If only coal is used, the optimal carbon tax follows from $U'(x) = b + \psi\tau$, where $x(t)$ as given in (13) follows from equation (8′) with $S(T_1) = 0$. With simultaneous use of oil and coal, the optimal $\tau(t)$ as given in (13′) follows from $U'(x + q) = b + \psi\tau$ with $x(t) + q(t)$ given by (9) and (9′). For the initial oil phase, note that for the optimal economy $U'\left(q(T_1) + x(T_1)\right) = b + \psi\mu(T_1)$ from (5b) gives $\tau(T_1) = \mu(T_1) = \left[U'\left(q(T_1) + x(T_1)\right) - b\right]/\psi.$ Integrating equation $\dot{\tau} = \rho\tau - D'(E_0 + S_0 - S)$ (from (5d)) backwards in time using this condition yields $\tau(t) = e^{-\rho(T_1 - t)}\tau(T_1) - \int_t^{T_1} D'(E_0 + S_0 - S(s))e^{-\rho(s-t)}ds,$ $0 \le t \le T_1,$ and thus the time path for $\tau = \mu$ given in (13″) for the oil phase. Q.E.D.

Hence, in the intermediate and final phases where either coal and oil are used alongside each other or only coal is used, the carbon tax increases over time as the marginal global warming damages increase over time (due to the absence of natural decay of the CO2 concentration in the atmosphere). In the initial phase where only oil is used, the use of oil as given by equations (10) and (11) of proposition 4 must diminish over time, but that is already encouraged by the rising oil extraction costs and the higher royalties. Hence, the positive carbon tax can be less high. However, fig. 3 indicates that the optimal carbon tax rises much more steeply during the oil-only phase than during the oil-coal and coal phases. This reflects that the carbon tax has to be designed in such a way as to ensure that the market is giving the correct incentive to switch from using coal to oil in the initial phase, and to ensure that the transition times for phasing in coal and phasing out oil are socially optimal. Also, the optimal carbon tax has to ensure that the market fully

[9] For our specific functional forms (12″) we have $\tau(T_1) = \left[\alpha - \beta\left(x(T_1+) + q(T_1+) - b\right)\right]/\psi$ and $\tau(t) = e^{-\rho(T_1 - t)}\tau(T_1)$ $-\kappa\left(\frac{(\alpha - \gamma) + \delta(E_0 + S_0)}{\rho\delta + \kappa}\right)\left(1 - e^{-\rho(T_1 - t)}\right) - \left(\frac{\kappa M_1}{\rho - \omega_1}\right)\left(e^{\omega_1 t} - e^{\rho(t - T_1) + \omega_1 T_1}\right) - \left(\frac{\kappa M_2}{\rho - \omega_2}\right)\left(e^{\omega_2 t} - e^{\rho(t - T_1) + \omega_2 T_1}\right).$

exhausts all oil reserves which requires that the market expects the net price of oil to rise at a rate smaller than the interest rate.

3.3. Sensitivity analysis

Although the qualitative nature of the simulation trajectories is unaffected by small changes in parameter values, table 2 indicates that the optimal transition times of phasing in coal and phasing out oil in the "laissez-faire" and optimal economies and the size of the optimal carbon tax do change. It also gives the long-run values of the optimal carbon tax, social price of energy and atmospheric CO2 concentration. The last panel of fig. 3 shows the sensitivity of the optimal carbon tax to key parameter values.

A discovery of new oil reserves today ($\Delta S_0 = 5$) does not affect the long-run social price of energy or the long-run optimal carbon tax. It does lead to a later phasing in of coal and a somewhat later phasing out of oil. The last panel of fig. 3 indicates that the optimal carbon tax still rises steeply during the oil-only phase, but will be lower during this initial phase as oil is less scarce and thus cheaper to extract from the earth. Consequently, fewer incentives are needed to persuade the market to use oil rather than coal in this initial phase. Note that the "laissez-faire" outcome leaves more oil in situ and phases coal in later.

Table 2: Optimal and "laissez-faire" transition times and the stock of oil left in situ by the market

	Benchmark	$\Delta S_0 = 5$	$\Delta b = -0.1$	$\Delta(1/\sigma) = -0.5$	$\rho = 0.025$	$\kappa = 0.00003$
Optimal economy:						
Long-run carbon tax	0.53	0.53	0.4	0.57	0.53	0.53
Long-run price of energy	1.2	1.2	0.9	1.26	1.2	1.2
Long-run CO2 concentration	124.4	124.4	93.3	133.4	222.2	248.9
Time of phasing in coal, T_1	19.7	25.6	24.1	17.3	11.4	12.2
Time of phasing out oil, T_2	118.4	120.6	85.0	97.9	164.9	205.4
"Laissez-faire" economy:						
Oil left in situ forever	10	10	10	10	10	10
Time of phasing in coal, T	6.3	7.9	5.5	5.6	6.3	6.3

A lower cost of coal ($\Delta b = -0.1$) brings forward the date when the economy starts to use coal in the "laissez-faire" economy. Cheaper coal lowers both the long-run social price of energy and the long-run value of the carbon tax. In fact, the last panel of fig. 3 shows that the whole time trajectory of the optimal carbon tax shifts down when coal is cheaper. Interestingly, cheaper coal leads to a later phasing in of coal and a much earlier phasing out of oil in the socially optimal outcome. The oil-coal phase lasts much shorter. The reason for these intertemporal shifts in the optimal outcomes is that it is better to use oil in the beginning whilst it is still relatively cheap to extract before having to inevitably shift to using the relative dirty, but cheaper coal. Neither the discovery of new oil reserves nor the lower cost of coal, leads to a bigger long-run build-up of CO_2 in the atmosphere.

A higher elasticity of intertemporal substitution ($\sigma = 2/3$) implies that the market can more easily substitute current consumption for future consumption; hence, coal is phased in earlier. It corresponds to a lower elasticity of intertemporal inequality aversion ($\Delta(1/\sigma) = -0.5$), so that the socially optimal outcome phases in coal and phases oil out more quickly. The long-run price of energy and the long-run carbon tax are somewhat higher. In fact, the last panel of fig. 3 indicates that the whole time path of the optimal carbon tax shifts up if society has a lower intertemporal inequality aversion. Still, the long-run concentration of CO_2 in the atmosphere is a bit higher. If society cares less about the impact of global warming on future generations, it phases in coal more quickly and ends up with a bigger climate problem.

If the government follows the market and does not apply for precautionary or other reasons a lower rate of discount rate, but a discount rate of say 0.025, it phases in coal more quickly and phases out oil much later. The optimal carbon tax is then initially much lower and takes much longer to reach its long-run value. Hence, there are more CO_2 emissions in the short run which exacerbates global warming. Finally, if the government has a too optimistic view on the costs of global warming ($\kappa = 0.00003$ instead of 0.00006), it phases in coal to early and phases out oil too late. As with the too impatient government, this leads a much bigger accumulation of CO_2 in the atmosphere.

4. Carbon-free renewables and the Green Paradox

4.1. Optimal carbon tax and transition times with carbon-free renewables

The optimal climate change policy is to set a carbon tax that reflects the marginal environmental damage from using the respective types of fossil fuels. Here we address the important issue of how our results are affected when there are carbon-free renewables as a potential source of energy supply as well. Consider therefore the presence of an infinitely elastic supply of such renewables at a cost *c* bigger than that of coal *b*. A necessary condition for renewables to be worthwhile introducing at some point is that $U'(0) = \alpha > c$. For *c* relatively low, i.e., close to *b*, we might not find a coal-only phase in the optimum with renewables, because renewables might take over before the only coal phase starts in view of the already accumulated CO_2 stock. However, we are interested in the case that the cost of renewables *c* is not relatively low and that renewables only take over after the oil-coal phase. Indeed, to have a solution with more accumulated CO_2 from coal at the end than at the beginning of the coal-only phase, we need that renewables are relatively expensive, $c > b + \psi(\gamma - b) = 0.775$ (see proposition 7). In fact, our simulations indicate that this lower bound on *c* is not high enough to ensure that renewables are only introduced after the coal-only phase has started. This requires that coal must be even more expensive. For our chosen parameter values this requires *c* > 1.0. For the core parameter values in our simulations (see footnote 8 in section 3), we thus require values of *c* in the range (1.0, 1.2).

Since renewables are more expensive than coal, they will not be phased in under the "laissez-faire" outcome unless there is a subsidy on them or their cost falls with time due to technical progress. However, the social planner which takes account of the rising costs of global warming does phase in carbon-free renewables eventually. Over time the use of coal rapidly increases the marginal damages of global warming and thus the social cost of coal. As soon, as the social cost of coal hits the cost of renewables, the economy stops using coal and switches to renewables. Granted that the introduction of renewables will affect both transition times T_1 and T_2, we use (13) in proposition 6 to see that renewables take over at the instant $T_3 > T_2$ where the cost of coal plus the carbon tax equals the cost of renewables:

(14) $$b + \psi\tau(T_3) = \alpha - \beta x(T_3) = c, \text{ where } \tau(T_3) = D'\left(E_0 + S_0 + Y(T_3)\right)/\rho.$$

The social cost of carbon at the end of the coal-only phase thus equals the present value of marginal global warming damages, which are constant as there is no CO2 pollution anymore once carbon-free renewables have been introduced. The following proposition characterizes the transition to the carbon-free economy.

Proposition 7: Suppose the optimal sequence is only oil until T_1, oil and coal from T_1 to T_2, only coal from T_2 to T_3, and renewables from T_3 onwards. Then, from the moment that oil is phased out, the accumulated stock of CO2 due to past burning of coal grows from $Y(T_2) = \frac{\rho}{\kappa}\left(\frac{\gamma - b}{\psi - 1}\right) - E_0 - S_0$ until the moment that renewables is introduced and then stays constant at the level $Y(T_3) = \frac{\rho}{\kappa}\left(\frac{c - b}{\psi}\right) - E_0 - S_0$. The long-run optimal carbon tax rate, renewables use, and stock of CO2 in the atmosphere from that moment on are given by:

(14′) $$\tau(T_3) = (c - b)/\psi, \quad x(T_3) = \frac{\alpha - c}{\beta}, \quad E(T_3) = E_0 + S_0 + Y(T_3) = \frac{\rho(c - b)}{\kappa\psi}.$$

During the oil-only and the oil-coal phase, the optimal carbon tax must rise.

Proof: The expression for $Y(T_2)$ follows from the requirement $\Omega(0, Y(T_2))=0$ as in section 2.4. The expressions for $x(T_3)$, $Y(T_3)$, $E(T_3)$ and $\tau(T_3)$ follow directly from equation (14). From T_3 onwards, the stock of CO2 remains constant. Hence, in order to satisfy the transversality constraint (5e) it is necessary that the carbon tax is constant from that moment on. If it would decrease before T_3, it would decrease always thereafter and the transversality condition would be violated. Q.E.D.

Hence, a high cost of renewables and a low extraction cost of coal necessitate a high long-run value of the optimal carbon tax. In that case, the final stock of accumulated stock of CO2 in the atmosphere resulting from burning coal will be higher as well, especially if the society is very impatient (high ρ) and the perceived cost of global warming is small (low κ, E_0 and S_0).

Since the social cost of carbon will be lower as a result of the availability of a carbon-free backstop, the optimal carbon tax, the final stock of CO2 in the atmosphere and thus the transition times to the oil-only and oil-coal phases will be affected.

To see how this works, note that as in section 2.3 $S(T_1) = \text{Max}\left[\dfrac{\rho(\gamma - b) - (\psi - 1)\kappa(E_0 + S_0)}{\rho\delta - (\psi - 1)\kappa}, 0\right]$,

$S(T_2) = 0$, and $Y(T_2) = \dfrac{\rho(\gamma - b)}{(\psi - 1)\kappa} - E_0 - S_0$. Now we use, respectively, the first and second part of (8) in proposition 2 to obtain the two boundary conditions $x(T_3-) = K_1 e^{\lambda_1 T_3} + K_2 e^{\lambda_2 T_3} = (\alpha - c)/\beta$ (from (14)) and $\dfrac{\rho(\gamma - b)}{(\psi - 1)\kappa} = \dfrac{\rho(\alpha - b)}{\psi\kappa} + \left(\dfrac{\psi K_1}{\lambda_1}\right) e^{\lambda_1 T_2} + \left(\dfrac{\psi K_2}{\lambda_2}\right) e^{\lambda_2 T_2}$. These two conditions can be solved for K_1 and K_2, and thus for $x(T_2+) = K_1 e^{\lambda_1 T_2} + K_2 e^{\lambda_2 T_2}$ given T_2 and T_3. As before, the continuity requirement and (9′) then give $x(T_2-) = \left[\dfrac{\rho\delta - (\psi - 1)\kappa}{\rho\delta + (\psi - 1)^2\kappa}\right] x(T_2+)$ and $q(T_2-) = \left[\dfrac{\psi(\psi - 1)\kappa}{\rho\delta + (\psi - 1)^2\kappa}\right] x(T_2+)$. We have from (9) that $S(T_2) = \dfrac{(\alpha - \gamma)(1 - \psi) - b + \gamma}{\psi\delta} + L_1 e^{\pi_1 T_2} + L_2 e^{\pi_2 T_2} = 0$, $S(T_1) = \dfrac{(\alpha - \gamma)(1 - \psi) - b + \gamma}{\psi\delta} + L_1 e^{\pi_1 T_1} + L_2 e^{\pi_2 T_1}$ and $-\pi_1 L_1 e^{\pi_1 T_2} - \pi_2 L_2 e^{\pi_2 T_2} = q(T_2-)$, which for given T_1 we can solve for L_1, L_2 and T_2. Again, we solve for $q(T_1+) = -\pi_1 L_1 e^{\pi_1 T_1} - \pi_2 L_2 e^{\pi_2 T_1}$, $x(T_1+) = \left(\dfrac{\rho\delta - (\psi - 1)\kappa}{\psi(\psi - 1)\kappa}\right) q(T_1+)$ and $q(T_1-) = q(T_1+) + x(T_1+)$. We can solve the three boundary conditions for the oil phase (11) and the terminal condition for the switch from the coal to the renewables phase, $\hat{K}_1 e^{\lambda_1 T_3} + \hat{K}_2 e^{\lambda_2 T_3} = \rho(c-\alpha)/\psi\kappa$ (from (14′) after substituting the second part of (8) for $Y(T_3)$), to solve for the constants M_1 and M_2 and the transition times T_1 and T_3. Equation (8) now gives coal use and accumulated CO2 from past coal use in the coal-only phase, i.e.,

$$x(t) = K_1 e^{\lambda_1 t} + K_2 e^{\lambda_2 t} \text{ and } Y(t) = \frac{\rho}{\psi\kappa}(\alpha - b) - E_0 - S_0 + \hat{K}_1 e^{\lambda_1 t} + \hat{K}_2 e^{\lambda_2 t}, T_2 \le t \le T_3.$$

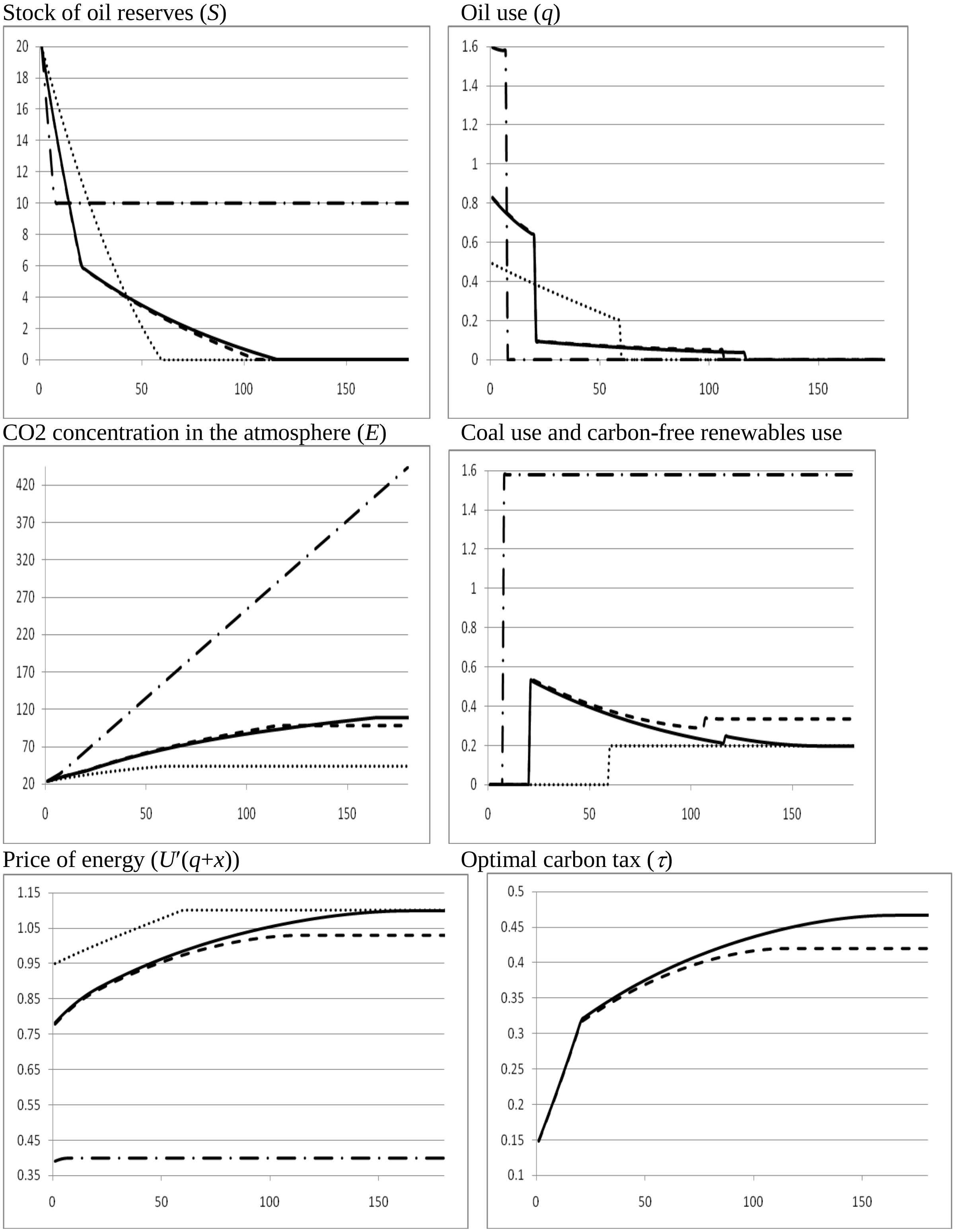
4.2. Simulation of the optimal carbon tax with renewables

Using the above, we have calculated the optimal regime for when there is a relatively expensive carbon-free renewable backstop available (at cost c = 1.1 > b = 0.4). The solid and dashed-dotted lines in fig. 4 present the resulting optimal and "laissez-faire" simulation trajectories and fig. 3 allows them to be compared with the outcomes when renewables are not available. The eventual introduction of carbon-free renewables lowers the social cost of carbon and hardly brings forward the phasing in of coal from instant T_1 =19.7 if there are no renewables to instant 19.6. The anticipation of the future introduction of renewables and the resulting reduction in the social cost of carbon also brings forward the date that oil is phased out and the economy relies on coal only slightly from instant T_2 = 118.4 to instant of time 115.2.

The length of the oil-coal phase is reduced somewhat from 98.7 to 95.6. Carbon-free renewables become profitable from instant T_3 = 163.3 onwards, so that the coal-only phase has a length of only 48.1 instead of infinity. Since a little more oil is used and less coal, the stock of accumulated CO2 is only 108.9 at the moment coal is phased out and stays at that level from thereafter. The optimal carbon tax rises from 0.14 to 0.47 at the time of the switch to renewables; hence, the presence of renewables means that the carbon tax can be lower, rising to 0.47 at the end of the coal-only phase and staying at that level thereafter. The optimal carbon tax and the social price of energy still rises steeply during the oil-only phase (from time zero to 19.6), rises less steeply during the oil-coal phase (from time 19.6 to 115.2) and the coal-only phase (from time 115.2 to 163.3). The final value of the optimal carbon tax has to be maintained at a constant level during the renewables-only phase; else, the economy has an incentive to switch to using coal again. Given the high cost of renewables, no drastic changes are observed during the oil-only and oil-coal phases compared to the situation when there are no renewables, but the coal-only phase comes to an end and thus the rise in the carbon tax comes to an end as well.

Now consider the implications of a lower cost of renewables, c = 1.03 instead of 1.1, which are indicated by the dashed lines in fig. 4. Coal is phased in marginally earlier at instant T_1 = 19.4 rather than 19.6, but oil is phased out rather earlier at instant T_2 = 105.0 instead of 115.2 and renewables are phased in much

Figure 4: Simulations with eventual introduction of renewables



Key: solid lines = socially optimum (c = 1.1); dashed-dotted lines = "laissez-faire" outcome; dashed lines = social optimum with subsidy (c = 1.03); dotted lines = prohibitive coal tax.

earlier at instant T_3 = 114.2 rather than 163.3. The lower cost of renewables thus leads initially to a little more and somewhat earlier pumping of oil, but much more importantly to coal being phased in and phased out more quickly. Interestingly, coal use is at a higher level but used for a much shorter period (9.2 instead of 48.1 periods). As a result and of renewables being phased in more quickly, the concentration of CO2 in the atmosphere is not reduced as much as it could have done, but still it falls from 108.9 to 98.0 (about 10 percent). If allowance is made for the optimal phasing in of coal on its own or alongside oil and the earlier phasing in of renewables, we see that the lower cost of renewables implies that it is optimal to have a lower carbon tax in the oil-coal and coal phases. The final level of the carbon tax is also lower (i.e., 0.42 instead of 0.47).

If the cost of renewables becomes low enough (c about 1.0), the coal-only phase vanishes completely. If the cost of renewables becomes lower than that, renewables will be introduced alongside oil and possibly coal. We leave this for further research.

4.3. Second-best policies: taxing coal or subsidizing carbon-free renewables?

A full-fledged carbon tax is politically infeasible in many countries, especially on CO2 emissions resulting from burning coal. In that case, governments often resort to subsidizing carbon-free renewables such as solar or wind energy. The idea behind such a second-best policy is that this will bring forward the date at which coal is phased out, so that this lowers CO2 emissions and curbs global warming. In earlier work on just coal and renewables, thus ignoring the impact of dirty backstops such as coal, concern has been voiced that such a subsidy may be counterproductive, since it merely encourages the market to pump up oil more quickly and thus exacerbate global warming (Sinn, 2008). However, if it is optimal to leave oil in situ and there is no dirty backstop, the subsidy encourages the market to leave more oil in situ and switch more quickly to renewables in which case global warming is curbed and there is no green paradox (van der Ploeg and Withagen, 2010a). To investigate what our context with a dirty backstop implies for the green paradox, we analyze the two second-best policies of either subsidizing renewable or taxing coal *without the introduction of an appropriate carbon tax*.

Subsidizing renewables so that they are marginally cheaper than the market price of coal leads to exactly the same oil-only phase, transition time and amount of oil left in situ as in the “laissez-faire” outcome, albeit once the backstop is introduced the stock of CO2 is stabilized.

Proposition 8: With renewables being subsidized to just below the extraction cost of coal, green welfare gains dominate the welfare losses resulting from the lump-sum taxes needed to finance the subsidy if

$c < b + \kappa\psi\left[E_0 + S_0 - (\gamma - b)/\delta + \psi(\alpha - b)/\rho\beta\right]$ holds. The subsidy then yields higher welfare than the "laissez-faire" outcome.

Proof: Since in the oil-only phase T and $S(T)$ are the same, we only need to consider the backstop phase. In the "laissez-faire" outcome, the green welfare loss for that phase is:

$$\Theta^{LF} \equiv \int_T^\infty D(E(t))e^{-\rho(t-T)}dt = \frac{\kappa}{2}\int_T^\infty \left[E_0 + S_0 - \frac{\gamma - b}{\delta} + \psi\left(\frac{\alpha - b}{\beta}\right)(t - T)\right]^2 e^{-\rho(t-T)}dt =$$

$$\frac{\kappa}{2\rho}\left[E_0 + S_0 - \frac{\gamma - b}{\delta}\right]^2 + \frac{\kappa}{\rho^2}\left[E_0 + S_0 - \frac{\gamma - b}{\delta}\right]\psi\left(\frac{\alpha - b}{\beta}\right) + \frac{\kappa}{\rho^3}\psi^2\left(\frac{\alpha - b}{\beta}\right)^2.$$

If renewables are subsidized until they are just economically viable, the green welfare loss is:

$$\Theta^{RS} \equiv \int_T^\infty D(E(t))e^{-\rho(t-T)}dt = \frac{\kappa}{2}\int_T^\infty \left[E_0 + S_0 - \frac{\gamma - b}{\delta}\right]^2 e^{-\rho(t-T)}dt = \frac{\kappa}{2\rho}\left[E_0 + S_0 - \frac{\gamma - b}{\delta}\right]^2 < \Theta^{LF}.$$

If the reduction in green welfare loss thus obtained outweighs the welfare cost of the lump-sum taxes needed to finance the subsidy, i.e., if

$$\Theta^{LF} - \Theta^{RS} = \frac{\kappa}{\rho^2}\left[E_0 + S_0 - \frac{\gamma - b}{\delta}\right]\psi\left(\frac{\alpha - b}{\beta}\right) + \frac{\kappa}{\rho^3}\psi^2\left(\frac{\alpha - b}{\beta}\right)^2 > \int_T^\infty (c - b)x(t)e^{-\rho(t-T)}dt = \left(\frac{c - b}{\rho}\right)\left(\frac{\alpha - b}{\beta}\right),$$

it makes sense to subsidize renewables in this way. This gives rise to the required condition. Q.E.D.

Hence, making renewables economically viable improves social welfare if renewables are cheap and coal is expensive (c low, b high), the climate challenge is acute (high E_0, S_0, ψ), society employs a prudently low discount rate (low ρ), the initial extraction cost of oil is low (low γ) and demand for energy is substantial (high α), although it might do somewhat less well if the subsidy has to be financed by distorting taxes). For our core parameters, this condition is satisfied because$c$ = 1.1 < 1.55. This is why in the simulations the subsidy on renewables yields higher (in fact, much higher) welfare than the "laissez-faire" outcome (-27.7 instead of -136.7).

The alternative second-best policy in the absence of a carbon tax is a prohibitive tax on coal, which makes coal marginally more expensive than renewable. This also has the effect that coal will never be introduced, but now renewables will be introduced at a later date. This policy is equivalent to a moratorium on coal-powered electricity and other uses of coal.

Proposition 9: With a prohibitive coal tax and the cost of renewable more than the cost of extracting the final drop of oil ($c > \gamma$), oil reserves will be fully exhausted. During the oil-only phase, oil reserves and oil use fall monotonically while coal use is zero. The paths for the fossil-fuel phase are given by:

$$(12a') \quad S(t) = \frac{\left[\frac{\alpha-\gamma}{\delta} - \left(S_0 + \frac{\alpha-\gamma}{\delta}\right)e^{\omega_2 T}\right]}{e^{\omega_1 T} - e^{\omega_2 T}} e^{\omega_1 t} + \frac{\left[\left(S_0 + \frac{\alpha-\gamma}{\delta}\right)e^{\omega_1 T} - \frac{\alpha-\gamma}{\delta}\right]}{e^{\omega_1 T} - e^{\omega_2 T}} e^{\omega_2 t} - \frac{\alpha-\gamma}{\delta},$$

$$(12b') \quad q(t) = -\omega_1 \frac{\left[\frac{\alpha-\gamma}{\delta} - \left(S_0 + \frac{\alpha-\gamma}{\delta}\right)e^{\omega_2 T} - S_0 e^{\omega_2 T}\right]}{e^{\omega_1 T} - e^{\omega_2 T}} e^{\omega_1 t} - \omega_2 \frac{\left[\left(S_0 + \frac{\alpha-\gamma}{\delta}\right)e^{\omega_1 T} - \frac{\alpha-\gamma}{\delta}\right]}{e^{\omega_1 T} - e^{\omega_2 T}} e^{\omega_2 t}, \quad 0 \le t \le T.$$

Oil reserves, oil use and renewables use in the carbon-free are given by:

$$(12c') \quad S(t) = 0, \quad q(t) = 0, \quad x(t) = \frac{\alpha - c}{\beta}, \quad t \ge T,$$

The transition time T follows from $q(T) = (\alpha - c)/\beta$, and is earlier if c is low.

Proof: The proof parallels that of proposition 5. Since the condition $G(S(T)) = c$ yields a negative value of $S(T)$ with $c > \gamma$, we must have $S(T) = 0$. We obtain (12a′) and (12b′) from (10) after solving for M_1 and M_2. Also, from $q(T) = \frac{(\omega_1 - \omega_2)\left(S_0 + \frac{\alpha-\gamma}{\delta}\right)e^{\rho T} - \left(\omega_1 e^{\omega_1 T} - \omega_2 e^{\rho T}\right)\frac{\alpha-\gamma}{\delta}}{e^{\omega_1 T} - e^{\omega_2 T}}$ we have $q'(T) > 0$, so T increases with . Q.E.D.

We thus establish that the transition to renewables occurs faster if renewables are cheap.

The dotted lines in fig. 4 show the effects of taxing coal so that it is just below the cost of renewables, hence the cost of coal has been raised from 0.4 to 1.1. The result of this second-best policy is that the market substantially delays the phasing out of oil from time 6.9 to time 58.7, albeit still quicker than the first-best outcome. So this second-best policy exhausts all oil reserves to a bigger extent (and in our simulation exhausts them fully) and uses oil for a much longer period. Afterwards, renewables rather than coal are phased in. The result of the tax on coal is to push up the price of energy very quickly to somewhere close to the cost of renewables (even higher than in first-best outcome). It is this that forces the market to be much more careful and persistent in the use of oil. The prohibitive tax on coal is extremely successful in terms of fighting global warming, since it reduces the long-run stock of carbon in the atmosphere from 108.9 in the first-best outcome to 44.0. However, social welfare is 7.8 in the second-best outcome compared with -136.7 in the "laissez-faire" outcome in which CO2 damages from coal go on forever, so a prohibitive coal tax is much better than doing nothing. Of course, the first-best policy yields higher social welfare (12.4) than the prohibitive coal tax because the latter damages the private component of social welfare even more than it curbs global warming damages. Interestingly, the first-best outcome does rely on coal both alongside oil and on its own to ensure that the private part of welfare does not fall too much. Finally, we see that in our simulations the subsidy on renewables does considerably worse than the coal tax (as welfare is -27.7 instead of 7.8).

5. Conclusions

We have investigated the optimal climate policy in the presence of an exhaustible resource, oil, and an abundant resource, coal (or tar sands), where coal contributes more CO2 per unit of energy than oil. We have characterized the various ways of sequencing oil and coal and departures from the Herfindahl rule that can occur in the social optimum. Indeed, in the social optimum the ordering does not just depend on extraction costs (which rise as oil reserves are depleted), but also on the social cost of global warming damages. With high enough demand for energy and with very dirty coal, coal will only be used after complete exhaustion of oil, even if coal is cheap to extract. For moderate CO2-emission coefficients of coal, an intermediate period of simultaneous use of oil and coal will be optimal. In the unlikely case where coal is relatively clean, the optimal economy might start with coal, after which there is a simultaneous phase and finally only coal again. So, a 'preference reversal' may occur. Simultaneous use only occurs if during the oil-only phase with oil reserves declining, oil extraction costs rise rapidly and CO2 emissions from coal are only moderate so that the comparative advantage of oil over coal decreases.

The "laissez-faire" market outcome never has a phase where coal and oil are used alongside each other. It either uses coal forever or more realistically starts with oil until it is no longer profitable to do so and then switches to using coal. We have shown that in the latter case the optimum policy requires a steeply rising CO2 tax during the oil-only phase and a less steeply rising CO2 tax during the subsequent oil-coal and coal-only phases to persuade the market to stop switching abruptly from oil to coal thereby leaving a substantial amount of oil in the crust of the earth.

Discovering new oil reserves means that the unfettered market leaves the same amount of oil in situ as before and phases in coal later. The long-run social price of energy and optimal carbon tax are unaffected, but in the optimum coal is phased in later and oil is phased out somewhat later. The optimal carbon tax is still lower during the oil-only phase. If coal is somewhat cheaper, it is phased in earlier by the market. It lowers the whole time path of the optimal carbon tax. There is later phasing in of coal and much earlier phasing out of oil in the socially optimal outcome and the oil-coal phase lasts much shorter..

If it is easier to substitute current consumption for future consumption, the market phases in coal earlier. The socially optimal outcome has a higher carbon tax which induces coal to be phased in and oil to be phased out more quickly. Since society cares less about the impact of global warming on future generations, it ends up with a bigger climate problem.

If the government is not precautionary and applies a higher market rate of discount rate, it phases in coal more quickly and phases out oil much later. The optimal carbon tax is then initially much lower and takes much longer to reach its long-run value. This results in more carbon pollution in the short run which exacerbates global warming. If the government is too optimistic about the costs of global warming, it phases in coal to early and phases out oil too late, and thus ends with more CO2 in the atmosphere.

If we allow for renewables and if they are optimally introduced after a coal-only phase, then during the new coal-only phase the social cost of carbon is smaller than before and thus the phasing in of coal and the phasing out of oil are brought forward somewhat. This shortens the length of the oil-coal phase. Since more oil is used and less coal, there will be less global warming and thus the optimal carbon tax can be lower. The time profile of the optimal carbon tax is similar, albeit that it stops increasing once renewables are introduced. A lower cost of carbon-free renewables implies that oil will be pumped slightly more vigorously and quickly, coal phased in more quickly albeit for a shorter period, and renewables phased in earlier. Although the optimal carbon tax will be lower, the CO2 concentration and global warming are curbed. If a carbon tax is infeasible, a prohibitive coal tax ensures that the market uses oil for much longer (albeit not as long as in the first-best outcome) and exhausts a larger portion of all oil reserves before switching to renewables. This second-best policy yields a much bigger improvement in green welfare than the first-best policy of a carbon tax, but of course not in total welfare. The boost to green welfare is thus not enough to offset the loss in the private component of social welfare which results from not using coal. Still, the coal tax does surprisingly well. Making renewables economically viable by subsidizing them leaves the oil-only phase, the transition time and the amount of oil left in situ unaffected, but does stabilize the stock of CO2 once renewables are introduced. This gain in green welfare dominates the welfare cost of the lump-sum taxes needed to finance the subsidy if the required subsidy is small and the global warming challenge is acute.

It may be worthwhile to extend our analysis in several directions. First, one might allow for imperfect substitution in the demand for oil and coal. This may arise from concerns with security of energy supplies, diversification and/or intermittence of backstops. Second, one might investigate what happens if there are various types of backstop available at the same time. If it is possible to rank them, e.g., clean but competitive (nuclear), clean and expensive (wind, solar, advanced nuclear) and dirty and expensive (tar sands), it is best to go for the cleanest and cheapest backstop. However, with dirty and cheap backstops, matters are more complicated especially if we allow for upward-sloping supply schedules of the backstop. Third, one could allow for the impact of exhaustibility of coal as well as oil on optimal climate policy. The challenge is then to offer a comprehensive analysis of what the optimal transition times are for phasing in oil, coal and the various types of renewables (cf., Chakravorty et al., 2008). Fourth, it is of interest to set our analysis within the context of a Ramsey growth model (cf., van der Ploeg and Withagen, 2010b) to investigate why developing countries have a bigger incentive to use coal than oil or renewables. Also, one could apply the theory of exogenous growth or that of endogenous growth and directed technical change to investigate how renewables can be introduced much more quickly (Bosetti et al., 2009; Aghion et al., 2009; Acemoglu et al., 2010). Finally, one could study international aspects such as carbon leakage and ways to sustain international cooperation within the context of a multi-country version of our model (cf., Hoel, 2008; Eichner and Pethig, 2010).

References

Acemoglu, D., P. Aghion, L. Bursztyn and D. Hemous (2010). The environment and directed technical change, Working Paper 15451, NBER, Cambridge, Mass.

Aghion, P., R. Veugelers and C. Serre (2009). Cold start for the green innovation machine, Policy Contribution 2009/12, Bruegel, Brussels.

Bosetti, V., C. Carraro, R. Duval, A. Sgobbi and M. Tavoni (2009). The role of R&D and technology diffusion in climate change mitigation: new perspective using the WITCH model, Department Working Paper 664, OECD, Paris.

Chakravorty, U. and D.L. Krulce (1994). Heterogeneous demand and order of resource extraction, *Econometrica*, 62, 6, 1445-1452.

Chakravorty, U., M. Moreaux and M. Tidball (2008). Ordering the extraction of polluting nonrenewable resources, *American Economic Review*, 98, 3, 1128-1144.

Edenhofer, O. and M. Kalkuhl (2009). Das Grünen Paradox – Menetekel oder Prognose, to appear.

Eichner, T. and R. Pethig (2010). Carbon leakage, the green paradox and perfect future markets, *International Economic Review*, forthcoming.

Gaudet, G., M. Moreaux and S.W. Salant (2001). Intertemporal depletion of resource sites by spatially distributed users, *American Economic Review*, 91, 4, 1149-1159.

Gerlach, R. (2009). Too much oil, *CESifo Economic Studies*, forthcoming.

Golosov, M., J. Hassler, P. Krusell and A. Tsyvinski (2009). Optimal taxes on fossil fuel in general equilibrium, mimeo.

Grafton, R.Q., T. Kompas and N.V. Long (2010). Biofuels subsidies and the Green Paradox, Working Paper No. 2960, CESifo, Munich.

Heal, G. (1976). The relationship between price and extraction cost for a resource with a backstop technology, *Bell Journal of Economics*, *7*, 371-378.

Heal, G. (1985). Interaction between economy and climate: A framework for policy design under uncertainty, in V. Smith and A. White (eds.), *Advances in Applied Microeconomics*, JAI Press, pp. 151-168.

Hoel, M. (1983). Monopoly resource extractions under the presence of predetermined substitute production, *Journal of Economic Theory*, 30, 201-212.

Hoel, M. (2008), Bush meets Hotelling: effects of improved renewable energy technology on greenhouse gas emissions, Working Paper 2492, CESifo, Munich.

Hoel, M. and Kverndokk, S. (1996). Depletion of fossil fuels and the impacts of global warming, *Resource and Energy Economics*, 18, 115-136.

Herfindahl, O.C. (1967). Depletion and economic theory, in M. Gaffney (ed.), *Extractive Resources and Taxation*, University of Wisconsin Press, Madison, WI.

Kemp, M. and N. Van Long (1980). On two folk theorems concerning the extraction of exhaustible resources, *Econometrica*, 48, 3, 663-673.

Lewis, T.R. (1982). Sufficient conditions for extracting least cost resources first, *Econometrica*, 50, 4, 1081-1083.

Livernois, J. and Martin, P. (2001). Price, scarcity rent, and a modified r per cent rule for non-renewable resources, *Canadian Journal of Economics*, 35, 827-845.

Ploeg, F. van der and C. Withagen (2010a). Is there really a Green Paradox?, Research Paper 35, Oxcarre, University of Oxford, Oxford, U.K.

Ploeg, F. van der and C. Withagen (2010b). Economic development, renewables and the optimal carbon tax, Research Paper 55, Oxcarre, University of Oxford, Oxford, U.K.

Sinn, H.W. (2008a). Public policies against global warming, *International Tax and Public Finance*, 15, 4, 360-394.

Sinn, H.W. (2008b). *Das Grüne Paradoxon. Plädoyer für eine Illusionsfreie Klimapolitik*, Econ, Berlin.

Stern, N. H. (2007). *The Economics of Climate Change: The Stern Review*, Cambridge University Press, Cambridge, U.K.

Tahvonen, O. (1997). Fossil fuels, stock externalities, and backstop technology, *Canadian Journal of Economics*, 30, 4a, 855-874.