Suboptimal Climate Policies – Green Paradox, Leakage and the Law of Small Abatements

DISSERTATION
of the University of St.Gallen
School of Management,
Economics, Law, Social Sciences
and International Affairs
to obtain the title of
Doctor of Philosophy in Economics and Finance

submitted by

Florian Habermacher

from

Malters (Luzern)

Approved on the application of

Prof. Dr. Dr. h.c. Gebhard Kirchgässner

and

Prof. Rick van der Ploeg, PhD

Dissertation no. 4219

Gutenberg AG, Schaan, 2014

The University of St. Gallen, School of Management, Economics, Law, Social Sciences and International Affairs hereby consents to the printing of the present dissertation, without hereby expressing any opinion on the views herein expressed.

St. Gallen, September 11, 2013

The President:

Prof. Dr. Thomas Bieger

Acknowledgements

I consider myself extremely lucky to have been able to write this dissertation. That I had been given the opportunity to receive substantial coursework education in the field of economics for more than two years and to write on the topics of my choice in environmental and resource economics for further two years, has certainly been more than I could have expected, given that I had brought with me mostly an education in engineering. That this has been possible and that it became an extremely rewarding experience academically and intellectually, but also socially, is the merit of many, notably the following.

Without Prof. Dr. Dr. h.c. Gebhard Kirchgässner, I might have never received the opportunity to write a dissertation in the field of my choice. Furthermore, as my employer, he has allowed me to almost entirely focus on my own education and on questions that I wanted to study. When I had questions or concerns, he always took the time to discuss the issues and to help me. He was also personally a very kind person, and I am overall exteremely thankful to him for everything he has done for me.

I am also very thankful to Prof. Dr. Rick van der Ploeg who agreed to host me during a year of academic research at the University of Oxford. Whilst the University of St. Gallen had been an academically very stimulating environment in general, being integrated into the Oxford Centre for the Analysis of Research Rich Economies has been very inspiring for my own research. I would like to thank Prof. van der Ploeg also for his helpful academic advice and for his personal concern.

Prof. Dr. Mark Schelker helped me often with good academic advice and I specially thank him and Prof. Dr. Simon Evenett, Prof. Dr. Reto Föllmi, Prof. Dr. Roland Hodler, Prof. Dr. Elena Irwin, Prof. Dr. Martin Kolmar, Prof. Dr. Marc Parlange, Prof. Dr. Ted Russell and Prof. Dr. Philippe Thalmann for having given me support when I needed it.

Writing the PhD wouldn't have been as rewarding without the intellectual stimulus, good advices, and friendship of my doctoral colleagues, among which I would like to thank particularly Thomas Davoine, Dario Fauceglia, Johannes Fritz, Berit Gerritzen, Teresa Körner, Fabian Schnell, Andreas Steinmayr and Martin Wermelinger. Similarly, Michèle Klarer and Gabriela Schmid, and all other members of the Swiss Institute for International Economics and Applied Economic Research, have been very helpful and pleasant company.

For having helped to keep my moral and physical health intact during different phases of those four years, I thank particularly Zoé Dumas, and Martin Brandenberger, Amir

Melzer, Devesh Rustagi and Bernhard Steubing, as well as my family, David, Franziska, Jonathan and Peter Habermacher, and, for their support during the year here on the other side of the Channel, notably Benedikte Bjerge, Simón Escoffier, Lara Fleischer, Rafa José Gude and Rachel Pascoe.

I hope that my future work will return something to the society whose taxes have funded me generously during so many years, and my thanks belong also to the developers of software which has allowed me to write this thesis very comfortably, notably LyX.

Oxford, March 2013
Florian Habermacher

Contents

List of Figures

List of Tables

Zusammenfassung

Der Klimawandel wird zunehmend als eine der grossen Bedrohungen für unsere Gesellschaft gesehen. Klimaforscher sind sich weitestgehend einig, dass die anthropogenen Treibhausgasemissionen – allen voran Kohlenstoffdioxid (CO_2) aus der Verbrennung fossiler Brennstoffe – die Atmosphäre um mehrere Grad Celsius zu erwärmen drohen, mit potenziell drastischen Auswirkungen auf die Lebensgrundlagen in vielen dicht bevölkerten Gebieten. Dies kann nur mit einer signifikanten Reduktion der Treibhausgasemissionen aus der Nutzung fossiler Brennstoffe während den kommenden Jahrzehnten abgewendet werden. Ohne wirkungsvolle politische Massnahmen dürften diese Emissionen in den nächsten Jahren stattdessen aber stark zunehmen.

Es ist ausserordentlich schwierig, die globalen Treibhausgasemissionen substanziell zu reduzieren. Im Gegensatz zu vielen anderen Schadstoffen gibt es keine einfache technologische Lösung; es ist keine ausgereifte Technik vorhanden, die es uns erlauben würde, die fossile Energie CO_2-frei zu nutzen, wenigstens nicht billig und in der benötigten Menge. Die Handlungsbereitschaft für Emissionsreduktionen wird stark dadurch beeinträchtigt, dass ein Land oder eine Region nur zu einem geringen Anteil von dem verbesserten Klima profitiert, wenn es seine eigenen Emissionen reduziert. Sogar wenn ein Land wie die Schweiz ab sofort gar keine Emissionen ausstossen würde, wäre es nicht möglich, den entsprechenden Effekt davon in zukünftigen Temperaturmessungen zu identifizieren, da die Temperatur grundsätzlich von den weltweiten Emissionen bestimmt wird. Diese sind ungleich höher als die heutigen inländischen Emissionen. Tendenziell leben die Emittenten und die Hauptleidtragenden örtlich und zeitlich weit voneinander entfernt. Die heutigen Emissionen tangieren hauptsächlich ungeborene Generationen, welche nicht auf ihr Recht auf ein geschütztes Klima pochen können. Die reichen Länder in der nördlichen Hemisphäre haben den grössten pro-Kopf Ausstoss, während die voraussichtlich am stärksten betroffenen die Armen in den südlichen Breitengraden sind. Dies erschwert die Aushandlung strenger globaler Klimaregulierungen erheblich, wie das Schicksal des Kyoto-Protokolls eindrücklich gezeigt hat. Obwohl von 190 Ländern unterzeichnet, hat es seit seiner Verabschiedung 1997 kaum zu substanziellen Emissionsreduktionen geführt.

Trotzdem sind zahlreiche Länder, Regionen, Unternehmen und Individuen bestrebt, zur Lösung des Problems beizutragen; entweder im Alleingang oder zusammen mit anderen, auch wenn nicht immer klar ist, zu welchem Grad diese Beiträge die globalen Emissionen beeinflussen, und obwohl global koordinierte und optimierte Beiträge das Kosten-Nutzen-Verhältnis stark verbessern könnten.

Das Fehlen optimal ausgestalteter Klimamassnahmen, sprich, die Grenzen politisch re-

alisierbarer Massnahmen, sind der Ausgangspunkt für alle vier Aufsätze dieser Dissertation. Die spezifischen Themengebiete mit welchen sich die einzelnen Aufsätze beschäftigen sind in der bestehenden (umwelt-)ökonomischen Literatur alle schon tiefergehend behandelt worden. Das Ziel der vorliegenden Arbeit ist es, die Kluft zwischen den bestehenden, oft relativ abstrakten ökonomischen Theorien und der wahren Welt zu verringern.

Die ersten zwei Aufsätze beschäftigen sich mit dem Grünen Paradoxon. Der Begriff wurde 2008 von H.-W. Sinn bekannt gemacht. Die damit verbundene Theorie besagt, dass eine politisch realistische Klimasteuer, welche mit einem tiefen Anfangssatz eingeführt wird, aber über die Zeit schnell ansteigen könnte, vorausschauende Ressourcenbesitzer dazu veranlassen könnte, ihre fossilen Brennstoffe noch schneller zu verkaufen. Dies würde den Klimawandel verschärfen statt abschwächen. Aufsatz 1 zeigt, dass dies unwahrscheinlicher ist, wenn in Betracht gezogen wird, dass eine anvisierte Klimasteuer nicht die einzige mögliche Klimamassnahme ist, und dass die Ressourcenbesitzer auch bei Nichteinführung der Steuer die Möglichkeit zukünftiger Massnahmen in ihr Kalkül miteinbeziehen. Aufsatz 2 benutzt die theoretischen Grundlagen um den Einfluss besagter Klimasteuern auf die CO_2-Emissionen in einem kalibrierten, dynamischen Modell numerisch zu simulieren. Insbesondere wenn die riesigen Vorkommen an (sehr emissionsintensiver) Kohle in die Berechnungen miteinbezogen werden, zeigt sich, dass die Steuer in erster Linie die Gesamtmenge mittel- und langfristiger Emissionen reduziert. Die Steuer wirkt sich – trotz einer möglichen leichten Erhöhung der kurzfristig auftretenden Emissionen – insgesamt vor allem positiv auf das Klima aus. Zudem zeigt eine Wohlfahrtsanalyse, dass die im Rahmen des Grünen Paradoxon hervorgehobene Reaktion der Ressourcenbesitzer auf die angekündigte Steuer den Wohlfahrtsgewinn für die brennstoffkonsumierenden Regionen sogar erhöhen könnte.

Aufsatz 3 geht der Frage nach, inwieweit allfällige CO_2-Emissionsreduktionen in einem Teil der Welt durch induzierte Steigerungen der Emissionen in der restlichen Welt wettgemacht werden (sogenanntes Carbon Leakage). Es zeigt sich, dass aufgrund der Erschöpfbarkeit der fossilen Brennstoffe zu befürchten ist, dass ein sehr grosser Teil der regionalen Emissionsreduktionen durch Emissionserhöhungen in anderen Regionen wettgemacht wird. Inwiefern dies tatsächlich zutrifft, hängt aber stark von zukünftigen technologischen, wirtschaftlichen und politischen Entwicklungen ab. Wie *relevant* das entsprechende Carbon Leakage ist, wiederum hängt stark davon ab, wie die zukünftigen Emissionen mit heutigen verrechnet werden, sprich von der Zeitdiskontrate. Diesen Faktoren wurde in der bestehenden Literatur zur Carbon Leakage-Frage nur ungenügend Rechnung getragen.

Zur marktnahen Regulierung von Treibhausgasemissionen stehen neben Emissionssteuern politisch vor allem auch mengenbasierte Regulierungen, sprich direkte Vorgaben der Emissionsmenge, wie z.B. beim Kyoto-Mechanismus, zur Debatte. Aufsatz 4 zeigt, dass sich die Unsicherheit über die Kosten und den Nutzen entsprechender Massnahmen bei regional beschränkten Regulierungen tendenziell so verhält, dass die Steuer gegenüber einer mengenbasierten Regulierung umso vorteilhafter wird, je kleiner die Region ist, welche die Klimamassnahme in Betracht zieht. Dieser Effekt ist so stark, dass sogar für relativ grosse Länder eine Steuer im einfachen ökonomischen Modell auf jeden Fall klar vorteilhafter sein dürfte als eine Mengensteuerung.

Introduction

Climate change is widely seen as a major threat to our society. There exists practically an unanimous consensus among climate scientists that anthropogenic emissions of greenhouse gases, mostly carbon dioxide (CO_2) stemming from the combustion of fossil fuels (coal, oil and gas), threaten to increase global average temperatures by several degrees, with potentially drastic effects on the livelihood of many densely populated regions in the world (IPCC, 2007). This can only be prevented if emissions from fossil fuel combustion are drastically reduced in the coming decades. These emissions are, however, projected to continue to rise at a rapid pace for a long time, unless stringent measures are taken to curb them.

The difficulty of reducing global greenhouse gas emissions is extraordinary. Contrary to many other pollutants, no simple technological fix exists for that problem; no cheap solution is available that would allow us to enjoy – on a large scale and at low costs – the fossil fuel energy without the undesirable CO_2 emissions. The willingness to act to reduce emissions is greatly reduced by the fact that a country or region reducing its emissions will itself enjoy only a small fraction of the climate benefits. For example, even if a country like Switzerland would immediately stop emitting greenhouse gases, it would never be possible to identify the effect of this change in the temperatures measured in Switzerland, as those are essentially determined by worldwide emissions which are incomparably higher than domestic ones. The emitters and those bearing the brunt of the climate change tend to be separated both in time and space. Current emissions will affect mostly generations which are not born yet and thus cannot insist on their right to an unaltered climate. The rich countries in the northern latitudes have the highest per capita emission rates, whilst those conceivably most vulnerable to climate change are the poor populations in the southern latitudes. Such relations tend to make the global coordination of stringent climate policy extremely difficult, as the fate of the Kyoto Protocol has demonstrated. Despite being signed by 190 countries, the protocol has, since his adoption in 1997, hardly produced any substantial emission reductions, and has lead to committed pledges to emission reductions mostly by European countries, responsible only for a limited (and decreasing) fraction of worldwide emissions. And even in that region, events such as the economic downturn in the latest years, and the general trend of the shift of industrial production to emerging economies, notably in Asia, seem of much larger relevance to the recent emission trend than the commitments related to the Kyoto Protocol. Nevertheless, numerous countries, regions, firms and individuals are undertaking non-negligible efforts to contribute to potential solutions to the problem; either individually or in coordination with others, even though it is not always clear to which extent these contributions will impact the global emissions, and even though globally coordinated and optimized efforts could improve the cost-benefit ratio of the efforts substantially.

The lack of optimally designed and perfectly informed climate policy, and the necessity to act under political limitations, is the starting point of the four essays in this dissertation. The specific topics these essays address have all received considerable attention in the climate economics literature. The aim of the essays is to narrow the gap between the often relatively abstract theories within which the topics were hitherto addressed, and the real world with its political, economic, and physical constraints.

Essays 1 and 2 address the Green Paradox. It has been brought forward that politically feasible carbon taxes would not simply be inefficient, but even counterproductive for the climate. As the taxes would, for political reasons, be introduced with a low initial level but rise rapidly over time, they would incite forward-looking fuel owners to speed up the extraction rather than to leave their fuels underground. While this could in theory indeed happen, Essay 1 explains that the possibility that other climate relevant developments, which may take place in future independently of whether a climate tax is introduced soon or not, can reverse or weaken the Green Paradox and tends to increase the climate benefits from the currently debated taxes. Essay 2 is based on a dynamic numerical model calibrated to the demand and the long-run supply of fossil fuels. Simulations of the effect of climate taxes suggest that the Green Paradox effects are, in general, hardly strong enough to make taxes undesirable for the climate. For the fuel owner anticipation effects to increase current extraction rates, the taxes would have to rise very rapidly over time. Consequently, these taxes would almost necessarily become so high after a few years that the induced longer-run emission reductions reduce medium and long-term climate damages very substantially. With these longer-run benefits, the net overall effect of the tax on the net present value of the climate damages is reduced, even for very high emission discount rates, despite potential increases of early emissions. Moreover, a welfare analysis shows that the reaction of the resource owners to the anticipated tax, emphasized in the theory of the Green Paradox, could increase rather than reduce the welfare benefits to fuel consuming countries.

Essay 3 addresses the issue of optimal climate policies when the policy, generally an emissions tax, is restricted regionally. It considers the reaction of the time-path of global emissions to current or future regional emission reductions. Because fossil fuels as the primary source of the emissions are scarce non-renewable resources, it appears that consumption reductions in one part of the world may partly be offset by consumption increases in the remainder of the world. In a stylized way, if one region reduces its imports of oil, it may take a bit longer until the reserves available in fuel exporting countries are used up by the importers in the remainder of the world, but the cumulative long-run consumption may not be reduced by very much. To which extent such a foreign offsetting of potential domestic emission reductions, called carbon leakage, undermines unilateral climate protection efforts, is subject to intense political and academic debates. The essay aims at deepening the understanding of the time-dimension of the problem, and points out a number of issues which have received scant attention in the existing carbon leakage literature. A main conclusion, derived again from a numerical calibrated fuel market model, is that the valuation of the response of global emissions to regional emission reductions varies very strongly with the strength delays in emissions are valued (the discount rate), and with how the fuel market is impacted in the future by technological (and also political) developments such as clean alternative technologies or the supply of synthetic oil from emission intensive coal liquefaction processes. When the dynamic fuel exhaustibility is taken into account, leakage effects may in the long-run be much stronger than the bulk of the literature, which often examined the question from a more static viewpoint, suggests. Moreover, the interplay between the more rapidly depleting oil and the more emission intensive and abundant coal can lead to counterintuitive effects with leakage rates (defined as the fraction of the regional emission reductions that are offset by emission increases abroad) above unity, or also below zero, depending on the details of the scenario. Furthermore, emission leakage rates may well differ in the order of magnitude or in their sign across fuels, warranting substantial differentiation of unilateral carbon tax rates across fuels, an issue which has barely been treated in the genuine leakage literature.[1]

Essay 4 is a shorter note on the optimality of quantity and price restrictions in climate change when abatement costs and benefits are uncertain. It explains that the seminal Weitzman (1974) result changes strongly in favor of a price regulation when climate policies of limited reach, such as a policy from a single country, are considered. It implies that for realistic climate policies, price regulations seem preferable even independently of the stock vs. flow pollutant question which had been addressed in the literature.

[1] Golombek et al. (1995), on which Essay 3 draws, provided a pioneering study on this topic, in a static framework. Michielsen (2011) considers the interplay between a scarce and an infinitely supplied resource, focusing mostly on the effect of a general carbon tax rather than one differentiated across fuels.

Essay 1

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

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 materialize. Moreover, the tax may allow the demanding countries to extract part of the resource rent, further increasing its desirability.

Authors: Florian Habermacher and Gebhard Kirchgässner

Keywords: climate change policy, greenhouse gas tax, Green Paradox, anticipation effects, exhaustible resources, fossil fuels market, backstop technology, uncertainty, resource rent.

JEL classification: Q54, Q31, Q38, Q41, Q42.

For helpful remarks and suggestions we mostly thank Robert King and Thomas Davoine, as well as the organizers and participants at the Venice Summer Institute Workshop on the Theory and Empirics of the Green Paradox, 20-21 July 2012.

1.1 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 feasibility. 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 (forthcoming) 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 optimizing 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 (2012a), 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 Green Paradox effect has originally been brought forward under the assumption of a fixed reservoir of stock, exploitable at limited costs (Sinclair, 1992; Sinn, 2008). We here take into account that the extraordinary diversity of fossil fuel deposits makes it unthinkable that the last dram of the physically existing fuels will ever be extractable economically, and that there exists a quasi-continuum of deposits in terms of per-unit extraction costs. Technically, this means that we thus assume a continuous extraction cost curve: the amount of resources exploitable for costs below a threshold price is strictly increasing in that price, at least for all prices for which some fuel demand exists.

The analyzes 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 materialize 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.[2] In a stylized 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

[2] 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 optimized or prohibitive tax paths, leaving aside the arbitrarily increasing taxes of the Green Paradox.

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 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.[3] 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 favorable 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.

[3] 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.

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' behavior 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 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 faster than the maximal rate which, according to the proposition of the Green Paradox, is compatible with a reduction in carbon emissions. We also briefly discuss the possible endogeneity of the future scheme switch. Section 7 provides a discussion and Section 8 concludes.

1.2 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 normalized 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 (cf. Herfindahl, 1967, as well as the next section where we derive this result).

The resource owners maximize 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' maximization problem can thus be written as

$$\begin{aligned} U_i \quad &= \quad \max_{r_{t,i}} \int_{t=\underline{t}}^{\bar{t}} \mathrm{e}^{-\rho t} r_{t,i} \cdot (p_t(r_t) - c(A_t) - \tau_t)\, \mathrm{d}t \\ &\text{s.t.} \quad \dot{A}_t = r_t \text{ and } A_0 = 0, \text{ i.e. } A_t = \int_{s=0}^{t} r_s \mathrm{d}s, \text{ and } r_t = \sum\nolimits_i r_{t,i}. \end{aligned} \tag{1.1}$$

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. (1.1), we thus have in the two considered variants of the model:

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

Taking this into account in the current-value Hamiltonian,

$$\mathcal{H} = r_t \cdot (p_t(r_t) - c(A_t) - \tau_t) - \lambda_t r_t, \tag{1.2}$$

we arrive at the following two first order conditions:

$$\frac{\partial \mathcal{H}}{\partial r_t} = 0 : \quad P_t(r_t) = c(A_t) + \tau_t + \lambda_t \tag{1.3}$$

$$\dot{\lambda}_t = \rho\lambda_t + \frac{\partial \mathcal{H}}{\partial A_t} : \qquad \dot{\lambda}_t = \lambda_t \rho - \dot{c}_t, \tag{1.4}$$

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. (1.4) become for any $\underline{t} < t < \bar{t}$,

$$\begin{aligned} \lambda_t &= \mathrm{e}^{\rho(t-\underline{t})}\lambda_{\underline{t}} - \int_{s=\underline{t}}^{t} \mathrm{e}^{\rho(t-s)}\dot{c}_s \mathrm{d}s \\ \lambda_t &= \mathrm{e}^{\rho(t-\bar{t})}\lambda_{\bar{t}} + \int_{s=t}^{\bar{t}} \mathrm{e}^{\rho(t-s)}\dot{c}_s \mathrm{d}s. \end{aligned} \tag{1.5}$$

The primary assumptions on which we will base our analysis of the supply behavior implicitly defined by the maximization 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.[4]

1.3 Economic Intuition and Green Paradox

Whilst standard dynamic programming rules allow to derive Eqs. (1.3) and (1.4) from the fuel owner's maximization problem, the economic intuition behind them is straightforward. It can be understood using a simple variational argument. Consider an owner

[4]This assumption seems largely unproblematic; an extended note on it is provided in Annex 1.

of a marginal unit of fuel extractable at a specific cost C. Taking the gross market price path p_t and the tax τ_t as given, he times his sale such as to maximize his net profit, $p_t - \tau_t - C$. Assuming smooth time paths, and a unique inner solution to exist, the seller must be indifferent between selling during his time of choice t, or marginally later or earlier. This is the case if the present value of his net revenue remains constant, that is, if a marginal delay increases the absolute net profit at the rate of interest,

$$\mathrm{d}\left[p_t - \tau_t - C\right]/\mathrm{d}t = \rho \cdot \left(p_t - \tau_t - C\right). \tag{1.6}$$

It is easy to see that Eq. (1.6) yields Eqs. (1.3) and (1.4), when λ_t is the (current value) net profit a seller gets when selling a unit of fuel, extractable at cost c_t, at time t.

For what follows in this section, we consider the exponentially increasing tax,

$$\tau_t = \tau_0 \exp(g_\tau t),$$

where $\tau_0 \geq 0$ and $g_\tau \geq 0$. For simplicity, the remainder of this section assumes price-taking fuel owners and a static inverse demand curve, $p(x)$, limited by a finite choke price, $p(0) < \infty$.

Eq. (1.6) can be rewritten as $\frac{\dot{p}_t - C}{p_t - C} = \rho + \frac{\dot{\tau}_t - \rho\tau_t}{p_t - C}$. Absent taxes, this expresses the famous Hotelling rule (Hotelling, 1931) that the resource price, net of the extraction costs, rises at the interest rate. The second term on the RHS suggests that the tax growing exponentially at the interest rate, $g_\tau = \rho$, may not affect the evolution of the price, as $\tau_t = \tau_0 \exp(\rho t)$ yields $\dot{\tau}_t = \rho\tau_t$, but that a tax growing faster may increase the slope of the price path.

Two simple additional properties allow to understand core issues about the Green Paradox central for this essay and the next. First, it is clear that the net-of-tax sales revenue, $p_t - \tau_t$ must strictly rise over periods with positive sales; otherwise a forward-looking, marginal seller could strictly increase his net present value profit, $p_t - \tau_t - C$, by anticipating the sale.

Second, the extraction costs must weakly increase over time, that is, the cheapest resources are sold first and the most expensive last. Intuitively, if one considers the sequential exploitation of two units with different extraction costs, a positive interest rate on profits implies that one fares better when selling the unit with lower extraction costs first, as this means more interest on profits overall. For analytical confirmation, consider two units, of costs C_1 and C_2, respectively, to be sold under prices P_1 and P_2 and taxes ς_1 and ς_2, at a time t_1 and $t_2 = t_1 + \delta$, with $\delta > 0$, respectively. In equilibrium, profit maximization requires $P_1 - \varsigma_1 - C_1 \geq (P_2 - \varsigma_2 - C_1) \exp(-\rho\delta)$ and

$(P_2 - \varsigma_2 - C_2)\exp(-\rho\delta) \geq P_1 - \varsigma_1 - C_2$, which is easily shown to require $C_2 \geq C_1$, that is, the extraction costs increase (weakly) over time, corresponding to the famous Herfindahl rule (Herfindahl, 1967). This ordering allows us to define a weakly increasing extraction cost curve $c(A)$ that indicates the marginal cost of extraction after A units have been extracted.

The impact of the tax on the extraction path depends on the assumption about the extraction cost curve. We first consider the case of an extraction cost curve with a finite slope in the region where the extraction costs reach the choke price: $c'(\bar{A}) < \infty$, where $\bar{A} \equiv \{A|c(A) = p(0)\}$, the amount of fuel economically extractable in the absence of a tax. As profit maximization implies that each unit of fuel is extracted if and only if the fuel price net of the tax covers the extraction costs, in this case even a small tax must unambiguously reduce the total extractions.

Because extraction stops (or, absent the tax, converges to zero) when the sum of extraction costs and the potential tax reaches (or, absent the tax, approaches) the demand choke price, rewriting Eq. (1.6) as

$$\dot{p}_t = \rho \cdot (p_t - c_t) + \dot{\tau}_t - \rho\tau_t, \tag{1.7}$$

we see that even a tax that rises at rate ρ reduces initial emissions and the emission rate throughout time:[5] Given that the tax terms cancel out in Eq. (1.7), if the initial price (and thus the initial consumption rate) would be unaffected by the tax, the price and extraction paths would be identical to the case without tax. This is, however, in contradiction with the fact that extraction must stop precisely when the sum of extraction cost and tax equals the choke price and that thus the tax necessarily reduces cumulative extraction. In addition, if the initial price would be reduced by the tax, and the consumption rate thus increased, this would weakly increase initial extraction costs. According to Eq. (1.7), the price would therefore rise less rapidly than without tax, and this would remain so throughout time.[6] The only remaining possibility is that the tax increases the initial price, and by symmetry to the case with a hypothetically lowered price we see that the higher initial price implies a reduction of the extraction rate throughout the extraction horizon.

Analogously, for a tax rising at a rate below ρ, Eq. (1.7) implies that the tax would necessarily reduce the slope of the price throughout the extraction horizon if the tax

[5] Cf. also Hoel (2010) who explains the following results, partially in slightly less detail.

[6] Intuitively, this can be understood as follows: we know that the slope of the price is the interest multiplied by the price net of the extraction cost. Reducing the initial price also slows down its increase over time, as the price net of extraction cost is reduced (by both, the lower level of the price as well as the resulting (weakly) higher extraction costs.

were to (weakly) reduce the initial price level. As this is again incompatible with the reduction of total extraction, the tax must also reduce the initial consumption.

Finally, it is trivial to see that a tax whose rate of increase is sufficiently above the interest rate may indeed increase the initial extraction: a tax with an arbitrarily low initial level but an arbitrarily high rate of increase does not inhibit initial resource sales but has the effect of an anticipated expropriation of resources not sold in the initial period, inciting fuel owners to sell at an increased rate initially. Only if the initial tax rate exceeds the initial resource rent from the no-tax scenario, does it necessarily reduce the initial consumption rate (see Eq. (1.3)).

In summary, with a continuous extraction cost curve, besides reducing total cumulative consumption, a tax rising exponentially

- at a rate $g_\tau = \rho$ reduces the extraction rate at all periods within the fuel extraction horizon,
- at a rate $g_\tau < \rho$ reduces the initial extraction rate,
- at a rate g_τ *sufficiently* larger than ρ *can* increase the initial extraction level *if and only if* the initial tax level is lower than the initial resource rent in the no-tax scenario.

This is in contrast to results from a simplified model with a fixed stock of a resource, aptly judged unrealistic in Hoel (2010). The case is characterized by a fixed stock A^* of existing resources, extractable at constant or weakly increasing extraction costs that remain strictly below the demand choke price: $c(A^*) < p(0)$. In that case of a fixed reservoir, even the last existing unit of resource may be extracted for costs strictly below the demand choke price, and a tax must thus not necessarily reduce the cumulative extraction. If the tax is low enough to leave total extractions unchanged, that is, if it is not too high to allow the economical extraction of the last unit of stock, Eq. (1.7) implies that the price and extraction paths are unchanged when $g_\tau = \rho$.[7] The fixed total extraction implies that the price paths with and without tax must cross when $g_\tau \neq \rho$, with a lower (higher) initial price for $g_\tau > \rho$ (for $g_\tau < \rho$). That is, a low enough tax increases (reduces) the initial extraction rate *iff* $g_\tau > \rho$ (for $g_\tau < \rho$). Nevertheless, if the tax exceeds the minimum level required to reduce to zero the resource rent from the last unit of stock, we are back to the situation where the tax reduces the total

[7] The tax would, however, imply a distortion-free shift of resource rents from the fuel owners to the taxing institution.

extraction and where Eq. (1.7) thus implies the three results emphasized for the case of the endogenous extraction.[8]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,[9] 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. (1.2) through (1.4) presents the suppliers' maximization 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. (1.5), this yields

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

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

On the other hand, *if no backstop is introduced*, we know that $\lim_{t\to\infty} \lambda_t \mathrm{e}^{-\rho t} = 0$. Using thus $\bar{t} = \infty$ and $\lim_{\bar{t}\to\infty} \mathrm{e}^{-\rho\bar{t}} \lambda_{\bar{t}} = 0$ in Eq. (1.5) we get

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

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

$$\lambda_T \begin{cases} = 0 & \text{for} \quad \text{BS (with backstop)} \\ > 0 & \phantom{\text{for}} \quad \text{BAU (no backstop).} \end{cases} \tag{1.8}$$

It is obvious that the introduction of the backstop at time T affects the resource owners' optimization problem exactly in the same manner as would a tax introduced at T if the

[8]The simplest way to understand this is to see that raising the tax leaves the extraction path unaffected until the tax reaches the level which extracts all the rent (in the case of heterogenous costs, that of the last unit of stock), and that to further increase the tax from this point on is equivalent to imposing a tax in the case of the fully endogenous extraction.

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

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 emphasized in Lemma 1 and proven in Annex 2, 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 3, 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. (1.8) 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 $\lambda_T = 0$.

1.4 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. (1.2) through (1.4) 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.[10] This is emphasized in Proposition 1 and proven in Annex 5.

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.

1.5 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

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

applies to the – 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 a few 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 a 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 maximize 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

$$\lambda_T(A_T^*) \stackrel{!}{=} A_T^{\text{inv}}(A_T^*),$$

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

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

$$\frac{\partial \lambda_T^{\text{pre}}(A_T^*)}{\partial A_T} \leq \frac{\partial \lambda_T^{\text{post}}(A_T^*)}{\partial A_T} \tag{1.9}$$

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

$$\begin{aligned}
\Pi(A_T^* + \varepsilon) - \Pi(A_T^*) &= \Pi^{\text{pre}}(A_T^* + \varepsilon) + \Pi^{\text{post}}(A_T^* + \varepsilon) - \Pi(A_T^*) \\
&\approx \varepsilon \lambda_T^{\text{pre}}(A_T^*) + \frac{\varepsilon^2}{2} \frac{\partial \lambda_T^{\text{pre}}(A_T^*)}{\partial A_T} - \varepsilon \lambda_T^{\text{post}}(A_T^*) - \frac{\varepsilon^2}{2} \frac{\partial \lambda_T^{\text{post}}(A_T^*)}{\partial A_T} \\
&\approx \frac{\varepsilon^2}{2} \left[\frac{\partial \lambda_T^{\text{pre}}(A_T^*)}{\partial A_T} - \frac{\partial \lambda_T^{\text{post}}(A_T^*)}{\partial A_T} \right]
\end{aligned} \tag{1.10}$$

for small deviations from A_T^*. Clearly, if Eq. (1.9) does not hold, Eq. (1.10) 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. 1.1, 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.

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. 1.2.

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. 1.3.

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. (1.9) 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. 1.4.

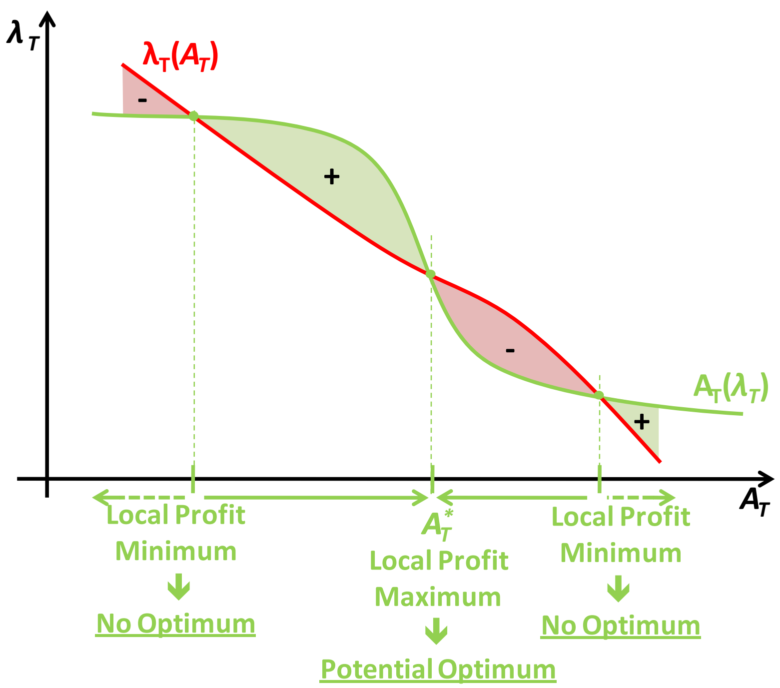

Figure 1.1: Possible equilibrium situations with flexible $\lambda_T(A_T)$

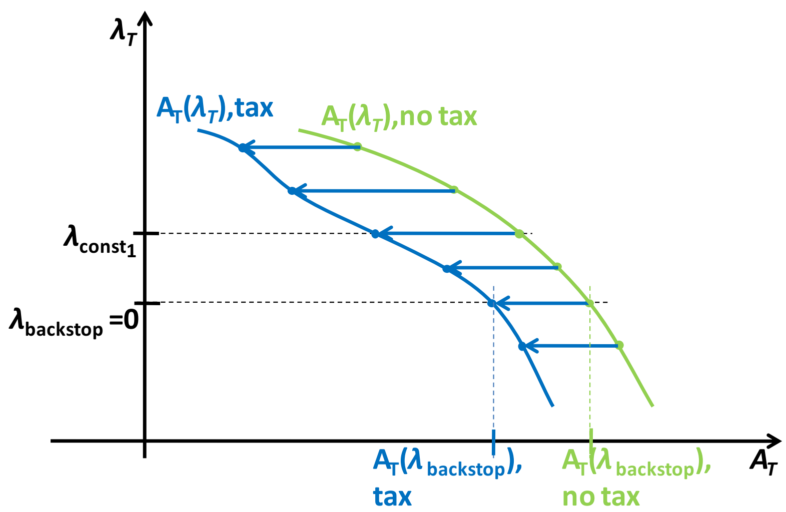

Figure 1.2: Tax reduces pre-T emissions A_T for constant λ_T

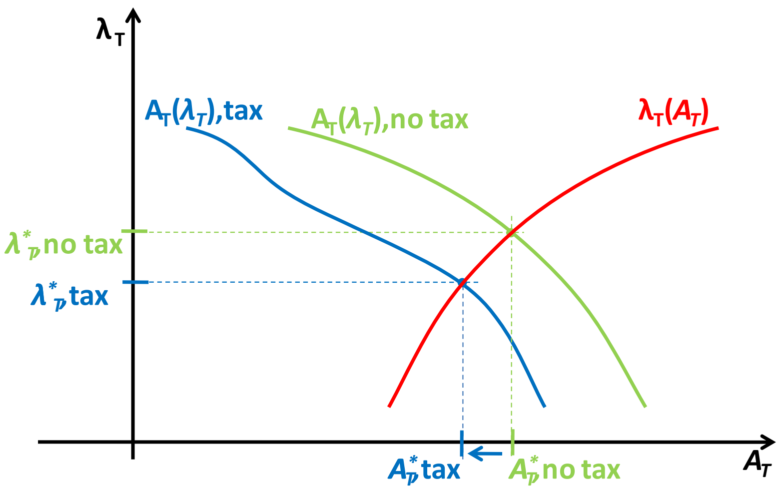

Figure 1.3: Tax reduces pre-T emissions A_T *for flexible* $\lambda_T(A_T)$ *when* $\lambda'_T(A_T) > 0$

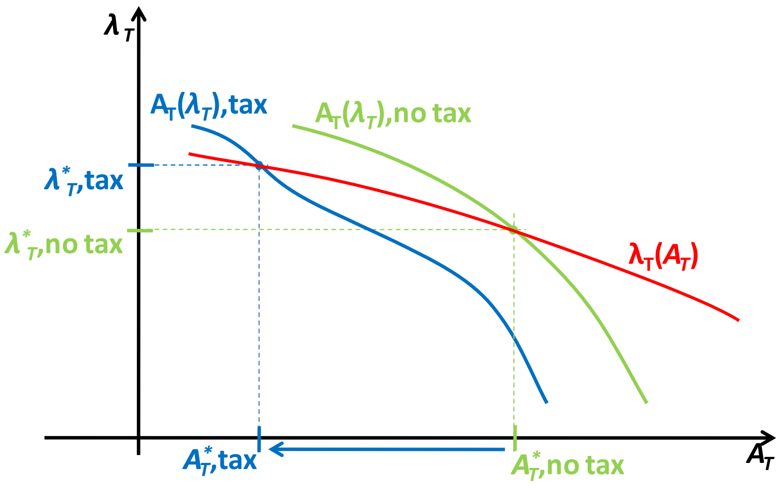

Figure 1.4: Tax reduces pre-T emissions A_T *for flexible* $\lambda_T(A_T)$ *when* $\lambda'_T(A_T) < 0$

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

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 8, 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.

1.6 Further Extensions

1.6.1 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 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. 1.5, the total demand for a given sales price and tax rate is

$$r(p_R, \tau) = r_1(p_1) + r_2(p_2) = r_1(p_R + \tau) + r_2(p_R). \tag{1.11}$$

Effect of the tax on the demand The demand curves of both regions are assumed to be continuous and strictly decreasing. Thus, Eq. (1.11) implies that the current worldwide 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.

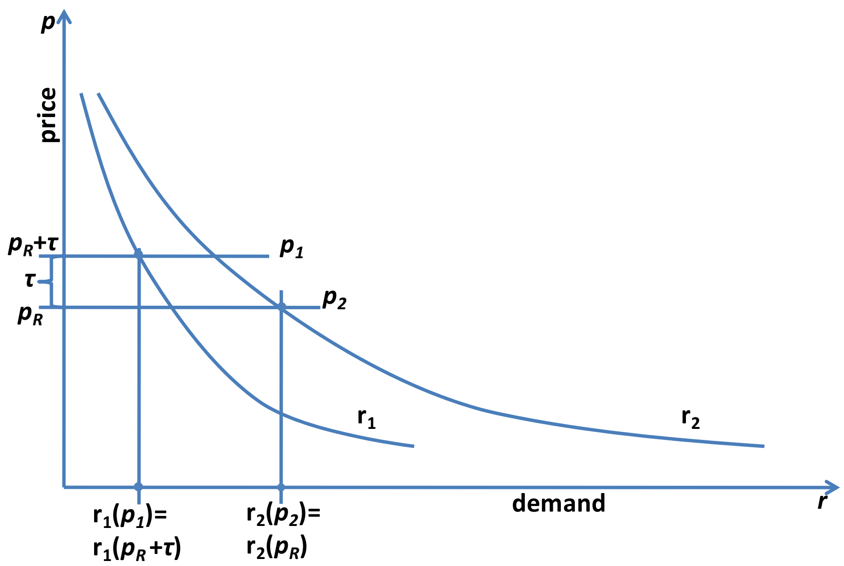

Figure 1.5: Regional demand with tax in Region 1

Given this new demand structure, the optimality condition Eq. (1.3) becomes

$$p(r_t, \tau_t) + r_t \frac{\partial p(r_t, \tau_t)}{\partial r} = c(A_t) + \lambda_t. \tag{1.12}$$

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. (1.12), 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 Annex 6), 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. (A1.18) 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 emphasize 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 maximization 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 Annex 6.

1.6.2 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).[11]

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

[11] 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.

constant. The additional discounting factor, ψ, which equals this periodic probability, inherits this 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 emphasize this claim with Proposition 4, and the analytical proof is given in Annex 7.

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 (forthcoming). Focusing on R&D subsidies rather than on a tax, the study uses a value function approach and shows that the subsidy can increase initial emissions but reduce expected long-run emissions.

1.6.3 Endogenous Future Regime Change

The introduction of a carbon tax influences the consumption prices for conventional energy and consequently changes incentives for the (decentralized) development of alternative technologies. However, the tax also affects carbon emissions and the climate and thus the political pressure to work on (centralized) 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. We omit the proof, but it seems clear that in this case the above results would even 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.

1.7 Interpretation

We assumed a world where the extraction cost curve for the fuels is a continuous function, implying an endogenous total amount of extraction rather than, as assumed in the original Green Paradox framework, a fixed amount of fuels. Given the diversity of existing fuel deposits, we consider our assumption as a better representation of reality. 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 21^{st} 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.[12] 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 lead to the potentially devastating warming effect of several degrees.

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 emphasizes – 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 stylized, 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.[13] We assumed the alternative measure to be independent of the decision on the current tax. This is a strong assumption. As

[12] 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.

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

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 modeling 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 favor of the tax raised in the literature necessarily invalidate all aspects of the Green Paradox. Hoel (2010) argues that 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 favor 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.[14] 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

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

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.[15] Under these conditions, the analysis of the profit-maximizing behavior 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 behavioral 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 utilized to fix the tax rate enough above the 'optimal'[16] rate so that the gross price exceeds the net price from the no-tax scenario by the value of the 'optimal' tax rate.[17] 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 behavioral 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.

[15] 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.

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

[17] 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.

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.

1.8 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 developments encompass not only technological innovations driven by increasing fossil 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 favorable for our climate than has been predicted by the Green Paradox:

First, if some of the previously mentioned future climate regime switches were to materialize 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.

For convenience and analytical tractability, we derived these results in a dynamic partial equilibrium model with exogenous fuel demand and interest rate, where the introduction of the alternative future measure – or its probability in the stochastic case – is correctly anticipated by the fuel owners. We believe that the underlying mechanism is prevalent also in the real world: As the quote from Sinn in our introduction shows, even advocates of the Green Paradox explain that there is a justified hope for stringent future climate measures to materialize. Fuel owners will in general anticipate the possibility of such alternative developments and the (partial) quasi-expropriation of their fuels which may come about from it. In this case, a current abstention from imposing a tax can resemble a baseline scenario with no tax today and a (potentially high) tax in future: if the future measure limits fuel consumption, it can have a similar effect on the future resource rents as a future tax. From the Green Paradox point of view, this situation with no measure at present and a strong measure in the future should be avoided in order to prevent an acceleration of fuel sales. Our work emphasizes that even if the tax to be introduced is rapidly increasing over time, it may in this case be desirable, as it would help to limit the amount of fuel that the owners may shift to the phase before the future measure comes about. This reasoning can work for both cases, where, as we assumed, the future alternative measure will come about independently of whether a tax is introduced in earlier periods (and replace the tax), or whether the future measure is implemented only in the absence of the current tax: in both situations, the future measure threatens the future resource rents and thus incites them to increase sales in the early periods, which the current tax mitigates.

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.

1.9 Annex

Annex 1 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.

Annex 2 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, $\Delta v = v_{\text{BAU}} - v_{\text{BS}}$. For the case where the BS scenario has the lowered final marginal resource value, the claim of Lemma 1 can then be stated as

$$\Delta\lambda_T < 0 \Rightarrow \Delta A_T > 0, \tag{A1.1}$$

with the considered time span being $t = [0, T]$. We will show by contradiction that the claim in Eq. (A1.1) holds unambiguously.

Assume thus the contrary,

$$\Delta\lambda_T < 0 \wedge \Delta A_T \leq 0, \tag{A1.2}$$

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. (A1.2) to hold, we thus know that the RHS in Eq. (1.3) (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:

$$\lim_{t\to T}\, \Delta r_t > 0 \tag{A1.3}$$

To have Eq. (A1.2) despite Eq. (A1.3) 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:

$$\exists\, t^* \,\text{s.t.} \begin{cases} \Delta r_{t^*} = 0 \\ \Delta r_t \geq 0 \,\forall\, t > t^* \\ \exists \varepsilon > 0 \,\text{s.t.}\, \forall\, t \in [t^* - \varepsilon, t^*)\, \Delta r_t < 0 \end{cases} \tag{A1.4}$$

Together with the second relation in Eq. (A1.4), Eq. (A1.2) implies $\Delta A_{t^*} < 0$. Together with the third relation in Eq. (A1.4), this implies that we either have

$$\exists \varepsilon > 0\,\text{s.t.}\, \Delta r_{t^*-\varepsilon} < 0 \,\wedge \begin{cases} \text{Case 1:} \quad \exists \delta > 0\,\text{s.t.}\, \underbrace{A_{t^*-\varepsilon,BAU} = A_{t^*+\delta,BS}}_{\equiv a}, \\ \quad \text{or} \\ \text{Case 2:} \qquad A_{t^*-\varepsilon,BAU} > A_{t,BS}, \end{cases} \tag{A1.5}$$

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. (A1.5). From Eq. (A1.4) 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. (A1.5)), 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. (A1.5), stating that the profit maximizing 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.[18]

Consider now case 2 in Eq. (A1.5), 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 maximizing 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,

$$(P_{t^*-\varepsilon,BAU} - a)\, e^{-r(t^*-\varepsilon)} \geq \lambda_{T,BAU}e^{-rT}, \tag{A1.6}$$

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

$$(P_{t^*-\varepsilon,BS} - a) \leq \lambda_{T,BS}e^{-rT}. \tag{A1.7}$$

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. (A1.6) and (A1.7).

Assuming Eq. (A1.2) thus yields a contradictory result in both elsewise possible cases, case 1 and case 2, which concludes our proof by contradiction. ■

Annex 3 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

[18]For a formal version of an analogous argumentation see the last part of the proof of Lemma 2 (Annex 3).

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,

$$\lim_{t \to T} r_{t,BS} < r_{t,BAU}, \tag{A1.8}$$

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 (A1.8) 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,

$$A_{T,BS} > A_{T,BAU}. \tag{A1.9}$$

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 (A1.8), 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 (A1.9) 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

$$r_{t',BS} > r_{t',BAU} \wedge r_{t'',BS} \leq r_{t'',BAU}. \tag{A1.10}$$

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 (A1.10), 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. A1.10 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 maximizing 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',

$$(P_{t',BS} - a)\, e^{-\rho t'} \geq (P_{t'',BS} - a)\, e^{-\rho t''}, \tag{A1.11}$$

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,

$$(P_{t',BAU} - a)\, e^{-\rho t'} \leq (P_{t'',BAU} - a)\, e^{-\rho t''}. \tag{A1.12}$$

Given decreasing marginal sales revenues (Property 5 from Section 2), Eq. (A1.10) does, however, imply, for the gross sales revenues

$$P_{t',BS} < P_{t',BAU} \wedge P_{t'',BS} \geq P_{t',BAU}. \tag{A1.13}$$

Eqs. (A1.12) and (A1.11) are, as an ensemble, however, in direct contradiction with Eqs. (A1.13). ■

Annex 4 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 $\rho > 0$,*

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

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 $\eta \equiv \int_0^T \mathrm{e}^{-\rho t} h(t)\mathrm{d}t$. Then, define h^m as the path that minimizes the discounted integral while respecting the imposed condition:

$$\begin{array}{c} \min_{h^m} \eta \\ s.t. \quad \int_0^t h^m(s)\mathrm{d}s > 0 \underset{t\epsilon(0,T)}{\forall} \end{array} \tag{A1.14}$$

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. (A1.14) 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 (A1.14) 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 minimized 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, $\int_0^T \mathrm{e}^{-\rho t}\,[g(t) - f(t)]\,\mathrm{d}t > 0$. ■

Annex 5 Proof of Proposition 1

Assume an exogenously given, fix λ_T.

From Eq. (1.5), we know for the monopolistic supplier

$$\lambda_t = \lambda_T \mathrm{e}^{\rho(t-T)} + \int_{s=t}^{T} \mathrm{e}^{(t-s)\rho} \dot{c}_s \mathrm{d}s. \tag{A1.15}$$

It will be intuitive that our analysis holds for the competitive case as well. Inserting Eq. (A1.15) in Eq. (1.3) yields

$$p_t(r_t) + r_t p_t'(r_t) \quad = \quad c_t + \tau_t + \lambda_T \mathrm{e}^{\rho(t-T)} + \int_{s=t}^{T} \mathrm{e}^{(t-s)\rho} \dot{c}_s \mathrm{d}s. \tag{A1.16}$$

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

$$A_{T,\text{tax}} \geq A_{T,\text{no}}, \tag{A1.17}$$

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. (A1.17) implies

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

We have $\lim_{t\to T} \int_{s=t}^{T} \mathrm{e}^{(t-s)\rho} \dot{c}_s \mathrm{d}s = 0$ and, from Eq. (A1.17), $\lim_{t\to T} c_{T,\text{tax}} \geq \lim_{t\to T} c_{T,\text{no}}$. Therefore, the RHS of Eq. (A1.16) 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

$$\lim_{t\to T} r_{t,\text{tax}} < \lim_{t\to T} r_{t,\text{no}}. \tag{A1.18}$$

Because all our variables evolve smoothly over time Eqs. (A1.17) and (A1.18) 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

$$r_{t^*,\text{tax}} = r_{t^*,\text{no}}, \tag{A1.19}$$

and

$$r_{t,\text{tax}} < r_{t,\text{no}} \,\forall_{t^* < t \leq T}. \tag{A1.20}$$

Relation Eq. (A1.20) 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. (A1.17) can only hold if

$$c_{t,\text{tax}} > c_{t,\text{no}} \,\forall_{t^* \leq t < T}. \tag{A1.21}$$

Eqs. (A1.19) and (A1.16), as well as the fact that $\tau_t \geq 0$ imply

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

and thus

$$\int_{t=t^*}^{T} \mathrm{e}^{\rho(t^*-t)}(\dot{c}_{t,\text{no}} - \dot{c}_{t,\text{tax}})\mathrm{d}t \geq c_{t^*,\text{tax}} - c_{t^*,\text{no}}. \tag{A1.22}$$

As according to Eq. (A1.21) the RHS of Eq. (A1.22) is strictly positive, it is easy to see that Lemma 4 (Annex 4) implies that Eqs. (A1.21) and (A1.22) cannot be reconciled, which concludes our proof by contradiction. ■

Annex 6 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 equalized. 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 equalized 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 equalized. 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:

$$\begin{aligned} r_1 + r_2 &= r \\ r_2(p) &= x \cdot r_1(p) \end{aligned} \tag{A1.23}$$

Eq. (A1.23) implies that all derivatives of the regional demand function differ by a factor x as well:

$$r_2^{(i)}(p) = x \cdot r_1^{(i)}(p), \tag{A1.24}$$

where the indice $(\cdot)^{(i)}$ denotes the i^{th} 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

$$r(p_R, \tau) = r_1(p_1) + r_2(p_2) = r_1(p_R + \tau) + r_2(p_R). \tag{A1.25}$$

As the demand curves in the two regions are continuous and strictly decreasing, Eq. (A1.25) 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,

$$\Delta p_1 + \Delta p_2 = \tau. \tag{A1.26}$$

Consider the illustration in Fig. A1.1.

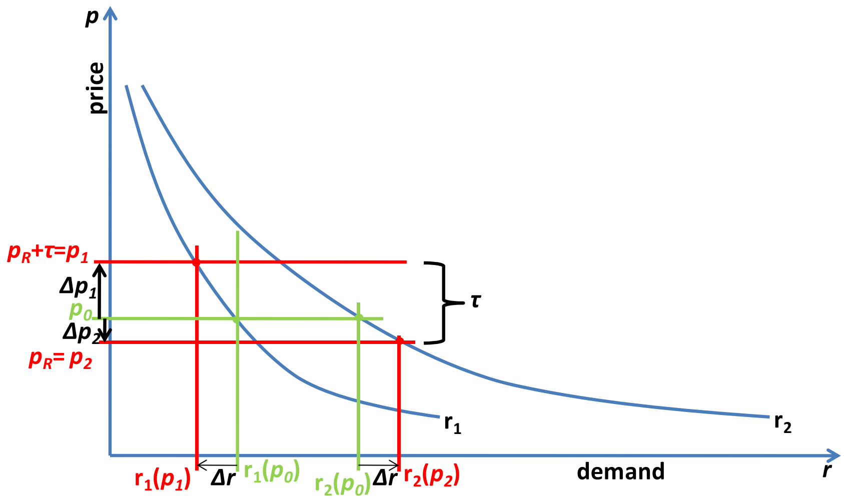

Figure A1.1: Hypothetical situation of regional tax which is neutral for global emissions

In a first approximation, we have:

$$\Delta r \approx \Delta p_1 \cdot r_1'(p_0) \tag{A1.27}$$

$$\Delta r \approx \Delta p_2 \cdot r_2'(p_0) \tag{A1.28}$$

With the inclusion of Eq. (A1.24), using Eqs. (A1.27) and (A1.28) in Eq. (A1.26) implies

$$\Delta p_2 \approx \frac{\tau}{1+x}, \tag{A1.29}$$

yielding

$$p_b \approx p_0 - \frac{\tau}{1+x}. \tag{A1.30}$$

Eq. (A1.30) 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. (A1.29) and (A1.26) follows

$$\Delta p_1 \approx x \cdot \Delta p_2. \tag{A1.31}$$

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:

$$\begin{aligned}
r_1'(p_1) &\approx r_1'(p_0) + \Delta p_1 \cdot r_1''(p_0) + \frac{(\Delta p_1)^2}{2} r_1'''(p_0) && \text{(A1.32)}\\
r_2'(p_2) &\approx r_2'(p_0) - \Delta p_2 \cdot r_2''(p_0) + \frac{(\Delta p_2)^2}{2} r_2'''(p_0) && \text{(A1.33)}\\
&\approx x \cdot r_1'(p_0) - \Delta p_1 \cdot r_1''(p_0) + \frac{(\Delta p_1)^2}{2x} r_1'''(p_0), && \text{(A1.34)}
\end{aligned}$$

where the minus sign for the second term on the right hand side in Eq. (A1.33) is due to the fact that Δp_2 is defined in absolute terms, and where Eq. (A1.34) follows from Eq. (A1.33) using Eq. (A1.24) as well as Eq. (A1.31).

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. (A1.32) and (A1.34), as well as Eq. (A1.24), 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. (A1.30):

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

Using Eqs. (A1.29) and (A1.31) 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:

$$\frac{\Delta\left[\frac{\partial p_R}{\partial r}\right]}{\Delta p} \approx \frac{\tau\cdot x\cdot r_1'''(p_0)}{2\,(1+x)^2\,r_1'(p_0)^2}, \tag{A1.35}$$

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

As the ratio in Eq. (A1.35) 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. ■

Annex 7 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' maximization 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

$$\begin{aligned} P_t(r_t) &= c(A_t) + \tau \mathrm{e}^{\theta t} + \lambda_t \\ \dot{\lambda}_t &= \lambda_t(\rho+\psi) - \dot{c}_t. \end{aligned} \tag{A1.36}$$

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

$$\lambda_t \;=\; \lambda_T \mathrm{e}^{\delta(t-T)} + \int_{s=t}^{T} \mathrm{e}^{\delta(t-s)}\dot{c}_s \mathrm{d}s. \tag{A1.37}$$

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: $\Delta\text{var} \equiv \text{var}_{\text{tax}} - \text{var}_{\text{no}}$, where var can be a single variable or a combined mathematical term.

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

$$\Delta A_{t_0} = 0 \tag{A1.38}$$
$$\text{and} \qquad \Delta A_t > 0 \underset{t_0<t<\infty}{\forall}.$$

This requires $\Delta r_{t_0} \geq 0$, and therefore, due to Eq. (A1.36) and Property 5, that $\Delta[\lambda_{t_0} + \tau_{t_0}] \leq 0$, wherewith we have

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

From Eq. (A1.37) (and the transversality condition), however, we know that $\lambda_{t_0} = \int_{t=t_0}^{\infty} e^{\delta(t_0-t)}\dot{c}_t dt$, which can be rewritten as

$$\lambda_{t_0} = e^{\delta t_0}\int_{t=t_0}^{\infty} e^{-\delta t}\dot{c}_t dt. \tag{A1.39}$$

It is, however, clear that Eqs. (A1.38) through (A1.39) are not reconcilable with Lemma 4 (Annex 4). 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.

$$\Delta A_{t_1} = 0 \wedge \Delta r_{t_1} \geq 0, \tag{A1.40}$$
$$\Delta A_{t_2} = 0 \wedge \Delta r_{t_2} \leq 0, \tag{A1.41}$$
$$\text{and} \qquad \Delta A_t \geq 0 \underset{t_1<t<t_2}{\forall}. \tag{A1.42}$$

Eqs. (A1.40) and (A1.41) imply

$$\Delta c_{t_1} = 0 \wedge \Delta c_{t_2} = 0, \tag{A1.43}$$

and therewith also

$$\begin{aligned} & \Delta[\lambda_{t_1} + \tau_{t_1}] \leq 0, \\ \text{and} \qquad & \Delta[\lambda_{t_2} + \tau_{t_2}] \geq 0. \end{aligned} \tag{A1.44}$$

Eq. (A1.42) indicates that

$$\Delta c_t > 0 \underset{t_1<t<t_2}{\forall}. \tag{A1.45}$$

From Eq. (A1.37), 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

$$\Delta[\lambda_{t_2} + \tau_{t_2}] \geq 0 \Rightarrow \mathrm{e}^{-\theta(t_2-t_1)} \Delta[\lambda_{t_2} + \tau_{t_2}] \geq 0 \Rightarrow \Delta[\mu_{t_2} + \tau_{t_1}] \geq 0. \tag{A1.46}$$

As $\Delta \tau_{t_1} > 0$ the last expression in Eq. (A1.46) implies

$$\Delta[a\mu_{t_2} + \tau_{t_1}] > 0 \underset{0 \leq a < 1}{\forall}. \tag{A1.47}$$

From Eq. (A1.44) we know $\Delta[\mu_{t_1} + \tau_{t_1}] \leq 0$, which we can rewrite as

$$\Delta[\mu_{t_2} \underbrace{\mathrm{e}^{[\delta-\theta](t_1-t_2)}}_{<1} + \int_{t_1}^{t_2} \mathrm{e}^{\delta(t_1-t)} \dot{c}_t \mathrm{d}t + \tau_{t_1}] \leq 0. \tag{A1.48}$$

Eqs. (A1.47) and (A1.48) imply

$$\Delta[\int_{t_1}^{t_2} \mathrm{e}^{\delta(t_1-t)} \dot{c}_t \mathrm{d}t] \leq 0. \tag{A1.49}$$

Eqs. (A1.49), (A1.43) and (A1.45), however, violate Lemma 4 (Annex 4), 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. (A1.36). 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. ■

Annex 8 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 stylized 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).

Essay 2

No Green Fuel-Tax Paradox on Earth?

Abstract

It has been postulated that a realistic carbon tax would be subject to a strong Green Paradox: Present discounted emissions would increase as anticipation effects would imply that an initially low and rapidly increasing climate tax would induce forward-looking fuel owners to speed up the extraction rather than to leave more fuel underground. While a tax may indeed increase initial emissions, it has been shown that under realistic assumptions long-run emissions are also reduced by the tax. A calibrated dynamic model of the fuel market, taking into account different fuels and the worldwide reserve (cost-)structure, is employed to investigate the plausibility that tax-induced early emission increases outweigh the desired later emission reductions, either in terms of the net present value of emissions or of the damages they induce. The results are clear-cut: Initial emission increases are so tiny and future emission (and damage) reductions are so large that, even with up to very high discount rates of around 10 %, taxes reduce the net value of the climate externality. This result holds independent of whether an endogenous clean backstop, liquefaction, convex damages based on a Nordhaus climate and damage module, or rate-dependent extraction costs are taken into account, and for a broad range of parameters and of Green Paradox-relevant tax paths. It remains valid also if the tax is regionally limited to the OECD. The resource owner's price response makes the carbon tax even more desirable.

Author: Florian Habermacher

Keywords: Green Paradox, climate change policy, greenhouse gas tax, anticipation effects, exhaustible resources, fossil fuels market, dynamic numerical simulation, backstop technology, resource rent, liquefaction.

JEL classification: Q54, Q41, Q48, Q42.

I am grateful to Gebhard Kirchgässner, Rick van der Ploeg, and participants at the Venice Summer Institute workshop on the Theory and Empirics of the Green Paradox in Venice, 20-21 July 2012, the International Association for Energy Economics European Conference in Venice, 9-12 September 2012, and the Brown Bag seminar at University of St. Gallen, 19 December 2012, for valuable comments and suggestions.

2.1 Introduction

Twenty years ago, Sinclair (1992) pointed out that rapidly increasing resource taxes could increase rather than slow down the speed at which resources are extracted. Recently, this idea has been popularized by Sinn (2008) with the Green Paradox, according to which realistic climate taxes (and potentially many other climate policies as well) would be counterproductive for the climate rather than mitigating its change.

The Green Paradox, with respect to climate taxes, has been brought forward in three main variants. First, because political constraints restrict the level of the tax to initially low values but the tax is likely to rise rapidly over time, fuel owners would sell their resources at an increased pace in the early periods after the introduction of the tax. This is the case because they prefer to sell them in periods with low taxes rather than when the tax has become very high. That is, early emissions would increase due to the introduction of the tax, a mechanism coined by Gerlagh (2011) as the 'weak' Green Paradox.

Second, if there exists a fixed amount of fuels that is exploitable at constant costs, and if the willingness to pay for those fuels is large enough (potentially infinitely large for the last tons of the fuel), the tax would in addition not have any effect on the total amount of fuels sold. In other words, not only would parts of the fuels be sold earlier due to the tax than they would have without it, but cumulative emissions would not be affected at all by the tax. This is what we call here the 'very strong' version of the Green Paradox because it would almost necessarily imply that the tax increases the climate damage independent of the strength by which we discount future emissions.

Third, by increasing early emissions, the tax may increase the net present value (NPV) of all future emissions, even if, under relaxed assumptions discussed below, it reduces long-term cumulative emissions. It will be an empirical question – and to the extent that no broadly agreed upon appropriate emission (or damage) discount rate exists, a somewhat normative one – whether we judge the overall effect of (i) increased near-term emissions vs. the (ii) reduction in the long-run cumulative emissions to be desirable or not. When the tax increases the NPV of the emissions or the induced damages we talk about a 'strong' Green Paradox, a notion, again, introduced by Gerlagh (2011). This study tries to give an answer to whether such a strong Green Paradox is likely to materialize for rapidly increasing carbon taxes.

When trying to apply the abstract Green Paradox theory to the case of the world, it is crucial to see that the very strong Green Paradox with respect to a carbon tax relies on a model with a reservoir of a fixed amount of fossil fuel that can be extracted at limited, potentially constant marginal costs, and under the premise that society would have the choice between (a) a rapidly rising fuel tax, and (b) none now or in future. Further, the willingness to pay for the fuel is assumed to be so large, and the tax level is supposed to not become too large too early, so that all the fuels would be exploited over time despite the tax.

In reality, extraction costs cannot be considered as constant. Instead the amount of extractable fuels is increasing in the maximal marginal extraction cost. As is intuitively understandable, and shown by Hoel (2010), in this case the total amount of fuel extracted is strictly decreasing in the longer-run tax level, that is, any positive tax reduces overall cumulative emissions.[19] Thus, the very strong Green Paradox is only an artifact of a simplifying model assumption of an exactly fixed fuel stock that is fully exploited for sure.

Besides limiting the overall amount of fuels extracted, a second goal of climate policy is, however, to delay current emissions, that is, future potential emissions are discounted. If this second aim did not exist, the Green Paradox would only be an issue about the criticized climate policies being potentially useless, but not counterproductive (as the tax is hardly expected to increase cumulative long-run emissions). Because of this second aspect of climate policy – the desire to postpone emissions for as long as possible – the Green Paradox, according to which the policies would make the emissions occur earlier, could imply that the policies are not only a wasted effort but that they are even counterproductive, aggravating instead of alleviating climate change. When this time-discounting is considered, it is – even with increasing extraction costs – not clear *a priori*, whether the sum of present discounted emissions (or damages) increases or decreases due to the carbon tax, that is, whether the relevant, strong version of the Green Paradox holds. Therewith, the question, whether a carbon tax with plausible characteristics can really be counterproductive, as Sinn (2008) has claimed in his original contribution on the Green Paradox, remains open even with increasing and prohibitive extraction costs. It is an empirical question. The answer depends primarily on the (evolution of) the worldwide fuel demand structure (which, among others, also depends on technological developments) and of the fuel supply cost curves (that is, how much fuels can be extracted at which marginal costs), as well as on the constraints, beliefs and objectives according to which the fuel owners act.[20]

[19] This assumes a finite demand choke price for the fuels and fuel extraction costs that surpass that choke price for parts of the fuels. These assumptions seem natural: an infinite willingness to pay for fossil fuels seems implausible notably as renewable substitutes are available at finite costs. Fuel extraction costs are bound to become prohibitively high at least when the last quantiles stored, e.g., in unconcentrated form at the edges of reservoirs, are accessed.

[20] That is, to which extent their behavior corresponds or deviates from that of forward-looking economic agents that maximize expected sales profits under standard discounting.

The aim of this contribution is to shed more light on this issue in probably the most natural approach: examining the fuel market in a dynamic setting where the fuel demand, the fuel supply cost curves, and the behavior of the fuel owners, as well as the climate dynamics (based on Nordhaus, 2008) are modeled. Analyzing the dynamic response of the fuel consumption – and the related emission – paths to the introduction of (different variants of) a carbon tax exhibiting the characteristics that Sinn had described, allows to find concrete answers to the question of whether such taxes alleviate or aggravate the climate problem. Or at least, it can lend itself to some judgment on whether the (Hotelling resource extraction) framework within which the Green Paradox had originally been brought forward in a theoretical analysis does really support the conclusion that the described climate taxes would exacerbate climate change as has been postulated. Besides using the standard Hotelling resource rent and extraction model with empirical cost curves for (cumulative) fuel supply, we test the effect of clean backstops, coal liquefaction, and extraction rate-dependent extra costs, and in addition to global taxes we also examine a regional tax.

The findings show that for a very broad spectrum of the controversial rapidly increasing carbon taxes the anticipation effects are very unlikely to increase present discounted climate damage. Especially if besides the often focused-on oil, coal – of which much larger reserves exist – is taken into account, a tax is most likely to reduce the sum of present discounted damages, increasing current emissions only slightly but reducing later damages drastically. Only an unusually rapid climate discount rate, combined with particular assumptions about other key-determinants of the fuel market outcome, would sustain the Green Paradox claim that carbon or fuel taxes could harm the climate. But even in this case, the result would not necessarily imply a clear policy recommendation against a carbon tax as a commitment problem must be accounted for. While the discounted emissions may theoretically be larger with a currently introduced tax compared to those in the business-as-usual (BAU) scenario of the perpetual absence of a tax, it is clear that the current policy cannot commit to the absence of the tax even in later periods. In fact, if the introduction of a tax is currently undesirable compared to a scenario without any tax ever, but the introduction of a tax becomes desirable at a later point in time,[21] there is no reason to believe that the tax (if currently abstained from due to the Green Paradox-effects) would not be introduced from that later point on, as had already been pointed out in Hoel (2010).[22] The fuel-owners will in this case already today anticipate this introduction of the tax later on, and therefore would try to sell a lot of their fuels on the current market before the tax is introduced. This in turn

[21] As extraction costs approach the choke price which caps the demand, the Green Paradox becomes less likely.

[22] Hoel (2010) did, however, focus on a much more stylized two-period model.

makes the current introduction of a tax more desirable. In a systematic analysis of this issue with a model closely resembling to that used in this paper, Habermacher (2012a) has shown that an unraveling becomes apparent; the abstention of the tax today lacks subgame-perfectness for even a much broader set of key-determinants than for which the current introduction of the tax dominates its perpetual absence in green-only welfare terms. This further increases the likelihood that the current introduction of a tax is beneficial rather than, as has been suggested with the Green Paradox, detrimental for the climate.

Section 2 describes the model and its calibration for the main model runs, as well as a number of extensions. Section 3 provides a note on the interpretation of the results. Section 4 shows the impact of the tax in the main scenario as well as in a large number of model runs with different setups, parameterization and taxes. Section 5 concludes.

2.2 Model

A small dynamic equilibrium model of the worldwide market for oil and coal – the two fossil fuels which are responsible for the large majority of energy related, manmade carbon dioxide emissions – is used. The model structure corresponds largely to that developed in Essay 3 in this dissertation to investigate interregional and intertemporal carbon leakage. A major difference of the model used here is that, rather than assuming a fixed fraction of carbon emissions to remain in the atmosphere forever, it uses the carbon cycle model by Nordhaus (2008), explicitly modeling the dynamic transfer of carbon between the atmosphere and the upper and lower ocean, and taking into account emissions other than those of endogenous fossil fuel use, to estimate NPV damages. Further, an extension of the model allows for a form of rate-dependent extraction costs. This presents a possible way to account for the conventional wisdom that particularly geological constraints prevent fuel owners from exploiting their reserves at arbitrarily high rates,[23] and it allows for the closure of a gap between the simulated initial fuel consumption levels – which tend to be too high in Hotelling extraction models without rate-dependent costs – and the regional consumption rates observed today.

[23] A more detailed discussion of this issue, and a more micro-funded way of addressing it than the more ad hoc manner used here is given in Venables (2011), who provide an analytically tractable model.

2.2.1 Setup

The model contains two fuel consuming regions (index r): the OECD and the rest of the world (ROW or Non-OECD). The carbon taxes may be imposed by the whole world or only by the OECD. The two fuels considered are oil and coal (index $i = \{1, 2\}$). A few words on the restriction to oil and coal as the two fossil fuels considered may be in order. The simulation results will prove complex already when we restrict the attention to oil and coal, with the interpretability presumably being further complicated when gas would be taken into account as well, whilst it is not clear whether relevant insights would be gained. Currently, 80 % of energy supply carbon emissions[24] stem from coal (43 %) and oil (36 %) burn, and only 20 % from gas. Moreover, whilst gas is occasionally considered as the fuel of the future, in reality, the current growth of total global carbon dioxide emissions is attributable by more than 50 % to coal, and by 2/3 to coal and oil, with the remainder attributable to other sources, including gas. Moreover, in the faster growing non-industrialized world, the share of coal and oil in the growth of all CO_2 emissions exceeds 75 % (IEA, 2012). Finally, because gas has many features similar to oil, especially in terms of the exhaustibility and the convertibility of coal through gasification or liquefaction, one may, to a certain degree, interpret 'oil' in our model as representative of the ensemble of oil and gas, an approach also used by van der Ploeg and Withagen (2011).[25]

The fuels are traded internationally at prices $p = [p_1, p_2]$. Regional fuel consumption is denoted by $x_r = [x_{r,1}, x_{r,2}]$.

The fuel consumers' energy consumption utility, u, is isoelastic in the regional consumption of energy Y_r, $u_r = \frac{1}{1-\alpha}\xi Y^{1-\alpha}$, $\alpha > 0$, implying an isoelastic demand for energy Y, with elasticity $\varepsilon = -\frac{1}{\alpha} < 0$, where ξ may grow over time. Energy Y is the sum of a constant elasticity of substitution (CES) aggregation of oil and coal consumption, $F(x_1, x_2) = \left(ax_1^\delta + (1-a)x_2^\delta\right)^{1/\delta}$, implying an elasticity of substitution $\sigma = 1/(1-\delta)$, plus, if allowed for, the consumption of a clean backstop B, $Y = F + B$. The backstop may be provided at any given demand rate (i.e., we assume an infinite supply elasticity) for an exogenous price which may vary over time.

As a property of the CES aggregation function, the unitary fuel aggregate cost, c_F, is $c_F(p_{x_1}, p_{x_2}) = \left(a^\sigma p_{x_1}^{1-\sigma} + (1-a)^\sigma p_{x_2}^{1-\sigma}\right)^{\frac{1}{1-\sigma}}$, and, for a specific aggregate fuel consumption F, oil and coal consumption becomes $x_1 = F \cdot \left(\frac{ac_F}{p_{x_1}}\right)^\sigma$, $x_2 = F \cdot \left(\frac{(1-a)c_F}{p_{x_2}}\right)^\sigma$. Supply

[24] Energy supply is responsible for 83 % of all anthropogenic greenhouse gas emissions (IEA, 2012).

[25] In a similar fashion, climate and energy CGE models tend to treat oil and gas as a separate constant elasticity of substitution (CES) sub-aggregate, nested under another CES aggregate in which the oil-gas sub-aggregate figures parallel to coal or even to different types of coal; see, e.g., Böhringer and Löschel (2004) and Böhringer et al. (2008).

of both, the aggregate fuel and (if not idled) the clean backstop, is readily modeled with a complementary slackness condition with respect to the weakly positive difference of their cost to the overall energy price p_Y,

$$
\begin{aligned}
F \geq 0 \quad &\perp \quad p_F - p_Y \geq 0 \\
B \geq 0 \quad &\perp \quad p_B - p_Y \geq 0.
\end{aligned}
$$

Because the regional calibration of the demand structure to today's regional oil and coal consumption rates implies regional share parameters a in the CES fuel aggregation function, the *aggregate* fuel energy price will differ across the regions, wherewith also the time of the introduction of the backstop will not exactly coincide in the two regions.[26]

The model can be run with an endogenous production of synthetic oil derived from coal by liquefaction as soon as the relative fuel prices make the process economic, given a specified overhead process cost and conversion efficiency, again implemented by a corresponding complementarity slackness condition analogous to that given for the backstop above, and assuming the synthetic fuel to be produced competitively.

Suppliers are assumed to sell their fuels on the international market under perfect competition. The exhaustibility of the fuels is modeled with an extraction cost curve that indicates the marginal cost of extraction after a specific cumulative amount of the fuel has been extracted. This depletion concept is the logical consequence of the Herfindahl rule (Herfindahl, 1967) which states that (given positive real-interest rates) profit-maximizing resource owners extract the fuels ordered in a sequence according to extraction costs: the resources with the lowest extraction costs are extracted first, and the ones with the highest extraction costs are extracted last.[27]

In addition to the 'pure' extraction costs, the model optionally allows accounting for an additional, extraction *rate* dependent cost component k that is increasing in the extraction rate r as well as in the difficulty of the access to the reserves, indexed by the overall advancement of extraction, A, $k(r, A) \geq 0$, $k_r(r, A) > 0$, $k_A(r, A) > 0$. The current-value Hamiltonian for the profit maximization problem for the fuel owners reads

$$
\begin{aligned}
\mathcal{H} \quad &= \quad r_t \cdot (p_t(r_t) - c(A_t) - k(r_t, A_t)) - \lambda_t r_t \qquad (2.1) \\
&\text{s.t.} \quad \dot{A}_t = r_t \text{ and } A_0 = 0, \text{ i.e. } A_t = \int_{s=0}^{t} r_s \mathrm{d}s,
\end{aligned}
$$

[26]In all model runs, the introduction-time of the backstop does, however, differ only to a limited extend across the regions.

[27]While in a simple theoretical framework this rule should hold not only in a monopolistic but also in a competitive framework, e.g. Beermann et al. (2011) give reasons why this rule is often only an approximation to reality.

where r_t is the amount of the fuel extracted at time t, A_t is the cumulative amount of the fuel extracted from the initial period up to time t, normalized to 0 for $t = 0$, $c(A)$ is the marginal extraction cost after the extraction of the A units of fuel that could be extracted at the lowest costs, and $p_t(r)$ is the inverse demand for the considered fuel at time t. The price p_t results on the international fuel market if r units of the fuel are supplied (with non-zero cross-price elasticities of fuel demand the price p_t may depend also on the amount of the other fuel supplied at time t). This model corresponds very closely to the framework within which the Green Paradox has been proposed and studied on a purely theoretical level (except that increasing extraction costs, the backstop and the possibility of liquefaction or of a regional limitation on the tax may have been abstracted from). Rate-dependent costs have generally been absent in the Green Paradox literature. As we will see, whilst not necessarily affecting the core results with respect to the judgment on the relevance of the Green Paradox, this additional cost component can reconcile the model predictions from the basic Hotelling framework with the currently observed fuel consumption rates and prices.

The FOCs of the Hamiltonian in Eq. (2.1) yield the following stationary condition and canonical equation:

$$\begin{aligned} \frac{\partial \mathcal{H}}{\partial r_t} = 0: &\quad p_t(r_t) \overset{!}{=} c(A_t) + k(r_t, A_t) + \lambda_t \\ \dot{\lambda}_t = \rho_{\text{res}}\lambda_t + \frac{\partial \mathcal{H}}{\partial A_t}: &\quad \dot{\lambda}_t \overset{!}{=} \lambda_t \rho_{\text{res}} - \dot{c}_t - r_t k_A(r_t, A_t), \end{aligned} \tag{2.2}$$

where we define $c_t \equiv c(A_t)$,[28] and λ_t is, at time t, the shadow value for a marginal unit of resource stock after the cumulative extraction of A_t previous units. As the stationary condition (Eq. (2.2)) shows, the resource shadow value is the difference between the price that the resources achieve on the market and the extraction costs, that is, the per-unit resource rent received by the resource-owner for sales at time t. Strictly speaking, the representation of rate-dependent extraction extra-costs, and their expression in Eq. (2.2), is consistent for the case of multiple, heterogeneous reserves offered competitively only in the sense of an additional cost that depends on the *global* extraction rate, which could theoretically be motivated by, e.g., higher prices of globally supplied, extraction related capital or services. It may, however, also be considered an approximation to the case of geology based field-specific extra costs.

[28]Note that as $r_t = \frac{\partial A_t}{\partial t}$, we have $\dot{c}_t \equiv \frac{\partial c(A_t)}{\partial t} = \frac{\partial A_t}{\partial t}\frac{\partial c(A_t)}{\partial A_t} = r_t \frac{\partial c(A_t)}{\partial A_t}$.

2.2.2 Calibration

The following presents the standard setup of the model used for the simulations, as well as how the major model extensions (liquefaction, rate-dependent extraction costs) are implemented. Various additional setups are considered in section 2.4.3.

The fuel demand (or utility) parameters are calibrated according to the current regional consumption of oil and coal at current prices in the OECD and the Non-OECD region and the desired direct- and cross-price elasticities of the demand.[29] Similarly to Golombek et al. (1995), a demand elasticity slightly below unity and a weak substitutability of the fuels are chosen in the main calibration, setting $\varepsilon = -0.9$ and $\sigma = 1.2$. The weak substitutability between oil and coal expresses mainly the difficulty of replacing oil, in its major applications,[30] by the solid fuel coal.[31] The possibility of deriving synthetic oil from coal liquefaction (also called the coal-to-liquids, CTL, process) is modeled as a separate process.

The curve of the extraction cost for oil as a function of cumulative extractions is implicitly defined through its inverse, the function giving the cumulative extraction A for a specific marginal cost, $A(c)$. The following functional form appears to allow a very good fit to the extraction cost curve by Rogner (1997),

$$A(c) = p_3/(1 + \exp((c - p_1)/p_2)) + p_4 \exp((c - 45)/p_5).$$

We thus calibrated the parameters p_1 through p_5 to the extraction cost curve by Rogner (1997). As Fig. 2.1 shows, this fit provides an almost perfect correspondence to the Rogner curve. Moreover, there is a very good correspondence between the more recent IEA (2008) and an acceptable correspondence to the IEA (2005) projections.[32]

Indicating extraction costs for up to 1740 gigatons (Gt) coal, the coal cost curve in Rogner (1997) covers only a relatively modest fraction of the totally estimated resources of 16 000 Gt coal (DERA, 2012). Moreover, as Rogner notes, he models coal reserves in less detail than oil, which likely is a reason for the roughness of his estimated cost

[29] The clean backstop is considered as absent or prohibitively expensive at this stage.

[30] As oil reserves are much more restricted than coal reserves, throughout our model simulations it will essentially be oil whose scarcity becomes relatively stronger, implying that the possibility of substitution of oil by coal is of relevance, rather than the inverse.

[31] Golombek et al. (1995) used -0.9 for the direct price elasticity for the fuel consumptions in the OECD and -0.75 for the ROW, and they used cross-price elasticities of 0.1, on average. In this study, the (in absolute terms) larger demand elasticity in the Non-OECD region represents the interpretation that as economies of the developing countries progress over time, their fuel demand structure may approach that of the developed countries. In an overview, Michielsen (2011) lists cross-price elasticities from various empirical studies, averaging 0.06 from oil (and gas) to coal and 0.12 from coal to oil., supporting the view of limited substitutability of the fuels.

[32] All curves are inflation adjusted to \$2012.

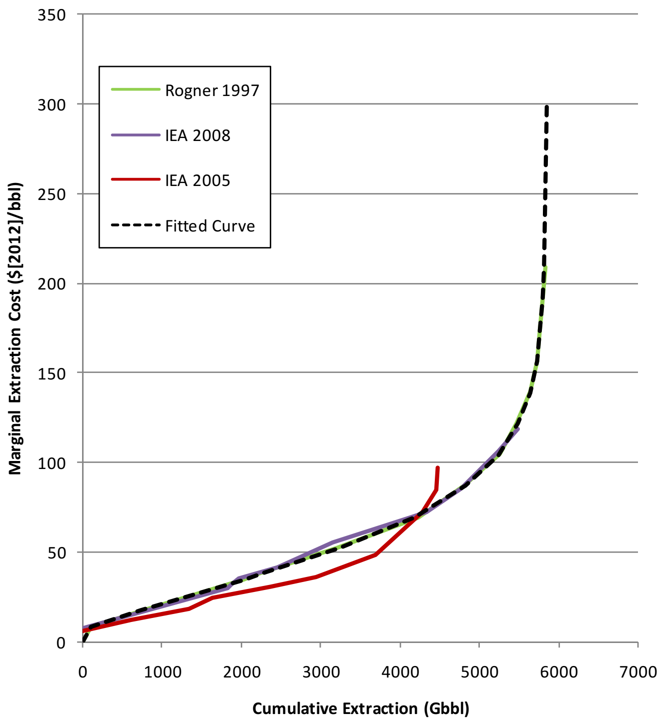

Figure 2.1: Oil extraction cost curves

curve, replicated in Fig. 2.2. Given that historically coal prices were relatively low, around 30 $/t in 2000, and today they fluctuate around 100 $/t (DERA, 2012; EIA, 2013a), with a relevant fraction of the currently rather high prices probably explained by the unprecedented growth of worldwide coal consumption in the current millennium[33] rather than a genuine long-term extraction cost increase, it was here decided to consider an actual coal extraction cost of 50 $/t, and to assume an exponentially increasing extraction cost curve that matches the extraction 'cost and cumulative amount'-data pair for the largest quantity considered in Rogner (marginal cost of 286 $/t after 1740 Gt extracted). This cost curve is given by $c = 50\,\$/\mathrm{t}\, e^{A/996\,\mathrm{Gt}}$. Fig. 2.2 shows how this curve provides a compromise between the general idea of a smooth, convex extraction cost curve, and the data points from the rough, convex and concave projection of Rogner (1997). Rate-dependent extraction costs are idled in the standard setup.

To be conservative in terms of refuting the Green Paradox, we assume a modest revenue discount rate for the fuel owners of $\rho = 3\,\%$.[34]

[33]Worldwide coal consumption used to stagnate before the beginning of this millennium, with annual growth rates averaging -0.3 %. The dash for coal, notably in Asia, has lead to an average coal consumption growth rate of 4.6 % per year from 2000 through 2011 (own calculations based on EIA, 2013a).

[34]Given historic real interests of often 5 % or above (see, e.g., Nordhaus, 2008, for an overview and discussion), plus conceivable substantial time-discounting due to political stability and property right issues in many fuel rich regions, this value is a conservative (low) estimate of the earning-impatience

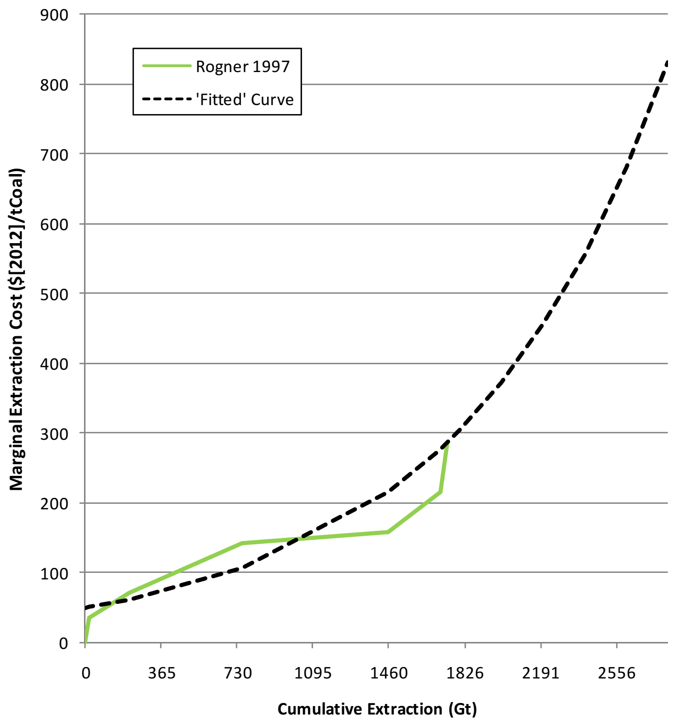

Figure 2.2: Coal extraction cost curves

The emission intensity is $0.43\,tCO_2/bbl$ for genuine oil and $2.8\,tCO_2/t$ for coal.

In the scenarios that consider liquefaction, the process is assumed to require 1 ton of coal per 2 barrels (bbl) of synthetic oil produced (DOE/NETL, 2006; Bartis et al., 2008).[35] Whilst the final product, the synthetic oil, has the same emission intensity at its final consumption as genuine oil (that is, in the domestic use direct emissions are the same), the use of half a ton of coal per barrel of oil produced, implies that excess emissions occur during the production (and thus, abroad) that exceed the final consumption emissions. In other words, overall synthetic fuel is more than twice as emission intensive. In addition to the input costs for this coal, the process is assumed to be subject to a constant additional fixed cost for each barrel of synthetic fuel produced, c_l, assumed to amount to c_l=15 \$/bbl (of produced synthetic oil). In the simulations that allow for liquefaction, the overall costs of the process represent an upper bound for the oil sales price such that any demand that cannot be met by the standard oil supply for that price will be provided as synthetic fuel from coal-liquefaction.

the major oil owners may exhibit. As the Green Paradox can only occur if the tax rate of increase exceeds the oil owner discount rate by a large enough margin, higher discount rates would tend to reduce the likelihood of a Green Paradox to occur.

[35] In reality, the conversion factor depends on the type of coal used. While a rule-of-thumb estimate for the coal-to-liquids yield from bituminous coal is 2 (barrels of oil per ton of coal), it is slightly lower for subbituminous coal, about 1.8 (Bartis et al., 2008).

When the clean backstop is considered, its price expressed in dollars per barrel of oil equivalent units of energy (bble) is assumed to approach $p_{b,\infty} = 200\,\$/\,\text{bble}$, with an initial price starting at $p_{b,\text{ini}} = 500\,\$/\,\text{bble}$ and the difference decaying exponentially at an annual rate of 2 %, as illustrated in Fig. 2.4 on p. 65. Nordhaus (2008) is more optimistic with respect to the possibility of replacing the fossil fuels. He assumes a backstop for 1170 \$/tC (319 \$/tCO_2) initially, converging to half of that price in the long run. In bble emission units this corresponds to an initial price of 137 \$/ bble. This price seems low as existing alternative technologies seem hardly competitive on a large scale even if oil prices currently tend to partly exceed 100 \$/bbl.[36]

Current fuel demand is calibrated to current regional consumption according to IEA World Energy Outlook 2010 data (IEA, 2010) of 16.4 and 14.3 bio. barrels (Gbbl) of oil per year in the OECD and the rest of the world (ROW), and coal consumption to the current 1.61 and 3.12 bio. tons (Gt) annual coal consumption in these two regions, at prices of 76 \$/bbl for oil and 83 \$/t for coal, according to World Bank (2011) Pink Sheet Data. Based on projections of the IEA World Energy Outlook 2009 for their reference scenario (IEA, 2009), we consider a scenario where fuel demand is constant in the OECD and growing at 2.6 % p.a. in the ROW during the first 25 years. In the long run we assume that after the first 25 years, as the ROW economies are maturing, their energy demand *growth rates* slowly decline over time by 0.05 % p.a., until the time the economies reach a state where (implicit) energy efficiency improvements set off any final demand increases; from then on the energy demand growth rate is zero.[37] Fig. 2.4

[36]Similarly, it is not clear whether a carbon price of some 300 \$/tCO_2 would enable the rapid and general adoption of carbon capture and storage especially on the large fraction of non-stationary or finely distributed emission sources such as in transport or residential consumption, with current technologies.

[37]In the World Energy Outlook 2009 reference scenario lasting through 2030, oil consumption is assumed to decline by 0.3 %p.a. between 2008 and 2030 in the OECD, while it increases by 2.3 % p.a. in the ROW (IEA, 2009, p. 81, Table 1.3). Correcting these *consumption* changes for the average annual oil consumption changes during the same period in a fuel market simulation with *constant demand* (these changes are -0.3 % p.a. in the OECD and -0.2 % p.a. in the ROW) to approximate *demand* changes, we find a constant oil *demand* in the OECD, and an increase in the ROW of 2.5 % p.a.

In the same World Energy Outlook scenario, coal consumption declines by 0.2 % p.a. between 2007 and 2030 in the OECD and increases by 2.8 % p.a. in the ROW (IEA, 2009, p. 90, Table 1.5). Correcting these *consumption* changes for the average annual coal consumption changes during the same period in our standard model with *constant demand* (these changes are +0.2 % p.a. in the OECD and -0 % p.a. in the ROW) to approximate *demand* changes, we find coal *demand* in the OECD changing by -0.4 % p.a., and increasing by approximately 2.8 % in the ROW.

Thus, in each region, average demand growth rates are very close to each other across the fuels during the period from 2007/2008 through 2030 and we approximate them by assuming a constant demand for both fuels in the OECD and an annual growth of 2.6 % for both fuels in the ROW. Note that for the coal demand in the OECD, the difference between our assumption (0 %) and what the World Energy Outlook data implies (0.4 %) is smaller in the medium-run than what the cited numbers suggest on first sight: OECD consumption in the World Energy Outlook is assumed to slightly decrease only until 2015, and from then on the projected consumption change is already approximately zero on p. 65 plots the demand growth path.

In the standard model runs, the backstop turns out to outcompete the fossil fuels within the next 200 years. We therefore consider a model horizon of 200 years in the scenarios with a backstop. In the scenarios without backstop, a simulation period of 400 years is used, in order to cover the period with an interesting dynamics, before the fuel consumption becomes low and largely irrelevant in NPV terms. The impact of the emissions on the climate damages is always modeled for 600 years.

When the extraction rate-dependent extra costs $k(r, A)$ are considered, they are assumed to be of the form $k(r_t, A_t) = \kappa \frac{r_t}{S_t}$, where S_t is the amount of the fuel remaining underground at time t, $S_t \equiv S_0 - A_t$, where we assume the amount of ultimately, theoretically extractable fuel, S_0, to amount to 2795 $GtCO_2e$ for oil and 42 000 $GtCO_2e$ for coal – closely corresponding to BGR (2009b) resource estimates as well as to approximately the amount of oil that may be exploitable for limited costs according to the extraction cost curve in Fig. 2.1. The parameters κ are calibrated such that the initial prices and regional extraction rates are within a 10 % range of today's oil and fuel prices and consumption.[38]

Nordhaus damages We examine the case for the Green Paradox in two ways. NPV emission changes resulting from a tax indicate whether the tax increases the value of the current emissions if a delay of emissions is valued directly using a specific emission time-discount rate. Damage changes, again present discounted, indicate whether the tax increases the climate damage if the emissions affect the climate and the economy according to the climate and damage module in the integrated assessment model from Nordhaus (2008). The damage flow is convex in the increase of the atmospheric temperature (roughly proportional to its square). The atmospheric temperature evolves dynamically as a result of net radiative forcing and heat exchange with the oceans, and the net radiative forcing is mainly driven by the atmospheric carbon concentration and the current temperature. The emitted carbon transfers between the atmosphere and the upper and lower ocean layers. In addition to the oil and coal emissions, anthropogenic emissions also comprise approximated gas combustion and land use change emissions; however, they are overall of limited importance compared to those from coal and oil combustion. We here model the gas combustion emission as starting at today's observed rate, evolving proportionally to the evolution of the sum of oil and coal emissions rel-

until 2030.

[38]For the model run used, with a 400 year simulation horizon and no backstop, this implied a value of $\kappa_{\text{oil}} = 8250\,\$\,\text{yr}\,\text{bbl}^{-1}$ and $\kappa_{\text{coal}} = 63\,500\,\$\,\text{yr}/\,\text{t}^{-1}$, yielding initial extraction rates of 32.2 Gbbl/yr for oil and 4.9 Gt/yr for coal, and initial extra extraction cost components of $k_{1,\text{oil}} = 41.7\,\$/\,\text{bbl}$ and $k_{1,\text{coal}} = 20.9\,\$/\,\text{t}$.

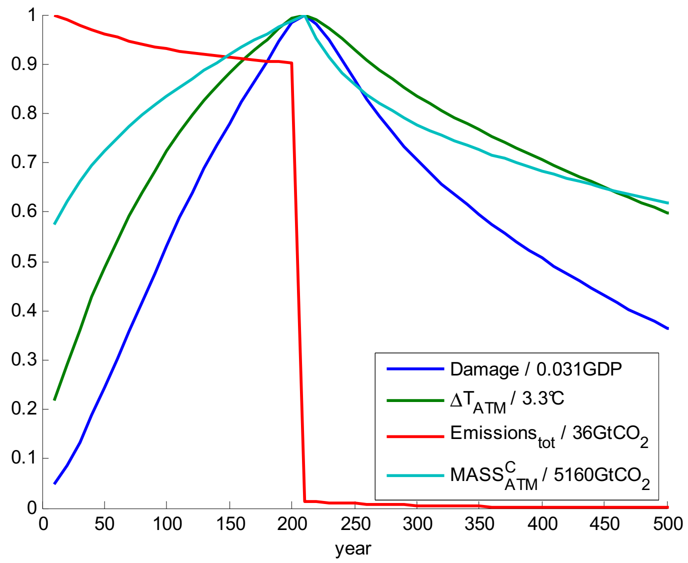

Figure 2.3: Illustration climate damages

Climate impact based on carbon cycle, temperature dynamics, and damage function as inNordhaus (2008).

ative to their initial value. A detailed report of the climate and damage equations is given in the Appendix of Nordhaus (2008).[39] For a given atmospheric temperature increase (compared to the level from 1900) T in °C, the current economic losses D are given by

$$D_t = \frac{0.0028\,T_t^2}{1 + 0.0028\,T_t^2}\Pi_t, \tag{2.3}$$

where Π is the global GDP. Fig. 2.3 illustrates the behavior of the key climate and damage variables for a constant GDP and fossil fuel emissions sustained at today's rate for 200 years and dropping to zero afterwards (land use change emissions are assumed to amount to 4.0 $GtCO_2$ per year initially and decline over time).

When calculating the relative NPV of the future damages, we do this estimation without augmenting the future damages by a potential GDP growth rate, that is, we neglect Π_t in Eq. (2.3). This provides a conservative result in terms of refuting the Green Paradox: as the tax generally reduces future damages but may increase initial ones, taking into account the fact that future damages are relatively larger since they are, in the Nordhaus model, proportional to the GDP would imply that the taxes are in general even more beneficial for the climate than calculated here, for given social discount rates, or, as explained in section 2.3, that the appropriate social discount rate to be applied

[39] The parameterization, which is also adopted here, can be found at http://www.econ.yale.edu/~nordhaus/DICE2007_programs/DICE_delta_v8_071707.GMS (accessed 27.2.13).

to the indicated damage path is smaller than the general social discount rate for the present discounting of *absolute* damages.[40]

As Fig. 2.3 illustrates, in the Nordhaus climate model, the CO_2 concentration, and along with it the temperature, tend to increase over a long time span even if the emission rate decreases over time, indicating a slow rate of decay of CO_2. As the damages are convex in the temperature this suggests that considering marginal damages instead of direct emissions, gives relatively more weight to future emissions. This suggests that when the emissions are mapped into the relevant damages, later emissions get a higher weight relative to earlier ones, and that the threshold discount rate required for the future climate benefits to be offset by potential early climate costs of the tax (cf. section 2.3), ρ^*, is larger when damages are considered than when directly the emissions are considered, that is, $\rho^*_d > \rho^*_e$. The results in section 2.4 confirm this prediction.[41]

Tax In order to be rather conservative in terms of refuting the Green Paradox, we assume, in the main scenario, a tax that respects all the salient tax-characteristics for which the Green Paradox has been brought forward: the tax starts at a very low initial level of $\tau_0 = 3\,\$/\,tCO_2$ and rises at a rapid pace, with a rate of increase of g_τ=6 %p.a. (all rates in this paper are indicated in per-year terms and we hereafter omit the *p.a.*).[42]

It seems unrealistic that the tax would keep rising infinitely beyond levels that could be justified by climate concerns. Instead, once the tax reaches or surpasses a level corresponding roughly to the perceived social cost of carbon emissions, even revenue-seeking governments will in general be better advised to consider increasing broader-based value-added taxes or income taxes rather than those on the limited residual

[40] For a unitary elastic (logarithmic) utility function, a GDP growth (if, as an approximation, assumed homogenous across a constant population) would not directly affect the utility losses due to the damages which are given as a fraction of GDP. In this case, the damage discount rate applicable to the here indicated results, ρ_d equals the pure-time preference and does not contain any growth related part. Cf. also the discussion of the discount rate in section 2.3.

[41] Depending on the details of the climate dynamics and the emission path, the accumulation and slow decay of the emissions in the atmosphere, coupled with a decreasing rate of emissions over time could in theory also have the inverse effect: if emissions are high initially but low later on, and the decay is not too slow, convex costs from temperature increases could imply that the dominant determinant of the aggregate climate costs is the maximal temperature achieved in the short-run or in the medium-run, and it could become primordial to delay parts of the early emissions and make them occur in times where the temperature has already declined. The simulation results suggest that this is not the relevant case in the scenarios considered.

[42] I consider a rate of increase of 6 % as one, one might choose when trying to 'elicit' a materialization of the paradox. A value much below 6 % seems less likely to lead to the paradoxical counterproductive results because it could be only marginally above the fuel owners' discount rate. A value significantly above 6 % would likely lead the tax level to 'explode' so rapidly after the tax introduction that the fuel owners would barely have time to sell a large fraction of their fuels before the tax reached prohibitive levels, making a relevant, i.e., strong, Green Paradox very unlikely as well.

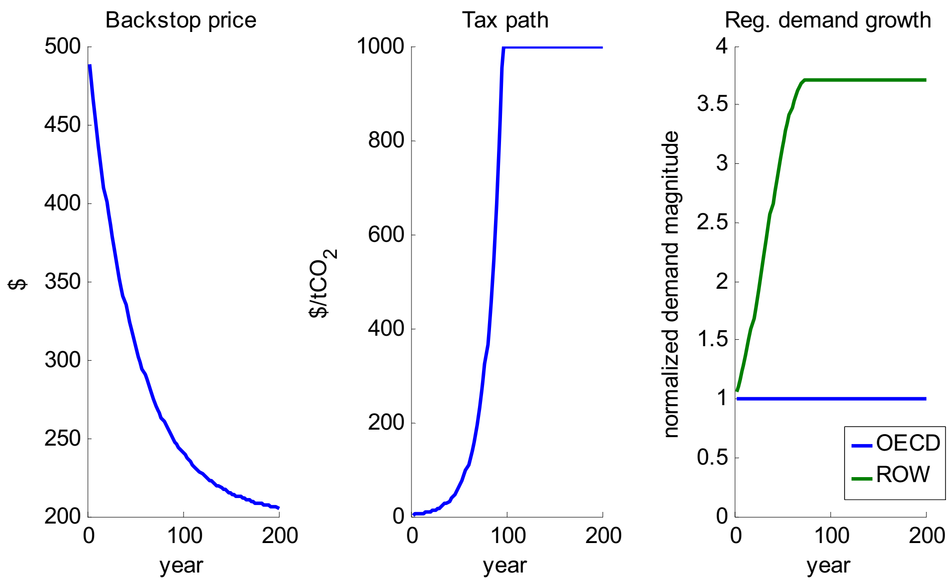

Figure 2.4: Backstop price, tax, and demand growth

demand of fossil fuels where the deadweight loss of additional net tax revenue would rise ever more rapidly as the increasing tax erodes the tax base and drives a large wedge between the social cost and the consumer cost of the fuels. Estimates of the social cost of carbon dioxide emissions are typically in the order of magnitude of up to a few hundred dollars per ton of CO_2.[43] In order to allow the tax level to exhibit its steep rate of increase for a long time, but to nevertheless account for the implausibility of a perpetual rise at a rate above the interest rate, the tax is thus here capped at a relatively high level of $\tau_{up} = 1000\,\$/\,tCO_2$; once the exponentially increasing tax has reached this value, it is held constant for the remainder of the simulation horizon.[44] It turns out that in the baseline calibration of the model, this upper limit is irrelevant for the model outcome, as the backstop replaces fossil fuel energy somewhat before the tax reaches the threshold value. Furthermore, alternative model runs provide results for lower and, for the case without the backstop, higher tax thresholds. The tax path of the main calibration is plotted in Fig. 2.4.

The extensive sensitivity analysis (see chapter 5) shows that the model results are

[43] According to a meta analysis by Tol (2008), the bulk of existing estimates for the social cost of carbon lies below $100\,\$/\,tCO_2$ for current emissions. This social cost of the emissions is, however, typically found to rise over time in the studies. For example, in his book, Nordhaus (2008), whose cost estimates tend to be low in comparison to estimates by other authors, indicates, as upper estimates of potential social costs in a no-policy business-as-usual scenario, a value of $12\,\$/\,tCO_2$ for the year 2005, but a value of $95\,\$/\,tCO_2$ in 2105.

[44] The $1000\,\$/\,tCO_2$ may on first sight seem like a large value. However, already today, even higher values are occasionally debated in political debates, such as is the case of the studies Ecoplan (2012) and Egger and Nigai (2013), who consider a tax of $1140\,\$/\,tCO_2$ necessary for meeting politically discussed carbon dioxide emission reduction targets. A tax of $1000\,\$/\,tCO_2$ corresponds to 2.3 $ per liter of gasoline.

largely robust to many of the parameter assumptions, including the level and the form of the tax path.

2.3 Note on Interpretation of Results

Since the Green Paradox has essentially been brought forward as a theorem on negative consequences for the environment rather than directly for overall welfare, the attention is in this section largely restricted to green-only welfare. That is, no strict statements will be made on the overall social desirability of the taxes investigated here. However, in the simplest possible world (notably with the same discount rate of the competitive fuel owners and of the fuel consumers), if a tax would indeed reduce green welfare, as postulated by the (strong) Green Paradox, it would also reduce overall welfare. This would be the case as the only economic justification for the tax could be that it reduces the pollution externality and else it essentially corresponds to an undesirable consumption distortion. If a tax alters the emission path in a way that increases the overall environmental quality, however, it is in general not necessarily clear to which extent the tax is also socially desirable, as the economic costs of the tax oppose the green welfare gains. Section 2.4.4 addresses welfare questions and provides a direct weighing of green and the more narrow economic welfare changes, but here we first strictly stick to the more narrow Green Paradox question. There are several reasons for this focus. First, Essay 3 in this dissertation considers welfare aspects of resource and climate taxes in an analytical and numerical framework that largely corresponds to the one used here. Furthermore, we think the green welfare-only results derived in this paper allow a relatively clear-cut answer to the question of the Green Paradox, whilst a definitive answer on the general social desirability of specific climate taxes hinges very strongly on the uncertainty and the controversy of the general severity of the damages suffered from different degrees of climate change. This difficulty to answer the question of the social desirability of a specific tax is also accentuated by the fact that those most vulnerable to climate change may be living in very poor regions, whilst mainly the wealthy produce the pollution. This separation makes the comparison of benefits and costs of climate mitigation for the real world especially complicated and prone to normative and political controversies, as will become clear in the welfare impact calculations in section 2.4.4.

The net present values (NPVs) of changes of emission E, and of damages D, induced by the climate tax,

$$\Delta E_{\text{NPV}} \equiv \Sigma_t (E_{t,\text{tax}} - E_{t,\text{no tax}}) \mathrm{e}^{-\rho_e t}, \text{ and}$$
$$\Delta D_{\text{NPV}} \equiv \Sigma_t (D_{t,\text{tax}} - D_{t,\text{no tax}}) \mathrm{e}^{-\rho_d t},$$

are calculated for a range of different emission and damage discount rates, ρ_e and ρ_d. As the initially low and rapidly rising tax will tend to increase emissions in early periods (that is, the 'weak' version of the Green Paradox holds), for large enough discount rates these NPVs will generally take on positive values, that is, for large ρ_e or ρ_d, overall climate change is considered be exacerbated by the tax. The tax will, however, generally reduce the absolute amount of cumulative long-run emissions, i.e., $\lim_{T\to\infty} \Sigma_{t=1}^{T} (E_{t,\text{tax}} - E_{t,\text{no tax}}) < 0$. For the results from the simulations, the same will generally turn out to be true for the summed climate damages. We thus have a positive climate effect of the tax for low enough emission or damage discount rates and a negative climate effect for large enough discount rates. In terms of refuting the Green Paradox, the crucial point is therefore whether the threshold discount rate for which the climate effect turns from positive to negative is large enough to be deemed unrealistically high.

Be thus ρ_e^* (or ρ_d^*) the emission (or damage) discount rates required for the initial emission (damage) increases to just offset the future, larger emission (or damage) reductions. That is, be ρ_e^* (or ρ_d^*) the threshold discount rate where higher discount rates $\rho_e > \rho_e^*$ (or $\rho_d > \rho_d^*$) imply that overall the tax introduction does exacerbate the climate problem in terms of how we evaluate it (i.e. there is a Green Paradox in the most important sense), and where lower discount rates $\rho_e < \rho_e^*$ (or $\rho_d < \rho_d^*$) imply that overall the tax does alleviate the climate problem (i.e., there is no Green Paradox in the most important sense). Analytically, we thus define

$$\rho_e^* = \{\rho_e \,|\, \Sigma_t (E_{t,\text{tax}} - E_{t,\text{no tax}}) \mathrm{e}^{-\rho_e t} = 0\}, \text{ and}$$
$$\rho_d^* = \{\rho_d \,|\, \Sigma_t (D_{t,\text{tax}} - D_{t,\text{no tax}}) \mathrm{e}^{-\rho_d t} = 0\},$$

and $\rho^* = \infty$ denotes the case where the tax increases the emissions (or the damage) for all positive discount rates.

An annual discount rate is a somewhat abstract measure. In order to get a better feeling about what the threshold discount rate signifies, the implied 'half-value time' of emissions, $t_{1/2}^* \equiv t_{1/2}(\rho_e^*)$, is reported: it expresses by how many years into the future emissions would have to be postponed for them to be valued only half as much as their amount in terms of emissions occurring today according to the discount rate, that is,

$t^*_{1/2} = \{t \,|\, (1-\rho^*_e)^t = 0.5\}$ or $t^*_{1/2} = \frac{\ln 0.5}{\ln(1-\rho^*_e)}$.

Given what is known about the extreme long-term characteristics of climate change,[45] it is here considered as unlikely that one judges the climate tax as detrimental for the climate if $t^*_{1/2}$ is less than 20 years. In other words, it is assumed to be accepted that any $t^*_{1/2}$ of less than 20 years would signify that the tax is beneficial in terms of climate change. Note that $t^*_{1/2} < 20\,\text{yrs}$ signifies $\rho^* > 3.47\,\%$. To consider such values as refutable is, of course, a somewhat subjective judgment. It is based on the belief that to hypothetically accept a doubling of emissions to delay them by less than 20 years would be considered as very impatient by most of those concerned about the long-term problem of climate change. Readers may of course make their own judgment and reinterpret the results reported below according to their own taste concerning the weighing of reducing and advancing emissions (or climate damages).

The standard method to present-discount emission damages in climate economics consists in combining a pure rate of social time preference, δ, expressing an intrinsic impatience, that is, the rate by which we (annually) discount what happens in the future for the mere reason that it does happen then instead of today, with an (economic) growth related second term expressing how much less severely a specific damage (output and consumption loss) will affect our utility in future when the marginal utility is smaller due to positive economic growth and decreasing marginal consumption utility. More precisely, the overall, so called social discount rate on the damages, R, is the sum of the pure time-discount rate, and the product $g_c\varepsilon_c$, with g_c the growth rate and ε_c the coefficient of relative risk aversion in the utility function. The here derived damages based on the Nordhaus climate damage module are given as fractions of the global GDP. To get the absolute amount of damage (expressed in terms of lost output) they must therefore be multiplied by the GDP, which implies that discounting the *absolute* damages at the rate $R = \delta + g_c\varepsilon_c$ is equivalent to discounting the here indicated damages as fraction of GDP at the rate $R_{\text{net}} = \delta + g_c(\varepsilon_c - 1)$. Taking the viewpoint of Nordhaus (2008) – an author famous and some times criticized for advocating high discount rates –, who assumes a medium-term growth rate of $g_c = 2\,\%$, a pure time-preference rate $\delta = 1.5\,\%$, and a consumption elasticity $\varepsilon_c = 2\,\%$, the net discount rate[46] to be applied to the here presented damage paths becomes $R_{\text{net}} = 3.5\,\%$. Alternative views favoring lower discount rates exist, as expressed notably in the Stern Review (Stern et al., 2006). There it is argued that morally, a low pure time discount rate, $\delta = 0.1\,\%$, is warranted,

[45] See, e.g., Solomon et al. (2009), for an illustration of the long-term character of the carbon emission induced climate change.

[46] Net in terms of accounting for the fact that we below compute damages as a fraction of GDP instead of in absolute terms, which therefore first would have to be augmented by the GDP before they ought to be discounted at the overall social rate of discount for the damages.

and a unitary consumption elasticity, $\varepsilon_c = 1$, is assumed. Therewith, the overall *net* discount rate for our purposes would collapse to a mere $R_{\text{net}} = 0.1\,\%$, given that the damages are expressed as a fraction of GDP. The Stern Review is sometimes criticized for assuming a social discount rate that is detached from the market interest rate even though in a simple economic model assuming complete markets the two rates do coincide (cf., e.g., Nordhaus, 2007, and **?**). However, there are compelling reasons why, in reality, an appropriate (long-term) social discount rate for climate damages might well be lower than the observed market interest (cf., e.g., Fleurbaey and Zuber, 2013, and Sterner and Persson, 2008). It is not the scope of this work to settle the debates around the discount rate. In any case, however, the conclusions from the results below do in general not hinge on whether a rather high net discount rate, such as the $R_{\text{net}} = 3.5\,\%$ derived above from parameters used by Nordhaus, or the very low net rate $R_{\text{net}} = 0.1\,\%$ derived from Stern's parameters, is used; as long as not rates very *much* beyond of what Nordhaus' view would imply are considered to aggregate the damage time paths, the main conclusions remain unchanged.

2.4 Results

In presenting the model results we proceed as follows: we start by illustrating the general model behavior (without any taxes). Then, in some detail, we show the effect of the tax on the emission path and on the net present discounted emissions and damages for the main model calibration. In the next subsection, an extensive robustness check is performed, confirming the robustness of the results beyond the particular main parameterization of the model.

2.4.1 Illustration of BAU simulations

The top left plot in Fig. 2.5 shows the fuel consumption path in the OECD and in the ROW. Demand growth initially leads to increasing oil and coal consumption in the ROW before the increased scarcity of oil (after around 50 years) and of coal (after around 70 years) reduces it. In both regions, the fossil fuels are replaced by the clean backstop when its price is reduced to around 230 \$/bble after around 120 years (the backstop price path is plotted in Fig. A2.1 in the Annex). The top right plot shows that the extraction costs (dashed curves) and the market prices of the fuels (solid curves) steadily increase over time, with the resource rent converging to zero as the time of the backstop emergence approaches. The bottom plots show cumulative fuel consumption (left) and the associated, instantaneous and cumulative emissions (right). Fig. A2.2

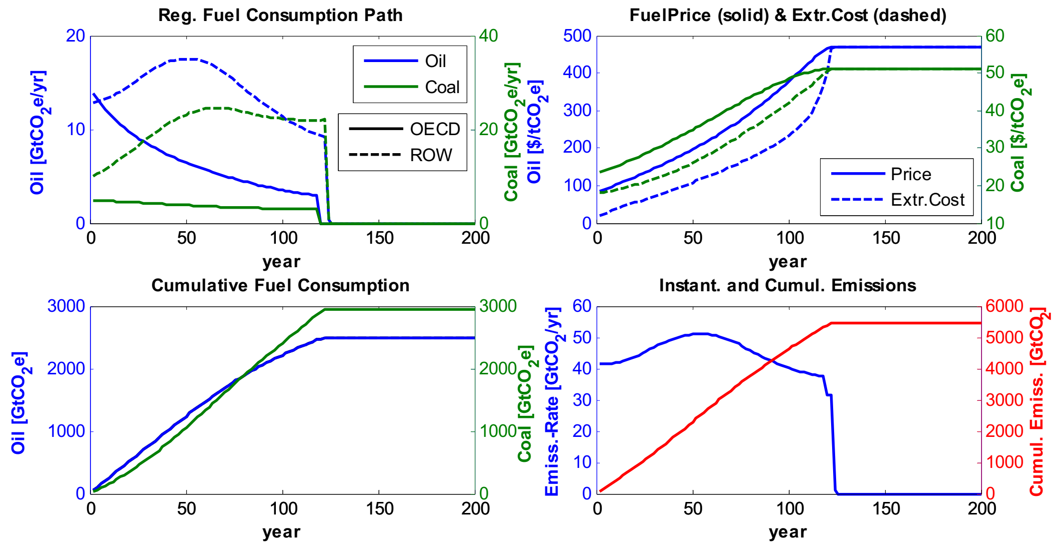

Figure 2.5: Simulation results base calibration

plots the same situation for a 400 year horizon simulation without a backstop. The climatic impact (carbon concentrations, temperature increase, and damages) is shown below in Fig. 2.7.

In the standard setting without rate-dependent extraction costs, the model projects current fuel prices that are lower than current real market prices. Correspondingly, initial consumption rates exceed the real values by a factor of around two. A major explanation for the difference between the modeled outcome and the real world may be that in reality the fuel owners cannot extract the fuels at any desirable speed without any extra effort. Indeed, the extension with an extra cost component that increases in the *rate* of extraction can be calibrated to yield today's real regional consumption rates in the simulation (cf. Fig. A2.3 in the Annex).

2.4.2 Tax Impact in Main Scenario

Fig. 2.6 shows, for the main setup without extra rate-dependent costs, the impact of the tax on the path of cumulative fuel consumption (lower plot) as well as the implied change of the NPV of both the emissions and the implied damages for a range of discount rates (upper plot). The upper plot shows that there is no strong Green Paradox even for discount rates up to 10 %. The NPV of emissions (solid red curve) or of the implied climate damages (dashed red) is lowered for all plausible discount ranges; the tax increases the NPV emissions only for discount rates of above 11.2 % and damages are reduced for even larger discount rates. The paths of the cumulative emissions in the lower plot explain this result: for oil, the initial emission increases are

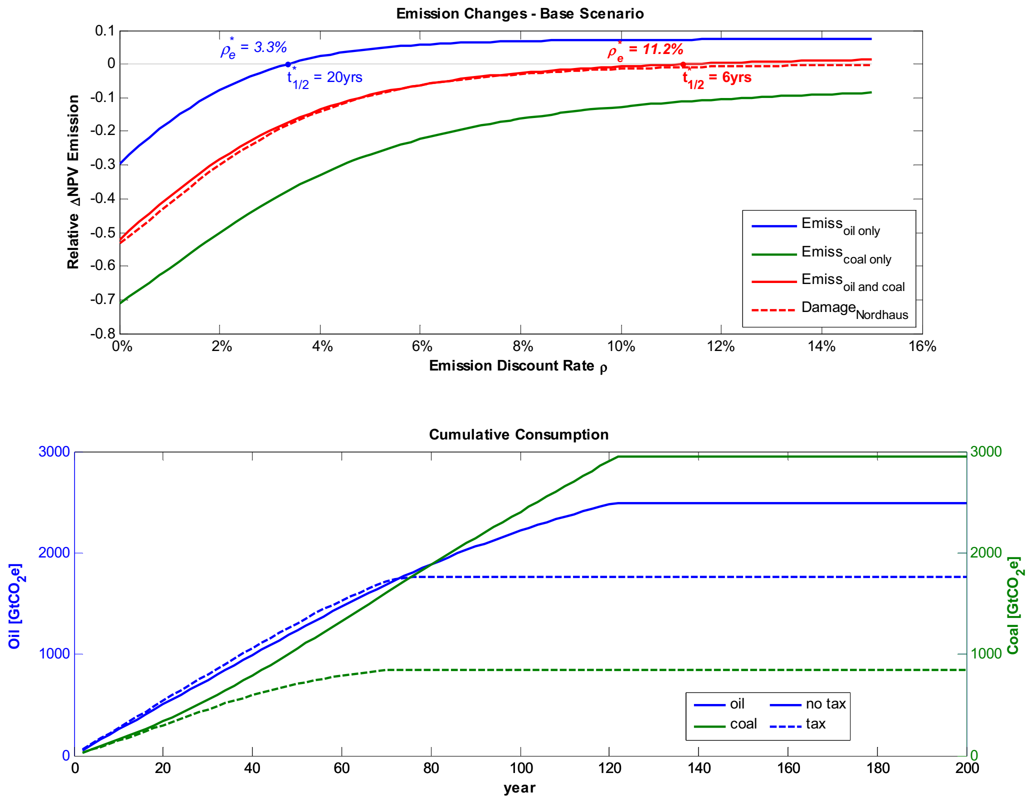

Figure 2.6: Tax impact

very small, and for coal emissions are reduced throughout the considered time horizon. The maximal cumulative emission increase from oil emissions amounts to 72 $GtCO_2$ and is outweighed by long-run cumulative reductions of 737 $GtCO_2$ for oil and more than 2000 $GtCO_2$ for coal. The reduction of coal use outweighs the increase of initial oil consumption very rapidly so that total cumulative emissions decrease already after 18 years. Therewith, for any arguably reasonable emission discount rate, there appears to be no 'strong' Green Paradox: the tax reduces the NPV emissions and damages.

Even if only oil – for which scarcity (and thus anticipation of future tax increases) plays a bigger role (cf. Fig. 2.5) and which is consequently is more likely subject to Green Paradoxes – is considered, emissions must be discounted strongly in order for the relevant Green Paradox to materialize ($\rho_e^* = 3.3\,\%$). Fig. 2.7 shows the estimated impact of the tax on the climate system and the corresponding economic damages. The tax reduces the maximal temperature increase by around a third, from 3.5 °C to 2.4 °C, and the damage by around 0.9 % of the global GDP to 0.8 % of global GDP, the peaks occurring around 2080 compared to 2140 in the BAU. Fig. A2.4 in the Annex shows additional details of the impact of the tax.

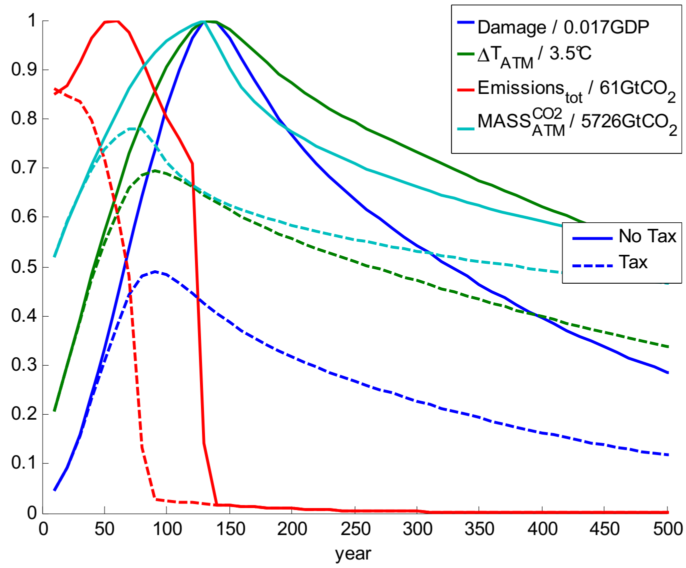

Figure 2.7: Details climate impact of the tax

Climate impact based on ten-year emission averages, and on carbon cycle, temperature dynamics, and damage function as in Nordhaus (2008).

2.4.3 Tax Impact Alternative Scenarios

This section examines the robustness of the main results with respect to departures from a number of key assumptions in the model. Table 2.1 lists the changes and their impact on the threshold half-value time and threshold discount factor for both, emissions directly (columns 4 and 5) and Nordhaus damages (last two columns). Fig. 2.8 plots the impact on the NPV of the emissions and of the damages for discount rates of up to 15 %, for all considered variants. Fig. A2.5 in the Annex shows the same plot for the NPV of changes in Nordhaus damages.

The names and descriptions of the variants listed in Table 2.1 are self-explaining; except for the indicated changes, the model (parameterization) used corresponds to the main setup from the previous subsection, explained in section 2.2.2.

The results (Table 2.1) are interesting in several ways. First, the only two variations which increase by a non-negligible amount the likelihood of a Green Paradox to occur, by reducing the threshold discount rate to values below $\rho_e^* \leq 6\,\%$ (red in Table 2.1), correspond to scenarios which are less realistic than the main scenario: real interest rates observed over the past decades, as well as uncertainty about property rights in many oil rich regions make oil owner discount rates of only 1.5 % or less annually (variant 11) highly implausible. Similarly, despite the substantial burdens regarding the introduction of a climate tax, an initial level of only 1 \$/ tCO_2 or less (variant 13) seems very low. Nevertheless, even in these two cases, the threshold discount rate

Table 2.1: Main results alternative scenarios

	Alternative setup	Emissions directly		Nordhaus damages	
		ρ_e^*	$t_{1/2}^*$	ρ_d^*	$t_{1/2}^*$
	Base	*11.2 %*	*6 y*	*>15 %*	*<4 y*
1	Rate dependent additional extraction costs	∞		∞	
2	Reduced upper tax limit $\tau_{up} = 300\,\$/\,tCO_2$	11.2 %	6 y	>15 %	<4 y
3	Cheaper backstop $p_{b,\text{ini}} = 300\,\$/\,\text{bble}$, $p_{b,\infty} = 100\,\$/\,\text{bble}$	13.9 %	5 y	∞	
4	No backstop	13.1 %	5 y	∞	
5	No backstop, reduced upper tax limit $\tau_{up} = 500\,\$/\,tCO_2$	>15 %	<4 y	∞	
6	No backstop, increased upper tax limit $\tau_{up} = 2000\,\$/\,tCO_2$	12 %	5 y	∞	
7	Low demand elasticity $\varepsilon = -0.4$	>15 %	<4 y	∞	
8	Complementary fuels, $\sigma = 0.3$	6.5 %	10 y	8.8 %	8 y
9	Strongly substitutable fuels, $\sigma = 1.8$	14 %	5 y	∞	
10	Liquefaction	9.4 %	7 y	14.7 %	4 y
11	Patient fuel owners, $\rho = 1.5\,\%$	4.2 %	16 y	5.7 %	12 y
12	Impatient fuel owners, $\rho = 5\,\%$	∞		∞	
13	Very low initial tax, $\tau_0 = 1\,\$/\,tCO_2$	6.0 %	11 y	7.8 %	9 y
14	Very rapidly increasing tax, $g_\tau = 8\,\%$	10.0 %	7 y	∞	
15	Less rapidly increasing tax, $g_\tau = 4\,\%$	>15 %	<4 y	>15 %	<4 y
16	Higher initial tax, $\tau_0 = 15\,\$/\,tCO_2$	>15 %	<4 y	∞	
17	OECD-only tax	>15 %	<4 y	∞	

Notes: Red emphasizes those variants for which the threshold emission discount rate ρ_e^* is substantially reduced, to values of 6 % or lower. Green emphasizes those variants for which the tax reduces the NPV emissions for the entire range of tested discount rates. '∞' indicates that the tax reduces cumulative emissions or damages for all periods.

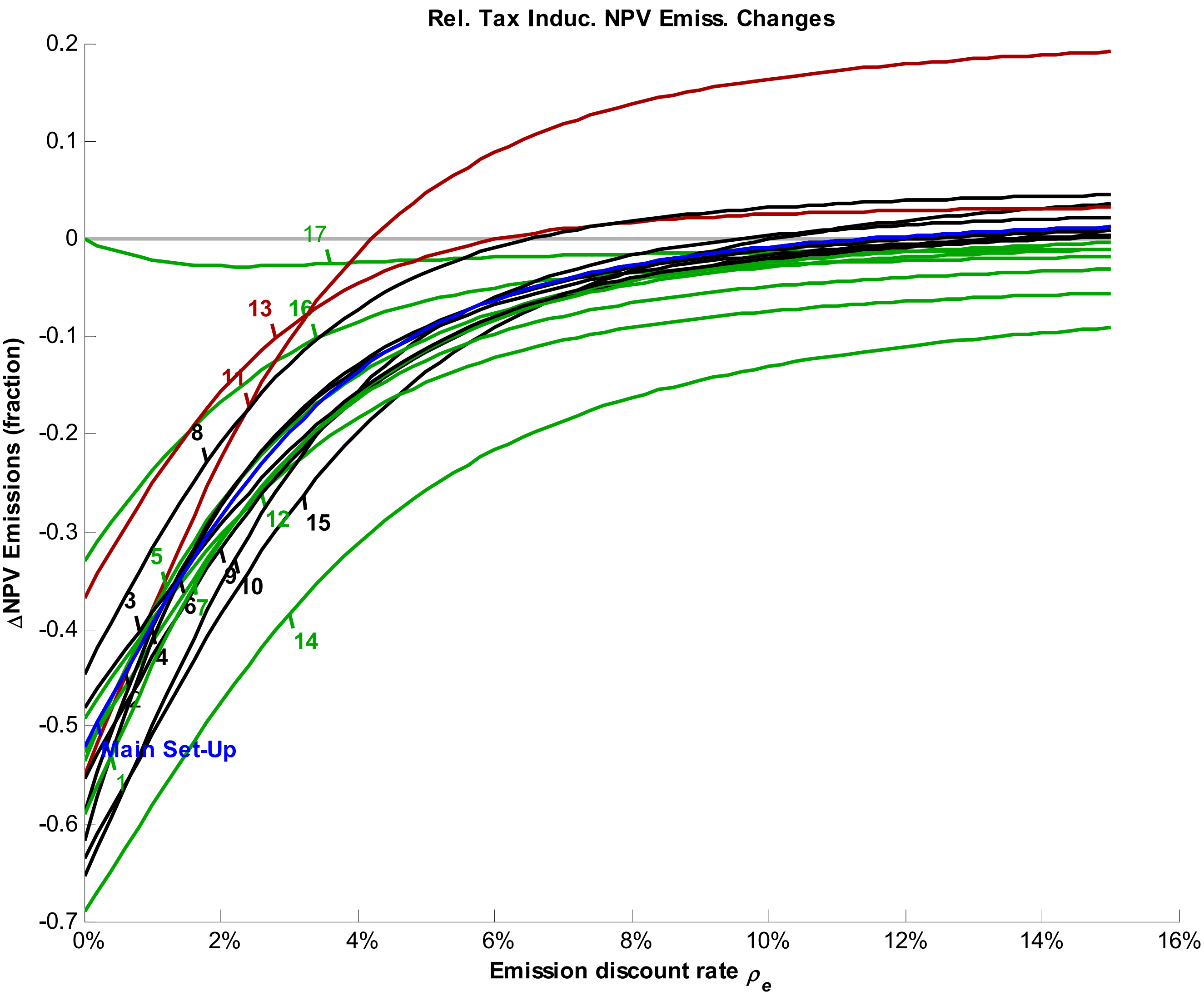

Figure 2.8: NPV emission changes alternative parameterization

Notes: Red emphasizes those variants for which the threshold emission discount rate ρ_e^* is substantially reduced, to values of 6 % or lower. Green emphasizes those variants for which the tax reduces the NPV emissions for the entire range of tested discount rates.

remains above a very high 4 % and the corresponding half-value times are 16 years or less, which, given the long-run processes of climate change, seems very short in terms of accepting a doubling of emissions in exchange of a delay by that period.

As a second observation, the convexity in the Nordhaus damage function increases the relative weight on the later emission reductions, increasing, as predicted, the threshold discount rate, $\rho_d^* > \rho_e^*$, in all variants for which emission and damage threshold rates are below 15 % (for the other variants we are ignorant of the exact values for the threshold discount rates).[47]

Third, assuming only the OECD imposes a tax (variant 17) affects the NPV of emissions in an interesting, non-monotonous way with respect to the discount rate (green line that almost converges to 0 for $\rho_e = 0$ in Fig. 2.8): the total amount of longer-run emissions is only slightly reduced by the regional tax (ΔNPV just below zero for $\rho_e = 0$). However, the tax does shift a substantial fraction of the medium-term emissions into the future, such that for intermediate discount rates of around 2 % the NPV of the emissions is reduced by a significant amount. In contrast to the cases of the worldwide tax scenarios, here the bulk of the fuels spared in the medium-run will generally not be saved forever. Instead it is consumed in the ROW later on. Overall, almost 100 % of the fuels saved in the OECD due to their regional tax will 'leak' to the remainder of the world.[48] In addition to this postponement of substantial amounts of the emissions the regional tax slightly increases the early emissions similarly to the case with the global taxes. Consequently, for extremely high discount rates, the early emission increases outweigh the delay of the medium-term emissions.

Fourth, allowing the extraction costs to increase not only in the cumulative extractions as in the standard setup, but as well in the rate of extraction (variant 1), makes the relevant Green Paradox less likely. A plausible explanation for this result seems to be that due to the 'convex instantaneous extraction costs'[49] the fuel owners' willingness to increase extractions in the early periods due to the future tax is dampened by the rapid decline of their net per-unit sales revenues that results from an increase in the fuel extractions during the initial periods.

Finally, the relation between the assumed fuel owner discount rate and the threshold discount rate below which the tax increases the considered green welfare offers a way

[47] Cf. section 2.2.2 for a rationale for this pattern.

[48] Cf. Essay 3 in this dissertation, as well as Habermacher (2012b) for a detailed discussion of the leakage effects responsible for that result. It is the minor negative time-derivative of the backstop price which implies that at least a tiny fraction of the medium-term domestic fuel savings translates into long-run global fuel savings.

[49] With the rate-dependent cost component, the extraction costs are convex in the sense that, for a specific period, the total cost associated to an amount extracted in that period is not linear but convex.

for a somewhat less normative, more positive interpretation of the results. In the base scenario and in the two variants with lower and higher fuel owner discount rates, these fuel owner rates ρ were 3 % and 1.5 % and 5 % respectively. The resulting threshold discount rates for emissions, ρ_e^*, (or for damages, ρ_d^*) were, in the same order, 11.2 % (or >15 %), 4.2 % (or 5.7 %), and ∞ (or ∞). Thus, the green discount rate would have to exceed the fuel owner discount rates very strongly in order for the strong Green Paradox to occur. This seems implausible not only in a simple world where borrowing between the fuel owners and the rest of the society would tend to align discount rates, but even more in a more realistic world where major fossil fuel owners face important expropriation risks, e.g., due to political instabilities, which may lift the discount rate of fuel owners to levels substantially above standard time-preference and interest rates used implicitly or explicitly elsewhere in the society.

In conclusion, for all at least more or less plausible variants examined,[50] the emission or damage discount rates required for a relevant Green Paradox to materialize are extremely high. The corresponding threshold half-value times are always substantially below 20 years (most often around 10 years or less), implying that the tax would exacerbate the climate problem only for emission time-preferences that seem myopic compared to the long-run problem of climate change; if one agrees that within a potential emission scenario it is probably undesirable to prevent some current emissions just to emit twice that amount before 20 years have passed, the carbon taxes would be desirable for the climate. Thus, in this sense there does not exist a (strong) Green Paradox for those taxes.

Moreover, as Fig. 2.8 shows, even for the variants where with extremely high discount rates the tax increases the NPV emissions, this increase is in general very small compared to the potential emission reductions resulting in the case of lower emission discount rates[51] (a pattern further accentuated when damages instead of emissions are considered, cf. Fig. A2.5 in the Annex).

Figs. A2.6 through A2.9 in the Annex plot detailed model and tax-impact results for the alternative model runs with the rate-dependent component in the extraction costs (variant 1), with the standard model without backstop (variant 4), with liquefaction (variant 10), and with an OECD-only tax (variant 17).

[50] We disregard only the case of the unrealistically low real revenue discount rate of 1.5 % for the fuel owners.

[51] To a certain degree, variant 11, based on the fuel owner discount rate $\rho = 1.5\,\%$, represents an exception to that rule. This fuel owner discount rate is extremely low and seems rather implausible.

2.4.4 Welfare Analysis

The analysis in the previous sections focused essentially on climate effects of the taxes. This does, however, not answer all questions about the welfare impact of the considered taxes. The overall welfare impact consists of the impact on the above discussed climate damage changes plus the changes of the net utility derived from the fuel-consumption. This section analyses this welfare impact of the taxes discussed above. It begins with a general discussion of what the resource extraction and taxation theory predicts regarding the welfare impact of the tax. It then shows the effect of the Green Paradox-type tax from the main calibration of the model with the Nordhaus climate damage module.

To understand this welfare impact, one has to bear in mind several basic effects of carbon taxation. First, for a given fuel (producer) price, and ignoring externalized pollution, the tax distorts the consumption of that fuel, reducing the consumption below the level which would equate marginal utility and the unitary fuel price. The tax therefore reduces the fuel consumption utility net of the fuel costs. However, except for this distortionary cost, the sum of taxes paid does itself not correspond to a direct loss, as it is recycled within the fuel consuming region. For the Green Paradox, the reaction of the producer price to the tax is, however, central. This makes it important to consider further effects. A second basic effect is that a coordinated consumption tax on scarce fuels allows an importer to extract parts of the fuel producer's scarcity profits. This rent shift, which corresponds closely to the case of the extraction of a monopolist's rents in the case of production with decreasing returns, opposes the distortionary costs of the tax, and as, for low taxes the latter are in general small in relative terms, a well conceived positive consumption tax can in general increase the (non-green) welfare of the fuel-importing region (cf., e.g., Dasgupta and Heal, 1979; Bergstrom, 1982; Kemp and Long 1980; Bergstrom et al., 1981). The dynamics of the Green Paradox and the corresponding suboptimal fuel tax has however implications that deviate from what has typically been considered in the literature on resource extraction theory about optimal resource taxes. More precisely, recall that we consider a tax that rises rapidly enough for a (weak) Green Paradox materializes in terms of increased *early* consumption. As we saw, this tax also tends to reduce consumption very strongly later on when the tax keeps rising rapidly during a longer period. As a third point, the here considered type of taxes influences the non-green welfare as follows: in an initial phase, as according to the Green Paradox the fuel price including the tax is *reduced* by the tax, non-green welfare is increased in two ways. On the one hand by the tax revenue and on the other hand by the lower gross fuel price. As has been pointed out in the previous sections, from a certain time on, the fuel consumption is reduced by the tax, which means that the producer price plus the tax is above the equilibrium price without tax. In the first years of this phase (immediately after the gross price with the tax has surpassed the price in the scenario without the tax) the net fuel consumption utility must still be increased, as the reduction of the producer price still almost fully compensates the tax, implying only a marginal reduction of the fuel consumption utility, but more than marginal tax revenues. As the tax rate increases at a high exponential rate, and especially as it grows at a higher rate than the resource rent, there may, however, come a time when the negative distortionary effects of the tax dominate the gains from the reduction of the producer price; that is, a time where the non-green welfare is lower with the tax than without. Nevertheless, even in this case, the fuel consumer welfare will, either forever or at least for a limited time, still be increased due to the reduction of the climate damages which opposes the direct economic costs of the tax. Should the tax become excessively large compared to the climate damage, it will, at some point, reduce the consuming region's welfare despite the climate gains. In presence of a backstop, the balance will change in favor of the tax for the remainder of the horizon at latest once the backstop (endogenously emerging but with an exogenous cost path) replaces the fossil fuels in the no-tax scenario;[52] from then on, the energy has the same price in both scenarios but the emissions stock, and thus climate damages, will be reduced with the tax. Similarly to the effect on the green-only welfare, it is not a priori clear to which extent the tax is beneficial for the overall welfare of the fuel-consumers. On the one hand, the tax allows the consuming regions to extract parts of the fuel resource rent, on the other hand its rapid rate of increase may bear excess costs as it disproportionally encourages early consumption and it can lead to rates that are beyond what can be justified on the basis of rent-extraction and climate protection.

Fig. 2.9 shows the impact of the tax on green-only, non-green, and overall welfare (and its annuitized value) for the case where the rents of the fuel owners are not taken into account, emphasizing the benefits of the majority of the fuel consuming countries when they extract scarcity rents from the fuel producers. The scenario is based on the main calibration of the model, where growth of the world economy is abstracted from, and where the backstop, with its exogenously decreasing price time path, becomes competitive against the depleting fuels in around year 120 in the BAU scenario. For the aggregation of the welfare impact over the different periods we assume a social discount rate of 3 %, corresponding to the assumed producer revenue discount rate in the main scenario.

As Fig. 2.10 shows, the tax reduces the producer price of the fuels throughout time

[52] As the backstop price decreases smoothly over time, this must happen already some time before the backstop really has emerged in the no-tax scenario.

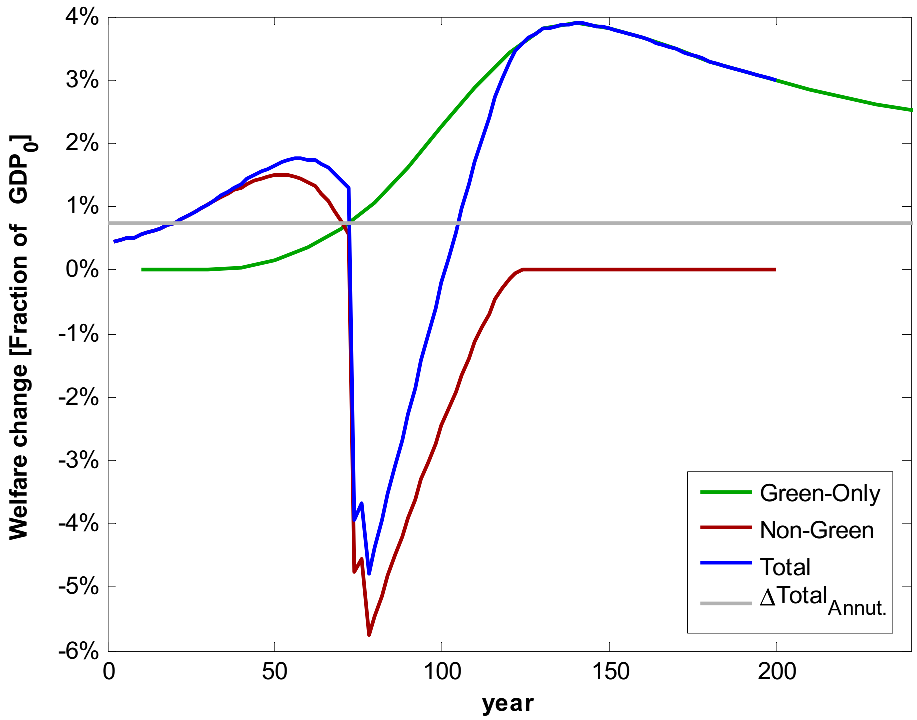

Figure 2.9: Welfare impact of tax

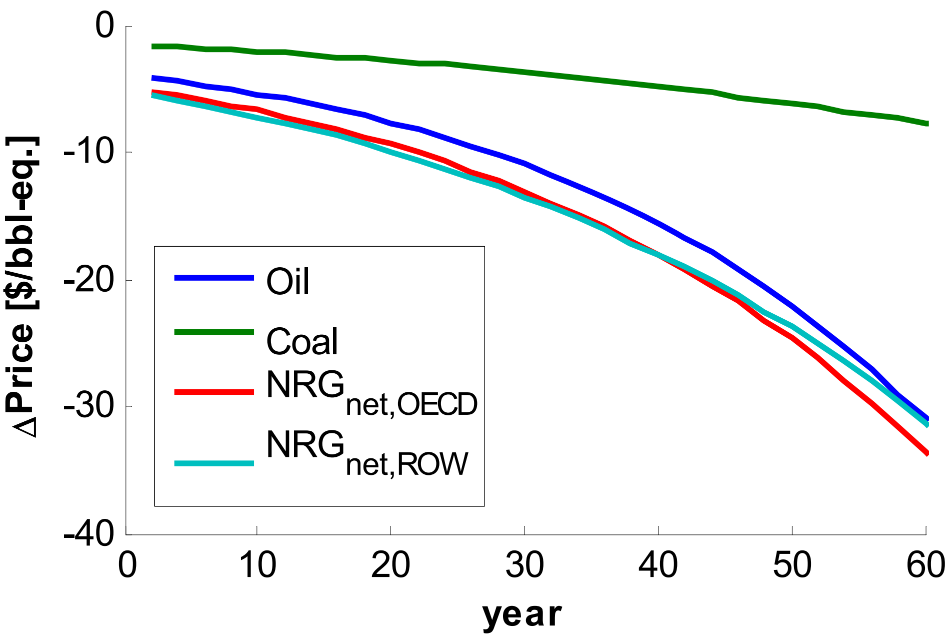

Figure 2.10: Impact tax on fuel and energy prices

(shown up to year 60, after which the fossil fuel is replaced by the backstop in the tax scenario). This explains that the low or modest tax during the first years increases the non-green welfare, whilst the green-only welfare is slightly reduced, due to slightly higher initial extraction rates. Since the initial emission reductions are small and cumulative emissions increase already after around 20 years (cf. Fig. 2.6), the change of green welfare becomes positive quickly. Nevertheless, when the tax induces the switch to the relatively expensive backstop after around 70 years, subsequently to which no tax revenues are generated anymore, a substantial loss of the overall welfare rate occurs (Fig. A2.10 in the Annex shows the details of the dynamic welfare paths for the BAU and the tax scenario, where the welfare drop associated with the endogenous introduction of the backstop under the tax is visible). The use of the backstop further increases the green welfare, but it takes a while until the convergence of the fossil fuel price in the no-tax case to the backstop price neutralizes the tax-induced loss of non-green welfare towards the year 120 where the backstop emerges also in the BAU (as the sum of green-only and non-green welfare, the overall welfare becomes positive again already somewhat before year 120, at around year 100). As Fig. 2.9 also shows, the summed net present value of the overall welfare changes is positive, corresponding to an increase of the world GDP now and forever by around 7 ‰ (of today's GDP). This beneficial welfare effect of the tax puts the Green Paradox into perspective. A comparison of this positive overall welfare impact resulting when the behavioral response of the fuel owners is taken into account, to the corresponding effect of the tax calculated assuming fixed producer prices, qualifies the Green Paradox even further. If the fossil fuel producer prices remained, throughout time, unaffected by the tax, the tax would have a negative effect on the overall fuel-consumers' welfare, corresponding to a loss of 3 ‰ of GDP each year, now and forever (cf. Fig. A2.11 in the Annex for the dynamic welfare impact path details). That is, rather than making the tax socially undesirable, the anticipation effects of the fuel owners may render the tax more desirable even when it is suboptimal and of the Green-Paradox type (that is, with a low initial level and a rapid rate of increase).

Classifying only parts of the resource owner rent as loss for the economy, naturally reduces the social desirability of the tax. As Fig. A2.12 in the Annex shows, the main-scenario tax remains beneficial even if up to around 50 % of the resource owner rents are accounted for as social gains. However, when a full accounting of the resource owner rents fully idles the rent-extraction motive, the social impact of the tax is an annuitized loss of almost 5 ‰ of GDP (the same loss would be 1.7 % in the counterfactual case with fixed resource prices).

2.5 Conclusions

Theoretical arguments have been brought forward showing that climate taxes could, under certain circumstances, exacerbate the climate problem rather than to alleviate it. Moreover, as the theoretical reflections suggested, the characteristics of the tax that could lead to these adverse effects could coincide with those required to make the taxes politically feasible: a low tax level at the time of introduction could enable political acceptance, and the level could rise relatively rapidly over time. Eventually it could approach the marginal climate damage from emissions, potentially even exceeding it, corresponding to governmental desires to generate tax revenues. If they were justified, these fears along the line of the Green Paradox would represent a heavy burden for most of the currently debated regional and global climate policies as even systems such as national and international cap-and-trade schemes lead to a pricing of the carbon emissions, and thus they indirectly correspond to taxes that may be subject to the same Green Paradox mechanism.[53]

This study tries to give an answer to the question whether (potential) tax induced adverse effects of causing emissions to occur earlier would really outweigh the benefits of later emission reductions that the same taxes would imply. It is taken into account that fossil fuels exist in a quasi-continuum extractable for increasing, up to probitively high costs. The calibrated, dynamic, numerical model considers an exogenous, downward sloping demand for energy. This energy can be provided as a constant elasticity of substitution aggregate of weakly substitutable oil and coal, or, potentially, by a clean backstop with infinite supply elasticity and a cost that diminishes over time. The fuel demand is split between the OECD and the rest of the world, calibrated according to current consumption and worldwide prices, and may grow over time. We exclude non-fuel trade between the two fuel consuming regions. The fossil fuels are extracted for marginal costs that increase with cumulative extraction according to empirical estimates. They are offered by forward-looking competitive suppliers who maximize their net present revenue subject to an exogenous interest rate – as corresponds to the Hotelling framework within which the Green Paradox has been proposed.

The main results from the calibrated model strongly suggest that fears of such counterproductivity are largely unjustified: climate taxes – even those with a low initial level

[53] Moreover, as has been explained by Sinn (2008) and others, the same effect could apply even more strongly to the case of policies supporting the development of alternative technologies: today's R&D support may provide alternative, clean technologies in the future, thus lowering profit expectations of fuel owners and inciting them to increase sales today. Contrary to the tax, which, even if initially low, may at least partly mitigate the increase in today's sales, the future technologies cannot mitigate initial sales increases.

rising rapidly over time – reduce global medium and long-term emissions so dramatically that any possibly induced increase of short-term emissions is by far outweighed by medium and long-term emission and damage reductions. This is still the case even if a very large discount rate is used to aggregate emissions or damages to their net present value. The analysis suggests that even for discount rates up to 11 % and beyond, a Green Paradox-type emissions tax tends to reduce the NPV of emissions and of climate damage. Various robustness checks for key model parameters largely confirm this result. Consequently, even with preferences corresponding to those of advocates of high climate damage discount rates based on market interest rates such as, e.g., Nordhaus,[54] the rapidly increasing emission taxes appear clearly beneficial for the climate.

Moreover, the behavioral response of the fuel owners to the carbon tax does not even necessarily mean that the tax becomes less desirable for the fuel consumers from a general welfare perspective. Taking into account that – by lowering fuel producer prices – the tax allows to extract resource rents, the main-calibration tax becomes desirable exactly *because* of the 'Green Paradox'-reaction of the fuel owners to the (rapidly increasing) tax: the reduction of the fuel price and of the climate damages more than compensates the consumers for the distortions from the suboptimal tax, and it does so even if up to around 60 % of the resource owner rents are accounted for as social gains, attenuating the rent-extraction motive. The calculations suggest that the tax from the main-calibration – rising to a relatively high level compared to damages from the Nordhaus damage module – would be undesirable if the fuel producer prices remained at their BAU scenario level; in this hypothetical case the economic distortion would exceed the environmental gain. Further analysis should consider to which extent these results extend to the case of regionally limited climate policies.

The present study tests the effect on the main results of an endogenous emergence of a clean backstop and of synthetic liquid fuel production through coal liquefaction, as well as of a rate-dependent component in the extraction costs, and of a geographical limitation of the tax. It does, however, abstract from the possibility of monopolistic fuel suppliers as well as the impossibility of current policymakers to commit to future taxes or to their absence. An analysis of these two issues in a model closely related to that developed here suggests that both effects would further reduce the likelihood of the strong Green Paradox (cf. Habermacher, 2012a).

An important driver for the present results seems to be the fact that a rapidly increasing tax, which is required for the anticipation effects to potentially outweigh the direct effect of the tax on current emissions, becomes so quickly extraordinarily high that fuel owners

[54] Cf., e.g., Nordhaus (2008, 2007).

hardly have time to profitably sell much more fuel in the initial years before the tax achieves a level that is larger than the resource rent in the BAU scenario. Once the tax has surpassed that BAU resource rent level, the tax cannot increase cumulative emissions anymore, independent of the residual resource rent.[55]

[55] Cf. Hoel (2010) for a detailed explanation of this point.

2.6 Annex – Additional simulation results

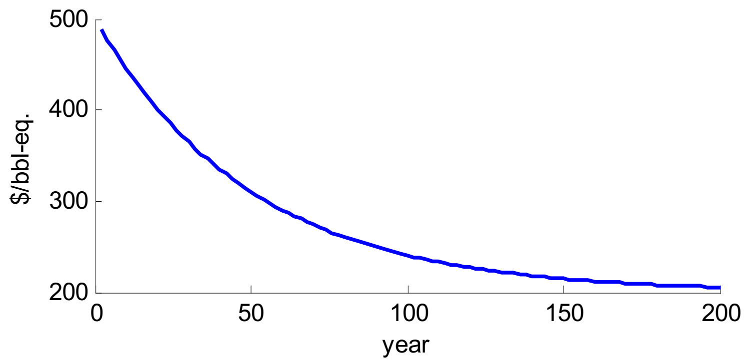

Figure A2.1: Backstop price path

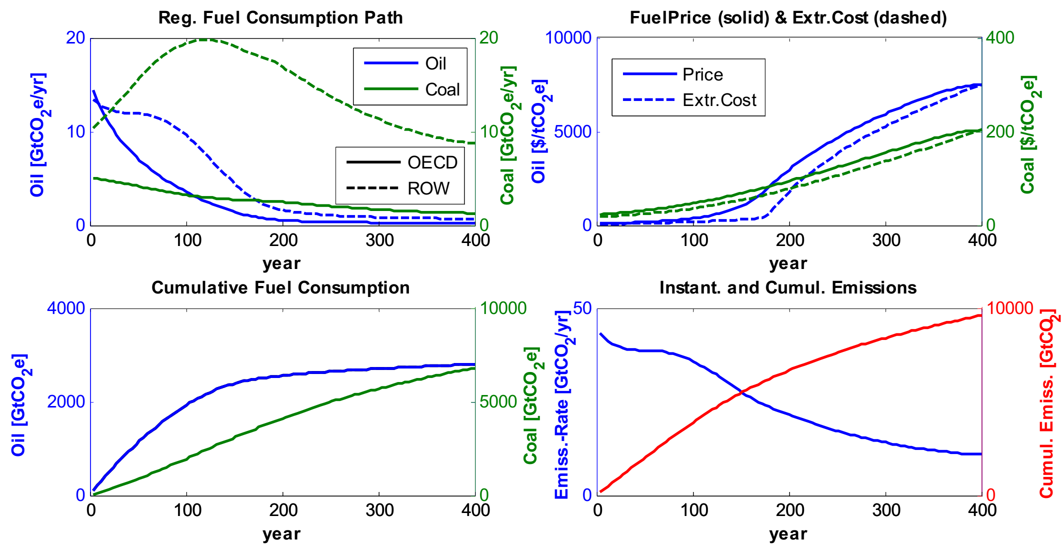

Figure A2.2: Simulation without backstop

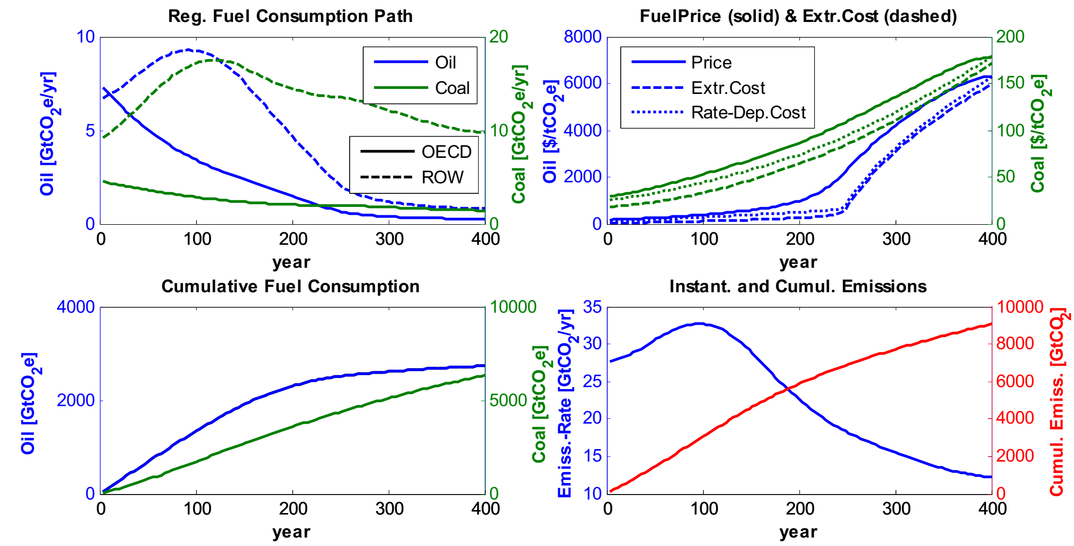

Figure A2.3: Simulation with rate-dependent costs (no backstop) (variant 1)

The ξ parameters used are 8250 \$/(bbl·yr) for oil and 12 700 \$/(bble·yr) for coal.

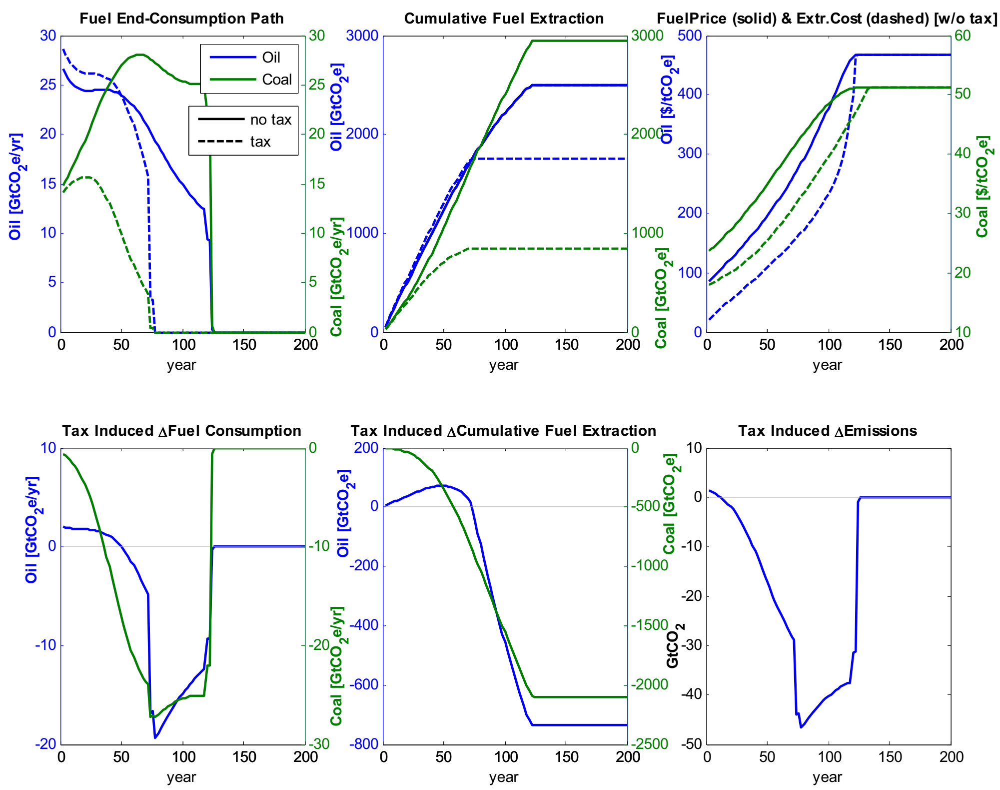

Figure A2.4: Tax impact details base calibration

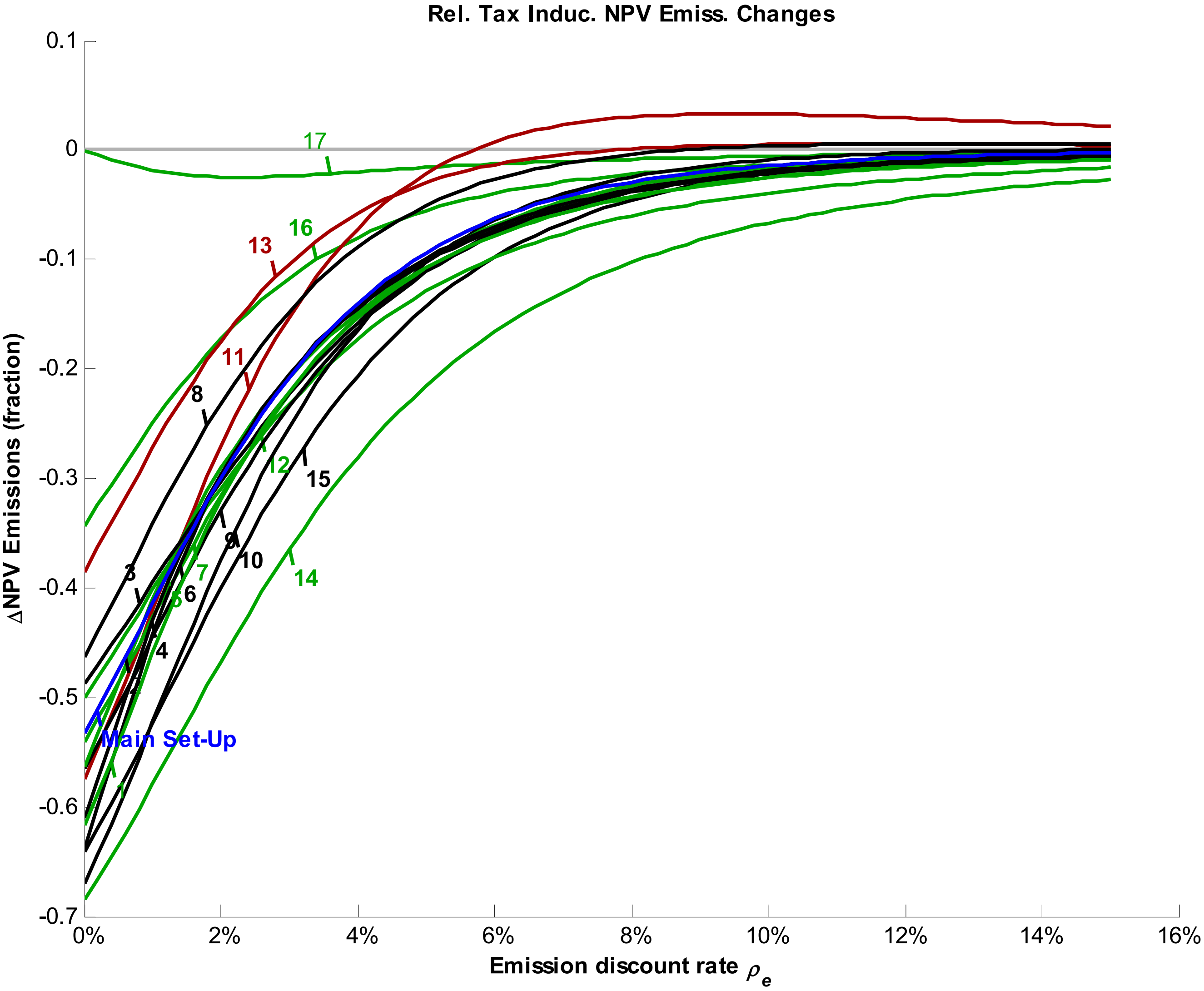

Figure A2.5: NPV damage changes alternative parameterization

Notes: Red emphasizes those variants for which the threshold emission discount rate ρ_e^* is substantially reduced, to values of 6 % or lower. Green emphasizes those variants for which the tax reduces the NPV emissions for the entire range of tested discount rates.

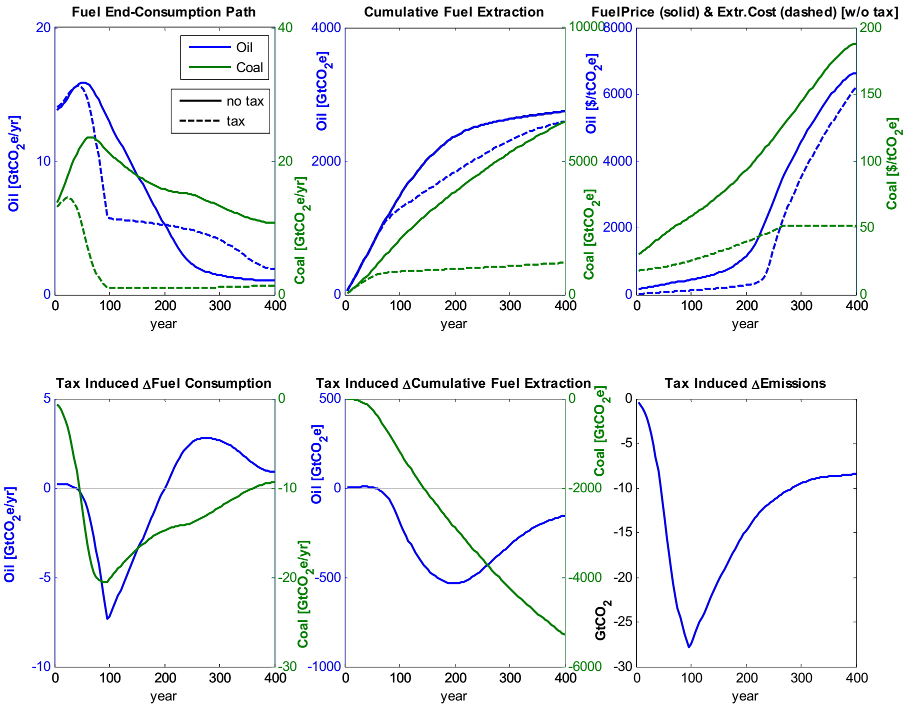

Figure A2.6: Tax impact details rate-dependent extra costs (variant 1)

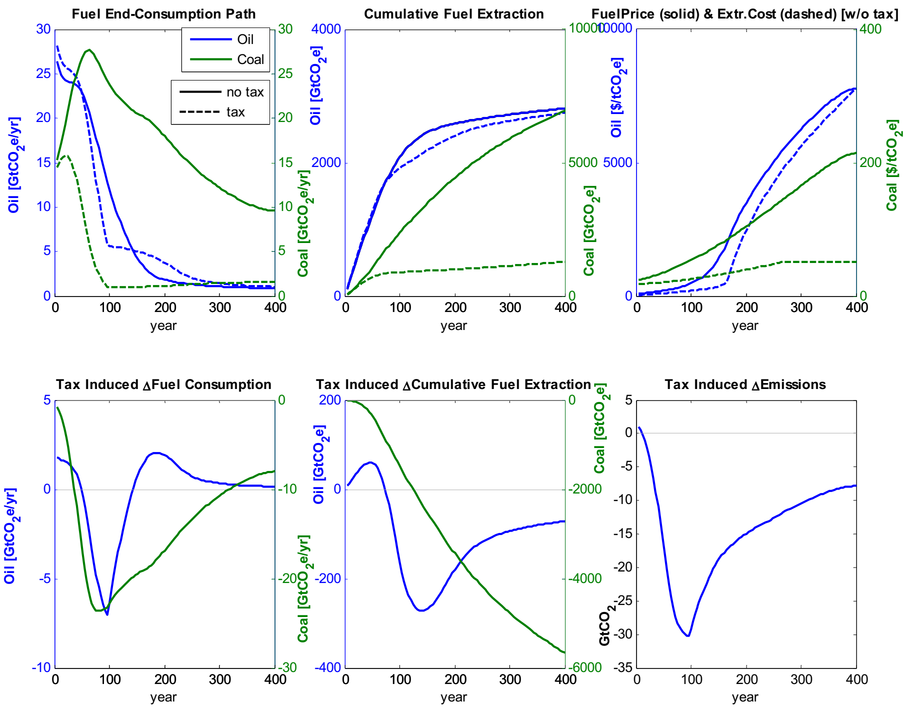

Figure A2.7: Tax impact details no backstop (variant 4)

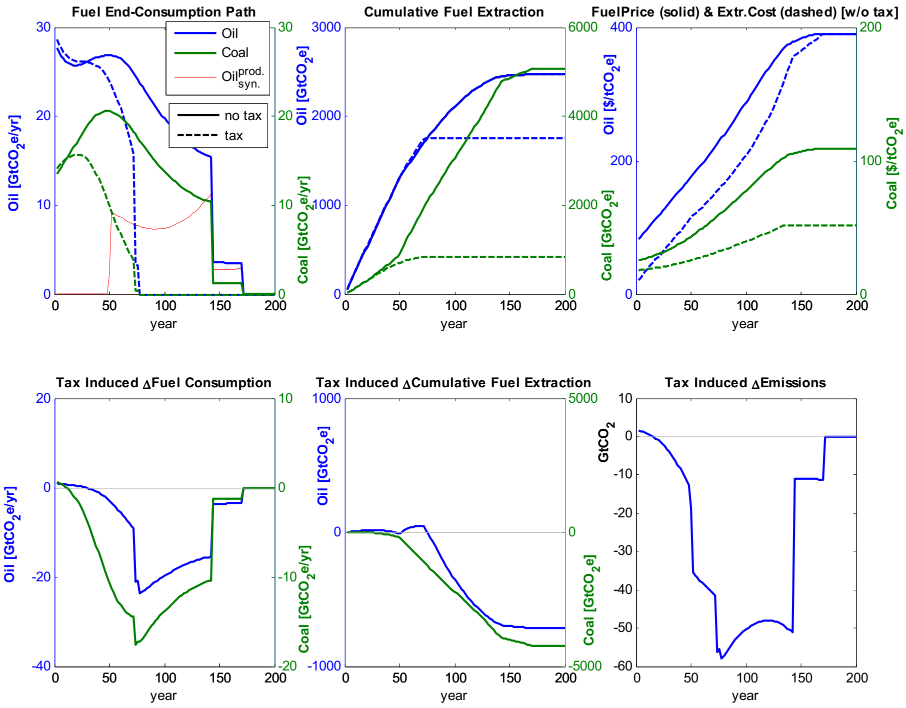

Figure A2.8: Tax impact details liquefaction (variant 10)

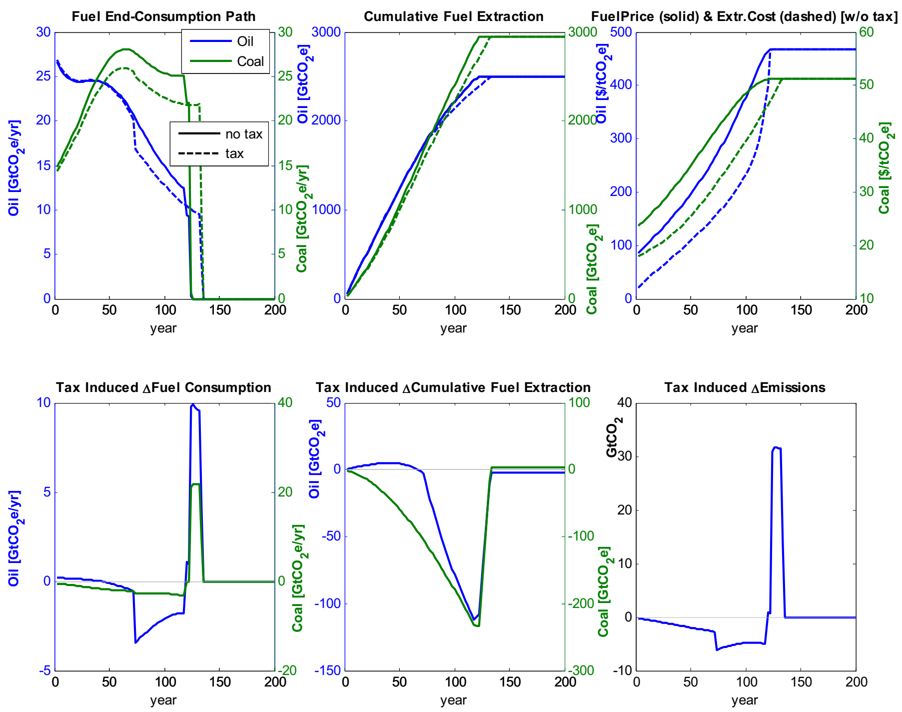

Figure A2.9: Tax impact details OECD-only tax (variant 17)

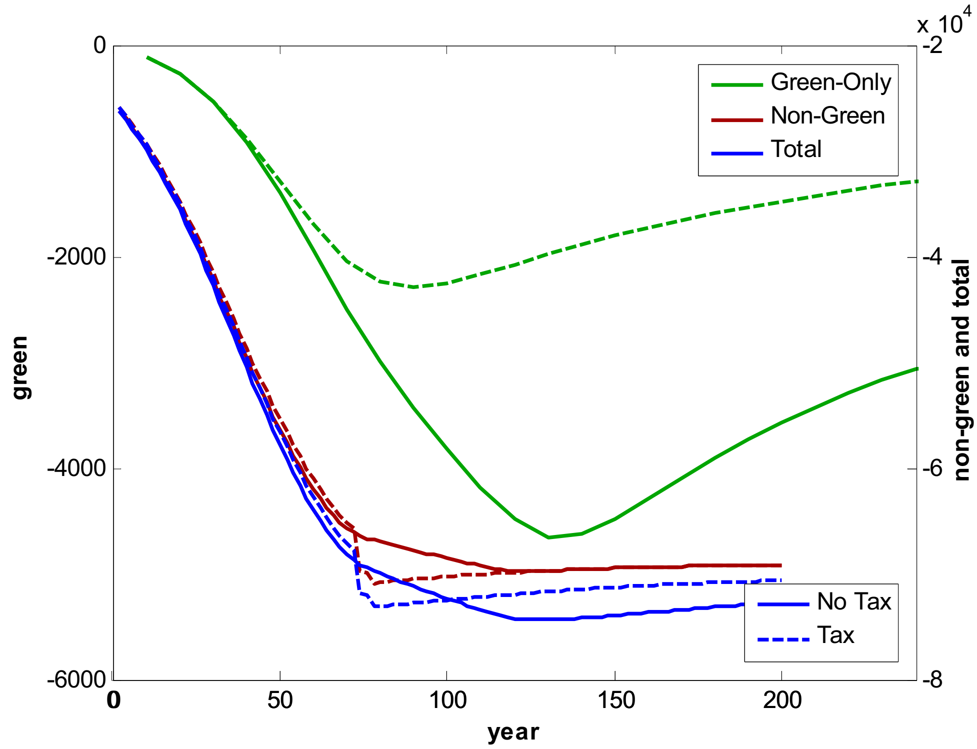

Figure A2.10: Welfare paths, BAU vs. tax

Main scenario, with welfare discount rate $\rho_w = 3\,\%$.

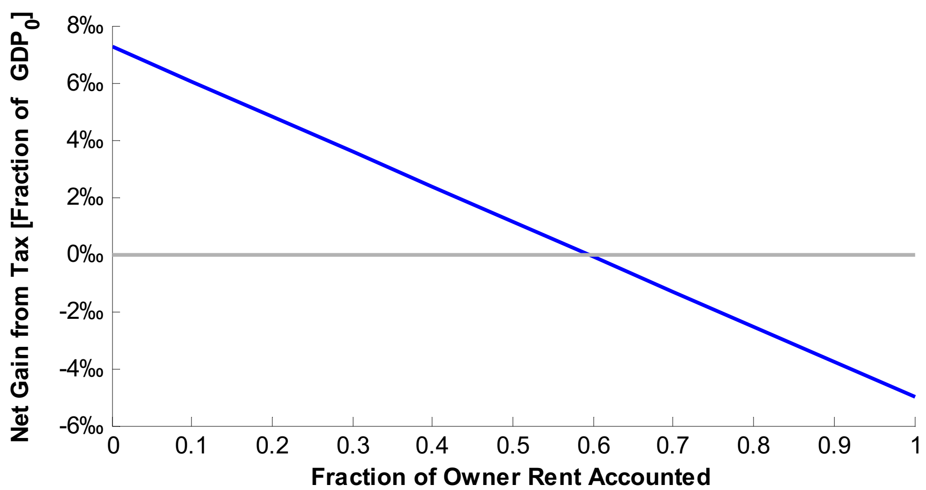

Figure A2.12: Welfare Impact of Tax, as Function of Resource Owner Rent Fraction Accounted for

Main scenario, with welfare discount rate $\rho_w = 3\,\%$.

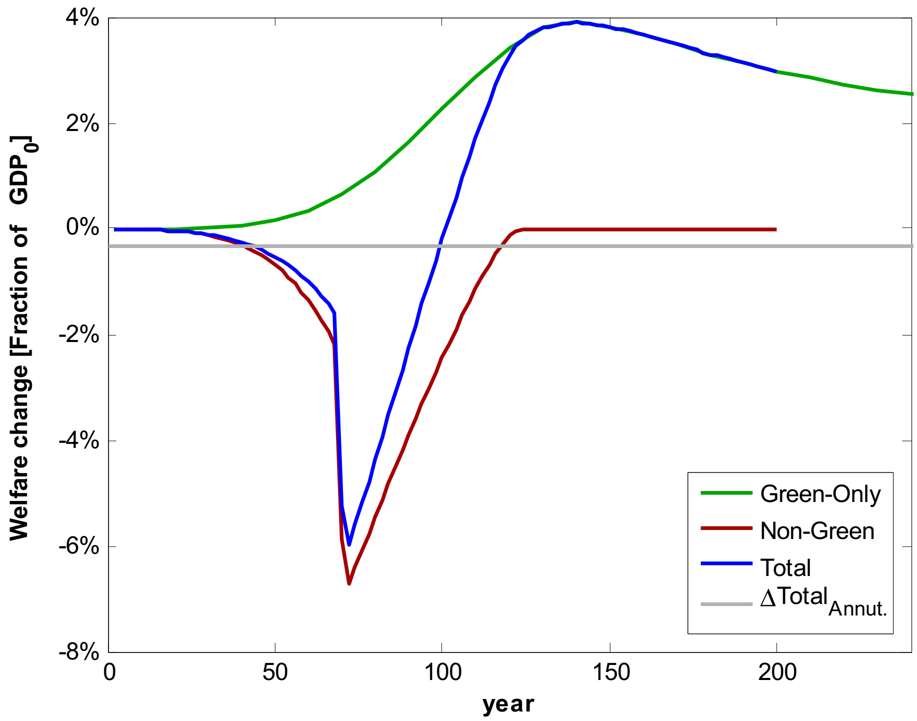

Figure A2.11: Welfare impact of tax, counterfactual with fixed price paths

Main scenario, with welfare discount rate $\rho_w = 3\%$. Price paths hypothetically fixed to those in the no-tax equilibrium.

Essay 3

Dynamic Carbon Leakage and Taxation with Depletion and Discounting

Abstract

This essay treats various aspects of unilateral carbon taxation in presence of exhaustible fuels. A method to disentangle terms-of-trade and pollution components of the committed optimal unilateral tax on exhaustible fossil fuels is provided. The method is used to replicate the optimal dynamic green tax path in a numerical model. In a calibrated, dynamic fuel market model with empirical fuel extraction cost curves we study leakage and optimal unilateral tax paths for the OECD. They vary strongly with model assumptions. The strong curvature of marginal oil extraction costs from empirical estimations, and coal liquefaction providing a dirty backstop specifically for oil, as well as a clean backstop for fossil fuels tend to have strong effects on the evolution of leakage rates. Leakage effects can be very large, even if future emissions are discounted. Whilst relatively insensitive to a scaling of the resource stock, leakage depends strongly on the curvature of extraction costs. The rates differ strongly across fuels and optimal unilateral oil and coal taxes can have opposite signs; not much is left of the idea that carbon taxes should be uniform. Notably, liquefaction can lead to negative leakage rates from oil emission reductions and consequently optimal oil emission taxes above the WTP for global emission reductions. In presence of an endogenous clean backstop, in contrast, oil savings tend to prolongate the fossil fuel era and increase global fossil fuel emissions. This can imply leakage rates above unity and negative optimal unilateral oil emission taxes, whilst for coal, limited leakage warrants positive taxes. Green Paradox effects tend to lead to increased present value emissions for anticipated taxes.

Author: Florian Habermacher

Keywords: unilateral climate policy, fuel specific carbon tax, fossil fuel depletion, dynamic carbon leakage, discounting, coal liquefaction, backstop, OECD.

JEL classification: Q54, Q41, H23, H21.

For helpful remarks and suggestions I am particularly thankful to participants at the Public Finance and Economic Policy Seminar (Finanzwissenschaftliches Forschungsseminar) at University of Munich, 22 July 2011, at the Energy Economics PhD Workshop at University Bern, 2 December 2011, at the Annual Congress of the International Institute of Public Finance (IIPF) at TU Dresden, 16-19 August 2012, as well as at the CESifo seminar at the Center for Economic Studies in Munich, 19 November 2012 and at the OxCarre seminar at University of Oxford, 22 November 2012.

3.1 Overview

"The static-equilibrium type of economic theory which is now so well developed is plainly inadequate for an industry in which the indefinite maintenance of a steady rate of production is a physical impossibility, and which is therefore bound to decline.", Hotelling, 1931.

What Hotelling wrote in 1931 with notably oil and coal in mind, 80 years later appears to still not have received due attention in the bulk of the literature that works to answer a question where the dynamics of the resource supply and of technological, and even political changes seem to be of overwhelming importance. This question is the following: how severely is a unilateral effort to contain climate change by reducing regional greenhouse gas (mostly carbon dioxide) emissions undermined by offsetting foreign emission reactions; that is, by the so called carbon leakage? A large fraction of the literature studying carbon leakage uses static models, or dynamic models with static fuel supply, and finds moderate to low leakage rates. This study presents theoretical and numerical results on carbon leakage from a fully dynamic perspective. Fuel exhaustibility implies that medium and long-run leakage can be much higher than previous studies suggested. The main reason for this is that fuels not consumed (imported) by a home region during a specific time-window may be sold by the fuel owners to other regions not only during that specific time-window, but they may instead also be sold at any point earlier or later as long as some demand exists for that fuel in the remainder of the world. In fact, the basic reason why competitively supplied fuels that can be extracted for costs of up to the current market price are not offered today is that the suppliers expect to sell them in future periods with even higher profits. With other words, if a policy is to prevent extraction of some fuels forever, it must necessarily reduce the net demand price for the fuels everywhere and always to a level below the fuels' extraction costs. If this was not the case, a fuel supplier would increase his profit by extracting at one of the periods with a demand price that exceeds the extraction costs. However, if the policy is regionally limited, it can not directly reduce that demand price in the remainder of the world except by increasing the foreign consumption rate. This means that domestic fuel consumption reductions are at least partially offset by foreign consumption increases, induced by a depression of the global fuel price. Assuming, as an approximation expressing the low costs with which fossil fuels are shipped over long distances, a completely globalized fossil fuel market, leakage could even fully offset domestic fuel consumption reductions in the long-run: if a fuel is spared from domestic consumption only due to a specific regional policy, no necessary reason why fuel consumption in the remainder of the world would stop before all of that fuel is extracted exists. In a simple world, a domestic fuel consumption reduction tends thus to mainly prolongate the fossil fuel consumption horizon instead of reducing the total amount consumed. Whether all initially spared fuel is consumed later on depends on whether technical or political developments will in future allow to replace the fossil energy also in the remainder of the world. Moreover, to which degree the leakage is considered a problem even if parts of the emission reduction offsetting occurs many decades later, depends on the time-discounting of emissions. Given a fixed amount of reserves of exhaustible fuels, a welfare assessment of the leakage problem is intrinsically related to the fuel depletion in the medium and long term future, to future developments in the fuel market framework, as well as to emission discounting. These aspects have received scant attention in the existing leakage literature. This chapter presents a step towards filling this gap, using analytical and numerical models to investigate how leakage can sensibly be represented and estimated in a fully dynamic setting.

Section 3.2 provides an extensive motivation for this study and refers to different strands of related literature. Part 1 analyzes optimal unilateral fuel taxes in the presence of global pollution. Among other things, it confirms that a hypothetical compensation method can disentangle terms-of-trade and climate tax components, and it provides expressions for the optimal climate tax and relates them to a welfare relevant damage leakage rate. The different leakage rate notions are also defined. Part 2 uses a calibrated dynamic market model of substitutable and exhaustible fuels, oil and coal. The two fuels are consumed as a constant elasticity of substitution energy aggregate. The energy demand curves are exogenous and modeled as independent functions for the OECD and the rest of the world (ROW), and may grow over time.[56] Fuels are extracted at costs that increase with cumulative extractions, by competitive, forward-looking suppliers which maximize their net present profits. Fuel consumption maps directly to global carbon emissions, and climate damage is assumed a linear or convex function of cumulative emissions. Model extensions allow the transformation of coal to synthetic oil, and a clean backstop. First, the model is used to replicate the findings from Part 1. Then, different leakage rates are analyzed in different scenarios. Key findings include that oil is subject to very large absolute leakage rates in a basic framework. Interestingly, and contrary to what one might think on first sight, this appears to be less due to the limited relative size of the reserves (compared, e.g., to the more abundant coal) but more to the strong curvature in the empirical oil extraction cost curve: it is shown that, for the example of linear extraction costs, a scaling of reserve sizes may not affect leakage rates; intuitively, however, with a strongly, convexly kinked cost curve, a regional fuel

[56] That is, we neglect direct industry dislocation effects of the tax. These would, in the basic model, tend to increase the overall leakage rate.

consumption reduction can mainly postpone by a few years the time until the cheap 'pre-kink' fuel is used up, rather than to delay fuel use for a long time. This is related to the Green Paradox effect, where under the assumption of a fixed reservoir, supplied at limited costs – as corresponds closely to the case of an 'infinitely strongly kinked' curve where costs rise infinitely rapidly after the extraction of a certain threshold amount – it is found that a tax may only affect the timing but not the absolute amount of fuel used (Sinclair, 1992; Sinn, 2008). Due to the foresight of the fuel owners, leakage, as a response to a specific time t's domestic consumption perturbation, tends to occur with a substantial spread across time and is centered around t in basic models: recognizing, e.g., that a future tax reduces the profitability of future sales, they will increase current sales (and potentially those after the period with the tax). With substantial discount rates, this implies that even when leakage offsets the major fraction of the domestic emission reductions in terms of absolute emissions, the relevant leakage rates may be low for today's taxes, but they may easily exceed unity for anticipated future taxes, where parts of the foreign emission offsets occur prior to the domestic reduction and are thus weighted relatively more (strong Green Paradox for future taxes). As therefore a regional oil tax may increase rather than reduce the net present value of global emissions, the optimal regional policy can contain negative taxes. This contrasts to an optimal global policy, where the positive emission damages warrant strictly positive tax rates throughout time. Above-unity leakage rates are, however, significantly less likely for coal, even with substantial discounting. As discussed, this seems to occur not as much because of the pure abundance of coal, but more because of the weaker curvature of the cost curve.

When the substitutability of the fuels is taken into account in a basic setup (no liquefaction and no backstop), the overall leakage rate for domestic oil emission reductions (in the following sections we tend to simplify the terminology, writing *oil leakage*) can become negative as the coal-to-oil substitution effect in the foreign CES demand dominates the direct oil price effect as oil becomes increasingly scarce over time. Introducing endogenous liquefaction in the model can lead to negative oil leakage rates already for today's tax: When liquefaction[57] supplements a synthetic substitute for crude oil in future, saving oil (today or in future) delays the start of the dirty liquefaction process employed abroad and therewith reduces global emissions even beyond the amount saved domestically. The increased demand for the input into liquefaction, i.e. coal, on

[57]Interestingly, Felder and Rutherford (1993) have also suggested negative leakage rates from a regional (not fuel-specific) climate tax during the years when liquefaction starts to play a role in the rest of the world. Besides the issues with the somewhat ad hoc representation of the fuel-extractions in their model (cf. below), they have restricted their attention to instantaneous leakage rates for each period rather than considering the (NPV) effect of current taxes on future emissions.

the other hand, increases the coal leakage rate. A further extension of the model contains an endogenously emerging clean backstop, available at costs that decrease over time and modeled as a perfect substitute for the fossil fuel aggregate. Its effect may surprise. Rather than reducing leakage by limiting the time available for the foreign offsetting of domestic reductions, the backstop implies very high leakage rates above unity – for current and future oil consumption reductions. The latter prolongate the time during which the fossil fuels can compete against the backstop. Given that during the final phase of fossil consumption the energy is very coal intensive, a bit more oil increases emissions strongly during that phase.

3.2 Motivation and Literature

A climate policy aimed at an economically efficient reduction of carbon dioxide (CO_2) emissions may take the form of a CO_2 tax or a cap-and-trade system. The level of the tax, or correspondingly the tightness of the allowances in the cap-and-trade system, expresses a willingness to pay (WTP) for climate protection; in other words for global greenhouse gas emission reductions. In a first-best world, where an optimal tax scheme can be imposed, all global emissions would be subject to an identical per-unit emission tax. Alternatively, in a second-best case, where a climate policy is implemented only in parts of the world (we refer to this as the policy region), a uniform tax level on emissions may still be optimal in the absence of relevant links between emissions in the policy region and those in the rest of the world. In this case, a regional emission reduction would translate one-for-one into reductions in global emissions, for which agents are willing to pay. However, both the first- and this second-best scenario are unlikely to correspond to the reality of current or near-future climate policies. First, all climate protection measures implemented thus far only cover a fraction of global emissions, and there is no global agreement in sight for at least the remainder of the decade. Second, major sources of fossil energy and anthropogenic CO_2, notably oil, natural gas and to some extent coal, are traded on global markets rather than only on regional markets (as are other goods whose production depends on the fuels). This implies that consumption reductions in one region will directly impact the resource availability and consumption in other regions, that is, the independence of emissions across regions is violated for the most important sources of anthropogenic CO_2 emissions. The global character of the fuel supply is a primary reason why a regional emission change does not generally mean a global emission change of the same magnitude. This is the well-known issue of carbon leakage (e.g., Felder and Rutherford, 1993, and Burniaux and Oliveira-Martins, 2012).

An efficient market measure motivated by climate protection implies uniform marginal emission costs for (indirect) *global* rather than regional emissions. However, as a regional policy can only sanction regional emissions, the second-best-efficient[58] policy must weight these regional emissions by the degree of influence they have on global emissions. The various primary fuels used in today's economy have strongly varying supply characteristics. For example, brown coal is often only consumed regionally;[59] coal reserves are often considered practically unlimited;[60] oil and gas are globally traded and exploitable in limited amounts at increasing costs; and locally or regionally consumed wood is, in some circumstances, renewable. Thus, a *regional* change in the consumption of one of the different fuels implies a *specific* variation in the *global* consumption of that fuel as well as other fuels. The optimal regional CO_2 price contains a proportionality factor that reflects the extent to which regional emission changes translate into global emission changes. Therefore, this price is likely to vary substantially across fuels. This implies that it is inappropriate for a regional market-oriented policy to weight (and thus, to price) all domestic emissions uniformly. This paper addresses the fuel dependency of optimal regional emission weighting, an issue that has received scant attention in existing literature.

Neglecting *fuel*-dependent prices, the traditional carbon leakage literature has largely restricted attention to economic sector-specific leakage and terms-of-trade factors that imply *sector*-specific carbon pricing and, potentially, sector-wide policy exemptions. Hoel (1996) provided an extensive analysis of sector-specific differentiation of a unilateral CO_2 tax considering a single aggregated fuel. More recently, Böhringer et al. (2010) introduced a specific technique to distinguish between the efficiency-related leakage motive and the terms-of-trade reason for sector-differentiation of a unilateral tax. In contrast to their analytical model, their numerical analysis of US and EU policies distinguishes between a number of different fuels. However, the considered tax was still wholly sector-specific, and fuel-specific taxes were not considered in their paper. Similarly, Kirchgässner et al. (1998) examines the importance of sectoral exemptions on the economic and environmental impacts of a unilateral climate tax. Kirchgässner (2001) discusses the reasons why the optimal climate taxes may be sector-specific if the objective, according to political economics or ordinary people's preferences, is to limit tax revenue rather than simply the excess burden. Finally, Burniaux and Oliveira-Martins (2012) extensively examine the differences between oil and coal in terms of supply elas-

[58] The policy is considered second-best because it is regional instead of global.

[59] See, e.g., IEA (2011), Part IV: The various Tables '3. Coal and peat production by type' and '8. Coal trade by type of coal' for the various surveyed regions.

[60] See, e.g., van der Ploeg and Withagen (2011, 2012b) and Burniaux and Oliveira-Martins (2012). The strong characteristic difference between oil and coal supply is also pointed out in Burniaux et al. (1992) and Golombek et al. (1995).

ticities and global market integration. While they identify the impact of these market characteristics on the leakage rate of unilateral climate policy, their focus remains on a uniform carbon price optimized not with respect to the carbon leakage but simply for respecting a specific *regional* emission threshold.

While Golombek et al. (1995) have addressed the issue of the optimal regional fuel-specific tax structure, the present analysis extends their study in two important ways. First, their focus remained on a static model, notably assuming an isoelastic, static supply of fossil fuels. This is in contrast to one of the most distinguishable features of the supply of non-renewable resources; that is, the fuels are exhaustible, with extraction costs that are, in the medium-term, increasing in the amounts previously extracted. In this study the exhaustibility of the fuels is explicitly considered within the framework of a numerical dynamic model of the fuel markets where suppliers strategically allocate the extraction of their fuels over time, maximizing their present discounted net revenues subject to the (increasing) extraction costs. As will be explained, this is crucial as the concept of a static leakage rate is inherently incompatible with exhaustible emission sources. Second, Golombek et al.'s static framework did not allow them to consider future developments in the fuel market. In reality, the supply of solid, liquid and gaseous fossil fuels may dramatically change from the currently observed pattern once the relative availability of specific fuels significantly changes due to advanced degrees of exhaustion. Clean backstops developed in the future is one example. Fuel transformation processes, such as coal-to-liquids (liquefaction), is another; they may become widespread if the extraction cost of oil increases further and coal remains abundant. Using a general equilibrium model with a detailed representation of the supply of petroleum, and other energy products in general, and a bottom-up implementation of coal-to-liquids processes, Chen et al. (2011) estimate that liquefaction could account for one-third of the global liquid fuel supply in 2050.[61] Allowing for such a fuel transformation process when the fuel prices render it economical, the model developed here is used to investigate the potential implications of these processes for the optimal unilateral climate tax structure.

The present study follows Golombek et al. (1995) by focusing on the market for fuels. This seems to be a suitable approach as, for example, McKibbin and Wilcoxen (2008), Böhringer et al. (2010) and Kuik and Hofkes (2010) have shown that the trade of non-energy goods is of lesser importance for both leakage and terms-of-trade effects – these

[61] South Africa produces 30 % of the liquid fuel that it consumes through such coal liquefaction processes (Sasol Synfuels International, 2005). While this currently makes South Africa the largest coal liquefactor, China has plans for a number of very large coal liquefaction plants, and proposals for plants exist in other countries as well (BGR, 2009a).

effects are dominated by the international trade in fuels.[62] Similarly, Oliveira-Martins (1995) and Burniaux and Oliveira-Martins (2012) find that the leakage effects are primarily determined by the fossil fuel market, while trade characteristics of consumer goods are less important.

The optimal regional, fuel-specific carbon taxes influence the time-path of the consumption of the various fossil fuels. These optimal time-paths are the central issue in studies by van der Ploeg and Withagen (2011, 2012b). Regarding the optimal carbon tax pattern, their analysis, on the one hand, is limited to a focus on *global* policies. On the other hand they disregard the issue of the imperfect substitutability of the fuels as inputs to specific end-uses. In reality, society does not simply have a demand for a specific amount of 'energy', but it has a demand for different forms of energy carriers that are to be used simultaneously. While, for example, liquid oil could be a valid substitute for many applications that currently feed on solid fuels, the inverse is not true with current technologies. In other words, the substitution would need specific fuel preparation, such as coal liquefaction or the switch from combustion engine-based mobility to vehicles powered by coal-derived grid electricity, with potential efficiency losses and overhead costs. This has important repercussions on the second-best time-paths of fuel consumption achieved with the second-best policy instrument of unilateral, fuel-specific carbon taxes, as we will demonstrate herein. In this sense, certain portions of the present paper can be considered as a synthesis of the static analysis about fuel-specific unilateral carbon pricing by Golombek et al. (1995) and van der Ploeg and Withagen's (2011) study on global policies and the optimal time-path given exhaustibility but without the issue of fuel-specific final energy demand. Michielsen (2011) is related to the present study in that it also studies regional and intertemporal leakage for two imperfectly substitutable fuels. One of the fuels is supplied infinitely elastically (coal) and the other depletes (oil). This provides important insights about Green Paradox and leakage effects, as well as about sensible climate policies. Michielsen does not, however, explicitly study optimal *fuel-specific* carbon taxes, and restricts his attention to a stylized two-period model. Eichner and Pethig (2011) also study leakage and Green Paradox effects in a two-period model. They consider a single fuel and assume a limited elasticity of intertemporal substitution in demand, i.e., consumption in one period has a direct effect on the demand in the other period.

The (substantial) uncertainty about the long-term climate damage induced by carbon emissions is not directly considered here. Golosov et al. (2011) develop an integrated dynamic stochastic general-equilibrium model to analyze optimal oil and coal taxes taking into account uncertainty about climate costs that is resolved only in the future.

[62] The simulation results of Fischer and Fox (2011) suggest the same conclusion.

Their analysis is, however, also limited to an optimal global climate policy and thus not concerned with leakage effects. Similarly to van der Ploeg and Withagen (2011), they assume oil and coal to be perfect substitutes, a view which is rejected here. Interestingly, Golosov et al. (2011) indicate the possibility of the use of liquefied coal in combustion engines as a reason for their assumption of the perfect substitutability between the fuels. In our view, however, while liquefaction is allowed for here as well, the fact that this process may become relevant in the future just shows that oil and coal are only imperfect substitutes: while in some applications the two fuels may be substitutable without large energy losses and overhead costs (consider, e.g., the replacement of coal by oil in stationary power stations), applications where coal can only be used after liquefaction imply substantial overhead costs in terms of capital, labor and energy.[63] While Golosov et al. allow for emission discounting, they use a fuel reserve model that is more stylized than that used here. They assume a fixed amount of oil available, worth around 30 years of current consumption and extractable without costs, and coal of limitless supply. Additionally, their model does not explicitly take into account the possibility of future climate measures.

In today's economic environment, the different uses of the various types of fuel suggest that demand characteristics vary considerably across fuels. For example, cheap coal can be used for electricity production and for some other immobile purposes, while particularly in the transport sector for explosion engines, and for simple apartment heating systems consumers rely on liquid (or gaseous) fuels. Clearly, there exists a certain substitutability. As an example, depending on the prices, one can heat an apartment with electricity (from coal) instead of directly burning oil (or gas). That the fuels are non-perfect substitutes seems logical as expressed by the large amounts of coal, oil and gas that have been simultaneously consumed for many decades, despite (short- and longer-term) shifts in relative prices over the past. While therewith the demands for the various fuels are complexly intertwined, corresponding cross-price elasticities should generally allow an acceptable approximation of the real demand structure. In the long run, however, it is important to consider, other than this substitutability in the final demand, that significantly large price differences may render the transformation of fuels profitable. Due to the large coal resources and the limited availability of oil, in the future this may lead to coal gasification or liquefaction (i.e., coal to oil transformation) as well as to gas to liquid processes.

The literature provides a considerable number of estimates of leakage rates for regional greenhouse gas emission reductions. The suggested rates cover the full range of imag-

[63] For example, energy losses in coal-to-liquids processes are very large. Overall energy efficiencies of CTL processes are close to 50 % (Bartis et al., 2008).

inable values. As an example, Böhringer et al. (2010) find leakage rates of 35–40 % for unilateral action for the EU, and 15–20 % for the US. Others find values as low as around 5 % (e.g., OECD, 2009). Still others argue that leakage may exceed 100 %. For example, Babiker (2005) finds leakage rates of up to 130 % when taking into account industry dislocation and economies of scale. Finally, Di Maria and van der Werf (2008) model how directed technical change in the climate policy region provides efficiency enhancements that may reduce emissions in the non-policy region even if the latter is not concerned about the climate. Overall, however, the bulk of the literature suggests very modest leakage rates. In an overview, Burniaux and Oliveira-Martins (2012) identify values ranging from 20 % to less than 5 %, and in Burniaux and Oliveira-Martins (2000) they conclude that "carbon leakages are likely to be small for the range of parameters most frequently quoted in the literature".

Independent of the large differences between these values, a policy maker interested in the medium or longer-run effects of unilateral action has a particular problem with the proposed leakage rates from most of these studies. They neglect the time dimension or treat it only inadequately, and therewith typically do not properly examine the underlying economic reasons why the leakage rates may be modest in reality. Instead, their models find limited leakage rates primarily for technical reasons. To see this, it is important to note that the models typically neither apply any discount rate for future emissions, nor assume any specific future technological or political climate relevant changes to drastically limit the scope for future emissions. If no technical or global political breakthrough in terms of climate protection is foreseeable, any unilateral carbon tax may, however, only postpone the time until which, for example, virtually all oil physically available and reasonably extractable is consumed. In this case, domestic oil consumption reductions from a unilateral climate policy are in the medium-term almost entirely compensated by emission increases throughout the rest of the world (ROW). Even if parts of this increase in ROW emissions occur somewhat later than the domestic emissions would have in the absence of any regulation (it is not *a priori* clear whether the time shift is large or small), the overall expected leakage is, in the absence of the discounting of future emissions, approximately 100 %. Therefore, modest emission leakage rates seem logical only under the assumption of future changes in the fuel market framework or if future emissions are discounted. Yet, the reasons for which most studies have come up with limited carbon leakage rates are of a different nature. For example, Böhringer et al. (2010), Oliveira-Martins (2012), Perroni and Rutherford (1993) and Babiker (2005) use static models. In such static models, the limited leakage rates typically stem from an *ad hoc* concept of a static fuel supply function. Correspondingly they do not capture that fuel consumption savings in one period may be offset in later periods when otherwise the fuel reserves would already have been depleted, i.e. the fuel simply lasts longer but will ultimately still be consumed. This even applies to the study of Di Maria and van der Werf (2008) who assume endogenous directed technological change but disregard the fuel-market channel of leakage and fossil fuels depletion.

Another strand of the leakage literature uses dynamic models but exhibits some shortcomings in the treatment of the time dimension. For example, the dynamic models in Bollen et al. (1999), Burniaux (2001), McKibbin et al. (1999), McKibbin and Wilcoxen (2008) and OECD (2009) seem not to feature endogenously depleting fossil fuel reserves, but instead make specific assumptions on the exogenously given resource availability in the different time-periods. Therewith their models still do not fully capture that lower fuel consumption in early periods may simply imply that the saved resources may be consumed later on. The reason for their modest leakage rates may thus also primarily be found in the negligence of the dynamic, endogenous depletion of the resources. That the (fuel) dynamics receives insufficient attention in a large fraction of the leakage studies is not only astonishing because of its obvious importance due to the long term character of climate change and the inherent exhaustibility of the fossil fuels, but also because early authors had already used dynamic models with at least partially endogenous fuel depletion mechanisms, for example Felder and Rutherford (1993) and Manne and Richels (1991). It should be noted that, however, the approach used in these two early works was rather a hybrid solution between an exogenous and an endogenous fuel depletion path, e.g. with constant ratio depletion elements, not allowing forward looking resource owners to choose a fully flexible fuel extraction path. Other examples of leakage studies that feature endogenously depleting fuels are Manne and Richels (2000) using the MERGE model, and Babiker and Jacoby (1999) using the EPPA model. Similarly to Felder and Rutherford (1993) and Manne and Richels (1991), they use simulation periods that end in 2050 or in 2100 and do neither discount emissions, nor assume that up to this point in time a definite technological or political solution to the carbon emission problem would be found.[64] Thus, it seems that even in these studies the modest leakage rates could be rather technical results. These may be reversed if the model horizons would be longer, allowing a major fraction of the domestically saved emissions to occur in the remainder of the world.[65] Thus, it appears

[64] Manne and Richels (2000) explain that any judgment on a Kyoto policy crucially depends on what happens in the decades after the initial commitment period centered around 2010, and they study scenarios until 2050. They do not model what happens beyond that period. As the dynamic model in section 5 shows, an important fraction of leakage from current policies may occur in the decades after 2050.

[65] Some studies assumed coal to be of infinitely elastic supply without depletion and allowed for replacement of liquid fuels by coal. In the absence of time-discounting of emissions – as well as specific alternative technologies to replace the fossil fuels – this approximation is valid only for the medium-

that the most important reasons for which leakage in the long run may be substantially below 100 %, typically are not explicitly addressed in literature.[66] The proposed leakage rates are thus, *per se*, only of limited value for forward-looking, concerned societies or their policy makers. This seems especially clear as the primary reason for concern about climate change is that *future* global warming is anticipated today. If one were to exhibit an overly strong time-discount rate with respect to future temperature changes, one would hardly be concerned about the climate problem at all. It seems obvious, then, that current policy evaluations must take into consideration the effect that the current policies will have on emissions also in (many) decades, and perhaps centuries, to come. In the present study, the time dimension, especially in terms of discounting for future emissions and the possibility of future market framework changes, is explicitly taken into account in a model that additionally features fully endogenously depleting fossil fuel reserves.

Part I Theory of optimal unilateral tax and decomposition

3.3 Optimal Unilateral Tax

Since Pigou (1920), we know that in a simple framework a uniform unit tax on emissions, corresponding to the level of the marginal damage d, leads to the optimal level of consumption of a polluting good. Another simple case is that of a perfectly global pollutant in a situation where a tax is regionally constrained and a unitary pollution reduction within the tax region increases pollution in the remainder of the world by α units (α is called the leakage rate), and where, besides this pollution leakage, no additional relevant interaction between the regions takes place. As is understandable, and as Annex 1 shows, in this case the regionally optimal, unilateral tax level is reduced to $(1-\alpha)\cdot d$ (see Proposition 1 in Annex 1).

In general, this regional pollution tax calculation is, however, not pertinent. This notably because the channel through which the domestic emission choice generally affects the foreign emissions is through price effects, and the presence of these price effects

term future, as in the long run even coal reserves deplete.

[66] A related point is made in Eichner and Pethig (2011, p. 768). They note that (from the perspective of the intertemporal theory of nonrenewable natural resources), "*the prevailing view on the effectiveness of demand-reducing policies is flawed because the public and academic discourse [...] has largely neglected the close link between the economics of global change and the economics of nonrenewable resources and has therefore failed to account for the supply side of the problem in an appropriate way*."

warrants special consideration in the analysis of the optimal unilateral tax. When regional consumption affects prices of interregionally traded goods, such as the fossil fuels which are the basis for the vast majority of anthropogenic carbon dioxide emissions, a regional importer or exporter has incentives to influence the terms-of-trade by distorting its domestic consumption (and production) of the good. Consequently, this affects the optimal total level of the tax on the polluting good's consumption.

The remainder addresses the interrelatedness of the terms-of-trade component and the pollution component of the optimal unilateral fossil fuel (emissions) tax. We assume that the climate policy region is able to commit throughout time to a specific, initially announced future tax path. The case where a region is restricted to time-consistent fuel taxes is discussed, e.g., in Karp (1984) and, for the case including pollution, in Beermann (2012). Cf. Habermacher (2013) for a discussion of parts of the results in Karp (1984).

Model A fuel produces emissions at an intensity normalized to 1. Two consuming regions buy all their fossil fuel from external, decentralized producers at price p.[67] Regional welfare is given as $U_r = \int_T e^{-\rho t} \left[u_r(e_{r,t}) - e_{r,t}p_t - D_{r,t}\right] \mathrm{d}t$, where $T = [0, \infty]$ is the considered time-horizon, $u_r(e_t)$ is a utility flow concavely increasing in current fuel consumption $e_{r,t}$, p_t is the fuel price, and D_t is climate damage convex in cumulative emissions E_t. We assume the simple situation where the interest rate corresponds to the time-discount rate, ρ. Be $r = \{h, f\}$ the indexes for the domestic and the foreign region. For simplicity we abstract from foreign emission disutility;[68] that is, we set $D_{f,t} \equiv 0$, and call domestic damage $D_{h,t} = D_t = D\left(E_t\right)$, where $E_t \equiv \int_0^t e_{w,s}\mathrm{d}s$, with $e_w = e_h + e_f$, the worldwide fuel consumption. We call d_t the marginal instantaneous damage from marginal emissions at time t, $d_t = d(E_t) = D'_t(E_t)$. Assuming convex damage from cumulative emissions, we have $d'(E) > 0$.

Foreign consumers are fully decentralized. We assume interior solutions, wherewith the foreign consumption choice is thus governed by the FOC

$$u'_f(e_{f,t}) = p_t. \tag{3.1}$$

Fuel producers extract fuels with increasing marginal cost $c_t = c(E_t)$, $c'(E) > 0$. Assuming a competitive market, dynamic programming shows that the maximization of

[67]This is a common setup used in the literature, e.g., in Eichner and Pethig (2011), and Karp (1984). As Karp suggests, for regions with domestic fuel production, the demand in this model may be considered as the regions' residual import-fuel demand.

[68]The extension to the case with foreign damage should be straightforward for most of what follows.

the net present value of sales profits implies[69] a pricing according to

$$p_t = c_t + \int_t^\infty e^{-\rho(s-t)} \dot{c}_s \mathrm{d}s, \tag{3.2}$$

where the second term on the RHS corresponds to the resource shadow value.

Unilateral Committed Policy We examine the case where the domestic region, h, considers to unilaterally tax fossil fuels, using a committed policy. The foreign region's consumption is governed by undistorted consumption decisions of decentralized fuel consumers.[70] In what follows, we often simplify the time index of variables, writing, e.g., e_t for $e(t)$. Additionally h is often omitted for the index of the *domestic* region, so e_t would stand for $e_h(t)$. Worldwide consumption is always indexed with w, i.e., we write e_w. Further, when we consider taxes, we always mean unit taxes.

The domestically optimal consumption rate is implicitly defined by the maximization problem

$$\max_{e_{h,T}} U = \int_T e^{-\rho t} \left[u(e_{h,t}) - e_{h,t} p_t - D_t \right] \mathrm{d}t,$$

where the paths p_T and $e_f(T)$ are functions of the choice variable path $e_{h,T}$ and implicitly defined by Eqs. (3.1) and (3.2).

Let e_h^* be the optimal domestic consumption path. The derivative of U for the FOC governing the optimal domestic consumption is

$$\begin{aligned} \frac{\mathrm{d}U(e_h^*)}{\mathrm{d}e_h(t)} &= e^{-\rho t} u'(e_{h,t}) - e^{-\rho t} p_t - \int_0^\infty e^{-\rho s} e_{h,s} \frac{\mathrm{d}p_s}{\mathrm{d}e_{h,t}} \mathrm{d}s \\ & \quad - \int_t^\infty e^{-\rho s} d_s \mathrm{d}s - \int_0^\infty e^{-\rho s} d_s \int_0^s \frac{\mathrm{d}e_{f,u}}{\mathrm{d}e_{h,t}} \mathrm{d}u \mathrm{d}s, \end{aligned} \tag{3.3}$$

where the last two terms are implied by $\frac{\mathrm{d}D_s}{\mathrm{d}e_{h,t}} = \underbrace{\frac{\partial D_s}{\partial E_s}}_{d_s} \frac{\mathrm{d}E_s}{\mathrm{d}e_{h,t}}$ and $\frac{\mathrm{d}E_s}{\mathrm{d}e_{h,t}} = \frac{\partial E_s}{\partial e_{h,t}} + \int_0^\infty \frac{\mathrm{d}e_{f,u}}{\mathrm{d}e_t} \frac{\partial E_s}{\partial e_{f,t}} \mathrm{d}u$, implying $\frac{\mathrm{d}E_{s<t}}{\mathrm{d}e_t} = \int_0^s \frac{\mathrm{d}e_{f,s}}{\mathrm{d}e_t} \mathrm{d}s$ and $\frac{\mathrm{d}E_{s\geq t}}{\mathrm{d}e_t} = \int_0^s \frac{\mathrm{d}e_{f,s}}{\mathrm{d}e_t} \mathrm{d}s + 1$, leading to $\int_0^\infty e^{-\rho s} \frac{\mathrm{d}D_s}{\mathrm{d}e_t} \mathrm{d}t = \int_t^\infty e^{-\rho s} d_s \mathrm{d}s + \int_0^\infty e^{-\rho s} d_s \int_0^s \frac{\mathrm{d}e_{f,u}}{\mathrm{d}e_{h,t}} \mathrm{d}u \mathrm{d}s$, and where the envelope theorem has allowed us

[69] By a simple variational argument: in equilibrium, if t is the optimal extraction period, we know that the time derivation of his NPV sales profit for extractions immediately before or after t must be zero, i.e., p_t must be such that for $C = c_t$ we have $\partial \frac{e^{-\rho t}(p_t - C)}{\partial t} \stackrel{!}{=} 0$. Solving for p_t yields Eq. (3.2). Alternatively, the problem can be solved using a Hamiltonian as below, Eqs. (3.13) through (3.14); the second term on the RHS in Eq. (3.2) corresponds to λ.

[70] The case where instead *both* consuming regions buy the exhaustible resource *strategically* from the competitive seller is treated in Karp and Newbery (1993); however, they do not take pollution into account.

to ignore the interdependence of the optimal *domestic* consumption rates from different time periods.

With decentralized consumers that equate private costs and benefits, and with a potential fuel or emissions tax τ_t, domestic consumption is governed by

$$u'(e_h(t)) = p_t + \tau_t,$$

which, for standard regularity conditions guaranteeing a single interior solution, provides an implicit one-to-one mapping between the tax τ_t and the domestic consumption $e_{h,t}$.

As the FOC requires $\frac{\mathrm{d}U}{\mathrm{d}e_{h,t}} = 0$, the tax path τ_T that sustains the optimal domestic consumption level $e_h(T)^*$ is thus defined by

$$\tau_t^* = \underbrace{\int_t^\infty e^{-\rho(s-t)} d_s \mathrm{d}s}_{\text{direct damage}} + \underbrace{\int_0^\infty e^{-\rho(s-t)} e_h(s) \frac{\mathrm{d}p_s}{\mathrm{d}e(t)} \mathrm{d}s}_{\text{terms-of-trade}} + \underbrace{\int_0^\infty e^{-\rho(s-t)} d_s \int_0^s \frac{\mathrm{d}e_{f,u}}{\mathrm{d}e_{h,t}} \mathrm{d}u \mathrm{d}s.}_{\text{leakage}} \quad (3.4)$$

Partial derivatives are to be considered with domestic consumption in the other periods, $e_{h,v\neq t}$, held fixed, and the foreign consumption path, $e_{f,T}$, and the fuel price p_T adjusting according to Eqs. (3.1) and (3.2).

The optimal tax is thus governed by three distinct effects. The natural interpretation of the first is the direct pollution effect: the fuel consumption at time t directly increases cumulative emissions for all subsequent times, implying a climate cost d_s for all periods from t on, $s \geq t$. It corresponds to the optimal global tax, τ^*_{Pigou} from Eq. (3.8). The consumption choice affects the price $p(s)$ paid for fuel imports $e(s)$, adding the second component, reflecting the terms-of-trade effect. Finally, the emissions in the foreign region, $e_f(s)$, are affected as well, leading to the third component, expressing the leakage effect.

3.4 Disentangling Climate and Terms-of-Trade Effect

The optimality condition Eq. (3.4) for the tax defines the optimal carbon tax only implicitly, giving the optimal tax path τ_T^* as a function of variables that themselves also depend on the tax path itself (or on the domestic consumption path which is influenced by the tax). For non-trivial, empirically calibrated models, it cannot be presumed that the optimal tax path can be calculated analytically. For a numerical model, the optimal carbon tax path, however, can be calculated as the tax path τ_T^* (or its consumption counterpart e_T^*) that maximizes the overall domestic welfare U.

When goods producers exert market power, importing (and exporting) regions may resort to 'terms-of-trade' tariffs in order to change the equilibrium prices (the 'terms-of-trade') of the goods to their advantage and to thereby increase domestic welfare (Brander and Spencer, 1984; Pomfret, 2008). This holds in particular for fossil fuels, which are produced with decreasing returns overall and whose owners reap scarcity rents. Consuming regions can extract parts of the fuel scarcity rents with (positive) taxes on their domestic fuel consumption (Brander and Djajic, 1983). This may lead to global welfare losses. Therefore, such beggar-thy-neighbour policies are in the general case in conflict with free trade principles, and it is a major aim of the WTO to prevent such distorting policies. Nevertheless, for a fuel importing region, the terms-of-trade motive theoretically increases the regionally optimal emissions tax to a value beyond what would be justified for purely environmental reasons; in addition to the environmental damage related terms, the optimal tax contains a terms-of-trade component, as shown in Eq. (3.4). For investigative, but also for policy purposes, it seems relevant to separate the environmental component of the optimal emissions tax from the terms-of-trade component, as it is widely acknowledged that taxes imposed genuinely for the protection of the global climate seem acceptable, whilst genuine terms-of-trade taxes, distorting trade at the expense of other parties, seem problematic. The remainder of this section analyzes how this separation can be implemented, and how the environmental-only component, which is used in the remainder of the paper, can be calculated. The next section shows that the optimal emissions tax that disregards terms-of-trade benefits is closely related to a net present damage value adjusted dynamic leakage rate.

Subtracting the Terms-of-Trade-only Tax The most straightforward attempt to split the numerically calculated optimal fuel consumption tax into a trade part τ_{tot}^* and a climate (or emissions) part τ_e^* would be to start by calculating the optimal fuel tax in absence of pollution damage. This would yield the optimal terms-of-trade tax τ_{tot}^*, and one could then use the difference between the optimal total tax with pollution, τ^*, and this trade-only tax, τ_{tot}^*, to define the climate component, or the optimal climate-only tax, as $\hat{\tau}_e \equiv \tau^* - \tau_{tot}^*$. However, this is not a very precise method as imposing the climate tax reduces domestic fuel consumption. Consequently, the optimal terms-of-trade tax, which, for fossil fuels offered with prices that increase in the demanded quantity, also decreases along with a reduction in the imported quantity. Annex 2 confirms this intuition with a simple analytical model for the case of the optimal tax in a static environment with an external fuel producer and a single importer. It shows that the pollution-only tax calculated with the proposed terms-of-trade subtraction method is always lower than the natural rate of the pollution-only tax, and that even for small pollution damages and correspondingly small environmental taxes, the bias of the pollution-only tax is non-marginal. Nevertheless, at least under certain circumstances, the approximation may still lead to results that are quite closely related to the truly optimal pollution-only tax τ_e^*, with, in the Annex example, the calculated tax $\hat{\tau}_e$ always deviating by less than a quarter from τ_e^* (cf. Proposition 2 in Annex 2).

Compensation Method An alternative way to neutralize terms-of-trade effects in the calculation of the optimal pollution tax is to hypothetically require the domestic region to compensate external actors for losses they incur due to the domestic consumption tax, ignoring, however, foreign damage from pollution. Taking this compensation into account, the domestic region no longer has a direct incentive to influence prices of the imported fuel. That this leads to the optimal pollution-only tax, corresponding to the level identified as optimal in the hypothetical presence of leakage in the absence of price-effects in section 1 in a static model with an external producer and a passive fringe consumer (and the corresponding carbon leakage) is shown in Annex 3, and corresponds largely to what Böhringer et al. (2010) have shown in their static framework with fuel consumption of industrial sectors.

The remainder of this section extends this result to the dynamic case with exhaustible fuels. It shows that compensation payments can be used to disentangle the terms-of-trade and the climate motive for the optimal unilateral fuel tax also in a dynamic framework with exhaustible fuels using a continuous time model with two fuel-consuming and one fuel-producing region (it is straightforward to extend the analysis to the case where fuel-production is distributed among the two consuming regions).

We use the same framework as in section 3, but with transfer payments. Thus, consider regional welfare as $U_r = z_r + \int_0^\infty e^{-\rho t} u_r(e_r(t)) - D_r(t)\mathrm{d}t$, where z is (present) consumption of a numeraire good, which, assuming perfect capital markets without borrowing constraints, can also be imagined as a shortcut for the NPV of a consumption path ζ_t, with $z = \int_0^\infty e^{-\rho t}\zeta_t \mathrm{d}t$; $u_r(e_r)$ as an instantaneous fuel consumption utility and D_r is the path of regional emission damages. For the domestic region, which may pay transfers Tr_f and Tr_e to the foreign region and the fuel producers, the budget constraint is $z_0 = z + Tr_f + Tr_e + \int_0^\infty e^{-\rho t} e(t) p_t \mathrm{d}t$. For the foreign region, potentially receiving the transfer Tr_f, it writes $z_{0f} + Tr_f = z_f + \int_0^\infty e^{-\rho t} e_f(t) p_t \mathrm{d}t$. The fuel producer's utility U_e is given as the level of consumption of a numeraire good, consisting of the NPV of fuel sales profits net of production costs plus a potential received transfer, Tr_e, $U_e = Tr_e + \int_0^\infty e^{-\rho t}\left(p_t - c_t\right) e_w(t)\mathrm{d}t$, with $e_w(t)$ the global fuel consumption, p_t the sales price, and c_t the extraction costs.

To calculate the climate effect separate from the terms-of-trade (t-o-t) effect, we switch off the t-o-t effect by requiring the domestic region to provide transfer payments that set off losses or gains the foreign region and the fuel producer would otherwise experience from the domestic fuel consumption (or emission) policy. The transfers compensate for changes in non-green welfare; that is, for a given policy, climate damage is not directly considered in the calculation of the compensation transfers.

We know that, absent any externality concerns, undistorted, decentralized consumption and production maximizes non-green overall output in terms of total present-discounted net output ignoring climate damages, derived from exhaustible resources. That is, social non-green surplus, $\int_0^\infty e^{-\rho t}\left[\underbrace{u(e(t)) + u_f(e_f(t))}_{\text{consumption value}} - \underbrace{(e(t) + e_f(t))\, c_t}_{\text{production costs}}\right] \mathrm{d}t$, is maximized without any policy influencing the regional consumers or distorting the fuel producer's behavior (Hotelling, 1931). The maximization problem for the domestic region implicitly accounting for the imposed transfers can thus be written as the problem of maximizing the sum of domestic and non-green foreign and producers' welfare normalized for the level of the transfer payments, denoted U^*. The fuel price p_t, paid by the consumers but received by the fuel producers, cancels out and only the extraction costs, c_t, as well as the climate costs for the domestic region, D_t, are overall subtracted from the regional consumption utilities:
$\max_{e_{h,T}} U^* = \int_T e^{-\rho t}\left[u_h(e_h(t)) + u_f(e_f(t)) - e_{w,t}c_t - D_t\right] \mathrm{d}t$, where both, the marginal extraction costs and the instantaneous damage defined as (increasing) functions of cumulative emissions, $c_t \equiv c\left(\int_0^t e_{w,s}\mathrm{d}s\right)$, $D_t \equiv D\left(\int_0^t e_{w,s}\mathrm{d}s\right)$, with e_f implicitly defined by Eqs. (3.1) and (3.2).

Assuming a single internal solution to obtain, the solution must satisfy the standard FOC. We thus develop

$$\begin{aligned}\frac{\mathrm{d}U^*}{\mathrm{d}e_h(t)} \quad = \quad & e^{-\rho t}u'(e_h(t)) + \int_0^\infty e^{-\rho s}u_f'(e_f(s))\frac{\mathrm{d}e_f(s)}{\mathrm{d}e(t)}\mathrm{d}s - e^{-\rho t}c_t \\ & - \int_0^\infty e^{-\rho s}\left[\frac{\mathrm{d}e_f(s)}{\mathrm{d}e_h(t)}c_s + e_{w,s}\frac{\mathrm{d}c_s}{\mathrm{d}e(t)}\right]\mathrm{d}s - \int_t^\infty e^{-\rho s}d_s\mathrm{d}s \\ & - \int_0^\infty e^{-\rho s}d_s\int_0^s \frac{\mathrm{d}e_{f,u}}{\mathrm{d}e_{h,t}}\mathrm{d}u\mathrm{d}s.\end{aligned}$$

From Eqs. (3.1) and (3.2) we have $u'_f(e_f(t)) = c_t + \int_t^\infty e^{-\rho(s-t)}\dot{c}_s \mathrm{d}s$. This implies

$$\frac{\mathrm{d}U^*}{\mathrm{d}e_h(t)} = e^{-\rho t}u'(e_h(t)) + \underbrace{\int_0^\infty e^{-\rho s}\left[\int_s^\infty e^{-\rho(u-s)}\dot{c}_u \mathrm{d}u\right]\frac{\mathrm{d}e_f(s)}{\mathrm{d}e(t)}\mathrm{d}s}_{I_1} - e^{-\rho t}c_t$$
$$-\underbrace{\int_0^\infty e^{-\rho s}e_{w,s}\frac{\mathrm{d}c_s}{\mathrm{d}e(t)}\mathrm{d}s}_{I_2} - \int_t^\infty e^{-\rho s}d_s\mathrm{d}s - \int_0^\infty e^{-\rho s}d_s\int_0^s \frac{\mathrm{d}e_{f,u}}{\mathrm{d}e_{h,t}}\mathrm{d}u\mathrm{d}s.$$

We now show that terms I_1 and I_2 in Eq. (3.3) cancel out, yielding Eq. (3.5). From the definition of the extraction costs we have $\frac{\mathrm{d}c_s}{\mathrm{d}e_h(t)} = c'_s \cdot \left[\{1 \text{ if } s \geq t \text{ else } 0\} + \int_0^s \frac{\mathrm{d}e_f(u)}{\mathrm{d}e_t}\mathrm{d}u\right]$. Therewith, I_2 rewrites $\int_t^\infty e^{-\rho s}e_{w,s}\left[c'_s \cdot 1\right]\mathrm{d}s + \int_0^\infty e^{-\rho s}e_{w,s}\left[c'_s \cdot \int_0^s \frac{\mathrm{d}e_f(u)}{\mathrm{d}e_h(t)}\mathrm{d}u\right]\mathrm{d}s$. Noting that $c_t \equiv c\left(\int_0^t e_w(t)\mathrm{d}t\right)$ implies $\dot{c}_t = c'_t \cdot e_w(t)$, I_2 simplifies to $\int_t^\infty e^{-\rho s}\dot{c}_s\mathrm{d}s + \int_0^\infty e^{-\rho s}\dot{c}_s \cdot \int_0^s \frac{\mathrm{d}e_f(u)}{\mathrm{d}e_h(t)}\mathrm{d}u\mathrm{d}s$. Seeing further that $\int_0^\infty e^{-\rho s}\dot{c}_s \cdot \int_0^s \frac{\mathrm{d}e_f(u)}{\mathrm{d}e_t}\mathrm{d}u\mathrm{d}s$ is a simple double integral over the open 'area' defined by $u \leq s$, we know $\int_0^\infty e^{-\rho s}\dot{c}_s \cdot \int_0^s \frac{\mathrm{d}e_f(u)}{\mathrm{d}e_t}\mathrm{d}u\mathrm{d}s = \int_0^\infty \int_u^\infty e^{-\rho s}\dot{c}_s \frac{\mathrm{d}e_f(u)}{\mathrm{d}e_t}\mathrm{d}s\mathrm{d}u$, which, switching u and s yields the same as I_1. Terms I_1 and I_2 thus cancel out in Eq. (3.3) and we get

$$\frac{\mathrm{d}U^*}{\mathrm{d}e(t)} = e^{-\rho t}u'(e_h(t)) - e^{-\rho t}c_t - \int_t^\infty e^{-\rho s}\dot{c}_s\mathrm{d}s - \int_t^\infty e^{-\rho s}d_s\mathrm{d}s - \int_0^\infty e^{-\rho s}d_s\int_0^s \frac{\mathrm{d}e_{f,u}}{\mathrm{d}e_{h,t}}\mathrm{d}u\mathrm{d}s. \tag{3.5}$$

The FOC of the maximization problem thus yields, with a multiplication by $e^{-\rho t}$ to switch from a present to current value expression,

$$u'(e_h(t)) = c_t + \int_t^\infty e^{-\rho(s-t)}\dot{c}_s\mathrm{d}s + \int_t^\infty e^{-\rho(s-t)}d_s\mathrm{d}s + \int_0^\infty e^{-\rho(s-t)}d_s\int_0^s \frac{\mathrm{d}e_{f,u}}{\mathrm{d}e_{h,t}}\mathrm{d}u\mathrm{d}s. \tag{3.6}$$

With a tax of rate τ_t, the decentralized consumer decisions are governed by the private FOC, equating private benefits and costs,

$$u'_h(e(t)) \stackrel{!}{=} p_t + \tau_t.$$

Recall from Eq. (3.2) that the competitive suppliers set $p_t = c_t + \int_t^\infty e^{-\rho(s-t)}\dot{c}_s\mathrm{d}s$. For the tax τ_t to sustain the optimal consumption level according to Eq. (3.6), we thus

require
$$\tau_t^* = \underbrace{\int_t^\infty e^{-\rho(s-t)} d_s \mathrm{d}s}_{\text{direct damage}} + \underbrace{\int_0^\infty e^{-\rho(s-t)} d_s \int_0^s \frac{\mathrm{d}e_{f,u}}{\mathrm{d}e_{h,t}} \mathrm{d}u \mathrm{d}s.}_{\text{leakage}} \tag{3.7}$$

The first term on the RHS in Eq. (3.7) is the direct domestic pollution component as the net current value[71] of the response of the damage in all periods from time t on to the change of cumulative emissions from increased domestic emissions at t, which also equals the optimal global tax, τ^*_{Pigou} from Eq. (3.8). The second term is the leakage component. Precisely, it is the net current value of all damage changes throughout time as a response to the cumulative change of foreign emissions in reaction to the domestic consumption change at time t. These direct pollution and leakage components, which were present also in the optimal unilateral tax without compensation, Eq. (3.4), together determine the optimal unilateral 'climate-only' tax level. Terms-of-trade effects are absent; as conceived, the hypothetical compensation payment has neutralized them.

Optimal Global Policy From its construction through the hypothetical compensation mechanism, which ensures that all agents' welfares are fully accounted for, it is clear that τ_t^* in Eq. (3.7) in the absence of leakage must correspond to the optimal worldwide pollution tax. That is, the optimal global policy is

$$\tau^*_{\mathrm{Pigou},t} = \int_t^\infty e^{-\rho(s-t)} d_s \mathrm{d}s > 0. \tag{3.8}$$

Indeed, Edenhofer and Kalkuhl (2010) show that this corresponds to the social planner's choice, equalizing the competitive and the normative extraction and emission path. Intuitively, the level of the optimal global tax at time t equals the net current value of all future marginal damages from a unit of emission added, d_t. Competition or climate-independent resource conservation effects do not enter the optimal tax because pricing issues only correspond to a shift of rents between the buyers and sellers without changing the overall rent, and the supplier's dynamic pricing behavior leads to the optimal conservation of the resource in absence of externalities.

Given that for convex damages we have $\dot{d}_s > 0$ during the fuel consumption phase, the tax is strictly growing, $\dot{\tau}^*_{Pigou} > 0$. Also, Eq. (3.7) implies that the tax grows at less than the interest rate (cf. also van der Ploeg, forthcoming): we have $\dot{\tau}^*_{Pigou,t} = \rho\tau^*_{Pigou,t} - d_t$, implying that $\tau^*_{Pigou,t}$ grows at a rate $g_{Pigou,t} = \rho - \underbrace{d_t/\tau^*_{Pigou,t}}_{>0} < \rho$. We thus emphasize:

Proposition 3a. Absent leakage effects, and given convex damages from cumulative

[71] See section 3.5 for a definition of the net *current* value.

emissions, the tax of the optimal pollution policy, τ^*_{Pigou}, is positive and strictly rising, growing at a rate $g_{\tau,Pigou}$ below the interest rate ρ,

$$\tau^*_{Pigou} > 0 \quad \text{and} \quad 0 < g_{\tau,Pigou} < \rho.$$

3.5 Definition of Leakage Rates and Terminology

The previous sections have shown that the tax rates of the optimal, green-only unilateral tax path are described in terms of the damage effect of (i) current domestic consumption, and (ii) the response of *foreign* consumption *at every period* to *current domestic* consumption changes (Eqs. (3.4) and (3.7)). That is, the optimal tax at time t does not directly depend on the response of *domestic* emissions at other periods, $e_{h,v\neq t}$, to changes in current emissions at time t, $e_{h,t}$.[72] Correspondingly, we here focus on leakage rates expressing the foreign offsetting of instantaneous domestic emission reductions when other domestic emissions are held constant. Given the results from the previous sections, these will be the leakage rates that are relevant for the optimal unilateral tax path.

First, as a concept that is probably the most compatible with both the existing literature on emission leakage as well as with a very casual interpretation of emission leakage, we define the absolute leakage rate (ALR), as the total fraction of some (anticipated) instantaneous emission savings that is offset by foreign emission changes,

$$\text{Absolute leakage rate: } \text{ALR}_t \equiv \int_T \frac{-\mathrm{d}e_{f,v}}{\mathrm{d}e_{h,t}} \mathrm{d}v,$$

where the considered time horizon T starts at the period from which tax t is anticipated, and, theoretically, lasts until infinity. In the numerical simulations below, which focus on committed policies, we will generally assume T to start at the present date (expressed as $t = 0$, i.e., $T = (0, \infty)$). This has the advantage of providing results both for unanticipated taxes (as the leakage rate for the initial period) as well as for anticipated taxes (as leakage rates for later taxes that are anticipated from now on).

Expressing the standard in climate economics to discount future emissions, we define the

[72] This is not necessarily a surprise given that the domestic taxes in the other periods are assumed to be optimal as well. From the point of view that leakage generally implies that the optimal taxes here fall short of the perfectly internalizing Pigouvian, this result may, however, still not necessarily have been expected. It is nevertheless intuitive in the sense that the optimal tax path is directly derived from the optimal domestic consumption path for which we know that the envelope theorem implies that derivatives of choice variables, from other time periods with respect to current choice variables, become irrelevant in the optimality condition.

NPV leakage rate, NLR, as the fraction of domestic emission reductions offset abroad in terms of the NPV value of emissions,

$$\text{NPV leakage rate: } \mathrm{NLR}_t \equiv \int_T e^{-\rho(v-t)} \frac{-\mathrm{d}e_{f,v}}{\mathrm{d}e_{h,t}} \mathrm{d}v, \tag{3.9}$$

with ρ the corresponding present-discount rate for the emissions.

Finally, the form of leakage that is truly relevant for welfare concerns and directly related to the optimal unilateral overall or green-only carbon tax, is what we here name the damage leakage rate (DLR). DLR is defined as the fraction, in NPV terms, of the direct damage reduction related to a domestic emission cut that is offset by damage increases implied by the response of foreign emissions throughout the considered time horizon,

$$\begin{aligned} \text{Damage leakage rate: } \mathrm{DLR}_t &\equiv \frac{\int_0^\infty \frac{-\mathrm{d}e_{f,u}}{\mathrm{d}e_{h,t}} \cdot \int_0^\infty e^{-\rho s} \frac{\partial D(s)}{\partial e_{f,u}} \mathrm{d}s \mathrm{d}u}{\int_0^\infty e^{-\rho u} \frac{\partial D(u)}{\partial e_{h,t}} \mathrm{d}u} \\ &= \frac{\int_0^\infty e^{-\rho(s-t)} d_s \int_0^s \frac{-\mathrm{d}e_{f,u}}{\mathrm{d}e_{h,t}} \mathrm{d}u \mathrm{d}s}{\int_t^\infty e^{-\rho(s-t)} d_s \mathrm{d}s}, \end{aligned} \tag{3.10}$$

where it is important to note that $\frac{\partial D}{\partial e_{r,t}}$ is the partial derivative (as opposed to the total derivatives $\frac{\mathrm{d}(\cdot)}{\mathrm{d}(\cdot)}$ taken elsewhere) of damage D with respect to emissions of region r at time t, $e_{r,t}$, holding emissions elsewhere (and in other periods) constant.[73] The second equality follows from what we noted for the FOC in the section on the unilateral committed policy in section 3.3 (Eq. (3.3)). With this definition of the DLR, the optimal pollution-only tax from the committed policy, Eq. (3.7) can be rewritten as

$$\tau^* = \underbrace{\int_t^\infty e^{-\rho(s-t)} d_s \mathrm{d}s}_{\text{direct damage}} \cdot (1 - \mathrm{DLR}_t) = \tau^*_{Pigou} \cdot (1 - \mathrm{DLR}_t), \tag{3.11}$$

confirming that DLR_t is the welfare relevant dynamic equivalent of the simple leakage rate of a static model (cf. Annex 1). We thus see, in combination with Proposition 3a, that

Proposition 3b. The tax of the optimal pollution policy, τ_t^*, is increasing over time when DLR_t decreases, and it can only be decreasing when DLR_t increases. The tax is negative when $\mathrm{DLR}_t > 1$:

$$\dot{\mathrm{DLR}}_t < 0 \implies \dot{\tau}_t^* > 0 \qquad \dot{\tau}_t^* < 0 \implies \dot{\mathrm{DLR}}_t > 0 \qquad \mathrm{DLR}_t > 1 \implies \tau_t^* < 0.$$

[73] The derivative $\frac{\partial D}{\partial e_{r,t}}$ does not depend on r.

Note that for a linear damage, that is, for constant marginal damage d, Eq. (3.10) simplifies to $\text{DLR}_t = \rho e^{\rho t}\int_0^\infty e^{-\rho s}\int_0^s \frac{\mathrm{d}e_{f,u}}{\mathrm{d}e_{h,t}}\mathrm{d}u\mathrm{d}s = \rho e^{\rho t}\int_0^\infty e^{-\rho s}\frac{\mathrm{d}e_{f,s}}{\mathrm{d}e_{h,t}}/\rho\mathrm{d}s$, which is equivalent to Eq. (3.9). For a linear damage function, we thus have $\text{NLR} = \text{DLR}$.

Further terminology used throughout the paper:

The terminology 'optimal' tax as used in this paper can be somewhat confusing. We essentially consider unilateral taxes which per se are economically inefficient compared to global taxes. When we write about the optimal tax, we typically simply mean the best among the unilateral taxes.

We use the term '*current*' for a specific time t as the value seen *from time t* on, i.e., opposed to the concept of '*present*' values which means that future values are expressed in their value seen *from today's time, time 'zero'*. To express it analytically, we have, for a utility or monetary real value V_s occurring at time s, a today's ($t = 0$) net present value $e^{-\rho s}V_s$ but a net current value at time t of $e^{-\rho(s-t)}V_s$.

To not complicate the descriptions unnecessarily, with a slight impreciseness we will usually use expressions such as a region's 'oil reduction'. This is always used as a shortcut to mean a regional reduction of the *consumption* of that fuel. Along the same line, we will in some cases use shortcuts such as 'oil leakage' when we mean the *overall emission* offsetting reactions in the remainder of the world as a reaction to the domestic reduction of *oil* consumption. That is, the term is not to be interpreted as concerning only the change in foreign *oil* consumption (emissions), but in the induced total foreign fuel consumption emission change as an equilibrium response to the domestic *oil* consumption change.

Part II Dynamic numerical model

This part analyzes optimal unilateral (OECD) carbon taxes, or the green component thereof, using a calibrated dynamic numerical fuel market model, accounting for the two dominant, and very distinct, fossil fuels – oil and coal. Sections 3.6 and 3.7 describe and illustrate the model, section 3.8 illustrates the results from the theory in Part 1. Finally, section 3.9 uses the model in a sequence of setups to examine the effect of various crucial elements of the model in order to provide an understanding of the mechanisms that drive the fuel-channel leakage rates and the optimal taxes. It also works to provide quantitative estimates of magnitudes of the leakage rate and the corresponding taxes for specific scenarios.

3.6 Model

Setup The model contains two fuel consuming regions, the OECD and the rest of the world (ROW or Non-OECD), indexed by $r = \{o, n\}$. The OECD is assumed to consider emission taxes while the remainder of the world abstains and consumes in a decentralized fashion. The two fuels considered are oil and coal, indexed by $i = \{1, 2\}$.

A few words on the restriction to oil and coal as the two fossil fuels considered may be in order. First, the simulation results will already prove to be complex when we restrict the attention to the oil and coal. The interpretability would presumably be further complicated if gas were taken into account as well, and it is not clear whether relevant further insights would be gained. Additionally, currently 80 % of energy supply carbon emissions[74] stem from burning coal (43 %) and oil (36 %), and only 20 % from gas. Moreover, whilst gas is occasionally considered as the fuel of the future, in reality more than 50 % of the current growth of total global carbon dioxide emissions is attributable to coal, and 2/3 to coal and oil, with the remainder attributable to other sources, including gas. Furthermore, in the faster growing non-industrialized world the share of coal and oil in the growth of all CO_2 emissions exceeds 75 % (IEA, 2012). Finally, because gas has many features similar to oil, especially in terms of the exhaustibility and the convertibility of coal through gasification or liquefaction, to a certain degree one may interpret 'oil' in our model as representative of the ensemble of oil and gas, an approach also used by van der Ploeg and Withagen (2011).[75] The Discussion (section 3.11) speculates on how gas, and notably the currently increasing production of shale gas, could influence the model results.

The fuels are traded internationally at prices $p = [p_1, p_2]$. Regional fuel consumption is denoted by $x_r = [x_{r,1}, x_{r,2}]$. Following Golombek et al. (1995), instantaneous regional welfare W_r is defined with three linearly separable terms: (i) utility from regional energy consumption Y, $u_r(Y_r)$, (ii) the total regional costs for energy provision $c_r(Y_r)$, and (iii) the regionally perceived environmental costs D_r, which we model as a function of cumulative global emissions, $D_r(E)$, where $E_t = \int_0^t e_s \mathrm{d}s$, with e_t global emissions at time t:[76]

$$W_r = u_r(Y_r) - c_r(Y_r) - D_r(E). \tag{3.12}$$

[74] Energy supply is responsible for 83 % of all anthropogenic GHG emissions (IEA, 2012).

[75] In a similar fashion, climate and energy CGE models tend to treat oil and gas as a separate constant elasticity of substitution (CES) sub-aggregate, nested under another CES where the oil-gas sub-aggregate figures parallel to coal or even to different types of coal, see, e.g., Böhringer and Löschel (2004) and Böhringer et al. (2008).

[76] See Golombek et al. (1995) for the derivation of this reduced form structure from a regional economic setting where fossil fuels are used also as an intermediate input for final goods production.

Each of the variables in Eq. (3.12) exists at each point in time $t \in [0..T]$, and the total regional welfare is defined as the present discounted integral of all instantaneous welfare values:

$$\mathbb{W} = \int_T e^{-\rho_u \cdot t} W_{r,t} \mathrm{d}t,$$

where $\rho_u < 1$ is the time discount rate of the consuming regions.

The fuel consumers' energy consumption utility is isoelastic in the consumption of energy Y, $u_r = \frac{1}{1-\alpha}\xi Y^{1-\alpha}$, $\alpha > 0$, implying an isoelastic demand for energy Y, with elasticity $\varepsilon = -\frac{1}{\alpha} < 0$. Section 3.9.6 extends the model to a growing regional demand. Energy Y is the sum of a constant elasticity of substitution (CES) aggregation of oil and coal consumption, $F(x_1, x_2) = \left(ax_1^\delta + (1-a)x_2^\delta\right)^{1/\delta}$, implying an elasticity of substitution $\sigma = 1/(1-\delta)$ plus, if allowed for, the consumption of a clean backstop B, $Y = F + B$. The backstop may be provided at any given demand rate (infinite elasticity) for an exogenous price which may vary over time.

As a property of the CES aggregation function, the unitary fuel aggregate cost, c_F, is $c_F(p_{x_1}, p_{x_2}) = \left(a^\sigma p_{x_1}^{1-\sigma} + (1-a)^\sigma p_{x_2}^{1-\sigma}\right)^{\frac{1}{1-\sigma}}$, and, for a specific aggregate fuel consumption F, oil and coal consumption becomes $x_1 = F \cdot \left(\frac{ac_F}{p_{x_1}}\right)^\sigma$, $x_2 = F \cdot \left(\frac{(1-a)c_F}{p_{x_2}}\right)^\sigma$. Supply of both the aggregate fuel and (if not idled) the clean backstop is readily modeled with a complementary slackness condition with respect to the weakly positive difference of their cost to the overall energy price p_Y,

$$\begin{aligned} F \geq 0 \quad &\perp \quad p_F - p_Y \geq 0 \\ B \geq 0 \quad &\perp \quad p_B - p_Y \geq 0. \end{aligned}$$

Because the regional calibration of the demand structure to today's regional oil and coal consumption rates implies regional share parameters a in the CES fuel aggregation function, the *aggregate* fuel energy price will in general differ across the regions, wherewith also the time of the introduction of the backstop will not exactly coincide in the two regions.

The model allows for an endogenous production of synthetic oil from coal by liquefaction as soon as the relative fuel prices make the process economic, given a specified overhead process cost and conversion efficiency, again implemented by a corresponding complementarity slackness condition.

An alternative way to model fuel consumption utility would be to consider a utility described by a polynomial quadratic in fuel consumption as Golombek et al. (1995) did when discussing fuel-specific leakage rates in their static model.[77] Here, our approach based on the CES fuel aggregate has two distinct advantages. First, it allows us to choose any desired value for each elasticity parameter when we calibrate the model to current fuel consumption/price pairs. Golombek et al. could only choose values for half of the (cross or direct price) elasticity parameters, and the observed consumption/price pairs implicitly defined the other half of the parameters. Second, and most importantly, the isoelastic fuel demand and the CES aggregation ensure that the imposed elasticities are valid also for fuel prices that are (far) away from the region of the original calibration point. As the progressive depletion of the fuels substantially increases fuel prices in the long run, this seems to be a crucial feature for our dynamic model.

Suppliers are assumed to sell their fuels on the international market under perfect competition. The exhaustibility of the fuels is modeled with an extraction cost curve that indicates the marginal cost of extraction after a specific cumulative amount of the fuel has been extracted. This depletion concept is the logical consequence of the Herfindahl rule (Herfindahl, 1967) which states that (given positive real-interest rates) profit-maximizing resource owners extract the fuels ordered in a sequence according to extraction costs: the resources with the lowest extraction costs are extracted first, and the ones with the highest extraction costs are extracted last.[78] Given this standard rule, and assuming the resource owners discount their net revenues with a revenue discount rate $\rho_{\text{res}} > 0$, a current-value Hamiltonian for the profit maximization problem for the owners of one specific fuel reads as follows:

$$\begin{aligned} \mathcal{H} &= r_t \cdot (p_t(r_t) - c(A_t)) - \lambda_t r_t \\ &\quad \text{s.t.} \quad \dot{A}_t = r_t \text{ and } A_0 = 0, \text{ i.e. } A_t = \int_0^t r_s \mathrm{d}s, \end{aligned} \tag{3.13}$$

where r_t is the amount of the fuel extracted at time t, A_t is the cumulative amount of the fuel extracted from the initial period up to time t, normalized to 0 for $t = 0$, $c(A)$ is the marginal extraction cost after the extraction of the A units of fuel that could be extracted at the lowest costs, and $p_t(r)$ is the inverse demand for the considered fuel at time t: the price p_t results on the international fuel market if r units of the fuel are supplied (with non-zero cross-price elasticities of fuel demand p_t may also depend on the amount of the other fuel supplied at time t).

The FOCs of the Hamiltonian in Eq. (3.13) yield the following stationary condition and

[77] See Habermacher (2011) for a dynamic application using also a quadratic utility function.

[78] While in a simple theoretical framework this rule should hold not only in a monopolistic but also in a competitive framework, e.g. Beermann et al. (2011) give reasons why this rule is often only an approximation to reality.

canonical equation:

$$\frac{\partial \mathcal{H}}{\partial r_t} = 0: \quad p_t(r_t) \overset{!}{=} c(A_t) + \lambda_t \tag{3.14}$$

$$\dot{\lambda}_t = \rho_{\text{res}}\lambda_t + \frac{\partial \mathcal{H}}{\partial A_t}: \qquad \dot{\lambda}_t \overset{!}{=} \lambda_t \rho_{\text{res}} - \dot{c}_t,$$

where we define $c_t \equiv c(A_t)$,[79] and at time t λ_t is the shadow value for a marginal unit of resource stock after the cumulative extraction of A_t previous units. As the stationary condition (Eq. (3.14)) shows, the resource shadow value is the difference between the price that the resources achieve on the market and the extraction costs, that is, the per-unit resource rent received by the resource-owner for sales at time t.

Calibration The fuel demand (or utility) parameters are calibrated according to the current regional consumption of oil and coal at current prices in the OECD and the Non-OECD region (see Annex 4) and to the desired direct- and cross-price elasticities of the demand.[80] Interestingly, whilst oil consumption is 15 % lower in the ROW than in the OECD, coal consumption in the ROW is almost twice that of the OECD. Similarly to Golombek et al. (1995), in the main calibration we choose an elasticity slightly below unity and a weak substitutability of the fuels, setting $\varepsilon = -0.9$ and $\sigma = 1.1$. The weak substitutability between oil and coal mainly expresses the difficulty of replacing oil, in its major applications, by the solid fuel coal. (Note, as oil reserves are much more restricted than coal reserves, throughout our model simulations it will essentially be oil whose scarcity becomes relatively stronger over time, implying that the possibility of substitution of oil by coal is of relevance rather than the inverse).[81] The possibility of deriving synthetic oil from coal liquefaction (also called the coal-to-liquids, CTL, process) will be modeled as a separate process.

The curve of the extraction cost for oil as a function of cumulative extractions is implicitly defined through its inverse, the function giving the cumulative extraction A for a specific marginal cost, $A(c)$. The following functional form appears to allow a very good fit to the extraction cost curve by Rogner (1997),

$$A(c) = p_3/(1 + \exp((c - p_1)/p_2)) + p_4 \exp((c - 45)/p_5).$$

[79] Note that as $r_t = \frac{\partial A_t}{\partial t}$, we have $\dot{c}_t \equiv \frac{\partial c(A_t)}{\partial t} = \frac{\partial A_t}{\partial t}\frac{\partial c(A_t)}{\partial A_t} = r_t \frac{\partial c(A_t)}{\partial A_t}$.

[80] The clean backstop is considered as absent or prohibitively expensive at this stage.

[81] Golombek et al. (1995) used -0.9 for the direct price elasticity for the fuel consumptions in the OECD and -0.75 for the ROW, and they used cross-price elasticities of 0.1, on average. Here, the larger demand elasticity (in absolute terms) in the Non-OECD region represents the interpretation that as economies of the developing countries progress over time their fuel demand structure may approach that of the developed countries. In an overview, Michielsen (2011) lists cross-price elasticities from various empirical studies, averaging 0.06 from oil (and gas) to coal and 0.12 from coal to oil.

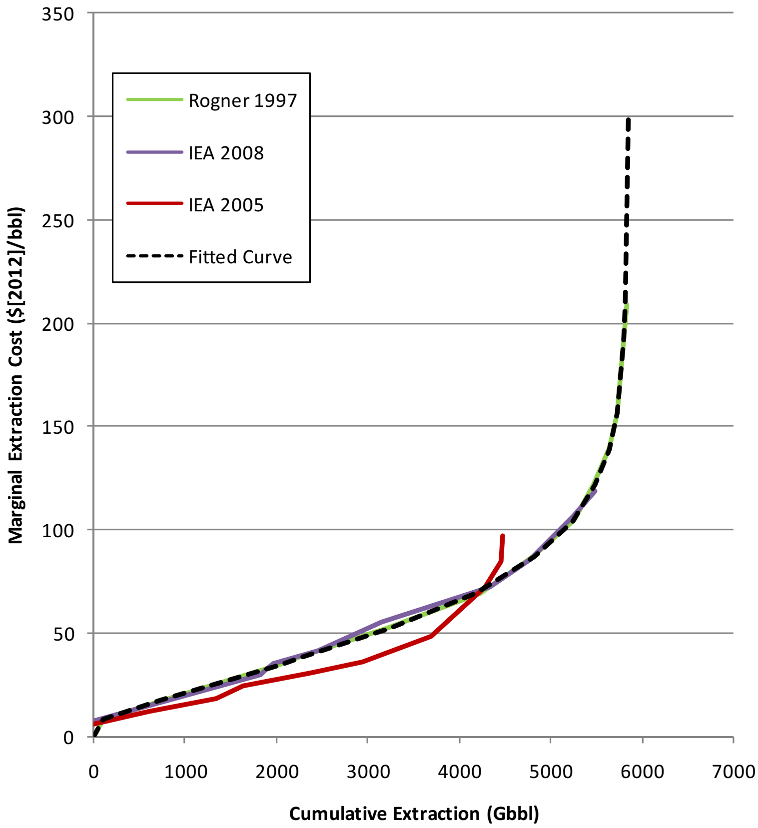

Figure 3.1: Oil extraction cost curves

We thus calibrated the parameters p_1 through p_5 to the extraction cost curve by Rogner (1997). As Fig. 3.1 shows, this fit provides an almost perfect correspondence to the Rogner curve. Moreover, there is a very good correspondence between the more recent IEA (2008) and an acceptable correspondence to the IEA (2005) projections.[82]

Indicating extraction costs for up to 1740 Gt coal, the coal cost curve in Rogner (1997) covers only a relatively modest fraction of the totally estimated resources of 16 000 Gt coal (DERA, 2012). Moreover, as Rogner notes, he models coal reserves in less detail than oil, which likely is a reason for the roughness of his estimated cost curve, replicated in Fig. 3.2. Given that historically coal prices were relatively low, around 30 \$/t in 2000, and today they fluctuate around 100 \$/t (DERA, 2012; EIA, 2013a), with a relevant fraction of the currently rather high prices probably explained by the unprecedented growth of worldwide coal consumption in the current millennium[83] rather than by a genuine long-term extraction cost increase, it was here decided to consider an actual

[82] All curves are inflation adjusted to \$2012.

[83] Worldwide coal consumption used to stagnate before the beginning of this millennium, with annual growth rates averaging -0.3 %. The dash for coal, notably in Asia, has lead to an average coal consumption growth rate of 4.6 % per year from 2000 through 2011 (own calculations based on EIA, 2013a).

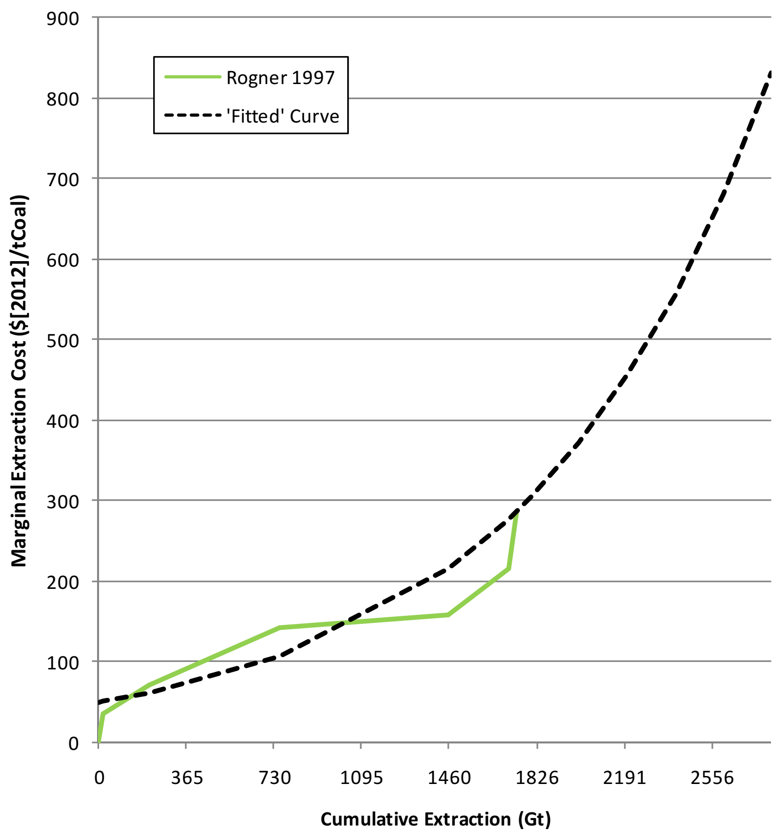

Figure 3.2: Coal extraction cost curves

coal extraction cost of 50 \$/t, and to assume an exponentially increasing extraction cost curve that matches the extraction 'cost and cumulative amount'-data pair for the largest quantity considered in Rogner (marginal cost of 286 \$/t after 1740 Gt extracted); that is, the curve given by $c = 50\,\$/\mathrm{t}\, e^{A/996\,\mathrm{Gt}}$. Fig. 3.2 shows how this curve provides a compromise between the general idea of a smooth, convex extraction cost curve, and the data points from the rough, convex and concave projection of Rogner (1997).

We generally consider two discounting schemes. Scheme 1 consists of equal discount rates for the fuel consumers and for the fuel producer, $\rho_u = \rho = 3\,\%$, as an attempted compromise between the in reality probably often rather large discount rates of fuel extractors and the potentially limited impatience of a regional planner in the fuel consuming regions.[84] Scheme 2 assumes a discount rate of only $\rho_u = 0.5\,\%$ in the consuming region and a higher fuel owner discount rate $\rho = 5\,\%$.[85] The emission intensity is

[84]Recall that the compensation method for disentangling the pollution and the terms-of-trade component of the optimal unilateral fuel tax described in Part 1 assumes equal discount rates for all involved actors.

[85]In the application of the model, the *consumers'* discount rate will essentially matter for the NPV calculation of future climate damages. At least for the case without growth, the modest 5 ‰ discount rate can be seen as a compromise between different, prominent views on climate discounting. Nordhaus

$0.43\,tCO_2$/bbl for genuine oil and $2.8\,tCO_2$ per ton for coal.

In the scenarios that consider liquefaction, the process is assumed to require 1 ton of coal per 2 barrels of synthetic oil produced (DOE/NETL, 2006; Bartis et al., 2008).[86] Whilst the final product, the synthetic oil, has the same emission intensity at its final consumption as genuine oil (that is, in the domestic use, direct emissions are the same), the use of half a ton of coal per barrel of oil produced implies excess emissions occurring during the production (and thus, abroad) that exceed the final consumption emissions. Consequently, overall synthetic fuel is more than twice as emission intensive. In addition to the input costs for this coal, the process is assumed to be subject to a constant additional fixed cost for each barrel of synthetic fuel produced, c_l, which takes on a value of c_l=15 \$/bbl (of produced synthetic oil) in the standard scenarios. In the simulations that allow for liquefaction, the overall costs of the process represent an upper bound for the oil sales price such that any demand that cannot be met by the standard oil supply for that price will be provided as synthetic fuel from coal-liquefaction.

When the clean backstop is considered, its price is assumed to approach an oil price equivalent of 200 \$/bbl-eq., with an initial price starting at 500 \$/bbl-eq., and the difference decaying exponentially at an annual rate of 2 %. To cover the period for which the considered processes imply an interesting dynamics, the simulation period stretches up to over 400 years. Shorter horizons are used in the cases where the clean backstop outcompetes all fuel sales earlier.

For the most part, we will restrict our attention to the case without growth in the energy demand, but we do treat an extended model with a demand growth roughly following IEA projections in the model extension in section 3.9.6. That section also treats the case of convex damages, whilst the NPV emission leakage rates examined in other sections can also be thought of as damage leakage rates under the implicit assumption of a linear damage function (cf. section 3.5).

3.7 Illustration of Model Results

Fig. 3.3 illustrates the model behavior in a standard setup with demand growth and the endogenous regional emergence of the backstop, as well as endogenous liquefaction

(2008) suggests a pure-time discount factor for the emission damages of 1.5 % and Stern (2007) suggests 1 ‰. Any extended discussion of the reasons for higher and lower values for the controversial and important discount factor is beyond the aim of the present study whose purpose is exploratory rather than to provide precise quantitative results.

[86] In reality, the conversion factor depends on the type of coal used. While a rule-of-thumb estimate for the coal-to-liquids yield from bituminous coal is 2 (barrels of oil per ton of coal), it is slightly lower for subbituminous coal, about 1.8 (Bartis et al., 2008).

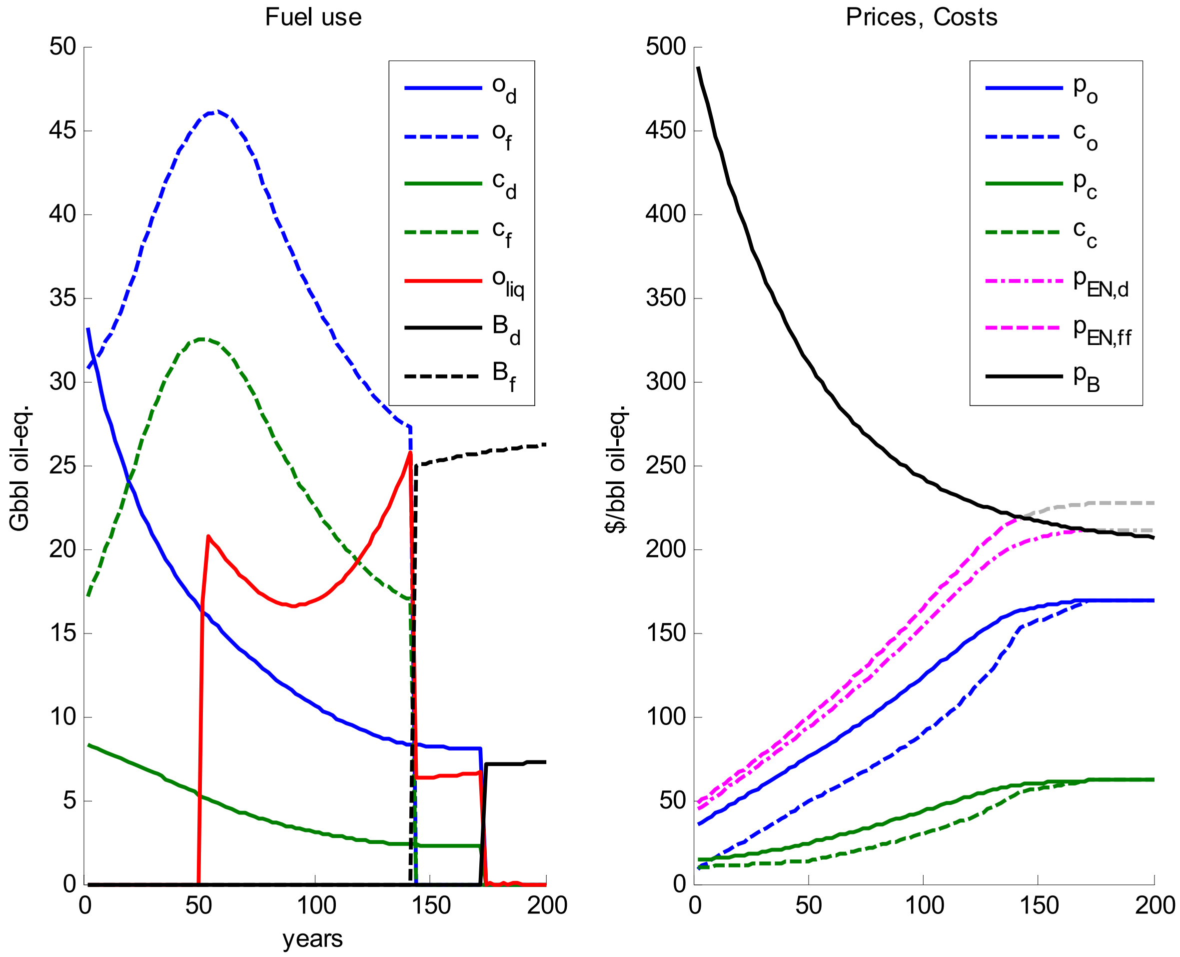

Figure 3.3: Simulation results with growth, backstop and liquefaction

(Annex 5 illustrates the outcome with constant demand and without liquefaction).

Plot 1 shows the fuel and backstop consumption paths. Blue shows domestic (i.e., OECD) and foreign (ROW) oil consumption, o_d and o_f. Green is coal consumption c_d and c_f, and black regional backstop consumption, B_d and B_f. Red is the production of synthetic liquid fuel through liquefaction, o_{liq}. Plot 2 shows the corresponding prices and extraction costs: blue and green are for oil and coal respectively, the market prices p_o and p_c, and extraction cost c_o and c_c. Profits λ_o and λ_c correspond to the difference between prices and costs. They converge towards zero as the backstop becomes competitive. Magenta shows the corresponding regional aggregate fossil energy prices $p_{EN,d}$ and $p_{EN,f}$, and the black line gives the exogenous backstop price, p_B.

Oil and coal consumption declines over time in the OECD. However, due to the large demand growth they increase initially in the ROW, peaking at around 50 years. With the relatively steep increase of the oil price compared to the more modest coal prices, liquefaction emerges in around year 50.[87] Because with coal the input to the CLT process is itself also exhaustible, crude and synthetic oil are thereafter used in parallel. When the regional aggregate fuel prices, $p_{EN,d}$ and $p_{EN,f}$, meet the backstop price p_B, in 140 years in ROW and in 170 years in the OECD, the regional fuel consumption stops and is replaced by the backstop. As the ROW dominates global consumption, resource rents already become very low by the time ROW fuel consumption stops, and they converge to zero by the time fossils also become redundant in the OECD. The slightly earlier switch from fossils to the backstop in ROW compared to the OECD is the natural consequence of the CES calibration to the current regional fuel consumption ratios, with the relatively larger coal consumption in the ROW. To see this, first note that oil represents the overwhelming share of fuel expenses today (in the OECD 90 % of total oil and coal costs, in the ROW 80 %, cf. Annex 4). In this case, the oil-share parameter approaches 1 in the CES function, and the closer to one the lesser coal is consumed in the benchmark for which the calibration is made. In parallel, the unitary aggregate energy cost becomes closer to the oil price, the fewer coal is consumed in the benchmark. In the price region for which the CES demand structure is calibrated in the model, modest increases of the amount of coal use in the benchmark (as is the case for the ROW compared to the OECD) do thus primarily increase the amount of coal used per unit of aggregate energy produced, and increase the unitary cost of aggregate energy, rather than to increase the implicit benchmark aggregate fuel energy consumption. This can make economic sense here if we believe that the relatively high coal consumption in the less developed ROW is partly explained by a less efficient use of coal rather than a higher overall need for final energy services. Given that the liquefaction process here prevents the oil price from diverging too far away from the lower coal price, the aggregate final energy costs in the OECD (which relies relatively more on oil) remain lower than those in the ROW (cf. plot 2 in Fig. 3.3). This explains that the switch to the backstop in the OECD occurs later than in the ROW. This is inverted in the case without liquefaction (Fig. A3.1 in Annex 5). Without the synthetic oil supply, the rapidly depleting oil becomes so expensive over time that the higher reliance of the OECD on oil makes its fossil energy aggregate pricier than in the ROW, implying that the OECD switches first to the clean backstop even in absence of a tax.

[87] The (quasi) coincidence with the peak of ROW fuel consumption is pure chance; varying the overhead costs of the liquefaction process, c_l, to values different from the here assumed 15 \$/bbl directly shifts the time of the liquefaction start-up whilst changing the ROW consumption peak only slightly.

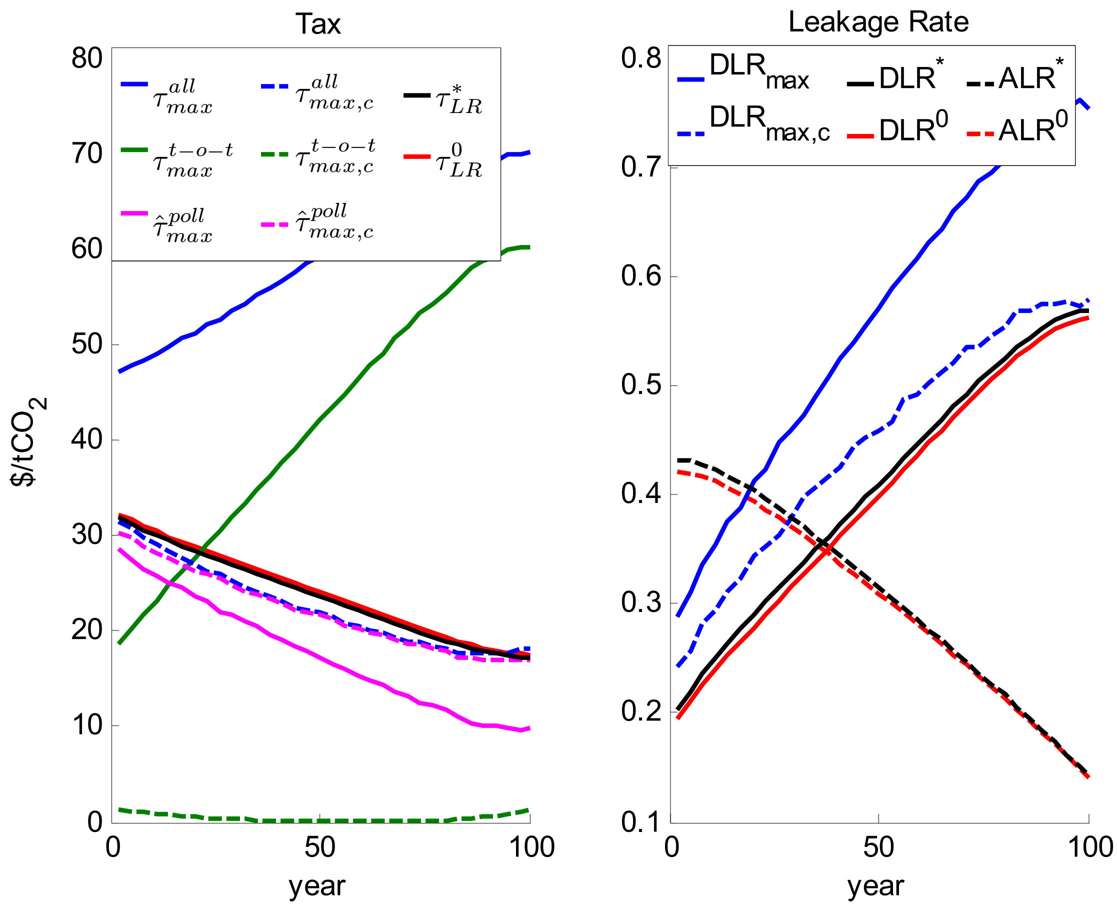

Figure 3.4: Illustrating the theory: taxes and leakage rates

3.8 Illustration of Tax Decomposition Method

Before we examine leakage rates in different scenarios in section 3.9, this section uses the numerical model to illustrate the main results of the theoretical findings from Part 1. For this purpose, consider Fig. 3.4, where the plot 1 (left) provides tax rate estimates, and plot 2 (right) shows leakage rates.

For simplicity we abstract from demand growth, from liquefaction or the backstop, and focus on a situation where damage is linear in cumulative emissions (implying, for the constant discount rate, that the calculated DLRs are equivalent to the NPV leakage rates, NLRs). We restrict the simulation horizon to 100 years assuming extractions to stop and cumulative emissions to remain constant thereafter. We consider discount scheme 1, i.e, $\rho_u = \rho = 3\,\%$. We assume an instantaneous marginal damage of $d = \rho \cdot 40\,\$/\text{tCO2}$, such that the current value of the future damage from a unit of emissions at time t is $\int_0^\infty e^{-\rho v} d dv = 40\,\$/\text{tCO2}$.

In the left plot, τ_{max}^{all} and τ_{max}^{t-o-t} (solid blue and green) indicate the OECD tax rates numerically calculated as those which maximize OECD utility with (index 'all') and without (index 't-o-t', for terms-of-trade-only) pollution. Solid magenta shows $\hat{\tau}^{poll}_{max} \equiv \tau^{all}_{max} - \tau^{t-o-t}_{max}$, the crude approximation of the pollution-only tax described in section 2. Dashed blue, green, and magenta, show $\tau^{all}_{max,c}$,$\tau^{t-o-t}_{max,c}$, and the corresponding difference, $\hat{\tau}^{poll}_{max,c}$, are the analogs for the case when the region hypothetically is forced to compensate the fringe consumers and the producers (cf. section 2).[88]

In the right plot, $\mathrm{DLR}_{\mathrm{max}}$ (blue solid) and $\mathrm{DLR}_{\mathrm{max,c}}$ (blue dashed) give the implicit leakage rates defined by the approximations of the pollution-only tax, $\hat{\tau}^{poll}_{max}$ and $\hat{\tau}^{poll}_{max,c}$, calculated as $\mathrm{DLR} = 1 - \hat{\tau}^{poll}/dam$ (cf. Eq. (3.11)), where dam is the net current value of all future damages from current emissions at time t, which due to the linear damage assumption is constant at $40\,\$/\mathrm{tCO}_2$.

The right plot also shows DLR^* (solid black) and DLR^0 (solid red), the directly calculated DLRs (here equivalent to the NLRs), based on the reaction of the foreign emission *path* to small variations of domestic emissions at t. DLR^* is calculated from $\tau^{all}_{max,c}$, the utility-maximizing tax path for pollution damage and the hypothetical compensation, that is, implicitly on the optimal 'green-only' tax,[89] and DLR^0 for the tax path $\tau_t = 0$. We note that these two leakage rate paths are almost identical despite the substantial difference in the base tax paths under whose fuel-market equilibrium the leakages are calculated. ALR^* (dashed black) and ALR^0 (dashed red) are the equivalents to DLR^* and DLR^0, in terms of absolute emissions.

Finally, τ^*_{LR} (black) and τ^0_{LR} (red) in the left plot give the optimal green tax rates implicitly implied by the damage leakage rates DLR^* and DLR^0, again according to Eq. (3.11).

The graphs confirm several key points from the theoretical part of this paper. The crude approximation of the pollution-only tax from the simple subtraction method, $\hat{\tau}^{poll}_{max}$, is substantially lower than the optimal pollution-only tax rates approximated by both, the compensation scheme calculation (with the corresponding numerical approximations $\hat{\tau}^{poll}_{max,c}$ and $\tau^{all}_{max,c}$) or the leakage-rate based calculation (τ^*_{LR}). This confirms the theoretical result that the subtraction method leads to an understatement of the optimal pollution tax (cf. section 3.4).[90] Further, that $\tau^{t-o-t}_{max,c}$ (no damage) is almost exactly zero confirms that the hypothetical compensation neutralizes terms-of-trade ef-

[88] It is calculated with sequential maximization for different ts, using two iterations (the simulations showed that the convergence is very rapid with differences between utility-maximizing tax values calculated starting from $\tau_0 = 0$ or by starting from the values after the first iteration being tiny already).

[89] Note that in theory $\tau^{t-o-t}_{max,c}$ is zero, and thus $\tau^{all}_{max,c}$ equal to $\hat{\tau}^{poll}_{max,c}$.

[90] Correspondingly, the implicit leakage rate calculated based on $\hat{\tau}^{poll}_{max}$, $\mathrm{DLR}_{\mathrm{max}}$, is too high.

fects as we have derived analytically.[91] Finally, the tax rates indicated by the solid cyan and red in the left – calculated based on the emission damage *dam* augmented by the factor (1-DLR), where DLRs were calculated directly using small perturbations of domestic consumption (solid black and red on the right) – correspond very closely to the directly calculated overall welfare maximizing domestic taxes (dashed green on the left). This confirms the theoretical finding that the optimal regional pollution-only tax (or the optimal regional overall tax under the compensation scheme) corresponds to $dam \cdot (1 - \text{DLR})$ (cf. Eq. (3.11)).

In the next section the form and values of the graphs are discussed along with the results from additional model setups.

3.9 OECD Leakage Rates

To ensure an understanding of the leakage rate paths in the more complex model, the following starts with a discussion of leakage rates in very basic setups, and sequentially adds complicating elements until, ultimately, models taking into account all the features described in section 3.6 are considered. Since the path of the foreign emission offset can best be understood when direct emission leakage, either in absolute or NPV terms, is considered, we start by considering solely ALR and NLR paths. The additional effect of the damage convexity by studying DLR rates is only taken into consideration at a later stage.

The various graphs of the model results in the following sections contain many details, each of which can stem from obvious or not so obvious effects in the modeled resource market. The aim here cannot be to fully describe all these details; instead, we try to focus on the most interesting aspects.

3.9.1 Single Fuels

We start by considering simplified variants of the model where only a single fuel is consumed, with constant demands and without liquefaction and backstop. In addition, we consider a variety of hypothetical extraction curves. In all other respects, the model corresponds to the original setup described above.

Fig. 3.5 shows NLR (solid) and ALR (dashed) leakage rate paths for five different single-fuel model setups, for discount schemes 1 and 2, the first four plots for oil, and plot 5

[91] We used a time-step duration of 1 year for the simulations.

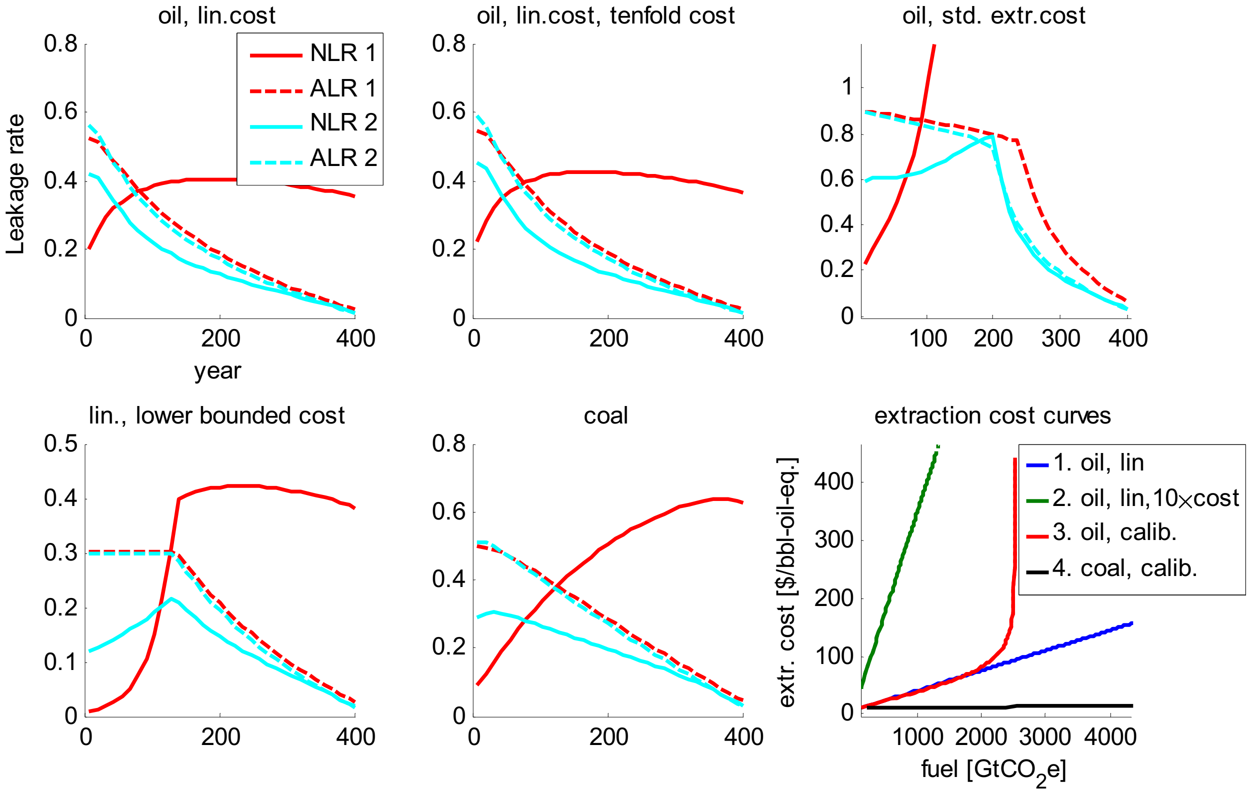

Figure 3.5: Leakage paths single fuels

for coal. Plot 1 is from a run where the highly convex oil cost curve is replaced by a linear cost curve constructed as the tangent to the initial slope of the original oil cost curve (plot 6 shows the cost curves: blue for the linear tangent curve and red for the original oil cost curve). Plot 2 is also based on a linear oil cost curve, but with oil made much scarcer, by multiplying by a factor 10 the costs from the first linear curve (green line in plot 6). Plot 3 considers the original cost curve, and plot 4 the first linear cost curve, but with all costs increased to minimally 70 $/bbl, resulting in a cost curve with a constant and a linearly increasing region. Finally, plot 5 is for the model with coal only, for standard coal extraction costs (black line in plot 6).

As already in Fig. 3.4, ALR paths are generally downward sloping. That is, an early domestic emission reduction leads to a larger overall foreign emission increase than a late domestic emission reduction of the same magnitude. The offset rates, ALRs, do not converge exactly to zero, but they become small as time approaches the end of the model horizon. This cannot simply be explained by the fuel owners' having, for early domestic reduction, more time to offset it abroad 'after' the time of that early reduction, because the model assumes perfect foresight (and commitment), that is, the fuel owners could per se react to *late* domestic emission reductions by increasing sales *prior* to the late reductions, symmetrically to their *future* reaction to *early* domestic emission reductions. Instead, the effect can be explained by the earlier fuel units being cheaper to extract. Intuitively, providing the fuel owners with an 'additional' unit of a *cheap* fuel increases fuel sales within a specific time-horizon more than if they were given a more expensive additional unit of fuel; if we hypothetically consider a unit of very cheap fossil fuel to be added to the reserves, this tends to increase overall fuel sales (within a limited sales period) by more than the addition of a unit whose extraction costs are so large that it is barely profitable to extract the unit. That this, rather than the timing of the domestic perturbation per se, explains the downward trend of the leakage rates is confirmed in the fourth oil cost curve scenario considered here (plot 4 in Fig. 3.5): assuming a cost curve that is constant for a specific amount of fuel and rises only after this amount is extracted the leakage rate remains constant during the phase of constant extraction costs, and it starts to decrease only once the fuel extraction costs have started to increase. Similarly, in oil cost scenario 3, when taking into account the calibrated, quasi-kinked oil cost curve that yields a substantial amount of very cheaply extractable oil but very rapidly rising costs after a certain amount when only difficult to extract units are left (Fig. 3.1), we find that during the whole initial phase with the modest costs the absolute leakage rate remains rather stable, and it starts to drop rapidly[92] after around 200 years when the phase with the rapidly rising extraction cost is reached (plot 3 in Fig. 3.5).[93] Finally, besides in this specific pattern in the case of the quasi-kinked oil cost curve, that the decrease of the ALRs is not simply explained by the approaching of the artificial ending of the simulation horizon, is also confirmed by the case where the fuel-sales period ends endogenously with the relative competitiveness of a backstop technology (cf., e.g., Fig. 3.11, plot 2).

A few additional patterns deserve our attention: First, the influence of the fuel-owner discount rate on the ALRs seems very limited.[94] In most cases, the ALR paths for the two discount schemes (dashed lines) are hardly distinguishable in Fig. 3.5. The only notable difference occurs for the case of the quasi-kinked oil cost curve (plot 3) where the difference seems to essentially stem from the fact that the more impatient fuel owners (cyan, $\rho = 5\,\%$) sell fuel more rapidly and thus that the quasi-kink in the extraction rate – and the corresponding steeper downward trend of the ALR path – is

[92]The very distinct kinks in the leakage rate paths are in reality slightly smoother if we use lower numerical time-steps for the calculation of the leakage rates; here we used leakage rate time-steps of 12 years duration.

[93]We omit the time-path of the extraction cost curve here. See, e.g., Figs. A3.2 and 3.7 where, for the standard model with oil and coal a similar oil leakage rate pattern results, with the kink in the leakage rate path in Fig. 3.7 corresponding to the time of the oil cost quasi-kink in Fig. A3.2 also a few decades after 200 years.

[94]The consumers here are acting non-strategically and so the consumer discount rate is irrelevant for absolute leakage.

reached earlier.

Second – and maybe most surprisingly –, a scaling down of the oil reserve sizes by a factor 10 hardly matters for the leakage rates, assuming linear oil extraction costs. All ALR and NLR paths are almost exactly identical between the corresponding plots 1 and 2, where scenario 2 was based on a 10 times increased cost curve. This reserve-scale invariance was not necessarily expected since a more pronounced scarcity is generally thought of as tending to increase fuel-price channel leakage rates: if a fuel is abundant enough, regional savings tend to have a small impact on the worldwide price, thus implying limited leakage rates. In contrast, for scarce enough resources one may, put bluntly, assume that what is not bought by one party will simply be bought by the other; in other words, (at least absolute) leakage rates to approach 100 % even for limited time horizons. In the present case, that a scaling of the reserve size only has a marginal effect, considering solely oil and linear extraction costs, can be explained by the scale invariance of the demand function. To see this, first note that in plot 6 in Fig. 3.5 the linear oil cost curve (blue line) passes almost through the origin, evidencing that the fuel extraction costs are almost proportional to cumulative extractions. Intuitively, having an isoelastic demand, an γ-fold increase of the cost of the extraction of a specific amount of fuel, may in this case simply imply that the fuel extraction is reduced by a constant factor and the fuel price (and extraction costs) augmented by another constant factor, with nothing else changing in the (time) pattern of the problem's solution. In this case, because the extraction cost increase has only scaled all variables proportionally by specific, constant factors, but left the problem and its solution unchanged elsewise, the leakage reactions to small domestic perturbations also should exhibit the same pattern in both cases. This intuition about a pure scale-only effect of the cost-augmenting factor on the fuel market outcome is analytically confirmed in Annex 6.

Third, and related to the second point, the absolute leakage rate is almost one for the initial periods in the case of the empirically calibrated oil cost curve (plot 3), and it does not vary with the fuel-owner discount rate. In all three other cases, the absolute leakage rate starts 'only' at just below 0.6 in the initial periods and becomes lower later on. This stark difference confirms the importance of Green Paradox type effects: with the pronounced quasi-kink in the oil extraction costs for the calibrated oil cost curve, the situation corresponds almost to one with a fixed amount of fuel extractable (the amount corresponding roughly to the cumulative production possible before the extraction cost become rapidly very high just after the curve's kink) relatively cheaply, and only a very limited additional amount extractable for even quite high costs. In this case, sparing some of the cheap fuels in one place will mostly relocate this consumption to the other, passively consuming, region rather than yielding long-run global savings.

Thus it will yield absolute leakage rates very close to unity. In this sense, and in relation to the previous observation about the resourse-scale-invariance of leakage rates, we note that in the basic dynamic resource exhaustion model the leakage rate pattern seems much more directly influenced by the curvature of the extraction cost curve than by the absolute amount of fossil fuels available (where the last point holds at least for linear costs with small intercepts[95]). The various scenarios examined here therefore also clearly indicate that whilst oil is – in the case of the calibrated cost curve (plot 3) – subject to a much higher leakage rate than the more abundant coal (plot 5), this is, at least in the simple model considered in this section, less due to the general relative abundance of the fuels. Instead it appears to be mainly due to the special form of the oil extraction cost curve, implying that with the quasi-kink a quasi-fixed amount of oil will be available.

Fourth, the time-path for NPV leakage rates for anticipated taxes is highly non-monotonous and has a tendency to stretch over large ranges from very low up to values exceeding unity, especially in the case of the higher emission (consumer) discount rate (NLR_1, $\rho_u = 3\,\%$). The initial NPV leakage rates are below their absolute counterparts. This is the necessary consequence of the foreign emission reaction being spread across time and thus to be discounted, i.e., they are weighted relatively less than the initial domestic emission changes. This difference between the NPV and the absolute leakage paths tends to become smaller over time, as the domestic emission reductions themselves become more discounted, and as parts of the foreign reaction take place in times prior to the domestic reductions. For the NLR_1, this leads to discount rates that even rise over time as the domestic emission reduction is discounted ever more while parts of the foreign reaction continue to take place in the earlier periods. For the linear oil extraction costs (plots 1 and 2), and for low discount rates (NLR_2), this effect is reversed in later periods where the absolute amount of leakage becomes so small (dashed lines) that the additional timing effect is outweighed by the decrease in the fraction of the emissions offset overall. The third oil case, with the calibrated oil cost curve, is interesting. For early emission reductions, a part of the offsetting abroad occurs later on. This implies that, whilst the absolute offsetting fraction is close to 1, especially NLR_1 is low as the initial domestic reduction achieves at least a partial delay of global emissions. The inverse holds for periods later on but still before the extraction cost's quasi-kink: relatively late domestic emission reductions are weighted against foreign offsetting reactions earlier on. In the case of a strong enough discounting, this implies

[95] In later scenarios with interdependent demand for several fuels, we will generally find that with convex extraction costs a higher demand (and thus, higher relative scarcity) tends to increase leakage rates (cf., e.g., discussion of coal leakage rates in the basic case with interdependent fuels, section 3.9.2 and in that with liquefaction, section 3.9.3).

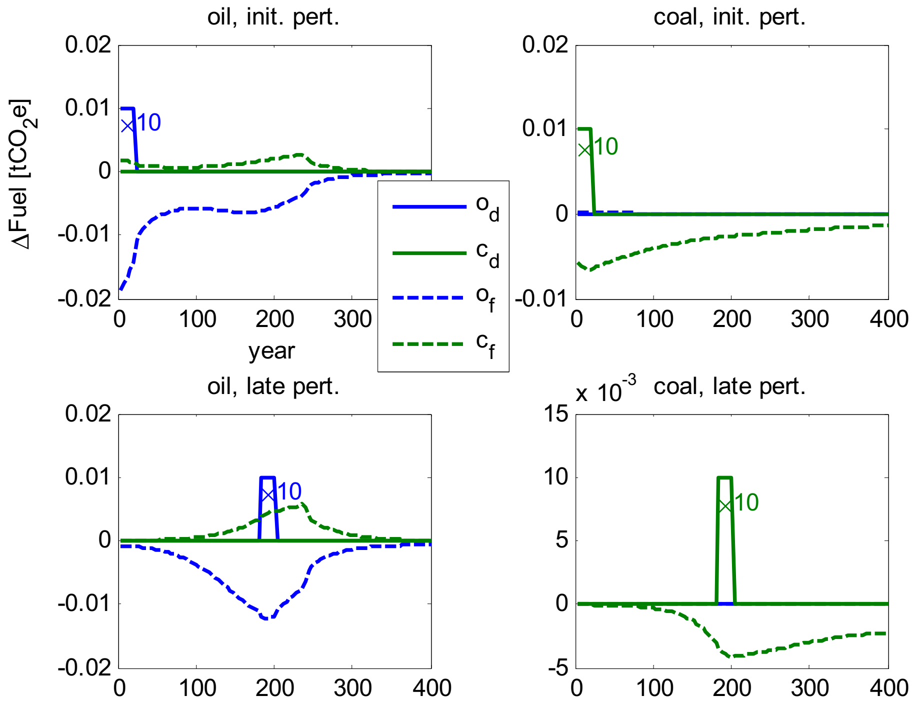

Figure 3.6: Emission reaction paths, basic setup

Foreign oil and coal consumption reactions (o_f, c_f) to domestic oil or coal (o_d or c_d) consumption changes, for weak substitutability ($\sigma = 1.1$) and discounting scheme 1 ($\rho = 3\,\%$). For visibility, domestic consumption changes are scaled down by a factor 10.

that the NPV leakage rate exceeds unity: we have a Green Paradox effect where the timing shift outweighs the (relatively small) overall emission reduction effect. Whilst plot 3 does not show the whole NLR_1 path, the path reaches very high levels above a factor of $\mathrm{NLR}_{1,t} = 10$. Similar effects drive the special shape of the NLR paths in plot 4. There, however, in a less accentuated form, as the slope of the cost curve after the kink is much less steep than in the calibrated oil cost curve, wherewith, the 'quasi-fixed reservoir' effect from the calibrated curve applies less strongly to case 4.

3.9.2 Basic Joint-Fuel Setup

We consider the model with joint fuel demand (substitution elasticity $\sigma = 1.1$), without growth, in the absence of liquefaction and backstop. Fig. A3.1 in Annex 7 provides the detailed model results. As the understanding of the leakage rate paths requires some insight in the foreign (cross-)fuel reactions to domestic consumption changes, we first illustrate and interpret the paths of foreign emission reactions to initial and later domestic fuel consumption reactions, plotted in Fig. 3.6.

Analysis The bars (exogenous increase of domestic consumption) always imply an opposite reaction of the dashed same-colored curves. They are spread in general over the whole time horizon (negative response of foreign consumption of the same fuel, as a direct price effect) with the reaction typically being strongest during the period of the exogenous perturbation, and in most cases they monotonously decrease in both time-directions. Foreign consumption of the other fuel – dashed, other-colored curve – reacts in the direction of the exogenous domestic perturbation. This is a consequence of the (weak) substitutability of the two fuels ($\sigma > 1$): if one fuel becomes more expensive on the marked and less consumed (due to the increase in foreign consumption of it), the other fuel is consumed more. This cross-fuel reaction is accentuated for larger substitutability σ, and reverted when the fuels are complements ($\sigma < 1$), as shown in Annex 7.

We note some further qualitative features of the reaction paths. First, the relative reaction of coal-to-coal is somewhat smaller than the oil-to-oil reaction.[96]

Second, whilst there is substantial coal-to-oil reaction (green dashed on left), the absolute oil-to-coal (blue dashed on right) appears to be small (this pattern persists, less accentuated, for the cases with σ further away from unity, cf. Figs. A3.3 and A3.4 in Annex 7). This may be a surprise given that the order of magnitude of absolute (foreign) emissions is the same for the two fuels, and we use a demand system based on a CES function with a single substitutability parameter σ for both fuels. An analysis of the properties of the CES demand system offers an answer. As a mathematical property, as the price of a fuel x becomes large, the relative reaction of the other fuel y to changes in the price p_x, $\partial y/\partial p_x$, becomes large relative to the direct price reaction $\partial x/\partial p_x$. Given the scarcity of oil and its much higher price (per unit of energy as well as per unit of emissions), the secondary reaction of coal to oil can thus be relatively large, whilst the more abundant coal can be replaced more easily by only a smaller amount of oil. This also offers an explanation for why the relative reaction of coal to oil becomes even smaller for the case of the late perturbation (plot 4) when the rapid depletion of oil has even further increased the relative scarcity of oil compared to coal.[97] Moreover, it explains why the coal-to-oil reaction is relatively larger for the later oil-perturbation (plot 3) than for the initial oil-perturbation (plot 1), and why the absolute magnitude of the coal-to-oil reaction peaks later than the oil-to-oil reaction itself (both, in plot 1

[96]Whilst the spread of the reaction over time is overall more or less comparable, with the same scaling applied in all graphs for coal the maximal amplitude of the reaction is always smaller than a tenth of the original perturbation whilst for oil the corresponding maximum is always larger than a tenth.

[97]This pattern persists in the results for $\sigma = 1.7$ and $\sigma = 0.3$, Figs. A3.3 and A3.4, Annex 7.

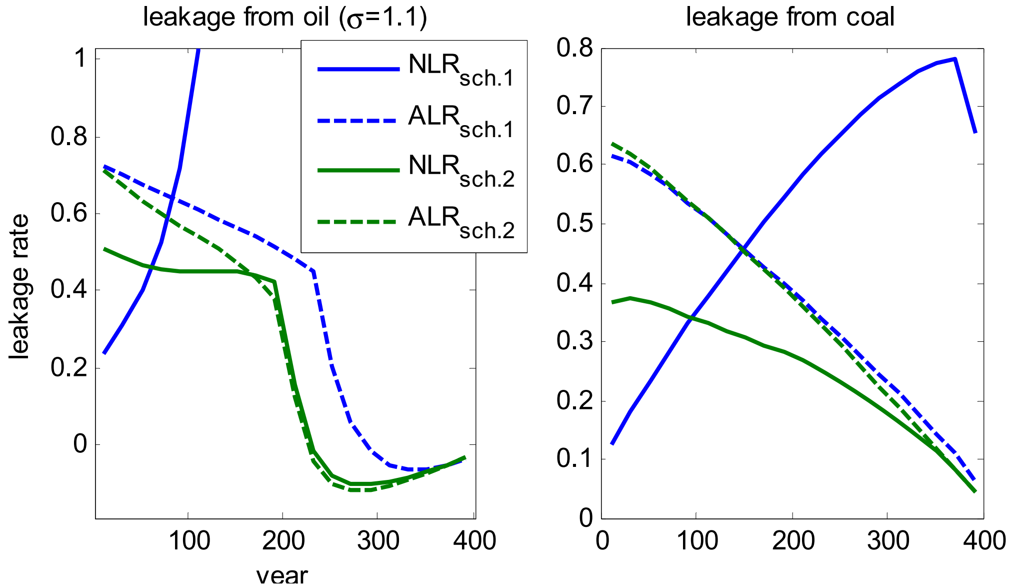

Figure 3.7: Leakage paths basic setup

and plot 3).[98]

Leakage rate paths Fig. 3.7 shows NLR (solid lines) and ALR paths (dashed lines) implied by the emission reactions illustrated in Fig. 3.6. Blue lines are for discounting scheme 1 ($\rho = \rho_u = 3\,\%$), green for scheme 2 ($\rho = 5\,\%$, $\rho_u = 0.5\,\%$).

Analysis Introducing fuel-interdependence (cf. Fig. 3.7) qualitatively changes things relatively little compared to the cases with single fuels (plots 3 and 5 in Fig. 3.5). For oil, LRs are slightly reduced (early and later, except for the NLR for discount scheme 1, which is hardly affected) as, given the (weak) substitutability, foreign coal emissions are somewhat reduced when domestic oil reduction increases foreign oil availability and consumption. We even find negative oil leakage for very late periods when the strong relative scarcity of oil implies that increased oil availability reduces coal consumption strongly, such that the oil-to-coal substitution dominates the original foreign oil emission increase (see the analysis to Fig. 3.6 for a detailed description of this mechanism). For coal, in contrast, LRs are increased, early and later. This seems to be readily explained as follows: with the fuels as substitutes, and oil becoming very scarce over time, the demand for coal increases relatively strongly over time compared to the single-fuel case where coal is modeled separately. Therefore, compared to coal modeled as single

[98]The magnitude of the oil-to-coal reaction is further reduced relative to that of the coal-to-oil reaction by the relative emission intensities of the fuels: coal is more emission intensive than oil, and in terms of the relative magnitude of the coal-to-oil vs. oil-to-coal effect ratios, and the relative size of the two cross-fuel emission changes is proportional even to the squared factor of the relative emission intensities.

fuel, future coal becomes relatively scarcer when the two fuels are modeled together. The higher relative scarcity of coal, increasing notably over time as oil becomes more and more rare and demand for coal increases progressively, in turn tends to increase the leakage rate for domestic coal reductions (note, coal extraction costs are exponential, not linear). As Fig. 3.6 shows, secondary coal-to-oil effects are small in terms of emissions and do not revert this result.

3.9.3 Liquefaction

We allow the endogenous emergence of coal liquefaction. As plot 1 in Fig. 3.8 shows, this provision of synthetic oil supplementing crude oil supply ramps up after around 60 years, to provide a substantial part of the overall oil consumption (sum of blue lines). The corresponding increase of coal consumption increases the coal cost and price (plot 2) compared to the case without liquefaction (Fig. A3.2 in Annex 7), whilst flattening and limiting the cost of (crude) oil. Fig. A3.5 in Annex 8 gives a zoomed view on the curves of plot 2 during the transition phase to the liquefaction process. It reveals that extraction cost and profit curves are kinked. The coal (oil) cost curve rises more (less) rapidly after the CTL onset, and the oil shadow value decreases rapidly as CTL is imminent, as was to be expected.

Fig. 3.9 shows the leakage rates, analogous to Fig. 3.6, for the case when the endogenous liquefaction production is allowed for.

The introduction of the liquefaction process has further increased the LRs for coal, compared to the case without liquefaction (cf. plots 2 in Figs. 3.9 vs. 3.7). Again, this is readily explained by coal (with convex extraction costs) becoming somewhat scarcer when its use, for the purpose of liquefaction, is increased – especially during the future periods after the initiation of the CTL process. For oil, LRs become much lower, and relatively soon even become negative, converging to below -1 in the longer run. This is explained as follows: with liquefaction and scarce crude oil and abundant coal domestic oil savings primarily lead to lower (or later) CTL globally; as synthetic oil is *more than twice* as emission intensive as genuine oil this implies strongly negative leakage rates, of up to *below -1*,[99] when fuel end-consumption leakage by itself is small. The latter is the case (i) especially during late periods (cf. section 3.9.1), but (ii) to a limited extent already earlier, given the abundance of coal and the corresponding abundance of the dirty, liquid backstop synthetic oil.

[99]Recall that the synthetic and genuine oil have similar emissions during their final consumption, but the production (abroad) of the synthetic fuel adds more than these emissions (section 3.6)

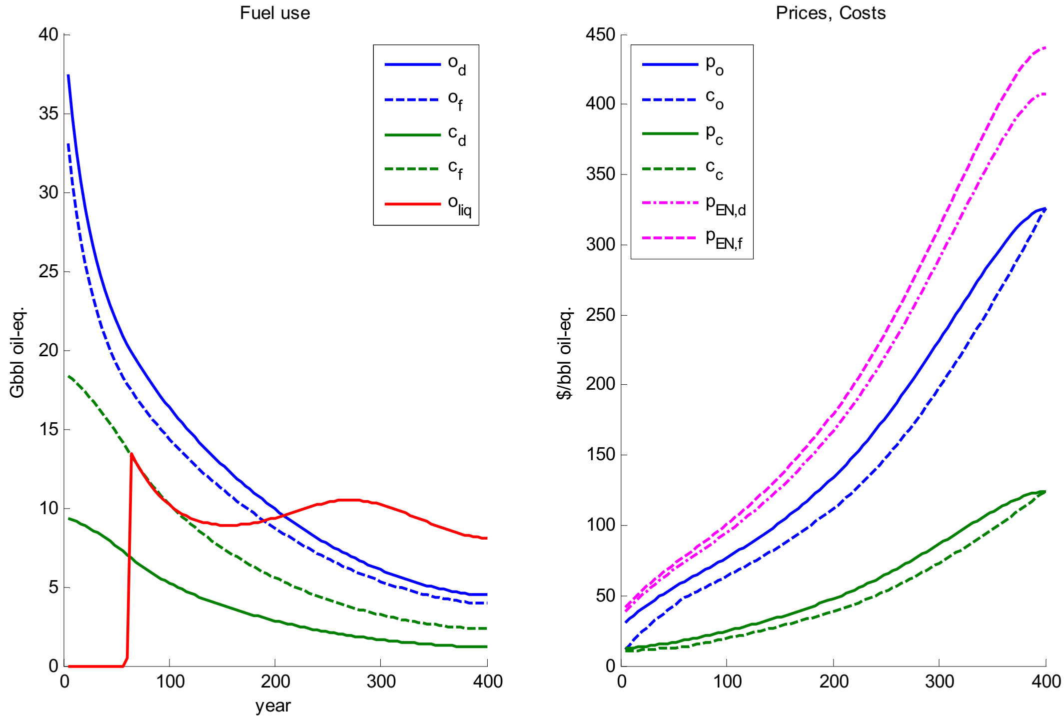

Figure 3.8: Simulation details with liquefaction

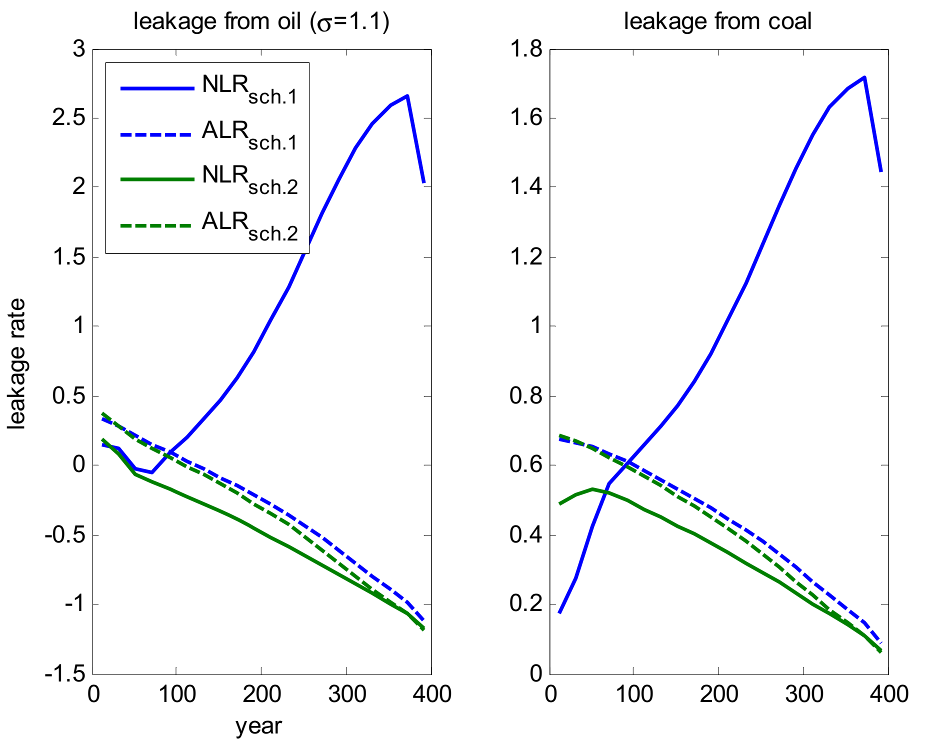

Figure 3.9: Leakage paths with liquefaction

However, both, absolute and NPV LRs are at least slightly positive for initial domestic oil reductions. This is explained by the fact that over the long time horizon considered even though coal is abundant it is still a depleting resource[100] with corresponding increases in foreign coal and synthetic oil consumption following a domestic oil reduction, offsetting the CTL-related negative leakage in part or even fully. It is the general decrease of this offsetting over time (again, cf. explanations section 3.9.1) which explains the downward slope of the leakage rates in general (the exception of the mainly increasing NLR for the strong time-discounting being, as usual, explained by later domestic emission changes implying *earlier* and thus relatively more weighted foreign reactions), as fuel-end consumption leakage decreases over time and the 'negative leakage' CTL-effect becomes more and more dominant. A negative NLR for initial oil savings results if overhead costs are small enough (shown in Fig. A3.8, Annex 8, for $c_l = 0$). Emission reaction paths, with details on the response of foreign fuel consumption to domestic changes and with liquefaction, are given in Annex 8 (Fig. A3.6 and, for cumulative changes, Fig. A3.7).

3.9.4 Clean Backstop

We allow the endogenous emergence of a clean backstop that replaces the fossils as soon as it is competitive in the different regions. Here, we consider the case in absence of liquefaction; section 3.9.5 allows also for liquefaction. As plot 1 in Fig. 3.10 shows, this alternative energy technology replaces the fossil fuel aggregate towards year 200, first in the OECD and soon thereafter in the ROW. Naturally, fuel profits (plot 2, difference between market prices, p, and extraction costs, c) converge to zero up to the time when the backstop fully replaces the fossil fuels. The backstop alters the problem in the sense that the stopping time of the fuel sales and emissions is no longer exogenously given by the end of the simulation horizon, but it is endogenous and dependent on domestic fuel emission changes.

Fig. 3.11 shows the leakage rates, analogous to Fig. 3.6, for the case when the backstop is considered.

In presence of the clean backstop for the fossil fuel aggregate, domestic *savings* of the scarcer and more rapidly depleting resource – oil – strongly *increase* the absolute amount of total emissions (plot 1, ALRs >1).[101] This in stark contrast to the case

[100] And its depletion is here non-negligible especially due to the large amount of coal used for liquefaction.

[101] Fig. A3.11 in Annex 9 confirms that this holds even for the case of where the resources are stronger substitutes or strong complements.

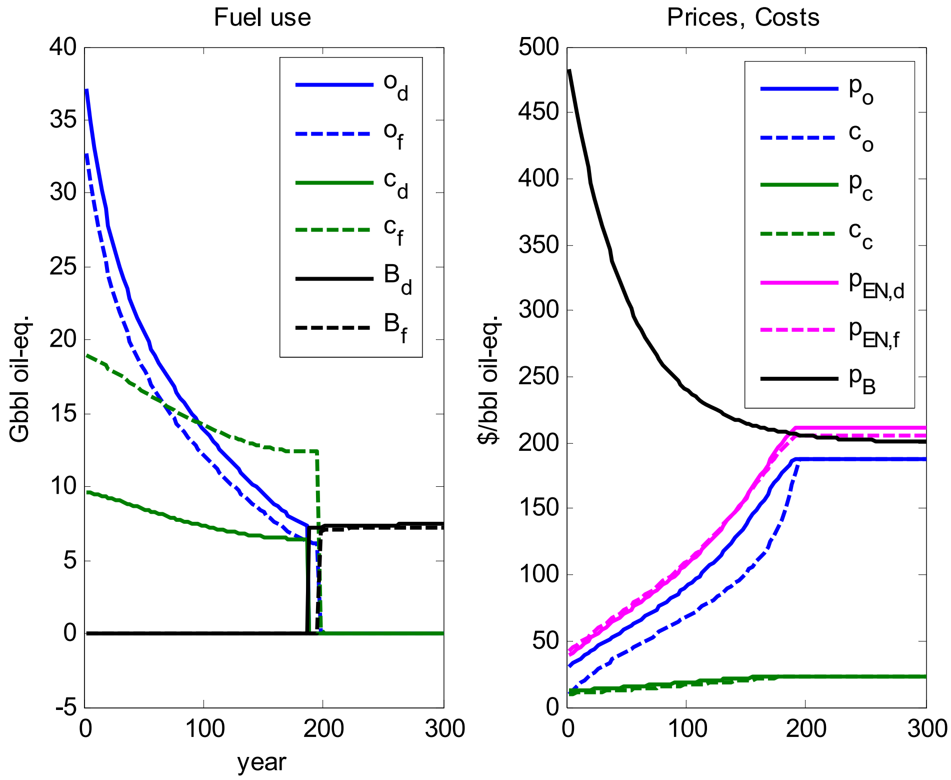

Figure 3.10: Simulation details with backstop

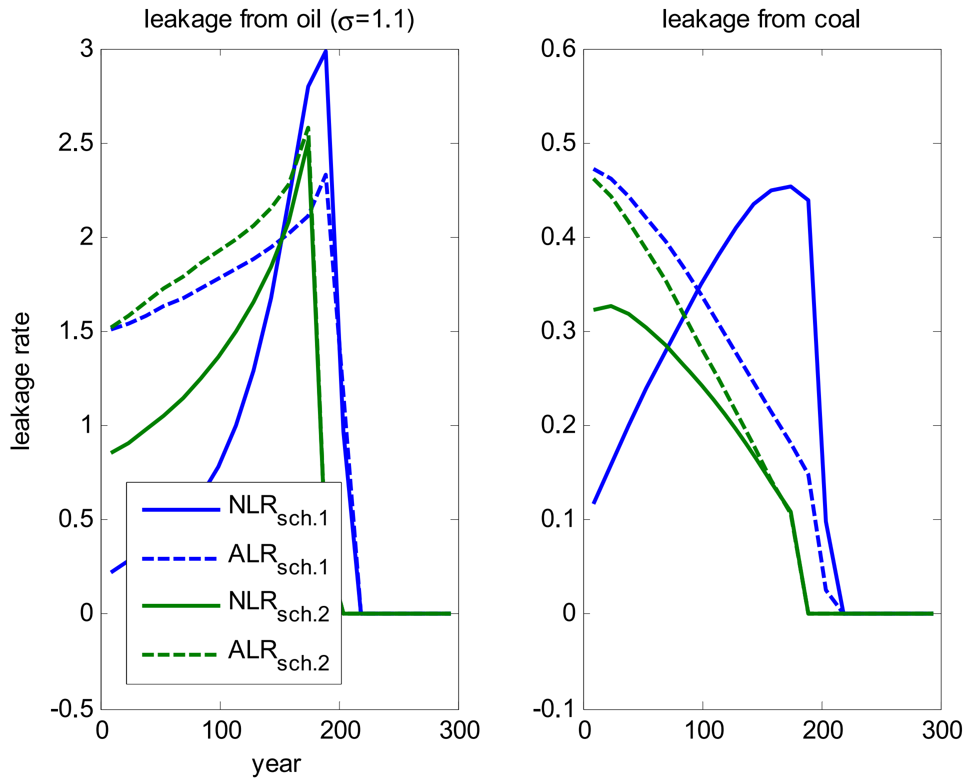

Figure 3.11: Leakage paths with backstop

without clean backstop, where absolute leakage rates are always below unity when the fuels are substitutes ($\sigma > 1$). Moreover, it also goes against the general predictions of the implications of a clean backstop for the effect of climate mitigation policies, where the backstop is typically found to limit leakage or Green Paradox effects, as the backstop facilitates the reduction of total fuel consumption (e.g., van der Ploeg, forthcoming; Essay 1 in this dissertation; Hoel, 2010[102]). The evolution of *cumulative* foreign emission changes in response to a domestic oil perturbation (Fig. A3.10, plot 1, Annex 9) helps to explain the above-unity leakage rate found here. The oil is scarce enough for a large fraction (almost 1 Gbbl) of the early domestic savings (of 1.5 Gbbl) to be offset abroad before the phasing in of the endogenous backstop, and the induced delay of the emergence of the backstop increases this fraction to almost exactly 100 %, the foreign oil offset reaching almost 1.5 Gbbl. What pushes the overall leakage rate to above 1 is the additional coal used during the additional time during which the fossil fuel aggregate is consumed. That is, by stretching the time during which the fossil-fuel aggregate is competitive against the clean backstop, an increased supply of the *scarcer* fuel tends to imply an overall emissions increase that exceeds the emissions of the additional amount of the scarce resource itself. The extra coal burn during the additional phase of fossil fuel use is large. This is again related to the relative scarcity of oil during these late periods: as oil is scarce (expensive) during that final period, the substitutability of the fuels implies that a unit of aggregate fossil energy to a larger degree consists of coal than, e.g., in the initial periods when the price of oil is not yet so much higher than that of coal. With the assumed weak substitutability, $\sigma = 1.1$, this effect is limited here and would conceivably be even stronger for a higher σ.

A very different picture results for coal leakage rates: domestic savings of the more abundant resource – coal – strongly reduce global emissions (Fig. 3.11, plot 2), with LRs < 0.5 for all variants and times. This corresponds to standard expectations for a relatively abundant[103] fuel in the presence of a clean backstop: given that the fuel is not too rapidly depleting and price effects of regional savings imply only relatively moderate (that is, 'slow') reactions of foreign emissions as well as only a limited delay of the phasing in of the backstop, a relevant fraction of the domestic savings is carried on until the phase-out of the fossils by the clean backstop. In fact, a comparison with the effect of an additional supply of oil confirms that it is mostly due to the depletion of oil that the backstop becomes competitive (Figs. A3.10 and A3.9, Annex 9). The

[102]The finite choke price in Hoel (2010) can readily be thought of as a backstop price.

[103]At first sight it may appear somewhat contradictory to the findings in section 3.9.1, where the importance of the absolute abundance of the fuel was qualified. Recall that this was the case for a setting with linear extraction costs and, importantly, single fuels. As we observe and explain here, when different fuels are consumed as substitutes (and especially in presence of a backstop), a higher scarcity of a fuel seems to imply a larger leakage rate for domestic savings of that fuel.

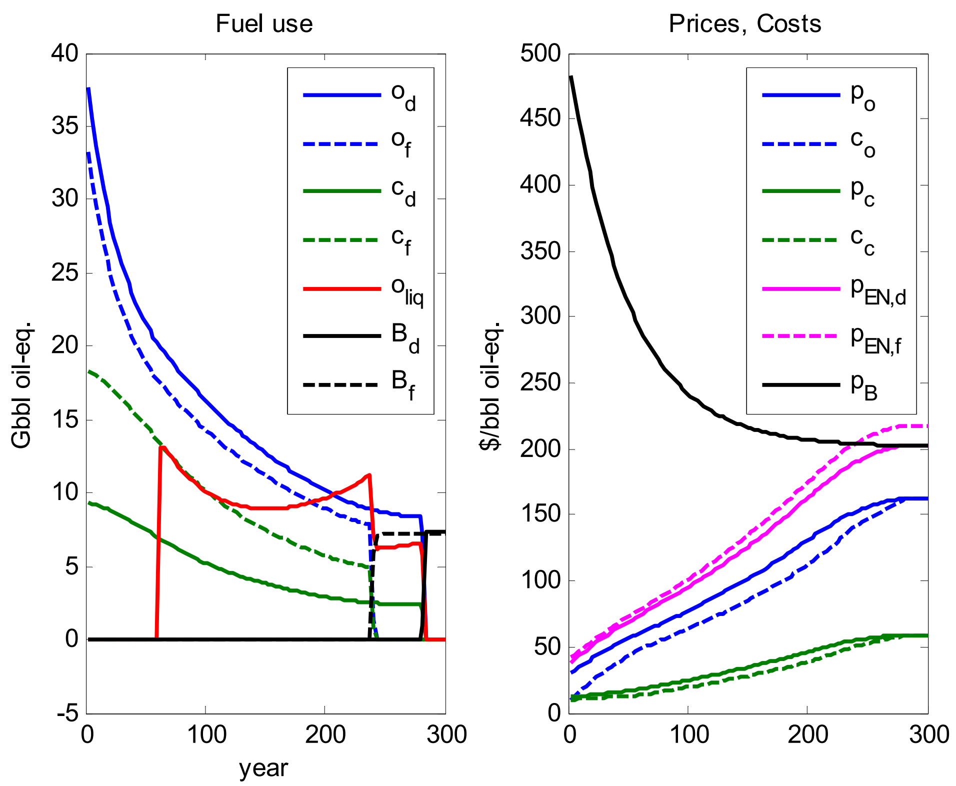

Figure 3.12: Simulation details with backstop and CTL

additional supply of coal has a much more negligible effect on the time the backstop replaces the fossil fuels and on the cumulative emission changes.

3.9.5 Clean Backstop with Liquefaction

Liquefaction is added to the model variant with the backstop from the previous section. Fig. 3.12 shows the detailed simulation results. Supplementing the supply of liquid fossil fuel, liquefaction stretches substantially the period during which the fossils outcompete the backstop, from just below 200 years for the case without liquefaction (Fig. 3.10) to almost 300 years here.[104] Given what we noted in the last section, this is no surprise: without CTL, it was essentially the rapidly depleting oil which allowed the backstop to replace the fossils, but here synthetic oil allows a sustained production of liquid fuel.

Fig. 3.13 shows the leakage rates, analogous to Fig. A3.9, for the case with liquefaction in addition to the backstop.

Besides prolonging the fossil fuel use, and therewith the path of non-zero leakage rates, liquefaction does not alter the main conclusions about the fuel-specific leakage rates

[104]Due to the mechanism described in section 3.7, liquefaction implies that the backstop replaces fossils first in the ROW and only later in the OECD.

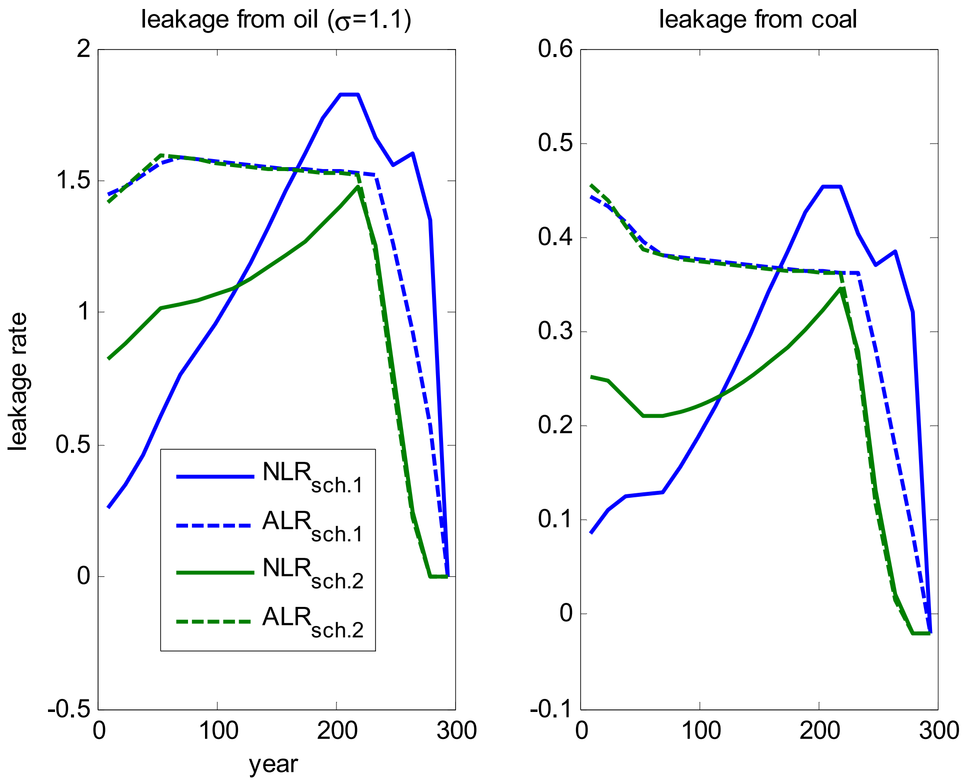

Figure 3.13: Leakage paths with backstop and CTL

in presence of the clean backstop. For oil, absolute leakage rates are still above unity, around 1.5 or even higher. CTL, as a relatively inefficient conversion and having overhead costs, leaves oil still as the relatively scarcer resource, implying that more oil significantly prolongates the fossil-fuel era. In addition, the initial 'negative' leakage (related to foreign emission savings if less synthetic oil is produced) is offset during the end phase of the fossil fuel era where CTL lasts slightly longer due to the initially saved coal from the initial CTL delay. Thus, overall CTL has a modest effect on initial ALRs. Also the NPV leakage is high, especially for the low emission discount rate, and NRL_1 is almost unity. Qualitatively also the coal leakage rates remain comparable to those from the case of the backstop without liquefaction with rates that are always below 50 %.

3.9.6 Extension

Here we consider the case where fuel demand is growing and where climate damages are convex in cumulative emissions.

Demand growth

Based on projections of the IEA World Energy Outlook 2009 for their reference scenario (IEA, 2009), we consider a scenario where fuel demand is constant in the OECD and growing at 2.6 % p.a. in the ROW during the first 25 years. In the long run we assume that after the first 25 years, as the ROW economies are maturing, their energy demand *growth rates* slowly decline by 0.05 % p.a. This ends when the economies reach a state where autonomous energy efficiency improvements set off any final demand increases; from then on the energy demand growth rate is zero.[105]

Quadratic damages

Whilst it also seems natural to address the question of carbon emission leakage in terms of the absolute fraction of emission offsetting in the rest of the world, it is clear that for the welfare impact of the leakage, if future utility flows are generally present-discounted, emissions should be weighted according to the time they occur, as we have done above for the NPV leakage rates. In addition, it is, however, also the case that marginal emissions at different points in time and in different scenarios can have very different utility impacts on the populations concerned. Indeed, it is generally assumed that the marginal damages from emissions can rise rapidly as cumulative emissions increase, that is, climate damage is considered strongly convex in cumulative emissions. The formulas we derived for the optimal 'green-only' policies, strictly speaking, support the definition of the optimal tax based on leakage estimates only if either we assume damages to be indeed linear in cumulative emissions, or if we use a new definition of the leakage rate that expresses a sort of a rate of leakage of the *damages* from (leaked) emissions rather than emissions directly. Accordingly, we calculate here what we defined in section 3.5 the 'damage leakage rate', DLR, defined as the fraction by which the

[105] In the World Energy Outlook 2009 reference scenario lasting through 2030, oil consumption is assumed to decline by 0.3 %p.a. between 2008 and 2030 in the OECD while it increases by 2.3 % p.a. in the ROW (IEA, 2009, p. 81, Table 1.3). Correcting these *consumption* changes for the average annual oil consumption changes during the same period in a fuel market simulation with *constant demand* (these changes are -0.3 % p.a. in the OECD and -0.2 % p.a. in the ROW) to approximate *demand* changes, we find a constant oil *demand* in the OECD, and an increase in the ROW of 2.5 % p.a.

In the same World Energy Outlook scenario, coal consumption declines by 0.2 % p.a. between 2007 and 2030 in the OECD and increases by 2.8 % p.a. in the ROW (IEA, 2009, p. 90, Table 1.5). Correcting these *consumption* changes for the average annual coal consumption changes during the same period in our standard model with *constant demand* (these changes are +0.2 % p.a. in the OECD and -0 % p.a. in the ROW) to approximate *demand* changes, we find coal *demand* in the OECD changing by -0.4 % p.a., and increasing by approximately 2.8 % in the ROW.

Thus, in each region, average demand growth rates are very close to each other across the fuels during the period from 2007/2008 through 2030 and we approximate them by assuming a constant demand for both fuels in the OECD and an annual growth of 2.6 % for both fuels in the ROW. Note that for the coal demand in the OECD, the difference between our assumption (0 %) and what the World Energy Outlook data implies (0.4 %) is smaller in the medium-run than what the cited numbers suggest on first sight: OECD consumption in the World Energy Outlook is assumed to slightly decrease only until 2015, and from then on the projected consumption change is already approximately zero until 2030.

foreign emission offsetting reduces the overall impact of the domestic emission change on present-discounted future climate damage. Since climate damage is often approximated as quadratic in emissions, we assume an instantaneous damage function which is proportional to the square of cumulative emissions.[106] Accounting for the approximately half a trillion tons of (anthropogenic) carbon (TtC), or 1835 $GtCO_2$ that have been emitted until today (Allen et al., 2009), the damages $D(E)$ after the cumulative emission E from today on are thus proportional[107] to $D(E) \propto (0.5\,\mathrm{TtC} + E_{TtC})^2$ and thus[108] $D'(E) \propto 2\,(1835\,\mathrm{GtCO2} + E_{GtCO2})$. With the numerical simulation ending at time T, and cumulative emissions taken into account up to that point, cumulative emissions, and therewith marginal damage, during the time beyond T is implicitly assumed constant, wherewith, for a discount rate ρ_u the cumulative emissions up to time T, $E(T)$ create, for the time after T, a NPV damage of $E(T) \cdot \int_0^\infty e^{-\rho_u(s+T)}\mathrm{d}s = E(T)\frac{1}{\rho_u}e^{-\rho_u T}$.

Results

In the following we present the usual emission leakage graphs (ALR and NLR paths) as well as the damage leakage rate (DLR) paths, taking demand growth into account. Whilst we consider the endogenous emergence of the backstop, we rule out liquefaction, thereby implicitly assuming that the international pressure prevents the non-climate coalition countries (ROW) from expanding the dirty CTL process.

Fig. 3.14 shows the simulation outcome details. The increased demand speeds up the depletion of the fuels and brings the switch to the backstop nearer, to around year 135. As a consequence of its relatively rapid initial growth, the ROW dominates global consumption quickly.

Leakage rate paths, including DLRs, are given in Fig. 3.15. Emission leakage rate paths (ALR and NLR) remain qualitatively very similar to cases with constant demand. However, the paths are 'squeezed' along the time-axis. This can be explained by the more rapid advancement of the depletion of the fuels due to the increased demand. Oil leakage remains typically above 1 (for discounting scheme 2, even the initial NLR exceeds unity, warranting a negative oil tax), and coal leakage is limited to values not much larger than 0.5. Interestingly, however, oil leakage has slightly decreased whilst

[106] This has the further advantage that (except for today's historic cumulative emissions) we do not need to define any additional parameter.

[107] Roughly half of the emitted carbon is absorbed quite rapidly and the other half stays in the atmosphere for hundreds of years. As this applies equally to the 0.5 TtC of historic emissions as to future emissions E, the proportionality is not affected by this factor of one half. Our formulation does, however, neglect that future emissions contain, besides those from oil and coal, additional carbon emissions from, e.g., gas and land use change, which also contribute a significant proportion. This is the case even though oil and coal contributes 80 % of manmade energy related emissions, cf. section 3.6.

[108] CO_2's molar mass is 3.67 times carbon's atomic mass.

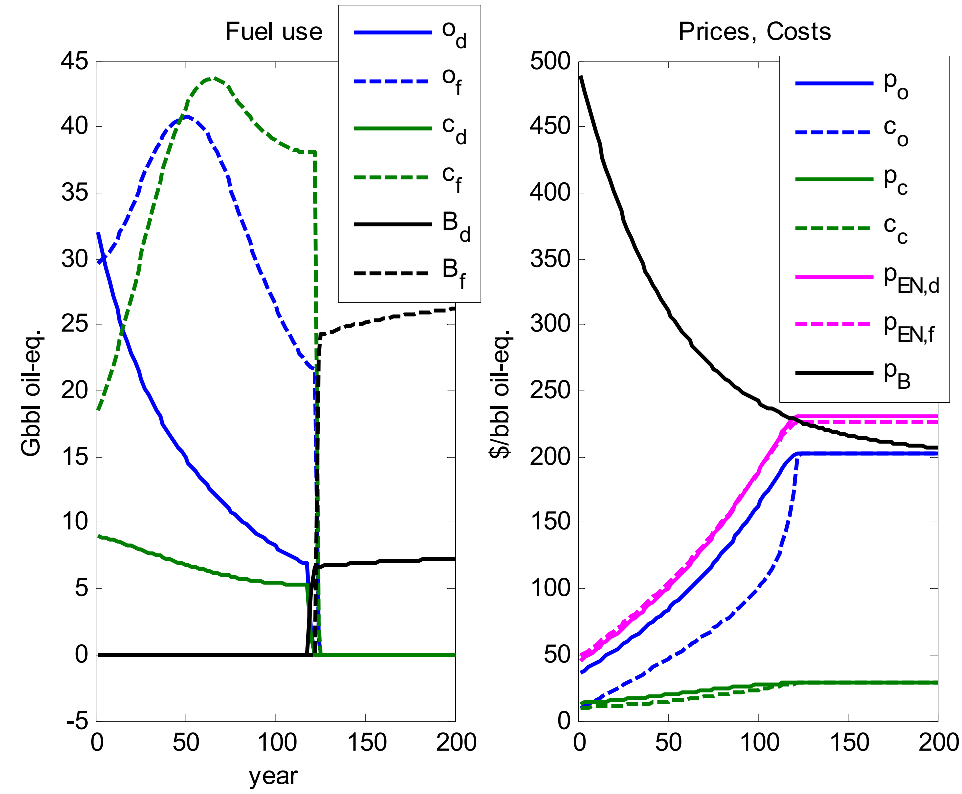

Figure 3.14: Simulation details with growth and backstop

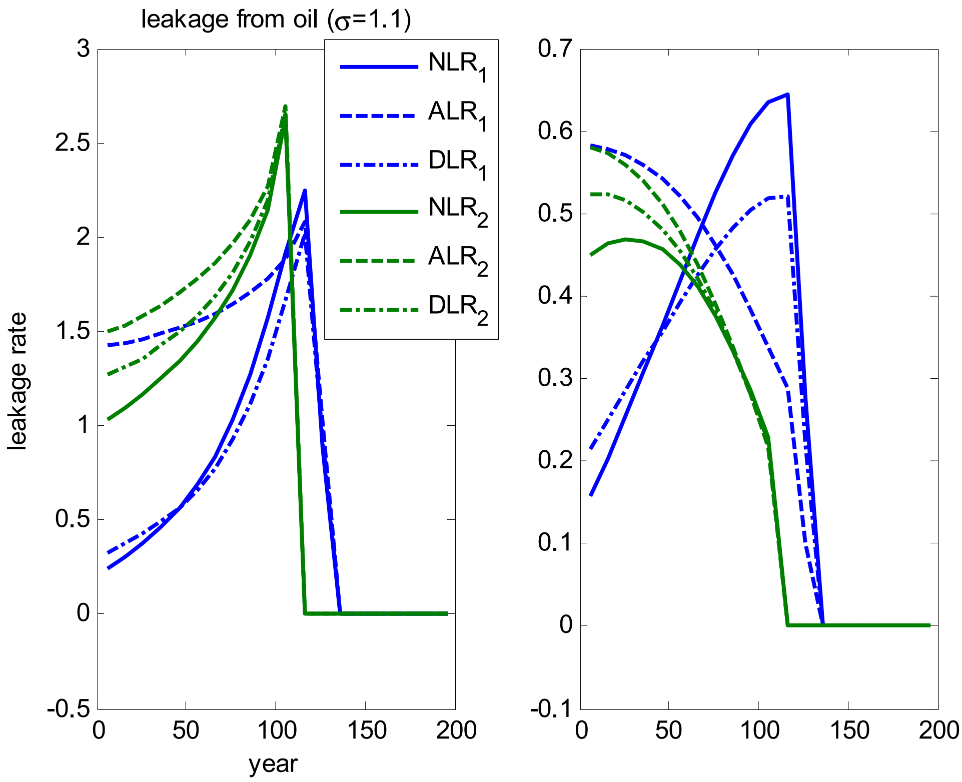

Figure 3.15: Leakage paths with growth and backstop

coal leakage has slightly increased compared to the case without demand growth (section 3.9.4; Fig. 3.11). Given the higher scarcity of oil and the very steep increase of the oil cost curve and the higher demand and thus accentuated scarcity effects, one might have expected oil leakage to increase relatively more compared to coal leakage. That this is not the case seems explained by the fact that the ROW relies relatively more on coal than the OECD (cf. plot 1 in Fig. 3.14), implying that the concentration of demand growth to the ROW tends to increase the global demand (and thus the scarcity) for coal more than for oil.[109]

Adding the damage convexity affects leakage rates only moderately. DLRs are slightly higher than the NLRs for early domestic perturbations but slightly lower for later ones. This is readily explained by earlier emissions being weighted relatively less and later ones relatively more, and the foreign reactions being spread across time. Of course, this modest effect on the leakage *rates* does not imply that the convexity would, per se, be unimportant for the climate policy; in fact, the convexity has itself a strong effect of increasing overall damage, and thus optimal taxes, relative to the case where marginal damages would remain constant at their present level.[110]

3.10 Optimal Tax Structure

Tax base for the evaluation

Part 1 discussed the relationship between the leakage rates and the optimal tax structure, with Eq. (3.11) relating the DLRs and the willingness to pay for global emission reductions to the optimal regional tax. The DLRs calculated in section 3.9 would thus allow the calculation of optimal tax paths, except for a general damage intensity (or global emission disutility) factor, which is exogenous to our analysis. The leakage rates were hitherto calculated, however, based on zero tax rates. Imposing substantial taxes could theoretically change the leakage rates and consequently the optimal tax values, even though the corresponding results from section 3.8 suggest that the influence on the leakage rates, and thus of the optimal taxes, will be very small.

To calculate the optimal tax path more precisely, we ran the simulations iteratively, calculating the leakage rates for an initial (zero) tax path, and derived the optimal

[109] This is confirmed by the ratio of the final coal extraction costs to increase by a higher fraction than for coal, when demand growth is added to the model.

[110] If we compare the NPV direct damage from present emissions in the case where marginal damages remain constant at today's level, to the case when this marginal damage increases over time according to the quadratic damage function and the historic emissions, the convexity increases the damage by a factor 1.7 or 3.1 in the situations with discount scheme 1 or 2, respectively (considering the case with growth and the clean backstop, illustrated in Fig. 3.14).

tax path according to Eq. (3.11) and a damage factor that yields – for today's level of cumulative emissions (hypothetically held constant throughout time) – a marginal long-run damage of 50 $/t$CO_2$. In the absence of leakage this would imply in the model an optimal tax of more than 50 $/t$CO_2$ that would rise over time; the increase would take place since additional emissions boost the marginal damage from (cumulative) emissions given that we consider a convex damage curve. In the next step, leakage rates were recalculated, starting from the just calculated tax path. We repeated this until convergence, which, as in section 3.8, was almost immediate; we thus stopped after four iterations where no notable changes were found anymore. Moreover, the differences between the leakage rates evaluated starting from zero tax rates and those from the optimal tax path hardly differ at all, confirming again what we found in section 3.8. Fig. 3.16 illustrates this with the example of the scenario without liquefaction and backstop (results for other scenarios show similarly small differences) for which we plot (the convergence of) DLRs and the corresponding optimal tax paths, shown for the first 100 years.

Optimal Tax

We spare a description or plot of the optimal tax paths for all the scenarios analyzed in section 3.9. Instead we recall that the optimal regional tax rate is, at each point in time t, proportional to $1-\mathrm{DLR}_t$, but, as given in Eq. (3.11), also to the net current value sum of all future marginal damages from point t on, which for the convex damage function increases over time as long as global emissions are not zero. As we have observed, in general the leakage rates given seem to be valid approximations for the leakage rates in the case where the OECD sets the regionally optimal tax rate.

3.11 Discussion

The results of the dynamic analysis show how emission leakage rates can have large magnitudes, how they strongly depend on the time spans (discounting) considered, how they vary dramatically with the exact scenario considered (assumptions about extraction cost curves, substitutabilities, presence of liquefaction or clean backstop), and how they can strongly differ across fuels. Whilst several key findings, some of which seem surprising even though they can be explained, are discussed above, admittedly, without strong assumptions on the exact fuel market setup of the future, it seems very hard to suggest a narrow range for realistic leakage rates, in relevant NPV (damage) terms. That is, if there is one overarching conclusion of the study, this may best be described as the uncomfortable implication that a more or less precise estimation of

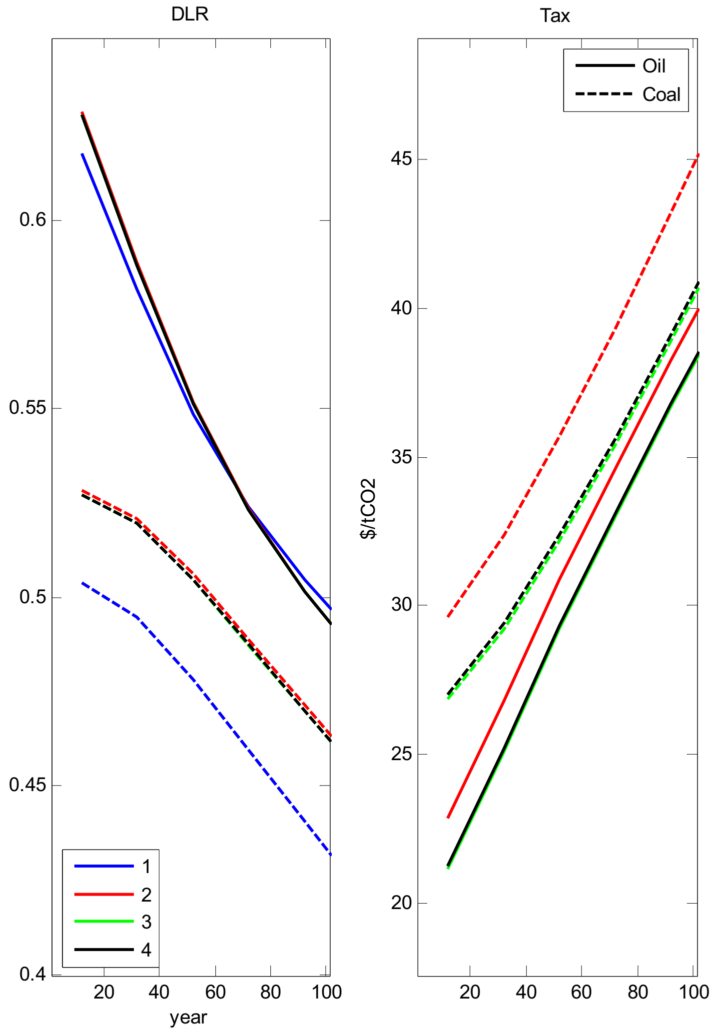

Figure 3.16: Convergence of tax path

The line colors 1-4 indicate the first through the fourth iteration.

welfare-relevant leakage rates will require not only considerable information about the current fuel market conditions but also significant information about the prospects for technical developments (e.g., the development of fuel transformation processes, alternative energy sources, technologies that may change our lifestyle and, consequently, the fuel demand pattern) or political developments (e.g., global climate treaties) concerning greenhouse gas emissions. Equally inconvenient is that any policy relevant leakage index will strongly depend on the controversial time-discounting of future greenhouse gas emissions. Explicitly making and stating assumptions about such parameters and about future developments, along with any proposed leakage rate, seems to be the only viable option. Nothing is gained from neglecting uncertainties and implicitly assuming these away, e.g., by relying on a dubious concept of a static fuel supply or considering only contemporaneous leakage during the next few years despite the long-term character of the climate problem. If, with an important though controversial probability, large-scale liquefaction will emerge in the future and if, in this case domestic oil savings could – as the present analysis suggests – be subject to a negative leakage rate, and if, without that liquefaction, domestic oil savings would be subject to a large positive leakage rate of around 50 %, potentially even above 100 % – as also suggested in our model –, then economic models should, at least insofar as they aim at estimating the truly relevant leakage rates, take these possibilities into account, despite the uncertainties attached, rather than solely focusing on a business-as-usual baseline and implicitly attributing a 100 % probability to its materialization. Finally, the skepticism expressed in this article against the traditional leakage literature may be rephrased as follows: many will agree that one cannot be sure whether a major fraction of the realistically exploitable fossil fuels will in the long run be left underground or whether practically all of these fuels will be consumed by future generations. In the latter case, it seems clear that regional emission savings during the next few decades are ultimately subject to a leakage of close to 100 % in terms of undiscounted emissions, at least if fuel is imported from a clearly globalized worldwide market, as it exists today for case of oil but increasingly also for the other fossil fuels.[111] The surveyed studies hardly provide any substantive economic reasons why this scenario should be impossible. Yet, they suggest deterministic, modest leakage rates. As far as sensible economic depletion models for the fossil fuels are used, those rates will, notably, depend on the time-horizon of the model simulations. In some cases it may indeed make some sense to assume limited horizons, rather than quasi-infinite ones, as one may attach a value to know that emissions be at least delayed for a couple of decades. This is a preference that may also be funded in the belief that technological or political progress hopefully will prevent the emissions from a certain

[111] See also Habermacher (2012b) for a discussion of this possibility.

point in time on. The judgment on the value of such a delay strongly depends on personal perceptions and beliefs about the future. A corresponding leakage rate should therefore be proposed together with explicit statements about the assumptions under which it is obtained and ideally tested for deviations from these assumptions. If this was broadly acknowledged, gradual discounting of emissions, rather than a simple and somewhat arbitrary cutoff of the simulation time-horizon, would surely be preferred.

The numerical analysis in the present study is based on a relatively simple fuel demand (utility) system with parameter values calibrated to fit current fuel consumption and prices, and some elasticity values inspired by the literature. Although a sensitivity analysis in a study closely related to the present one, using a slightly different model,[112] has found model results to be rather robust to changes in a variety of parameters and assumptions, it would be interesting to further examine the core issues of this paper – the time dimension of carbon leakage from a market-based regional climate policy and the fuel-dependent structure of the optimal regional policy – within a more detailed model in a multisectoral framework. An adequate representation of the fuel substitutabilities (in specific applications) and fuel transformation processes, such as coal-to-liquids, would be crucial for accurate modeling (cf., e.g., Lanz and Rausch, 2011, who show that the inclusion of bottom-up elements is necessary for a general equilibrium model to accurately represent the electricity sector and its emissions). For example, instead of the here considered clean backstop that directly replaces the fossil fuel *aggregate*, a more detailed characterization of different alternative energy technologies could make the model more realistic, potentially also qualifying some of the present findings. Thus, complementing a multisectoral top-down model with bottom-up elements concerning the substitutability of fossil fuels in the major fuel-consumption domains could be an interesting point for future research on the topic addressed in this paper (see, e.g., Chen et al., 2011, for a dynamic model in which a top-down approach is coupled with a bottom-up representation of coal liquefaction processes). However, clearly this should not come at the price of giving up the here specifically considered supply aspects with forward-looking resource owners.

Finally, that leakage effects would imply that fuels not consumed in a climate-protecting region would be consumed elsewhere in the world is one of the strongest political arguments against stringent unilateral climate policy. Thus, properly accounting for such leakage effects in the dimension of *fuel*-specific carbon policies may not only imply an efficiency gain but specifically increase the political acceptance of unilateral action.

[112] Habermacher (2011), conducted the study using a dynamic model to calculate optimal constant tax rates based on some simplifying assumptions and a calibrated utility quadratic in oil and coal, largely analogous to Golombek et al. (1995).

The present analysis does not explicitly take natural gas into account, and therewith ignores a fossil fuel which contributes 20 % of the carbon emissions from energy supply worldwide. No obvious reason as to why the main results of the present analysis should be fundamentally altered when natural gas is modeled as well exists. Nevertheless, taking into account this third most important fossil fuel in terms of current consumption may still have a significant effect on the response of the energy market and emissions to a unilateral policy. First, gas may be a better substitute to both, oil and coal, than those two fuels are between themselves: Gas is widely used to feed power stations, which are also the most important consumers of coal, but much less of oil.[113] And gas is also a relatively good substitute for oil in major applications such as domestic heating and transport, where the direct use of solid coal is less trivial. Accounting for gas would thus not only add a fuel to the model which could be substituted easily for one of the other fuels, but it could indirectly increase also substitutability between oil and coal, as changing the use of one of these two major fuels would affect the demand for gas which in turn would affect the demand for the other major fuel. An increased substitutability between the fuels can narrow the gap between the leakage rates across the different fuels. At the same time, it would not directly reduce the overall leakage rate for a general emissions tax. Second, gas is relatively expensive to transport. Pipeline transportation is much more costly for gas than for oil. Both oil and coal can be shipped over short and long distances much more cheaply than gas, which requires capital and energy intensive liquefied natural gas (LNG) fascilities. This seems to be a main reason why the shale gas 'revolution' in the US leads to very low current prices for natural gas in the US compared to overseas, with a current production of shale gas in the US (and worldwide) that remains limited (less than 200 bcm in 2010) compared to worldwide gas production (more than 3000 bcm).[114] From a medium or long-run perspective, the current US gas price anomaly and the significant transportation costs are, however, not indicating a true segregation of the global gas market: Even though LNG-transportation costs are substantial compared to current energy prices, they are not prohibitive for an interregional gas trade when large and sustained price differentials are foreseeable.[115] A third point concerns the greenhouse-gas intensity of the increasingly important unconventional gas resources. A

[113]Coal, natural gas, and oil contribute 41 %, 21 %, and 5 % of worldwide power generation (OECD, 2012). Some see (shale) gas as an important step to reduce the emission intensity of the energy system, notably due to substitution of coal (e.g., Helm, 2012). This view is supported by the strong decrease of US energy related carbon emissions 2007-2012 which seems at least partly due to the substitution of coal by gas in power production.

[114]In total, unconventionals (shale gas, tight gas and coal bed methane) have a share of around 12 % in global gas production (IEA, 2010).

[115]As is evidenced, for example, by LNG exports from South-East Asia, the Middle East, Africa and South America to Europe and Japan, and, at least until recently, to the US.

first study that included methane emissions suggested that shale gas may even be much more emission intensive than coal (Howarth et al., 2011). The study has been widely criticized, and more recent peer-reviewed studies conclude, on the contrary, that the overall emission-intensity of unconventional gas hardly exceeds that from conventional natural gas, and that shale gas powered electricity is substantially more climate friendly than coal power (e.g., Hultman et al., 2011, and Jiang et al., 2011). The controversy [116] is far from settled, and it is thus unclear what the differential in the emission intensity between conventional and unconventional gas could imply for the leakage rate. Overall, it seems plausible that an increased availability of gas in the medium and longer-term – the inclusion of shale-gas has increased the technically recoverable worldwide resources by 47 % according to EIA (2013b), a number which is, however, subject to very high uncertainty[117] – affects the quantitative results to a non-negligible degree, but it cannot be said *a priori* in which direction.

3.12 Conclusion

We provide a method to disentangle the terms-of-trade and the pollution part of an optimal regional climate policy in a dynamic framework with an exhaustible fuel, and define a welfare relevant leakage rate related to it.

We calculate fuel-specific leakage rates for current and anticipated future taxes, in a stylized, calibrated model of the dynamic market for major fossil fuels. The model considers an exogenous, downward sloping demand for energy, which can be provided as a constant elasticity of substitution aggregate of weakly substitutable oil and coal, or, potentially, by a clean backstop with infinite supply elasticity and a cost that diminishes over time. The fuel demand is split between the OECD and the rest of the world, and calibrated according to current demand and worldwide prices; held constant over time in the main setup, demand is allowed to grow in an extension of the model. We exclude non-fuel trade between the two fuel consuming regions and focus only on the fuel channel of leakage. The fossil fuels are extracted for marginal costs that increase with cumulative extraction according to empirical estimates, and they are offered by forward-looking competitive suppliers, corresponding to a standard Hotelling framework with an exogenous discount rate. We find that leakage rates for OECD fuel emission reductions may vary strongly in magnitude, and even sign, across fuels and as

[116] See Stevens (2012) for an overview of the debate.

[117] Stevens (2012) provides a brief overview on the uncertainties attached with estimates of technically recoverable shale gas. For example, in 2012 Poland reduced it's estimate of technically recoverable resources to around one-tenth of the initially indicated figure.

a function of the considered scenarios. For example, when coal-liquefaction supplements the supply of liquid fuel, domestic oil savings can be subject to negative leakage rates: increasing the availability of oil on the global market delays the use of the very emission intensive liquefaction process and therewith implies that even foreign emissions are reduced when less domestic oil is consumed; this can imply negative leakage rates even for oil reductions during times before liquefaction has become economically viable. In the same scenarios with liquefaction, coal emission leakage rates remain positive, as more coal available on the global market implies more foreign coal consumption, both for direct use and for use as an input to liquefaction. The variations across scenarios are so large that it would seem questionable to indicate here a specific guess for the exact value of the real expected leakage rate for an emission tax. Depending on the scenario considered, they may be as low as 10 % for a fuel, such as found for coal taxed in early periods in the setups without liquefaction: the relative abundance of coal compared to the more 'limiting' factor oil implies that a bit more coal hardly changes the rate (and, in the presence of the backstop technology emerging in future, duration) of aggregate fossil energy use, so that a domestic reduction implies only a small foreign increase in coal use and emissions. As in addition, the foreign emission reaction is spread over time, the NPV leakage rate becomes especially low for early domestic coal reductions, starting at around 10 % for immediate reductions. In other cases, the leakage rates may exceed unity, as found for oil in presence of the clean backstop and without liquefaction: given that oil depletes very fast, it becomes essentially the limiting factor among the two fuels aggregated with a relatively weak substitutability to overall fossil energy. Increasing the availability of oil thus prolongates the time that the fossil-fuel aggregate remains competitive against the clean backstop technology, and because relatively much coal is burned per unit of oil used for the aggregate fossil energy provision, this extension of the fossil fuel use implies a lot of additional foreign emissions per unit of oil-emissions saved at home, overall leading to a leakage rate that tends to exceed unity, especially for ALR, where the emissions from the additional use of the fossil aggregate are not discounted. The time dimension of the problem appears to be of overwhelming importance; the rates depend very strongly on discount rates and on future developments in the fuel market, such as the emergence of liquefaction processes or a clean backstop replacing the fossils.

The fact that the leakage rates depend so strongly on future developments qualifies numerous semi-empirical estimates provided in the literature without any explicitly stated (or discussed) assumptions about future technical or political developments on the fossil fuel markets; the concept of static leakage rates, and of undiscounted leakage throughout a specific and limited time-period, must be reconsidered. Contrary to what most leakage studies, often focused on a static fuel demand, suggest, leakage may offset an overwhelming fraction of domestic emission reductions.

3.13 Annex

Annex 1 Optimal Pollution Tax, Partial Equilibrium

Consider a numeraire good z and a polluting good x that costs p and whose global consumption leads to a proportional pollution damage, for the avoidance of which the home region has a marginal WTP d.

Consider a domestic (h) and a foreign (f) region, two regions $r = \{h, f\}$, with domestic utility $U_h = z + \log x_h - d \cdot X$, subject to the budget constraint $z = z_0 - p \cdot x_h$, with global consumption $X = x_h + x_f$, and some leakage, which can be written as $x_f = x_{f0} - \alpha x_h$, where α is the leakage rate.

The domestic planner's FOC for domestic consumption x_h writes

$$p \stackrel{!}{=} \frac{1}{x_h} - d(1 - \alpha). \tag{A3.1}$$

Domestic decentralized consumption decisions, subject to a potential tax τ_h imposed by the domestic government, are given by the FOC which takes into account that private consumption has a negligible effect on the regional consumption level (as well as on the redistributed tax proceeds), that is, the direct marginal consumption utility must equal the private costs, $\frac{1}{x_h} \stackrel{!}{=} p + \tau$. In this simple setup, the optimal level of domestic consumption implicitly given by Eq. (A3.1) can thus be sustained in a decentralized market by imposing a domestic pollution tax of the level

$$\tau_h^* = d(1 - \alpha).$$

Proposition 1. *At constant prices, if only global pollution matters and if foreign consumption of a polluting good increases proportionally at rate α when domestic consumption is reduced, i.e., we have a leakage rate of α, the regionally optimal level of the unilateral pollution tax τ_h^* is $\tau_h^* = d(1-\alpha)$, where d is the domestic WTP for global pollution reductions.*

Whilst it surely makes quite some intuitive sense that the optimal tax may be proportional to 1 minus the leakage rate α, as the region has a WTP for global emission reductions and those emission reductions, after all, are $1 - \alpha$ per unit of domestic pollution avoided, the proposition is based on the assumption of fixed prices. The inconvenient truth about this assumption is that leakage naturally occurs exactly via price effects – it is just, e.g., by affecting the global price of fuels (or of that traded goods) that domestic demand changes affect foreign fuel consumption.

Annex 2 Optimal Pollution Tax and Terms-of-Trade Effects

Assume decreasing returns in production of a good X, supplied by external supplier for a price equaling marginal costs,

$$p \equiv a + bX \quad \text{with } a, b > 0. \tag{A3.2}$$

Be domestic utility U^* linearly separable in the consumption of the fuel, with

$$U^* \equiv z + \log X, \tag{A3.3}$$

implying the budget constraint being $z = z_0 - pX = z_0 - aX - bX^2$. We define $U \equiv U^* - z_0$, implying $U = \log X - aX - bX^2$, and maximizing U is equivalent to maximizing U^*.

The regional planner's FOC is $\frac{1}{X} \stackrel{!}{=} a + 2bX$, implying $X^* = \frac{-a+\sqrt{a^2+8b}}{4b}$. In contrast, the decentralized consumer takes the fuel price as a given and chooses according to the decentralized FOC for $U = \log X - pX$, namely $\frac{1}{X} \stackrel{!}{=} p$, which, in equilibrium, implies $\frac{1}{X} \stackrel{!}{=} a + bX$, yielding the (suboptimally high) free market consumption $X_m = \frac{-a+\sqrt{a^2+4b}}{2b} > X^*$.

A correcting (unit) tax τ ensures that even the market outcome yields the optimal consumption level X^*: With the tax, the decentralized FOC implies $\frac{1}{X} \stackrel{!}{=} a + bX + \tau$ and thus $X_{m,\tau} = \frac{-(a+\tau)+\sqrt{(a+\tau)^2+4b}}{2b}$. Requiring the tax to bring the consumption level to the optimum, we have τ implicitly defined by $\frac{-(a+\tau)+\sqrt{(a+\tau)^2+4b}}{2b} \stackrel{!}{=} \frac{-a+\sqrt{a^2+8b}}{4b}$, yielding $\tau^* = \frac{1}{4}\left(\sqrt{a^2+8b} - a\right) > 0$, which is increasing in b and decreasing in a.

Adding external pollution damage, d, we have $U = \log X - aX - bX^2 - dX$. By analogy to the case without damage, it is trivial to see that this yields the regionally optimal level $X_d^* = \frac{-(a+d)+\sqrt{(a+d)^2+8b}}{4b} < X^*$. As the decentralized actors ignore the damage their consumption induces when choosing X, the market consumption is still $X_{m,\tau}$ from above, and the optimal tax is thus implicitly given by $X_{m,\tau} \stackrel{!}{=} X_d^*$, i.e., $\frac{-(a+\tau)+\sqrt{(a+\tau)^2+4b}}{2b} \stackrel{!}{=} \frac{-(a+d)+\sqrt{(a+d)^2+8b}}{4b}$, yielding

$$\tau_d^* = \frac{1}{4}\left(\sqrt{a^2+8b+2ad+d^2} - a + 3d\right). \tag{A3.4}$$

Recognizing that $\frac{\partial}{\partial d}\sqrt{a^2+8b+2ad+d^2} = \frac{1}{\sqrt{1+\frac{8b}{a+d}}} \in (0,1)$, we see that Eq. (A3.4) implies

$$\tau^* + \frac{3}{4}d < \tau_d^* < \tau^* + d,$$

that is, the optimal overall tax rises less rapidly in presence of a pollution externality than the pollution externality itself, with a rate that is smaller the larger the slope of the price, b, is compared to both, the level of the price (a) and the externality (d). We emphasize this result in the proposition 2. The increase of the overall tax does, however, in this framework amount to more than $\frac{3}{4}$ of the pollution externality.

Proposition 2. *Consider a good offered and consumed according to Eqs.* (A3.2) *and* (A3.3)*. A specific level of marginal pollution-externality attached to the consumption of that good increases the optimal tax levied by the region on that good by less than the level of the marginal pollution-externality. The difference between the level of the externality level and the optimal environmental tax level is of first order degree, i.e., the ratio* $d/(\tau_d^* - \tau)$ *is non-marginally above 1 even for an asymptotically small environmental externality level* d*, when* τ_d^* *is the optimal buyer tax with, and* τ^* *that without pollution.*

Proof: given above.

An intuition for this result is as follows: In absence of pollution, the region optimally levies a tax τ^* on the consumption that corrects for the distortion arising from that fact that an individual consumer takes into account only the direct price she pays for her individual consumption but ignores the price-increasing effect of her purchase on the price the others pay for their consumption. As we see in the expression for the optimal tax without environmental damage, this correction is stronger when a lower a yields a larger amount of consumption (more consumers suffer from the price-increasing effect of additional consumption by an individual), and when the price-increase by a marginal quantity consumed is larger, i.e., when b is larger). Given that with environmental damage, and the corresponding additional climate tax component, the amount of the good purchased overall is reduced, the price-increase-effect of marginal consumption by an individual harms less other consumers, i.e., the optimal import tariff component is reduced. As the optimal import tariff component is thus reduced as the environmental damage increases, the optimal overall tax level rises less rapidly than the environmental externality level.

If terms-of-trade effects imply positive optimal unilateral taxes even in the absence of pollution, it can thus be non-trivial to disentangle the non-pollution and the pollution component of an optimal overall tax. As we have the inequality $\tau_d^* < \tau^* + d$ for the case of zero leakage, we can not expect $\tau_d^* = \tau^* + d(1 - \alpha)$ to hold either. In fact, from the first-order deviation, without further investigation that at least for low enough leakage rates we know that we have $\tau_d^* < \tau^* + d(1 - \alpha)$: note that for an arbitrarily small foreign country and thus an arbitrarily small leakage rate, the optimal overall home tax is arbitrarily close to the above τ_d^*, and in the case without pollution it would be arbitrarily close to τ^*. Thus, for small enough α, in this case we know that the optimal overall tax falls short of the sum of the optimal tax in the absence of pollution and $1-\alpha$ times the externality rate.

Annex 3 Compensation in Static Framework

Consider regional welfare functions of the form $U = z_0 - p(E)\cdot e + u(e) - D(E)$, where z_0 is regional numeraire good consumption, e and E regional and global fuel consumption, $p(E)$ the fuel price as a function of the global consumption $E = e_h + e_f$, $u(e)$ is the fuel consumption utility and $D(E)$ the damage from global emissions.

For decentralized decisions of foreign consumers we know $u'_f(e_f) = p(E)$. Knowing the domestic region compensates the foreigners for the tax induced changes (ignoring environmental damage), we can directly take the foreign consumption and the fuel producers' profit (U_p) into account in our maximization and omit to write the transfer, $U^* \equiv U + U_f + U_p$. The maximization problem writes $\max_e U^{**} = z_0 - C(E) + u(e) - D(E) + u_f(e_f)$, s.t. $u'_f(e_f) = c(E)$, and $E = e + e_f$ and with $C(E) = \int_0^E c(E)\mathrm{d}E$, yielding $C'(E) > 0$ and $C''(E) > 0$, i.e. marginal costs of extraction increase as $c'(E) > 0$.[118] Similarly, we write total damage $D(E) = \int_0^E d(E)\mathrm{d}E$. As usual we assume $p(E)$ to be increasing: $p'(E) > 0$.

The FOC writes

$$\begin{aligned}
\frac{\partial C(E)}{\partial e} + \frac{\partial D(E)}{\partial e} &= u'(e) + \frac{\partial u_f(e_f)}{\partial e} \\
c(E)\frac{\partial E}{\partial e} + d(E)\frac{\partial E}{\partial e} &= u'(e) + u'_f(e_f)\frac{\partial e_f}{\partial e} \\
c(E)\frac{\partial E}{\partial e} + d(E)\frac{\partial E}{\partial e} &= u'(e) + c(E)\frac{\partial e_f}{\partial e} \\
c(E)\left(1 + \frac{\partial e_f}{\partial e}\right) + d(E)\left(1 + \frac{\partial e_f}{\partial e}\right) &= u'(e) + c(E)\frac{\partial e_f}{\partial e} \\
c(E) + d(E)\underbrace{\left(1 + \frac{\partial e_f}{\partial e}\right)}_{\text{leak.adj.emiss.}} &\stackrel{!}{=} u'(e). \qquad \text{(A3.5)}
\end{aligned}$$

Thus, the marginal utility of energy consumption is to equate marginal social costs of energy: extraction cost plus damage with emission factor adjusted for the relative impact of own emissions on foreign emissions (the leakage rate).

[118] Note the fuel owners' profit canceled out since, as a transfer it does not affect the sum of welfare overal all actors.

Decentralized domestic consumption under a tax is given by the FOC

$$p(E) + \tau \stackrel{!}{=} u'(e), \tag{A3.6}$$

where we know $p(E) = c(E)$, wherewith Eq. (A3.6) shows that, according to Eq. (A3.5), for the tax to sustain the optimal level of consumption with compensating transfers to the producer and to the other region, the tax required is

$$\tau^* = d(E) \underbrace{\left(1 + \frac{\partial e_f}{\partial e}\right)}_{\text{leak.adj.emiss.}},$$

that is, it is fully independent of the terms-of-trade effects which would depend on the *change* of the price (and thus the *change* of the marginal extraction cost) induced by consumption changes. We thus have confirmed that the compensation of both, the producer and the foreign consuming country, allows, in this setting, to isolate the optimal emissions tax from any terms-of-trade tax component. This corresponds to what Böhringer et al. (2010) have shown in their static framework for the analysis of sector-specific leakage.

Annex 4 Details Numeric Calibration

Current Prices and Regional Consumption of Fuels for Calibration are given in Table A3.1.

Table A3.1: Current fuel consumption and prices

Current Consumption		
Using IEA WEO2010 Data	Oil (bio.bbl/yr)	Coal (bio. t/yr)*
OECD	16.4	1.61
Non-OECD	14.3	3.12
World	30.7	4.74
* in Mtce		
Relevant Current Prices (average from 2006-2010, in US 2010 $)		
Using Worldbank Pink Sheet Data (2011)	Price	
Oil ($/bbl)	76	
Coal ($/t)	83	

Sources: IEA (2010) and World Bank (2011)

Annex 5 Model Run Basic Setup

Fig. A3.1 shows the model results for a basic setup, with a backstop but without liquefaction and with constant demand.

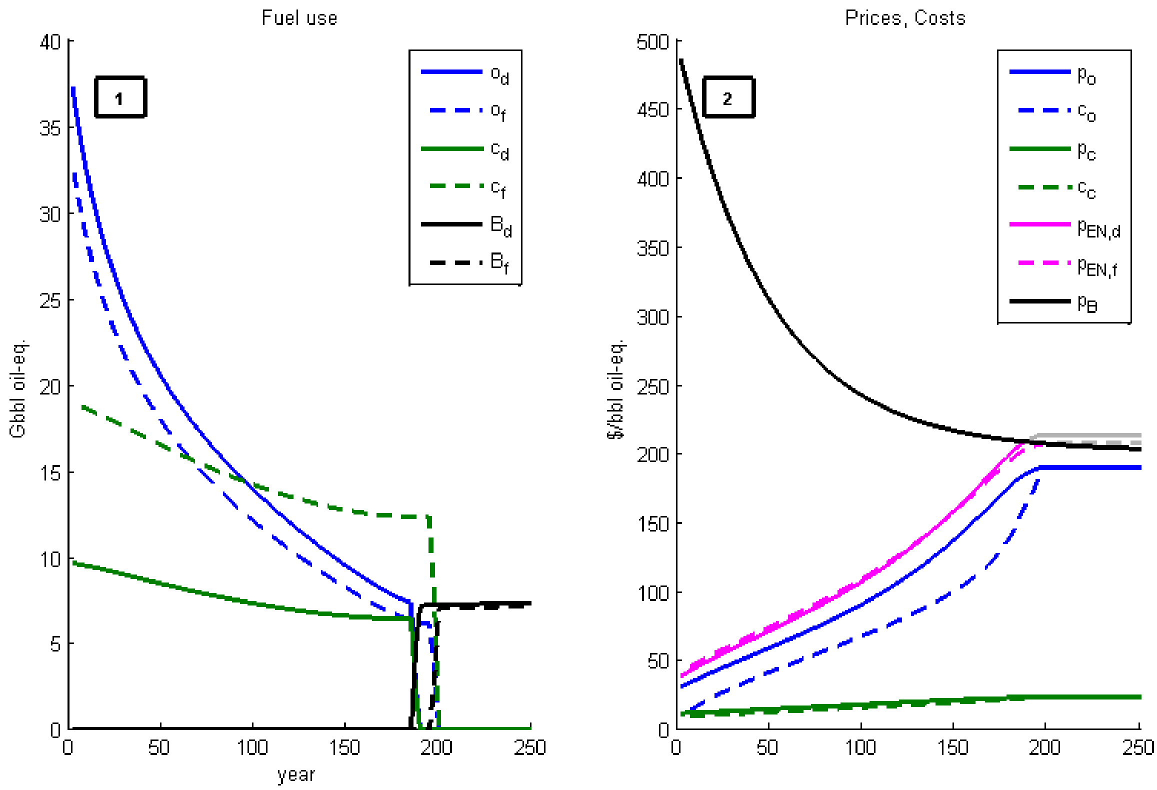

Figure A3.1: Simulation results basic setup

Annex 6 Scale-Invariability for Linear Costs and Isoelastic Demand

The supplier's price their fuel on the market according to

$$p_t = c_t + \int_t^{\infty} e^{-\rho(s-t)} \dot{c}_s \mathrm{d}s. \tag{A3.7}$$

We have an isoelastic demand,[119]

$$x_t = \xi p_t^{\varepsilon}, \tag{A3.8}$$

and extraction cost proportional to cumulative extractions, $c_t = b \int_0^t x_s \mathrm{d}s$.

Consider an γ-fold increase of the extraction costs, i.e., $c_t^* = \gamma b \int_0^t x_s \mathrm{d}s$. Assume this increases the resulting cost path by a constant factor α, that is,

$$c_t^* = \alpha c_t \, \forall t.$$

This respects the extraction cost curve when we have

$$x_t^* = \frac{\alpha}{\gamma} x_t. \tag{A3.9}$$

The supplier's pricing equation, Eq. (A3.7), remains respected if we have

$$p_t^* = \alpha p_t. \tag{A3.10}$$

The remaining condition which must be fulfilled is the demand equation, Eq. (A3.8). With Eqs. (A3.9) and (A3.10) this implies $\frac{\alpha}{\gamma} = \alpha^{\varepsilon}$, which holds *iif* $\alpha = \gamma^{1/(1-\varepsilon)}$.

In the linear oil-cost curve models in section 3.9.1, we have $\varepsilon = -0.9$ and, neglecting the relatively small intercept in the extraction cost curve, an inverse reserve-scaling factor $\gamma = 10$, implying that p_t and c_t increase by a factor $\alpha = 10^{1/1.9} = 3.4$ and extraction is reduced by the factor $\frac{\gamma}{\alpha} = 2.98$. Despite the not perfectly negligible intercept of the extraction cost curve, this approximates extremely well what happens in the simulation, where the price and the extraction costs increase by a factor 3.34 and extractions are reduced by the factor 2.90 (values calculated for the last simulated period; plots omitted here have shown these values to remain almost perfectly constant also in other periods).

[119] In the leakage model, we have two regions demanding the fuel. This can readily be accounted for by assuming the here used ξ to be the sum of the corresponding regional demand parameters, $\xi = \xi_d + \xi_f$, not affecting the remainder of the analysis here.

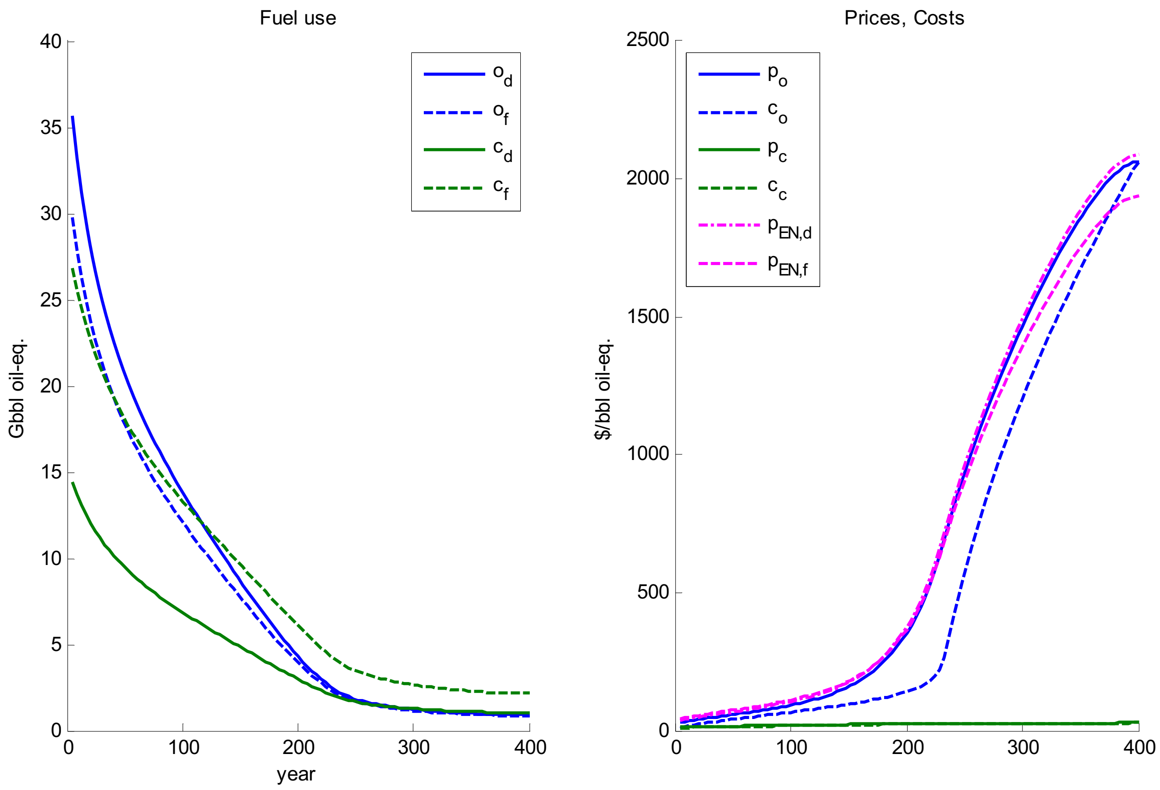

Figure A3.2: Simulation results basic setup

Scenario with constant demand, no backstop and no liquefaction. Discounting scheme 1.

Annex 7 Supplementary Graphs Basic Setup

Simulation results for the basic setup: constant demand, no liquefaction, no backstop, considering discounting scheme 1.

Figs. A3.2 through A3.4 give detailed simulation results, for unperturbed domestic equilibrium consumption.

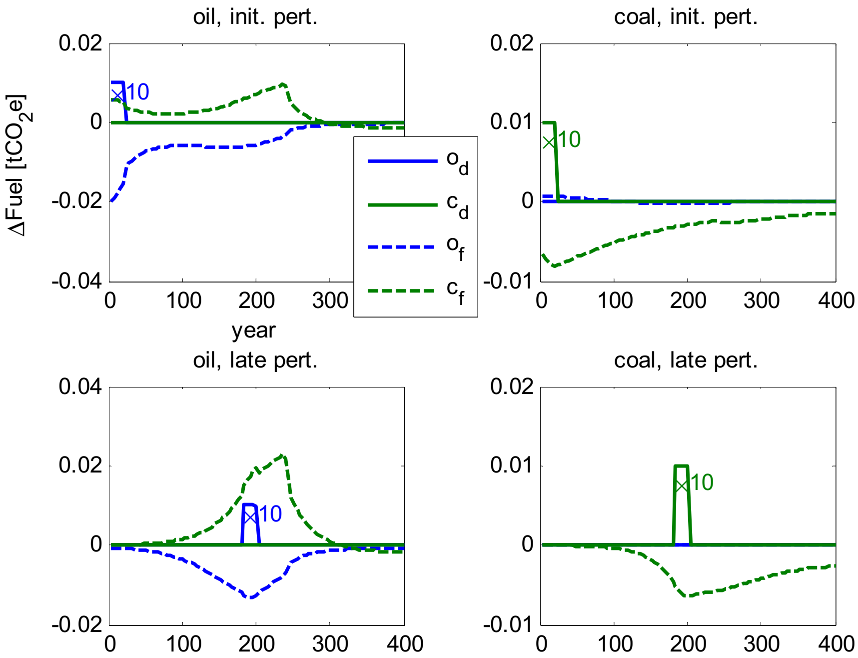

Figure A3.3: Emission reaction, basic setup, stronger substitutability

Scenario with constant demand, no backstop and no liquefaction. Substitutability increased to $\sigma = 1.7$. Discounting scheme 1.

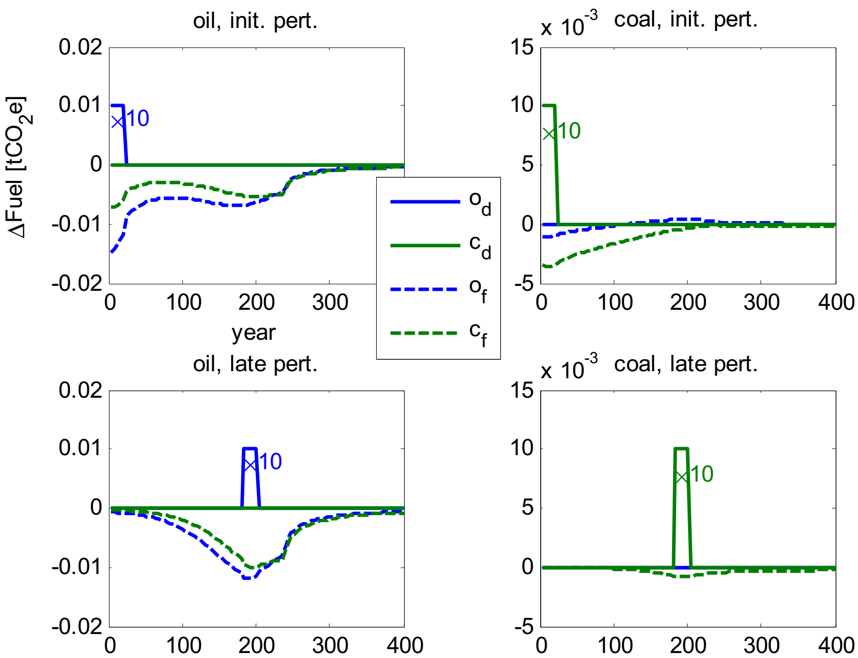

Figure A3.4: Emission reaction, basic setup, strong complementarity

Scenario with constant demand, no backstop and no liquefaction. Complementary fuels, $\sigma = 0.3$. Discounting scheme 1.

Annex 8 Supplementary Graphs Liquefaction

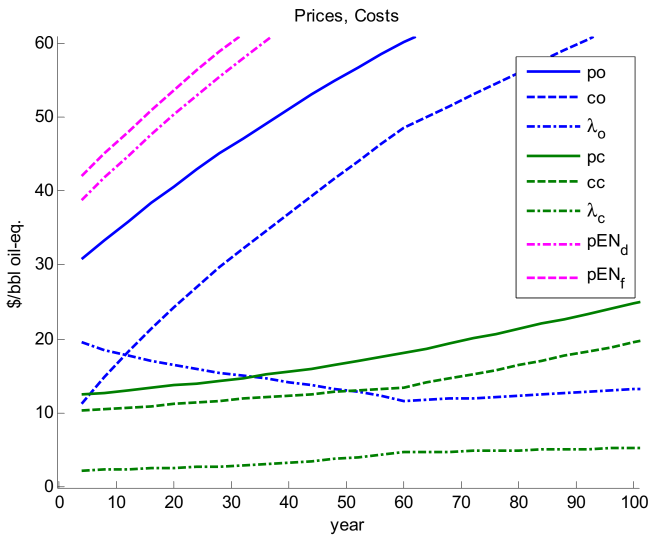

Figure A3.5: Zoom, simulation details with liquefaction

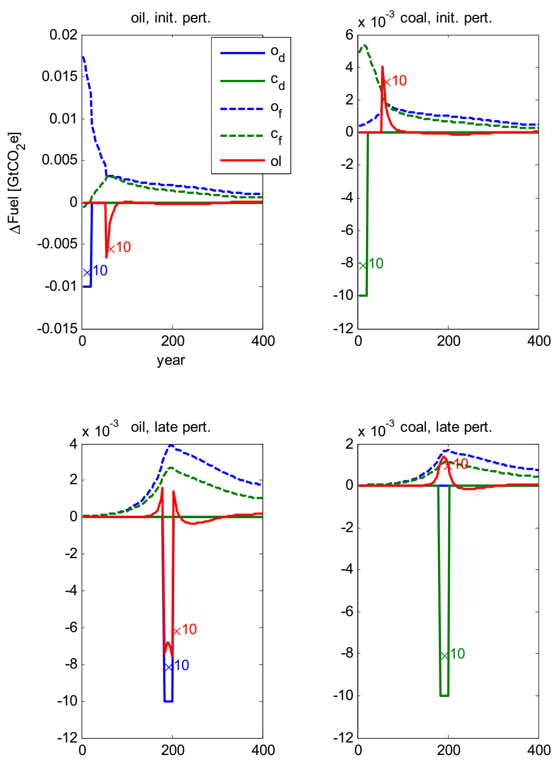

Figure A3.6: Emission reaction paths, liquefaction

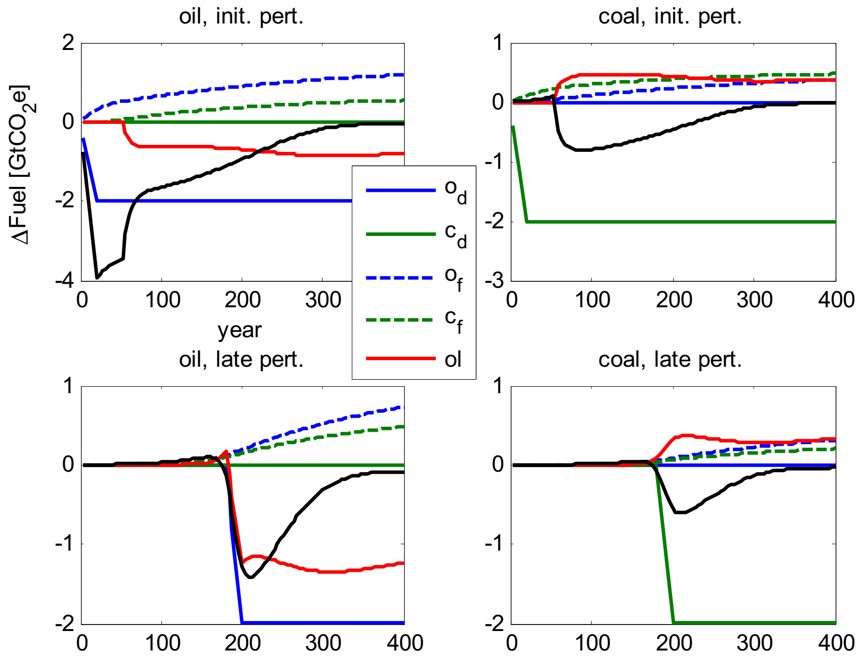

Figure A3.7: Cumulative emission reactions, liquefaction

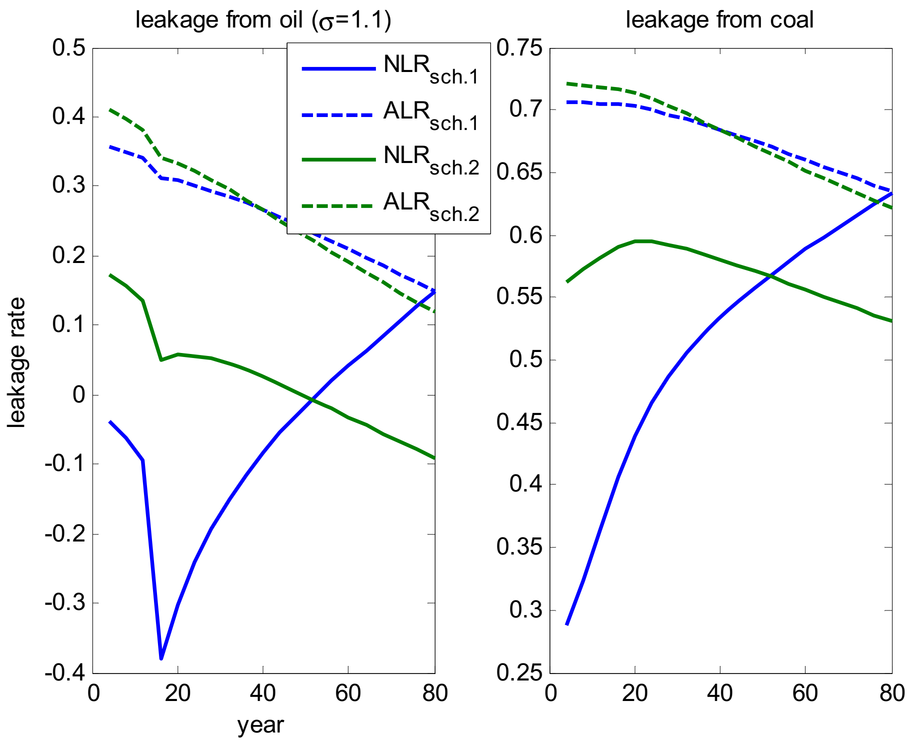

Figure A3.8: Leakage paths with liquefaction, for c_l=0 \$/bbl

Annex 9 Supplementary Graphs Backstop

Figs. A3.9 through A3.11 plot instantaneous and cumulative emission reactions for domestic medium-term consumption reductions, as well as leakage rate paths for stronger and weaker than standard fuel substitutabilities.

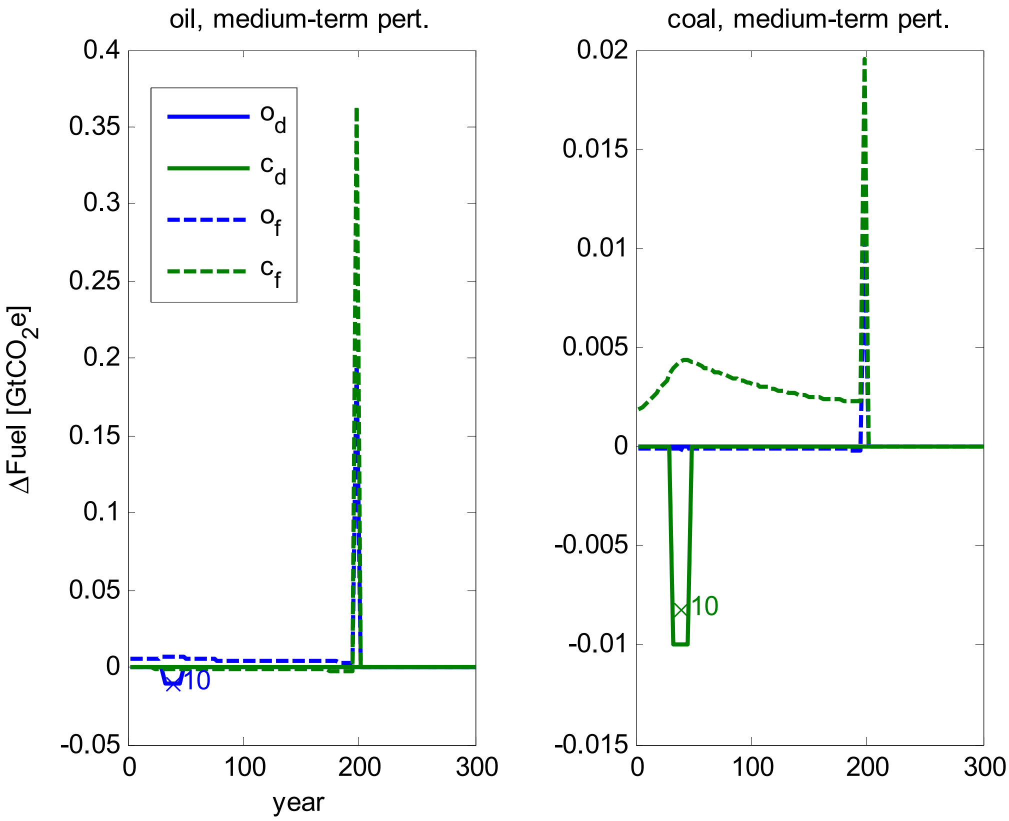

Figure A3.9: Emission reaction paths, backstop

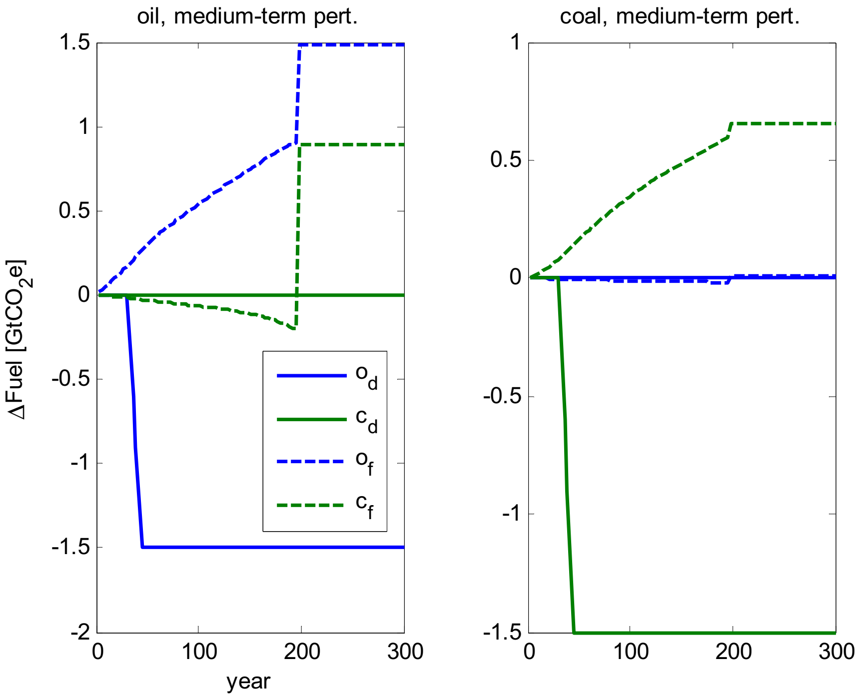

Figure A3.10: Cumulative emission reactions, backstop

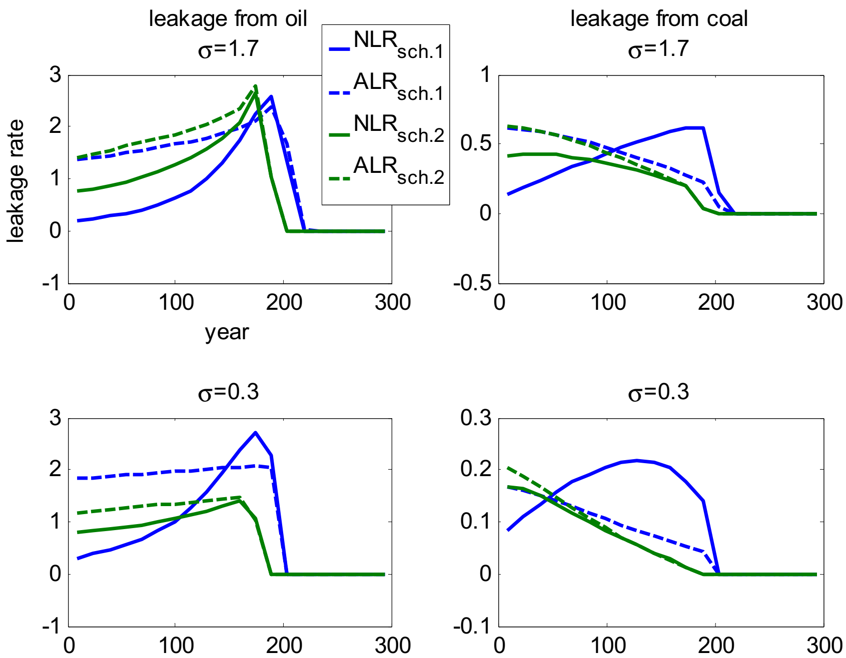

Figure A3.11: Leakage paths with backstop, alternative substitution elasticities σ

Essay 4

The Law of Small Abatements: Prices over Quantities for Realistic Climate Policies

Abstract

A fundamental and highly relevant question for climate policy design is whether price controls, such as CO_2 taxes, or quantity restrictions, such as emission quotas, should be preferred.

This paper shows that as the reach of climate policies is limited either geographically or in terms of suboptimally low reduction targets, the likelihood of price measures to be more advantageous in terms of minimizing uncertainty-related welfare losses increases. The increase in the relative advantage of the price mechanisms over quantity measures may be more than proportional to the regional limitations of policies, suggesting that even for relatively important climate coalitions, the identified factor implies a clear advantage for price measures.

Unlike previous theoretical literature addressing the question of prices versus quantities, which typically relied on the assumption of first best (i.e., global) policies, this paper examines local or regional climate policies, as corresponds to any currently realistic accord. Illustrating the main thought of the analysis, I explain why, in the example of policies corresponding to the current Kyoto mechanism, the simple theoretical weighting of the price versus the quantity approach seems to favor price mechanisms independently of the exact form of the global abatement cost and benefit curves.

Author: Florian Habermacher

Keywords: climate policy, prices versus quantities, uncertainty, greenhouse gas tax, emission quotas, abatement costs, abatement benefits, unilateral policy.

JEL classification: Q54, Q52, D81, Q40..

I mostly thank participants of the 21st International Climate Policy PhD Workshop at ETH Zurich, 22-23 October 2010 for comments and suggestions.

4.1 Introduction

For decades, economists have been weighing the advantages and drawbacks of price and quantity measures in market based policies for the regulation of public goods, such as the climate, in both theoretical as well as empirical frameworks. In theory, both measure types could be equivalent in a framework with complete knowledge, i.e., if both the cost and the benefit curve for pollution abatements were known without any uncertainty.

In his seminal contribution, using a theoretical framework with uncertainty about the costs and benefits from externality reductions, Weitzman (1974) showed four decades ago that price and quantity measures lead to different welfare losses and that it may not be clear a priori which type of measure achieves lower expected losses. In his analysis, he assumed that both possible policies could be implemented in an optimal way with respect to the available knowledge, i.e., where the planner governed over the whole economy responsible for the externality and where the planner was unrestricted in choosing the ex ante level of – depending on which of the two options is opted for – prices, i.e., a pollution tax, or quantities, i.e., a system of (tradable) pollution permits. His results suggested that if the curvature of the abatement-benefits function exceeded the curvature of its abatement-cost counterpart, a quantity measure would be preferable; otherwise the price mechanism would be preferable.

Both types of instruments discussed here are thought of as market based policies, that is, *quantity* measures refer to a fixation of the *overall* limit of pollution, with the emitting polluters being allowed to trade corresponding emission permits among each other. For a treatment of the choice between a tax and fixed administrative standards imposed on individual agents without possibility for pollution trade (both, in a deterministic and a stochastic framework), cf., e.g., Baumol and Oates (1975).

Current literature still largely considers the Weitzman result in its raw format as a theoretical rule for the decision between the two policy options in the climate change debate, even though questions about the practical operability are in addition considered as well. A major extension of Weitzman's theoretical considerations has been the recognition that greenhouse gases in the atmosphere represent a stock pollutant consisting of aggregate emissions over a long time period, being determined only to a limited extent by 'current' emissions over a short period. This fact tends to smooth the abatement benefit curve and thus to favor price mechanisms (see, e.g., Nordhaus, 2005). However, without substantial information about the exact abatement benefit and abatement cost curves, this may not be as strong a point in favor of tax schemes as the argument may imply on first sight. Policies that are seriously concerned with uncertainty-induced welfare losses may need to commit to measures that are stable over a period of several decades to prevent large economic losses. The reason for this necessity is that industry makes certain investments; for instance, power plants may run for 40 years, and they have high adaptation costs to new policies once the plants are constructed in a way that is optimized for a certain policy framework. Hepburn (2006) mentions that, while the stock argument implies taxes to be clearly favorable for short policy periods, permits could be favorable for commitments over several decades. While therewith there could eventually still be some requirement to analyze cost and benefit curvatures for specific climate policies in detail, the following shows that for realistic climate policies, such analysis may not be necessary, as the advantage of tax measures in limiting the expected uncertainty-induced welfare losses seems indeed unambiguous.

Climate policies implemented so far and those likely to be implemented in the near future systematically depart from optimal policies considered in Weitzman (1974) in at least two crucial ways. First, none of the measures implemented so far covers nearly all global greenhouse gas emissions. Second, none of the politically implemented measures is stringent enough to be considered to correspond to the ex ante conceivable optimal abatement strategies. Nonetheless, Weitzman's still frequently cited analysis has not been extended to account for these two differences. For example, Hepburn (2006) reviews the literature on the prices versus quantities question and even mentions the absence of a global regulator. He does not consider that the theoretical analysis of the uncertainty-induced welfare losses undergoes an important change when, instead of the global policy, a unilateral action of a single country or other entity is considered. The following analysis fills this gap and shows that the classic Weitzman result is fundamentally altered and that prices may generally be preferable in the cases where either (i) policies are applied only in parts of the global economy, such as is the case of unilateral or regional climate policies, or (ii) where policies result in only limited abatements compared with the optimal policy, which is likely to be the case even for future global treaties, should they ever be agreed upon. I briefly illustrate this result with the example of the Kyoto accord, showing that even without an analysis of the exact form of the cost and benefit functions, the uncertainty about abatement costs and benefits seems to favor a price mechanism. This result shows an even stronger trend when more local unilateral climate policy is considered.[120]

The present analysis is concerned with a country or region acting *unilaterally* and independently of other countries' climate strategies. For a globally coordinated introduction

[120]In an extension, Weitzman (1974) considered the case where multiple machines are producing, but he assumes the regulator to control them all, whilst the here discussed case of unilateral policy corresponds to the situation where only a single unit is covered by the policy. This notably changes the effect of the correlation of the abatement costs across the units.

of policies in the different countries or regions, the original Weitzman rule can remain the relevant one, when all parties agree on using the same instrument.

4.2 The Model and Weitzman's Result

Three characteristics of both climate benefits and economic costs induced by emission reductions are of primary importance for the theoretical decision about the optimal abatement level and whether price measures, such as a carbon tax, or quantity measures, such as a cap-and-trade scheme, may be more desirable. The first characteristic concerns marginal costs and marginal benefits from initial abatement. For a strictly positive abatement level to be desirable, the initial marginal benefits must exceed the marginal abatement costs in the laissez faire situation. The second and third elements, which are crucial for the question between price and quantity measures to be important even in a model neglecting most practical issues related to policy implementation, are the uncertainty about the marginal abatement costs and benefits and the non-linearity of the cost and benefit curves, i.e., the non-constancy of their marginal values. These characteristics appear to reflect reality. Energy is an essential primary input for all sectors of economic activity, and most of the energy consumed today depends on the combustion of fossil fuels, which results in CO_2 emissions that currently can be released into the atmosphere for free. Therefore there must almost certainly be processes for which emissions could be reduced at marginal abatement costs close to zero. On the other hand, the current intensity of global warming already seems to imply costs in various parts of the world. Thus, the first condition – that initial marginal abatement benefits exceed initial marginal abatement costs – seems to be given. Furthermore, there is substantial uncertainty both about the different climate change costs associated with greenhouse gas emissions and correspondingly about the benefits resulting from emission abatements, as well as about the economic costs for abatements. Therefore the second characteristic is clearly demonstrated as well. Finally, substantial non-linearity of climate costs and emission abatement benefits is widely acknowledged. Abatement costs cannot realistically be constant over a longer range, and it can generally be expected that cheaper abatements are chosen first and that ever more expensive measures are implemented as the policy becomes more stringent. Therefore, the abatement cost curve is likely to be convex. In fact, this trend has been confirmed by applied studies. Estimated carbon abatement cost curves show that marginal costs for different abatement levels tend to spread over a whole set of orders of magnitude with a relatively large fraction of the possible abatement levels eventually being achievable with only a small fraction of the costs for a larger abatement, leaving little doubt about the highly convex nature of any typical CO_2 emission abatement scenario. The abatement benefits, on the other hand, seem very likely to describe a concave curve, according to climate economists.

The remainder of this section sets up the model in which the uncertainty-induced welfare losses of the two policy types are compared and derives the Weitzman rule about the preferable policy choice between tax and cap-and-trade mechanisms (Eq. (4.3)).

In line with Weitzman's and subsequent contributors' approach, and to account for the three previously mentioned characteristics, we assume quadratic abatement benefit and cost functions (see, e.g., Weitzman, 1974; Pizer, 1997; Pizer, 1999), with some uncertainty about the values of the linear parameters. To point out how in the case of a policy with limited reach the theoretical decision rule departs from the original result concerning optimal policies, we start with the functions that map global emission abatement q_g into global benefits from the corresponding limitation of the climate change, $b_g(q_g)$, as well as corresponding global economic abatement costs, $c_g(q_g)$, and then explain the Weitzman result.

The cost and benefit functions are

$$c_g(q_g) = \alpha_c q_g + \frac{\beta_c}{2} q_g^2 \tag{4.1}$$

$$b_g(q_g) = \alpha_b q_g - \frac{\beta_b}{2} q_g^2, \tag{4.2}$$

where, in order for an inner solution with a well-defined strictly positive optimal abatement level to exist, and corresponding to climate economists' broad agreement, the parameters have the following properties: $\alpha_b > \alpha_c > 0$, $\beta > 0$, i.e., the abatement benefits describe a concave curve and the costs a convex curve, both initially being upward sloping, with a steeper initial slope for the benefits curve.

Corresponding to Weitzman (1974), we assume some uncertainty about the marginal abatement costs and benefits, allowing for some limited variability of the linear parameters, α_c and α_b. As Annex 1 graphically illustrates, (i) welfare losses from uncertainty only about climate benefits, i.e., from the variability of α_b, are the same for price as for quantity measures, and (ii) the expected welfare losses arising from uncertainty about α_b and about α_c are purely additive. Therefore, it is sufficient to examine the welfare losses expected from uncertainty about the abatement costs to determine which type of measure yields lower expected welfare losses.

Thus, from here on, we will simply attribute all uncertainty to the cost curve. In this case, the situation can be depicted as in Fig. 4.1. Given the marginal abatement cost curve (mc) and the marginal benefit curve (mb), the optimal abatement level q^*, or the

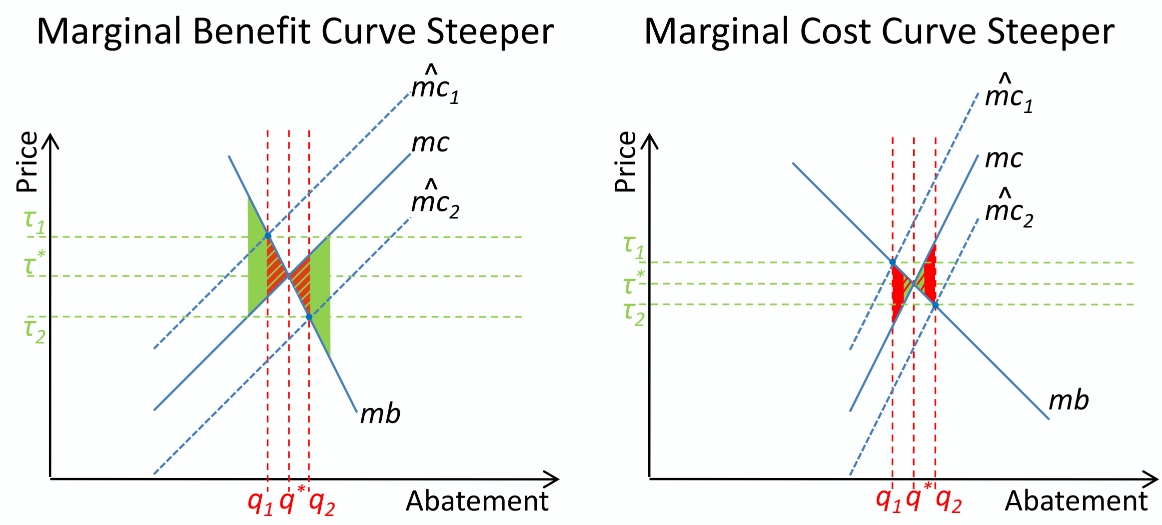

Figure 4.1: Welfare losses under uncertainty

Notes: Uncertainty-induced welfare losses for permits (red triangles) and for taxes (green triangles). Losses for permits are smaller if the marginal benefits curve is steeper (left plot), while losses for the tax are smaller if the marginal cost curve is steeper (right plot). The dashed lines represent two examples of imprecise estimates ($\hat{mc}$) of the real marginal abatement costs (mc). See also the related graph in Pizer (1997).

optimal tax level τ^*, are defined by the intersection of these two curves. However, given some uncertainty about the marginal cost curve, we may overestimate or underestimate the marginal costs ($\hat{mc}_1$ or $\hat{mc}_2$). In this case, the estimate of the optimal tax level or the emission cap will also be affected; instead of being fixed at the optimal levels, they will be determined by the intersection of the marginal benefits curve and the realized estimate of the marginal cost curve, corresponding to the green and the red curves for the estimates $\hat{mc}_1$ or $\hat{mc}_2$ in the figures. Correspondingly, instead of the optimal values, the abatements q_1 or q_2, or the taxes τ_1 or τ_2, may be chosen.

Given this imprecision, deadweight losses will accrue in the form of unrealized potential for further abatement gains (left of the optimum, given the cost estimate $\hat{mc}_1$) or excessively large abatements (right of the optimum, given the estimate $\hat{mc}_2$), corresponding to the colored triangles in the plots of Fig. 4.1 (green for the tax measure and red for the quotas). Besides depending on the deviation of the marginal cost estimates from the true value, these losses' sizes are influenced by the slopes of the curves.

When considering a tax, if the estimated marginal cost curve deviates vertically by an amount Δ_{mc} from its real position, basic geometry shows that the estimate of the optimal tax will deviate by $\Delta_t = \Delta_{mc}\frac{\beta_b}{\beta_c+\beta_b}$, where β_c and β_b indicate the absolute values of the slopes of the marginal cost and the marginal benefit curves. The corresponding deadweight loss can be characterized by the deviation of the resulting emission abatement from the optimal value. This deviation can again be determined from Fig. 4.1 using some simple geometry, and is equal to $\Delta_q^{tax} = \frac{\Delta_t}{\beta_c} = \Delta_{mc}\frac{\beta_b}{(\beta_c+\beta_b)\beta_c}$.

As for the quota, the deviation of the estimate of the optimal value from its true value can be directly deduced from the graphs in Fig. 4.1, and it is equal to $\Delta_q^{quota} = \Delta_{mc}\frac{1}{\beta_c+\beta_b}$.

Thus, given some specific deviation of the estimate of the marginal cost curve from the real marginal costs, the ratio of the implied deviation in the resulting abatement level between the case where a price measure is used and where a quantity measure is used becomes $\frac{\Delta_q^{tax}}{\Delta_q^{quota}} = \frac{\beta_b}{\beta_c}$. As can also be seen in Fig. 4.1, higher deviation of the abatement level from the optimum yields greater welfare losses, implying the following rule:

$$\left.\begin{array}{c}\text{quantity scheme}\\ \text{tax scheme}\end{array}\right\} \text{ preferred if } \quad \left|\frac{b_g''(q_g)}{c_g''(q_g)}\right| \left\{\begin{array}{l} >1 \\ <1 \end{array}\right. , \tag{4.3}$$

where $b_g''(q_g)$ and $c_g''(q_g)$ express the second derivatives of the abatement benefit and the abatement cost curves in the vicinity of the opted for abatement level q_g. Eq. (4.3) corresponds to the famous result of Weitzman (1974).

4.3 Policy with Limited Reach

Currently, all odds are against a global climate policy targeted at an approximately optimal emission level being implemented in the near future. Instead, only a few countries seem willing to make a noticeable contribution to limiting climate change, and even these countries seem to be doing so only if they are sure that the associated direct economic costs are small, preventing measures that imply broad-based and substantial carbon prices. In this section, this is taken into account by deriving what the different characteristic limitations of the climate policies have on the result in Eq. (4.3).

The first point conciders that a policy is not implemented globally, but only regionally. To illustrate the implications of this restriction, we will assume that the policy covers a specific fraction of the global economy, or of its emissions, and that Eqs. (4.1) and (4.2) represent the abatement cost and benefit curves that would be relevant if the policy were a global and optimal one. The following analysis shows how the nature of the respective curves to be considered for the regional policy can be derived from the curves for the global policy.

Global emission abatement cost curves are generally strongly convex, because of the large range of very different possibilities of substituting various types of inputs and processes or even products of an economy. The *relative* form of such curves should not primarily depend on whether the global or regional economy is considered: even a national economy is likely to offer a large range of the different types of abatement possibilities. Therefore, a global cost function with the global emission reductions as a parameter, $c_g(q_g)$, may be transformed into a regional abatement cost function for a specific region as a function of its abatement as

$$c(q) = c_g(\frac{q}{a}) \cdot a, \tag{4.4}$$

where a is $0 < a < 1$, and represents the fraction of the worldwide emissions attributed to the considered region, and q represents the regional abatement. The reasoning behind Eq. (4.4) is as follows: if the region covers a fraction a of the globe's emissions, a reduction by 10 % of its emissions will typically cost the region approximately a times as much as a reduction of 10 % of worldwide emissions would cost the whole world. Replacing the example of 10 % by a continuum of reductions and using q for absolute values rather than percentages then directly yields Eq. (4.4).[121]

The situation is different when considering the benefit. If the region of size a covers only a fraction of the global emissions, the benefit curve to be considered is not primarily scaled down by a specific constant factor, but remains the same: climate benefits from emission reductions do not depend on the place where they arise; the global benefits from a regional abatement are the same as they are from abatements taking place globally, implying

$$b(q) = b_g(q). \tag{4.5}$$

In the following, we examine the implications of this analysis for the case of different limits of climate policies.

Fully altruistic region. This section considers a region that constitutes a fraction a of the world (or of its emissions) and that implements a climate policy by fully taking into account the global emission externalities rather than only the fraction borne by its own citizens. For this case, Eqs. (4.4) and (4.5) can be used directly. Applying them to the specification of the global costs and benefits as of Eqs. (4.1) and (4.2) yields the

[121] One may note that Weitzman (1974) contains an equation similar to Eq. (4.4), for the case where a production is made by a number of similar units in parallel. The focus there, however, lies on the question whether a planner should use price or quantity measures when he can impose rules on all units simultaneously instead of, as here, on how one single unit should be treated independently of the other units.

following cost and benefits curves:

$$
\begin{aligned}
c(q) &= \left(\alpha_c\left(\frac{q}{a}\right)+\frac{\beta_c}{2}\left(\frac{q}{a}\right)^2\right)a=\alpha_c q+\frac{\beta_c}{2a}q^2\\
b(q) &= \alpha_b q-\frac{\beta_b}{2}q^2,
\end{aligned}
$$

which yields the second derivatives

$$
\begin{aligned}
c''(q) &= \frac{\beta_c}{a}\\
b''(q) &= -\beta_b.
\end{aligned}
\tag{4.6}
$$

Eq. (4.6) shows that limiting the scope of the policy to a fraction a of the planet augments the cost function's concavity by the factor $\frac{1}{a}>1$, which tends to support the price policy over the quantities policy, at least relative to the basic Weitzman result.

Region with some degree of egoism. The situation becomes even more clear if the region discounts the remainder of the world's utility, considering it only as a fraction $d<1$: in this case, while the cost curve is the same as for the altruistic region, the considered benefit function becomes $b(q)=(a+(1-a)d)\left(\alpha_b q-\frac{\beta_b}{2}q^2\right)$, and the second derivative is therefore $b''(q)=-\left(a+(1-a)d\right)\beta_b$. Recalling from the previous case that $c''(q)=\frac{\beta_c}{a}$, we see that here the ratio of the benefit over the cost curvature is decreased through the multiplication by an overall factor $a\left(a+(1-a)d\right)$, which is smaller than a and becomes as small as a^2 for a perfectly egoistic country,[122] yielding an even stronger case for a price measure rather than a quantity measure in this case.

Suboptimally low climate targets. Finally, consider the case where a policy may extend across the globe but has a limited reach for various reasons that eventually limit the ability of the society to fully account for the climate benefits when choosing the policy. This situation can readily be accounted in a similar manner to the case of a policy in a region with limited empathy for the rest of the world by assuming only a fraction e, $0<e<1$, of the climate benefits to be considered for the policy choice. In this case, the whole benefit undergoes a discounting by the factor e, yielding $b(q)=e\cdot\left(\alpha_b q-\frac{\beta_b}{2}q^2\right)$, and implying $b''(q)=-e\beta_b$. Again, the abatement cost curve remains unchanged. This type of limitation of the reach of the policy likewise favors the price scheme, in this case by augmenting the ratio by the factor $e<1$.

This factor appears quite relevant when considering realistic climate policies. Climate

[122] The latter case of the pure quadratic term may not be the most relevant as a perfectly egoistic country of a small size is unlikely to implement *any* substantial climate protection measures. Nevertheless, it can surely be expected that countries will generally discount the remaining world's benefits to a certain extent when implementing a climate policy.

protection measures that are accepted are typically so limited that it seems that only a limited fraction of the predictable climate damages are taken into account as costs. Another explanation could be that abatement costs are perceived as higher than they are according to economic analysis. In this case, the abatement costs curve may be augmented by a factor $\frac{1}{e} > 1$ instead. Dividing the costs by a factor e is, however, largely equivalent to multiplying the benefits curve by the same factor e and thus would lead to the same conclusions for the present analysis of the prices versus quantities question.

Clearly, the previous subsection has already shown that an egoistic region will opt for a suboptimally low reduction target. Thus, one might argue that the fact that realistic climate policies are not very stringent can already be explained by the regionality of the policy. However, this does not fully explain the hesitation of people or institutions to take into account all of the benefits, hypothetically assuming that once climate policies are implemented worldwide, there would most likely be plenty of room for people favoring an abatement level that is, from an economic perspective, considerably below the optimal abatement. Now, one could still argue that, while this may indeed be the case, we could still opt for policy that takes this problem into account and that the whole benefits should therefore be considered. However, the policy should be implemented according to the preferences of the deciding institutions, and it should be optimized accordingly in a positivist manner.

Policy with simultaneous limitations. Finally, the case that yields the least stringent measures favors price measures the most because of the culmination of the various points mentioned previously, and, last but not least, to the detriment of the climate, is the most realistic. Currently, climate agreements of only a regional extent may be introduced, the corresponding policies may not fully account for the externality, and they may do so even less for the part of the world outside of the mentioned region. In this case, the cost and benefit functions to be considered become

$$\begin{aligned} c(q) &= \alpha_c q + \frac{\beta_c}{2a} q^2 \\ b(q) &= (a + (1-a)d) \cdot e \cdot \left(\alpha_b q - \frac{\beta_b}{2} q^2 \right), \end{aligned}$$

and the respective second derivatives are

$$\begin{aligned} c''(q) &= \frac{\beta_c}{a} \\ b''(q) &= -\left(a + (1-a)d\right) e\beta_b. \end{aligned}$$

The overall factor by which the ratio of the *relative* curvature of the benefit curve over

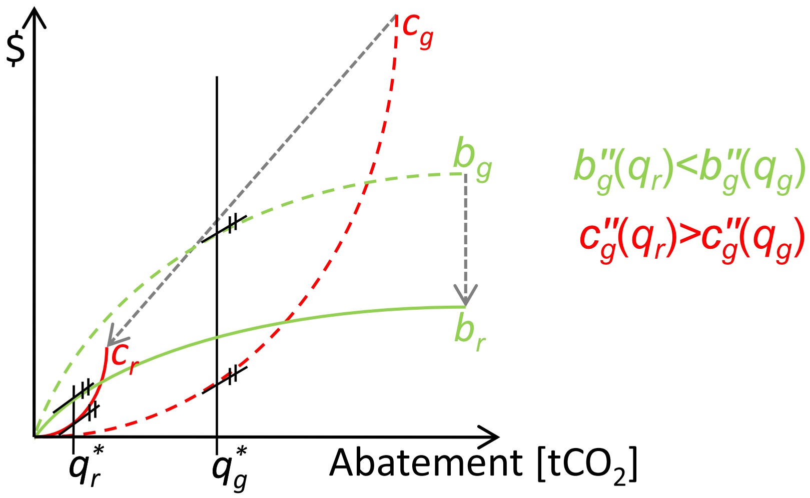

Figure 4.2: Curve transformation for regional policy

Notes: Regionality of a policy reducing the 2nd derivative of the abatement benefits curve and increasing it for the abatement cost curve. c and b stand for abatement costs and benefits, q^* for the optimal abatement level, and the indexes g and r indicate global and regional variables.

the cost curve differs from the original ratio for the globally optimal policy is then $a \cdot e \cdot (a + (1 - a)d))$, which eventually becomes a very small number as $a < 1$, $e < 1$, and $d < 1$.

We thus have that

$$\frac{b''(q)}{c''(q)} = a \cdot e \cdot (a + (1 - a)d) \frac{b''_g(q_g)}{c''_g(q_g)}, \tag{4.7}$$

which is the main result of this paper.

Fig. 4.2 illustrates the difference between the global and the limited regional policy.

4.4 Prices over Quantities for Kyoto

The example of the Kyoto protocol, which is thus far the single most important regional climate treaty, provides a good example for the illustration of the results of this study. Clearly, the way it is implemented, it is able to reduce global medium-term emissions only by a small percentage. Similarly, the effects of the protocol on global warming are very modest. Wigley (1998) projected that if sustained for the entire century, the protocol accord itself may lead to a reduction of the global temperature increase by roughly 0.1-0.2 °C, or eventually 6 % of the expected increase through 2100, and if the protocol were to be strengthened progressively throughout the century in a reasonable way, it could eventually reduce global warming by some 14 %. Aside from the question of whether these numbers might require some correction, their orders of magnitude seem fairly realistic. So what do they suggest about the curvatures of the relevant curves for climate benefits and economic costs? Clearly, even if the curvature of the climate damage costs is substantial, the *marginal* climate costs of emissions in a world where 100 % of business-as-usual warming occurred are very close to the *marginal* costs of additional units in a case with only 90 % of this business as usual warming, that is, business as usual for the year 2100 minus the mean of reductions of a Kyoto protocol that is unsustained, -6 %, or sustained, -14 %. If the climate cost curve is convex, the climate costs attributed to an additional unit of emissions may be very high, but they are unlikely to differ substantially between the two different amounts of total emissions or warming.[123] Things look different for the economic abatement costs: compared with their projected business-as-usual scenario, the emissions reductions agreed upon by the Annex I countries amounted to somewhere between 0 and 20 % for most of the countries. Clearly, the costs of marginal emission reductions of zero versus 10 % or 20 % emission reductions are likely to substantially differ as a result of the fact that the most inexpensive emission abatement measures are opted for initially, followed by the implementation of ever more expensive measures.

The relevant abatement benefit curve is thus very likely to be substantially flatter than the abatement cost curve. According to the adapted rule of Eq. (4.7) this implies that price mechanisms such as a carbon tax may have reduced the expected losses that accrue because of imprecision in the estimates of the costs and the benefits. Thus, even for a relevant policy measure such as the Kyoto accord, there seems to be little possibility for quantity measures to be theoretically as efficient as price measures. The same argument illustrates that for a unilateral policy affecting an even smaller region, such as a small or medium-sized single country, it becomes virtually certain that the uncertainty under consideration favors price measures over quantity measures.

4.5 Robustness

The above analysis is based on quadratic functions. While these functions are consistent with relevant literature and convenient to work with in the context of uncertainty-induced welfare losses, they are subject to a specific feature that ignores a certain aspect that may be relevant in reality. The constancy of these functions' second derivatives over the argument hides the fact that even for specific curves the ratio of the cost and benefit curves' second derivatives may also vary as the optimal versus the opted

[123] The study of Wigley (1998) also predicted that Kyoto may reduce business as usual sea level rise of 50 cm until 2100 by a mere 2.5 cm, strengthening the view of hardly varying marginal abatement benefits.

for abatement level changes. Here, the aforementioned change from the global to the regional policy will affect not only the cost and benefit curves but also the chosen abatement levels. Because the regional marginal abatement costs are higher than the global ones and because the considered benefits curve is either unchanged (completely altruistic region) or reduced (partially egoistic region), the total abatement level considered in the regional policy is most likely considerably lower than that of an (optimal) global policy. This endogeneity of the abatement level would be relevant if, instead of the quadratic cost and benefit functions, other forms were considered.

A natural alternative is to consider isoelastic functions for both curves, $c_g(q_g) = \gamma_c q_g^{\delta_c}$ and $b_g(q_g) = \gamma_b q_g^{\delta_b}$. Preservation of convexity of the cost curve and of the concavity of the benefits curve requires $\delta_c > 1$ and $\delta_b < 1$. Interestingly, for this functional form, the previously discussed effect of the limitation of the climate policy on regional measures vanishes. At the endogenous optimal abatement levels the ratio of the benefit and cost curves under consideration turn out to be independent of the size a of the region implementing the policy, of its altruism index d, and of the cost discounting factor e, as is shown in Annex 2. At first glance, this result may appear to limit the validity of the analysis in the previous sections. Assuming isoelastic cost and benefit curves, the theoretical answer to the question of prices versus quantities is unaffected by either of the proposed limitations of climate policies in this paper. Yet, a closer examination shows that rather than the result from the quadratic functions, the result for isoelastic functions is only an artifact based on specific function properties that are at odds with reality. With a concave isoelastic specification, the absolute magnitude of the curvature of the benefit curve rapidly grows to infinity as the abatement level approaches zero. It is this growth of the benefit curve's curvature that offsets the increase of the curvature of the cost function as the region that implements the policy becomes smaller.

In reality, however, the curvature of the benefits curve is clearly considered to be finite. To clarify, it may be considered that the emission abatement benefit curve is simply a mirror image[124] of the climate damage curve for specific emissions, normalized to zero for the business-as-usual emissions. There is no apparent reason why this damage curve should exhibit a singularity at the level of business-as-usual emissions. For the atmosphere, the amount of emissions in the business-as-usual scenario is an amount like any other potential amount as well. In any case, the expected damage curve is usually considered to be a steadily increasing function with finite first and second derivatives, i.e. the curvature of the emission damages curve – and thus the curvature of the abatement benefit curve – is bounded upwards. In the following argument, it is shown that this boundedness of the curvature of the abatement benefits strongly suggests that the main

[124] Meant is an inversion of the sign of both axes of the curve.

point of this paper holds independently of specific functional form assumptions.

We focus on the case of a regional climate policy designed in a fully altruistic way in terms of accounting for abatement benefits, i.e., the original global emission abatement curve is considered. Concerning the abatement costs of the region, its curvature is to increase proportionally as the region – assumed to contain a representative fraction of the global economy's abatement possibilities – becomes smaller. This can readily be understood as one considers the lower bound, $\underline{k_g}$, of the curvature of the global abatement cost curve, $\underline{k_g} = \min_{q_g} c_g''(q_g)$. As has been explained in the *Policy with Limited Reach* section, a positive non-zero lower bound on the curvature is likely to exist, because of the large range of different abatement possibilities, i.e., $\underline{k_g} > 0$. From the same section, we know that $c(q) = ac_g(\frac{q}{a})$. It is easy to verify that then $\underline{k} = \frac{1}{a}\underline{k_g}$, where $\underline{k} = \min_q c''(q)$ the lower bound of the regional abatement cost curvature, i.e., the (minimal) curvature of the regional abatement cost function increases proportionally as the size of the region decreases. Thus, for a small enough region the curvature of the abatement cost necessarily becomes larger than that of the benefit curve (whose value has a finite upper bound, as argued above), independently of the specific functional forms that we assume for either curve.[125]

This also suggests that the above example of the isoelastic curves failed to identify the effect of the policy's regionality on the price versus quantity outcome simply because the assumed form of the benefit curve was unrealistic, as it implied a benefit curvature that becomes huge when approaching low abatement values, rather than being bounded upwards.

4.6 Discussion

The general validity of the main result – that the uncertainty under consideration clearly favors price mechanisms over quantity measures for small economies – is also confirmed by a more intuitive argumentation without any specific mathematical support. It may be worthwhile to do so in order to prevent suspicion of whether the main result will remain if functional forms other than those considered in this study were taken into account, respectively of whether the curvature-boundedness conditions relied upon in

[125]Strictly speaking, the crucial relation of Eq. (4.4) hardly holds for incredibly tiny regions a, which in the extreme could comprise only a single production place with little scope for relevant emission abatement possibilities with marginal costs ranging over a number of orders of magnitudes. However, even if we limit our attention to the smaller among the developed countries (countries from the developed world currently appear more likely to implement substantial unilateral climate measures than others), they are large enough for their economies to contain the relevant large range of activities leading to strongly convex abatement cost curves.

the robustness analysis are misleading. Consider that global emission reductions provide some global economic benefits as opposed to some emission abatement costs. These abatement costs and benefits can also be considered functions of relative abatement efforts. If the country or region reduces its emissions by a specific fraction, this will lead to specific benefits and costs. If, from a global perspective, the world abates a small fraction of its emissions, the marginal abatement costs are likely to be very low. However, the costs may become exorbitantly high if almost all emissions are to be abated, i.e., the cost curve exhibits substantial convexity.

The same holds for a small region abating its emissions. Abating a small percentage of its emissions will induce small marginal costs, and abating almost all its own emissions will imply huge marginal abatement costs. This similarity does not hold for the benefits curve. Climate change costs are widely seen as being largely convex. However, the global abatement benefit curve, which is the mirror image of the climate change costs, is concave. If the world reduces its emissions by a small fraction, the per-unit climate benefits are expected to be large, whereas when abating almost all of its emissions, additional emission savings will induce only very limited additional benefits. The situation is different for small regions. If small enough, a region faces a virtually flat curve of (global or regional) climate benefits from its own relative abatements. If it abates only a small fraction of its emissions, the marginal climate benefits may be substantial, and then they will remain so also if almost all of its emissions are abated, with the marginal benefit in both cases primarily depending on the amount of emissions of the remainder of the world, which is largely independent of the small region. Thus, as the abatement benefits from a small region are virtually linear, the abatement benefits curve under consideration is clearly more curved than the abatement cost curve.

Hoel and Karp (2002) suggests that, as cumulative emissions matter more than instantaneously exhausted gases, $b_g(\cdot)$ may be quite linear in the relevant short run.[126] Together with the fact that abatement cost curves are strongly curved, this has already limited the likelihood of $\left|b_g''(\cdot)\right| > \left|c_g''(\cdot)\right|$. Considering in addition Eq. (4.7) implies that for any policy whose extent is seriously limited - either as it covers only a fraction of the world's emissions, or as it is substantially less ambitious than an optimal policy would have to be – the probability of $|b''(\cdot)| > |c''(\cdot)|$ seems indeed to become vanishingly small.

The benefits 'discounted' by the factors e and $a + (1 - a)d$ do not reflect the real

[126] The long run cost-benefit is relevant, but as the policy can be updated more or less regularly if new information is available, the mentioned paper studies the problem in a dynamic setting, with the result that the curvature of the abatement benefit from one policy period to the next is relevant rather than the curvature over the long run.

global long-term benefits of climate protection (which may arguably be calculated, e.g., by attributing at least nearly the same weight to future generations as to the present population). It is important to note that this fact does not alter the fundamental result in Weitzman (1974) of the optimality of a price versus a quantity measure for given cost and benefit curvatures. For example, consider a country that only takes the domestic benefits of a policy into account (implying $a < 1$ and $d = 0$) and does so by strongly discounting the future population's benefits (eventually leading to a low value for e). Policies are legitimized by policymakers by virtue of the fact that they are acting according to the population's preferences. They must implement a policy that is optimal with respect to the climate benefits as the population considers them, regardless of whether these are shortsighted or somewhat egoistic. This implies that when using the Weitzman result of Eq. (4.3), the *considered* cost and benefit curves must be taken into account. Ultimately, we arrive at our result, expressed in Eq. (4.7) for quadratic functions, and to a conclusion that continues to confirm the main result, according to our robustness analysis, if other realistic functional forms are assumed as well.

It seems quite common that individual, small or middle-sized countries unilaterally set quantitative domestic emission (abatement) targets, and it may for some seem counter-intuitive that one should, as argued here, use a tax to regulate the domestic emissions rather than a direct 'quantity' approach which (in theory) exactly allow to achieve a predefined emission reduction, such as with a cap-and-trade system. The crucial point for the here found result is that from a welfare perspective, when a country's contribution is really driven by the desire to abate optimally according to the trade-off between climate damage and the (non-green) economic welfare, either strictly within the country or by taking foreign welfare (partly) into account as well, there exists no simple economic reason why a specific reduction target should be aimed at. Instead, given that the marginal welfare benefit from emission abatements is practically unaffected by the domestic contribution, this target should depend directly on the (uncertain) economic costs that different abatement levels imply; the concern for the environmental utility should rather imply a willingness to pay *per unit* of emission avoided, as each unit of emission reduction has a (quasi-)constant impact in terms of green-only welfare. To avoid a discrepancy between the abatement costs and the willingness to pay per unit of emissions avoided, it is therefore efficient to set a tax (with the level corresponding to the per-unit willingness to pay) rather than a specific quantitative target for which it is unknown how much the 'last' unit of emission avoided costs the economy.

4.7 Conclusions

Weitzman's (1974) result provided a still often cited theoretical basis for the potential equivalence of price and quantity measures, in terms of minimizing welfare losses induced by the uncertainty about the cost and benefit curves in the area of climate policy. His result stated that the optimal choice depends on the relative size of the second derivatives of the marginal abatement cost and benefit. Intuitively, it may not seem clear which type of measure will minimize the uncertainty related to welfare losses in climate policy. Not only do emission abatement cost functions seem highly convex, but the global climate damage function also implies strongly concave abatement benefit functions. However, realistic climate policies are unlikely to cover all of the world's emissions and seem to be much less stringent than seems economically optimal. The analysis in this paper shows how these limitations of climate policies imply that the cost and benefit curves under consideration differ from the global estimates in a way that strongly suggests that price measures seem preferable for most climate measures, notably the Kyoto protocol. Surely, such measures would be preferable for any conceivable unilateral policy implemented by small- or medium-sized countries. As explained, with respect to their own potential emission abatements, small economies face virtually linear climatic abatement benefit curves in any case, implying that rather than considering a truly curved benefit function, they should exhibit a willingness to pay for avoided emissions that is independent of their own abatement level. In the case of a small economy, uncertainty-driven welfare losses are thus found to be unambiguously lower for tax measures compared to quota measures. This holds for rational small economies independently of whether their concern for the climate is driven primarily by egoistic or altruistic motives. The result is based on the assumption that the considered region has cost and benefit curves that are approximately proportional to their global counterparts, as well as that it acts alone or independently of other regions. It extends to the case where the whole world would implement policies simultaneously but independently, to the degree that the abatement cost curves of the different countries are uncorrelated to each other.

Clearly, there exist additional advantages and disadvantages attached to any specific climate policy, often important for practical implementability, as discussed, e.g., by Nordhaus (2005). He also explains why the to be considered climate benefit curve is anyway relatively flat. This is due to the fact that what matters is the stock of emissions in the atmosphere rather than the rate of exhaustion, supporting this paper's claims of the superiority of the price schemes for unilateral or otherwise limited policies.

4.8 Annex

Annex 1 Additional Graphical Illustrations

(i) Equivalence of Prices and Quantities for Uncertainty about Climate Benefits

Consider Fig. A4.1.

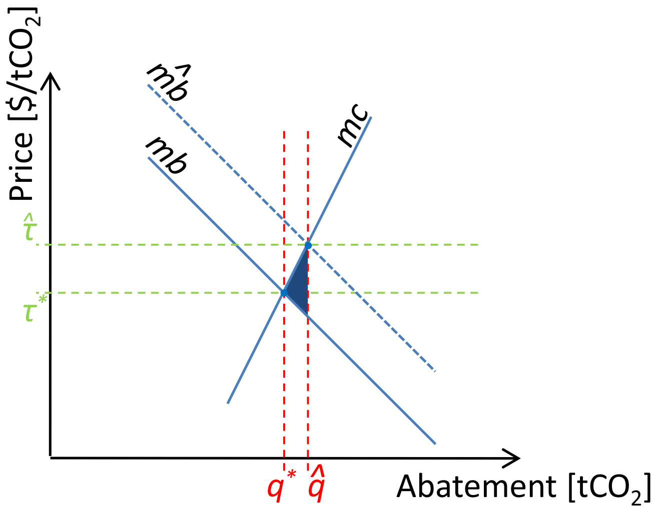

Figure A4.1: Tax or quota induced welfare loss from imprecision in abatement benefit estimation

Be the amount of tradable permits or the tax rate chosen based on an imprecise estimate $\hat{mb}$ of the marginal abatement benefits instead of the true marginal benefits curve mb. When the marginal cost curve (mc) is precisely estimated, the welfare losses (blue surface) accruing due to the imprecision in the fixation of the amount of tradable permits are the same as the losses from the imprecision in the level of the tax.

(ii) Additivity of Expected Losses from Uncertainty about Climate Benefits and from Uncertainty about Abatement Costs

The left part of Fig. A4.2 illustrates welfare losses (corresponding to the black surfaces) accruing due to some variability in the estimation of the marginal climate benefit curve. The right part of Fig. A4.2 illustrates the expected welfare loss (light green surface) accruing if the marginal cost curve is overestimated, as well as by how much this loss increases if in addition there is some variability in the estimation of the marginal benefit curve – it is easy to see that this increase corresponds to the size of the black surfaces in the right figure. Also, the geometry of the graphs implies that black surfaces in the left part of Fig. A4.2 must have the same size as the ones in the right part, indicating

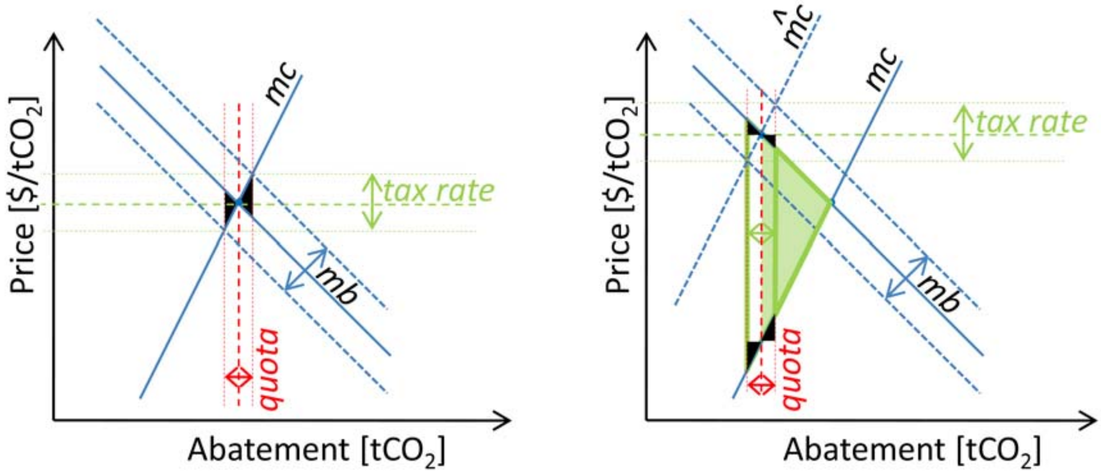

Figure A4.2: Additivity of welfare losses from uncertainty about climate benefits and welfare losses from uncertainty about abatement costs

that the welfare losses accruing under some imprecision in costs as well as of benefits simply correspond to the sum of the welfare losses accruing under imprecision only in the cost estimate plus the losses accruing under imprecision only in the benefits.

Annex 2 Constancy of Relative Curvatures for Isoelastic Cost and Benefit Functions

Consider isoelastic global abatement cost (c_g) and benefit (b_g) curves,

$$\begin{aligned} c_g(q_g) &= \gamma_c q_g^{\delta_c} \\ b_g(q_g) &= \gamma_b q_g^{\delta_b}, \end{aligned}$$

where $\delta_c > 1$ and $0 < \delta_b < 1$ and strictly positive γs would guarantee that the curves are strictly increasing, the cost curve convex and the benefit curve concave.

The second derivatives of the respective functions take the form $\delta(1-\delta)\gamma q_g^{\delta-2}$, and thus depend on the abatement level q_g - notably they become infinitely large as q_g reaches zero.

The optimal abatement level q_g^* is implicitly given by the efficiency condition

$$c'_g(q_g^*) \stackrel{!}{=} b'_g(q_g^*),$$

implying

$$q_g^* = \left(\frac{\delta_b \gamma_b}{\delta_c \gamma_c}\right)^{\frac{1}{\delta_c - \delta_b}}.$$

For this abatement level, the ratio of the second derivatives of the two functions becomes

$$\begin{aligned} \frac{b_g''(q_g^*)}{c_g''(q_g^*)} &= \frac{\delta_b(\delta_b - 1)\gamma_b}{\delta_c(\delta_c - 1)\gamma_c} \left(\left(\frac{\delta_b \gamma_b}{\delta_c \gamma_c}\right)^{\frac{1}{\delta_c - \delta_b}}\right)^{\delta_b - \delta_c} \\ &= \frac{\delta_b - 1}{\delta_c - 1}. \end{aligned}$$

The second equality shows that the curvature ratio is independent of the scale parameters γ at the optimal global abatement level q_g^* which itself does depend on these parameters γ.

It is straightforward to see that this implies that the ratio of the curvatures does not change if a regional policy is considered instead of the global policy: the transformation $c(q) = c_g(\frac{q}{a}) \cdot a$ as well as a scaling of the benefits curve $b(q) = (a + (1 - a)d) \cdot e \cdot b_g(q)$ (see section *Policy with Limited Reach* in the main text) correspond to simple changes in the parameters γ_c and γ_b, but do not affect the exponents δ_c or δ_b. As we just have shown, the relative curvatures do not depend on the values γ_c and γ_b in the region of the rationally opted for abatement level. Thus, if the costs and benefits of emission reductions would both follow isoelastic curves, the ratio of the curvatures would not depend on the regionality of the climate policies for the endogenous relevant abatement target.

References

M.R. Allen, D.J. Frame et al. (2009), Warming caused by cumulative carbon emissions towards the trillionth tonne, *Nature* 458(7242), pp. 1163 – 1166.

M.H. Babiker (2005), Climate Change Policy, Market Structure, and Carbon Leakage, *Journal of International Economics* 65(2), pp. 421 – 445.

M. Babiker and Jacoby H.D. (1999), Developing Country Effects of Kyoto-Type Emissions Restrictions, *MIT Joint Program Report Series* No. 53, October.

J.T. Bartis, F. Camm and D.S. Ortiz (2008), Producing Liquid Fuels from Coal – Prospects and Policy Issues, *Rand Corporation*, Santa Monica CA et al.

W.J. Baumol and W.E. Oates (1975), The Theory of Environmental Policy – Externalities, Public Outlays, and the Quality of Life, Prentice-Hall, New Jersey.

Ch. Beermann (2012), Unilateral Carbon Taxation in Presence of Carbon Leakage, *Working Paper*.

Ch. Beermann, D. Jus and M. Zimmer (2011), Resource Competition and the Herfindahl Rule, *Working Paper*.

T.C. Bergstrom (1982), On Capturing Oil Rents with National Excise Tax, *American Economic Review* 71, pp. 194 – 201.

T.C. Bergstrom, J.G. Cross and R.C. Porter (1981), Efficiency Inducing Taxation for a Monopolistically Supplied Depletable Resource, *Journal of Public Economics* 15, pp. 23 – 32.

BGR (Bundesanstalt für Geowissenschaften und Rohstoffe) (2009a), Energierohstoffe 2009: Reserven, Ressourcen, Verfügbarkeit, BGR, Paris.

BGR (Bundesanstalt für Geowissenschaften und Rohstoffe) (2009b), Energierohstoffe 2009: Reserven, Ressourcen, Verfügbarkeit – Tabellen, BGR, Paris.

C. Böhringer, A. Lange and T.F. Rutherford (2010), Optimal Emission Pricing in the Presence of International Spillovers: Decomposing Leakage and Terms-of-Trade Motives, *NBER Working Paper* No. 15899, Cambridge MA, April.

C. Böhringer and A. Löschel (2004), Climate Policies: Trade Spillovers, Joint Implementation and Technological Spillovers, Market Power, Investment Risks, in C.Böhringer and A.Löschel (ed.), *Climate Change Policy and Global Trade*, pp. 231 – 296.

C. Böhringer, A. Löschel and H. Welsch (2008), Environmental Taxation and Induced Structural Change in an Open Economy: the Role of Market Structure, *German Economic Review* 9(1), pp. 17 – 40.

J. Bollen, T. Manders and H. Timmer (1999), Kyoto and Carbon Leakage: Simulations in WorldScan, in Intergovernmental Panel of Climate Change (IPCC) - Working Group III (ed.), *Economic Impacts of Mitigation Measures - Proceedings of IPCC Expert Meeting on Economic Impact of Mitigation Measures*, pp. 93 – 116.

J. Brander and S. Djajic (1983), Rent-Extracting Tariffs and the Management of Exhaustible Resources, *Canadian Journal of Economics* 16(2), pp. 288 – 298.

J.A. Brander and B.J. Spencer (1984), Trade Warfare: Tariffs and Cartels, *Journal of International Economics* 16(3), pp. 227 – 242.

J.M. Burniaux (2001), International Trade and Investment Leakage Associated with Climate Change Mitigation, *Paper Presented at the Fourth Annual Conference on Global Economic Analysis*, June. See also: `https://www.gtap.agecon.purdue.edu/resources/download/503.pdf` (accessed 2011-12-4).

J.M. Burniaux, J.P. Martin and J. Oliveira-Martins (1992), The Effect of Existing Distortions in Energy Markets on the Costs of Policies to Reduce CO2 Emissions: Evidence from Green, *OECD Economic Studies*, pp. 141 – 165.

J.M. Burniaux and J. Oliveira-Martins (2000), Carbon Leakages: A General Equilibrium View, *OECD Working Paper* 15.

J.M. Burniaux and J. Oliveira-Martins (2012), Carbon Leakages: A General Equilibrium View, *Economic Theory* 49, pp. 473 – 495.

Y.-H.H. Chen, J.M. Reilly and S. Paltsev (2011), The Prospects for Coal-To-Liquid Conversion: A General Equilibrium Analysis, *MIT Joint Program Report Series* No. 197, April. See also: `http://globalchange.mit.edu/pubs/abstract.php?publication_id=2151` (accessed 2011-11-20).

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

P.S. Dasgupta and G.M. Heal (1979), Economic Theory and Exhaustible Resources.

DERA (Deutsche Rohstoffagentur) (2012), Energiestudie 2012 – Reserven, Ressourcen und Verfügbarkeit von Energierohstoffen, *DERA Rohstoffinformationen*.

C. Di Maria and E. van der Werf (2008), Carbon Leakage Revisited: Unilateral Climate Policy with Directed Technical Change, *Environmental and Resource Economics* 39(2), pp. 55 – 74.

DOE/NETL (2006), Economic Impacts of U.S. Liquid Fuel Mitigation Options, *National Energy Technology Laboratory – Department of Energy Report*, USA, July.

R. Dullieux, L. Ragot and K. Schubert (2011), Carbon tax and OPEC's rents under a ceiling constraint, *The Scandinavian Journal of Economics* 113(4), pp. 798 – 824.

Ecoplan (2012), Volkswirtschaftliche Auswirkungen einer Ökologischen Steuerreform – Analyse mit einem Berechenbaren Gleichgewichtsmodell für die Schweiz.

O. Edenhofer and M. Kalkuhl (2010), Prices vs. Quantities and the Intertemporal Dynamics of the Climate Rent, *CESifo Working Paper* No. 3044, Munich, May.

P. Egger and S. Nigai (2013), Energy Reform in Switzerland: A Quantification of Carbon Taxation and Nuclear Energy Substitution Effects, *KOF Working Papers* 327, January.

EIA (U.S. Energy Information Administration) (2013a), International Coal Data, `http://www.eia.gov/coal/data.cfm#intl` (accessed 2013-02-06).

EIA (U.S. Energy Information Administration) (2013b), Technically Recoverable Shale Oil and Shale Gas Resources: An Assessment of 137 Shale Formations in 41 Countries Outside the United States, June 2013.

T. Eichner and R. Pethig (2011), Carbon leakage, The Green Paradox, And Perfect Future Markets, *International Economic Review* 52, pp. 767 – 805.

S. Felder and T.F. Rutherford (1993), Unilateral CO2 Reductions and Carbon Leakage: The Consequences of International Trade in Oil and Basic Materials, *Environmental Economics and Management* 25(2), pp. 162 – 176.

C. Fischer and A.K. Fox (2011), Comparing Policies to Combat Emissions Leakage: Border Carbon Adjustments versus Rebates, *Ressources for the Future Discussion Paper* No. 09-02-REV.

M. Fleurbaey and S. Zuber (2013), Climate Policies Deserve a Negative Discount Rate, *Chicago Journal of International Law* 13, pp. 565 – 685.

R. Gerlagh (2011), Too Much Oil, *CESifo Economic Studies* 57(1), pp. 79 – 102.

R. Golombek, C. Hagem and M. Hoel (1995), Efficient Incomplete International Climate Agreements, *Resource and Energy Economics* 17(1), pp. 25 – 46.

M. Golosov, J. Hassler et al. (2011), Optimal Taxes on Fossil Fuel in General Equilibrium, *NBER Working Paper* No. 17348, Cambridge MA, August.

F. Habermacher (2011), Optimal Fuel-Specific Carbon Pricing and Time Dimension of Leakage, *HSG Economics Working Paper* No. 1144.

F. Habermacher (2012a), No Green Fuel-Tax Paradox on Earth?, *mimeo*, available at http://www.habermacher.net/attachments/File/No_Green_Fuel-Tax_Paradox_on_Earth_New.pdf.

F. Habermacher (2012b), Is Carbon Leakage Really Low? A Critical Reconsideration of the Leakage Concept, in L.Kreiser, A.Y.Sterling et al. (ed.), *Critical Issues in Environmental Taxation XI – Carbon Pricing, Growth and the Environment*, pp. 247 – 261, Edward Elgar.

F. Habermacher (2013), Consistent Tax on Non-Renewable Resources Revisited, *mimeo*, January.

D. Helm (2012), The Carbon Crunch, Yale University Press.

C. Hepburn (2006), Regulation by Prices, Quantities, or Both: a Review of Instrument Choice, *Oxford Review of Economic Policy* 22(2), pp. 226 – 247.

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

M. Hoel (1996), Should a Carbon Tax be Differentiated across Sectors?, *Journal of Public Economics* 59, pp. 17 – 32.

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

M. Hoel and L. Karp (2002), Taxes versus Quotas for a Stock Pollutant, *Resource and Energy Economics* 24(4), pp. 367 – 384.

H. Hotelling (1931), The Economics of Exhaustible Resources, *Journal of Political Economy.*

R.W. Howarth, R. Santoro and A. Ingraffea (2011), Methane and the Greenhouse-Gas Footprint of Natural Gas from Shale Formations, *Climatic Change* 106(4), pp. 679 – 690.

N. Hultman, S. Rebois et al. (2011), The Greenhouse Impact of Unconventional Gas for Electricity Generation, *Environmental Research Letters* 6(4).

IEA (International Energy Agency) (2005), Resources to Reserves – Oil & Gas Technologies for the Energy Markets of the Future, OECD/IEA, Paris.

IEA (International Energy Agency) (2008), World Energy Outlook 2008, OECD/IEA, Paris.

IEA (International Energy Agency) (2009), World Energy Outlook 2009, OECD/IEA, Paris.

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

IEA (International Energy Agency) (2011), Coal Information, *IEA Statistics*, OECD/IEA, Paris.

IEA (International Energy Agency) (2012), CO2 Emissions from Fuel Combustion – Highlights, *IEA Statistics*, OECD/IEA, Paris.

IPCC (2007), Climate Change 2007: Impacts, Adaptation and Vulnerability (AR4), `http://www.ipcc.ch/publications_and_data/publications_and_data_reports.shtml` (accessed 2013-03-04).

M. Jiang, W. M. Griffin et al. (2011), Life Cycle Greenhouse Gas Emissions of Marcellus Shale Gas, *Environmental Research Letters* 6(3).

L. Karp (1984), Optimality and Consistency in a Differential Game with Non-Renewable Resources, *Journal of Economic Dynamics and Control* 8, pp. 73 – 97.

L. Karp and D.M. Newbery (1993), Intertemporal Consistency Issues in Depletable Resources, in A.V.Kneese and SweeneyJ.L. (ed.), *Handbook of Natural Resource and Energy Economics* Vol. III, Elsevier.

M.C. Kemp and N.V. Long (1980), Optimal Tariffs and Exhaustible Resources, in M.C.Kemp and N.V.Long (ed.), *Exhaustible Resources, Optimality and Trade*, North-Holland, Amsterdam.

G. Kirchgässner (2001), Trade Neutrality of National Environmental Policy: Some Theoretical Considerations and Simulation Results for Switzerland, in P.J.J.Welfens (ed.), *Internationalization of the Economy and Environmental Policy Options*, pp. 125 – 152.

G. Kirchgässner, U. Müller and M. Savioz (1998), Ecological Tax Reform and Involuntary Unemployment: Simulation Results for Switzerland, *Swiss Journal of Economics and Statistics* 134(3), pp. 329 – 353.

O. Kuik and M. Hofkes (2010), Border Adjustments for European Emissions Trading – Competitivesness and Carbon Leakage, *Energy Policy* 38, pp. 1741 – 1748.

B. Lanz and S. Rausch (2011), General Equilibrium, Electricity Generation Technologies and the Cost of Carbon Abatement, *MIT Joint Program Report Series* No. 194, February. See also: `http://globalchange.mit.edu/pubs/abstract.php?publication_id=2140` (accessed 2011-11-20).

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

A.S. Manne and R.G. Richels (1991), Global CO2 Emission Reductions – the Impacts of Rising Energy Costs, *The Energy Journal* 12(1), pp. 87 – 107.

A.S. Manne and R.G. Richels (2000), The Kyoto Protocol: A Cost-Effective Strategy for Meeting Environmental Objectives?, in C.Carraro (ed.), *Efficiency and Equity of Climate Change Policy*, pp. 43 – 62.

W. McKibbin, M. Ross et al. (1999), Emission Trading, Capital Flows and the Kyoto Protocol, *Paper Presented at the IPCC Working Group III Expert Meeting*, The Hague, May.

W.J. McKibbin and P.J. Wilcoxen (2008), The Economic and Environmental Effects of Border Tax Adjustments for Climate Policy, *Brookings Trade Forum: Climate Change, Trade, and Competitiveness: Is a Collision Inevitable?* (2008/2009), Washington, June.

Th. Michielsen (2011), Brown Backstops versus the Green Paradox, *CentER Discussion Papers* No. 2011-076.

W.D. Nordhaus (2005), Life After Kyoto: Alternative Approaches to Global Warming, *Yale University Manuscript*, December.

W.D. Nordhaus (2007), A Review of the 'Stern Review on the Economics of Climate Change', *Journal of Economic Literature* 54(3), pp. 686 – 702.

W.D. Nordhaus (2008), A Question of Balance – Weighing the Options on Global Warming Policies, *Yale University Press*, New Haven & London.

OECD (2009), The Economics of Climate Change Mitigation: Policies and Options for Global Action Beyond 2012, Paris.

OECD (2012), OECD Factbook 2011-2012: Economic, Environmental and Social Statistics.

J. Oliveira-Martins (1995), Unilateral Emission Control, Energy-Intensive Industries and Carbon Leakages, *Global Warming: Economic Dimensions and Policy Responses*, OECD, Paris.

D. Pearce (1991), The Role of Carbon Taxes in Adjusting to Global Warming, *The Economic Journal* 101(2), pp. 938 – 948.

P. Perroni and T.F. Rutherford (1993), International trade in Carbon Emission Rights and Basic Materials: General Equilibrium Calculations for 2020, *Scandinavian Journal of Economics* 95(3), pp. 257 – 278.

A.C. Pigou (1920), The Economics of Welfare, New York.

W.A. Pizer (1997), Price vs. Quantities Revisited: The Case of Climate Change, *Resources for the Future Working Paper* 98(2), Washington, October.

W.A. Pizer (1999), Optimal Choice of Policy Instrument and Stringency under Uncertainty: The Case of Climate Change, *Resource and Energy Economics* 21(3-4), pp. 255 – 287.

S. Polborn (2011), The Green Paradox and Increasing World Energy Demand, *SSRN eLibrary.*

R.W.T. Pomfret (2008), Lecture Notes on International Trade Theory and Policy.

H.-H. Rogner (1997), An Assessment of World Hydrocarbon Resources, *Annual Review of Energy and the Environment* 22, pp. 217 – 262.

SASOL SYNFUELS INTERNATIONAL (2005), Unlocking the Potential Wealth of Coal – Introducing Sasol's Unique Coal-to-Liquids Technology, *Information Brochure*, see also: http://www.sasol.com/sasol_internet/downloads/CTL_Brochure_1125921891488.pdf (accessed 2011-11-13).

P. SINCLAIR (1992), High Does Nothing and Rising is Worse: Carbon Taxes Should Keep Falling to Cut Harmful Emissions, *The Manchester School* 60(1), pp. 41 – 52.

H.W. SINN (2008), Public Policies Against Global Warming: A Supply Side Approach, *International Tax and Public Finance* 15(4), pp. 360 – 394.

S. SOLOMON, G.K. PLATTNER ET AL. (2009), Irreversible Climate Change due to Carbon Dioxide Emissions, *Proceedings of the National Academy of Sciences* 106(6), pp. 1704 – 1709.

N. STERN (2007), Stern Review on the Economics of Climate Change, *HM Treasury Cabinet Office – Independent Reviews*. See also: http://www.hm-treasury.gov.uk/sternreview_index.htm (accessed 2011-12-16).

N. STERN, S. PETERS ET AL. (2006), Stern Review: The Economics of Climate Change.

TH. STERNER AND U.M. PERSSON (2008), An even Sterner Review: Introducing Relative Prices into the Discounting Debate, *Review of Environmental Economics and Policy* 2(1), pp. 61 – 76.

P. STEVENS (2012), The 'Shale Gas Revolution': Developments and Changes, *Chatham House Briefing Paper* 2012/4, August.

J. STRAND (2007), Technology Treaties and Fossil-Fuels Extraction, *The Energy Journal* 28(4), pp. 129 – 141.

R. TOL (2008), The Social Cost of Carbon: Trends, Outliers and Catastrophes, *Economics: The Open-Access, Open-Assessment E-Journal* 2.

R. VAN DER PLOEG (forthcoming), Cumulative Carbon Emissions and the Green Paradox, *Annual Review of Resource Economics*.

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 (2012a), Is there really a Green Paradox?, *Journal of Environmental Economics and Management* 64(3), pp. 342 – 363.

R. van der Ploeg and C. Withagen (2012b), Too Much Coal, Too Little Oil, *Journal of Public Economics* 96, pp. 62 – 77.

A. Venables (2011), Depletion and Development: Natural Resource Supply with Endogenous Field Opening, *OxCarre Research Paper* 62, Oxford, June.

M.L. Weitzman (1974), Prices vs. Quantities, *The Review of Economic Studies* 41(4), pp. 477 – 491.

T.M.L. Wigley (1998), The Kyoto Protocol: CO2 CH4 and Climate Implications, *Geophysical Research Letters* 25(13), pp. 2285 – 2288.

World Bank (2011), World Bank Commodity Price Data (Pink Sheet), `http://www.worldbank.org/prospects/pinksheets` (accessed 2012-03-03).

Short Curriculum Vitae

FLORIAN HABERMACHER Born March 6, 1982 in Lucerne, Switzerland

florian.habermacher@gmail.com, www.habermacher.net

Education

02/2009 – 7/2013	PhD in Economics and Finance, specialization in Economics and Econometrics, University of St. Gallen (HSG), Switzerland
10/2012 – 9/2013	Visiting Researcher, Oxford Centre for the Analysis or Resource Rich Economies (OxCarre), Dept. of Economics, University of Oxford
2/2010 – 1/2011	Swiss Program for PhD Students in Economics, Study Center Gerzensee
8/2001 – 5/2006	Masters in Environmental Science and Engineering, specialization in Industrial Ecology, ETH Lausanne, Switzerland
8/2003 – 6/2004	Exchange Year in Environmental and Civil Engineering, Indian Institute of Technology in Delhi

Professional Experience

From 9/2013	Head of Energy Systems Modeling, Aurora Energy Research, Oxford
11/2008 – 9/2012	Research Assistance to Prof. Dr. Dr. hc. G. Kirchgässner (chair economic policy) at Swiss Institute for International Economics and Applied Economic Research (SIAW)
10/2011 – 9/2012	Head of Section Environment, Energy and Transport, *foraus* – Thinktank for Swiss foreign Policy
10/2007 – 10/2008	Research Assistance Chair of Environmental Policy and Economics, ETH Zurich
Between 2004 and 2007	Research & Work at: Lab for Economics and Management of the Environment, ETH Lausanne; Lab for Atmospheric Modeling, GeorgiaTech, Atlanta; Lab Environmental Fluid Mechanics, ETH Lausanne; Regional Air Pollution Agency, Basel; Triform SA environmental engineering office, Fribourg

Selected Achievements & Affiliations

CESifo Research Network Affiliation
Swiss Association of Energy Economics Stud. Paper Award 2012, 1st Prize
Research Grant Swiss National Science Foundation, 12 month visit at University of Oxford 2013
Quantitative Section of GRE: 800/800
Certificate of Proficiency in English