Universität St.Gallen

Climate Effects of Carbon Taxes, Taking into Account Possible Other Future Climate Measures

Florian Habermacher, Gebhard Kirchgässner

School of Economics and Political Science,
Department of Economics

University of St. Gallen

Climate Effects of Carbon Taxes, Taking into Account Possible Other Future Climate Measures[1]

Florian Habermacher, Gebhard Kirchgässner[2]

Author's address: Florian Habermacher
SIAW-HSG
Bodanstrasse 8
CH-9000 St. Gallen
Phone   +41 71 224 31 59
Fax   +41 71 224 22 98
Email   florian.habermacher@unisg.ch
Website www.siaw.unisg.ch

Prof. Dr. Gebhard Kirchgässner
SIAW-HSG
Bodanstrasse 8
CH-9000 St. Gallen
Phone   +41 71 224 23 40
Fax   +41 71 224 22 98
Email   gebhard.kirchgaessner@unisg.ch
Website www.siaw.unisg.ch

Abstract

Increasing fuel extraction costs and global temperatures make it likely that in the medium-term future, technological or political measures against global warming will be implemented. In assessments of current climate policy, possible medium-term future developments, such as backstop technologies, are largely neglected, but such developments may crucially affect policy impacts. If such measures are implemented, a carbon tax introduced now may mitigate climate change to greater effect than recent reflections along the lines of the Green Paradox would suggest. Notably, the weak and the strong version of the Green Paradox, related to current and longer-term emissions, may not materialise. Moreover, the tax may allow the demanding countries to extract part of the resource rent, further increasing its desirability.

Keywords

JEL Classification

1 Introduction

In his seminal contribution, Pearce (1991) discussed the advantages of a carbon tax as an efficient policy instrument to reduce carbon dioxide emissions. He considered only the demand side, implicitly assuming a fixed, exogenous energy supply.

Today, a large fraction of climate economics research still exhibits the same limitation, reducing the supply side of the energy market to a static process. However, at least since the contribution of Sinn (2008), there is a growing awareness that supply-side effects can be crucial in assessing carbon emission reduction strategies. According to the claim that Sinn entitled 'Green Paradox', a realistic carbon tax, which for political reasons deviates from the optimal tax and is introduced at a low initial level but rapidly increasing over time, at a rate higher than the interest rate, might be counterproductive for the climate. Rather than delaying or reducing the exploitation of limited resources, the tax could accelerate their combustion. Sinn does not claim that the *optimal* carbon tax would rise at a rate higher than the interest rate. Instead, the relevance of the rapidly rising tax is explained in terms of political feasiblity. Past experiences have confirmed that realistically governments are not willing or not able to impose carbon taxes with a high initial level. In addition, because governments seek ways to increase their revenues and popular resistance against the tax may fall once it is introduced and its general principle becomes accepted, it is not necessarily implausible that the tax may rise rather rapidly over time. Therefore, here we start from the assumption that the tax would indeed rise at the postulated high rates, even if, as van der Ploeg (2013) shows, the optimal tax would rise less rapidly. That is, we clearly consider a second-best world. In this case, the Green Paradox occurs if in the early periods optimising owners of fuel anticipate the tax to be higher in the future, which causes them to sell more of their fuels today rather than on the highly taxed future markets. While controversial, Sinn's analysis has impressively demonstrated the importance of supply-side effects for the assessment of greenhouse gas policies.

A growing body of literature addresses the mentioned counterproductive effects of climate policies. Gerlagh (2011) examines the impact of supplier anticipation on the climate benefits of cheaper future backstop technologies. A similar approach is taken by van der Ploeg and Withagen (2012), who also show that a specific tax which is not rapidly increasing could be beneficial for the climate but do not discuss effects of other non-optimal taxes. Polborn (2011) concludes that intensifying research on carbon capture and storage has the advantage of reversing the negative anticipation effects that research on backstop technologies would have in terms of near-term carbon emissions.

The analyses by Sinn and subsequent contributors assumed a world in which the debated policy would be the only potential relevant climate measure and that this would hold today and forever. Abstaining from a carbon tax today, however, does not imply that neither a carbon tax, nor any alternative climate relevant development, such as backstop technologies, global fuel demand cartels à la Kyoto, or carbon capture and storage systems, will materialise in the future. Rather, absent substantial measures today, the unlimited growth of the climate threat may increase the urgency of future measures, requiring even more stringent future measures than if the carbon tax were introduced today.

Hoel (2010) considers this issue. He notes that purposely avoiding the introduction of a current tax influences only the probability of having a tax in the medium or long-term future, rather than strictly preventing any potential future tax.[1] In a stylised two-period model with an endogenous carbon tax in the second period he finds that the impossibility of long-term commitments that typifies current politics increases the (environmental) desirability of introducing a carbon tax today. Sinn (2008) himself pointed out a number of potential remedies that mankind should attempt to apply in future, some of which could be used to also reduce cumulative long-term extractions or emissions. He clearly believes that certain measures may be possible in future as he writes:

> "the good thing about the Kyoto Protocol is that it did show that world-wide cooperative agreements are possible. Integrating the three big countries mentioned [India, China and the US] and Australia, which recently announced that it wants to sign, would mean that another 45% of carbon consumption, in total three quarters of world consumption, would be captured. This share in itself would be substantial, and there could be hope that the remaining quarter could also be disciplined by political means. If the world acts quickly, before the resource owners have time to react, it might be possible to establish a world-wide trading system without loopholes."
>
> *Sinn, 2008.*

However, he has not considered how, if implemented in future, these measures may alter the conclusion about the effect of current taxes or measures.

We assess the impact of current carbon taxes given that even if a tax is currently avoided, other climate measures, such as backstop technologies, global fuel demand cartels à la Kyoto, carbon capture and storage systems, or alternative carbon taxes, may be introduced at some point in the future, that is, the relevant baseline scenario is not a perpetual business-as-usual (BAU). To keep the model tractable, we generally assume that these

---

[1] van der Ploeg and Withagen (2011) discuss the effect of dirty and clean backstops on optimal carbon taxation; however, they focus on backstops that are already available today, and, more importantly, they consider only the choice of optimised or prohibitive tax paths, leaving aside the arbitrarily increasing taxes of the Green Paradox.

future measures will be introduced independently of the current tax. We discuss the endogeneity of such measures, which could strengthen our results (section 6.3).

In the case of a future regime change, such as the emergence of a backstop technology, any presently implemented positive tax path that bridges the time until the future measure becomes effective will reduce cumulative emissions not only in the long term but already in the medium term, suggesting that the strong version of the Green Paradox may not hold.[2] This result holds for worldwide taxes and a future measure that becomes effective at a specific, anticipated future time. We show at least for limited tax levels that the results remain valid in the case of regional taxes. The exact type of the future scheme does not affect our findings. According to recent estimates, the warming effect of emissions in the current century will remain almost unchanged over the next 1,000 years (Solomon et al., 2009). This finding suggests that in terms of medium-term emissions the *cumulative* emissions are of primary importance, and the exact path of the emissions across the decades is only of limited additional relevance. Thus, by reducing cumulative medium-term emissions, the tax is very likely to have favourable effects on the climate.

Even if the point of the backstop introduction is stochastic rather than fixed and known, under these conditions the weak version of the Green Paradox does not necessarily hold; taxes that increase at rates higher than the real interest rate can reduce not only cumulative emissions for some future period but reduce current and near-term emissions as well.

This analysis has important implications for climate policy assessment. There are numerous assessments of climate policy measures, but these studies typically compare scenarios with the measure in question to a business-as-usual scenario including no alternative climate policy measures. There is, however, no reason to believe that a decision made about a particular climate policy will be decisive for every other potential climate measure as well. Taking the possibility of alternative future climate measures into account may be necessary to prevent strongly biased results.

Section 2 describes the model for the resource owners' intertemporal decision problem, and Section 3 shows how the anticipation of a backstop implemented in the medium-term future affects the resource suppliers' behaviour in the business-as-usual scenario with no present tax. We explain that the anticipation of the future regime change creates a situation that is comparable to a future high tax. These findings imply that anticipation effects pointed out by Sinn (2008) make the introduction of a present tax especially urgent. Section 4 shows that a tax bridging the time from today to the introduction of the backstop will unambiguously reduce cumulative medium-term emissions. Section 5 explains how

---

[2]Following Gerlagh (2011) we use the notions of weak and strong versions of the Green Paradox to differentiate between the increase in current emissions (weak) or of the net present value of cumulative emissions (strong) due to the anticipation of cheaper clean energy.

the analytical derivation in the previous section extends to alternative future schemes. In Section 6, we discuss possible extensions of the model and show the robustness of our analysis to a regional tax. In addition, we show that a stochastic point of emergence of the backstop implies that a tax can unambiguously reduce short- and medium-term emissions even if it increases at a rate higher than the revenue discount rate of the fossil fuel owners. We also briefly discuss the possible endogeneity of the future scheme switch. Section 7 provides a discussion and Section 8 concludes.

2 Model

Considering the different categories of fossil fuels as one *resource*, we assume a world in which consumers' instantaneous demand rate r_t, which equals the extraction rate, is a continuous, strictly decreasing, and potentially time-varying function of its price, p_t. Thus, we have the demand curve, $r_t(p_t)$, as well as its inverse, $p_t(r_t)$, as two strictly decreasing functions, $r_t'(\cdot) < 0$, $p_t'(\cdot) < 0$, where the strict inequalities may only cease to apply if the values of r or p reach their respective upper or lower boundaries, should these exist.

Instantaneous extraction rates integrate to cumulative extractions, A_t, which are normalised to zero at the starting time, $A_0 \equiv 0$, $A_t = \int_0^t r_s \mathrm{d}s$. Extraction costs, c, are assumed to be strictly increasing in the cumulative extractions, $c'(A) > 0$. This relationship implies that the most easily extractable resources are extracted first – a standard assumption that has been shown to be a necessary condition for the optimality of an extraction path (Herfindahl, 1967).

The resource owners maximise the present value of expected total net revenues, applying a positive discount rate ρ. Given a specific carbon tax path τ_t, the revenue flow for a specific seller i at time t is $r_{t,i} \cdot (p_t - c_t - \tau_t)$, where $r_{t,i}$ is seller i's extraction rate, and the suppliers' maximisation problem can thus be written as

$$U_i = \max_{r_{t,i}} \int_{t=\underline{t}}^{\bar{t}} \mathrm{e}^{-\rho t} r_{t,i} \cdot (p_t(r_t) - c(A_t) - \tau_t)\, \mathrm{d}t \tag{1}$$

$$\text{s.t.} \quad \dot{A}_t = r_t \text{ and } A_0 = 0, \text{ i.e. } A_t = \int_{s=0}^t r_s \mathrm{d}s, \text{ and } r_t = \sum_i r_{t,i}.$$

In the competitive ('comp') case, suppliers' individual rates are so small that each considers the market price as independent of his own supply, while the monopolistic ('mono') supplier will take the effect of his extraction rate on prices into account because the total rate equals his own supply rate, $r_t \equiv r_{t,i}$. Defining P_t as the considered rate of change of the gross sales revenues, $r_t \cdot p_t(r_t)$ in Eq. (1), we thus have in the two considered variants

of the model:

$$P_{t,\text{mono}}(r_t) \equiv \left. \frac{\partial \left[r_{t,i} p_t(r_t) \right]}{\partial r_{t,i}} \right|_{\text{mono}} = p_t(r_t) + r_t p_t'(r_t)$$

$$P_{t,\text{comp}}(r_t) \equiv \left. \frac{\partial \left[r_{t,i} p_t(r_t) \right]}{\partial r_{t,i}} \right|_{\text{comp}} = p_t(r_t)$$

Taking this into account in the current-value Hamiltonian,

$$\mathcal{H} = r_t \cdot (p_t(r_t) - c(A_t) - \tau_t) - \lambda_t r_t, \tag{2}$$

we arrive at the following two first order conditions:

$$\frac{\partial \mathcal{H}}{\partial r_t} = 0 : \quad P_t(r_t) = c(A_t) + \tau_t + \lambda_t \tag{3}$$

$$\dot{\lambda}_t = \rho \lambda_t + \frac{\partial \mathcal{H}}{\partial A_t} : \quad \dot{\lambda}_t = \lambda_t \rho - \dot{c}_t, \tag{4}$$

where we defined $c_t \equiv \mathrm{c}(A_t)$ and used the fact that $r_t = \frac{\partial A_t}{\partial t}$, to develop $\dot{c}_t \equiv \frac{\partial c(A_t)}{\partial t} = \frac{\partial A_t}{\partial t} \frac{\partial c(A_t)}{\partial A_t} = r_t \frac{\partial c(A_t)}{\partial A_t}$, and where λ_t is the shadow value at time t for a marginal unit of resource stock after the cumulative extraction of A_t previous units. This multiplier λ_t is a non-negative value; as with a larger resource stock, the producer's future extraction costs will be reduced and the future achievable profit will therefore be (weakly) higher.

The backward- and forward-looking explicit solutions for the multiplier in Eq. (4) become for any $\underline{t} < t < \bar{t}$,

$$\lambda_t = e^{\rho(t-\underline{t})} \lambda_{\underline{t}} - \int_{s=\underline{t}}^{t} e^{\rho(t-s)} \dot{c}_s \mathrm{d}s$$

$$\lambda_t = e^{\rho(t-\bar{t})} \lambda_{\bar{t}} + \int_{s=t}^{\bar{t}} e^{\rho(t-s)} \dot{c}_s \mathrm{d}s. \tag{5}$$

The primary assumptions on which we will base our analysis of the supply behaviour implicitly defined by the maximisation problem are as follows:

- Property 1: $p(0) > c(0)$, i.e., in the absence of a tax, there will be a strictly positive extraction rate, at least at the start.

- Property 2: $p(0) < \infty$, i.e., the choke-price is finite. This assumption is intuitive, notably as surrogates such as renewable wood or plant oils lend themselves as natural substitutes.

- Property 3: $c(A) < p(0) \Rightarrow 0 < c'(A) < \infty$, i.e., as long as some resources are profitably extractable, the rate of increase of the extraction costs is strictly positive and finite.

- Property 4: $\lim_{r \to \infty} p(r) = 0$, i.e., when the supply rate tends to infinity, the demand price becomes zero.

- Property 5: Single crossing in the first order conditions for the monopolistic supplier: the marginal revenue of a monopolist's resource sales at a specific period is decreasing in the current rate of extraction, i.e., $\frac{\partial[p(r)+p'(r)r]}{\partial r} < 0$ holds in the case of the globally homogenous market, and $\frac{\partial[p(r,\tau)+p'(r,\tau)r]}{\partial r} < 0$ in the case of the regional tax.[3]

3 Future Regime Change in the Baseline

In our analysis, the baseline (BS) scenario refers to the case in which no present tax is introduced. However, we generally assume it to contain a relevant future regime switch at an exogenously given time,[4] that is, it explicitly differs from a business-as-usual (BAU) continuation of the current situation. This section compares this baseline to the BAU case in which no future regime change would occur.

Without taking taxes into consideration, the Hamiltonian formulation from Eqs. (2) through (4) presents the suppliers' maximisation problem using $\tau_t = 0$.

As a first step we assume that a technological breakthrough provides a (cheap enough) *backstop* at time T that prevents future sales. In this case, the impossibility of profitable post-T sales implies $\lambda_T = 0$. Using $\bar{t} = T$ in Eq. (5), this yields

$$\lambda_t = \int_{s=t}^{T} e^{\rho(t-s)} \dot{c}_s \mathrm{d}s.$$

This also signifies that λ_t approaches zero for $t \to T$.

On the other hand, *if no backstop is introduced*, we know that $\lim_{t \to \infty} \lambda_t e^{-\rho t} = 0$. Using thus $\bar{t} = \infty$ and $\lim_{\bar{t} \to \infty} e^{-\rho \bar{t}} \lambda_{\bar{t}} = 0$ in Eq. (5) we get

$$\lambda_t = \int_{s=t}^{\infty} e^{(t-s)\rho} \dot{c}_s \mathrm{d}s.$$

Therefore, for any time t prior to the extraction of the last unit, the multiplier λ_t will take on a strictly positive value in the BAU without backstop. This will notably be the case for the time of the implementation of the backstop in the alternative baseline scenario, i.e., at T: defining the backstop-scenario as a case where the backstop is *relevant* implies

---

[3]This assumption seems largely unproblematic; an extended note on it is provided in Part A of the Annex.

[4]See Section 6 for a discussion of this assumption of exogeneity and why its relaxation would generally strengthen our results.

that it would be introduced at a time before the resource extraction would otherwise have stopped. We thus have

$$\lambda_T \begin{cases} = 0 & \text{for} \qquad \text{BS (with backstop)} \\ > 0 & \qquad \text{BAU (no backstop).} \end{cases} \tag{6}$$

It is obvious that the introduction of the backstop at time T affects the resource owners' optimisation problem exactly in the same manner as would a tax introduced at T if the tax rate were to be high enough to prevent any oil sales from time T onwards. This effect suggests the primary mechanism by which we will find that, in expectation of alternative future schemes, the introduction of stringent present carbon dioxide taxes is more likely to be beneficial than detrimental. The suppliers anticipate a future tax-resembling measure; if no tax is introduced today, the situation for the suppliers will correspond to one with a high future tax but with no tax in the present. This is exactly the case in which the Green Paradox would (correctly) predict a counterproductive outcome, at least in terms of short-term emissions. From this point of view, a tax introduction today becomes even more urgent the more anticipation effects drive resource owners.

As emphasised in Lemma 1 and proven in Part B of the Annex, the anticipated introduction of the (cheap) backstop – or theoretically of any other measure that leads to a reduction of the marginal value of the unexploited resources at the introduction time T to a specific value below its no-backstop counterpart – in the baseline scenario will increase the pre-T emissions (compared to the BAU).

Call the natural sales horizon the period up to the time when the fuel owner(s) sell the last unit of ever extracted fuel if they face an unrestricted choice of when to sell how much of the fuels in the BAU scenario, subject to the extraction cost and demand curves.

Lemma 1. *If at a time T strictly within the natural sales horizon, an anticipated measure implies a reduction (increase) of the marginal value of the unexploited resources to a constant marginal value $\lambda_T(A_T) = \lambda_{T,BS} = const$ below (above) the BAU value without the measure, $\lambda_{T,BAU}$, the cumulative extraction during the period up to T increases (decreases).*

Section 5 provides a note on the interpretation of a time-T marginal rent that does not vary with cumulative extractions, $\lambda_T(A_T) = const$.

Further, be the theoretically exploitable fuel the amount of fuel for which there exists some demand for a price covering the extraction costs, i.e., the amount of fuel extractable for costs lower than the demand choke price. We are going to show that an (anticipated) prevention of fuel sales after a time T strictly within the natural sales horizon implies that some of the theoretically exploitable fuel will be left underground:

Lemma 2. *An external shortening of the sales horizon below its natural duration implies that the fuel owners leave some of the theoretically exploitable fuels underground.*

Whilst we provide the analytical proof in Annex C, the intuition behind Lemma 2 is straightforward. In the business-as-usual scenario without external shortening of the sales horizon ending at a finite or infinite t_{last}, a strictly positive rate of extraction obtains during the periods within the natural sales horizon. Given the strictly falling demand, this means that consumer prices are strictly below the choke price during that natural fuel sales period. Now, for assumed external reasons, fuel sales are prevented after time $T < t_{\text{last}}$. As is intuitive, and as we have shown with Lemma 1, the fuel owners will in this case increase fuel sales in the pre-T phase (recall, from Eq. (6) we have $\lambda_{T,\text{BAU}} > 0$, but $\lambda_{T,\text{BS}} = 0$ for the baseline case where all fuel sales are prevented after T), and they will have lower opportunity costs for selling fuels during a specific pre-T period. That is, fuels will be sold for lower prices. As the market prices during the time approaching T were already lower than the demand choke price (otherwise no fuel would have been sold then), and the shortening of the sales horizon further lowers market prices, we have fuels sold at prices strictly below the choke price during the periods immediately preceding the final period T. Therefore, fuel owners do not extract fuels that have extraction costs of the level of the choke price, as else the extraction costs would exceed the gross sales revenues.

Note that while Lemma 2 is provided for the case where the time-T measure leads to $\lambda_T = 0$, it is intuitive, and known from Lemma 1, that if the fuels are to keep a certain value after period T, $\lambda_T > 0$, the amount of fuel left underground at time T must be at least as large as for $\lambda_T = 0$.

4 Introducing a Tax before the Backstop

Here, we consider the case for a present tax when the baseline contains a cheap backstop technology introduced in the medium-term future at time T. The Hamiltonian formulation with the corresponding first order conditions for the dynamic problem is given in Eqs. (2) through (4) in Section 2.

Recall from the previous section that the backstop is assumed to be cheap enough for it to drive the resource rent down to zero at the time of its introduction, $\lambda_T = 0$. This approximation, which has been used in earlier literature (see e.g. Dasgupta and Heal, 1974), may not be as far from reality as it may seem. As the backstop may replace fossil fuels in all major energy-related applications, the residual demand for fossil fuels (for example, in chemical applications), will only amount to a limited fraction of prior consumption. This will drastically reduce the expected achievable resource rent. The

smaller demand could limit the scope for monopolies as even owners of small stocks could become relevant competitors. Section 5 considers the case for a residual value λ_T for the post-T period that is non-zero and can vary with the amount of resources left at time T.

As the tax generally reduces the possible net revenues from resource sales, it seems intuitive that positive tax rates will lead to reduced *cumulative* extractions, given that the fuel owners freely choose the amount of fuel they sell, and how much they leave underground in the pre-T period.[5] This is emphasised in Proposition 1 and proven in Part E of the Annex.

Proposition 1. *If at a specific time $T > 0$, a breakthrough implies that the marginal value of resources for post-T sales becomes zero, or a different constant that does not vary with changes in cumulative extractions, $\lambda_T(A_T) = \lambda_T \geq 0$, such as is the case for the disruptive development of a cheap enough backstop preventing future sales and thus implying $\lambda_T = 0$, any scheme of positive carbon taxes up to time T, with strictly positive rates as time approaches T, leads to a reduction of cumulative emissions up to time T.*

If a regime change such as the introduction of a backstop technology is anticipated, a carbon tax thus yields a decrease of total medium-term consumption up to time T, largely independently of the form of the tax path or of the demand and production cost structure. According to our argument in the introduction, reducing cumulative medium-term emissions is of primary importance compared to the exact path of the emissions as long as relatively limited time-spans are considered. Thus, under the assumption of a future backstop in the baseline scenario, quite any path of nonnegative tax rates seems beneficial for the climate – at least in the case where the backstop is cheaper than the extraction of fuels from time T on. The case where the future measure allows for some fuel use after T is addressed in the next section.

5 Extension to Alternative Future Schemes

Proposition 1 has established that when the resource stock looses its value at a given time T – e.g., because an emerged cheap backstop prevents future fuel sales – a positive tax before this period T reduces total pre-T fuel sales. In this case, the prevention of future sales implies that the marginal value of unexploited reserves at time T, λ_T, becomes zero, independently of the amount of fuel that remains underground, A_T. For the following analysis, it is important to bear in mind that Proposition 1 further applies to the –

---

[5]Recall from Lemma 2 that a strictly positive amount of the theoretically exploitable fuels is left underground until time T.

admittedly somewhat more abstract – case where instead of being zero independently of A_T, the value of the marginal reserve available for post-T sales would, whilst still independent of A_T, be positive rather than zero.

In reality, this value of the marginal remaining resource unit for post-T sales, λ_T, is, if above zero at all, unlikely to be independent of the size of the stock of remaining resources. That is, without a (perfect) backstop, we generally expect λ_T to vary with the cumulative extractions at time T, rather than being constant. This can be expected, for example, if the post-T scheme is a demand cartel or an extremely high tax that still allow for a certain amount of extractions: If the post-T regime does not prohibit all lucrative sales of the resource, fuel-owners will derive profits from sales, which, for the marginal resource, correspond to λ_T and can positively or negatively depend on the amount of resources left underground. Satiation tends to decrease the marginal profit derived from additional resources, but the lower extraction costs for added remaining resources tends to increase them. Thus, without making further assumptions about the exact nature of the post-T resource market framework or about extraction costs or the demand function, it cannot be known *a priori* whether $\lambda_T(A_T)$ is upward or downward sloping. Using Lemma 1, we show that one can rule out one case for the relationship between λ_T and A_T in the region of the optimally chosen amount of cumulative extractions, A_T^*. We then discuss why the derived restriction on the relationship between λ_T and A_T signifies that Proposition 1 extends to cases with a flexible implied final multiplier $\lambda_T(A_T)$.

First, it seems useful to provide afew additional words on the relationships between the marginal value and the amount of cumulative exploitations at the time of the introduction of the new regime, λ_T and A_T. We have introduced the function $\lambda_T(A_T)$ as the value of a marginal additional unexploited unit of resource at time T that is available for the post-T period. This value is defined as the additional (expected) profit the resource owner can make in the post-T future if he has a marginally increased stock of remaining exploitable resources at time T. Thus, it depends on how the post-T fuel market framework looks. Conversely, the function $A_T(\lambda_T)$ designates the cumulative amount of pre-T sales the resource owner chooses when the value of a marginal unit left underground at T is λ_T. The function therefore corresponds to the amount of pre-T sales for which the sale of an additional marginal unit in the pre-T period would yield exactly λ_T additional corresponding units of pre-T profits (ignoring the influence on the post-T situation). To maximise his overall profits, the resource owner will choose an amount A_T^* of pre-T sales for which the marginal additional pre-T profit for another sold marginal unit in the pre-T period just equates the marginal forgone profit from post-T sales due to the increase of the pre-T exploitations. In other words, if A_T^* denotes the chosen (optimal) amount of pre-T sales, the following condition is satisfied

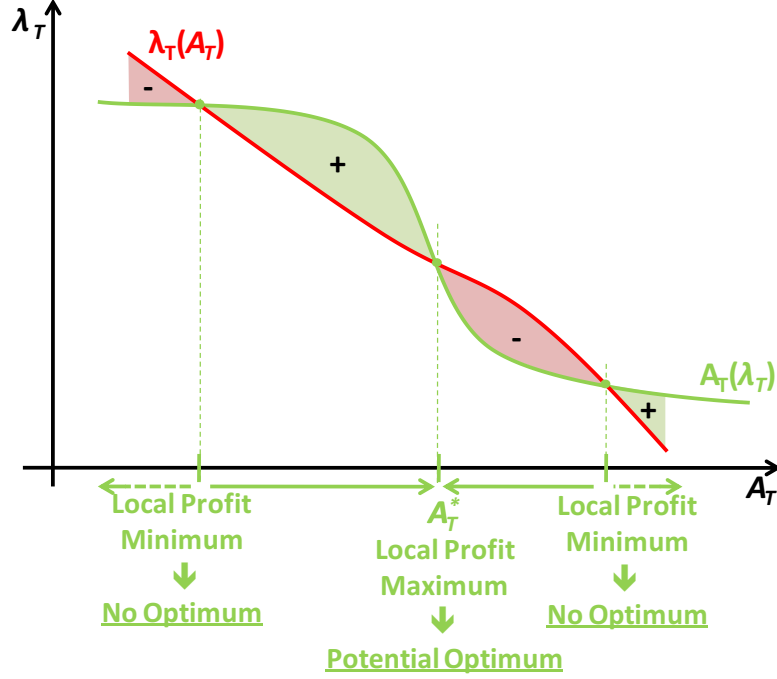$$\lambda_T(A_T^*) \overset{!}{=} A_T^{\mathrm{inv}}(A_T^*),$$

11

Figure 1: *Possible equilibrium situations with flexible* $\lambda_T(A_T)$

where $A_T^{\mathrm{inv}}(\cdot)$ is the inverse function of $A_T(\lambda_T)$.

Let $\lambda_T^{\mathrm{pre}}(A_T) \equiv A_T^{\mathrm{inv}}(A_T)$, whose simple interpretation is the marginal pre-T profit from additional pre-T sales given A_T units sold until T. For clarity, let $\lambda_T^{\mathrm{post}}(A_T) \equiv \lambda_T(A_T)$.

Recall from Lemma 1 that $A_T(\lambda_T)$ is decreasing in λ_T. For the optimal amount of cumulative exploitations A_T^*, the condition

$$\frac{\partial \lambda_T^{\mathrm{pre}}(A_T^*)}{\partial A_T} \leq \frac{\partial \lambda_T^{\mathrm{post}}(A_T^*)}{\partial A_T} \tag{7}$$

must hold, as otherwise it would be lucrative for the resource owner to increase A_T^*: the change in overall discounted profits, $\Pi = \Pi^{\mathrm{pre}} + \Pi^{\mathrm{post}}$, can be approximated as

$$
\begin{aligned}
\Pi(A_T^* + \varepsilon) - \Pi(A_T^*) &= \Pi^{\mathrm{pre}}(A_T^* + \varepsilon) + \Pi^{\mathrm{post}}(A_T^* + \varepsilon) - \Pi(A_T^*) \\
&\approx \varepsilon \lambda_T^{\mathrm{pre}}(A_T^*) + \frac{\varepsilon^2}{2} \frac{\partial \lambda_T^{\mathrm{pre}}(A_T^*)}{\partial A_T} - \varepsilon \lambda_T^{\mathrm{post}}(A_T^*) - \frac{\varepsilon^2}{2} \frac{\partial \lambda_T^{\mathrm{post}}(A_T^*)}{\partial A_T} \\
&\approx \frac{\varepsilon^2}{2} \left[\frac{\partial \lambda_T^{\mathrm{pre}}(A_T^*)}{\partial A_T} - \frac{\partial \lambda_T^{\mathrm{post}}(A_T^*)}{\partial A_T} \right]
\end{aligned}
\tag{8}
$$

for small deviations from A_T^*. Clearly, if Eq. (7) does not hold, Eq. (8) would imply profits that increase for any small value of ε, i.e., A_T^* would not be a profit-maximizing choice. This result is illustrated graphically in Fig. 1, where the pluses indicate regions in which it would be optimal for the resource owner to increase pre-T sales, and minuses indicate where it would be optimal for him to decrease sales.
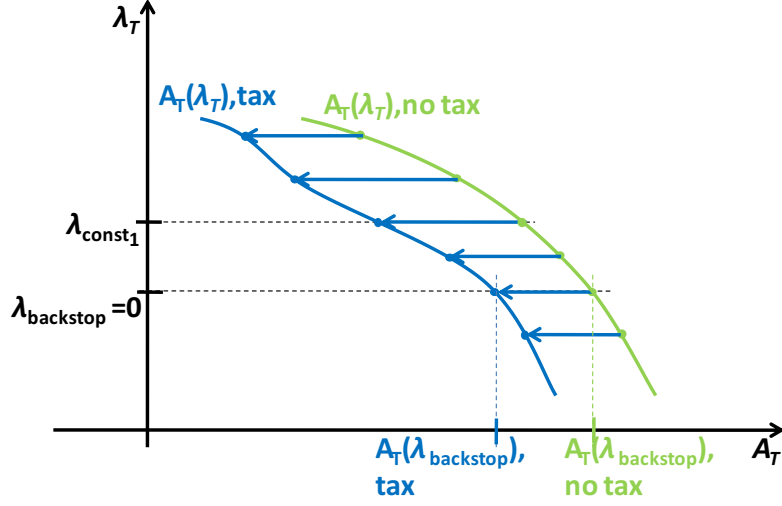
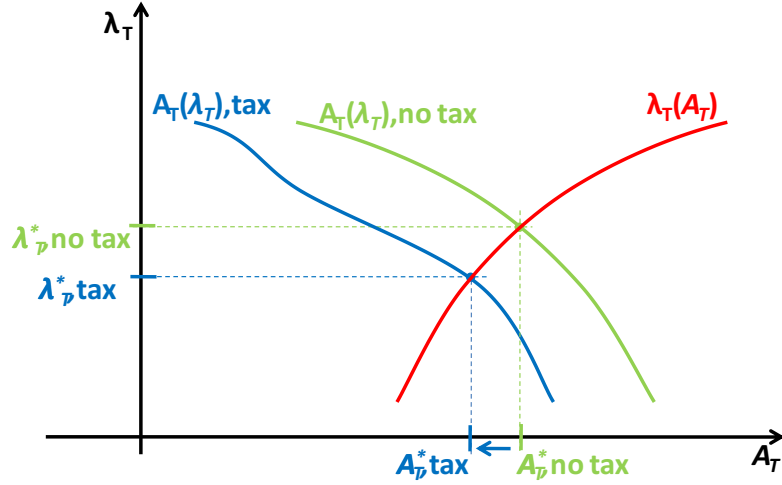Figure 2: Tax reduces pre-T emissions A_T for constant λ_T

Figure 3: Tax reduces pre-T emissions A_T for flexible $\lambda_T(A_T)$ when $\lambda'_T(A_T) > 0$

Second, recall from Proposition 1 that the tax unambiguously reduces pre-T sales for any given fixed λ_T. As the function $A_T(\lambda_T)$ remains the same here as when $\lambda_T^{\text{post}}(A_T)$ was constant, we thus know that in a diagram with A_T on the horizontal axis, $A_{T,\text{tax}}$ must lie strictly to the left of $A_{T,\text{no}}$ in all relevant ranges, as is shown in Fig. 2.

In the case where $\lambda_T^{\text{post}\prime}(A_T^*) > 0$ it is implied that the tax reduces the optimal amount of pre-T sales A_T^*. This prediction is illustrated in Fig. 3.

As $\lambda_T^{\text{post}\prime}(A_T^*) > 0$, $\lambda_T^{\text{pre}\prime}(A_T) < 0$ and $A_{T,\text{tax}}(\lambda_T) < A_{T,\text{no}}(\lambda_T)$, we have $A_{T,\text{tax}}^* < A_{T,\text{no}}^*$.

By a similar argument and using Eq. (7) it becomes clear that even if $\lambda_T^{\text{post}\prime}(A_T) < 0$, $A_{T,\text{tax}}^* < A_{T,\text{no}}^*$ holds. This situation is depicted in Fig. 4.

Therefore the proposition from the previous section extends to the case of a flexible final multiplier, $\lambda_T^* = \lambda_T(A_T^*)$, which we summarise in Proposition 2.
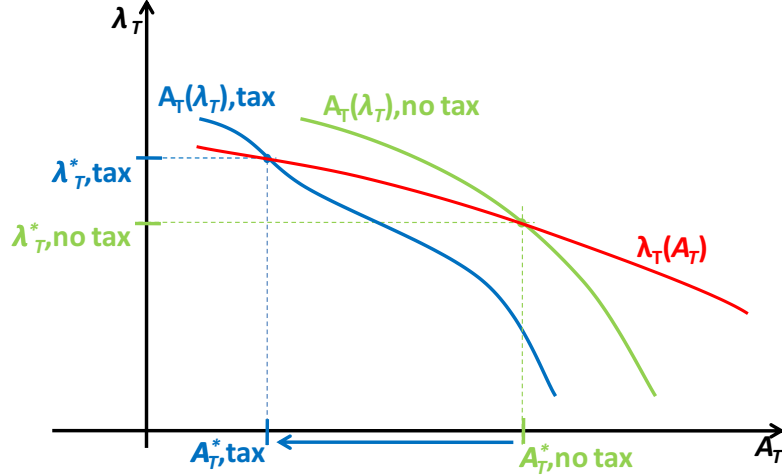
13

Figure 4: Tax reduces pre-T emissions A_T for flexible $\lambda_T(A_T)$ when $\lambda'_T(A_T) < 0$

Proposition 2. *If at a specific future time T an alternative climate measure, such as a Kyoto-like demand capping or a mandatory carbon capture and storage scheme, is introduced, implying that the marginal value of resources for post-T sales is a continuous differentiable function of the cumulative extractions up to T, $\lambda_T = \lambda_T(A_T)$, any scheme of positive CO_2 taxes up to time T leads to a reduction in cumulative emissions up to time T.*

Thus, for an independent political or technological development replacing a potential initial tax at an exogenously given time T, the debated pre-T tax yields a decrease of total medium-term consumption up to time T. This is generally independent to the extent to which fuel owners will be able to make use of their resources left underground after time T. We leave it open which – if any – exact future measure will lead to a regime switch in the future. After all, many different technical or political developments are theoretically possible, and, as history teaches us, even not yet thought-about developments may become relevant rather abruptly. Nevertheless, rather generally applicable reasonings, detailed in Annex H, suggest that a fuel saving up to time T may well imply a sustained strict reduction of cumulative emissions for a substantial time beyond T, potentially perpetually.

6 Further Extensions

6.1 Regional Tax

So far, experiences with climate protection discussions suggest that if an international carbon tax is introduced in the near future, not all countries will necessarily be willing to participate in such a treaty. We therefore examine the effect of a bridging tax that

14

remains limited to a part of the world. Analytically, this implies that the world, with respect to its demand for fossil fuels, is split into two regions: Region 1, which imposes a tax on its carbon emissions, and Region 2, which will not take any such regulatory action in the close future. In our model for this divided world we first assume that the ratio by which the worldwide demand is split is (for a consumption price that is the same in both regions) fixed and constant over time. We explain below why the derived conclusions extend to the case in which the fractions composing the two regions of the world change over time. This last point may be relevant as the parts of the world that have been revealed as the leaders or the laggards, with respect to commitments in the current political climate debate, exhibit not only distinctive climate intensities but also different rates of growth in demand.

Demand structure   A demand for fossil fuels that is split into two fixed regions implies that the demand for a specific price in one region can be expressed as a multiple of the corresponding demand in the other region. Accordingly, we introduce the variable x as the following ratio:

$$r_2(p) = x \cdot r_1(p),$$

i.e. x indicates which multiple of the demand in Region 1, r_1, corresponds to demand in Region 2, r_2.

The worldwide demand is the sum of both regions' demands,

$$r = r_1 + r_2.$$

When a tax is levied in Region 1, the *consumption price* for the resource, p_1, will be the sum of the *consumption price* of Region 2, p_2, and the tax level, τ. The price p_2 corresponds to the *sales price* for the resource owner, p_R:

$$
\begin{aligned}
p_1 &= p_2 + \tau = p_R + \tau \\
p_2 &= p_R
\end{aligned}
$$

The demands of the two regions, $r_1(p_1)$ and $r_2(p_2)$, can thus be expressed as $r_1(p_1) = r_1(p_R + \tau)$ and $r_2(p_2) = r_2(p_R)$. Thus, as shown in Fig. 5, the total demand for a given sales price and tax rate is

$$r(p_R, \tau) = r_1(p_1) + r_2(p_2) = r_1(p_R + \tau) + r_2(p_R). \tag{9}$$

Effect of the tax on the demand   The demand curves of both regions are assumed to be continuous and strictly decreasing. Thus, Eq. (9) implies that the current worldwide
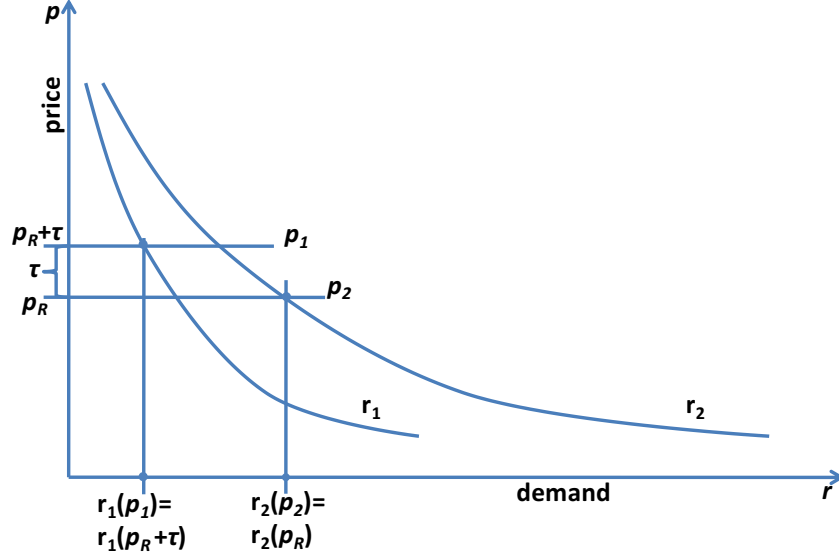
Figure 5: Regional demand with tax in Region 1

demand decreases as well in the current sales price p_R as in the current tax rate τ. Therefore, the inverse demand curve (here, the sales price which yields a specific demand, $p_R(r,\tau)$), is strictly decreasing in r.

Given this new demand structure, the optimality condition Eq. (3) becomes

$$p(r_t, \tau_t) + r_t \frac{\partial p(r_t, \tau_t)}{\partial r} = c(A_t) + \lambda_t. \tag{10}$$

While the answer eventually seems intuitive, without any further analytical inspection it is not necessarily clear whether the LHS of the supplier's adapted FOC, Eq. (10), decreases unambiguously in τ_t in the case of the regional tax. Hence, it is proven and stated as a general result in Lemma 3 (which is stated below Proposition 3 and its proof given in Part F of the Annex), at least for limited tax levels.

Thus, according to Lemma 3, a regional tax levied on Region 1's consumption at time t reduces worldwide consumption at the same time t for given multiplier λ_t and extraction costs c_t. Given this result, it is evident that the proof for Proposition 1 extends to the case of the regional tax – Lemma 3 ensures that Eq. (A.18) holds in the proof.

Accordingly, a regional tax also leads to a reduction in cumulative emissions up to time T, a conclusion that is formulated as Proposition 3. Note that while we are not aware of any particular reasons for which the statement should not extend to larger taxes as well, the proven validity of our analytical derivations for Lemma 3, and thus for Proposition 3, is restricted to certain smoothness conditions for the tax as well as to tax rates that are not too large.

The analysis remains valid in the case in which the demand-ratio between the regions, x, varies over time: Lemma 3 is not affected at all, and the proof of Proposition 1 allows for

time-varying $p_R(r, \tau)$.

We thus emphasise the following result:

Proposition 3. *If an alternative climate measure is introduced at a specific future time T, any scheme with positive carbon taxes covering a (eventually non-constant) fraction of the world's demand up to time T leads to a reduction in cumulative worldwide emissions up to T, at least for limited tax rates.*

Lemma 3. *If an interior solution to the profit maximisation problem with the first order condition*

$$p(r_t, \tau_t) + r_t \frac{\partial p(r_t, \tau_t)}{\partial r} = c(A_t) + \lambda_t$$

exists, then a current regional tax at time t levied on Region 1's consumption reduces current worldwide consumption for a given multiplier λ_t and extraction costs c_t, at least for not too large tax rates.

The proof of Lemma 3 is provided in Part F of the Annex.

6.2 Stochastic Introduction of the Future Scheme

It cannot be predicted with certainty which climate change mitigation policy may prevent the release of the remaining carbon stored in fossil fuels into the atmosphere. It would be even more unrealistic to pretend to know when exactly such a breakthrough will occur. In addition, the change may come gradually, over several years, rather than at one specific point in time, and there may be substantial uncertainty about the time of the future regime change. Finally, it could also be the case that there will be no real regime change at all. To account for these uncertainties, a stochastic model has to be considered, complicating the analysis considerably.

An analytical investigation of the stochastic case may be possible to a certain extent, especially with a backstop, at the emergence of which the resources left underground at time T will lose all their value. In such a case, the stochastic end time can readily be accounted for by augmenting the discounting rate ρ by an appropriate term ψ_t and otherwise using the deterministic model, as has been shown by Dasgupta and Heal (1974).[6]

For simplicity, we consider here the case where the probability of the emergence of a backstop, conditional on no prior occurrence (further called periodic probability), is constant. The additional discounting factor, ψ, which equals this periodic probability, inherits this

---

[6]See also Strand (2007), where the stochastic introduction of an alternative technology is shown to augment the overall discount rate considered by the resource owners.

constancy, i.e., $\psi_t = \psi$. This result implies that the analytical structure of the model does not differ from the deterministic case at all. Note that the underlying (unconditional) probability density for the introduction of the backstop at date t is then $f(t) = \psi e^{-\psi t}$.

The additional discount factor from the possible emergence of the backstop alters the conclusion about the taxes' impact on the emissions. In the case where no backstop was considered, the Green Paradox would hold to a certain extent, implying that a tax rising more rapidly than with the real interest rate could lead to larger current emissions; however, this finding is not valid in the case of the possible backstop. In this case, taxes that exponentially rise at any rate lower than $\rho + \psi$ necessarily imply reductions of current emissions and lower cumulative emissions at any future time period. We emphasise this claim with Proposition 4, and the analytical proof is given in Part G of the Annex.

Proposition 4. *Any positive tax exhibiting a rate of increase, θ, that figures between 0 and the sum of the real interest rate, ρ, and the periodic probability of the emergence of a backstop technology, ψ, leads to a reduction in the expectancy of the cumulative emissions and never yields increased potential cumulative emissions.*

The assumption that the backstop is a perfect substitute and, after the breakthrough, cheaper than the fuel, is relaxed in van der Ploeg (2013). Focusing on R&D subsidies rather than on a tax, they use a value function approach and show that the subsidy can increase initial emissions but reduce expected long-run emissions.

6.3 Endogenous Future Regime Change

The introduction of a carbon tax influences the consumption prices for conventional energy and consequently changes incentives for the (decentralised) development of alternative technologies. However, the tax also affects carbon emissions and the climate and thus the political pressure to work on (centralised) additional measures. It is clear that a present tax may therefore influence the likelihood or the timing of the implementation of future measures. The assumption that the latter is perfectly exogenous is thus a simplification of reality, and it seems important to address the possible endogeneity of the future climate regime. This problem is beyond the scope of the present work. It can, however, be foreseen that the direction of the effect on the expected results is ambiguous: lower political pressure resulting from eventual tax-induced emissions reductions could lower the probability that early alternative measures will be implemented, but technological development boosted by the higher carbon prices could imply earlier development of substitute technologies.

Especially for alternatives such as advancements in backstop or carbon capture and storage technologies, for which the time of the (potentially gradual) introduction will depend on the consumer price of the fuels, the tax path will have a direct influence on the time of the regime switch (which, for the case of a gradual introduction, may be considered as the time at which the new technology makes the standard combustion completely or almost completely redundant). We thus have T as a function of the tax path τ_t. In this case, the above results can in general even be expected to be strengthened: In the late periods before T (where λ approaches 0 even in the case without the tax), the positive tax increases the consumer price of the conventional fuel combustion, meaning the alternative technology becomes economic – and thus replaces the conventional fuel combustion – earlier. That is, the tax leads to an earlier $T_{\text{tax}} < T_{\text{BAU}}$, leaving less time to sell fuels overall, thus reducing cumulative emissions already before the time of the introduction of the alternative technology in the BAU-scenario T_{BAU}. In this sense, the endogeneity of the time when the future development becomes effective makes the strong Green Paradox even less likely to occur than in our simplified analysis above.

7 Interpretation

We saw that if the baseline scenario contains future, alternative climate-relevant measures – which cannot only be cheap backstop technologies but also, e.g., efficient global cap-and-trade systems that replace the tax in the medium-term future – the current tax reduces not only long-run but already cumulative medium-term emissions. Reducing the time during which emissions may increase, these future developments decrease the potential relevance of Green Paradox type anticipation effects with respect to future tax rate rises; increases of the net present value of discounted future emissions become less likely.

At the current rate of consumption, the well-assessed, worldwide oil reserves last for another 46 years. Given past growth rates of the worldwide fuel consumption, it is plausible that without any relevant political or technological developments the large majority of the oil resources, which exceed the reserves by a factor of around three, would be burned well before the end of the 21st century. Gas and coal reserves-to-production ratios exceed those for oil, but growth rates of their consumption have exceeded even those for oil in the last decades.[7] It is therefore foreseeable that in a BAU future a large fraction of the overall extractable hydrocarbon reserves will be transformed into atmospheric carbon dioxide before the end of the century, which will leadto the potentially devastating warming effect of several degrees.

---

[7]For details about past, current and projected fuels consumption see the World Energy Outlook 2010, Fig. 2.4 on p. 84 (IEA, 2010). The outlook reports reserve-to-production ratios for oil, gas and coal of 46, 58 and 150 years.

If there is some hope that the climate can be saved from this scenario, this hope must be based on some stringent climate-protecting measures becoming effective well within the current century. It must be admitted that there is no ready-made solution at hand right now and that some pessimism may be justified given the small fruits efforts of the last decades have produced. It cannot be denied, however, that the hope to find some solutions generally exists – or else the money and political and personal efforts spent all around the world to find solutions to the climate problem would hardly be accepted. As the quote in the introduction shows, this valid hope is even explicitly acknowledged by Sinn (2008) in his seminal contribution on the Green Paradox.

At least when a measure such as a future global Kyoto is introduced, the resource owners will not sell a strictly *fixed* quantity of fuels in the time prior to its introduction. Instead – and this is what our analysis emphasises – any path of positive taxes prior to that medium-term measure will strictly reduce the amount of fuels sold in the medium-term.

For a stylised, analytically tractable illustration, we mostly assume the medium-term change to occur with certainty. We thus adopt a position fully opposite of the framework within which the Green Paradox was originally brought forward and where the probability of future alternative developments has implicitly been assumed to be zero. Reality lies between these two extremes, and our results can be considered as an illustration of the way the Green Paradox results change if the positive probability of future measures is accounted for.

The medium-term measure must not necessarily be a backstop technology replacing all fossil fuels quasi-instantaneously. Instead, it may also be a non-perfect and gradually developing new technology or an (almost) global cap-and-trade system successful enough to replace the currently introduced tax.[8] We assumed the alternative measure to be independent of the decision on the current tax. This is a strong assumption. As we explained, it is unclear in which direction our results change if the endogeneity of political and technological developments (both difficult to quantify) is taken into account. However, the tax tends to shorten the pre-regime switch period, which in general strengthens our results.

Arguments questioning the relevance of the Green Paradox have been raised previously, but the possibility of an exogenous future regime switch as part of the baseline scenario so far has received scant attention in the literature. This additional element in the modelling of the effects of a carbon tax renders the predictions more accurate and shows that a carbon tax may be more desirable than previous studies have suggested. This finding is important, especially because not all of the other points in favour of the tax raised in the literature necessarily invalidate all aspects of the Green Paradox. Hoel (2010) argues that

---

[8]Note that the analysis allowed for a time-dependent fuel demand, which readily allows modelling, e.g., the case when a backstop technology is gradually developed, gradually reducing the net fuel demand.

any positive tax rate would reduce overall (i.e. long-term) emissions in any case. This is a conclusion that can intuitively be understood given the smoothly increasing extraction costs together with a demand-price limited by a finite choke-price: while without any tax the last unit of fuel exploited would be the one for which extraction costs correspond to the choke-price, this ultimate price would be reduced by any positive tax. Because of the increasing extraction cost curve, this implies that total extractions would decrease as well. For two reasons this insight may not be considered a decisive argument in favour of a carbon tax for every case: first, depending on the form of the demand function and the extraction costs curve defining the available resource quantities, without any externally given future regime switch, the time at which the last unit of the resource would be exploited may theoretically be far enough in the future. Therefore, the timing of the emissions could no longer be considered of lesser importance compared to the absolute emissions. Second, and even more importantly, if alternative technologies are not effectively developed, the choke-price of the demand may be large enough for the extraction cost curve to already be steep at the corresponding point: given that the total amount of the fossil reserves is limited and that an important fraction is exploitable at relatively low costs, while the last drops somewhere deep in the ground would be exploitable only at very high costs, it may seem plausible that the cost curve in the region of the choke-price may be steep.[9] Such a result implies that the overall exploited quantity may vary only to a small extent in reaction to a limited tax.

Beyond what the above analysis reveals, there is an additional reason why anticipation effects could increase the desirability of a carbon tax rather than reduce it. Without the external climatic effects of the combustion of carbon-containing fuels, one may generally presume that an eventual carbon tax would be associated with negative economic effects on the taxed region. This negative effect on the economy may increase with the level of the tax, and only the negative climatic externalities may justify an eventual carbon tax. For a fixed net fuel price and in a first approximation, the optimal compromise between climate protection and economic activity should be achieved by a tax level that corresponds to the level of the marginal climate costs of an emitted unit of carbon. In this case, the demand for fuels should be reduced until an additional reduction would yield economic costs that exceed the additional benefit from increased climate protection.[10] Under these conditions, the analysis of the profit-maximising behaviour of the resource owners shows that they will reduce the net price they demand for their goods if a climate tax is introduced. Thus, the previously described 'optimal' climate tax would reduce consumption by less than the climate policy maker may have expected, should he have neglected this behavioural

---

[9]According to IEA predictions, fuel extraction cost curves indeed become very steep as the amount of fuel extracted increases.

[10]It is beyond the scope of this paper to address the numerous practical problems of the introduction of such a tax. While they are crucial for any project of a carbon tax in general, these problems seem to be of lesser importance to our specific argumentation.

adaptation; the gross price does not increase by the full amount of the tax rate but only by part of it. In this sense, at first sight one could be tempted to consider the tax as inefficient. The reduction in the demanded net sales price by the resource owners could, however, also be utilised to fix the tax rate enough above the 'optimal'[11] rate so that the gross price exceeds the net price from the no-tax scenario by the value of the 'optimal' tax rate.[12] In this case the result is the previously described demand reduction and the originally mentioned economic costs. Despite the tax rate that was higher than originally described, the costs for the economy increase only by the originally targeted value. In addition, however, the taxing region generates higher extra tax revenues, corresponding to a transfer of parts of the resource rent from the resource owners to the consumer countries. The tax-induced behavioural adaptation of the resource owners can thus be used to the advantage of the fuel importing countries and increases the economic attractiveness of such a tax. In this sense, the profit-expectation-reductions related to the anticipatory effects of the resource owners, as well as the associated attenuation of the impact of the tax on sales price and the demanded quantity, should not be considered as a tax efficiency problem. Rather, these results should be understood as a possible means to efficiently reduce simultaneously the cumulative demand as well as the import costs of the oil.

8 Conclusions

The claim that carbon taxes with rapidly increasing rates would exacerbate the climate problem rather than alleviate it cannot be sustained as generally as has been suggested. This is still the case even if one departs from Sinn's (2008) starting point that rather than opting for the socially optimal climate tax, legislators will in reality choose carbon taxes that start at a low level and rise rapidly over time.

This paper details two primary limitations with regard to the claims proposed with the Green Paradox, based on the fact that the perpetual business-as-usual is the wrong baseline scenario against which the tax must be compared. Even if we were to abstain from introducing a carbon tax today, other future climate-related developments may influence the resource market in the future and consequently the carbon emission path. Such possible developmentsencompass not only technological innovations driven by increasing fossil

---

[11]Optimal is used here in the sense of the level that would be desirable if no supply-side adaptation occurred.

[12]In this case, the result could be even further improved when the tax rate is not exactly fixed in this way. Although such a discussion is beyond the scope of this article, a discussion of the optimal tax accounting for the strategic consumer-owner interaction on the resource market can be found in Liski and Tahvonen (2004), who examine a first best climate taxation in general, and Dullieux et al. (2011) who examine the optimal tax given a 2 °C warming equivalent emission constraint. Both studies find that under certain conditions, the optimal tax may contain an import tariff component, i.e., it is larger than the pure Pigou tax.

fuel extraction costs but also political movements driven by ever-increasing emissions and climate damages. Potential measures include backstop technologies, demand cartels, carbon capture and storage systems or prohibitively high future carbon taxes. Given such possible future measures, a currently introduced carbon tax may be more favourable for our climate than has been predicted by the Green Paradox:

First, if some of the previously mentioned future climate regime switches were to materialise at the specified time in the medium term, then the cumulative emissions may become more relevant than the detailed evolution of the emission path, and any path of positive taxes can be expected to reduce these cumulative emissions up to the time of the regime switch.

Second, if a future regime switch (such as the introduction of a backstop technology) is stochastic, our model suggests that even the weak version of the Green Paradox does not hold. Even current emissions can be reduced by carbon taxes whose levels increase more rapidly than at the real interest rate.

In addition to the impact of the taxes on the climate, the anticipation effects can even be beneficial for the consumer countries as the tax allows these countries to extract part of the suppliers' resource rent, which may increase the carbon tax-related welfare gains for the consumer countries.

There are some caveats regarding the presented analytical findings. First, in the framework of the stochastic regime switch, our result provides a clear indication only for a tax whose maximal rate of increase is still limited, even if (because of the possibility of the backstop) this limit may be substantially higher than the one originally suggested by the Green Paradox. It is clear, however, that, consistent with our argumentation brought forward in the deterministic case, the examination of the stochastic case should not stop here. Even a tax that may increase faster than our elevated threshold rate of increase identified in the stochastic analysis may overall be beneficial. Such a tax may slightly increase emissions in the initial period but lead to substantial emission reductions later. In the case in which the probability distribution for the occurrence of the regime switch may indicate that the latter is likely to occur in the medium term, our argument for the primary relevance of the *cumulative* emissions should be considered as well: if the tax leads to substantial cuts in future emissions, these reductions may more than compensate for the smaller increases in earlier emissions.

Second, we mostly ignored the potential endogeneity of the future climate scheme change. This is a severe limitation as it is clear that the eventual carbon tax affects virtually all variables influencing the potential future regime switch, e.g., the temperature path, the consumer price, the general economic development, or the technical progress with alternative energies.

Finally, we explained that especially for the here relevant medium-term future the cumulative emissions may be more important than the detailed emission path. This is only a simplified view. Ideally, one would more properly weigh increases in current emissions against reductions in cumulative medium- or long-term emissions. For this case, a more realistic model for total net present damage would be desirable. Some limited discounting of future damages, coupled with a non-linear mapping of cumulative emissions (or concentrations) to damages, would ideally be considered.

An encompassing *analytical* examination of all these issues seems infeasible. In order to address them, it would therefore be interesting to explore the case for the Green Paradox by means of numerical simulations. Even if many of the relevant parameters for such an undertaking – especially those about the future climate regime switch – may be subject to large uncertainties, such a model can allow at least some approximate quantitative assessment of the qualitative claims of the Green Paradox (and of our analysis).

From a broader perspective, we would like to conclude by stressing the implications of this analysis for climate policy evaluation beyond the question of the Green Paradox. While we have shown here that future independent climate-relevant developments may strongly influence how the consequences of a carbon tax may fit the predictions of the Green Paradox, the potential future climate developments may be crucial for the net impact of any currently debated climate measure. These potential future developments should therefore be taken into account when assessing the desirability and impacts of current measures in general, which is hardly being done so far. Predictions about future climate-relevant developments, be they policy measures or technological developments, are intrinsically linked to large uncertainty and complicating reflections. However, the uncertainty of predictions is not truly reduced by simply ignoring its cause. Rather, the latter approach introduces a potentially large bias which, as shown here, may crucially affect conclusions about possible policies.

References

P. Dasgupta and G. Heal (1974), The Optimal Depletion of Exhaustible Resources, *Review of Economic Studies* 41, pp. 3 − 28.

R. Dullieux, L. Ragot and K. Schubert (2010), Carbon Tax and OPEC's Rents under a Ceiling Constraint, *The Scandinavian Journal of Economics*, 113, pp. 798 − 824.

R. Gerlagh (2011), Too Much Oil, *CESifo Economic Studies* 57, pp. 79 − 102.

O.C. Herfindahl (1967), Depletion and Economic Theory, in M.H. Gaffney (ed.), *Extractive Resources and Taxation*, pp. 63 − 69.

M. Hoel (2010), Is There A Green Paradox?, *CESifo Working Paper* No. 3168, Munich, September.

IEA (International Energy Agency) (2010), World Energy Outlook 2010, OECD/IEA, Paris.

M. Liski and O. Tahvonen (2004), Can Carbon Tax Eat OPEC's Rents? *Journal of Environmental Economics and Management*, 47, pp. 1 − 12.

D. Pearce (1991), The Role of Carbon Taxes in Adjusting to Global Warming, *Economic Journal* 101, pp. 938 − 948.

S. Polborn (2011), The Green Paradox and Increasing World Energy Demand, mimeo, University of Aarhus, January.

H.-W. Sinn (2008), Public Policies Against Global Warming: A Supply Side Approach, *International Tax and Public Finance* 15, pp. 360 − 394.

S. Solomon, G.K. Plattner, R. Knutti and P. Friedlingstein (2009), Irreversible Climate Change due to Carbon Dioxide Emissions, *Proceedings of the National Academy of Sciences* 106, pp. 1704 − 1709.

J. Strand (2007), Technology Treaties and Fossil-Fuels Extraction, *The Energy Journal* 28, pp. 129 − 141.

R. Van der Ploeg (2013), Cumulative Carbon Emissions and the Green Paradox, forthcoming in *Annual Review of Resource Economics* 5.

R. Van der Ploeg and C. Withagen (2011), Optimal Carbon Tax with a Dirty Backstop: Oil, Coal, or Renewables?, *CESifo Working Paper* No. 3334, Munich, January.

R. Van der Ploeg and C. Withagen (2012), Is There Really A Green Paradox?, *Journal of Environmental Economics and Management* 64, pp. 342 − 363.

Annex

(A) Single-crossing property for monopolist's revenue

In order to rule out some theoretically possible multiple local maxima that would be difficult to deal with analytically, we assume that the demand functions $r(p)$, or their inverses $p(r)$, exhibit the property that the marginal revenue of a monopolist's resource sales at a specific period is falling in the current rate of extraction, i.e., that $\frac{\partial[p(r)+p'(r)r]}{\partial r} < 0$, over the full range of considerable extraction rates. This condition guarantees that $P_t(r_t)$ is a strictly decreasing function not only in the competitive but also in the monopolistic case. The condition notably implies that, should the value of $P_t(r_t)$ decrease, its argument r_t increases, and vice versa. Note that the property represents only an absolutely mild assumption: typically considered demand functions, be they linear, quadratic, isoelastic, or exponential, all meet this assumption in any case. For the case of the world with a monopolist and a tax in a region covering only a fraction of the worldwide demand, the stringency of the analytically derived conclusions will require an extension of this assumption. In this case, we will assume that for any considered regional tax level τ, the worldwide demand $r(p,\tau)$, which is the sum of the demand $r_1(p+\tau)$ in Region 1 that levies the tax and the demand in the second, non-taxing region, $r_2(p)$, is such that $\frac{\partial[p(r,\tau)+p'(r,\tau)r]}{\partial r} < 0$. This condition is likely to hold as well in most cases. It can analytically be shown that the condition holds for all linear, exponential and quadratic demand forms for which the corresponding condition from the worldwide tax case holds. This conclusion applies to the quadratic demand, at least for limited tax levels. Exceptions are, however, possible for a limited subset of situations with isoelastic demand in the case of the regional tax.

(B) Proof of Lemma 1

Consider two situations, indexed BAU and BS, in the same model but with notably differing final multipliers, λ_T. We define, for a variable v, the Δv as the difference between the two situations' values, $\Delta v = v_{\mathrm{BAU}} - v_{\mathrm{BS}}$. For the case where the BS scenario has the lowered final marginal resource value, the claim of Lemma 1 can then be stated as

$$\Delta\lambda_T < 0 \Rightarrow \Delta A_T > 0, \tag{A.1}$$

with the considered time span being $t = [0, T]$. We will show by contradiction that the claim in Eq. (A.1) holds unambiguously.

Assume thus the contrary,

$$\Delta\lambda_T < 0 \wedge \Delta A_T \leq 0, \tag{A.2}$$

which we will prove to be inconsistent.

All considered variables, A_t, λ_t, r_t and c_t, exhibit continuous time paths.

This implies that $\lim_{t \to T} \lambda_t = \lambda_T$ and $\lim_{t \to T} A_t = A_T$, i.e., $\lim_{t \to T} \Delta\lambda_t = \Delta\lambda_T$ and $\lim_{t \to T} \Delta A_t = \Delta A_T$. Assuming Eq. (A.2) to hold, we thus know that the RHS in Eq. (3) (with $\tau = 0$) will be smaller for $t \to T$ in the case of the reduced final multiplier, i.e., ΔRHS < 0. Therefore, Property 5 implies that the chosen extraction rates become larger in the region where t is close to T:

$$\lim_{t \to T} \Delta r_t > 0 \tag{A.3}$$

To have Eq. (A.2) despite Eq. (A.3) requires that we have some earlier periods with lowered r_t and thus a 'last time', t^*, where Δr_t just converges from a negative level to zero and stays at or above zero until T:

$$\exists\, t^* \,\text{s.t.} \begin{cases} \Delta r_{t^*} = 0 \\ \Delta r_t \geq 0 \,\forall\, t > t^* \\ \exists\, \varepsilon > 0 \,\text{s.t.} \,\forall\, t \in [t^* - \varepsilon, t^*)\, \Delta r_t < 0 \end{cases} \tag{A.4}$$

Together with the second relation in Eq. (A.4), Eq. (A.2) implies $\Delta A_{t^*} < 0$. Together with the third relation in Eq. (A.4), this implies that we either have

$$\exists\, \varepsilon > 0 \,\text{s.t.}\, \Delta r_{t^* - \varepsilon} < 0 \;\wedge\; \begin{cases} \text{Case 1:} \quad \exists\, \delta > 0 \,\text{s.t.}\, \underbrace{A_{t^* - \varepsilon, BAU} = A_{t^* + \delta, BS}}_{\equiv a}, \\ \quad \text{or} \\ \text{Case 2:} \qquad A_{t^* - \varepsilon, BAU} > A_{t, BS}, \end{cases} \tag{A.5}$$

where case 1 refers to the situation where the unit of fuel exploited at time $t^* - \varepsilon$ in the BAU scenario is exploited after t^* in the alternative baseline scenario (BS) with the lower final multiplier, and case 2 refers to the situation where that unit of fuel is not extracted at all in that baseline BS.

Consider first case 1 in Eq. (A.5). From Eq. (A.4) we know that $\Delta r_{t^* + \delta} > 0$. Together with strictly decreasing marginal sales revenues (Property 5 in Section 2) and $\Delta r_{t^* - \varepsilon}$ (Eq. (A.5)), this implies that if a marginal unit of resource with a given, fixed extraction cost yields the (weakly) highest marginal sales profit in situation BAU by being sold at time $t^* - \varepsilon$, selling the same marginal unit must yield a strictly higher marginal profit when, in situation BS, sold also at time $t^* - \varepsilon$ rather than at $t^* + \delta$. This is, however, in contradiction with case 1 in Eq. (A.5), stating that the profit maximising resource owner chooses to sell the same marginal resource unit, defined by its extraction costs a, is sold

at time $t^* - \varepsilon$ in situation BAU but at $t^* + \delta$ in situation BS.[13]

Consider now case 2 in Eq. (A.5), for which a very similar argument leads to a contradictory result, based, however, on the refusal to extract the specific marginal unit rather than, as in case 1, a delay. As the profit maximising owner in situation BAU decides to sell the marginal fuel unit extractable at marginal costs $a < A_{T,BAU}$ at time $t^* - \varepsilon$, we know he gets a net present-value profit exceeding $\lambda_{T,BAU} e^{-rT}$ as the latter would be the present value of the net revenue from a more expensively extracted marginal unit sold at time T. Thus,

$$(P_{t^*-\varepsilon,BAU} - a) \, e^{-r(t^*-\varepsilon)} \geq \lambda_{T,BAU} e^{-rT}, \tag{A.6}$$

where P_t is the marginal gross sales revenue from marginal resource sales as defined in Section 2. Knowing, in case 2, that in situation BS the resource owner refrains from extraction of the marginal resource unit a, whilst he gets, in net present value terms, a profit of $\lambda_{T,BS} e^{-rT}$ for the marginal resource at time T, we have

$$(P_{t^*-\varepsilon,BS} - a) \leq \lambda_{T,BS} e^{-rT}. \tag{A.7}$$

As we have $\Delta r_{t^*-\varepsilon} < 0$, strictly decreasing marginal gross sales revenues (Property 5 in Section 2) imply, however, $P_{t^*-\varepsilon,BAU} < P_{t^*-\varepsilon,BS}$, which is in direct contradiction to the ensemble of Eqs. (A.6) and (A.7).

Assuming Eq. (A.2) thus yields a contradictory result in both elsewise possible cases, case 1 and case 2, which concludes our proof by contradiction. ∎

(C) Proof of Lemma 2

For all theoretically exploitable fuels to be extracted, i.e., for final extraction costs to reach the demand choke price and thus A_T to reach a level such that $c(A_T) = p(0)$, the extraction rate during the final fuel sales phase must reach zero, since only in this case the demand price for the fuels reaches $p(0)$ and only then the fuel owners can sell resources extracted at marginal costs $c = p(0)$ without incurring direct net losses. Since in the BAU without the restriction on the sales horizon, sales before T were strictly positive, the sales price was strictly below $p(0)$. In order for the market price to rise to reach $p(0)$ in the case of the shortened sales horizon, the extraction rate must thus fall to below the original level during the final periods,

$$\lim_{t \to T} r_{t,BS} < r_{t,BAU}, \tag{A.8}$$

---

[13]For a formal version of an analogous argumentation see the last part of the proof of Lemma 2 (Annex C).

where the index BAU (business-as-usual) stands for the scenario without external shortening of the sales horizon and BS (baseline) for the scenario with the externally limited sales horizon (consider, e.g., the introduction of a cheap enough, perfect backstop).

We provide a proof by contradiction that Lemma 2 must hold, assuming (A.8) to hold and showing that this leads to an inconsistency.

We know that with the shortened horizon, cumulative extractions up to T exceed those from the BAU scenario,

$$A_{T,BS} > A_{T,BAU}. \tag{A.9}$$

For this to hold, we know that there must exist some periods $t < T$ for which $r_{t,BS} > r_{t,BAU}$, and, given smooth functions and (A.8), thus also an inner period t^*, $0 < t^* < T$, for which the extraction rate is unchanged from the BAU extraction rate, $r_{t^*,BS} = r_{t^*,BAU}$, and immediately before which extraction in BS exceeds the BAU extraction rate for the last time, $\lim_{t \to t^{*-}} r_{t,BS} > r_{t,BAU}$ and $r_{t,BS} \leq r_{t,BAU} \, \forall t \in [t^*, T]$. Together with (A.9) this implies

$$A_{t^*,BS} > A_{t^*,BAU}.$$

We therefore know that there exists a time t' strictly smaller than t^* for which the then sold marginal unit of fuel in the BS scenario, identified by its unitary extraction cost $A_{t',BS}$, would have been extracted at a time t'' strictly later than t^* in the BAU scenario:

$$\exists t' < t^*, t'' > t^* \,\text{s.t.}\, A_{t',BS} = A_{t'',BAU}.$$

At the same time we know that

$$r_{t',BS} > r_{t',BAU} \, \wedge \, r_{t'',BS} \leq r_{t'',BAU}. \tag{A.10}$$

We denote $a = A_{t',BS} = A_{t'',BAU}$.

As a last step for the proof, we show that the differences in the sales price paths for the fuels between the two scenarios, pointed out in (A.10), are incompatible with the fuel owner choice to allocate the specified marginal unit of fuel to time period t' in the BS scenario whilst allocating it to period t'' in the BAU scenario:

From the strictly falling marginal sales revenues (Property 5 in Section 2), Eq. A.10 implies that selling a marginal amount more of the fuel with a fixed extraction cost a during the earlier period t' in the BS scenario is relatively less profitable than selling it at t'', compared to the BAU scenario. However, since the profit maximising fuel owner(s) do sell fuel with marginal extraction cost a only at t' in the BS scenario, and only at t'' in the BAU scenario, we know that marginal profits for marginal sales of that resource are highest at t' in the BS scenario and at t'' in the BAU scenario, a contradiction.

Formally, this concluding step of the proof can be seen as follows:

Since in the BS case the fuel owner chose to extract fuels that have a marginal cost a at time t' and at no time else, we know that present discounted net sales revenues from marginal sales of fuel extracted at cost a is highest at time t',

$$\left(P_{t',BS} - a\right)e^{-\rho t'} \geq \left(P_{t'',BS} - a\right)e^{-\rho t''}, \tag{A.11}$$

where P_t stands for the marginal (gross) sales revenue for the fuel owner(s) at time t when they marginally vary their sales around the rate of their choice, as defined specifically for monopolistic and for competitive fuel owners in Section 2.

The similar reasoning shows that for the BAU scenario the inverse holds,

$$\left(P_{t',BAU} - a\right)e^{-\rho t'} \leq \left(P_{t'',BAU} - a\right)e^{-\rho t''}. \tag{A.12}$$

Given decreasing marginal sales revenues (Property 5 from Section 2), Eq. (A.10) does, however, imply, for the gross sales revenues

$$P_{t',BS} < P_{t',BAU} \wedge P_{t'',BS} \geq P_{t',BAU}. \tag{A.13}$$

Eqs. (A.12) and (A.11) are, as an ensemble, however, in direct contradiction with Eqs. (A.13). $\blacksquare$

(D) Lemma 4

Lemma 4. *For any two continuous and differentiable functions $G(t)$ and $F(t)$ and their finite derivatives $g(t)$ and $f(t)$, and any $T > 0$ and $\rho > 0$,*

$$\left.\begin{array}{c} G(0) = F(0) \\ G(t) > F(t) \underset{t\epsilon(0,T)}{\forall} \end{array}\right\} \quad \Rightarrow \quad \int_0^T e^{-\rho t}\left[g(t) - f(t)\right]\mathrm{d}t > 0 \ .$$

Proof. Define $H(t) \equiv G(t) - F(t)$ and $h(t) \equiv g(t) - f(t)$. We thus have $H(0) = 0$ and $H(t) > 0 \underset{t\epsilon(0,T)}{\forall}$. Use further $\eta \equiv \int_0^T e^{-\rho t}h(t)\mathrm{d}t$. Then, define h^m as the path that minimises the discounted integral while respecting the imposed condition:

$$\min_{h^m} \eta$$
$$s.t. \quad \int_0^t h^m(s)\mathrm{d}s > 0 \underset{t\epsilon(0,T)}{\forall} \tag{A.14}$$

By the following reasoning h^m cannot contain any periods with negative values:

- If $h^m(t)$ were to contain any negative values that were not preceded (in terms of lower values of t) by some positive values, the condition in Eq. (A.14) would be violated: the integral over consequentially negative values with at least some of them being strictly negative is necessarily negative.

- If $h^m(t)$ were to contain some strictly negative values that are preceded only by positive values, then simultaneously reducing some of the preceding positive values and increasing some of the mentioned negative values will on one hand, leave unaffected the condition (A.14) and on the other hand reduce the value of η, as the reduction of the earlier-occurring positive values is discounted less than the increase in the later-occurring negative values, leaving a net reduction in η and therefore contradicting that the initial h^m minimised η.

As the path $h^m(t)$ can thus not contain any negative values, and as in order for $H(t)$ to take on *strictly* positive values on the integral $(0, T)$, it is clear that η must be positive, as it is an integral of weighted positive values among which some are strictly positive, and where all weights $e^{-\rho t}$ are strictly positive. Thus, $\int_0^T e^{-\rho t} \left[g(t) - f(t) \right] \mathrm{d}t > 0$. $\blacksquare$

(E) Proof of Proposition 1

Assume an exogenously given, fix λ_T.

From Eq. (5), we know for the monopolistic supplier

$$\lambda_t = \lambda_T e^{\rho(t-T)} + \int_{s=t}^{T} e^{(t-s)\rho} \dot{c}_s \mathrm{d}s. \tag{A.15}$$

It will be intuitive that our analysis holds for the competitive case as well. Inserting Eq. (A.15) in Eq. (3) yields

$$p_t(r_t) + r_t p_t'(r_t) \;\; = \;\; c_t + \tau_t + \lambda_T e^{\rho(t-T)} + \int_{s=t}^{T} e^{(t-s)\rho} \dot{c}_s \mathrm{d}s. \tag{A.16}$$

In the following, we are going to prove by contradiction that the tax necessarily reduces cumulative extractions up to T.

Suppose thus hypothetically that the contrary would be the case, i.e., that

$$A_{T,\mathrm{tax}} \geq A_{T,\mathrm{no}}, \tag{A.17}$$

in which we introduced the indexes *tax* and *no* to designate the variable, here A_T, in the case with and without the tax respectively.

Eq. (A.17) implies

$$c_{T,\mathrm{tax}} \geq c_{T,\mathrm{no}}.$$

We have $\lim_{t \to T} \int_{s=t}^{T} e^{(t-s)\rho} \dot{c}_s ds = 0$ and, from Eq. (A.17), $\lim_{t \to T} c_{T,\text{tax}} \geq \lim_{t \to T} c_{T,\text{no}}$. Therefore, the RHS of Eq. (A.16) is strictly larger in the tax case (note that $\lim_{t \to T} \lambda_t = \lambda_T$ in both, the tax as well as the no-tax case), and thus Property 5 (see Section 2) implies

$$\lim_{t \to T} r_{t,\text{tax}} < \lim_{t \to T} r_{t,\text{no}}. \tag{A.18}$$

Because all our variables evolve smoothly over time Eqs. (A.17) and (A.18) imply that there is a t^* that meets the definition that *the two variants' extraction rates equal each other for the last time* in the pre-T period, i.e., such that

$$r_{t^*,\text{tax}} = r_{t^*,\text{no}}, \tag{A.19}$$

and

$$r_{t,\text{tax}} < r_{t,\text{no}} \, \forall_{t^* < t \leq T}. \tag{A.20}$$

Relation Eq. (A.20) implies that the difference $A_{t,\text{tax}} - A_{t,\text{no}}$ is strictly decreasing during the time between t^* and T, which, considering Eq. (A.17) can only hold if

$$c_{t,\text{tax}} > c_{t,\text{no}} \, \forall_{t^* \leq t < T}. \tag{A.21}$$

Eqs. (A.19) and (A.16), as well as the fact that $\tau_t \geq 0$ imply

$$c_{t^*,\text{tax}} + \int_{t=t^*}^{T} e^{\rho(t^*-t)} \dot{c}_{t,\text{tax}} dt \leq c_{t^*,\text{no}} + \int_{t=t^*}^{T} e^{\rho(t^*-t)} \dot{c}_{t,\text{no}} dt,$$

and thus

$$\int_{t=t^*}^{T} e^{\rho(t^*-t)} (\dot{c}_{t,\text{no}} - \dot{c}_{t,\text{tax}}) dt \geq c_{t^*,\text{tax}} - c_{t^*,\text{no}}. \tag{A.22}$$

As according to Eq. (A.21) the RHS of Eq. (A.22) is strictly positive, it is easy to see that Lemma 4 (Part D of the Annex) implies that Eqs. (A.21) and (A.22) cannot be reconciled, which concludes our proof by contradiction. ∎

(F) Proof of Lemma 3

First, note that from Property 5 (see Section 2) we know that, for a fixed tax, an increase in the value of the RHS of the first order condition yields a decrease of the momentary extraction rate.

Suppose that the value of the LHS expression in the FOC decreases when r_t is fixed and the tax τ_t is increased from zero to a positive value. Consider further a no-tax case, where the RHS has an initial value, called RHS_0, yielding an initial extraction rate $r_{t,0}$ at which the RHS and the LHS of the FOC are equalised. As we suppose, adding a tax

τ_t decreases the value on the LHS of the FOC when $r_{t,0}$ is hypothetically held constant in a first step. We thus would need a lower hypothetical RHS-value, RHS_1, in order for the FOC to be equalised in the new situation with the tax. Now, the RHS-value is, however, given and will not really be reduced to RHS_1 but will remain at RHS_0. In order to see what this implies for the instantaneous extraction rate, we then consider in a second step a hypothetical re-increase of the RHS-value from RHS_1 to RHS_0. Along with this hypothetical re-increase of the RHS, we, however, will have to decrease the instantaneous extraction rate for both sides of the FOC to remain equalised. This result shows that *if* adding an instantaneous tax τ_t decreases the LHS-value of the FOC, then the extraction rate at that time will have to decrease, given that the value on the RHS remains unchanged. We now show that the tax τ_t will indeed decrease the LHS-value at time t, which therefore implies that it will decrease the extraction rate r_t. This demonstration will conclude our proof. Note that showing this property is not as obvious as it may seem at first sight, as adding a tax in Region 1 and leaving *worldwide* demand unchanged does not simply mean to decrease a demand, but to, eventually, simultaneously decrease demand in Region 1 and increase the demanded quantity in Region 2.

While $p_R(r_t, \tau_t)$ unambiguously decreases with an increasing tax for a given r_t, this cannot be claimed to necessarily be the case for the second term of the LHS of the corresponding FOC, $\frac{\partial p_R(r_t, \tau_t)}{\partial r}$, without any further assumptions about the demand function. Here, we show that the reduction of p_R induced by a tax, i.e., $- [p_R(r_t, \tau_t) - p_R(r_t, 0)]$, unambiguously dominates the potential increase of the second term, i.e., $\left[r_t \frac{\partial p_R(r_t, \tau_t)}{\partial r} - r_t \frac{\partial p_R(r_t, 0)}{\partial r} \right]$, at least for not too large tax levels and demand curves with finite derivatives, wherewith the direct effect of the tax at time t unambiguously reduces the extraction rate in the current period, r_t.

Be $r(p)$, the worldwide demand curve for the resource, a continuous, strictly decreasing function with a third derivative that is finite for any $p > 0$. The worldwide demand is split into the regional demands r_1 and r_2, such that for a worldwide equal price, demand in Region 2 corresponds to x times the demand in Region 1:

$$
\begin{aligned}
r_1 + r_2 &= r \\
r_2(p) &= x \cdot r_1(p)
\end{aligned}
\tag{A.23}
$$

Eq. (A.23) implies that all derivatives of the regional demand function differ by a factor x as well:

$$
r_2^{(i)}(p) = x \cdot r_1^{(i)}(p),
\tag{A.24}
$$

where the indice $(\cdot)^{(i)}$ denotes the ith derivative.

When Region 1 introduces a tax, the *consumer price* for the resource in that region, p_1,

exceeds the *consumer price* in the tax free Region 2, p_2, as well as the *sales price* for the resource owners, p_R, by the tax rate τ:

$$p_1 = p_2 + \tau = p_R + \tau$$

The aggregate demand for a given sales price and a specific tax rate is

$$r(p_R, \tau) = r_1(p_1) + r_2(p_2) = r_1(p_R + \tau) + r_2(p_R). \tag{A.25}$$

As the demand curves in the two regions are continuous and strictly decreasing, Eq. (A.25) directly implies that the worldwide demand is strictly decreasing as well in p_R as in τ. It is therefore clear that the inverse demand curve, here the sales price that for a given tax yields a specific aggregate demand, $p_R(r, \tau)$, is strictly decreasing in r.

In the following, we will use the syntax Δvar in order to express the discrete change of the value of the variable 'var' resulting from the introduction of the tax:

$$\Delta \text{var} \equiv \text{var}_{\text{tax}} - \text{var}_{\text{no}},$$

where the indexes *tax* and *no* stand for the situation with and without the tax. Consider the hypothetical case in which a consumer tax is introduced in Region 1 and the sales price demanded by the resource owners is adapted accordingly in such a way that, overall, the introduction of the tax implies an unchanged global consumption. In this case, demand in Region 1 would have to decline by exactly the same amount as the demand in Region 2 would increase, and the corresponding changes in the regions' sales, denoted Δr, would have to have exactly the size that implies that the price difference between the two regions amounts to the level of the tax,

$$\Delta p_1 + \Delta p_2 = \tau. \tag{A.26}$$
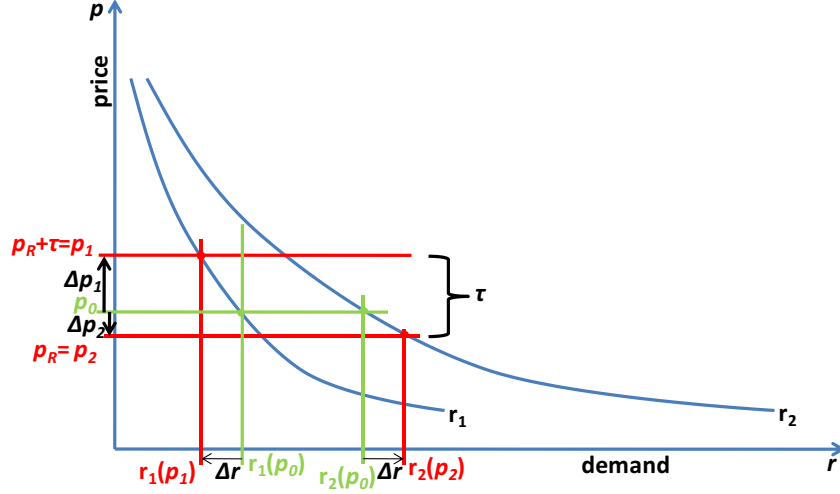
Consider the illustration in Fig. 6.

Figure 6: Hypothetical situation of regional tax which is neutral for global emissions

In a first approximation, we have:

$$\Delta r \approx \Delta p_1 \cdot r'_1(p_0) \tag{A.27}$$

$$\Delta r \approx \Delta p_2 \cdot r'_2(p_0) \tag{A.28}$$

With the inclusion of Eq. (A.24), using Eqs. (A.27) and (A.28) in Eq. (A.26) implies

$$\Delta p_2 \approx \frac{\tau}{1+x}, \tag{A.29}$$

yielding

$$p_b \approx p_0 - \frac{\tau}{1+x}. \tag{A.30}$$

Eq. (A.30) expresses that, in order to keep aggregate demand constant, the sales price for the resource owner must decease by a value that is approximately proportional to the tax rate.

From Eqs. (A.29) and (A.26) follows

$$\Delta p_1 \approx x \cdot \Delta p_2. \tag{A.31}$$

In order to be able to make a statement about the corresponding change of the global

demand, $\frac{\partial p_R(r,\tau)}{\partial r}$, we again develop two Taylor approximations:

$$r_1'(p_1) \approx r_1'(p_0) + \Delta p_1 \cdot r_1''(p_0) + \frac{(\Delta p_1)^2}{2} r_1'''(p_0) \qquad (A.32)$$

$$r_2'(p_2) \approx r_2'(p_0) - \Delta p_2 \cdot r_2''(p_0) + \frac{(\Delta p_2)^2}{2} r_2'''(p_0) \qquad (A.33)$$

$$\approx x \cdot r_1'(p_0) - \Delta p_1 \cdot r_1''(p_0) + \frac{(\Delta p_1)^2}{2x} r_1'''(p_0), \qquad (A.34)$$

where the minus sign for the second term on the right hand side in Eq. (A.33) is due to the fact that Δp_2 is defined in absolute terms, and where Eq. (A.34) follows from Eq. (A.33) using Eq. (A.24) as well as Eq. (A.31).

As $\frac{\partial r}{\partial p} = \frac{\partial r_1}{\partial p} + \frac{\partial r_2}{\partial p}$, relying on continuity of all relevant functions we know that $\frac{\partial p}{\partial r} = \left[\frac{\partial r_1}{\partial p} + \frac{\partial r_2}{\partial p} \right]^{-1}$ and we can therefore write

$$\Delta \left[\frac{\partial p_R}{\partial r} \right] = \frac{1}{r_1'(p_R + \tau) + r_2'(p_R)} - \frac{1}{r_1'(p_0) + r_2'(p_0)}.$$

By using Eqs. (A.32) and (A.34), as well as Eq. (A.24), we can thus approximate this response of the first derivative of the selling-price, $\frac{\partial p_R}{\partial r}$, to the introduction of the tax as

$$\Delta \left[\frac{\partial p_R}{\partial r} \right] \approx \frac{1}{(1 + x)r_1'(p_0) + (\Delta p_1)^2 \left(1 + \frac{1}{x} \right) r_1'''(p_0)/2} - \frac{1}{(1 + x)r_1'(p_0)}.$$

For relatively small $(\Delta p_1)^2$ this approximates to

$$\Delta \left[\frac{\partial p_R}{\partial r} \right] \approx \text{-} (\Delta p_1)^2 \frac{r_1'''(p_0)}{2x(1 + x)r_1'(p_0)^2},$$

which is proportional to the square of the tax induced price change.

The response of the *seller price that leaves global demand unchanged* to the introduction of the tax, $\Delta p = p_R - p_0$, can be approximated using Eq. (A.30):

$$\Delta p \approx \text{-} \frac{\tau}{1 + x}$$

Using Eqs. (A.29) and (A.31) we therefore have the following ratio between the direct effect of the tax on the *seller price which leaves global demand unchanged* and the corresponding change of the price's derivative with respect to r:

$$\frac{\Delta \left[\frac{\partial p_R}{\partial r} \right]}{\Delta p} \approx \frac{\tau \cdot x \cdot r_1'''(p_0)}{2 (1 + x)^2 r_1'(p_0)^2}, \qquad (A.35)$$

whose sign depends on the not specified sign of $r'''(p_0)$.

As the ratio in Eq. (A.35) is proportional to the tax rate, and the factor by which this tax rate is multiplied cannot be infinite due to the boundedness of our derivatives of the demand function, we thus know that Δp is larger in absolute terms than any finite multiple of $\Delta \left[\frac{\partial p_R}{\partial r} \right]$ for taxes that are not too large, which proves our claim. ■

(G) Proof of Proposition 4

Having a constant periodic probability (ψ) that a backstop technology may emerge, we know from Dasgupta and Heal (1974) that the resource owners' maximisation problem differs from the deterministic case without backstop solely by a corresponding increase of the discount factor. The first order conditions governing the fuel owners' behavior can thus be written as

$$P_t(r_t) = c(A_t) + \tau e^{\theta t} + \lambda_t \tag{A.36}$$
$$\dot{\lambda}_t = \lambda_t(\rho + \psi) - \dot{c}_t.$$

Defining $\delta \equiv (\rho + \psi)$, we get

$$\lambda_t = \lambda_T e^{\delta(t-T)} + \int_{s=t}^{T} e^{\delta(t-s)} \dot{c}_s \mathrm{d}s. \tag{A.37}$$

We are considering an exponentially increasing tax, $\tau_t = \tau_0 e^{\theta t}$, where θ may exceed ρ, as long as $\theta < \delta$.

We use the same syntax as in the proof for Lemma 3: $\Delta \mathrm{var} \equiv \mathrm{var}_{\mathrm{tax}} - \mathrm{var}_{\mathrm{no}}$, where var can be a single variable or a combined mathematical term.

Note that, as the no-tax case corresponds to simply setting $\tau_{t,\mathrm{no}} = 0 \underset{t \geq 0}{\forall}$ and in the tax case we have $\tau_{t,\mathrm{tax}} > 0 \underset{t \geq 0}{\forall}$, we know that $\Delta \tau_t > 0 \underset{t \geq 0}{\forall}$.

In the next step we are going to show by contradiction that the described tax path cannot lead to increased cumulative emissions for any point in time:

Assume thus, hypothetically, that the contrary holds, i.e., $\Delta A_t > 0$ for some t.

We treat the two possible subcases separately:

-*Subcase 1:* Suppose, $\exists t_0$ s.t.

$$\Delta A_{t_0} = 0 \tag{A.38}$$
$$\text{and} \quad \Delta A_t > 0 \underset{t_0 < t < \infty}{\forall}.$$

This requires $\Delta r_{t_0} \geq 0$, and therefore, due to Eq. (A.36) and Property 5, that $\Delta[\lambda_{t_0} + \tau_{t_0}] \leq 0$, wherewith we have

$$\Delta \lambda_{t_0} < 0.$$

37

From Eq. (A.37) (and the transversality condition), however, we know that $\lambda_{t_0} = \int_{t=t_0}^{\infty} \mathrm{e}^{\delta(t_0-t)} \dot{c}_t \mathrm{d}t$, which can be rewritten as

$$\lambda_{t_0} = \mathrm{e}^{\delta t_0} \int_{t=t_0}^{\infty} \mathrm{e}^{-\delta t} \dot{c}_t \mathrm{d}t. \tag{A.39}$$

It is, however, clear that Eqs. (A.38) through (A.39) are not reconcilable with Lemma 4 (Part D of the Annex). Thus, it is shown by contradiction that subcase 1 is impossible. ∎*Subcase 1.*

-*Subcase 2:* Suppose $\exists t_1, t_2, t_1 < t_2$, s.t.

$$\Delta A_{t_1} = 0 \wedge \Delta r_{t_1} \geq 0, \tag{A.40}$$

$$\Delta A_{t_2} = 0 \wedge \Delta r_{t_2} \leq 0, \tag{A.41}$$

$$\text{and} \quad \Delta A_t \geq 0 \underset{t_1 < t < t_2}{\forall}. \tag{A.42}$$

Eqs. (A.40) and (A.41) imply

$$\Delta c_{t_1} = 0 \wedge \Delta c_{t_2} = 0, \tag{A.43}$$

and therewith also

$$\Delta[\lambda_{t_1} + \tau_{t_1}] \leq 0, \tag{A.44}$$

$$\text{and} \quad \Delta[\lambda_{t_2} + \tau_{t_2}] \geq 0.$$

Eq. (A.42) indicates that

$$\Delta c_t > 0 \underset{t_1 < t < t_2}{\forall}. \tag{A.45}$$

From Eq. (A.37), we know

$$\lambda_{t_1} = \lambda_{t_2} \mathrm{e}^{\delta(t_1-t_2)} + \int_{t_1}^{t_2} \mathrm{e}^{\delta(t_1-t)} \dot{c}_t \mathrm{d}t.$$

Defining $\mu_t \equiv \lambda_t \mathrm{e}^{-\theta(t-t_1)}$, which yields $\mu_{t_1} = \lambda_{t_1}$ and $\lambda_{t_2} = \mu_{t_2} \mathrm{e}^{\theta(t_2-t_1)}$, we can write

$$\mu_0 = \lambda_0 = \mu_{t_2} \mathrm{e}^{[\delta-\theta](t_1-t_2)} + \int_{t_1}^{t_2} \mathrm{e}^{\delta(t_1-t)} \dot{c}_t \mathrm{d}t.$$

Consider

$$\Delta[\lambda_{t_2} + \tau_{t_2}] \geq 0 \Rightarrow \mathrm{e}^{-\theta(t_2-t_1)} \Delta[\lambda_{t_2} + \tau_{t_2}] \geq 0 \Rightarrow \Delta[\mu_{t_2} + \tau_{t_1}] \geq 0. \tag{A.46}$$

As $\Delta\tau_{t_1} > 0$ the last expression in Eq. (A.46) implies

$$\Delta[a\mu_{t_2} + \tau_{t_1}] > 0 \underset{0 \leq a < 1}{\forall}. \tag{A.47}$$

38

From Eq. (A.44) we know $\Delta[\mu_{t_1} + \tau_{t_1}] \leq 0$, which we can rewrite as

$$\Delta[\mu_{t_2}\underbrace{e^{[\delta-\theta](t_1-t_2)}}_{<1} + \int_{t_1}^{t_2} e^{\delta(t_1-t)}\dot{c}_t dt + \tau_{t_1}] \leq 0. \tag{A.48}$$

Eqs. (A.47) and (A.48) imply

$$\Delta[\int_{t_1}^{t_2} e^{\delta(t_1-t)}\dot{c}_t dt] \leq 0. \tag{A.49}$$

Eqs. (A.49), (A.43) and (A.45), however, violate Lemma 4 (Part D of the Annex), a contradiction. ∎*Subcase 2.*

If the tax were to increase cumulative emissions for some period, either subcase 1 or subcase 2 would have to hold; we have $A_0 = 0$ in any case, and for any t^* where $\Delta A_{t^*} > 0$ there must be a latest preceding period, $\underline{t}$, $\underline{t} < t^*$, for which the tax does not impact the cumulative emissions, $\Delta A_{\underline{t}} = 0$ ($\underline{t}$ may be 0). Then, there are two possibilities: either the tax will increase cumulative emissions for all periods after time $\underline{t}$ – this is subcase 1 –, or there is a future period for which the cumulative emissions are not affected by the tax – this is subcase 2. Therefore, the demonstrated inconsistency of both subcases proves that the considered taxes cannot increase the cumulative emissions, A_t, for any period t.

In addition, it is impossible that the tax would not change emissions in any period: if this were the case, then λ_t would be unchanged as well, but in this case, the tax τ_t would affect the extraction rate r_t in Eq. (A.36). Thus, the considered tax necessarily reduces emissions, at least in some periods.

We conclude that the considered tax (i) does not increase any period's cumulative emissions, (ii) reduces cumulative emissions at least for some periods and (iii) thus unambiguously reduces the expectancy of the cumulative emissions, QED. ∎

(H) Lastingness of medium-term emission reductions

The following reflections suggest that a tax-induced fuel saving up to time T, as identified in sections 4 and 5 in presence of time-T regime switch, implies a sustained strict reduction of cumulative emissions for a substantial time beyond T, potentially perpetually. First, as the net demand for fossil fuels is finite even without the tax or any additional carbon-limiting measure, it will require an non-marginal period of time until the amount of emissions saved up to time T could be offset by increases in the post-T era. Second, for some stylised scenarios of different conceivable alternative measures, cumulative emissions may (i) fully converge to the BAU emissions only after the time when all fuels would have

been used in the BAU scenario without pre-T tax, or they may (ii) not converge at all after T, or they may (iii) converge only partially overall. Cases (ii) and (iii) imply eternal overall emission savings. Case (i) is the natural outcome in a framework with a global cap-and-trade scheme after time T with exogenous and constant allowances, or if we had a backstop from time T on, supplied infinitely elastically at a fixed price b below the choke price of a demand which we assume to be constant over time. In this case, it is easy to see that it would take a strictly positive time Δ after T until cumulative emissions in the case with the tax correspond to the emissions until T in the no-tax baseline scenario. As this would naturally mean that the emissions path after time $T + \Delta$ in the pre-T-tax case corresponds to the emissions path after time T with a constant shift of Δ periods, it is clear that cumulative emissions would converge only once all fuel extraction has stopped in the tax scenario (which would in this case be Δ periods later than when extraction would have stopped in the no-tax baseline). Given that marginal damages are lower for lower atmospheric carbon stocks and that the social value of extraction is higher if extraction costs are lower, one may expect the emission allowances in a global cap-and-trade scheme to be larger in the case with the pre-T tax. If the emission allowances are adapted dynamically according to the social value of emissions as a function of cumulative emissions and extractions, it is still the case that, in the simplest world, it would require a non-marginal time Δ after T until cumulative emissions in the case with the tax reach the level of emissions up to T in the no-tax baseline, and that afterwards the emission (and allowance) paths would be the same, shifted by Δ in time, still leading to strictly lower cumulative emissions for each period of time from T up to beyond the baseline duration of extractions. In a more dynamic world with technological progress that could reduce the social value of emissions (for a specific amount of cumulative emissions and extractions) over time, one could even expect the tax-induced delay of fuel consumption from the pre-T phase to result in total emission reductions for all times after T (case iii). Finally, if the future measure is a globally enforced carbon capture and storage mechanism, post-T emissions would not have to be larger in the case with the pre-T tax than in the baseline case, implying that no convergence of emissions after T takes place (case ii).