Christian-Albrechts-Universität zu Kiel

Economics of using fossil fuels and tackling climate change

Inaugural-Dissertation

zur Erlangung des akademischen Grades eines Doktors

der Wirtschafts- und Sozialwissenschaften

der Wirtschafts- und Sozialwissenschaftlichen Fakultät

der Christian-Albrechts-Universität zu Kiel

vorgelegt von

Felix Dominique Meier, M.Sc.

aus Moosinning

Leipzig, 2020

Gedruckt mit der Genehmigung
der Wirtschafts- und Sozialwissenschaftlichen Fakultät
der Christian-Albrechts-Universität zu Kiel

Dekan:	Prof. Dr. Kai Carstensen
Erstbegutachtung:	Prof. Dr. Martin F. Quaas
Zweitbegutachtung:	Prof. Dr. Till Requate

Tag der Abgabe der Arbeit:	20.08.2020
Tag der mündlichen Prüfung:	25.11.2020

Acknowledgements

This doctoral thesis was written under the supervision of Prof. Martin Quaas who encouraged me to take on this project, and provided me with exceptional conditions to pursue my research. I am deeply grateful to Martin for the trust he placed in me and for his continuous support since my master studies. My research has benefited immensely from countless discussions with him. His insightful advice has been invaluable. I would also like to thank Prof. Till Requate for kindly agreeing to serve as second referee.

In the course of this project, I had the pleasure to be invited by Prof. Christian Traeger for a semester of research at the University of Oslo. I am extremely grateful to Christian for this memorable experience, and for his generous and kind support ever since. I have benefited in many ways from his help, and the numerous inspiring and fun discussions with him. I also want to express my thanks to Dr. Wilfried Rickels for many valuable comments and for giving me the opportunity to continue my research.

Several parts of this thesis are joint work. I am lucky to have such great collaborators and I am looking forward to many more exciting projects in the future. Every chapter of this thesis has gained from comments and discussions at conferences, workshops, and summer schools. I would also like to acknowledge the DFG for funding the project LEAC II which allowed me to pursue this research.

Without my wonderful colleagues from Kiel University, iDiv, and Kiel Institute the last years would have not been the same. Thank you for your feedback and comments, your moral support, and all the fun moments. Finally, a special thanks goes to my family and friends for keeping me motivated over all these years. Thank you Joelina for always believing in me!

Contents

List of Figures

List of Tables

1 Introduction

The use of natural resources has improved living conditions for many people on this planet. Consequently, the question of how natural resources should be managed over time is of great importance. Hotelling's (1931) and Herfindahl's (1967) work provide basic implications for the price development and optimal extraction of nonrenewable resources. The models are based on several simplifying assumptions including the absence of technological change (Krautkraemer, 1998). Progress in extraction technology, however, has turned out to play a decisive role. One example is the recent U.S. shale gas boom where technological innovations in horizontal drilling and hydraulic fracturing have enabled the large-scale production of natural gas from shale formations that were so far considered to be uneconomical (Joskow, 2015; Mason et al., 2015; Wang et al., 2014).

Despite the economic benefits that the extraction and combustion of fossil fuels provides (e.g. Kinnaman, 2011; Paredes et al., 2015; Weber, 2012), it has left current and future generations with one of the biggest challenges of the 21st century. The accumulation of greenhouse gases in the atmosphere causes global mean temperature to rise (IPCC, 2013), leading to potentially severe impacts on economies, ecosystems and species worldwide (Carleton and Hsiang, 2016; Hoegh-Guldberg and Bruno, 2010; Urban, 2015). Although most countries acknowledge the need for cooperation (UN, 1992), it has proven difficult to achieve stable international agreements (Nordhaus, 2015). Free-riding incentives have led to the concern that worldwide abatement efforts might not be sufficient to limit the increase in global mean temperature to well below 2°C in accordance with the Paris Agreement (Field and Mach, 2017; Lawrence et al., 2018).

Recently, alternative measures to address global warming have gathered attention. Climate engineering, or geoengineering, has been proposed as a means to counteract increasing temperatures by either intentionally interfering in the earth's

solar radiation budget or the large-scale removal of carbon dioxide from the atmosphere (Crutzen, 2006; Heutel et al., 2016; Keith and MacMartin, 2015; Klepper and Rickels, 2014; Lawrence et al., 2018; Rickels et al., 2018). The deployment of geoengineering technologies is, however, subject to many uncertainties, and potential side-effects (Field and Mach, 2017; Fuss et al., 2014; McMartin et al., 2016; Robock, 2008).

This cumulative thesis consists of three individual research articles and primarily contributes to the environmental and resource economics literature. The second chapter studies the recent U.S. shale gas boom by introducing endogenous technological progress in a Hotelling-Herfindahl model of resource extraction. The third and fourth chapters investigate the economic implications of two climate engineering measures, stratospheric aerosol injections and carbon dioxide removal, based on the recently emerging literature on analytic integrated assessment models (Gerlagh and Lsiki, 2018; Golosov et al., 2014; Traeger, 2018). In the following, I briefly introduce each chapter separately and emphasize my own contribution.

The second chapter is titled **"Booming gas – A theory of endogenous technological change in resource extraction"**. Motivated by the recent U.S. shale gas boom, the article aims to improve the understanding of endogenous technological change in models of resource extraction. We introduce optimal forward-looking technology investments in a standard Hotelling-Herfindahl model, and study the implications for price development and optimal order of resource extraction. Further, we analyze how the introduction of a carbon tax affects the timing of the shale gas boom. Results show that technology investments increase during the extraction of conventional gas, and start to decrease once production shifts towards shale gas. Consistent with current trends, the theory explains how gas prices can follow a U-shaped path. We show that the introduction of a carbon tax could drastically change the temporal patterns of shale gas extraction. The forward-looking behaviour of firms is crucial for such an effect, which does not occur in models that

treat the improvement in extraction technology as an unanticipated shock to the industry.

The second chapter is joint work with my supervisor Martin Quaas, and has been submitted to the "Journal of Environmental Economics and Management" (current status: Revise & resubmit). The paper benefited from comments by Hassan Benchekroun, the audiences at WCNRM 2020, VfS-Conference 2019, WCERE 2018, and the participants at the Green-Econ Spring School 2018 at AMSE. I have contributed substantially to this paper during all major stages of the research process, including the literature review, the development of the resource economic model, the calibration of the model to the U.S. natural gas industry, and the quantitative analysis.

The third chapter is titled **"Solar geoengineering and strategic interactions in a regional analytic climate economy"**. The article introduces the option of stratospheric sulfur injections into the Analytic Climate Economy (ACE) model from Traeger (2018). For this purpose, we derive a new class of solutions to analytic Integrated Assessment Models (IAMs) that allows us to solve an IAM with sulfur-based geoengineering and damages in closed-form, and to model realistic strategic interactions between two active regions, and a passive rest of the world. We analyze the resulting Markov game, and derive analytic solutions for the Social Cost of Carbon (SCC). The model suggests a globally optimal sulfur deployment strategy that increases linearly in the atmospheric carbon concentration, as well as a high sensitivity of geoengineering measures to potential damages. Current damage guesstimates suggest that solar geoengineering could reduce the SCC of a global social planner by 12-22%. The regional SCC can both increase or decrease depending on the heterogeneity of climate and geoengineering damages across regions.

The third chapter is joint work with Christian Traeger. The paper benefited from discussions with Martin Quaas and comments at EAERE 2019, Toulouse Conference on The Economics of Energy and Climate 2019, and the ASSA meeting 2020. I

have contributed substantially to this paper during all major stages of the research process, including the literature review, the development of the global and regional model, and the discussion of the results.

The fourth chapter is titled **"Carbon dioxide removal in a global analytic climate economy"**. The paper introduces the option of Carbon Dioxide Removal (CDR) and storage in different reservoir types into an analytic integrated assessment model of climate change. It derives the optimal level of CDR deployment, and analyzes the implications for traditional carbon dioxide mitigation incentives. I show that the introduction of CDR lowers net energy input and net emissions over the entire time path. Furthermore, CDR affects the SCC via changes in total economic output but leaves the analytic structure of the SCC unchanged. The SCC is first lower and then higher compared to a standard climate-economy model. CDR leads to higher emissions in the first years and lower emissions in the later years. The quantitative analysis implies that the effect of CDR on the SCC is minor, and that CDR is needed on top of traditional mitigation efforts.

The fourth chapter is single-authored but benefited from comments and discussions with Martin Quaas, Wilfried Rickels, Christian Traeger, and Jasper Meya.

References

Aguilera, Roberto F and Radetzki, Marian. The shale revolution: Global gas and oil markets under transformation. Mineral Economics, 26(3):75–84, 2014.

Carleton, Tamma A and Hsiang, Solomon M. Social and economic impacts of climate. Science, 353(6304), 2016.

Crutzen, Paul J. Albedo enhancement by stratospheric sulfur injections: A contribution to resolve a policy dilemma? Climatic Change, 77(3-4):211–219, 2006.

Field, Christopher B and Mach, Katharine J. Rightsizing carbon dioxide removal. Science, 356(6339):706–707, 2017.

Fuss, Sabine; Canadell, Josep G; Peters, Glen P; Tavoni, Massimo; Andrew, Robbie M; Ciais, Philippe; Jackson, Robert B; Jones, Chris D; Kraxner, Florian; Nakicenovic, Nebosja, and others. Betting on negative emissions. Nature Climate Change, 4(10):850–853, 2014.

Gerlagh, Reyer and Liski, Matti. Consistent climate policies. Journal of the European Economic Association, 16(1):1–44, 2018.

Golosov, Mikhail; Hassler, John; Krusell, Per, and Tsyvinski, Aleh. Optimal Taxes on Fossil Fuel in General Equilibrium. Econometrica, 82(1):41–88, 2014.

Herfindahl, Orris C. Depletion and economic theory. Extractive resources and taxation, pages 63–90, 1967.

Heutel, Garth; Moreno-Cruz, Juan, and Ricke, Katharine. Climate engineering economics. Annual Review of Resource Economics, 8:99–118, 2016.

Hoegh-Guldberg, Ove and Bruno, John F. The impact of climate change on the world's marine ecosystems. Science, 328(5985):1523–1528, 2010.

Hotelling, Harold. The economics of exhaustible resources. Journal of Political Economy, 39(2):137–175, 1931.

IPCC, The Intergovernmental Panel on Climate Change. Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2013.

Joskow, Paul L. The Shale Gas Revolution: Introduction. Economics of Energy and Environmental Policy, 4(1), 2015.

Keith, David W and MacMartin, Douglas G. A temporary, moderate and responsive scenario for solar geoengineering. Nature Climate Change, 5(3):201–206, 2015.

Kinnaman, Thomas C. The economic impact of shale gas extraction: A review of existing studies. Ecological Economics, 70(7):1243–1249, 2011.

Klepper, Gernot and Rickels, Wilfried. Climate engineering: Economic considerations and research challenges. Review of Environmental Economics and Policy, 8(2):270–289, 2014.

Krautkraemer, Jeffrey A. Nonrenewable resource scarcity. Journal of Economic literature, 36(4):2065–2107, 1998.

Lawrence, Mark G; Schäfer, Stefan; Muri, Helene; Scott, Vivian; Oschlies, Andreas; Vaughan, Naomi E; Boucher, Olivier; Schmidt, Hauke; Haywood, Jim, and Scheffran, Jürgen. Evaluating climate geoengineering proposals in the context of the Paris Agreement temperature goals. Nature Communications, 9(1):3734, 2018.

Mason, Charles F; Muehlenbachs, Lucija A, and Olmstead, Sheila M. The economics of shale gas development. Annu. Rev. Resour. Econ., 7(1):269–289, 2015.

MacMartin, Douglas G; Kravitz, Ben; Long, Jane CS, and Rasch, Philip J. Geoengineering with stratospheric aerosols: What do we not know after a decade of research? Earth's Future, 4(11):543–548, 2016.

Nordhaus, William. Climate clubs: Overcoming free-riding in international climate policy. American Economic Review, 105(4):1339–70, 2015.

Traeger, Christian P. ACE - Analytic Climate Economy (with Temperature and Uncertainty), 2018. URL https://ssrn.com/abstract=3307622.

Paredes, Dusan; Komarek, Timothy, and Loveridge, Scott. Income and employment effects of shale gas extraction windfalls: Evidence from the Marcellus region. Energy Economics, 47:112–120, 2015.

Rickels, Wilfried; Reith, Fabian; Keller, David; Oschlies, Andreas, and Quaas, Martin F. Integrated assessment of carbon dioxide removal. Earth's Future, 6(3):565–582, 2018.

Robock, Alan. 20 reasons why geoengineering may be a bad idea. Bulletin of the Atomic Scientists, 64(2):14–18, 2008.

UN, United Nations. Agenda 21. 1992. URL https://sustainabledevelopment.un.org/content/documents/Agenda21.pdf.

Urban, Mark C. Accelerating extinction risk from climate change. Science, 348(6234):571–573, 2015.

Wang, Qiang; Chen, Xi; Jha, Awadhesh N, and Rogers, Howard. Natural gas from shale formation–the evolution, evidences and challenges of shale gas revolution in United States. Renewable and Sustainable Energy Reviews, 30:1–28, 2014.

Weber, Jeremy G. The effects of a natural gas boom on employment and income in Colorado, Texas, and Wyoming. Energy Economics, 34(5):1580–1588, 2012.

2 Booming gas – A theory of endogenous technological change in resource extraction

Felix D. Meier[1,2] and Martin F. Quaas[2,3]

[1] *Kiel Institute for the World Economy*

[2] *German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig*

[3] *Department of Economics, Leipzig University*

Abstract: This paper introduces endogenous technological change in a Hotelling-Herfindahl model of natural resource use to study the recent developments in the U.S. natural gas industry. We consider optimal forward-looking technology investments, and study implications for the order of extraction of conventional and shale gas, and a backstop technology, and characterize the development of gas prices. We find that technology investments increase during the extraction of conventional gas. Once production shifts towards shale gas, investments decline. Consistent with current trends, our theory explains how gas prices can follow a U-shaped path. The calibrated model suggests that U.S. shale gas production continues to grow and prices continue to decrease until 2050. We show that a carbon tax could drastically change the temporal patterns of shale gas extraction. The forward-looking behaviour of firms is crucial for such an effect, which does not occur in models that treat the improvement in extraction technology as an unanticipated shock to the industry.

Keywords: shale gas, endogenous technological change, optimal order of extraction, natural gas prices, extraction costs, renewable backstop, optimal transition

We thank Hassan Benchekroun, the audiences at WCNRM 2020, VfS-Conference 2019, WCERE 2018 and the participants at the Green-Econ Workshop 2018 for valuable comments. We gratefully acknowledge the support of the German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig funded by the German Research Foundation (FZT 118).

2.1 Introduction

Resource booms have become a reoccurring phenomenon across the world, including the natural gas and oil sector (Carter et al., 2011; Jacobsen and Parker, 2016). However, the Hotelling-Herfindahl workhorse model of natural resource economics is unable to explain this phenomenon of increasing resource use (Hotelling, 1931; Herfindahl, 1967). This paper studies endogenous technological change as an important driver of resource booms. We propose a simple, novel resource-economic theory that can explain the recent developments in the U.S. natural gas market.

Within the last decade, the U.S. has experienced a major shift in its energy supply. From 2007 to 2018 the share of shale gas in total natural gas production grew from 8 to 69 percent (EIA, 2019b, 2014). At the same time, overall natural gas production increased by more than 50 percent and gas prices declined significantly (EIA, 2018b,a).

These recent developments are not due to the exploration of new reserves. The existence of abundant U.S. shale gas resources has been known for many years (Newell et al., 2019; Asche et al., 2012). In 1821, the first well was drilled in the Devonian Dunkirk Shale in Chautauqua County, New York. However, due to low permeability, the extraction from shale formations was regarded as technically difficult and not profitable. Consequently, production stayed on low levels for a long time (Wang et al., 2014).

The two main technologies that have allowed for the large scale production of shale gas in the early 21st century are hydraulic fracturing and horizontal drilling. Both technologies were developed by the natural gas and oil industry over decades. Hydraulic fracturing was first used in the 1940s but its application was quite limited. Horizontal drilling was common in the natural gas industry by the late 1970s. In the 1980s and 1990s, pioneering companies invested in the development of both techniques with the goal to make the production of shale formations profitable.

It took until the turn of the century to reduce extraction costs sufficiently by facilitating a combination of hydraulic fracturing and horizontal drilling as well as new monitoring techniques using micro-seismic data (Aguilera and Radetzki, 2014; Joskow, 2015, 2013; Mason et al., 2015; Wang et al., 2014).[1] Although costs have strongly declined, production started when average extraction costs of shale gas were substantially higher than for conventional gas (Aguilera, 2014; IEA, 2013).[2]

The observation of increasing production and decreasing prices contradicts standard resource-economic theory. The classical model of Hotelling (1931) predicts monotonically increasing prices and decreasing production over time. For the case of multiple resource types, Herfindahl (1967) established the rule that deposits should optimally be exploited in ascending order with respect to their marginal extraction costs.

In this paper, we develop a novel variant of the Hotelling-Herfindahl model by endogenizing progress in extraction technology for one type of resource (shale gas) as a result of costly investments into research and development. We study the consequences for price development and order of extraction when another resource type (conventional gas) with mature extraction technology, or a renewable backstop can be used as well. Our theory is not restricted to the natural gas industry, but can also be applied to other exhaustible resources such as oil where similar technology developments are taking place.

The above mentioned stylized facts indicate that the development of shale extraction technology required continuous investments from natural gas and oil producers over decades. Yet, empirical studies often base their identification strategy on the assumption that the shale gas boom was an exogenous and unexpected shock to the industry (e.g. Arezki et al., 2017; Wakamatsu and Aruga, 2013). Our model

[1]Shale formations contain natural gas and oil. The major technological innovations that have made production profitable are similar for both resources (Joskow, 2015).

[2]Other countries have not yet shifted to the extensive use of shale gas resources. China, for example, holds the world's largest shale gas reserves but still lacks in advanced extraction technologies (Lee and Sohn, 2014; Hu and Xu, 2013). China is expected to rise to the second largest shale gas producer by 2040 (EIA, 2017a).

puts this assumption in question, as it shows that the observed patterns of the shale gas boom are fully consistent with forward-looking behavior of firms investing in the development of extraction technology depending on expected future returns.

Our theory contributes to two strands of literature. First, we add to the literature that deals with the optimal order of resource extraction. This literature origins with Herfindahl (1967) who shows that under the assumption of constant marginal extraction costs, deposits with low costs should optimally be exploited first.

Various authors have set up more generalized models by relaxing different assumptions of Herfindahl's model (Kemp and Long, 1980; Lewis, 1982; Amigues et al., 1998; Holland, 2003; Chakravorty et al., 2008; Chakravorty and Krulce, 1994; Chakravorty et al., 2005; Gaudet et al., 2001; Gaudet and Lasserre, 2011; Gaudet and Salant, 2016). This literature has shown that Herfindahl's (1967) rule is part of a more general principle, according to which deposits should optimally be extracted in sequence of their full marginal cost, i.e. marginal extraction cost plus opportunity cost implied by the scarcity of the resource. This result has become known as the Generalized Herfindahl Principle (Gaudet et al., 2001). We find that a variant of the Generalized Herfindahl Principle also holds when taking endogenous technological change into account, but that the shadow values of both resources depend on the patterns of technological change.

Second, we contribute to the literature that tries to explain why prices of nonrenewable resources do not monotonically increase as implied by Hotelling (1931). Economic theory offers two explanations for this. For many nonrenewable resources, the discovery of additional deposits has exceeded extraction so that reserves have actually increased (Adelman, 1990). Pindyck (1978) develops a model in which he allows for the exploration of new reserves in the presence of stock dependent extraction costs. He shows that due to exploration, the price of a resource can follow a U-shaped path. As U.S. shale gas resources have been known long before extraction started, our model assumes that resource stocks are known from the beginning.

The second explanation for declining prices is technological progress and its decreasing effect on extraction costs. Slade (1982) adds exogenous technological change to a Hotelling model and predicts a U-shaped relative price curve. Farzin (1995) looks at exogenous technological change where marginal extraction costs depend on cumulative production and current extraction rate. His model also shows the possibility of a U-shaped price path. Lin and Wagner (2007) account for exogenous technological change and stock dependent extraction costs. Their model allows for a steady state solution and thus can explain constant prices over time. Rausser (1974) develops a model of endogenous technological progress via learning by doing in the extraction industry. As technological improvements come automatically with resource extraction, the implications of the model for resource prices are similar as the ones derived from the assumption of exogenous technological progress. Holland (2008) summarizes different models of U-shaped price paths and points out that modeling technological progress exogenously is insufficient as it leaves the main point of interest unexplained.

Our paper addresses this gap in the literature and offers insights on how technological progress develops over time. We show that technology investments increase during conventional gas extraction. Once production shifts completely towards shale gas, investments decline. The technology stock evolves in a S-shaped fashion, similar to the literature on technology diffusion (Davies, 1978; Jaffe and Stavins, 1994; Helm and Mier, 2019). Unlike Hotelling (1931), but consistent with current trends, our theory explains how gas prices can follow a U-shaped path. Further, we find that Herfindahl's least-cost-first principle does no longer apply. Endogenous technological progress allows for additional orders of resource extraction. In particular, even if firms have already shifted to the use of the renewable backstop, production can switch back to a nonrenewable resource for a period of time, when extraction technology has sufficiently advanced. Our calibrated model suggests that U.S. shale gas production continues to grow and prices continue to decrease until

2050. We further find that the introduction of a carbon tax could have had a strong postponing effect on the U.S. shale gas boom. This time-delaying effect critically depends on the forward-looking behaviour of producers, and does not occur in models that threat the improvement in shale extraction technology as an exogenous and unanticipated shock to the industry. For policy analysis it thus makes a big difference whether technical progress is exogenous or endogenous.

We proceed as follows. In the next section, we include endogenous technological change in a standard Hotelling-Herfindahl model. Section 2.3 presents our propositions on technological progress, order of resource extraction, and price development. In section 2.4, we calibrate the model to the U.S. natural gas industry. Section 2.5 considers two variations of the calibrated baseline model. The last section concludes.

2.2 Resource economic model

We consider the extraction of a homogeneous resource (natural gas) from two types of deposits, S (shale gas) and C (conventional gas). The corresponding extraction quantities are denoted by $q^s(t) \geq 0$ and $q^c(t) \geq 0$ at any point in time, t. In the following, we omit the time dependency unless needed for clarity. Stocks S and C of the two resources thus evolve according to

$$\dot{S} = -q^s, \quad \text{with } S(0) > 0 \text{ given}, \tag{2.1}$$

$$\dot{C} = -q^c, \quad \text{with } C(0) > 0 \text{ given}. \tag{2.2}$$

Alternatively, consumers can use a renewable backstop at quantity $q^b(t) \geq 0$. All three are perfect substitutes, such that gross consumer benefit can be written as

$$U(q^s + q^c + q^b) = \int_0^{q^s+q^c+q^b} P(j)\, dj, \tag{2.3}$$

where $P(j)$ is the inverse demand for the resource with $P'(j) < 0$. Marginal cost for the backstop are constant and denoted by $k^b > 0$. Extraction cost for deposit type C, $K^c(q^c)$, are increasing and weakly convex in q^c. The extraction technology for deposit type C is assumed to be mature and fixed. Extraction cost for deposit type S, by contrast, depend on the state of technology, $Z(t)$, and are denoted by $K^s(q^s, Z)$. Marginal extraction costs are positive and non-decreasing, $K^s_q > 0$, and $K^s_{qq} \geq 0$. Cost and marginal cost are decreasing with the state of technology, $K^s_Z < 0$, and $K^s_{qZ} < 0$. Extraction cost and the marginal cost reduction from an increase in the technology stock are zero, if there is no extraction from deposit type S, $K^s(0, Z) = 0$, and $K^s_Z(0, Z) = 0$.

The state Z of the extraction technology for deposit type S can be improved by technology investments, $w(t) \geq 0$. Choosing units of measurement, the technology stock evolves according to

$$\dot{Z} = w, \quad \text{with } Z(0) \geq 0 \text{ given.} \tag{2.4}$$

The cost of technology investments is given by $L(w)$ with positive and increasing marginal cost, $L'(w) > 0$, and $L''(w) > 0$.

Competitive firms are assumed to maximize the present value of revenues minus extraction and investment cost,

$$\max_{\{q^s, q^c, q^b, w\}} \int_0^\infty \left[p\,(q^s + q^c + q^b) - K^s(q^s, Z) - K^c(q^c) - k^b\,q^b - L(w)\right]\, e^{-\delta\,t}\,dt \tag{2.5}$$

subject to (2.1), (2.2), and (2.4), and non-negativity constraints for all variables. The discount rate is denoted by δ.

As in Pindyck (1984), market equilibrium requires that $P(q^s(t)+q^c(t)+q^b(t)) = p(t)$ at each point in time. To guarantee an interior solution to the market equilibrium conditions, we adopt the standard assumption that the difference between consumer benefit and costs, $W = U(q^s+q^c+q^b) - K^s(q^s, Z) - K^c(q^c) - k^b\,q^b - L(w)$,

is weakly concave in its arguments. In addition to the above-stated assumptions on cost functions the weak concavity of W in q^s and Z requires

$$\left(P'(q^s + q^c + q^b) - K^s_{qq}(q^s, Z)\right)\, K^s_{ZZ}(q^s, Z) + K^s_{qZ}(q^s, Z)^2 \leq 0. \tag{2.6}$$

Note that Condition 2.6 is always fulfilled for standard specifications of the extraction cost function, e.g. for the specification $K^s(q^s, Z) = \left(k_0 + \frac{k^s}{Z}\, q^s\right)\, q^s$.

2.3 Theoretical results

The shadow prices of the resource stocks are denoted by $\gamma(t)$ for the shale gas deposits and by $\lambda(t)$ for the conventional gas deposits, and capture the current value of an additional unit of resource in situ. The shadow price of the technology stock is denoted by $\phi(t)$ and captures the current value of technological progress in the shale gas industry.

With this notation, the optimality conditions for resource production and technology investment are given by

$$P(q^s + q^c + q^b) = p \leq K^s_q(q^s, Z) + \gamma, \tag{2.7a}$$

$$P(q^s + q^c + q^b) = p \leq K^c_q(q^c) + \lambda, \tag{2.7b}$$

$$P(q^s + q^c + q^b) = p \leq k^b, \tag{2.7c}$$

$$-L'(w) + \phi \leq 0, \tag{2.7d}$$

where conditions hold with equality whenever the corresponding variable is positive and with inequality if the non-negativity constraint is binding. In (2.7) we have already inserted the market equilibrium condition that inverse demand $P(\cdot)$ is equal to the supply price p of the resource.

Equations (2.7a), (2.7b), and (2.7c) state that the resource price should equal the full marginal cost of the resource whenever a positive amount is used. The full

marginal cost of each deposit type is characterized by marginal extraction cost plus opportunity cost associated with the scarcity of the resource stock. Following the literature, we define the sum of these costs as augmented marginal costs ($AMC^s(t)$ for deposit type S and $AMC^c(t)$ for deposit type C). Equation (2.7d) shows that investment in extraction technology is determined by the condition that marginal cost of technology development should equal the shadow price of technology.

The dynamic optimality conditions require that the shadow prices γ and λ of the two deposit types increase exponentially at the discount rate (Hotelling, 1931). Using $\gamma(0) = \gamma_0$ and $\lambda(0) = \lambda_0$ to denote the – endogeneous – initial values of the two shadow prices, the shadow prices at time t thus are

$$\gamma(t) = \gamma_0 \, e^{\delta \, t}, \tag{2.8a}$$

$$\lambda(t) = \lambda_0 \, e^{\delta \, t}. \tag{2.8b}$$

Extraction from a resource stock ends when either the deposit type has been exhausted or its shadow value has turned to zero. As shadow prices are always positive, both stocks will ultimately be exhausted.

The optimal development for the shadow price of technology is determined by

$$\dot{\phi} = \delta \, \phi + K_Z^s(q^s, Z), \tag{2.8c}$$

with transversality condition

$$\lim_{t \to \infty} e^{-\delta \, t} \, \phi(t) \, Z(t) = 0. \tag{2.8d}$$

Equation (2.8c) shows very different dynamics for the shadow price of technology for phases without extraction from deposit type S, i.e. when $q^s(t) = 0$ and thus $K_Z^s(0, Z) = \ 0$, and during phases with $q^s(t) > 0$ and thus $K_Z^s(q^s, Z) < 0$.

2.3.1 Technological progress

The dynamics of technological progress in the shale gas industry depend on how the marginal benefit of technology, $-K_Z^s(q^s, Z)$, develops over time. By assumption, it is equal to zero if there is no resource extraction from deposit type S. If there is resource extraction exclusively from deposit type S, it is decreasing over time.

Lemma 2.1. *The marginal benefit of technology, $-K_Z^s(q^s, Z)$, monotonically decreases over time during a period of resource extraction exclusively from deposit type S.*

Proof. See Appendix A.1.1. □

The decline in the marginal benefit of technology results from a combination of two effects: First, the technology stock may continue to grow if investment stays positive during resource extraction, and this decreases the marginal benefit of further technology improvements. Second, resource extraction may increase or decrease over time, which tends to increase or decrease the marginal benefit of technology. Due to discounting and the concavity of the objective function, the net effect is towards decreasing marginal benefit of technology, under optimal resource extraction, cf. Appendix A.1.1.

Lemma 2.1 allows us to characterize the dynamics of technological progress in the period without and in the period with resource extraction exclusively from deposit type S.

Proposition 2.1. *Investment in extraction technology (weakly) monotonically increases during a period without extraction from deposit type S, and (weakly) monotonically decreases during a period with extraction exclusively from deposit type S.*

Proof. See Appendix A.1.2. □

The proof of Proposition 2.1 utilizes the connection between optimal investment and the shadow price of technology ϕ given by (2.7d). If this condition holds with

equality, we can differentiate it to obtain $\frac{w\,L''(w)}{L'(w)}\,\frac{\dot{w}}{w} = \frac{\dot{\phi}}{\phi}$. As $L''(w) > 0$, there is a monotonic relationship between w and ϕ. The growth rate of w, multiplied by the elasticity of marginal investment cost, equals the growth rate of the shadow price of technology. Accordingly, the statements in Proposition 2.1 hold with strict monotonicity during a period where investment is positive.

During the period without shale gas extraction, (2.8c) implies $\dot{\phi} = \delta\,\phi$. The shadow price of technology increases over time at the rate of interest – provided it is positive to begin with. Due to the monotonic relationship between w and ϕ, investments increase over time as well and the technology stock grows in a convex fashion, $\ddot{Z} = \dot{w} > 0$. There are no revenues in the shale gas industry during this period. However, producers already have an investment incentive since the present value of investment is positive, $\phi\,w - L(w) = L'(w)\,w - L(w) > 0$.

In the period in which shale gas is the exclusive source of production, investment declines and the technology function grows in a concave fashion, $\ddot{Z} = \dot{w} < 0$. As the shale gas stock approaches exhaustion, investment goes to zero since technological progress becomes worthless.

Over the two periods, before shale gas is used and after deposit type S is used exclusively, the technology stock develops over time in an S-shaped fashion, first convex and then concave. This is a pattern familiar from the literature on technology diffusion (Davies, 1978; Jaffe and Stavins, 1994; Helm and Mier, 2019).

2.3.2 Order of resource extraction

To obtain clear-cut statements about the temporal order of resource use, we assume constant marginal costs of extraction for both resources,

$$\begin{aligned} K^c(q^c) &= k^c\,q^c, \\ K^s(q^s, Z) &= k(Z)\,q^s. \end{aligned} \tag{2.9}$$

Under this assumption, only one type of resource will be used at each point in time, as shown below. We return to the case of increasing marginal extraction costs in the quantitative analysis of a calibrated version of the model in Section 2.4. The general results on the order of resource extraction hold for the calibrated version as well, but are blurred by phases of simultaneous resource extraction.

We focus on the relevant case in which unit extraction cost for both nonrenewable resources (may) be low enough to make them competitive to the backstop technology, $k^c < k^b$ and there exists some $\underline{Z} \geq 0$ such that $k(Z) < k^b$ for all $Z \geq \underline{Z}$.

We further assume a linear marginal investment cost function, i.e. that the elasticity of marginal investment cost is equal to one,

$$\frac{w\, L''(w)}{L'(w)} = 1. \tag{2.10}$$

This assumption implies that the growth rate of investment is equal to the growth rate of the shadow price of technology.[3]

The optimality conditions (2.7a), (2.7b), and (2.7c), with assumption (2.9), imply that resources are used according to the Gerneralized Herfindahl Principle. Since the resources are perfect substitutes, and marginal costs are constant by assumption (2.9), optimality requires that only one resource will be used at a time. At each point in time, this is the resource with the smallest augmented marginal cost. For the backstop, the augmented marginal cost are equal to the constant marginal cost of production, $AMC^b = k^b$. For shale and conventional gas, augmented marginal cost is equal to the marginal cost of production plus the shadow price of the resource stock, $AMC^s = k(Z) + \gamma$ and $AMC^c = k^c + \lambda$. Thus, the decision whether a resource type is produced or not is not only governed by the change in shadow values over time but also by the dynamics of the technology stock.

[3]We assume that marginal investment cost is sufficiently small so that deposit type S will be used. This is for example the case, when marginal cost is zero as investments are zero, $L'(0) = 0$. Here, optimality condition (2.7d) always holds with equality since the shadow price of technology is non-negative.

Proposition 2.2. *The only possible orders of resource extraction are the following:*

(CSB) *deposit type C → deposit type S → Backstop*

(CBSB) *deposit type C → Backstop → deposit type S → Backstop*

(CSCB) *deposit type C → deposit type S → deposit type C → Backstop*

(SCB) *deposit type S → deposit type C → Backstop*

Proof. See Appendix A.1.3. □

In Herfindahl (1967) producers first extract the deposit type with the highest rent. In our model, this rule does not hold. Producers might start with the lower rent deposit type since the rent from the higher rent deposit type cannot yet be realized.

In Appendix A.1.3 we show that extraction order SCB from Proposition 2.2 can only result, if $\lambda_0 < \gamma_0$. Since resource extraction follows the Generalized Herfindahl Principle, SCB implies that the initial technology stock is high and $k(Z(0)) < k^c$. As this seems not to be a sensible assumption for shale gas, in the following we focus on the first three extraction orders.

2.3.3 Price development

Investment decisions not only affect the optimal order of extraction but also price development. Optimality requires that the resource price equals augmented marginal costs for the type of deposit extracted. Hence, during the extraction of deposit type C price monotonically increases over time as in Hotelling (1931),

$$\dot{p} = \delta \ (p - k^c) > 0 \quad \text{whenever } q^c > 0. \tag{2.11}$$

During periods where the backstop is used, price stays constant, $p = k^b$. In periods where deposit type S is produced, price behavior depends on the technology stock,

and develops according to

$$\dot{p} = \delta\ (p - k(Z)) + k'(Z)\, w \quad \text{whenever } q^s > 0. \tag{2.12}$$

There are two opposite effects on price development during shale gas extraction. The first term represents the positive effect from the scarcity of the resource, just as it is the case for extraction from deposit type C. The second term shows the effect that technological progress reduces marginal extraction costs over time. This effect may dominate the effect of increasing scarcity. Endogenous technological progress in resource extraction may thus explain decreasing resource prices, and increasing resource extraction, especially if the discount rate is low enough. We formally state this result in the following proposition.

Proposition 2.3. *(i) After a transition from extraction of deposit type C to extraction of deposit type S, the resource price decreases if the discount rate is small enough. (ii) After a transition from backstop to extraction of deposit type S, the resource price decreases, irrespective of the discount rate.*

Proof. (i) At the switch from C to S, $p - k(Z) < k^b$. This is because price can never be higher than the marginal cost of the backstop. In the limit of zero discounting, from (2.12) it follows

$$\lim_{\delta \to 0} \dot{p} < \lim_{\delta \to 0} \left(\delta\, k^b + k'(Z)\, w\right) = \lim_{\delta \to 0} k'(Z)\, w \leq 0. \tag{2.13}$$

By continuity, there exists an interval of positive discount rates where still $\dot{p} < 0$ at the switch point between conventional and shale gas.

(ii) At the transition from backstop to deposit type S, price must be decreasing, as otherwise it could not be optimal to switch away from the backstop. □

Denoting by $\underline{T}^s$ the time when extraction from deposit type S starts, by $\overline{T}^s$ the time when it ends, and by T^b the time of the switch from C to B, we can summarize

the price dynamics for the first three extraction orders from Proposition 2.2 as follows:

$$CSB: \qquad p(t) = \begin{cases} AMC^c(t) = \lambda_0\, e^{\delta\, t} + k^c, & t \in [0, \underline{T}^s] \\ AMC^s(t) = \gamma_0\, e^{\delta\, t} + k(Z), & t \in [\underline{T}^s, \overline{T}^s] \\ AMC^b = k^b, & t \in [\overline{T}^s, \infty] \end{cases} \tag{2.14a}$$

$$CBSB: \qquad p(t) = \begin{cases} AMC^c(t) = \lambda_0\, e^{\delta\, t} + k^c, & t \in [0, T^b] \\ AMC^b = k^b, & t \in [T^b, \underline{T}^s] \\ AMC^s(t) = \gamma_0\, e^{\delta\, t} + k(Z), & t \in [\underline{T}^s, \overline{T}^s] \\ AMC^b = k^b, & t \in [\overline{T}^s, \infty] \end{cases} \tag{2.14b}$$

$$CSCB: \qquad p(t) = \begin{cases} AMC^c(t) = \lambda_0\, e^{\delta\, t} + k^c, & t \in [0, \underline{T}^s] \\ AMC^s(t) = \gamma_0\, e^{\delta\, t} + k(Z), & t \in [\underline{T}^s, \overline{T}^s] \\ AMC^c(t) = \lambda_0\, e^{\delta\, t} + k^c, & t \in [\overline{T}^s, T^b] \\ AMC^b = k^b, & t \in [T^b, \infty]. \end{cases} \tag{2.14c}$$

During the extraction of deposit type C, price equals augmented marginal costs of conventional gas and thus monotonically increases over time. Note that the initial shadow values of deposit types C and S are endogenous and in general different for extraction orders CSB, $CBSB$ and $CSCB$. During the use of deposit type C, technology investments increase over time and shale extraction technology is build up at an increasing pace. In extraction order CSB, producers switch directly to shale gas after conventional gas is depleted, and price follows augmented marginal costs of shale gas (see 2.14a). In extraction order $CBSB$, the conventional deposit type is depleted before shale gas becomes competitive and the market switches to the renewable first. During the period in which the backstop is used, price stays constant. Investments continue to increase and extraction costs of shale gas continue

to decline. As augmented marginal costs become small enough, production finally switches to shale gas (see 2.14b). In extraction order $CSCB$, the conventional deposit type is not depleted as shale gas becomes competitive. Producers switch back to deposit type C after deposit type S is used up (see 2.14c).

During the extraction of shale gas, production can increase and price can decrease for a period of time, especially if the discount rate is low enough (see Proposition 2.3). Eventually, price must rise again since technological progress becomes worthless as investments go to zero and deposit type S approaches exhaustion. In the end, resource scarcity dominates the price behavior, giving rise to a U-shaped price path (for a schematic illustration see Figure 2.1). Note that in extraction order $CBSB$, resource price will always follow a U-shaped path during shale gas production as otherwise the market would not switch away from the backstop. In the end, as all nonrenewable resources are used up, production switches (again) to the renewable and price stays constant at the marginal cost of the backstop.

2.4 Calibration to the U.S. gas industry

To verify that the rather simple model developed here is able to reproduce patterns of actual resource use, we specify and calibrate the model for the U.S. natural gas industry. The transition from conventional to shale gas in the U.S. shows a period of simultaneous extraction. To capture this, we allow for increasing marginal extraction costs for both types of deposits. For conventional gas we assume an extraction cost function of the form

$$K^c(q^c) = \left(k_0 + k^c \cdot (q^c)^\beta\right) \cdot q^c, \tag{2.15}$$

with (positive) parameters k_0, k^c, and β to be calibrated. We use the same type of cost function for the extraction of shale gas, but the non-constant marginal cost

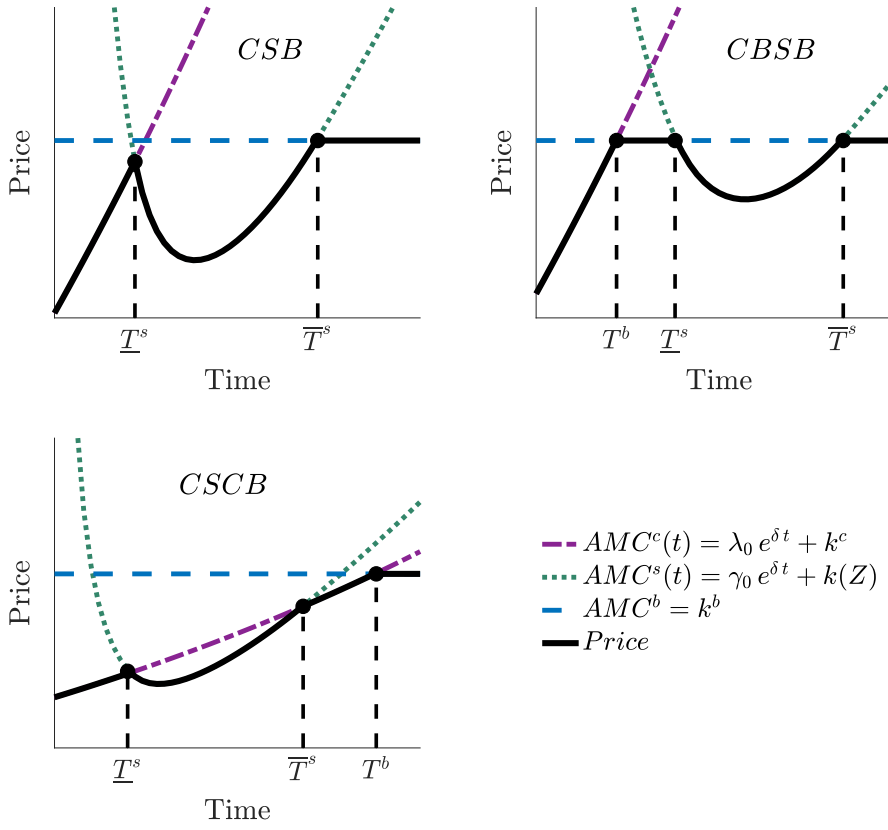

Figure 2.1: Price and augmented marginal cost dynamics for extraction orders CSB (top left graph), $CBSB$ (top right graph) and $CSCB$ (bottom left graph). Price (in black) always follows the deposit with the smallest augmented marginal costs (shown in different colours). The graph assumes that the discount rate is sufficiently small (see Proposition 2.3).

term depends on the state of technology, i.e.

$$K^s(Z, q^s) = \left(k_0 + \frac{k^s}{Z} \cdot (q^s)^\beta\right) \cdot q^s, \tag{2.16}$$

where k^s is a further parameter to be calibrated. This calibration assures that the marginal extraction cost for shale gas will not be below the lower bound of extraction cost for conventional gas. This seems a sensible assumption for shale and conventional gas, since shale gas extraction requires the extra effort of horizontal drilling and hydraulic fracturing which comes at a cost even with the most advanced technology.

The marginal cost of the backstop is constant and given by k^b. For technology investments we assume a quadratic cost function with cost parameter $l > 0$ to be calibrated,

$$L(w) = \frac{l}{2} w^2. \tag{2.17}$$

We assume a linear inverse demand function:

$$P(q^s + q^c + q^b) = d_0 - d_1 \left(q^s + q^c + q^b\right), \tag{2.18}$$

where d_0, with $d_0 > k^b > 0$, is a choke price for natural gas use, and $d_1 > 0$ is the slope of the inverse demand function.

We calibrate this model to yearly production data of shale and conventional gas in the U.S. natural gas industry from 1997 to 2017. The data is publicly available on the EIA (2017b) website. Our calibration procedure is to specify some starting parameter set, and numerically optimize investment in shale gas extraction technology, and optimal extraction of both types of natural gas. The dynamic optimization model is numerically implemented in AMPL with Knitro, which implements state of the art interior-point and active-set algorithms for large-scale nonlinear programming problems (Byrd et al., 1999, 2006).[4] We then adjust the parameter set to minimize the distance of extraction quantities derived from the model and the data. The resulting calibration of the model's parameter values is given in Table 2.1. Note that the values for the cost parameters k_0, k^c, k^s, k^b are given in billion USD, and the initial shale and conventional gas stocks $S(1997)$ and $C(1997)$ in trillion cubic feet (Tcf).

Overall, the calibrated parameter values seem to be in a plausible range. The calibration suggests that the overall reserves of shale gas, at the beginning of the model period in 1997, have been considerably larger than the reserves of conventional gas. To capture the rather strong effect of increased gas supply in a model

[4]Programming codes are available from the authors upon request.

that does not include any exogenous influences on the gas price, we obtain a rather large value for d_1, the slope of the inverse demand function. Also the discount rate is rather high, although not outrageous for an industry that is operating under various risks. In the calibration, the rather high discount rate allows to reproduce the fast dynamics of transition from conventional to shale gas.

Table 2.1: Calibrated parameter values

k_0	k^c	k^s	β	k^b	l	d_0	d_1	δ	$S(1997)$	$C(1997)$	$Z(1997)$
0.3	3	30	0.3	8	55	15	0.4	0.1	1700	340	0.1

Figure 2.2 shows the output of the calibrated model, along with the extraction data on which the model is calibrated. The calibrated model reproduces the observed decrease in conventional gas extraction and the strong increase in shale gas extraction after 2007. Due to the strong increase in shale gas extraction, the gas price decreases, also in line with observations. Our calibrated model suggests that the decrease in the gas price will continue for the next three decades, before the price starts to monotonically increase up to the marginal cost of the backstop.

2.5 Quantitative analysis

In this section, we use the calibrated model for three quantiative analyses. First, we compare the outcomes of the model with endogenous, forward-looking technological change to an alternative model specification where the improvement in extraction technology comes as an unanticipated shock, as often assumed in the empirical literature. Second, we use the model with endogenous technological change to study how the extraction of natural gas would have been influenced by the introduction of a carbon tax. Last, we also include a carbon tax in the alternative model specification and compare the results of both model types.

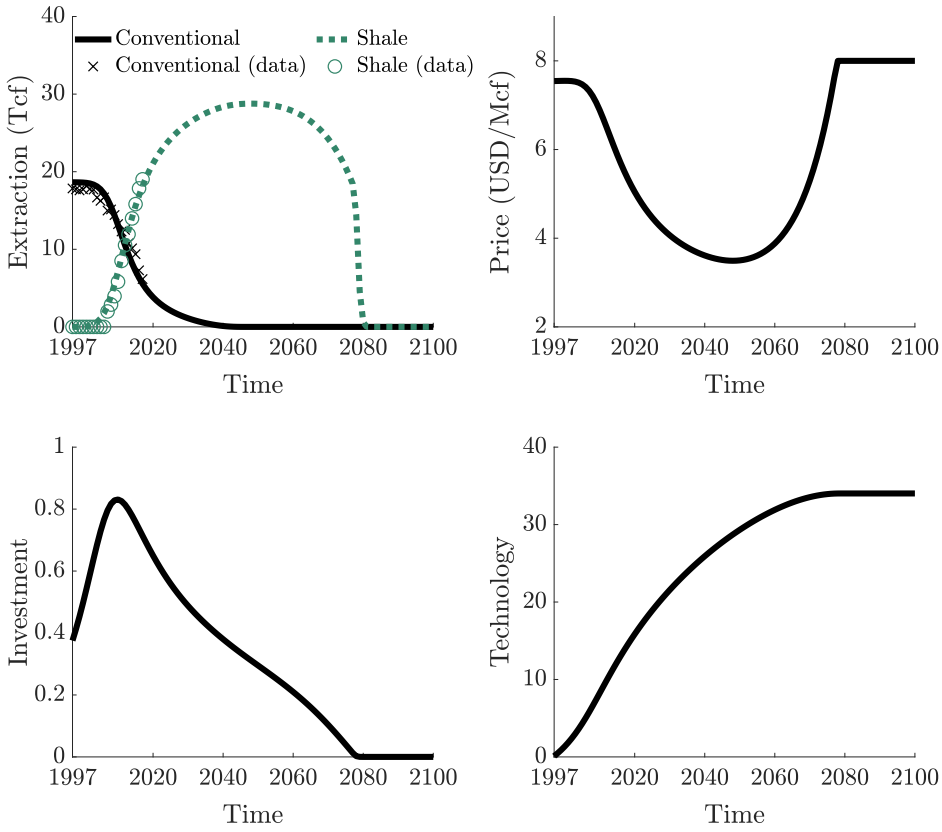

Figure 2.2: Development of conventional and shale gas extraction (top left graph), gas price (top right), investment into shale gas extraction technology (bottom left) and technology stock (bottom right) for the calibrated model. The crosses and dots in the top left graph show the EIA (2017b) data on which the model is calibrated.

2.5.1 Endogenous technological change vs. technology shock

We now contrast the outcome of the calibrated model presented in the previous section to the outcome under the alternative assumption that the improvement in shale extraction technology (and the decrease in extraction costs) comes from an exogenous and unanticipated shock rather than from the forward-looking behaviour of firms that invest in research and development to maximize profits. We assume that in 2007 technology improved from the initial to the final level of the technology

stock in our calibrated model with endogenous technological change:

$$Z(t) = \begin{cases} Z(0) \ = 0.1, & t \in [1997, 2006], \\ Z(\overline{T}^s) = 34, & t \in [2007, \infty]. \end{cases} \tag{2.19}$$

Figure 2.3 shows how natural gas extraction and price dynamics change due to this modification. Whereas at first glance both models seem to generate similar patterns of resource extraction, there are pronounced differences, clearly visible especially in the price development. In the model with exogenous technology shock, the price is continuously increasing, except for the discontinuous drop in the year where the technology shock hits. In particular, the price path follows the Hotelling rule of exponentially increasing prices before shale gas becomes available, and again after shale gas extraction starts. Moreover, shale gas extraction continuously decreases after the technology shock. All these patterns are inconsistent with observations. By contrast, the model with endogenous, forward-looking technological change shows a continuous increase of shale gas extraction, before it peaks a few decades after shale gas extraction started. Accordingly, the price continuously decreases for a period of several years, and only eventually starts to increase again.

2.5.2 Effects of a carbon tax on natural gas extraction

We now turn back to our model with endogenous technological change and include a carbon tax (τ) for producers. Assuming a carbon content of natural gas of 117 pounds of CO_2 per million British thermal units (EIA, 2019a), we convert the tax from USD/tCO_2 into USD/Tcf. This allows us to include the tax in the production cost functions for shale and conventional gas in the following way:

$$K^c(q^c) = \left(k_0 + k^c \cdot (q^c)^\beta + \tau\right) \cdot q^c \tag{2.20}$$

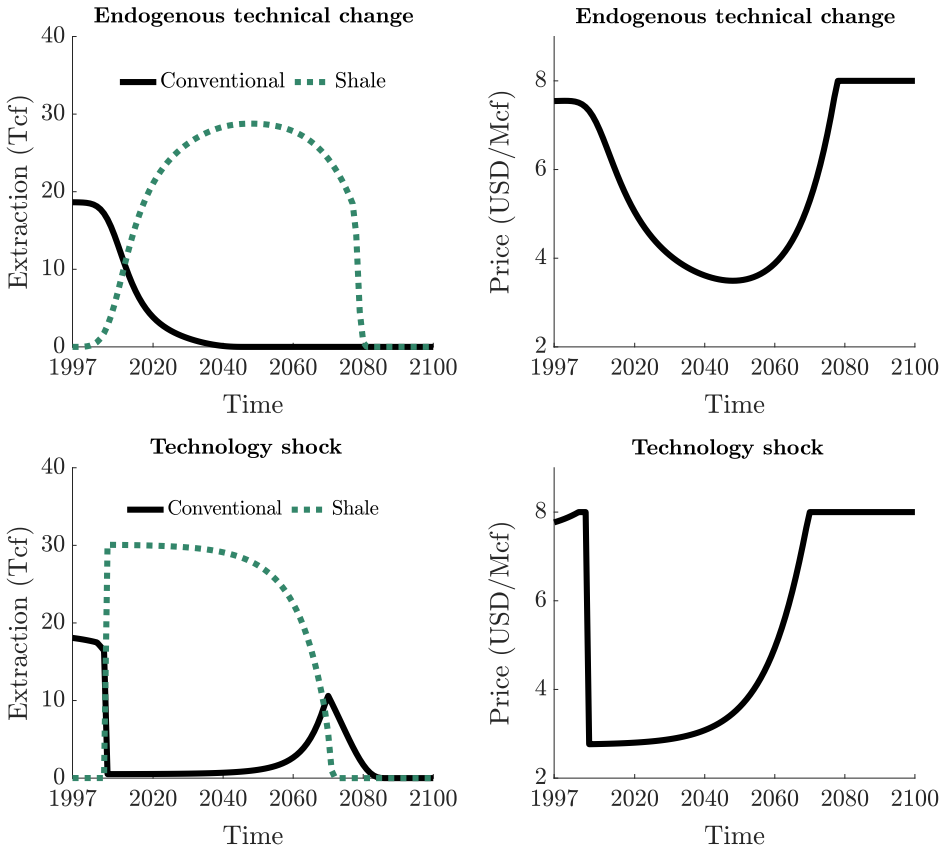

Figure 2.3: Natural gas extraction and price dynamics in the model with endogenous technological change (top graphs) and under the alternative assumption of an exogenous shock to shale extraction technology in the year 2007 (bottom graphs).

$$K^s(Z, q^s) = \left(k_0 + \frac{k^s}{Z} \cdot (q^s)^\beta + \tau \right) \cdot q^s \tag{2.21}$$

Figure 2.4 shows how a carbon tax affects the point in time when shale gas production starts and the point in time when shale gas production peaks. We define the start year as the year in which shale gas production exceeds 3 Tcf for the first time. This implies that in our baseline model, without a carbon tax, shale gas production starts in 2008 and peaks in 2048. If the carbon tax is small, the shale gas boom is only delayed by a couple of years. However, as the carbon tax is scaled up, the effect gets more pronounced. For a carbon tax of 45 USD/tCO_2, the start of the shale gas boom is pushed far into the future. Production starts in the year 2061 and peaks in 2112. Considering an even higher carbon tax of 46 USD/tCO_2

pushes the start year and peak year outside the here considered time horizon of 2120.

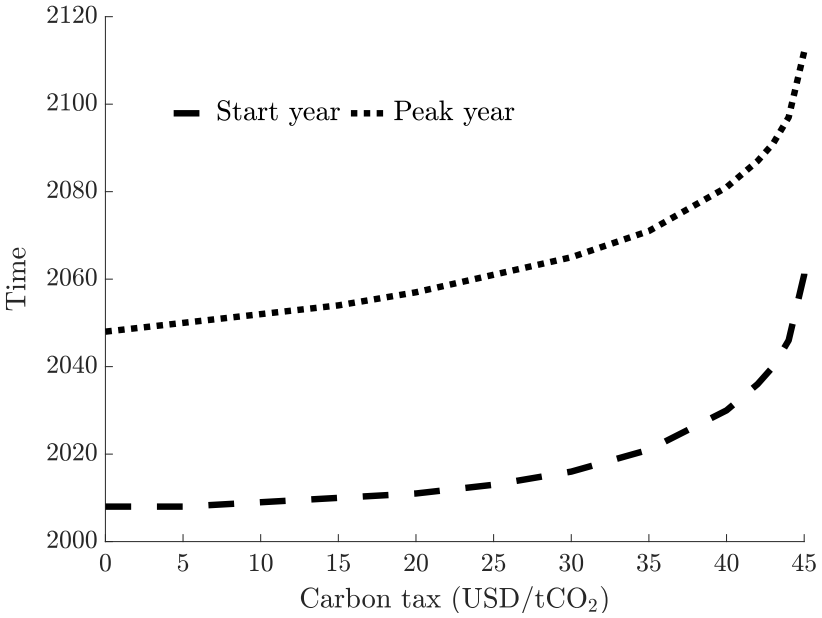

Figure 2.4: The graph shows the year in which shale gas production starts (dashed line) and the year in which shale gas production peaks (dotted line) as a function of the carbon tax. The start year is defined as the year in which shale gas production exceeds 3 Tcf for the first time. For a carbon tax higher than 45 USD/tCO_2, the year of the start and peak lie outside the here considered time horizon.

Figure 2.5 shows the sensitivity of the start year 2008 and 2050 with respect to a change in the carbon tax and the discount rate. We can observe that both, an increase in the carbon tax, but also an increase in the discount rate, induces a shift of technology investments into the future. Due to this, the production of shale gas is postponed. Considering for example the year 2050, the introduction of a carbon tax of 45 USD/tCO_2 is equivalent to a increase in the discount rate of about 8 percent. The sensitivity of the timing of the shale gas extraction to the discount rate underscores the importance of forward-looking behavior in technology development.

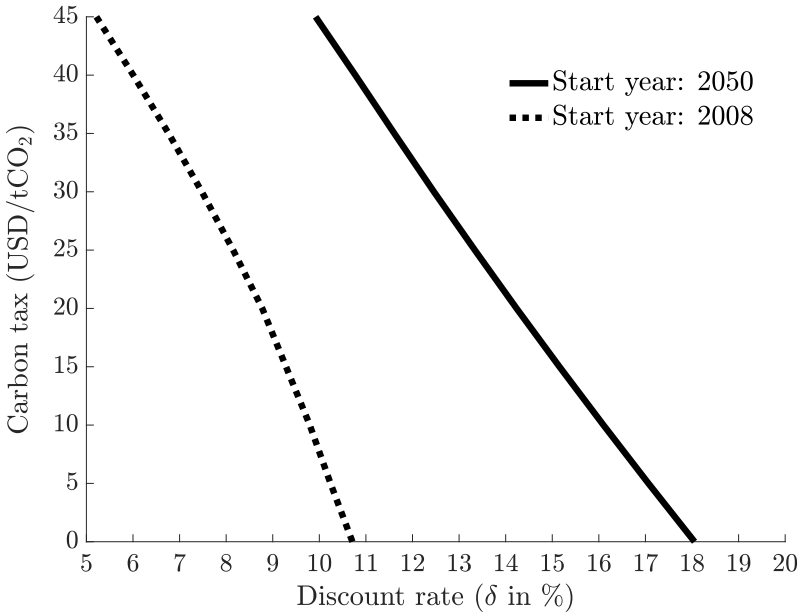

Figure 2.5: Sensitivity of the start year 2008 (dotted line) and 2050 (solid line) to a change in the carbon tax and the discount rate.

2.5.3 Effects of a carbon tax in models with endogenous technological change vs. technology shock

Last, we also include a carbon tax in our alternative model specification from section 2.5.1 where the improvement in extraction technology comes as an unanticipated shock to the industry. Afterwards, we compare the results from this model to the results from the model with endogenous technological change from section 2.5.2. Figure 2.6 shows how a carbon tax of 45 USD/tCO_2 would have affected natural gas extraction and price development for both model types.

The results are strikingly different. As already discussed in the previous section, in the model with endogenous technological change, the year of the start and peak of shale gas extraction is sensitive to the level of the carbon tax. Due to the tax of 45 USD/tCO_2, the shale gas boom is pushed into the future. Extraction starts in 2061 and peaks in 2112. This strong time-delaying effect is a specific result of the forward-looking behaviour of firms and stands in sharp contrast to the other model specification where the improvement in extraction technology comes as an

unanticipated shock. Under this alternative assumption, the start year of shale gas extraction is completely insensitive to the carbon tax. Similar to the model without a tax, extraction starts and price drops in 2007 when the shock hits. Afterwards, extraction continuously decreases and price continuously increases until shale and conventional gas are depleted. This demonstrates that for policy analysis it makes a big difference to endogenize technological change in the model.

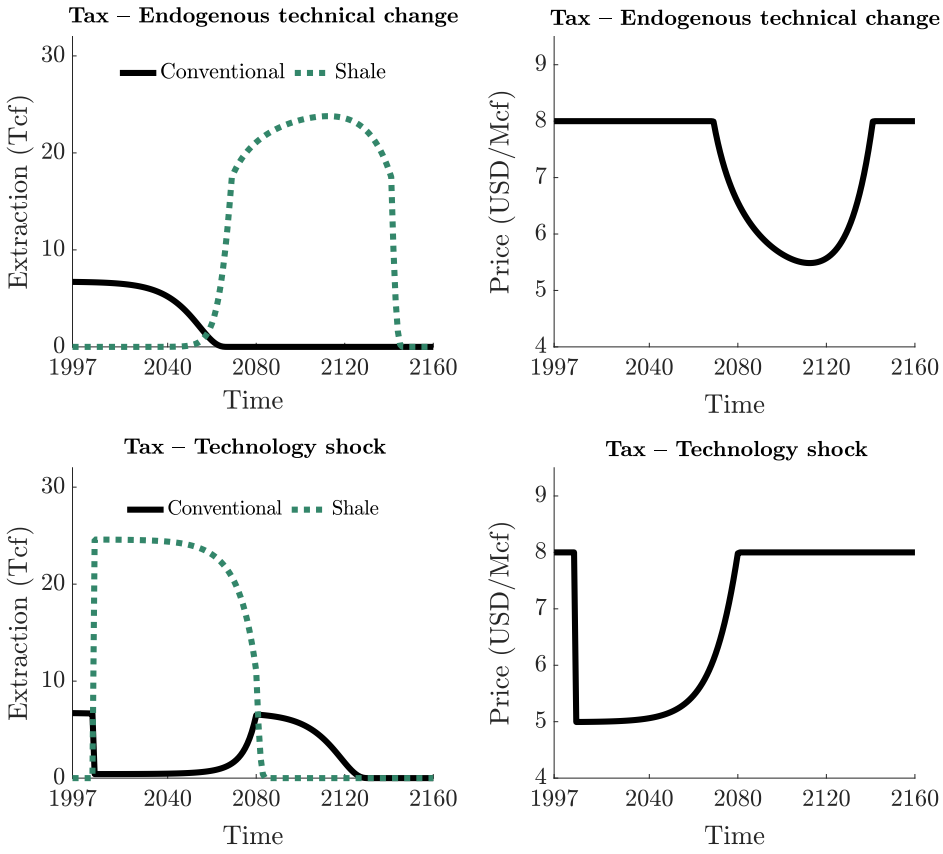

Figure 2.6: Natural gas extraction and price development with a carbon tax of 45 USD/tCO_2 in the calibrated model with endogenous technological change (top graphs) and under the alternative assumption of an exogenous shock to shale extraction technology in the year 2007 (bottom graphs).

2.6 Summary and conclusions

This paper presents a novel resource-economic theory that describes the main technological developments in the U.S. natural gas industry over the last two decades

as a result of endogenous decisions on technology investments.

We find that investments in extraction technology for shale gas start slowly and are increasing before shale gas extraction begins. As production shifts completely towards shale gas, our theory predicts that investments will start to decline. Similar to the literature on technology diffusion, the technology stock develops over time in a S-shaped fashion, first convex and then concave. Our theory also provides a simple explanation of the recent decrease in gas prices and increase in gas production. According to the model, resource price can follow a U-shaped path after shale gas production starts. This finding is in contrast to Herfindahl (1967) where price always increases after producers switch from a low to a high cost deposit type.

The calibrated model reproduces well the past development of U.S. shale gas production. The model suggests that natural gas production will continue to grow up to 30 Tcf and prices continue to decrease until 2050. We have also shown that for policy analysis it makes a big difference whether technical change is the endogenous result of forward-looking investment decisions or assumed to be an exogenous shock. The model with endogenous technical change suggests that a carbon tax could have postponed the shale gas boom considerably. This time-delaying effect is specific to the model with forward-looking investment behaviour, and does not occur in models that treat the improvement in extraction technology as an exogenous shock.

Our theory is not restricted to the natural gas industry but can also be applied to other exhaustible resources such as oil where similar technology developments are taking place. For all these resources, our model offers theoretical insights in the order of extraction. We find that by endogenizing technological investments in the Hotelling-Herfindahl model, production decisions do not only depend on initial shadow values but also on the state of extraction technology. This allows for additional orders of resource extraction. In particular, even if firms have already shifted to the use of the renewable backstop, production can switch back to a nonrenewable resource for a period of time, when extraction technology has sufficiently advanced.

References

Adelman, Morris A. Mineral depletion, with special reference to petroleum. The Review of Economics and Statistics, 72(1):1–10, 1990.

Aguilera, Roberto F. Production costs of global conventional and unconventional petroleum. Energy Policy, 64:134–140, 2014.

Aguilera, Roberto F and Radetzki, Marian. The shale revolution: Global gas and oil markets under transformation. Mineral Economics, 26(3):75–84, 2014.

Amigues, Jean-Pierre; Favard, Pascal; Gaudet, Gérard, and Moreaux, Michel. On the optimal order of natural resource use when the capacity of the inexhaustible substitute is limited. Journal of Economic Theory, 80(1):153–170, 1998.

Arezki, Rabah; Fetzer, Thiemo, and Pisch, Frank. On the comparative advantage of US manufacturing: evidence from the shale gas revolution. Journal of International Economics, 107:34–59, 2017.

Asche, Frank; Oglend, Atle, and Osmundsen, Petter. Gas versus oil prices the impact of shale gas. Energy Policy, 47:117–124, 2012.

Byrd, Richard H; Hribar, Mary E, and Nocedal, Jorge. An Interior Point Method for Large Scale Nonlinear Programming. SIAM Journal of Optimization, 9(4):877–900, 1999.

Byrd, Richard H; Nocedal, Jorge, and Waltz, Richard A. KNITRO: An integrated package for nonlinear optimization. In Di Pillo, G. and Roma, M., editors, Large-Scale Nonlinear Optimization, pages 35–59. Springer, New York, NY, 2006.

Carter, Colin A; Rausser, Gordon C, and Smith, Aaron. Commodity Booms and Busts. Annual Review of Resource Economics, 3(1):87–118, 2011.

Chakravorty, Ujjayant and Krulce, Darrell L. Heterogeneous Demand and Order of Resource Extraction. Econometrica, 62(6):1445–1452, 1994.

Chakravorty, Ujjayant; Krulce, Darrell, and Roumasset, James. Specialization and non- renewable resources: Ricardo meets Ricardo. Journal of Economic Dynamics and Control, 29(9):1517–1545, 2005.

Chakravorty, Ujjayant; Moreaux, Michel, and Tidball, Mabel. Ordering the extraction of polluting nonrenewable resources. American Economic Review, 98(3):1128–1144, 2008.

Davies, Stephen W. Inter-firm diffusion of process innovations. European Economic Review, 12(4):299–317, 1978.

EIA, Energy Information Administration. Shale gas provides largest share of U.S. natural gas production in 2013. 2014. URL https://www.eia.gov/todayinenergy/detail.php?id=18951.

EIA, Energy Information Administration. International Energy Outlook 2017. 2017a. URL https://www.eia.gov/outlooks/ieo/pdf/0484(2017).pdf.

EIA, Energy Information Administration. Natural Gas Gross Withdrawals and Production. 2017b. URL https://www.eia.gov/dnav/ng/NG_PROD_SUM_DC_NUS_MMCF_A.htm.

EIA, Energy Information Administration. Henry Hub Natural Gas Spot Price. 2018a. URL https://www.eia.gov/dnav/ng/hist/rngwhhdM.htm.

EIA, Energy Information Administration. U.S. Dry Natural Gas Production. 2018b. URL https://www.eia.gov/dnav/ng/hist/n9070us2A.htm.

EIA, Energy Information Administration. FAQ: How much carbon dioxide is produced when different fuels are burned? 2019a. URL https://www.eia.gov/tools/faqs/faq.php?id=73&t=1.

EIA, Energy Information Administration. How much shale gas is produced in the United States? 2019b. URL https://www.eia.gov/tools/faqs/faq.php?id=907&t=8.

Farzin, Y Hossein. Technological change and the dynamics of resource scarcity measures. Journal of Environmental Economics and Management, 29(1):105–120, 1995.

Gaudet, Gérard and Lasserre, Pierre. The efficient use of multiple sources of a nonrenewable resource under supply cost uncertainty. International Economic Review, 52(1):245–258, 2011.

Gaudet, Gérard and Salant, Stephen W. Modeling Nonrenewable Resources Use with Multiple Demands and Multiple Sources. Environmental and Resource Economics, 70(4):737–755, 2016.

Gaudet, Gérard; Moreaux, Michel, and Salant, Stephen W. Intertemporal depletion of resource sites by spatially distributed users. American Economic Review, 91(4):1149–1159, 2001.

Helm, Carsten and Mier, Mathias. On the efficient market diffusion of intermittent renewable energies. Energy Economics, 80:812–830, 2019.

Herfindahl, Orris C. Depletion and economic theory. Extractive resources and taxation, pages 63–90, 1967.

Holland, Stephen P. Extraction capacity and the optimal order of extraction. Journal of Environmental Economics and Management, 45(3):569–588, 2003.

Holland, Stephen P. Modeling Peak Oil. The Energy Journal, 29(2):61–79, 2008.

Hotelling, Harold. The economics of exhaustible resources. Journal of Political Economy, 39(2):137–175, 1931.

Hu, Desheng and Xu, Shengqing. Opportunity, challenges and policy choices for China on the development of shale gas. Energy Policy, 60:21–26, 2013.

IEA, International Energy Agency. Resources to Reserves – Oil Gas and Coal Technologies for the Energy Markets of the Future. 2013. URL https://www.oecd-ilibrary.org/energy/resources-to-reserves-2013_9789264090705-en.

Jacobsen, Grant D and Parker, Dominic P. The Economic Aftermath of Resource Booms: Evidence from Boomtowns in the American West. Economic Journal, 126(593):1092–1128, 2016.

Jaffe, Adam B and Stavins, Robert N. The energy-efficiency gap What does it mean? Energy Policy, 22(10):804–810, 1994.

Joskow, Paul L. Natural Gas: From Shortages to Abundance in the United States. American Economic Review: Papers and Proceedings, 103(3):338–343, 2013.

Joskow, Paul L. The Shale Gas Revolution: Introduction. Economics of Energy and Environmental Policy, 4(1):1–4, 2015.

Kemp, Murray C and Long, Ngo Van. On Two Folk Theorems Concerning the Extraction of Exhaustible Resources. Econometrica, 48(3):663–673, 1980.

Lee, Woo Jin and Sohn, So Young. Patent analysis to identify shale gas development in China and the United States. Energy Policy, 74:111–115, 2014.

Lewis, Tracy R. Sufficient Conditions for Extracting Least Cost Resource First. Econometrica, 50(4):1081–1083, 1982.

Lin, C-Y Cynthia and Wagner, Gernot. Steady-state growth in a Hotelling model of resource extraction. Journal of Environmental Economics and Management, 54(1):68–83, 2007.

Mason, Charles F; Muehlenbachs, Lucija A, and Olmstead, Sheila M. The economics of shale gas development. Annu. Rev. Resour. Econ., 7(1):269–289, 2015.

Newell, Richard G; Prest, Brian C, and Vissing, Ashley B. Trophy hunting versus manufacturing energy: The price responsiveness of shale gas. Journal of the Association of Environmental and Resource Economists, 6(2):391–431, 2019.

Pindyck, Robert S. The Optimal Exploration and Production of Nonrenewable Resources. Journal of Political Economy, 86(5):841–861, 1978.

Pindyck, Robert S. Uncertainty in the Theory of Renewable Resource Markets. The Review of Economic Studies, 51(2):289–303, 1984.

Rausser, Gordon C. Technological change, production, and investment in natural resource industries. The American Economic Review, 64(6):1049–1059, 1974.

Slade, Margaret E. Trends in natural-resource commodity prices: an analysis of the time domain. Journal of Environmental Economics and Management, 9(2):122–137, 1982.

Wakamatsu, Hiroki and Aruga, Kentaka. The impact of the shale gas revolution on the US and Japanese natural gas markets. Energy Policy, 62:1002–1009, 2013.

Wang, Qiang; Chen, Xi; Jha, Awadhesh N, and Rogers, Howard. Natural gas from shale formation–the evolution, evidences and challenges of shale gas revolution in United States. Renewable and Sustainable Energy Reviews, 30:1–28, 2014.

A.1 Appendix to Chapter 2

A.1.1 Proof of Lemma 2.1

The marginal benefit of the technology develops according to

$$\frac{d}{dt}\left(-K_Z^s(q^s, Z)\right) = -K_{ZZ}^s(q^s, Z)\, w - K_{qZ}^s(q^s, Z)\, \dot{q}^s. \tag{2.22}$$

During the phase of shale gas extraction, (2.7a) holds with equality. Differentiating that condition with respect to time, and using (2.4) yields

$$\begin{aligned}\Big(P'(q^s + q^c + q^b) - K_{qq}^s(q^s, Z)\Big)\, \dot{q}^s \\ = K_{qZ}^s(q^s, Z)\, w + \delta\, \Big(P(q^s + q^c + q^b) - K_q^s(q^s, Z)\Big).\end{aligned} \tag{2.23}$$

Using this in (2.22) and collecting terms delivers

$$\begin{aligned}-\Big(P'(q^s + q^c + q^b) - K_{qq}^s(q^s, Z)\Big)\, \frac{d}{dt}\left(-K_Z^s(q^s, Z)\right) \\ = \Big(\Big(P'(q^s + q^c + q^b) - K_{qq}^s(q^s, Z)\Big)\, K_{ZZ}^s(q^s, Z) + K_{qZ}^s(q^s, Z)^2\Big)\, w \\ + \delta\, K_{qZ}^s(q^s, Z)\, \Big(P(q^s + q^c + q^b) - K_q^s(q^s, Z)\Big).\end{aligned} \tag{2.24}$$

The first term on the left-hand-side of this equation is positive, $-\left(P'(q^s + q^c + q^b) - K_{qq}^s(q^s, Z)\right) > 0$, as the inverse demand function is downward-sloping and the marginal extraction costs are non-decreasing. The first term on the right-hand-side of the equation is non-positive by the concavity assumption, cf. (2.6). The second term on the right-hand side of the equation is negative, as $K_{qZ}^s < 0$ and $P - K_q = \gamma > 0$.

A.1.2 Proof of Proposition 2.1

By (2.7d), the shadow price of technology is non-negative. Whenever investment is positive, such that (2.7d) holds with equality, the temporal change of investment is monotonic in the change in the shadow price of technology, as $L''(w)\, \dot{w} = \dot{\phi}$ and $L''(w) > 0$ by assumption. In a corner solution $w = 0$, $\dot{w} = 0$ by definition.

Whenever $q^s(t) = 0$, we have $K_Z^s(q^s, Z) = 0$ and thus (2.8c) implies $\dot{\phi} = \delta\,\phi$. The shadow price of technology, and, hence, investment, monotonically increases over time – provided it is positive to begin with.

Consider the time interval $t \in [\underline{T}^s, \overline{T}^s]$ with extraction from deposit type S, $q^s(t) > 0$, where $\overline{T}^s$ is the point in time where shale gas extraction ends. It must be $\phi(t) = 0$ for all $t \geq \overline{T}^s$, as otherwise $\phi(t)$ would grow at the discount rate after $\overline{T}^s$, whereas $Z(t) > 0$, which would violate the transversality condition (2.8d).

Thus, during the phase of shale gas extraction,

$$\phi(t) = -\int_t^{\overline{T}^s} e^{-\delta\,(\tau - t)}\, K_Z^s(q^s(\tau), Z(\tau))\, d\tau. \tag{2.25}$$

By Lemma 2.1 we have

$$\delta\,\phi(t) < -K_Z^s(q^s(t), Z(t)) \int_t^{\overline{T}^s} \delta\, e^{-\delta\,(\tau - t)}\, d\tau \tag{2.26}$$

$$= -K_Z^s(q^s(t), Z(t)) \left(1 - e^{-\delta\,(\overline{T}^s - t)}\right) < -K_Z^s(q^s(t), Z(t)). \tag{2.27}$$

Thus,

$$\dot{\phi} = \delta\,\phi + K_Z^s(q^s(t), Z(t)) < 0. \tag{2.28}$$

A.1.3 Proof of Proposition 2.2

A switch in production can occur when augmented marginal costs of two resource types are equal. Thus, we have three switch functions.

i) Switch points between shale and conventional gas are determined by the condition that the difference of AMC^s and AMC^c is zero. The switch points are given as the roots of the switch function

$$\psi^{sc}(t) = AMC^s(t) - AMC^c(t) = k(Z(t)) - k^c - (\lambda(0) - \gamma(0))\, e^{\delta\, t}. \tag{2.29}$$

Note that at a switch point from conventional to shale gas, investments in the technology

stock are positive. The derivative of the switch function with respect to time is

$$\dot{\psi}^{sc} = k'(Z)\, w - \delta\,(\lambda(0) - \gamma(0))\, e^{\delta\, t}. \tag{2.30}$$

The first term is negative. The second term is negative as well if $\lambda(0) - \gamma(0) > 0$, whereas it is positive for $\lambda(0) - \gamma(0) < 0$. In case $\lambda(0) - \gamma(0) \geq 0$, the switch function ψ_t^{sc} is monotonically decreasing at any switch point. As the switch function is continuous, there is at most one switch point between regimes. Furthermore, the switch must be from conventional to shale gas. The only possibility is first deposit type C, then deposit type S. In case $\lambda(0) - \gamma(0) < 0$, there may be more than one switch point. As the switch function is continuous, it must assume a maximum or minimum in between any two switch points. At this maximum or minimum, the curvature of the switch function is

$$\ddot{\psi}^{sc} = k''(Z)\, w^2 + k'(Z)\, \dot{w} + \delta^2\ (\gamma(0) - \lambda(0))\ e^{\delta\, t} \tag{2.31a}$$

$$= k''(Z)\, w^2 + k'(Z)\, \frac{w}{\phi}\,(\delta\, \phi + k'(Z)\, q^s) + \delta^2\ (\gamma(0) - \lambda(0))\ e^{\delta\, t} \tag{2.31b}$$

$$= \underbrace{k''(Z)\, w^2 + (k'(Z))^2\, \frac{w}{\phi}\, q^s}_{>0} + \delta\ \underbrace{(k'(Z)\, w + \delta\ (\gamma(0) - \lambda(0))\ e^{\delta\, t})}_{=\dot{\psi}^{sc}} > 0. \tag{2.31c}$$

Thus, $\ddot{\psi}^{sc} > 0$ for any point in time where $\dot{\psi}^{sc} = 0$, which implies that $\dot{\psi}^{sc}$ can change signs only once. Overall, there are at most two switch points between the regimes. The only possible orders of production are SC, CSC, and CS.

ii) Switch points between conventional gas and the backstop are determined by

$$\psi^{cb} = k^c + \lambda(0)\, e^{\delta\, t} - k^b = 0. \tag{2.32}$$

We have $k^c - k^b < 0$ is constant. Thus, if there is a switch point at all, the switch function is monotonically increasing at this point. As the switch function is continuous, there is exactly one switch point and the switch must be from deposit type C to the backstop.

iii) Switch points between shale gas and the backstop are determined by

$$\psi^{sb} = k(Z) + \gamma(0)\, e^{\delta\, t} - k^b = 0. \tag{2.33}$$

The derivative of the switch function with respect to time is

$$\dot{\psi}^{sb} = k'(Z)\, w + \delta\, \gamma(0)\, e^{\delta\, t}. \tag{2.34}$$

The first term is negative and the second term is positive, i.e. $\dot{\psi}^{sb}$ may switch signs, and thus $\psi^{sb} = 0$ may have more than one solution. Consider the curvature of ψ^{sb} at the minimum or maximum, i.e. when $\dot{\psi}^{sb} = 0$:

$$\ddot{\psi}^{sb} = k''(Z)\, w^2 + k'(Z)\, \dot{w} + \delta^2\, \gamma(0)\, e^{\delta\, t} \tag{2.35a}$$

$$= k''(Z)\, w^2 + k'(Z)\, \frac{w}{\phi}\, (\delta\, \phi + k'(Z)\, q^s) + \delta^2\, \gamma(0)\, e^{\delta\, t} \tag{2.35b}$$

$$= \underbrace{k''(Z)\, w^2 + (k'(Z))^2\, \frac{w}{\phi}\, q^s}_{>0} + \delta\, \underbrace{(k'(Z)\, w + \delta\, \gamma(0)\, e^{\delta\, t})}_{=\dot{\psi}^{sb}} > 0. \tag{2.35c}$$

Thus, $\ddot{\psi}^{sb} > 0$ for any point in time where $\dot{\psi}^{sb} = 0$, which implies that $\dot{\psi}^{sb}$ can change signs only once. Overall, there are at most two switch points between regimes. The only possible orders of production are BSB, and SB. Note that BS is not possible since w goes to zero as the shale deposits approaches depletion. Hence, in the end ψ^{sb} must be positive.

Looking at all combinations of the switch functions yields the four extraction orders from Proposition 2.2. Note that $\psi^{sc}(0) > \psi^{sb}(0)$ since $k^b > k^c + \lambda(0)$. Hence, BSB can never occur in combination with SC.

3 Solar geoengineering and strategic interactions in a regional analytic climate economy

Felix D. Meier[1,2] and Christian P. Traeger[3,4,5]

[1] *Kiel Institute for the World Economy*

[2] *German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig*

[3] *Department of Economics, University of Oslo*

[4] *ifo Institute, Munich*

[5] *Frisch centre, Oslo*

Abstract: The paper analyzes solar geoengineering and strategic interactions in an integrated assessment model (IAM) of climate change. For this purpose, we (i) derive a new class of solutions to analytic IAMs that allows us to (ii) solve an integrated assessment model with sulfur-based geoengineering and damages in closed-form, and to (iii) model realistic strategic interactions between two active regions, and a passive rest of the world. We examine the determinants of a region's engagement in geoengineering, and analyze the Markov game and its equilibria. Our model suggests that sulfur deployment is highly sensitive to potential geoengineering damages. For a global social planner, solar geoengineering could cut the Social Cost of Carbon (SCC) into half, if damages turn out negligible. However, current damage guesstimates would reduce the global SCC by only 12-22%. In the regional model, the availability of geoengineering can both increase or decrease the SCC, in both the active regions and the rest of the world, depending on the heterogeneity of climate and geoengineering damages across regions.

Keywords: climate change, integrated assessment, solar geoengineering, dynamic games, strategic conflicts, free-driver, free-rider, carbon tax, social cost of carbon

3.1 Introduction

Worldwide greenhouse gas emissions are still on the rise (Tollefson, 2018) leading to potentially severe consequences for the world and its economy. Future warming will substantially reduce future output and may reduce global economic growth rates (Carleton and Hsiang, 2016) apart from destroying ecosystems and driving species to extinction (Urban, 2015; Hoegh-Guldberg and Bruno, 2010). In light of these developments, engineering a cooler climate evolves to be a hot topic. Observations from large volcanic eruptions suggest that the deliberate injection of sulfur aerosols into the stratosphere by airplanes can cool our planet by reflecting sunlight back into space (Crutzen, 2006).

In this paper, we build a full-fledged Integrated Assessment Model (IAM) with state of the art climate dynamics and an economy that transforms nonrenewable and renewable resources into energy, final goods, and emissions. Introducing a new class of analytic solutions, we are able to integrate the temperature response to stratospheric sulfur injections into our model, calibrate the cooling efficiency of sulfur well to recent scientific work of Kleinschmitt et al. (2018), and model realistic strategic interactions between regions. We separate climate change damages into damages from rising temperatures and damages from increasing atmospheric carbon concentrations (Moreno-Cruz et al., 2017; Klepper and Rickels, 2014). Geoengineering implies a third type of damages, side-effects of the employed chemicals and changes in the radiative spectrum (Heckendorn et al., 2009; Crutzen, 2006; Keith and MacMartin, 2015; Robock, 2008; Kravitz et al., 2009), which are at this point in time mostly unknown (Emmerling and Tavoni, 2018b).

We provide an analytic description of a global planner's optimal sulfur deployment, and show how it depends on climate and geoengineering damages, climate dynamics, and the cooling efficiency of sulfur. Despite several non-linearities and efficiency losses, a quantitatively well-fitted model gives rise to a sulfur deployment

strategy that increases linearly in the atmospheric carbon concentration. The resulting reduction in the Social Cost of Carbon (SCC) and thus in the incentives to abate carbon dioxide is a delicate balance between the effectiveness of sulfur-based cooling and potential damages from geoengineering.

We split the world into two climate zones inhabited by non-cooperatively acting regions, and a passive rest of the world that only affects the decisions of the active regions by its contributions to the atmospheric carbon concentration. Each region prefers a world with preindustrial temperatures but has to balance the trade-offs between climate change impacts and geoengineering damages. Geoengineering affects regions via two channels. First, sulfur particles spread around the globe leading to heterogeneous cooling and damage spill-overs between the active regions and the rest of the world.[5] Second, a region's temperature adjusts to that of other regions, directly through heat exchange across the planet and via a common ocean. Both of these interaction channels give rise to strategic behavior. We allow each region to either contribute to the cooling, stay inactive, or to engage in a potential countermeasure that neutralizes part of the cooling and damage spill-overs from the other region (Parker et al., 2018; Heyen et al., 2019).

The game solves in linear strategies. This linear dependence on the atmospheric carbon concentration not only permits an analytic solution of the game, but also coincides with the structure of the optimal carbon dioxide response in the non-strategic model. In the resulting dynamic Markov game, we analyze how regional sulfur deployment and carbon dioxide mitigation incentives depend on the characteristics of both regions. We derive a set of subgame-perfect Nash equilibria and characterize three qualitatively different types of equilibria. In the climate match, both regions contribute to global cooling. Under unilateral action, only one region

[5]If sulfur is injected along the equator, particles spread effectively towards the poles leading to an almost uniform cooling effect across the globe (Lawrence et al., 2018). We deviate from this common assumption in the regional model as recent literature suggests that the geographic distribution of the cooling effect can be optimized by varying the latitude and altitude of injections (Jones et al., 2018; Kravitz et al., 2017; MacMartin et al., 2017).

engages in geoengineering and we shed some light on whether and when free-riding or free-driving best describes such situations. Finally, we permit counter-measures to geoengineering giving rise to another equilibrium where regional interests clash. The introduction of counter-geoengineering also allows us to identify whether an inactive region is free-riding or if it is opposed to the other region's actions and therefore engaging in counter-measures.

Our model contributes to two fields of literature. First, we add to the literature that studies strategic incentives from solar geoengineering. Low operational costs (Smith and Wagner, 2018; McClellan et al., 2012) and heterogeneous side-effects (Ricke et al., 2010) have led to the concern that a country could implement geoengineering unilaterally at the expense of others; the so called "free-driver" incentive (Weitzman, 2015; Pasztor et al., 2017). Heyen et al. (2019) find that the free-driver outcome becomes unstable once counter-geoengineering is available. However, the option of counter-geoengineering might instead lead to a "climate clash" when no moratorium treaty (countries abstain from climate engineering) and no cooperative deployment is realized.

A further concern is that the prospects of geoengineering could undermine traditional mitigation efforts and push society onto a "slippery slope" (Lawrence and Crutzen, 2017; Morrow, 2014; Quaas et al., 2017). Moreno-Cruz (2015) shows that the impact of geoengineering on mitigation depends on the similarity between countries. If countries are similar with respect to climate and geoengineering damages, the option of geoengineering leads to lower mitigation levels. However, if countries differ with respect to damages, mitigation in both countries can increase.

Our paper contributes to this literature by taking the strategic incentives governing geoengineering into a dynamic integrated assessment model of climate change. We characterize Markov perfect equilibria where regions' interests either match, clash, or exhibit free-riding, or free-driving behavior. Moreover, we derive analytic formulas for the regional SCC to study how the availability of geoengineering affects

mitigation incentives in the active regions, and whether the passive rest of the world makes the slope of geoengineering even more slippery, or turns the slippery slope uphill.

Second, we contribute to the recently emerging literature on analytic integrated assessment models of climate change that derive closed-form solutions for the globally optimal carbon tax (Traeger, 2018; Golosov et al., 2014; Gerlagh and Lsiki, 2018). We explain the components that reduce earlier formulas for the optimal carbon tax because of our ability to cool the planet.

Our study is among the first to analyze the strategic interaction of regions within an integrated assessment model. Other game-theoretic models include Nordhaus and Yang (1996), who develop a regional version of the DICE model (Nordhaus and Sztorc, 2013) which is called the RICE model. Tol (2002a,b) uses the multi-regional FUND model to estimate the damages of climate change. Nordhaus (2015) studies free-riding incentives in climate agreements and introduces the idea of climate clubs with trade penalties for non-participants. Hassler and Krusell (2012) develop a dynamic stochastic general-equilibrium model with multiple regions that can be solved in closed-form, and show that only taxes on oil producers can mitigate climate change. Taxes on oil consumers have no effect. Hambel et al. (2018) analyze the role of international trade in a regional analytic climate economy model and find that the SCC increases in trade volume.

Our paper is structured as follows. The next section introduces sulfur-based geoengineering into the analytic climate economy model from Traeger (2018), and derives the optimal sulfur strategy and the optimal carbon tax for a global planner. In section 3.3, we split the model into two non-cooperatively acting regions, and a passive rest of the world. We first discuss the results of a baseline model version without heat flows, and analyze how the rest of the world reacts to the decisions of the active regions. Afterwards, we allow heat to equilibrate across the globe, and show how this affects regional strategies, and mitigation incentives.

3.2 Global model

This section introduces geoengineering into the analytic climate economy model ACE (Traeger, 2018). First, we summarize a slightly simplified version of the ACE model. Then, we introduce geoengineering and calibrate the forcing effect of sulfur to scientific data. Finally, we discuss the optimal cooling strategy of the social planner and the difference that geoengineering makes for the optimal carbon tax.

3.2.1 Economy

Production. Gross output is a function of vectors of dimension I_j with $j \in \{A, N, K, E\}$. The technology levels $\boldsymbol{A}_t$ are exogenous. Capital is optimally distributed over the different sectors, resulting in the capital levels summarized by the vector $\boldsymbol{K}_t$. Labor distribution is denoted by $\boldsymbol{N}_t$ and energy inputs by $\boldsymbol{E}_t$.

$$Y_t = \mathcal{F}(\boldsymbol{A}_t,\, \boldsymbol{K}_t,\, \boldsymbol{N}_t,\, \boldsymbol{E}_t) \tag{3.1}$$

The production function is homogeneous of degree κ in capital such that

$$\mathcal{F}(\boldsymbol{A}_t,\, \lambda\, \boldsymbol{K}_t,\, \boldsymbol{N}_t,\, \boldsymbol{E}_t) = \lambda^{\kappa} \mathcal{F}(\boldsymbol{A}_t,\, \boldsymbol{K}_t,\, \boldsymbol{N}_t,\, \boldsymbol{E}_t) \quad \forall \quad \lambda \in \mathbb{R}_+. \tag{3.2}$$

Population size is normalized to unity $\sum_{i=1}^{I_N} N_{i,t} = 1$.

Damages. We denote the atmospheric carbon stock (or concentration) by $M_{1,t}$. It is convenient to measure it relative to the preindustrial stock level as $m_t = \frac{M_{1,t}}{M_{\text{pre}}}$. Global atmospheric temperature $T_{1,t}$ measures the temperature increase over 1900 levels (in degree Celsius). Atmospheric temperature increase, atmospheric carbon concentration, as well as the level of the cooling agent S_t cause (net) damages $D_t(T_{1,t}, S_t, m_t)$ that we measure as a fraction of output. They are of the form

$$D_t(T_{1,t}, S_t, m_t) = 1 - \exp\left[-D_T(T_{1,t}) - D_G(S_t) - D_m(m_t)\right]. \tag{3.3}$$

We take the temperature-based damages

$$D_T(T_{1,t}) = \xi_0 \exp(\xi_1 T_{1,t}) - \xi_0 \tag{3.4}$$

from the ACE model's calibration to DICE. Our global model assumes

$$D_G(S_t) = d\, S_t \tag{3.5}$$

making d the semi-elasticity of damages from stratospheric sulfur injections (the percentage loss of output resulting from an additional ton of sulfur injections). The parameter d includes linear operational costs. Our regional version of the model in section 3.3 further refines the damage parameter into damages from geoengineering, counter-geoengineering, and the costs of injecting the chemical agents into the stratosphere. The net costs of atmospheric carbon are

$$D_m(m_t) = a\,(m_t - 1), \tag{3.6}$$

where a is the semi-elasticity of production with respect to changes in the carbon dioxide concentration. Costs include ocean-acidification and benefits the fertilizer effect that increases plant production and crop yields (Proctor et al., 2018).

Emissions, resources and capital. The first I^d energy inputs $E_1, ..., E_{I^d}$ are fossil fuels causing CO_2 emission and are collected in subvector $\boldsymbol{E}_t^d$. These energy inputs are measured in terms of their CO_2 content so that total emissions from production are $\sum_{i=1}^{I^d} E_{i,t}$. Fossil fuel resource stocks are collected in vector $\boldsymbol{R}_t \in \mathbb{R}_+^{I^d}$. The dynamics of the resource stock are

$$\boldsymbol{R}_{t+1} = \boldsymbol{R}_t - \boldsymbol{E}_t^d \tag{3.7}$$

with initial stock size $\boldsymbol{R}_0 \in \mathbb{R}_+^{I^d}$. Renewable energies are indexed by I^{d+1} to I_E.

We assume full depreciation of capital over the course of a decade, the model's time step. We choose this simplifying assumption noting that Traeger's (2018) extension for capital persistence would also go through in our setting. The global economy's capital stock evolves as

$$\begin{aligned} K_{t+1} &= Y_t \left[1 - D_t\left(T_{1,t}, S_t, m_t\right)\right] - C_t \\ &= Y_t \exp\left[-\xi_0 \exp\left(\xi_1 T_{1,t}\right) + \xi_0 - d\, S_t - a(m_t - 1)\right] - C_t. \end{aligned} \tag{3.8}$$

3.2.2 Climate

Carbon dioxide. Similar to DICE, we consider three carbon reservoirs, atmosphere (carbon content M_1), upper ocean (carbon content M_2) and lower ocean (carbon content M_3) which we summarize in the vector $\boldsymbol{M}$. The extension to additional carbon reservoirs is straight-forward. The dynamics of the carbon reservoirs is

$$\boldsymbol{M}_{t+1} = \boldsymbol{\Phi} \boldsymbol{M}_t + \tilde{\boldsymbol{e}}_t, \tag{3.9}$$

with transition matrix $\boldsymbol{\Phi}$. Further we define $\tilde{\boldsymbol{e}}_t = \boldsymbol{e}_1 E_t^{tot}$, with total CO_2 emissions $E_t^{tot} = \sum_{i=1}^{I^d} E_{i,t} + E_t^{\text{exo}}$ resulting from industrial fossil fuel burning and other exogenous processes including lang use change and forestry.

Greenhouse effect and cooling. Greenhouse gases in the atmosphere as well as geoengineering affect our planet's temperature. The net heating with respect to preindustrial times is summarized by the resulting radiative forcing F_t (measured in W/m^2). It increases (logarithmically) in the increase of atmospheric CO_2 relative to preindustrial levels m_t and it falls as a consequence of geoengineering measures $G_t(S_t)$ that inject sulfur aerosols S_t into the stratosphere

$$F_t^{exact} = \frac{\eta}{\log 2} \log(m_t) - G_t(S_t) = \frac{\eta}{\log 2} \log\left(m_t \exp\left(-\frac{\log 2}{\eta} G_t(S_t)\right)\right).$$

We ignore non-CO_2 greenhouse gases, which can easily be added as in the ACE model. The next subsection will fit a data-based approximation F_t to F_t^{exact} that will replace the generic formulation above. It is common to express radiative forcing in CO_2 equivalents, which corresponds to the argument of the logarithm on the right side of the equation. We note that, in terms of CO_2 equivalents, the forcing from sulfur and CO_2 are no longer independent (see section 3.2.3).

Temperature dynamics. In the medium to long run a new level of radiative forcing implies the new atmospheric equilibrium temperature $T_{0,t} = \frac{s}{\eta} F_t$. Following ACE, we model the evoloution of atmospheric temperature $T_{1,t}$ as a generalized mean of last period's atmospheric temperature (persistence), the last period's ocean temperature (currently cooling), and the new equilibrium temperature corresponding to radiative forcing $T_{0,t}$. Similarly ocean temperature $T_{2,t}$ evolves as a generalized mean of own past and atmospheric temperature

$$\begin{aligned} T_{1,t+1} &= \frac{1}{\xi_1} \log\Big((1 - \sigma^{forc} - \sigma_{21}) \exp(\xi_1 T_{1,t}) + \sigma^{forc} \exp(\xi_1 T_{0,t}) + \sigma_{21} \exp(\xi_1 T_{2,t}) \Big) \\ T_{2,t+1} &= \frac{1}{\xi_1} \log\Big((1 - \sigma_{12}) \exp(\xi_1 T_{2,t}) + \sigma_{12} \exp(\xi_1 T_{1,t}) \Big) \end{aligned} \tag{3.10a}$$

with $\xi_1 = \frac{\log 2}{s}$. We rewrite these equations in terms of transformed temperatures $\tau_{it} = \exp(\xi_1 T_{i,t})$ as

$$\begin{pmatrix} \tau_{1,t+1} \\ \tau_{2,t+1} \end{pmatrix} = \underbrace{\begin{pmatrix} 1 - \sigma_{\text{forc}} - \sigma_{21} & \sigma_{21} \\ \sigma_{12} & 1 - \sigma_{12} \end{pmatrix}}_{\equiv \boldsymbol{\sigma}} \begin{pmatrix} \tau_{1,t} \\ \tau_{2,t} \end{pmatrix} + \begin{pmatrix} \sigma_{\text{forc}} \exp\left(\frac{\log(2)}{\eta} F_t \right) \\ 0 \end{pmatrix}. \tag{3.10b}$$

3.2.3 Geoengineering

Based on the experience of many volcanic eruptions scientists have learned that the injection of small sulfur particles (aerosols) into the atmosphere reflects sunlight

back into space cooling our planet. Yet, at high injection rates, sulfur particles lump together which decreases their cooling efficiency. As a result, scientists expect that the cooling from stratospheric aerosol injections has an asymptotic limit (Lawrence et al., 2018). The uncertainty governing the forcing efficiency is high and the instantaneous radiative forcing effect of sulfur injections varies strongly across different climate models (Kleinschmitt et al., 2018; Lawrence et al., 2018; Niemeier and Schmidt, 2017; Niemeier and Timmreck, 2015). Table 3.1 shows model data on the effective radiative forcing effect from sulfur injections from two recent studies. Note that the effective radiative forcing effect, which also includes rapid adjustments such as changes in atmospheric temperature, is larger than the instantaneous radiative forcing effect (Boucher et al., 2017).[6]

Table 3.1: Effective radiative forcing effect from sulfur injections

Kleinschmitt et al. (2018)

TgS	2	5	10	20	50
W/m^2	-1.11	-1.64	-2.91	-4.34	-5.63

Niemeier and Schmidt (2017)

TgS	4	6	8	10	30	40	50
W/m^2	-0.34	-1.30	-1.54	-1.78	-4.04	-4.76	-5.18

Units. A negative 6 W/m^2 (Watts per square meter) is approximately double the cooling power that we have currently produced in terms of warming a result of antropogenic greenhouse gas emissions since preindustrial times (IPCC, 2013). 1 TgS/yr (Tera grams sulfur per year) are 1 Million tons of sulfur annually deployed into the stratosphere. Each TgS/yr corresponds to approximately 25 Boeing 747 loads deployed daily for a year. We note that a Boeing cannot make it into the stratosphere, but it can fuel fighter jets that deploy the sulfur in the necessary altitude. Other options to deploy the sulfur include stratospheric balloons (Robock et al., 2009).

[6] In addition, the literature has proposed alternative aerosols like alumina and diamond particles (Weisenstein et al., 2015; Dykema et al., 2016), calcite or limestone (Keith et al., 2016). Given the lack of a natural experiment with such aerosols, our knowledge about the resulting forcing effect is even more limited.

Calibration. We calibrate our model to the recent study by Kleinschmitt et al. (2018).[7] For this purpose, we develop a new functional form with several degrees of freedom that will permit an analytic solution of the dynamic programming problem. We approximate exact radiative forcing for the use of stratospheric sulfur injections by

$$F_t^{exact} \approx F_t = \frac{\eta}{\log(2)} \log \left(\underbrace{f_0 + f_1\, m_t + \left(f_2 - f_3 \left(\frac{m_t}{S_t} \right)^n \right) S_t}_{\equiv F_t^{CO2}} \right). \tag{3.11}$$

The expression F_t^{CO2} characterizes the effect of sulfur on radiative forcing in CO_2 equivalents. In the absence of geoengineering, the term m_t would capture the CO_2 forcing. The round inner bracket reduces the forcing in response to sulfur injections. The main contribution derives from the term $f_3 \left(\frac{m_t}{S_t} \right)^n$. Sulfur forcing is more efficient relative to CO_2 the larger the atmospheric carbon concentration and the lower the sulfur concentration. For high levels of sulfur, particles lump together reducing their cooling efficiency. The higher the CO_2 concentration, the lower the warming implied by the marginal ton of CO_2 and the higher the relative forcing reduction of sulfur, which we measure in CO_2 equivalents. We summarize both of these nonlinearities in the joint term whose level effect is captured by f_3 and whose nonlinearity is captured by $n > 0$.

We fit the function to Kleinschmitt et al.'s (2018) data on effective radiative forcing from sulfur injections (Table 3.1).[8] Our fit combines Kleinschmitt et al.'s (2018) forcing data for sulfur injections with the well-known forcing from atmospheric carbon dioxide over the interval $m_t \in [1.5, 3]$, i.e. up to a tripling of preindustrial carbon dioxide emissions. Our fit minimizes the squared differences for those 80 data points. We list the resulting parameters in Table 3.2.

[7] Note that the publication only cites the instantaneous radiative forcing impact of sulfur. We obtained the effective radiative forcing effect in Table 3.1 from the authors in personal correspondence.

[8] The function F_t can also be calibrated well to data from other studies (see Appendix B.1.1).

Table 3.2: Estimated forcing parameters

f_0	f_1	f_2	f_3	n
0.254	1.16	0.014	0.46	0.69

In Appendix B.1.1, we show radiative forcing resulting from a given annual flow of sulfur injections at a given atmospheric carbon dioxide concentration in a 3D graph. Figure 3.1 illustrates the goodness of our fit, slicing the 3D graph in the two dimensions and adding the data points from Kleinschmitt et al. (2018). We assume that total radiative forcing remains positive (above preindustrial levels). Based on our empirical fit of the radiative forcing equation we take the following assumption.

Assumption 3.1. *The (fit-)parameters* f_i, $i \in \{0, ..., 3\}$, *are positive and* $0 < n < 1$. *Radiative forcing remains above the preindustrial level,* $F_t > 0$, *and sulfur injections are between 2 and 50 TgS.*

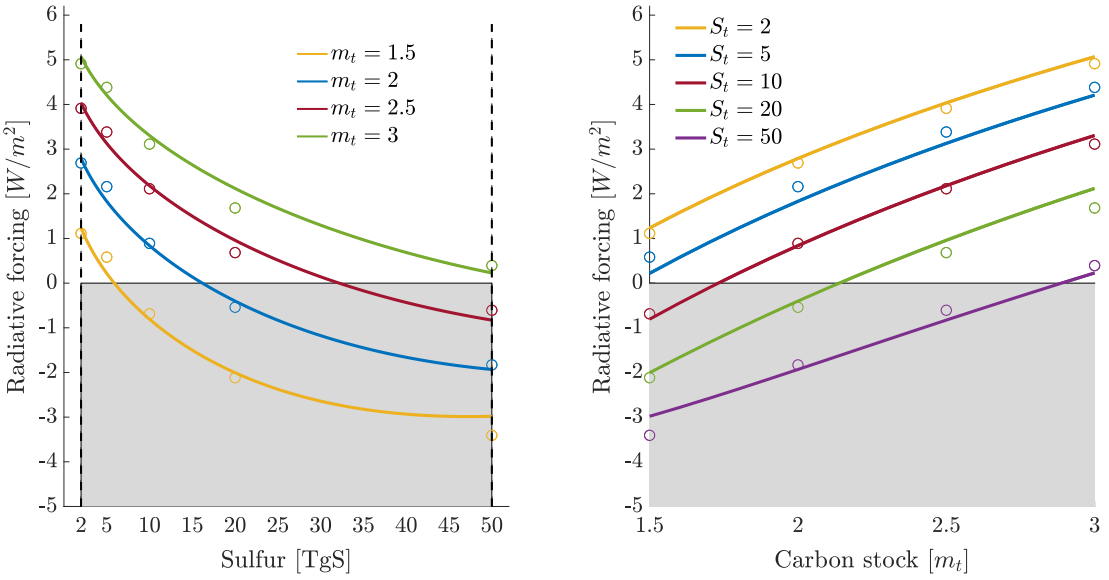

Figure 3.1: Approximation of radiative forcing F_t to model data (shown by circles) from Kleinschmitt et al. (2018) for sulfur injections between 2 and 50 TgS, and positive radiative forcing levels (see Assumption 3.1). The left graph shows radiative forcing as a function of sulfur for different atmospheric carbon concentrations. The right graph shows radiative forcing as a function of the atmospheric carbon concentration for different sulfur injection rates.

Operational costs and damages of geoengineering. Table 3.3 shows recent cost estimates on stratospheric sulfur injections by newly designed airplanes. Estimates are given for a reduction in radiative forcing (W/m^2) or the quantity of sulfur injected into the stratosphere (Mt). Stars denote the original values from the study. The values in Table 3.3 suggest average operational costs of $d = 0.0017\%$.[9]

Table 3.3: Annual operational costs of stratospheric sulfur injections

Authors	**Estimate**					
Klepper and Rickels (2012)	billion USD	2-18	for	-1* W/m^2	$\approx$ 2	TgS
Moriyama et al. (2017)	billion USD	10	for	-2* W/m^2	$\approx$ 7	TgS
McClellan et al. (2012)	billion USD	1-3	for		1*	TgS
	billion USD	2-8	for		5*	TgS
Smith and Wagner (2018)	billion USD	1.5	for		1*	TgS

Assessments of the economic impacts from geoengineering (e.g. from acid rain) suffer from insufficient observation. Some authors, for example Moreno-Cruz and Keith (2013), therefore analyze optimal policy as a function of the damage parameter while others make explicit assumptions, acknowledging that there is limited or non-existent empirical bases. We show several of those estimates in Table 3.4. Heutel et al. (2018) assume a cost of 3% of world output for resetting radiative forcing to its preindustrial level, independent of the prevailing forcing level. In general, it is more expensive cooling the planet down to preindustrial forcing levels when CO_2 concentrations are higher. We interpret their costs as the average of neutralizing the forcing of carbon concentrations of $m \in \{1.5, 2, 2.5, 3\}$ and find an approximate cost guesstimate of $d_H \approx 0.21\%$.[10] Emmerling and Tavoni's (2018b)

[9] At an annual world output of 135 trillion USD (purchasing power parity, IMF 2018), a deployment cost of 1 billion USD translates into a fractional output cost of $d = 7.4 \times 10^{-4}\,\%$. In compiling the average, we give equal weight to authors (putting only half the weight on each of McClellan et al.'s (2012) estimates).

[10] To translate the value from Heutel et al. (2018) into our model, we denote by S_m^{pre} the sulfur levels required to neutralize the antropogenic forcing of carbon concentrations of $m \in \{1.5, 2, 2.5, 3\}$. This range of carbon concentrations corresponds approximately to the concentrations along the simulated paths in Heutel et al. (2018). Then a damage of 3% of output implies $d_H = \frac{1}{4} \sum_{m=1.5}^{3} \frac{0.03}{S_m^{pre}} \approx 0.21\%$.

guesstimate is for a forcing reduction of 3.5 W/m^2, independent of the prevailing forcing level. Our model captures a decreasing efficiency of sulfur deployment, and a 3.5 W/m^2 reduction at low levels of cooling implies lower sulfur injections than the same forcing reduction at an already high level of cooling. Using an intermediate value, we convert Emmerling and Tavoni's (2018b) guesstimate into a damage of 0.1%.[11] The doubling of CO_2 forcing in Goes et al. (2011) is similar to the forcing increase of 3.5 W/m^2 and the range of damages of 0-5% translate into a range for the damage parameter of 0-0.17%.

Table 3.4: Damages from solar geoengineering

Authors	**Best guess**
Emmerling and Tavoni (2018b)	Consumption loss of 3% compensating each 3.5 W/m^2 of forcing
Heutel et al. (2018)	GDP loss of 3% for setting forcing back to the preindustrial level
Goes et al. (2011)	GDP loss between 0 and 5% per forcing equivalent to a doubling CO_2 forcing

3.2.4 Global planner solution

In the present section, a social planner maximizes the infinite stream of consumption flows

$$\max_{C_t, E_t, S_t} \sum_{t=0}^{\infty} \beta^t \log(C_t) \tag{3.12}$$

subject to model equations (3.1)-(3.11) and Assumption 3.1. The parameter β denotes the utility discount factor (pure time preference). In the regional model, each region follows the analogous objective for their own region's welfare (no trade). In Appendix B.1.2 we solve the inter-temporal optimization problem and derive the global optimal level of sulfur deployment.

[11] Figure 3.1 shows that at low initial cooling level, 25 TgS reduce cooling e.g. along the red curve by approximately 3.5 W/m^2. In contrast, reducing forcing by 3.5 W/m^2 along the green curve (triple preindustrial concentration) requires over 40 TgS. Using a value of 30 TgS, we obtain the damage guesstimate of $d_E = 0.1\%$.

Proposition 3.1. *The optimal level of sulfur deployment is*

$$S_t^* = \left(\frac{(1-n)\,\gamma\, f_3}{d + \gamma\, f_2} \right)^{\frac{1}{n}} m_t \tag{3.13}$$

with climate change impact $\gamma = \beta\, \xi_0\, \tilde{\sigma}_{11}\, \sigma_{forc}$ *and temperature dynamics contribution* $\tilde{\sigma}_{11} = [(\mathbf{1} - \beta\, \boldsymbol{\sigma})^{-1}]_{11}$*, where* $[\cdot]_{11}$ *denotes the first element of the inverted matrix in square brackets.*

Proof. See Appendix B.1.2.2. □

The optimal level of sulfur deployment increases linearly in the atmospheric carbon concentration. We refer to the proportionality factor

$$z \equiv \left(\frac{(1-n)\,\gamma\, f_3}{d + \gamma\, f_2} \right)^{\frac{1}{n}} \tag{3.14}$$

as the *geoengineering propensity*. It will reappear in the strategic setting and characterizes the drivers and moderators of the cooling effort given the atmospheric carbon concentration. This cooling propensity *increases* in the discount factor β, the temperature damage coefficient ξ_0, and the sulfur efficiency f_3. Sulfur deployment *decreases* in geoengineering damages d and the non-linear efficiency loss of sulfur cooling n.[12]

Using the fit parameters from Table 3.2 and the parameter values from the baseline calibration of the ACE model (see Appendix B.1.2.2) we find the optimal sulfur deployment level

$$S_t^* = \left(\frac{1.65}{16\% + 10^3 d} \right)^{1.45} m_t, \tag{3.15}$$

as a function of the damage (semi-)elasticity of sulfur, and the atmospheric carbon concentration (relative to preindustrial levels).

[12] $\frac{d}{dn} \log\left(\frac{(1-n)\,\gamma\, f_3}{d+\gamma\, f_2} \right)^{\frac{1}{n}} = -\frac{\log\left(\frac{(1-n)\,\gamma\, f_3}{d+\gamma\, f_2} \right)}{n^2} - \frac{1}{n(1-n)} < 0$ since Assumption 3.1 requires $S_t \geq 2$ for $m_t \in [1.5, 3]$, and thus $z^n = \frac{(1-n)\,\gamma\, f_3}{d+\gamma\, f_2} \geq 1.22$.

The calibrated formula shows that the optimal sulfur deployment is extremely sensitive to damages from geoengineering. The left graph in Figure 3.2 shows the optimal level of sulfur as a function of the damage parameter d for different carbon stocks. Inserting S_t^* into equation (3.11) yields the optimal level of radiative forcing as a function of the damage parameter d and the atmospheric carbon stock, which we show in the right graph of Figure 3.2. The higher the damages, the higher the tolerated forcing and, thus, warming levels.

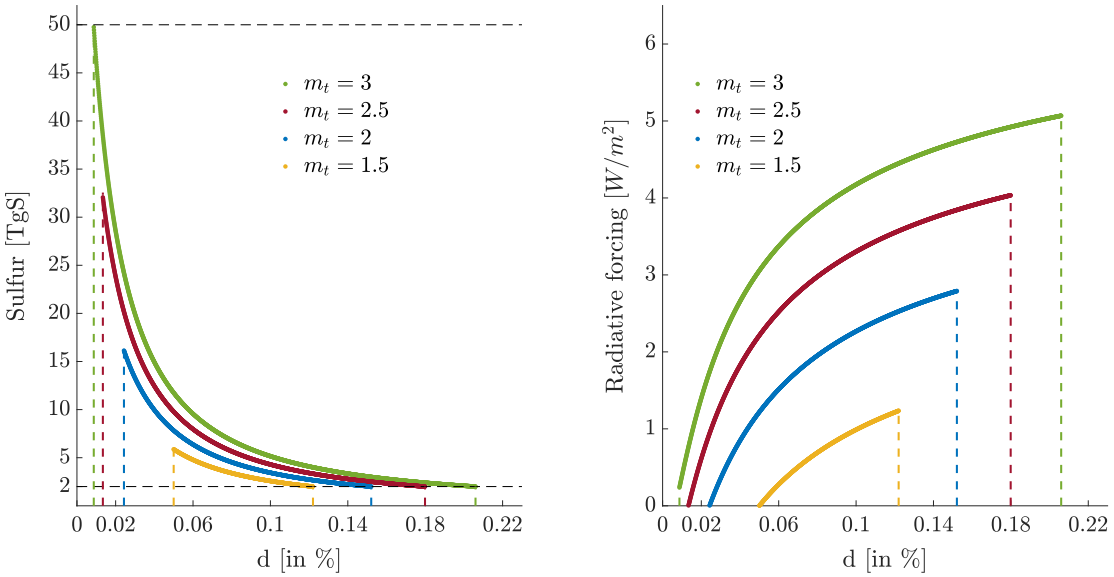

Figure 3.2: The graph shows the optimal level of sulfur (left) and the optimal level of radiative forcing (right) as a function of the damage parameter d (in %) for different atmospheric carbon stocks m_t.

We restrict our model to "well-calibrated" intervals. These intervals depend on the carbon stock $m_t \in [1.5, 3]$ and result from Assumption 3.1 restricting sulfur levels between 2 and 50 TgS and requiring a positive radiative forcing level. For lower than the depicted damage levels either $S_t > 50$ TgS or $F_t < 0$. For higher damages it must be that $S_t < 2$.

We now study the SCC and, thus, optimal carbon tax in the presence of solar geoengineering. The SCC reflects the long-term damage from releasing a marginal ton of CO_2 into the atmosphere. Proposition 3.1 shows that an increase of atmo-

spheric carbon increases the level of geoengineering. As a result, geoengineering partially mitigates the damages from the release of a ton of CO_2.

Proposition 3.2. *The SCC in money-measured consumption equivalents is*

$$SCC = \frac{Y_t^{net}}{M_{pre}} \left[a + f_1 \gamma - \left(\left(\frac{f_3}{z^n} - f_2 \right) \gamma - d \right) z \right] \tilde{\phi}_{11} \tag{3.16}$$

with the carbon dynamics contribution $\tilde{\phi}_{11} = \left[(\mathbf{1} - \beta \, \mathbf{\Phi})^{-1} \right]_{11}$, *where* $[\cdot]_{11}$ *denotes the first element of the inverted matrix in square brackets. As defined in connection to Proposition 3.1, the geoengineering propensity is* $z = \left(\frac{(1-n)\gamma f_3}{d+\gamma f_2} \right)^{\frac{1}{n}}$ *and the climate change impact is* $\gamma = \beta \, \xi_0 \, \tilde{\sigma}_{11} \, \sigma_{forc}$.

Proof. See Appendix B.1.2.3. □

The fraction $\frac{Y_t^{net}}{M_{pre}}$ sets the scale and units of the SCC. The square brackets characterize net damages and the term $\tilde{\phi}_{11}$ amplifies the SCC as a result of the long life-time of atmospheric CO_2 (carbon cycle).[13] Earlier analytic integrated assessment models like ACE only contain a term corresponding to our $f_1\gamma$ reflecting the cost resulting from a temperature increase in the absence of climate engineering. First, formula (3.16) adds the term a reflecting the direct net damages from atmospheric CO_2 caused by ocean-acidification net of the land-based fertilization effect. Second, it introduces the term in round brackets, which reduces the SCC as a result of geoengineering (the bracket is always positive). This reduction of the SCC increases in the level of geoengineering, but at a falling rate; the geoengineering propensity z in the denominator of $\frac{f_3}{z^n}$ reflects that the effectiveness of sulfur-based cooling decreases in the level of geoengineering. This reduction of the SCC increases in the effectiveness of geoengineering f_3. Third, the damages from geoengineering d add to the SCC, and more so the higher the sulfur deployment per unit of carbon (geoengineering propensity z).

[13] See Traeger (2018) for a detailed interpretation of the term $\tilde{\phi}_{11} = \left[(\mathbf{1} - \beta \, \mathbf{\Phi})^{-1} \right]_{11}$. Instead of a simple decay, it captures how much carbon inserted into the atmosphere remains in or returns to the atmosphere over the discounted infinite time horizon. A similar interpretation applies to the heat flow contribution $\tilde{\sigma}_{11}$.

The SCC's composition in Proposition (3.2) explains how the SCC responds to the geoengineering propensity. Yet, the geoengineering propensity is itself a function of damages and the sulfur's cooling effectiveness. Breaking the geoengineering propensity down into its various contributions, we find

$$\left(\left(\frac{f_3}{z^n} - f_2\right)\gamma - d\right)z = n\frac{(1-n)^{\frac{1-n}{n}}\ (\gamma\, f_3)^{\frac{1}{n}}}{(d+\gamma\, f_2)^{\frac{1-n}{n}}}. \tag{3.17}$$

The qualitative dependence of this SCC-reduction on damages d and cooling effectiveness f_3 remains as discussed above. We observe that *the SCC-reduction is less responsive to geoengineering damages than is the geoengineering propensity* z (comparing equations 3.14 and 3.17 the power $\frac{1-n}{n} < \frac{1}{n}$ as $0 < n < 1$). Lower damages increase the geoengineering propensity and, thus, the cooling response to atmospheric carbon. As a result, the marginal direct impact of releasing carbon decreases, but the indirect impact by triggering a larger deployment of sulfur partially offsets this benefit from the reduced temperature impact.

Quantifying the SCC, we follow ACE using a time step of 10 years and the parameter specification summarized in Table 3.5.[14] Together with our radiative forcing estimates from Table 3.2 we obtain the optimal carbon tax in (USD-2018-) money-measured consumption equivalents as a function of the damage parameters.

Table 3.5: Parameter values from ACE re-calibrated for 2 temperature layers

Y_t^{net}	M_{pre}	β	ξ_0	σ_{forc}	$\tilde{\sigma}_{11}$	$\tilde{\phi}$
135×10^{13}	$3.667 \times 600 \times 10^9$	0.986^{10}	0.021	0.52	1.22	4.26

Figure 3.3 graphs this SCC as a function of the geoengineering damages d for a given damage parameter $a = 0$ (no direct damages from an increase in the atmospheric carbon stock) and $a = 0.1\%$.

[14] We re-calibrated ACE's temperature dynamics to use two rather than three temperature layers. This common simplification hardly affects the model's ability to replicate the temperature dynamics of scientific climate models and substantially eases the presentation of the regional model in Section 3.3. The calibration follows the same method as in Traeger (2018).

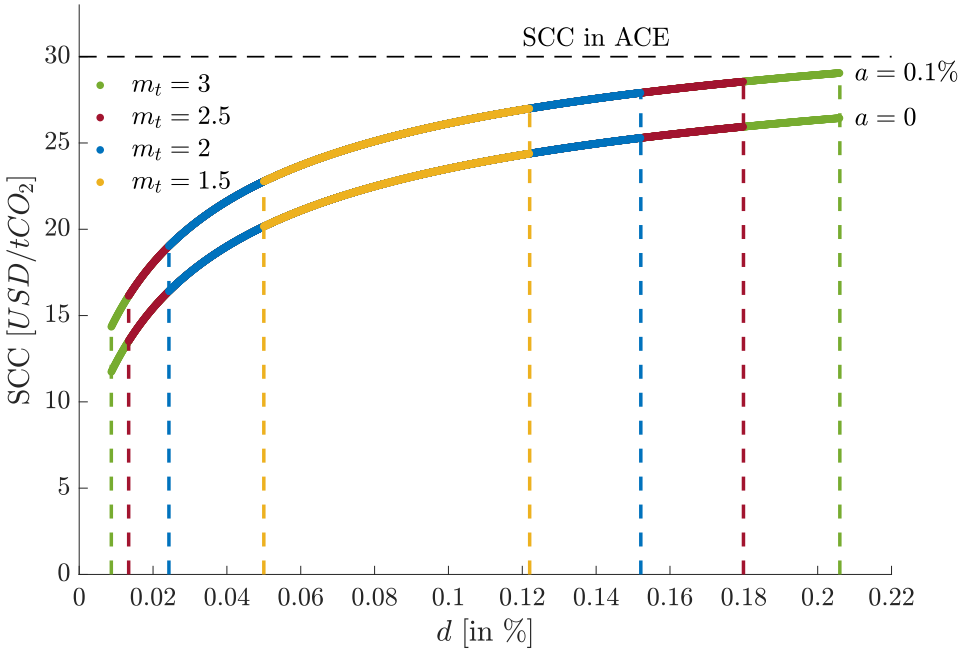

Figure 3.3: The graph shows the global SCC (in USD/tCO_2) as a function of the geoengineering damage parameter d (in %) without ocean acidification damages, $a = 0$, and with damages of $a = 0.1\%$, and compares it to the SCC in ACE.

3.3 Non-cooperative regions

We now split the world into regions that act non-cooperatively. We focus on two active regions A and B that consider engaging in climate engineering, either sulfur-based cooling or counter-geoengineering. The rest of the world only affects the decisions of regions A and B through their contributions to the global carbon stock.

3.3.1 Regional economies and climate dynamics

This section explains the changes required to split up the world in several economic and climate regions. It also introduces more detailed damage specifications and the option to engage in counter-geoengineering.

Regional economies and emissions. The regional economies follow equations (3.1)-(3.8), where functional forms and parameters are idiosyncratic to the regions. CO_2 mixes globally and the CO_2 dynamics are still described by equa-

tion (3.9). However, the total CO_2 emissions are now resulting from region A, region B, and the rest of the world W:

$$E_t^{tot} = \sum_{i=1}^{I^d} E_{A,i,t} + \sum_{i=1}^{I^d} E_{B,i,t} + \sum_{i=1}^{I^d} E_{W,i,t} + E_t^{\text{exo}}$$

Regional temperature dynamics. We characterize two climate zones by temperature levels $T_{1,t}^A$ and $T_{1,t}^B$, containing the two active regions. For simplicity, we assume that the rest of the world is part of region B's temperature zone. Analogously to the global case, we define transformed temperatures $\tau_{1,t}^A = \exp(\xi_1 T_{1,t}^A)$, $\tau_{1,t}^B = \exp(\xi_1 T_{1,t}^B)$, and $\tau_{2,t} = \exp(\xi_1 T_{2,t})$ with $\xi_1 = \frac{\log 2}{s}$. Then the regional temperature dynamics evolve according to

$$\begin{pmatrix} \tau_{1,t+1}^A \\ \tau_{1,t+1}^B \\ \tau_{2,t+1} \end{pmatrix} = \underbrace{\begin{pmatrix} \sigma_A^A & \sigma_B^A & \sigma_O^A \\ \sigma_A^B & \sigma_B^B & \sigma_O^B \\ \sigma_A^O & \sigma_B^O & \sigma_O^O \end{pmatrix}}_{\equiv \boldsymbol{\sigma}^A} \begin{pmatrix} \tau_{1,t}^A \\ \tau_{1,t}^B \\ \tau_{2,t} \end{pmatrix} + \begin{pmatrix} \sigma_{\text{forc}} \exp(\frac{\log 2}{\eta} F_t^A) \\ \sigma_{\text{forc}} \exp(\frac{\log 2}{\eta} F_t^B) \\ 0 \end{pmatrix},$$

where $\sigma_A^A = 1 - \sigma_B^A - \sigma_O^A - \sigma_{\text{forc}}$, $\sigma_B^B = 1 - \sigma_A^B - \sigma_O^B - \sigma_{\text{forc}}$, and $\sigma_O^O = 1 - \sigma_A^O - \sigma_B^O$. To preserve symmetry in notation, we define the corresponding matrix $\boldsymbol{\sigma}^B$ by swapping the first and second rows and columns, characterizing the identical dynamics from the perspective of region B. In anticipation of a similar climate impact as in the global case, we define

$$\tilde{\sigma}_{ij}^A = \left[(\mathbf{1} - \beta^A \boldsymbol{\sigma}^{\boldsymbol{A}})^{-1}\right]_{ij} \text{ and } \tilde{\sigma}_{ij}^B = \left[(\mathbf{1} - \beta^B \boldsymbol{\sigma}^{\boldsymbol{B}})^{-1}\right]_{ij} \text{ for } i, j \in \{1, 2\}.$$

The term $\tilde{\sigma}_{11}^A$ characterizes the discounted heat increase in region A over the infinite time horizon resulting from a heat influx into region A's atmosphere in the present, and $\tilde{\sigma}_{12}^A$ characterizes the discounted heat increase in region A over the infinite time horizon resulting from an influx into region B's atmosphere in the present.[15]

[15]We obtain this interpretation from expanding the inverse $(\mathbf{1} - \beta^A \boldsymbol{\sigma}^{\boldsymbol{A}})^{-1} = \sum_{l=0}^{\infty} \beta^{A^l} \boldsymbol{\sigma}^{\boldsymbol{A}^l}$

We spell out the dynamics in terms of actual temperatures $T^A_{1,t}$, $T^B_{1,t}$, and $T_{2,t}$ in Appendix B.1.3.1.

Regional forcing. Radiative forcing in region A is a function of global atmospheric carbon and the geoengineering undertaken in the two regions. We let $\alpha_B\, S^B_t$ and $\alpha_A\, S^A_t$ denote the spill-over of the cooling or the counter-geoengineering agent from one region to the other.[16] The parameter $\alpha_A < 1$ ($\alpha_B < 1$) determines the share of region A's (B's) injection of the cooling or counter-geoengineering agent that travels to region B (A). The magnitude of the α parameter depends in particular on the geographic location of the two regions. For example, the α parameters will be relatively low if one region is located on the northern and the other region on the southern hemisphere (e.g. the US and Brazil). It will be close to unity if both regions are located on the same hemisphere and at similar latitude (e.g. Europe and North America). It will be asymmetric if one region lies North of the other on the same hemisphere (e.g. Canada would get perfect spill-over from the US, but the US much less spill-over from Canada). Then we have the following regional forcing levels:

$$F^A_t = \frac{\eta}{\log(2)}\, \log\left[f_0 + f_1\, m_t + \left(f_2 - f_3\, \left(\frac{m_t}{(S^A_t + \alpha_B\, S^B_t)}\right)^n\right)(S^A_t + \alpha_B\, S^B_t)\right]$$

$$F^B_t = \frac{\eta}{\log(2)}\, \log\left[f_0 + f_1\, m_t + \left(f_2 - f_3\, \left(\frac{m_t}{(S^B_t + \alpha_A\, S^A_t)}\right)^n\right)(S^B_t + \alpha_A\, S^A_t)\right].$$

We assume that a counter-measure ($S_t < 0$) exists and can be used to neutralize

using the von Neumann series. The entry i-j of the l^{th} power of the transition matrix σ^A, i.e. $\left[\boldsymbol{\sigma^A}^l\right]_{ij}$, characterizes how much of the heat flow entering temperature layer j in the present still remains in or returns to layer i after l periods.

[16] The cooling potential of sulfur does not only depend on the injection rate but also on the location of the injections. If sulfur particles are injected in the tropics, they spread effectively towards the poles (Lawrence et al., 2018). As a consequence, it is impossible to do regional climate management using stratospheric geoengineering without spillovers to the other regions. Recent studies suggest that it might be possible to optimize the geographic distribution of the cooling by varying the altitude, latitude and season of injections (Visioni et al., 2019; Dai et al., 2018; Jones et al., 2018; MacMartin et al., 2017; Kravitz et al., 2017).

(part of) the cooling imposed by the other region (counter-geoengineering). Based on our radiative forcing approximation and Assumption 3.1 from the global model we assume 2 TgS $\leq S_t^A + \alpha_B\, S_t^B \leq 50$ TgS, and 2 TgS $\leq S_t^B + \alpha_A\, S_t^A \leq 50$ TgS.

Regional damages and counter-geoengineering. Geoengineering creates damages and operational costs which, for region A, we summarize in d_{AA} for the marginal costs of the region's own action, and d_{BA} for the marginal damages imposed on region A by region B. Damages as a fraction of output in region A are given by

$$D_t^A(\tau_{1,t}^A, S_t, m_t) = 1 - \exp\left[\xi_0^A\, \left(1 - \tau_{1,t}^A\right) - (d_{AA} S_t^A + d_{BA}\, \alpha_B\, S_t^B) - a^A\, (m_t - 1)\right].$$

Region A's self-imposed marginal costs d_{AA} depend on whether it is cooling or conducting a neutralizing measure

$$d_{AA}(S_t^A) = \begin{cases} d_{AA}^g + \epsilon_A^g & \text{for} \quad S_t^A > 0 \\ d_{AA}^c - \epsilon_A^c & \text{for} \quad S_t^A < 0 \\ 0 & \text{for} \quad S_t^A = 0 \end{cases}$$

where d_{AA}^g is the damage from sulfur-based cooling and $\epsilon_A^g > 0$ is the cost of injecting the sulfur into the stratosphere. The parameter d_{AA}^c characterizes the damage *reduction* (noting that $S_t^A < 0$) from employing counter-geoengineering, and $\epsilon_A^c \geq 0$ is the cost of counter-geoengineering. For the damages imposed by region B onto region A we distinguish whether region B engages in sulfur-based cooling or counter-geoengineering

$$d_{BA}(S_t^B) = \begin{cases} d_{BA}^g & \text{for} \quad \alpha_B\, S_t^B > 0 \\ d_{BA}^c & \text{for} \quad \alpha_B\, S_t^B < 0 \\ 0 & \text{for} \quad \alpha_B\, S_t^B = 0 \end{cases}$$

where d_{BA}^c is again a partial offsetting of the damages from region A's climate engi-

neering in case region B is countering it. While the direct radiative forcing damage will be offset, the damages caused by the chemical agent sulfur will probably not be offset, only partially offset, or maybe even enhanced by the counter-geoengineering agent. Thus d^c_{AA} will generally be strictly lower than d^g_{AA} in the real world (and it could potentially even be negative). Therefore, we restrict the damage parameters in the active regions as follows.

Assumption 3.2. *The damage relief from counter-geoengineering is smaller than the damage caused by geoengineering:* $d^c_{ij} \leq d^g_{il}$ *for* $i, j, l \in \{A, B\}$.

By symmetry we obtain the same damage definitions for region B (see Appendix B.1.3.1). We note that Assumption 3.2 combines necessary assumptions to guarantee a unique optimal response ($d^c_{AA} \leq d^g_{AA}$ and $d^c_{BB} \leq d^g_{BB}$), and assumptions that we only impose because they seem economically sensible.

3.3.2 Results of the base model

For ease of presentation, the present section turns off the direct heat transfer between the regions.

Assumption 3.3. *The heat flow coefficients* $\sigma^A_B, \sigma^B_A, \sigma^O_A$, *and* σ^O_B *are zero.*

As a result, the regional climates interact only through the spill-over of the cooling and, potentially, counter-geoengineering agents. The assumption simplifies the functional expressions without changing the qualitative results as we show in section 3.3.4 where we relax Assumption 3.3.

The section identifies a set of equilibrium strategies, characterizes the resulting subgame-perfect Nash equilibria and derives corresponding formulas for the resulting non-cooperative SCC levels in the different regions. Appendix B.1.3 solves the corresponding dynamic Markov game. In every period, the regions control their sulfur (and CO_2) emissions optimally, anticipating the future reaction of the other

region to its own actions; the solution is sometimes referred to as a feedback equilibrium (as opposed to an open-loop equilibrium).

Proposition 3.3 (Strategies)**.** *The following strategies characterize a Markov perfect Nash equilibrium of the dynamic game. If (i)* $S_t^B = 0$ *region A chooses* $S_t^A(m_t) = z_A^g m_t$ *and if (ii)* $S_t^B \neq 0$ *region A chooses*

$$\begin{aligned} S_t^A(m_t) &= \frac{m_t}{1-\alpha_A\,\alpha_B}\left(z_A^g - \alpha_B\, z_B\right) \quad \textit{for} \quad S_t^A > 0 \\ S_t^A(m_t) &= \frac{m_t}{1-\alpha_A\,\alpha_B}\left(z_A^c - \alpha_B\, z_B^g\right) \quad \textit{for} \quad S_t^A < 0 \\ S_t^A &= 0 \quad \textit{otherwise} \end{aligned} \tag{3.18}$$

with geoengineering propensity and counter-geoengineering reluctance

$$z_A^g = \left(\frac{(1-n)\,f_3\,\gamma_A}{f_2\,\gamma_A + (d_{AA}^g + \epsilon_A^g)}\right)^{\frac{1}{n}}, \quad z_A^c = \left(\frac{(1-n)\,f_3\,\gamma_A}{f_2\,\gamma_A + (d_{AA}^c - \epsilon_A^c)}\right)^{\frac{1}{n}}$$

and climate change impact $\gamma_A = \beta^A\,\xi_0^A\,\tilde{\sigma}_{11}^A\,\sigma_{forc}$*. By Assumption 3.2,* $z_A^g < z_A^c$*. Swapping region indices characterizes region B's strategies. Note that in equation* (3.18) $z_B \in \{z_B^g, z_B^c\}$ *depending on whether* $S_t^B \gtrless 0$*.*

Proof. See Appendix B.1.3.6. □

The game solves in linear strategies. This linear dependence on the carbon stock not only permits an analytic solution of the game, but also coincides with the structure of the (unique) optimal CO_2 response in a non-strategic model (see Proposition 3.1).

One active region. In the case where one of the regions remains inactive ($S_t^B = 0$), the other region's optimal cooling effort is structurally the same as in the social planner setting (Proposition 3.2); sulfur deployment increases proportional to the atmospheric carbon concentration and to the geoengineering propensity z_A^g. As in the social planner setting, this cooling propensity *increases* in the climate change impact components summarized in γ_A and the sulfur efficiency f_3. Sulfur

deployment *decreases* in geoengineering damages and the non-linear efficiency loss n of sulfur cooling. In difference to the social planner, the active region only accounts for its own climate impact γ_A and for its own damages d^g_{AA} and costs ϵ^g_A from geoengineering (damages and costs were combined into a single term d in the social planner's problem).

Both regions cooling. In the case where both regions engage in cooling, the strategic interaction adds two new aspects to sulfur deployment. First, a region acknowledges the other region's contribution and reduces its own sulfur deployment accordingly. In equation (3.18), this direct response subtracts the spill-over-weighted geoengineering propensity $\alpha_B z^g_B$ of the other region from its original geoengineering propensity z^g_A. Second, the region anticipates that also the other region will respond to its own action. Because of the linear response functions, this higher order reaction leads to the multiplier $\frac{1}{1-\alpha_A\alpha_B}$; it scales up region A's action as a result of its anticipation that region B reduces its sulfur deployment ("free-rides") in response to A's action. This higher order response counteracts the initial free-riding incentive. The higher order response increasing the sulfur deployment is always lower in magnitude than the initial decrease.[17] *In summary, with both regions cooling, each region's sulfur deployment (i) decreases strategically as a result the joint action (or "free-riding"), but this decrease is (ii) partially offset by a region's anticipation of the other region's cooling reduction (or "free-riding").*

Counter-geoengineering. In the case where the regions' interests clash, one region, say region B, is cooling. Region B's "excessive" geoengineering propensity z^g_B spilling over into region A drives region A to engage in counter-geoengineering

[17]The finding is less obvious at second thought. The initial free-riding incentive grows in the other region's geoengineering propensity, whereas a region's anticipation of the other regions free-riding is (analogously) driven by its own geoengineering propensity. In principle, a region's anticipation effect can therefore dominate the initial free-riding effect as the following calculation shows: $\frac{z^g_A-\alpha_B z^g_B}{1-\alpha_A\alpha_B}m_t > z^g_A m_t \Leftrightarrow z^g_A - \alpha_B z^g_B > z^g_A - \alpha_A\alpha_B z^g_A \Leftrightarrow z^g_B < \alpha_A z^g_A$. However, if region A's spill-over weighted geoengineering propensity indeed dominates region B's own geoengineering propensity ($\alpha_A z^g_A > z^g_B$) the two regions would be in an equilibrium where only region A is cooling and region B free-rides (or counter-geoengineers) from the start as we will establish in Proposition 3.4.

($S_t^A < 0$). A higher z_A^c still states that region A would prefer relatively more cooling, but less than what region B is going for. Thus, we call z_A^c the counter-geoengineering reluctance (rather than the geoengineering propensity). A higher z_A^c pushes out the point where region A starts to engage in counter-geoengineering and, once it does counter-geoengineering, a higher z_A^c reduces its level. The sign in front of the deployment costs ϵ_A switches because reducing S_t^A now imposes deployment costs of the counter-geoengineering agent (rather than reducing sulfur deployment costs). *This discrete jump from reducing sulfur deployment costs to creating counter-geoengineering costs ensures a non-trivial (z_B^g-)interval where A simply remains inactive.*

In this climate-clash setting, the direct strategic response to the other region's cooling drives counter-geoengineering. The higher order response leads once again to the multiplier $\frac{1}{1-\alpha_A\alpha_B}$. Here, it scales up region A's action as a result of its anticipation that region B increases its sulfur deployment in response to A's counter-geoengineering. *In the "riding" metaphor, A presses down the gas pedal extra-hard because it anticipates that B presses down the gas pedal further when it realizes that A tries to pull the planet into the opposite direction. Yet, the outcome in our setting is more favorable than the metaphor suggests. Because spill-overs are incomplete ($\alpha_A\alpha_B < 1$), the simultaneous warming and cooling in the two regions brings both regions closer to their desired temperatures.*

Figure 3.4 shows the deployment of sulfur or the counter-geoengineering agent for region A and B as a function of region B's geoengineering propensity z_B^g by varying d_{BB}^g (and assuming $d_{BB}^c = 0.5\, d_{BB}^g$). To get a unique parameterization in terms of propensities, the graph shows z_B^g even if the chosen strategy is z_B^c. We assume a fixed geoengineering propensity and counter-geoengineering aversion for region A, and define $\Delta\, z_A = \alpha_A\, z_A^g - (z_B^c - z_B^g)$. In case each region's propensity to cool is higher than the spillovers it receives from the other region, $\alpha_A\, z_A^g < z_B^g < z_A^g\, \alpha_B^{-1}$, both regions deploy sulfur. If region B's geoengineering propensity is smaller

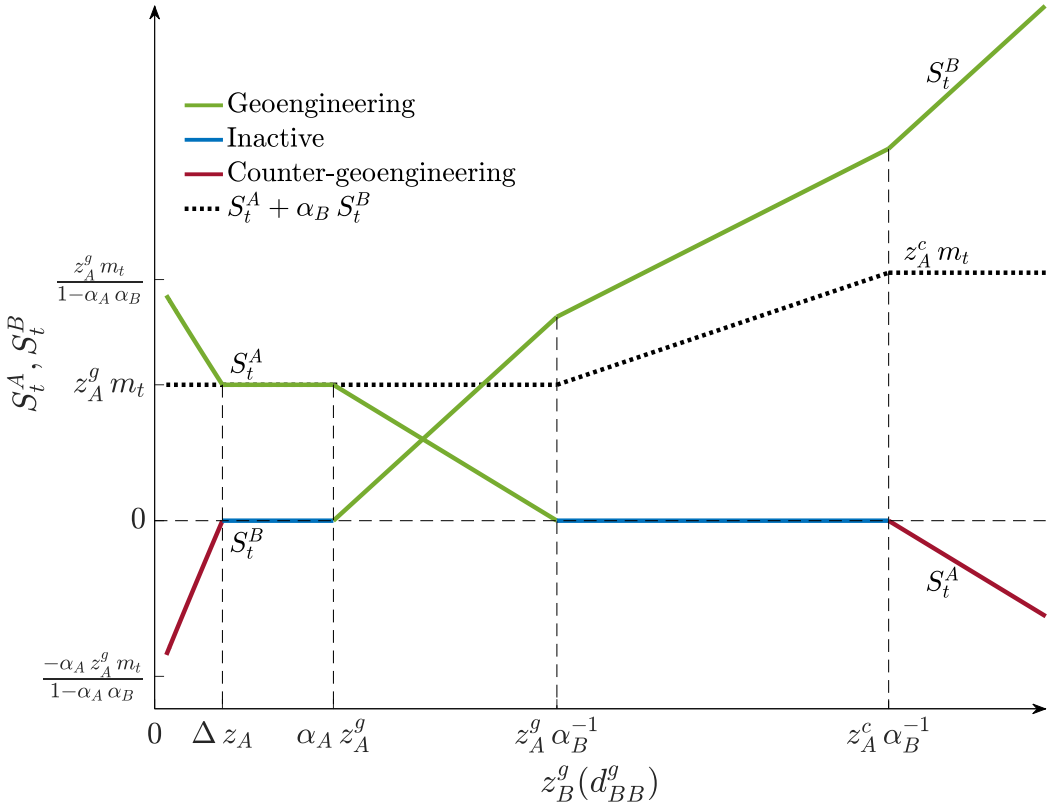

Figure 3.4: The graph shows the regional strategies as a function of region B's geoengineering propensity z_B^g by varying d_{BB}^g. Each region either deploys sulfur (shown in green), the neutralizing agent (shown in red), or stays inactive (shown in blue). The dotted black line depicts the target level of sulfur in region A ($S_t^A + \alpha_B\, S_t^B$). We defined $\Delta\, z_A = \alpha_A\, z_A^g - (z_B^c - z_B^g)$.

than the spillovers from region A, $z_B^g < \alpha_A\, z_A^g$, there are two possibilities: If region B's reluctance to engage in counter-geoengineering is low enough, and $z_B^g < \Delta\, z_A$, it will react by deploying the neutralizing agent. Otherwise, region B remains inactive. Similarly, the same happens if region A's geoengineering propensity is smaller than the spillovers it receives from region B, $z_B^g > z_A^g\, \alpha_B^{-1}$: Region B cools while region A either remains inactive (if its reluctance to do counter-geoengineering is too high, $z_B^g < z_A^c\, \alpha_B^{-1}$) or deploys the counter-geoengineering agent. As long as $z_B^g < z_A^g\, \alpha_B^{-1}$ region A keeps the (target) level of sulfur in its region (given by $S_t^A + \alpha_B\, S_t^B$) at the constant level $z_A^g\, m_t$. Since counter-geoengineering is not able to perfectly offset the damages from spillovers, for $z_B^g > z_A^g\, \alpha_B^{-1}$ region A increases its (target) level of

sulfur towards $z_A^c\, m_t$.

We now identify the parameter ranges that give rise to the different equilibria and show that they are mutually exclusive and cover the full parameter domain.

Proposition 3.4 (Equilibria)**.** *The strategies in Proposition 3.3 give rise to 5 qualitatively different Nash equilibria. They are mutually exclusive and classified based on fundamentals as follows:*

Climate clash	$S_t^A > 0, S_t^B < 0:$	$\alpha_A^{-1} < h$
Free-driver/rider	$S_t^A > 0, S_t^B = 0:$	$h \leq \alpha_A^{-1} \leq H$
Climate match	$S_t^A > 0, S_t^B > 0:$	$\alpha_B < H < \alpha_A^{-1}$
Free-driver/rider	$S_t^A = 0, S_t^B > 0:$	$H \leq \alpha_B \leq \hat{H}$
Climate clash	$S_t^A < 0, S_t^B > 0:$	$\hat{H} < \alpha_B$

where

$$h = \frac{z_A^g}{z_B^c} \quad , \quad H = \frac{z_A^g}{z_B^g} \quad , \text{ and } \quad \hat{H} = \frac{z_A^c}{z_B^g}.$$

We note that $h \leq H \leq \hat{H}$ *and that* $\alpha_B \leq \alpha_A^{-1}$.

Proof. See Appendix B.1.3.7. □

Two fully symmetric regions would desire the same geoengineering target, implying $H = 1$. They would be in the climate match equilibrium where both contribute equally to the overall cooling (assuming imperfect spill-overs $\alpha_A, \alpha_B < 1$). As region A's perceived damages from geoengineering increase, it's cooling costs become larger (relative to GDP), or its climate impacts would be lower, H falls along with its geoengineering propensity z_A^g. Initially, this fall merely implies that region A contributes less to the cooling, but eventually region A will become inactive ("free-ride"). This switch occurs when the spill-overs from region B dominate region A's own geoengineering propensity, $z_A^g \leq \alpha_B z_B^g$. Alternatively, one could call this scenario the free-driving of region B, which unilaterally sets the global temperature

distribution. However, if region A's perceived damages from geoengineering are even higher and/or climate damages even lower as compared to those of region B, region A will no longer tolerate the geoengineering level imposed by region B. Then, it is worthwhile for region A to pay the costs of counter-geoengineering. This shift in equilibria happens when the spillovers from region B not only surpass its geoengineering propensity, but also its (higher) counter-geoengineering reluctance, $z_A^c < \alpha_B z_B^g \ (\Leftrightarrow \hat{H} < \alpha_B)$.

The left graph in Figure 3.5 illustrates how the domain of the Nash equilibria shifts under an increase in the cost of geoengineering (damages or operational costs). Dashed lines show the domain before the change occurs, solid lines and color coding mark the domain of equilibria after the change. An increase in the cost of geoengineering for region A decreases its geoengineering propensity z_A^g. As a result, the domain where region A free-rides on region B expands (blue area shifting down together with H). At the same time, the domain where region B cools together with region A rather than B free-riding expands to the right, and so does the domain where region B free-rides rather than engaging in counter-geoengineering. These shifts to the right state that region B is not free-riding or engaging in counter-geoengineering even for higher spill-overs α_A from region A.

The right graph in Figure 3.5 illustrates how the domain of the Nash equilibria shift under an increase in the cost of counter-geoengineering (increase in operational costs or decrease in effectiveness). It increases region A's reluctance to engage in counter-geoengineering z_A^c, and expands the domain where region A stays inactive ("free-rides") and region B unilaterally sets the temperature distribution ("free-drives"). We note that, depending on the realization of h, H, and $\hat{H}$, the constraint that $\alpha_A, \alpha_B < 1$ will select a subdomain of the graphs shown in Figure 3.5 eliminating parts of the climate clash and free-driving domains by either reducing the domain from above or from the right.

The strategic interactions characterize how the regions try to increase or defend

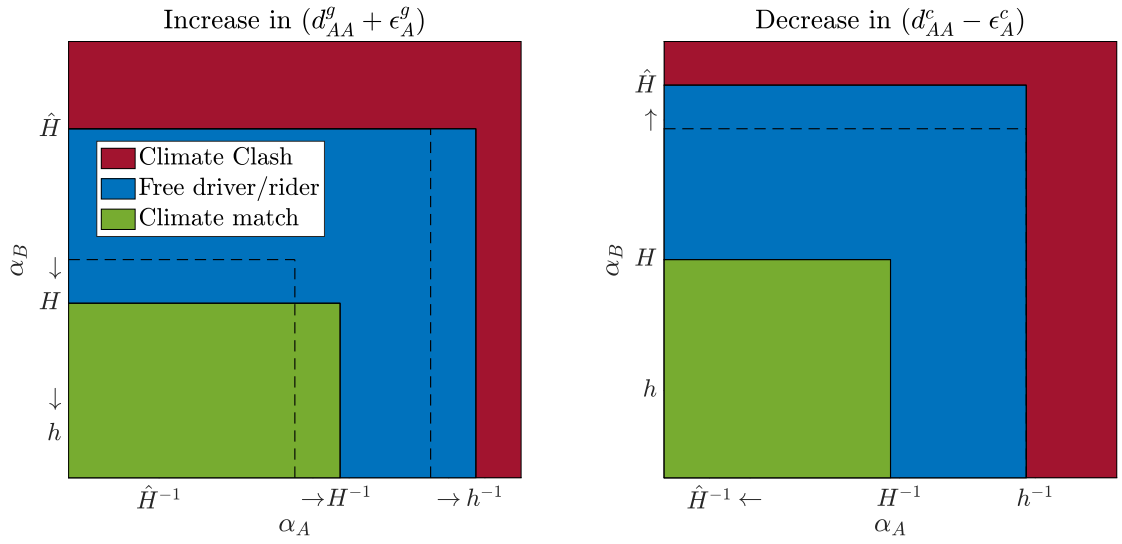

Figure 3.5: The graph shows how the domain of Nash equilibria is affected by a increase in the costs or damages of geoengineering (left graph), and an increase in the cost or a decrease in the effectiveness of counter-geoengineering (right graph) for region A.

their own well-being given the availability of geoengineering. We will now analyze the implications for a region's incentives to undertake classic climate change mitigation, the reduction of greenhouse gas emissions. These incentives are characterized by the regional SCCs, whose structure differs by type of equilibrium.

Proposition 3.5. *If $S^B_t = 0$, the SCC in region A is*

$$SCC^A = \frac{Y^{net}_{A,t}}{M_{pre}} \left[a^A + f_1\,\gamma_A - \left(\left(\frac{f_3}{(z^g_A)^n} - f_2 \right) \gamma_A - (d^g_{AA} + \epsilon^g_A) \right) z^g_A \right] \tilde{\phi}^A_{11}.$$

If region A itself is inactive ($S^A_t = 0$) its SCC is

$$SCC^A = \frac{Y^{net}_{A,t}}{M_{pre}} \left[a^A + f_1\,\gamma_A - \left(\left(\frac{f_3}{(\alpha_B\, z^g_B)^n} - f_2 \right) \gamma_A - d^g_{BA} \right) \alpha_B\, z^g_B \right] \tilde{\phi}^A_{11}.$$

If both regions are active ($S^B_t \neq 0$ and $S^A_t \neq 0$) the SCC is

$$SCC^A = \frac{Y^{net}_{A,t}}{M_{pre}} \left[a^A + f_1\,\gamma_A - \left(\left(\frac{f_3}{(z_A)^n} - f_2 \right) \gamma_A - d_{AA} \right) z_A - \frac{\alpha_B (z_B - \alpha_A\, z_A)(d_{AA} - d_{BA})}{1 - \alpha_A\, \alpha_B} \right] \tilde{\phi}^A_{11}$$

with $z_A \in \{z_A^g, z_A^c\}$, $z_B \in \{z_B^g, z_B^c\}$, $d_{AA} \in \{d_{AA}^g + \epsilon_A^g, d_{AA}^c - \epsilon_A^c\}$, *and* $d_{BA} \in \{d_{BA}^g, d_{BA}^c\}$ *depending on whether the corresponding region engages in geoengineering (g) or counter-geoengineering (c). Swapping region indices characterizes region B's SCC.*

Proof. See Appendix B.1.3.8. □

Before discussing the SCC-formulas in detail, we summarize some consequences in the following corollary.

Corollary 1. *The availability of geoengineering reduces the SCC of a unilaterally acting region (similar to the global model). In all other equilibria, geoengineering can either increase, decrease or leave the SCC unchanged relative to the regional world without geoengineering, depending on the heterogeneity of damages and the spillovers.*

We note that the corresponding regional SCC in the absence of geoengineering is given by the first two terms in square brackets (and coincides with setting the geoengineering propensities to zero).

Unilateral action. If region B is inactive ($S_t^B = 0$), region A's SCC has the same structure as in the global model. In difference to the social planner, region A only accounts for its own climate impact γ_A and for its own damages d_{AA}^g and costs ϵ_A^g. As in the global social planner case, *the availability of geoengineering always reduces the active region's SCC.*[18]

Inactive region. If region A is inactive ($S_t^A = 0$), the structure of its SCC resembles that of the previous case where it is acting unilaterally. However, the SCC is no longer driven by its own geoengineering specifiers z_A and d_{AA} but by

[18] Analogously to equation (3.17), the cooling contribution $\left(\left(\frac{f_3}{(z_A^g)^n} - f_2\right)\gamma_A - (d_{AA}^g + \epsilon_A^g)\right) z_A^g = n\frac{(1-n)^{\frac{1-n}{n}}\,(\gamma_A\, f_3)^{\frac{1}{n}}}{\left((d_{AA}^g+\epsilon_A^g)+\gamma_A\, f_2\right)^{\frac{1-n}{n}}} > 0$ always reduces the SCC.

the spillovers resulting from region B, the geoengineering propensity $\alpha_B\, z_B^g$ and the damages d_{BA}^g. For most regions, the damage spillovers from a ton of sulfur deployment in the other region are less than damages plus deployment costs from a ton of own geoengineering ($d_{BA}^g < d_{AA}^g + \epsilon_A^g$). Then, *also the inactive region's SCC falls for a sufficiently low geoengineering propensity of the active region*, i.e., as long as the marginal spillover damages are less than the marginal cooling benefits $d_{BA}^g < \left(\frac{f_3}{(\alpha_B\, z_B^g)^n} - f_2\right)\gamma_A$. *If the active region's geoengineering propensity becomes too high, region A's spillover damages dominate its cooling benefits. Then, its SCC increases under geoengineering because more CO_2 emissions trigger more of the (net-)damaging geoengineering.* Eventually region A will abandon its passive role and engage in counter-geoengineering.[19]

Both regions active: the spill-over term. If both regions are active ($S_t^A \neq 0$, and $S_t^B \neq 0$), an additional term enters region A's SCC. It results from the spill-overs between regions A and B and we call it the

$$\text{spill-over term:} \quad -\frac{\alpha_B(z_B - \alpha_A\, z_A)(d_{AA} - d_{BA})}{1 - \alpha_A\, \alpha_B} = -\alpha_B \frac{S_t^B(m_t)}{m_t}(d_{AA} - d_{BA}).$$

Climate-match. At the core of the spill-over term lie the *excess costs* $d_{AA} - d_{BA}$ *from region A's own as compared to region B's geoengineering action.* It captures the difference between the marginal damages and deployment costs ($d_{AA} = d_{AA}^g + \epsilon_A^g$) resulting from region A's own cooling as compared to the marginal damages resulting from geoengineering in the other region (d_{BA}^g). If damages are independent of where geoengineering is deployed, then excess costs equal the positive deployment costs ϵ_A^g; as compared to unilateral action, region A benefits from B's support

[19]Formally, the geoengineering terms in SCC^A result in a cooling if $\alpha_B z_B^g < \left(\frac{f_3\,\gamma_A}{f_2\,\gamma_A + d_{BA}^g}\right)^{\frac{1}{n}}$. The unilateral action equilibrium with $S_t^A = 0$ requires $z_A^g \leq \alpha_B z_B^g \leq z_A^c$. If $d_{BA}^g < d_{AA}^g + \epsilon_A^g$ then there exists a non-empty z_B^g-interval where $z_A^g \leq \alpha_B z_B^g < \left(\frac{f_3\,\gamma_A}{f_2\,\gamma_A + d_{BA}^g}\right)^{\frac{1}{n}}$.

We note that region A will move to the climate clash equilibrium before its SCC dominates that of a world without geoengineering if spillover damages are very high and counter-geoengineering is very cheap and effective, $d_{BA}^g < d_{AA}^c - \epsilon_A^c$.

in cooling the planet. The benefit from B's engagement grows with the spill-over weighted sulfur emissions of region B, $\alpha_B \frac{S_t^B(m_t)}{m_t}$, here expressed per unit of atmospheric carbon. CO_2 emissions increase the atmospheric carbon concentration and trigger additional geoengineering in both regions. The spillover term represents the benefits from the other region's cooling support, which reduces the costs of (partially) offsetting the damages from releasing a unit of CO_2. *In summary, possible cost and damage savings from the other region's cooling support can reduce the SCC. If the spillover damages are substantially larger than a region's self-imposed damages, e.g. because of the sulfur distribution and resulting acid rain patterns, then excess costs are negative and the spillover term can also increase the region's SCC. For extremely heterogenous damages, the spillover term can dominate the ("unilateral") cooling term, increasing the SCC over the one prevailing in a world without geoengineering, even in the climate match equilibrium.*

Climate-clash. In the case where region A is cooling and region B engages in counter-geoengineering, region B's deployment term $\frac{S_t^B(m_t)}{m_t}$ turns negative. Moreover, the own excess costs $d_{AA} - d_{BA} = d_{AA}^g + \epsilon_A^g - d_{BA}^c$ are positive because region B's counter-geoengineering can at most offset damages in region A (Assumption 3.2). As a result, the spill-over term always increases the SCC for the cooling country. In the case where region A uses counter-geoengineering and region B is cooling, $\frac{S_t^B(m_t)}{m_t}$ remains positive and the own excess costs $d_{AA} - d_{BA} = d_{AA}^c - \epsilon_A^c - d_{BA}^g$ turn negative as (now A's) counter-geoengineering can at most offset the damages. Again, the spillover term increases the SCC. Thus, *in a climate clash, the spillover always increases the SCC in both regions. The intuition is that the other region always interferes with what a region would like to do if it was acting alone. This interference grows stronger for higher CO_2 concentrations, thereby increasing the cost of emitting another ton of carbon.* Depending on whether the positive spill-over term dominates the reduction from cooling, the net effect on region A's SCC can be positive or negative, and therefore the SCC can be higher or lower compared to

a regional world without geoengineering.

Figure 3.6 illustrates region A's SCC across the different regions. On the horizontal axis, we increase region B's geoengineering propensity by varying d^g_{BB}. As in Figure 3.4, region A's geoengineering propensity and counter-geoengineering aversion are fixed. Again, we set $d^c_{BB} = 0.5\, d^g_{BB}$, and define $\Delta\, z_A = \alpha_A\, z^g_A - (z^c_B - z^g_B)$. Choosing a simple example we let $d^g_{AA} = d^g_{BA}$. If B's cooling propensity is sufficiently small ($z^g_B < \Delta\, z_A \Leftrightarrow z^c_B < \alpha_A\, z^g_A$), region A is cooling and region B uses counter-geoengineering. As a result, the spillover contribution in region A's SCC is positive; the clash between the regions grows with each unit of carbon emissions raising their social costs. As region B's reluctance to engage in counter-geoengineering increases ($z^c_B \to \alpha_A\, z^g_A \Leftrightarrow z^g_B \to \Delta\, z_A$), this spillover effect declines, implying that A's SCC falls. Once the counter-geoengineering reluctance of region B matches the spillover-weighted geoengineering propensity of region A ($z^c_B = \alpha_A\, z^g_A$), region B turns inactive ($S^B_t = 0$). Then, A's SCC remains constant because it is independent of region B's characteristics; region A does as it pleases with no interference from the other region, strategic interactions are absent. Once also region B's geoengineering propensity matches region A's spillover-weighted propensity ($z^g_B = \alpha_A\, z^g_A$), region B starts cooling. The spillover effect in region A's SCC turns negative ($d_{AA} - d_{BA} = \epsilon^g_A > 0$). As a result, region A's SCC decreases; the impact of an additional unit of carbon emitted by region A will increasingly be offset by region B's cooling (saving A the effort of sulfur deployment). When z^g_B approaches $z^g_A\, \alpha_B^{-1}$ region A becomes inactive. From this point onward, only region B cools. The SCC of region A continues to decline until $\alpha_B\, z^g_B = \left(\frac{f_3(1-n)\gamma_A}{f_2\gamma_A + d^g_{AA}}\right)^{\frac{1}{n}}$.[20] Afterwards, A's SCC begins to rise. Finally, when $z^g_B \to z^c_A\, \alpha_B^{-1}$ region A starts to engage in counter-geoengineering. The spillover term becomes positive and the SCC of region A increases further.

[20] $\frac{\partial\, SCC^A(S^A_t=0)}{\partial\, z^g_B} = \frac{Y^{net}_{A,t}}{M_{pre}} \left[-\alpha_B \left(\frac{(1-n) f_3\, \gamma_A}{(\alpha_B\, z^g_B)^n} + f_2\gamma_A + d^g_{BA}\right)\right] \tilde{\phi}^A_{11} < 0$ for $\alpha_B\, z^g_B < \left(\frac{f_3(1-n)\gamma_A}{f_2\gamma_A + d^g_{AA}}\right)^{\frac{1}{n}}$ since $d^g_{BA} = d^g_{AA}$. Moreover, the unilateral action equilibrium with $S^A_t = 0$ requires $z^g_A \leq \alpha_B\, z^g_B$. Thus, SCC^A declines in the non-empty z^g_B-interval where $z^g_A \leq \alpha_B\, z^g_B < \left(\frac{f_3(1-n)\gamma_A}{f_2\gamma_A + d^g_{AA}}\right)^{\frac{1}{n}}$.

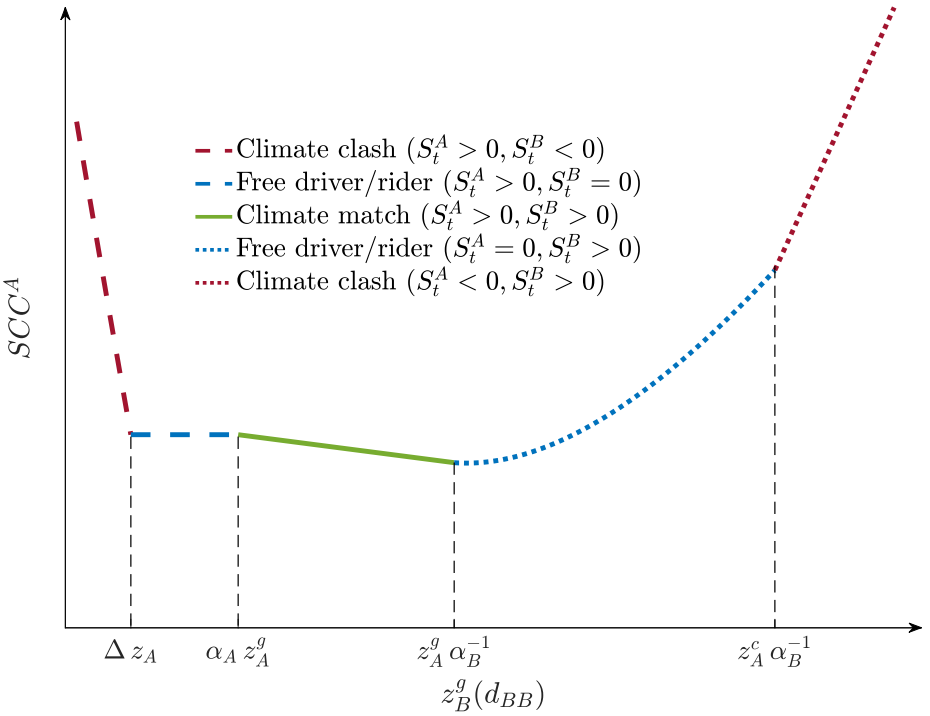

Figure 3.6: The graph shows region A's SCC as a function of region B's geoengineering propensity z_B^g (by varying d_{BB}^g) for the intuitive case where $d_{AA}^g = d_{BA}^g$. We defined $\Delta\, z_A = \alpha_A\, z_A^g - (z_B^c - z_B^g)$.

3.3.3 Rest of the world

We discussed the strategic geoengineering and mitigation decisions in technologically advanced and politically powerful regions that are willing and able to engage in climate engineering. The large part of the world will not be able or willing to perform such planetary alterations. Yet, even a single region's activity will change the mitigation incentives around the globe. This section discusses how the availability of geoengineering (and possibly counter-geoengineering) for some region changes the mitigation incentives around the world.

The "rest of the world" does not engage in climate engineering and affects regions A and B only indirectly through its emissions and the resulting changes of atmospheric carbon concentrations affecting geoengineering levels. The rest of the world has an aggregate economy similar to that of regions A and B with idiosyncratic

production function, damages, and resulting capital and emission dynamics. We denote the sulfur related geoengineering damages (or damage reductions) caused to the rest of the world by region $i \in \{A, B\}$ through its deployment of sulfur (or the counter-geoengineering agent) by $d_{iW} \in \{d^g_{iW}, d^c_{iW}\}$. If region i remains inactive, the respective damage parameter is zero. The resulting overall damages incurred by the rest of the world, as a fraction of its output, are

$$D^W_t(\tau^B_{1,t}, S_t, m_t) = 1 - \exp\left[\xi^W_0 \left(1 - \tau^B_{1,t}\right) - \left(d_{BW} S^B_t + d_{AW}\, \alpha_A\, S^A_t\right) - a^W\,(m_t - 1)\right].$$

As we assumed that the rest of the world is part of climate zone B, we also assume that sulfur levels are the same as in region B, including both local deployment S^B_t and spillovers $\alpha_A\, S^A_t$ from region A. We now characterize the impact of geoengineering on the mitigation incentives in the rest of world.

Proposition 3.6. *If region $i \in \{A, B\}$ acts unilaterally, the SCC in the rest of the world is given by*

$$SCC^W(z) = \frac{Y^{net}_{W,t}}{M_{pre}}\left[a^W + f_1\,\gamma_W - \left(\left(\frac{f_3}{z^n} - f_2\right)\gamma_W - d^g_{iW}\right) z\right]\tilde{\phi}^W_{11},$$

where $z = \alpha_A\, z^g_A$ for $i = A$ (other climate zone active) and $z = z^g_B$ for $i = B$ (same climate zone active). If both regions are active ($S^A_t \neq 0$ and $S^B_t \neq 0$), the SCC in the rest of the world is

$$SCC^W = \frac{Y^{net}_{W,t}}{M_{pre}}\left[a^W + f_1\,\gamma_W - \left(\left(\frac{f_3}{z^n_B} - f_2\right)\gamma_W - d_{BW}\right) z_B - \frac{\alpha_A(z_A - \alpha_B\, z_B)(d_{BW} - d_{AW})}{1 - \alpha_A\,\alpha_B}\right]\tilde{\phi}^W_{11},$$

with $z_A \in \{z^c_A, z^g_A\}$, $z_B \in \{z^c_B, z^g_B\}$, $d_{AW} \in \{d^c_{AW}, d^g_{AW}\}$, and $d_{BW} \in \{d^c_{BW}, d^g_{BW}\}$.

Proof. See Appendix B.1.3.9. □

The structure of the SCC in the rest of the world is similar to that of the active regions discussed in Proposition 3.5. As in the previous section, the first two contributions characterize the SCC in the absence of geoengineering, and a similar corollary follows.

Corollary 2. *The availability of geoengineering in regions A and B can increase, decrease or leave the SCC in the rest of the world unchanged relative to the regional world without geoengineering.*

One region acts unilaterally. If only one region is active, the SCC in the rest of the world is the exact analogue of that of the inactive region in Proposition 3.5. The one difference lies not in the structure of the SCC, but in the fact that the rest of the world will not match the cooling no matter how beneficial, and it will not be able to engage in counter-geoengineering no matter how damaging geoengineering is. As a result, if climate engineering is beneficial, the rest of the world can free-ride much more than a region with geoengineering potential (it has the credible "strategy" not to engage). But it can also be hit by geoengineering without any ability to respond. As we point out in section 3.3.2, avoiding the climate clash does not have to be a good thing; the clash allows both regions to get close to their desired climate targets. Without the ability to clash, the rest of the world is forced to inhabit an unfavorable environment.

Without loss of generality we assume that region A is the active region. If the rest of the world's benefit-cost ratio of (passively incurred) geoengineering exceeds that of the active region $\frac{f_3\,\gamma_W}{f_2\,\gamma_W+d^g_{AW}} > \alpha_A^n \frac{(1-n)\,f_3\,\gamma_A}{f_2\,\gamma_A+(d^g_{AA}+\epsilon^g_A)}$ $\left[=(\alpha_A z^g_A)^n\right]$ then the rest of world benefits from geoengineering; its marginal cost of emitting carbon is reduced by the cooling term $\big(\big(\frac{f_3}{(\alpha_A\,z^g_A)^n}-f_2\big)\gamma_W - d^g_{AW}\big)\alpha_A\,z^g_A > 0$. Unsurprisingly, this reduction grows with climate impact γ_W in the rest of the world and falls with the incurred geoengineering damages d^g_{AW}. We note that the benefit-cost ratio of geoengineering for the active region has to account for the deployment costs

ϵ_A^g and the decreasing effectiveness of geoengineering[21] both of which are absent for the passive region. Thus, *if the regions are somewhat symmetric, the rest of the world will tend to benefit (in the short run) from geoengineering and, as a consequence, increase its CO_2 emissions (worsening the long-run situation). We could consider this result a formal version of the slippery slope argument, which warns that going down the road of geoengineering makes us emit more and, thereby, makes us increasingly dependent on geoengineering.*[22] *We point out that even a rest of the world that has no interest in ever engaging in climate engineering can make this slope a lot more slippery if only some country is willing to engage in geoengineering. Slipping becomes worse the higher the assumed effectiveness of the cooling, the higher the climate damages, and the lower the geoengineering damages to other regions.*

However, it is reasonable to assume that the incentive to do climate engineering are largest for those countries that benefit the most. If the benefit-cost ratio and, thus, geoengineering propensity of the active region is much higher than in the rest of the world, the above inequalities flip and the SCC in the rest of the world increases, making CO_2 emissions even more pricey than in the absence of geoengineering. Then the slippery slope argument turns around, at least on the planetary scale. While the active region slides it down a little bit, overall emissions decrease (also) to avoid more geoengineering damages from the unilaterally cooling region, and as global emissions decrease, also the cooling region will voluntarily reduce its geoengineering level.

Climate match. With both regions active, the SCC in the rest of the world resembles that of the active regions (see Proposition 3.5). However, the rest of the world does not participate in the costly cooling. If its marginal damages from

[21] Sulfur deployment scales with the CO_2 concentrations. The active decision maker takes the decreasing effectiveness of sulfur into account and, therefore, has a higher opportunity cost of sulfur deployment.

[22] And maybe even worse, merely expecting that some region will engage in geoengineering in the future, already reduces the expected damages from today's emissions, thereby increasing the pressure to truly engaging in geoengineering in the future.

geoengineering are independent of the origin, $d^g_{BW} = d^g_{AW}$, then the spillover term is zero. In this case, the situation is identical to the scenario of unilateral action. The rest of the world does not care whether one or both regions engage in geoengineering. Moreover, its SCC reduction (or increase) is proportional to the geoengineering propensity of region B, which shares the climate zone and sets the "local" climate target. If both active regions have similar geoengineering propensities and the marginal damages from geoengineering in region A are lower than those from geoengineering in region B, then the spillover term will reduce the rest of the world's SCC relative to that based on region B's unilateral action.

Climate clash. If region A cools and region B engages in counter-geoengineering, then $\alpha_A(z^g_A - \alpha_B\, z^c_B) > 0$. Provided that counter-geoengineering cannot perfectly offset the damages from geoengineering, $d^c_{BW} < d^g_{AW}$, the spillover term increases the SCC. In the opposite scenario where region B cools and region A engages in counter-geoengineering, $\alpha_A(z^c_A - \alpha_B\, z^g_B) < 0$ and $d^g_{BW} > d^c_{AW}$. Again, the spillover term is positive and increases the SCC. *Thus, in a climate clash the spill-over term is always positive, and increases the SCC in the rest of the world.* If region B cools and region A is countering it, then the positive spill-over term increases the SCC compared to case where region B acts unilaterally. If geoengineering damages from region B's cooling are high and the effectiveness from counter-geoengineering is low, the SCC in the rest of the world can be higher than in a world without geoengineering, *helping to turn the slippery of geoengineering uphill increasing global mitigation.*[23]

3.3.4 Heat transfer

This section lifts Assumption 3.3 of the baseline model, introducing direct heat exchange between the regions; temperature change in one region now directly af-

[23]The climate clash equilibrium where $S^A_t < 0$ and $S^B_t > 0$ requires $z^c_A - \alpha_B\, z^g_B < 0$ and $z^g_B - \alpha_A\, z^c_A > 0$. Thus, the availability of geoengineering increases the SCC if $d^g_{BW} > \frac{(1-\alpha_A\,\alpha_B)\left(\frac{f_3}{(z^g_B)^n} - f_2\right)\gamma_W\, z^g_B - \alpha_A(z^c_A - \alpha_B\, z^g_B)d^c_{AW}}{z^g_B - \alpha_A\, z^c_A} > 0$.

fects the temperature in the other region. As a result, we find adjustments to the geoengineering targets as well as the strategic interactions. Propositions 3.7 states that under according modifications of the geoengineering propensity (and counter-geoengineering reluctance), Propositions 3.3 and 3.4 characterizing the subgame-perfect Nash equilibrium remain valid. Proposition 3.8 shows how heat transfers alter the SCC.

In our baseline model, interaction between regions was reduced to the spill-over of the cooling sulfur deployment (or counter-geoengineering agent). Now, cooling one region also directly alters the temperature in the other region, even without sulfur spill-over, a natural consequence of heat exchange across the globe.

Proposition 3.7. *With heat transfers, Propositions 3.3 and 3.4 remain valid under the following modifications of the definitions of the geoengineering propensities z_A^g and counter-geoengineering reluctance z_A^c:*

If (i) $S_t^B = 0$, then

$$z_A^g = \left(\frac{a_A + b_A \, \alpha_A^{-n}}{\delta_A^g} \right)^{\frac{1}{n}},$$

and if (ii) $S_t^B \neq 0$, then

$$z_A^g = \left(\frac{a_A - b_A \frac{b_B}{a_B}}{\delta_A^g - b_A \frac{\delta_B}{a_B}} \right)^{\frac{1}{n}}, \quad z_A^c = \left(\frac{a_A - b_A \frac{b_B}{a_B}}{\delta_A^c - b_A \frac{\delta_B}{a_B}} \right)^{\frac{1}{n}},$$

where

$$a_A = \quad (1-n) f_3 \, \gamma_A \quad , \qquad \delta_A^g = (d_{AA}^g + \epsilon_A^g) + \gamma_A \, f_2 + \alpha_A \, \gamma_A^{heat} \, f_2,$$

$$b_A = \alpha_A (1-n) f_3 \, \gamma_A^{heat}, \qquad \delta_A^c = (d_{AA}^c - \epsilon_A^c) + \gamma_A \, f_2 + \alpha_A \, \gamma_A^{heat} \, f_2,$$

and direct climate change impact $\gamma_A \equiv \beta^A \, \xi_0^A \, \tilde{\sigma}_{11}^A \, \sigma_{forc}$, as well as heat transfer driven climate change impact $\gamma_A^{heat} \equiv \beta^A \, \xi_0^A \, \tilde{\sigma}_{12}^A \, \sigma_{forc}$. Swapping region indices characterizes region B's strategies.

Proof. See Appendix B.1.3.6. □

The new term γ_A^{heat} reflects the climate change impact from heat exchange with the other region. Its component $\tilde{\sigma}_{12}^A = \left[(\mathbf{1} - \beta^A \boldsymbol{\sigma}^{\boldsymbol{A}})^{-1}\right]_{12}$ characterizes the discounted long-term heat flux from region B to region A resulting from a present heating (forcing) change in region B. This heat exchange affects both, heat increase as a result of greenhouse gas emissions and heat reduction as a result of geoengineering.

We note that the geoengineering propensities and counter-geoengineering reluctance without heat transfers ($\gamma_A^{heat} = 0$) in Propositions 3.3 and 3.4 are simply

$$\text{w/o heat transfer:} \quad z_A^g = \left(\frac{a_A}{\delta_A^g}\right)^{\frac{1}{n}}, \quad z_A^c = \left(\frac{a_A}{\delta_A^c}\right)^{\frac{1}{n}},$$

where a_A captures sulfur's direct cooling efficiency f_3 in region A, and the damage impact of global warming γ_A, both of which increase the region's geoengineering propensity. The component δ_A^g captures the damages from and costs of geoengineering, which reduces the geoengineering propensity (as well as a negligible efficiency loss in forcing captured by the close-to-zero f_2, which here gains an additional term capturing heat spill-over from region B caused by region A's cooling). Comparing a model with heat transfer to one without heat transfers implies a re-calibration of the heat transfer matrix $\boldsymbol{\sigma}$. In order to describe the same equilibrium temperature response, the matrix elements $\left[(\mathbf{1} - \boldsymbol{\sigma})^{-1}\right]_{ij}$ for $i, j \in \{1, 2\}$ have to coincide. For perfectly patient decision makers, this condition implies that the parameters $\tilde{\sigma}_{ij}$ and, thus, γ_A, γ_B are directly comparable across models. For the impatient decision maker, the climate impact parameters γ_A, γ_B would slightly differ, a minor difference that we ignore in the subsequent discussion.

One active region. If only region A is active, the geoengineering propensity increases by the term $b_A \alpha_A^{-n}$, where b_A *is the heat transfer benefit.* It characterizes the climate–impact–γ_A–weighted cooling benefits in region A that result from its cooling of region B through sulfur spill-overs α_A at efficiency f_3. These spill-overs then feed back into region A through the heat transfer captured in $\tilde{\sigma}_{12}^A$. The

geoengineering propensity adjustment weighs this heat transfer benefit by α_A^{-n}. A higher sulfur spill-over has two implications. First, it increases the marginal benefit from deploying sulfur through the heat transfer benefit (part of b_A). But second, it also implies that the level of sulfur in region B is already high, and the effectiveness of deploying yet another ton of sulfur to cool region B is lower. The net effect of the spill-over α_A on the additional geoengineering propensity as a result of heat transfer is positive ($\alpha_A^{-n}\alpha_A = \alpha_A^{1-n}$, which increases in α_A), but it is lower than one might expect when merely considering the heat transfer benefits. It might be more intuitive to call this (negative) heat transfer a "cooling transfer". *In summary, under unilateral action, the cooling transfer gives the region an additional incentive for cooling as it benefits from the global forcing impact of its sulfur deployment.*

Both regions cooling. With two active regions, strategic effects set in. We focus on the additional strategic effects implied by heat transfer. Other strategic implications discussed in Proposition 3.3 apply alongside. We explain the adjustment of the geoengineering propensity in case (ii) in two steps. First, we assume that region B does not experience any heat transfer benefit ($b_B = 0$). Then, region B's geoengineering propensity remains $z_B^g = \left(\frac{a_B}{\delta_B}\right)^{\frac{1}{n}}$ as in the case without heat transfer. Region A increases its geoengineering propensity according to the contribution $b_A\frac{\delta_B}{a_B}$; its geoengineering increases in response to a higher heat transfer benefit b_A, but less so if region B already has a high geoengineering propensity, which reduces the efficiency of additional cooling.[24] Second, let us turn back on region B's heat transfer benefit b_B. Now also region B benefits from the cooling in region A. Similarly to region A in the first case, it will ramp up sulfur deployment. Anticipating this response of region B, region A lowers its own target as captured by the term $b_A\frac{b_B}{a_B}$. This strategic "free-riding" response increases in the heat transfer benefit (connectedness) of both regions.[25] A short calculation shows that region

[24]The term $\frac{\delta_B}{a_B} = \left(\frac{1}{z_B^g}\right)^n$ falls in region B's geoengineering propensity z_B^g.

[25]This additional "free-riding" as a result of heat transfer falls in a_B, which characterizes the climate impact weighted cooling efficiency within region B. A higher a_B relative to b_B makes region

A's cooling increases as a result of heat (or cooling) transfer, if its own benefit-cost ratio before heat transfer $\frac{a_A}{\delta_A^g}$ is larger than the spill-over benefit over region B's geoengineering damages $\frac{b_A}{\delta_B^g}$.[26] *In summary, the heat transfer increases the incentive for cooling. With both regions cooling, efficiency loss and free-riding incentive counter this additional cooling incentive. A region's cooling incentive increases overall under heat transfer whenever its benefit-cost ratio without heat transfers exceeds the spill-over benefits to other region's costs ratio (otherwise the free-riding incentive will dominate).*

Counter-geoengineering. Now, let us consider the climate clash scenario where region B cools, and region A deploys the counter-geoengineering agent. Following the two step interpretation, first, suppose region B does not experience any heat transfer ($\Rightarrow b_B = 0$). Then region B's geoengineering propensity remains $z_B^g = \left(\frac{a_B}{\delta_B}\right)^{\frac{1}{n}}$. Region A's reluctance to engage in counter-geoengineering increases (=less counter-geoengineering) with heat transfer as the denominator in z_A^c decreases by the term $b_A \frac{\delta_B}{a_B}$; region A gets more cooling for less damages and results less contrarian to region B's cooling. We note that the reluctance to counter-geoengineering is more sensitive to heat transfers than the geoengineering propensity in the case above where both regions are cooling since $\delta_A^c > \delta_A^g$. Second, turning on heat transfer for region B ($b_B > 0$) leads to an increase in region B's geoengineering propensity (as in the case above). Region A anticipates the reaction of region B and decreases its target sulfur level (as the numerator in z_A^c decreases by the term $b_A \frac{b_B}{a_B}$). The strategic component renders the regions more contrarian under heat transfers. Overall, the effect of heat transfer on deployment levels depends on the cost-benefit

B relatively less responsive to switching on heat transfer. Then, region A lowers its geoengineering propensity less than in the case where it anticipates a stronger response of region B.

[26] Region A's propensity increases if $\frac{a_A - b_A \frac{b_B}{a_B}}{\delta_A^g - b_A \frac{\delta_B}{a_B}} > \frac{a_A}{\delta_A^g} \Leftrightarrow \frac{a_A}{\delta_A^g} > \frac{b_B}{\delta_B^g}$, and decreases if $\frac{a_A}{\delta_A^g} \leq \frac{b_B}{\delta_B^g}$. Analogously, region B's propensity increases if $\frac{a_B}{\delta_B^g} > \frac{b_A}{\delta_A^g}$, and decreases if $\frac{a_B}{\delta_B^g} \leq \frac{b_A}{\delta_A^g}$. Assuming $\tilde{\sigma}_{11}^i > \tilde{\sigma}_{12}^i$ implies $\gamma_i > \gamma_i^{heat}$, and thus $a_i > b_i$ for region $i \in \{A, B\}$. As a result $\frac{a_A}{\delta_A^g} \leq \frac{b_B}{\delta_B^g}$ can never occur in combination with $\frac{a_B}{\delta_B^g} \leq \frac{b_A}{\delta_A^g}$ since the former implies $\frac{a_A}{\delta_A^g} < \frac{a_B}{\delta_B^g}$ and the latter implies $\frac{a_B}{\delta_B^g} < \frac{a_A}{\delta_A^g}$. Thus, a decrease in both regions' propensities cannot occur.

ratios of geoengineering and counter-geoengineering, and the heat benefits in both regions. If heat transfer increases region B's propensity to cool and decreases region A's counter-geoengineering reluctance, both regions increase their deployment levels (the climate clash gets worse). The opposite occurs if heat transfer decreases region B's propensity and increase A's reluctance. In this case, both regions lower their deployment levels, and the climate clash turns less extreme. If heat transfer increases region B's cooling propensity, and increases region A's reluctance the net effect on deployment levels is ambiguous. We note that a decrease in the geoengineering propensity of region B and a decrease in the counter-geoengineering reluctance of region A cannot occur in combination.[27] *In summary, in the clash scenario, the non-strategic implications of heat transfer reduce counter-geoengineering. However, they also increase the cooling region's geoengineering propensity. Thus, the strategic incentives boost the clash. Overall, heat transfer can both turn the climate clash worse or improve the situation, depending on the regions' cost-benefit ratios for geoengineering and counter-geoengineering.*

We now derive the regional SCC characterizing the mitigation implications of heat exchange. We focus on the new contributions resulting from heat exchange and abbreviate the original SCC formula of our baseline model as $SCC^A_{w/o}(\cdot)$, indicating as arguments the engineering propensities on which it depends. These propensities change as discussed above, whereas the structure of $SCC^A_{w/o}(\cdot)$ remains identical to that observed in Proposition 3.5.

---

[27] Region B's geoengineering propensity increases if $\frac{a_B - b_B \frac{b_A}{a_A}}{\delta^g_B - b_B \frac{\delta^c_A}{a_A}} > \frac{a_B}{\delta^g_B} \Leftrightarrow \frac{a_B}{\delta^g_B} > \frac{b_A}{\delta^c_A}$, and decreases if $\frac{a_B}{\delta^g_B} \leq \frac{b_A}{\delta^c_A}$. Analogously, region A's counter-geoengineering reluctance increases if $\frac{a_A - b_A \frac{b_B}{a_B}}{\delta^c_A - b_A \frac{\delta^g_B}{a_B}} > \frac{a_A}{\delta^c_A}$ $\Leftrightarrow \frac{a_A}{\delta^c_A} > \frac{b_B}{\delta^g_B}$, and decreases if $\frac{a_A}{\delta^c_A} \leq \frac{b_B}{\delta^g_B}$. Assuming $\tilde{\sigma}^i_{11} > \tilde{\sigma}^i_{12}$ implies $\gamma_i > \gamma^{heat}_i$, and thus $a_i > b_i$ for region $i \in \{A, B\}$. As a result $\frac{a_B}{\delta^g_B} \leq \frac{b_A}{\delta^c_A}$ and $\frac{a_A}{\delta^c_A} \leq \frac{b_B}{\delta^g_B}$ cannot occur in combination since the former implies $\frac{a_B}{\delta^g_B} < \frac{a_A}{\delta^c_A}$ and the latter implies $\frac{a_B}{\delta^g_B} > \frac{a_A}{\delta^c_A}$. Thus, a decrease in region B's geoengineering propensity cannot occur in combination with a decrease in region A's counter-geoengineering reluctance.

Proposition 3.8. *If $S_t^B = 0$, with heat transfers the SCC in region A is*

$$SCC^A = SCC^A_{w/o}(z_A^g) + \frac{Y^{net}_{A,t}}{M_{pre}}\left[f_1\,\gamma_A^{heat} - \left(\frac{f_3}{(\alpha_A z_A^g)^n} - f_2\right)\gamma_A^{heat}\,\alpha_A z_A^g\right]\tilde{\phi}^A_{11}$$

If region A itself is inactive ($S_t^A = 0$) its SCC is

$$SCC^A = SCC^A_{w/o}(z_B^g) + \frac{Y^{net}_{A,t}}{M_{pre}}\left[f_1\,\gamma_A^{heat} - \left(\frac{f_3}{(z_B^g)^n} - f_2\right)\gamma_A^{heat}\,z_B^g\right]\tilde{\phi}^A_{11}$$

If both regions are active ($S_t^B \neq 0$ and $S_t^A \neq 0$) the SCC is

$$SCC^A = SCC^A_{w/o}(z_A, z_B) + \frac{Y^{net}_{A,t}}{M_{pre}}\left[f_1\,\gamma_A^{heat} - \left(\frac{f_3}{z_B^n} - f_2\right)\gamma_A^{heat}\,z_B\right]\tilde{\phi}^A_{11},$$

with $z_A \in \{z_A^c, z_A^g\}$, and $z_B \in \{z_B^c, z_B^g\}$, depending on whether the corresponding region engages in geoengineering (g) or counter-geoengineering (c). Swapping region indices characterizes region B's SCC. We note that Corollary 1 from the base model still holds.

Proof. See Appendix B.1.3.8. □

All cases still contain the original terms abbreviated by $SCC^A_{w/o}(\cdot)$. Their contributions depends on the geoengineering propensities and counter-geoengineering reluctancies, which now change because of the strategical responses to heat transfer that we discussed above in Proposition 3.7. In addition, heat transfer introduces two new terms. The first in the square brackets reflects the heat flow across regions related to greenhouse gas emissions. It is positive and increases in γ_A^{heat} characterizing the heat transfer based climate change impact. This contribution also arises in a regional model with heat flows but without geoengineering. Other damage terms are absent because heat transfer does not affect ocean acidification or sulfur-based damages. The second term in the square brackets is negative and reflects the cooling from geoengineering under heat transfer.

3.4 Summary and conclusions

We have introduced solar geoengineering into a state of the art integrated assessment model of climate change. For this purpose, we found a new solution class to closed-form dynamic models which allows us to incorporate the current scientific knowledge about the radiative forcing response to stratospheric sulfur injections.

The global model shows that the social planner's optimal sulfur deployment is very sensitive to potential damages from geoengineering, which are mostly unknown. The size of the SCC reduction increases in the sulfur-based cooling efficiency and falls with geoengineering damages. We find that the SCC reduction is less responsive to geoengineering damages than is the geoengineering activity. Solar geoengineering could cut the SCC into half if damages turn out negligible. However, current damage guesstimates would reduce the globally optimal SCC by only 12-22%.

Our strategic model assumes that regional deployment follows a linear Markov strategy, which coincides with the unique optimal course of action of a global planner. We analytically characterize the regional strategies and identify three qualitatively distinct types of equilibria; unilateral action where only one region is active, a climate match where both regions cool the world, and a climate clash where one region cools and the other region engages in a counter-measure. We show that these equilibria are mutually exclusive and depend on the regions' damage characteristics. Further, we analyze how changes in geoengineering damages and the effectiveness of counter-measures affect the equilibria domains.

If the active regions are somewhat symmetric, they will both contribute to the global cooling. As they become heterogeneous, the region with lower climate damages or higher geoengineering damages will increasingly free-ride on the other region's actions. Given its ability to free-ride, the region's SCC will fall and it will emit more greenhouse gases as compared to a situation where the regions are symmetric. Eventually, the free-riding region will stop to contribute entirely. In this

equilibria, there exists a small domain where its SCC decreases further until the inactive region reaches its free-riding "bliss point". As the regions' benefit-cost ratios of geoengineering become more asymmetric, the free-riding region will eventually start to dislike the active region's cooling. Initially, it is still better off than in a world without cooling, but it will already make use of the option to neutralize some of the active region's cooling, if costs for the counter-measure are sufficiently low. In this case, we call the active region a free-driver. It imposes a special type of externality on the other region. The overall externality on the inactive region is positive, however, the marginal externality is negative. Eventually, as the climate clash gets more extreme, also the overall externality turns negative.

For simplicity of presentation, we first limited the interactions to cooling spill-overs. We then show that introducing heat transfer leaves the qualitative results unchanged, but adds two additional layers. First, heat exchange gives rise to two additional terms in the SCC. Second, heat exchange increases a region's incentive to cool the planet because it benefits not only from its own direct cooling, but also from the spill-over cooling in the neighboring regions. Under unilateral action, heat transfer therefore increases a region's cooling incentive. However, if both regions are active, they anticipate the other's response, resulting in a free-riding incentive that reduces or even inverts the original incentive to do more.

The rest of the world, which is either not willing or has not the ability to engineer the planet's climate, will always be able to free-ride, but never be able to counter. Thus, the rest of the world has more to gain from having similar geoengineering preferences as an active region and more to lose from being dissimilar. This powerfully represents a version of the slippery slope argument. A single region's cooling action can reduce the mitigation incentives, represented by the SCC, in all or large parts of the rest of the world, which will further increase the geoengineering level of the active region. However, it is reasonable to assume that the incentives to engage in climate engineering are largest for regions that benefit the most. If

these regions have a sufficiently higher benefit-cost ratio of geoengineering than the rest of the world, the SCC in large parts of the world increases. As a result CO_2 emissions become even more pricey than in the absence of geoengineering, turning the slippery slope of geoengineering uphill. Such a mitigation enhancing implication of geoengineering can arise in all three types of equilibria. As overall emissions decrease (also) to avoid more geoengineering, the active regions voluntarily scale down their efforts.

Last, we note that our framework permits the introduction of uncertainty, which is ubiquitous in the economics of geoengineering (Heutel et al., 2018; Emmerling and Tavoni, 2018a). Numeric assessment of the uncertainty in regional (including strategic) models is severely challenged by the curse of dimensionality in dynamic programming. Our analytic model will be able to integrate uncertainty into the high-dimensions space of a regional integrated assessment model, even though we currently have to leave it for future work.

References

Boucher, Olivier; Kleinschmitt, Christoph, and Myhre, Gunnar. Quasi-Additivity of the Radiative Effects of Marine Cloud Brightening and Stratospheric Sulfate Aerosol Injection. Geophysical Research Letters, 44(21):158–11, 2017.

Carleton, Tamma A and Hsiang, Solomon M. Social and economic impacts of climate. Science, 353(6304), 2016.

Crutzen, Paul J. Albedo enhancement by stratospheric sulfur injections: A contribution to resolve a policy dilemma? Climatic Change, 77(3-4):211–219, 2006.

Dai, Zhen; Weisenstein, Debra K, and Keith, David W. Tailoring Meridional and Seasonal Radiative Forcing by Sulfate Aerosol Solar Geoengineering. Geophysical Research Letters, 45(2):1030–1039, 2018.

Dykema, John A; Keith, David W, and Keutsch, Frank N. Improved aerosol radiative properties as a foundation for solar geoengineering risk assessment. Geophysical Research Letters, 43(14):7758–7766, 2016.

Emmerling, Johannes and Tavoni, Massimo. Climate Engineering and Abatement: A 'flat' Relationship Under Uncertainty. Environmental and Resource Economics, 69(2):395–415, 2018a.

Emmerling, Johannes and Tavoni, Massimo. Exploration of the interactions between mitigation and solar radiation management in cooperative and non-cooperative international governance settings. Global Environmental Change, 53:244–251, 2018b.

Gerlagh, Reyer and Liski, Matti. Consistent climate policies. Journal of the European Economic Association, 16(1):1–44, 2018.

Goes, Marlos; Tuana, Nancy, and Keller, Klaus. The economics (or lack thereof) of aerosol geoengineering. Climatic Change, 109(3-4):719–744, 2011.

Golosov, Mikhail; Hassler, John; Krusell, Per, and Tsyvinski, Aleh. Optimal Taxes on Fossil Fuel in General Equilibrium. Econometrica, 82(1):41–88, 2014.

Hambel, Christoph; Kraft, Holger, and Schwartz, Eduardo S. The carbon abatement game. Technical report, National Bureau of Economic Research, 2018.

Hassler, John and Krusell, Per. Economics and climate change: integrated assessment in a multi-region world. Journal of the European Economic Association, 10(5):974–1000, 2012.

Heckendorn, Patricia; Weisenstein, Debra; Fueglistaler, Stephan; Luo, Beiping P; Rozanov, Eugene; Schraner, Martin; Thomason, Larry W, and Peter, Thomas. The impact of geoengineering aerosols on stratospheric temperature and ozone. Environmental Research Letters, 4(4):045108, 2009.

Heutel, Garth; Moreno-Cruz, Juan, and Shayegh, Soheil. Solar geoengineering, uncertainty, and the price of carbon. Journal of Environmental Economics and Management, 87:24–41, 2018.

Heyen, Daniel; Horton, Joshua, and Moreno-Cruz, Juan. Strategic Implications of Counter-Geoengineering: Clash or Cooperation? Journal of Environmental Economics and Management, 95:153–177, 2019.

Hoegh-Guldberg, Ove and Bruno, John F. The impact of climate change on the world's marine ecosystems. Science, 328(5985):1523–1528, 2010.

IPCC, The Intergovernmental Panel on Climate Change. Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2013.

Jones, Anthony C; Hawcroft, Matthew K; Haywood, James M; Jones, Andy; Guo, Xiaoran, and Moore, John C. Regional Climate Impacts of Stabilizing Global Warming at 1.5 K Using Solar Geoengineering. Earth's Future, 6(2):230–251, 2018.

Keith, David W and MacMartin, Douglas G. A temporary, moderate and responsive scenario for solar geoengineering. Nature Climate Change, 5(3):201–206, 2015.

Keith, David W; Weisenstein, Debra K; Dykema, John A, and Keutsch, Frank N. Stratospheric solar geoengineering without ozone loss. Proceedings of the National Academy of Sciences, 113(52):14910–14914, 2016.

Kleinschmitt, Christoph; Boucher, Olivier, and Platt, Ulrich. Sensitivity of the radiative forcing by stratospheric sulfur geoengineering to the amount and strategy of the SO2injection studied with the LMDZ-S3A model. Atmospheric Chemistry and Physics, 18(4):2769–2786, 2018.

Klepper, Gernot and Rickels, Wilfried. The Real Economics of Climate Engineering. Economics Research International, 2012:1–20, 2012.

Klepper, Gernot and Rickels, Wilfried. Climate engineering: Economic considerations and research challenges. Review of Environmental Economics and Policy, 8(2):270–289, 2014.

Kravitz, Ben; Robock, Alan; Oman, Luke; Stenchikov, Georgiy, and Marquardt, Allison B. Sulfuric acid deposition from stratospheric geoengineering with sulfate aerosols. Journal of Geophysical Research Atmospheres, 114(14):1–7, 2009.

Kravitz, Ben; MacMartin, Douglas G; Mills, Michael J; Richter, Jadwiga H; Tilmes, Simone; Lamarque, Jean-Francois; Tribbia, Joseph J, and Vitt, Francis. First simulations of designing stratospheric sulfate aerosol geoengineering to meet multiple simultaneous climate objectives. Journal of Geophysical Research: Atmospheres, 122(23):12,616-12,634, 2017.

Lawrence, Mark G and Crutzen, Paul J. Was breaking the taboo on research on climate engineering via albedo modification a moral hazard, or a moral imperative? Earth's Future, 5(2):136–143, 2017.

Lawrence, Mark G; Schäfer, Stefan; Muri, Helene; Scott, Vivian; Oschlies, Andreas; Vaughan, Naomi E; Boucher, Olivier; Schmidt, Hauke; Haywood, Jim, and Scheffran, Jürgen. Evaluating climate geoengineering proposals in the context of the Paris Agreement temperature goals. Nature Communications, 9(1):3734, 2018.

MacMartin, Douglas G; Kravitz, Ben; Tilmes, Simone; Richter, Jadwiga H; Mills, Michael J; Lamarque, Jean-Francois; Tribbia, Joseph J, and Vitt, Francis. The climate response to stratospheric aerosol geoengineering can be tailored using multiple injection locations. Journal of Geophysical Research: Atmospheres, 122(23):12,574–12,590, 2017.

McClellan, Justin; Keith, David W, and Apt, Jay. Cost analysis of stratospheric albedo modification delivery systems. Environmental Research Letters, 7(3):034019, 2012.

Moreno-Cruz, Juan B. Mitigation and the geoengineering threat. Resource and Energy Economics, 41:248–263, 2015.

Moreno-Cruz, Juan B and Keith, David W. Climate policy under uncertainty: A case for solar geoengineering. Climatic Change, 121(3):431–444, 2013.

Moreno-Cruz, Juan B and Smulders, Sjak. Revisiting the economics of climate change: the role of geoengineering. Research in Economics, 71(2):212–224, 2017.

Moriyama, Ryo; Sugiyama, Masahiro; Kurosawa, Atsushi; Masuda, Kooiti; Tsuzuki, Kazuhiro, and Ishimoto, Yuki. The cost of stratospheric climate engineering revisited. Mitigation and Adaptation Strategies for Global Change, 22(8):1207–1228, 2017.

Morrow, David R. Ethical aspects of the mitigation obstruction argument against climate engineering research. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 372(2031):20140062, 2014.

Niemeier, Ulrike and Timmreck, Claudia. What is the limit of climate engineering by stratospheric injection of SO2? Atmospheric Chemistry and Physics, 15(16):9129–9141, 2015.

Niemeier, Ulrike and Schmidt, Hauke. Changing transport processes in the stratosphere by radiative heating of sulfate aerosols. Atmospheric Chemistry and Physics, 17(24):14871–14886, 2017.

Nordhaus, William. Climate clubs: Overcoming free-riding in international climate policy. American Economic Review, 105(4):1339–70, 2015.

Nordhaus, William and Sztorc, Paul. DICE2013R: Introduction and User's Manual. 2013. URL `http://www.econ.yale.edu/~nordhaus/homepage/homepage/documents/DICE_Manual_100413r1.pdf`.

Nordhaus, William D and Yang, Zili. A regional dynamic general-equilibrium model of alternative climate-change strategies. The American Economic Review, 86(4):741-765, 1996.

Parker, Andy; Horton, Joshua B, and Keith, David W. Stopping Solar Geoengineering Through Technical Means: A Preliminary Assessment of Counter-Geoengineering. Earth's Future, 6:1058–1065, 2018.

Pasztor, Janos; Scharf, Cynthia, and Schmidt, Kai-Uwe. How to govern geoengineering? Science, 357(6348):231–231, 2017.

Proctor, Jonathan; Hsiang, Solomon; Burney, Jennifer; Burke, Marshall, and Schlenker, Wolfram. Estimating global agricultural effects of geoengineering using volcanic eruptions. Nature, 560(7719):480–483, 2018.

Quaas, Martin F; Quaas, Johannes; Rickels, Wilfried, and Boucher, Olivier. Are there reasons against open-ended research into solar radiation management? a model of intergenerational decision-making under uncertainty. Journal of Environmental Economics and Management, 84:1–17, 2017.

Ricke, Katharine L; Morgan, M. Granger, and Allen, Myles R. Regional climate response to solar-radiation management. Nature Geoscience, 3(8):537–541, 2010.

Robock, Alan. 20 Reasons Why Geoengineering May Be a Bad Idea. Bulletin of the Atomic Scientists, 64(2):14–18, 2008.

Robock, Alan; Marquardt, Allison; Kravitz, Ben, and Stenchikov, Georgiy. Benefits, risks, and costs of stratospheric geoengineering. Geophysical Research Letters, 36(19), 2009.

Smith, Wake and Wagner, Gernot. Stratospheric aerosol injection tactics and costs in the first 15 years of deployment. Environmental Research Letters, 13(12):124001, 2018.

Tol, Richard S. Estimates of the damage costs of climate change. Part 1: Benchmark estimates. Environmental and resource Economics, 21(1):47-73, 2002.

Tol, Richard S. Estimates of the damage costs of climate change, Part II. Dynamic estimates. Environmental and Resource Economics, 21(2):135-160, 2002.

Tollefson, Jeff. Global industrial carbon emissions to reach all-time high in 2018. Nature News, 2018. URL `https://www.nature.com/articles/d41586-018-07666-6#references`.

Traeger, Christian P. ACE - Analytic Climate Economy (with Temperature and Uncertainty). 2018. URL `https://ssrn.com/abstract=3307622`.

Urban, Mark C. Accelerating extinction risk from climate change. Science, 348(6234):571–573, 2015.

Visioni, Daniele; MacMartin, Douglas G; Kravitz, Ben; Tilmes, Simone; Mills, Michael J; Richter, Jadwiga H, and Boudreau, Matthew P. Seasonal Injection Strategies for Stratospheric Aerosol Geoengineering. Geophysical Research Letters, 46(13):7790–7799, 2019.

Weisenstein, Debra K; Keith, David W, and Dykema, John A. Solar geoengineering using solid aerosol in the stratosphere. Atmospheric Chemistry and Physics, 15(20):11835–11859, 2015.

Weitzman, Martin L. A Voting Architecture for the Governance of Free-Driver Externalities, with Application to Geoengineering. Scandinavian Journal of Economics, 117(4):1049–1068, 2015.

B.1 Appendix to Chapter 3

B.1.1 Further results on radiative forcing

Model data from Kleinschmitt et al. (2018). Figure 3.7 shows approximated radiative forcing F_t as a 3D function of the relative atmospheric carbon concentration m_t and sulfur injections S_t.

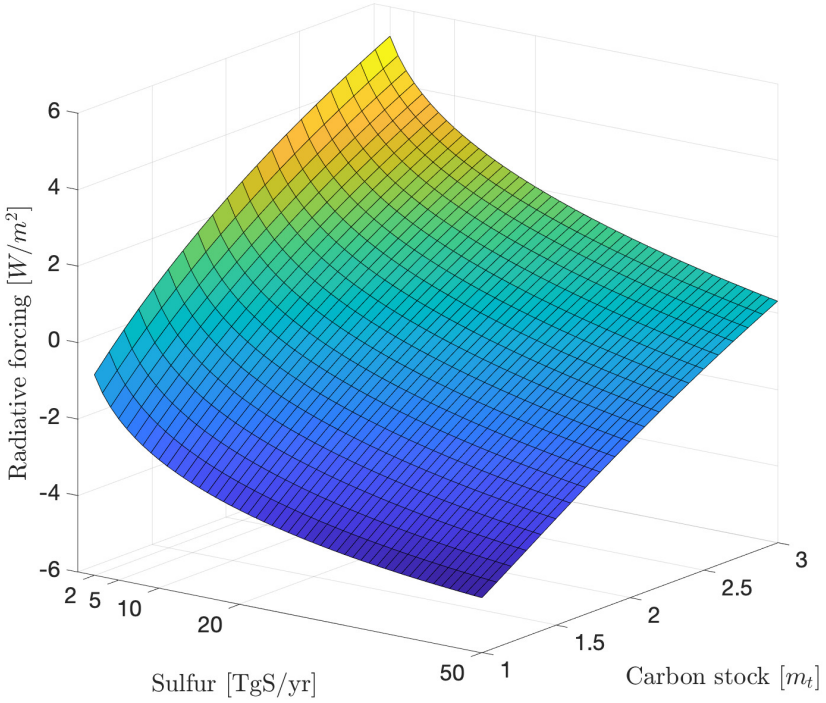

Figure 3.7: Radiative forcing F_t as a function of the relative atmospheric carbon concentration m_t and sulfur injections S_t.

Model data from Niemeier and Schmidt (2017). We also calibrate radiative forcing equation F_t to data from Niemeier and Schmidt (2017) over the relative atmospheric carbon interval $m_t \in [1.5, 3]$. Table 3.6 shows that the estimated forcing parameters also fulfill Assumption 3.1.

Table 3.6: Estimated forcing parameters

f_0	f_1	f_2	f_3	n
0.4	2.9	0.004	2.08	0.9

Figure 3.8 shows approximated radiative forcing based on the data from Niemeier and Schmidt (2017) as a function of sulfur injections (left graph), and as a function of the relative atmospheric carbon concentration (right graph).

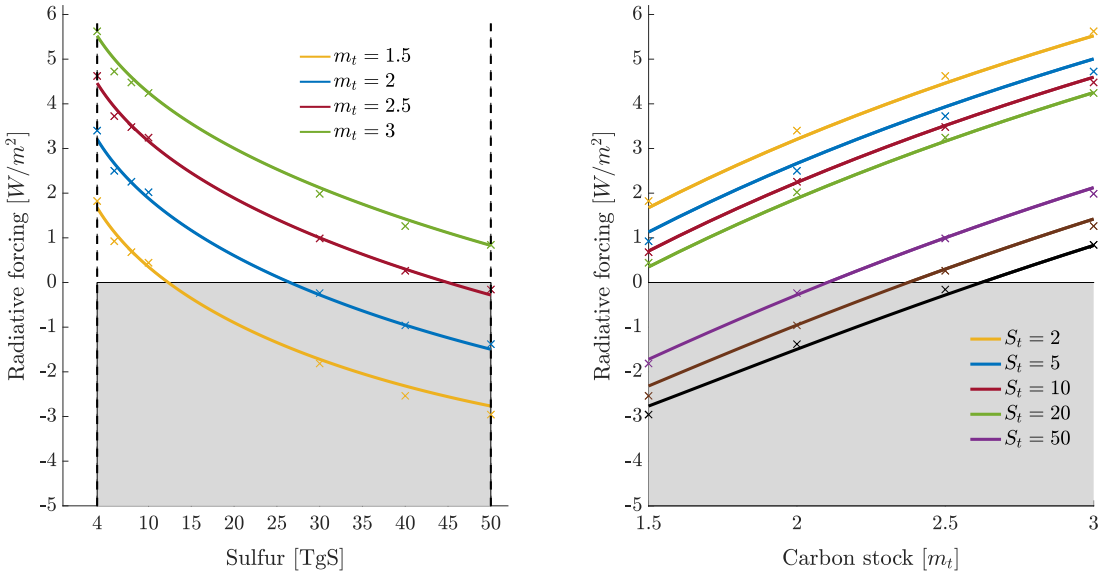

Figure 3.8: Approximation of radiative forcing F_t to model data (shown by cross markers) from Niemeier and Schmidt (2017) for sulfur injections between 4 and 50 TgS, and positive radiative forcing levels. The left graph shows radiative forcing as a function of sulfur for different atmospheric carbon concentrations. The right graph shows radiative forcing as a function of the atmospheric carbon concentration for different sulfur injection rates.

B.1.2 Global model

B.1.2.1 Solving the Bellman equation

Definitions. We note that aggregate capital $K_t = \sum_{i=1}^{I_K} K_{i,t}$ and that the share of capital in industry i is $\mathcal{K}_{i,t} = \frac{K_{i,t}}{K_t}$. We define the consumption rate as

$$x_t = \frac{C_t}{Y_t \left[1 - D_t\left(T_{1,t}, G_t(S_t), m_t\right)\right]}.$$

Homogeneity of the production function implies

$$Y_t = \mathcal{F}(\boldsymbol{A}_t, \boldsymbol{K}_t, \boldsymbol{N}_t, \boldsymbol{E}_t) = K_t^{\kappa} \mathcal{F}(\boldsymbol{A}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{N}_t, \boldsymbol{E}_t),$$

such that

$$\log C_t = \log x_t + \kappa \log K_t + \log \mathcal{F}(\boldsymbol{A}_t,\, \boldsymbol{\mathcal{K}}_t,\, \boldsymbol{N}_t,\, \boldsymbol{E}_t) + \xi_0\ (1-\tau_{1,t}) - d\, S_t - a(m_t - 1).$$

We transform the optimization problem into its dynamic programming form (Bellman equation)

$$V(k_t, \boldsymbol{\tau}_t, \boldsymbol{M}_t, \boldsymbol{R}_t, t) = \max_{x_t, \boldsymbol{N}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{E}_t, S_t} \Bigg\{ \log x_t + \kappa\, k_t + \log \mathcal{F}(\boldsymbol{A}_t,\, \boldsymbol{\mathcal{K}}_t,\, \boldsymbol{N}_t,\, \boldsymbol{E}_t) \\ + \xi_0\ (1-\tau_{1,t}) - d\, S_t - a(m_t - 1) + \beta\, V(k_{t+1}, \boldsymbol{\tau}_{t+1}, \boldsymbol{M}_{t+1}, \boldsymbol{R}_{t+1}, t+1) \Bigg\}$$

where $k_t = \log K_t$ with the equation of motion

$$k_{t+1} = \kappa\, k_t + \log \mathcal{F}(\boldsymbol{A}_t,\, \boldsymbol{\mathcal{K}}_t,\, \boldsymbol{N}_t,\, \boldsymbol{E}_t) + \log(1-x_t) + \xi_0\ (1-\tau_{1,t}) - d\, S_t - a(m_t - 1). \tag{3.19}$$

The linear affine guess for the value function

$$V(k_t, \boldsymbol{\tau}_t, \boldsymbol{M}_t, \boldsymbol{R}_t, t) = \varphi_k\, k_t + \boldsymbol{\varphi}_\tau^T\, \boldsymbol{\tau}_t + \boldsymbol{\varphi}_M^T\, \boldsymbol{M}_t + \boldsymbol{\varphi}_{R,t}^T\, \boldsymbol{R}_t + \varphi_t \tag{3.20}$$

turns the Bellman equation into the form

$$\varphi_k\, k_t + \boldsymbol{\varphi}_\tau^T\, \boldsymbol{\tau}_t + \boldsymbol{\varphi}_M^T\, \boldsymbol{M}_t + \boldsymbol{\varphi}_{R,t}^T\, \boldsymbol{R}_t + \varphi_t \\ = \max_{x_t, \boldsymbol{N}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{E}_t, S_t} \Bigg\{ \log x_t + \kappa\, k_t + \log \mathcal{F}(\boldsymbol{A}_t,\, \boldsymbol{\mathcal{K}}_t,\, \boldsymbol{N}_t,\, \boldsymbol{E}_t) + \xi_0\ (1-\tau_{1,t}) - d\, S_t - a(m_t - 1) \\ + \lambda_t^{\mathcal{K}}(1 - \sum_{i=1}^{I_K} \mathcal{K}_{i,t}) + \lambda_t^{N}(1 - \sum_{i=1}^{I_N} N_{i,t}) + \beta\, \varphi_k \left(\kappa\, k_t + \log \mathcal{F}(\boldsymbol{A}_t,\, \boldsymbol{\mathcal{K}}_t,\, \boldsymbol{N}_t,\, \boldsymbol{E}_t) + \log(1-x_t) \right. \\ \left. + \xi_0\ (1-\tau_{1,t}) - d\, S_t - a(m_t - 1) \right) + \beta\, \boldsymbol{\varphi}_\tau^T \left(\boldsymbol{\sigma}\boldsymbol{\tau}_t + \tilde{\boldsymbol{F}}_t \right) + \beta\, \boldsymbol{\varphi}_M^T\ (\boldsymbol{\Phi}\boldsymbol{M}_t + \tilde{\boldsymbol{e}}) + \beta\, \boldsymbol{\varphi}_{R,t+1}^T \left(\boldsymbol{R}_t - \boldsymbol{E}_t^d \right) \\ + \beta\, \varphi_{t+1} \Bigg\}.$$

We show that the system is linear in states and that the affine value function, (3.20), solves the system. Inserting the trial solution and the next periods states (equations 3.7,

3.9, 3.10b and 3.19) into the Bellman equation delivers

$$
\begin{aligned}
&\varphi_k\, k_t + \boldsymbol{\varphi}_\tau^T\, \boldsymbol{\tau}_t + \boldsymbol{\varphi}_M^T\, \boldsymbol{M}_t + \boldsymbol{\varphi}_{R,t}^T\, \boldsymbol{R}_t + \varphi_t \\
&= \max_{x_t, \boldsymbol{N}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{E}_t, S_t} \Bigg\{ \log x_t + \kappa\, k_t + \log \mathcal{F}(\boldsymbol{A}_t,\, \boldsymbol{\mathcal{K}}_t,\, \boldsymbol{N}_t,\, \boldsymbol{E}_t) + \xi_0\, (1 - \tau_{1,t}) - d\, S_t - a(m_t - 1) \\
&+ \lambda_t^{\mathcal{K}}(1 - \sum_{i=1}^{I_K} \mathcal{K}_{i,t}) + \lambda_t^N(1 - \sum_{i=1}^{I_N} N_{i,t}) + \beta\, \varphi_k \left(\kappa\, k_t + \log \mathcal{F}(\boldsymbol{A}_t,\, \boldsymbol{\mathcal{K}}_t,\, \boldsymbol{N}_t,\, \boldsymbol{E}_t) + \log(1 - x_t)\right. \\
&\left. + \xi_0\, (1 - \tau_{1,t}) - d\, S_t - a(m_t - 1)\right) + \beta\, \boldsymbol{\varphi}_\tau^T \left(\boldsymbol{\sigma}\boldsymbol{\tau}_t + \tilde{\boldsymbol{F}}_t\right) + \beta\, \boldsymbol{\varphi}_M^T\, (\boldsymbol{\Phi}\boldsymbol{M}_t + \tilde{\boldsymbol{e}}_{\boldsymbol{t}}) + \beta\, \boldsymbol{\varphi}_{R,t+1}^T \left(\boldsymbol{R}_t - \boldsymbol{E}_t^d\right) \\
&\qquad\qquad + \beta\, \varphi_{t+1} \Bigg\}.
\end{aligned}
$$

First order conditions. Maximizing the right hand side over x_t yields

$$
\frac{1}{x_t} - \beta\, \varphi_k \frac{1}{1 - x_t} = 0 \quad \Longrightarrow \quad x_t = \frac{1}{1 + \beta\, \varphi_k}. \tag{3.21}
$$

Maximizing the right hand side over $\mathcal{K}_{i,t}$ yields

$$
(1 + \beta\, \varphi_k) \frac{\frac{\partial \mathcal{F}(\boldsymbol{A}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{N}_t, \boldsymbol{E}_t)}{\partial \mathcal{K}_{i,t}}}{\mathcal{F}(\boldsymbol{A}_t,\, \boldsymbol{\mathcal{K}}_t,\, \boldsymbol{N}_t,\, \boldsymbol{E}_t)} = \lambda_t^{\mathcal{K}}
$$

which is equivalent to

$$
\mathcal{K}_{i,t} = \frac{\sigma_{Y,\mathcal{K}_i}(\boldsymbol{A}_t,\, \boldsymbol{\mathcal{K}}_t,\, \boldsymbol{N}_t,\, \boldsymbol{E}_t)}{\sum_{i=1}^{I_K} \sigma_{Y,\mathcal{K}_i}(\boldsymbol{A}_t,\, \boldsymbol{\mathcal{K}}_t,\, \boldsymbol{N}_t,\, \boldsymbol{E}_t)}
$$

with

$$
\sigma_{Y,\mathcal{K}_i}(\boldsymbol{A}_t,\, \boldsymbol{\mathcal{K}}_t,\, \boldsymbol{N}_t,\, \boldsymbol{E}_t) \equiv \frac{\partial \mathcal{F}(\boldsymbol{A}_t,\, \boldsymbol{\mathcal{K}}_t,\, \boldsymbol{N}_t,\, \boldsymbol{E}_t)}{\partial \mathcal{K}_{i,t}} \frac{\mathcal{K}_{i,t}}{\mathcal{F}(\boldsymbol{A}_t,\, \boldsymbol{\mathcal{K}}_t,\, \boldsymbol{N}_t,\, \boldsymbol{E}_t)}.
$$

Similarly, the first order conditions for the labor input is

$$
(1 + \beta\, \varphi_k) \frac{\frac{\partial \mathcal{F}(\boldsymbol{A}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{N}_t, \boldsymbol{E}_t)}{\partial N_{i,t}}}{\mathcal{F}(\boldsymbol{A}_t,\, \boldsymbol{\mathcal{K}}_t,\, \boldsymbol{N}_t,\, \boldsymbol{E}_t)} = \lambda_t^N
$$

and hence

$$N_{i,t} = \frac{\sigma_{Y,N_i}(\boldsymbol{A}_t,\, \boldsymbol{\mathcal{K}}_t,\, \boldsymbol{N}_t,\, \boldsymbol{E}_t)}{\sum_{i=1}^{I_N} \sigma_{Y,N_i}(\boldsymbol{A}_t,\, \boldsymbol{\mathcal{K}}_t,\, \boldsymbol{N}_t,\, \boldsymbol{E}_t)}$$

with

$$\sigma_{Y,N_i}(\boldsymbol{A}_t,\, \boldsymbol{\mathcal{K}}_t,\, \boldsymbol{N}_t,\, \boldsymbol{E}_t) \equiv \frac{\partial \mathcal{F}(\boldsymbol{A}_t,\, \boldsymbol{\mathcal{K}}_t,\, \boldsymbol{N}_t,\, \boldsymbol{E}_t)}{\partial N_{i,t}} \frac{N_{i,t}}{\mathcal{F}(\boldsymbol{A}_t,\, \boldsymbol{\mathcal{K}}_t,\, \boldsymbol{N}_t,\, \boldsymbol{E}_t)}$$

The first order condition for the optimal input of fossil fuels is given by

$$(1+\beta\,\varphi_k)\frac{\frac{\partial \mathcal{F}(\boldsymbol{A}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{N}_t, \boldsymbol{E}_t)}{\partial E_{i,t}}}{\mathcal{F}(\boldsymbol{A}_t,\, \boldsymbol{\mathcal{K}}_t,\, \boldsymbol{N}_t,\, \boldsymbol{E}_t)} = \beta(\varphi_{R,i,t+1} - \varphi_{M1})$$

which is equivalent to

$$E_{i,t} = \frac{(1+\beta\,\varphi_k)\sigma_{Y,E_i}(\boldsymbol{A}_t,\, \boldsymbol{\mathcal{K}}_t,\, \boldsymbol{N}_t,\, \boldsymbol{E}_t)}{\beta(\varphi_{R,i,t+1} - \varphi_{M1})}$$

with

$$\sigma_{Y,E_i}(\boldsymbol{A}_t,\, \boldsymbol{\mathcal{K}}_t,\, \boldsymbol{N}_t,\, \boldsymbol{E}_t) \equiv \frac{\partial \mathcal{F}(\boldsymbol{A}_t,\, \boldsymbol{\mathcal{K}}_t,\, \boldsymbol{N}_t,\, \boldsymbol{E}_t)}{\partial E_{i,t}} \frac{E_{i,t}}{\mathcal{F}(\boldsymbol{A}_t,\, \boldsymbol{\mathcal{K}}_t,\, \boldsymbol{N}_t,\, \boldsymbol{E}_t)}.$$

So far, our results are equivalent to those of the ACE model. Next, we spell out the part of the Bellman equation that depends on sulfur

$$B_t = \beta\,\varphi_{\tau 1}\,\sigma_{\text{forc}}(f_2\,S_t - f_3\,m_t^n\,S_t^{1-n}) - (1+\beta\,\varphi_k)\,d\,S_t$$

and find the first order condition for optimal sulfur deployment

$$\beta\,\varphi_{\tau 1}\,\sigma_{\text{forc}}(f_2 + (n-1)f_3\,m_t^n\,S_t^{-n}) - (1+\beta\,\varphi_k)\,d = 0.$$

Solving the first order condition for S_t gives the optimal level of sulfur deployment

$$S_t^* = \underbrace{\left(\frac{\beta\,\varphi_{\tau 1}\,\sigma_{\text{forc}}\,(n-1) f_3}{(1+\beta\,\varphi_k)\,d - \beta\,\varphi_{\tau 1}\,\sigma_{\text{forc}}\,f_2} \right)^{\frac{1}{n}}}_{\equiv z} m_t. \tag{3.22}$$

Solving the system of first order conditions gives us $\boldsymbol{N}_t^*(\boldsymbol{A}_t, \varphi_k, \boldsymbol{\varphi}_M, \boldsymbol{\varphi}_{R,t+1})$, $\boldsymbol{\mathcal{K}}_t^*(\boldsymbol{A}_t, \varphi_k, \boldsymbol{\varphi}_M, \boldsymbol{\varphi}_{R,t+1})$, and $\boldsymbol{E}_t^*(\boldsymbol{A}_t, \varphi_k, \boldsymbol{\varphi}_M, \boldsymbol{\varphi}_{R,t+1})$ which are independent of the states and $S_t^*(\varphi_k, \varphi_{\tau 1}, M_{1,t})$ which depends on the atmospheric carbon stock. In the following we show that given these optimal controls the maximized Bellman equation is linear in all states.

Shadow values. Inserting the optimal control rules into the maximized Bellman equation gives us

$$\begin{aligned}
&\varphi_k\,k_t + \boldsymbol{\varphi}_\tau^T\,\boldsymbol{\tau}_t + \boldsymbol{\varphi}_M^T\,\boldsymbol{M}_t + \boldsymbol{\varphi}_{R,t}^T\,\boldsymbol{R}_t + \varphi_t \\
&= \log x_t^* + \kappa\,k_t + \log \mathcal{F}(\boldsymbol{A}_t,\,\boldsymbol{\mathcal{K}}_t^*,\,\boldsymbol{N}_t^*,\,\boldsymbol{E}_t^*) + \xi_0\,(1-\tau_{1,t}) - d\,S_t^* - a(m_t-1) + \beta\,\varphi_k\,\big(\kappa\,k_t \\
&+ \log \mathcal{F}(\boldsymbol{A}_t,\,\boldsymbol{\mathcal{K}}_t^*,\,\boldsymbol{N}_t^*,\,\boldsymbol{E}_t^*) + \log(1-x_t^*) + \xi_0\,(1-\tau_{1,t}) - d\,S_t^* - a(m_t-1)\big) + \beta\,\boldsymbol{\varphi}_\tau^T\,\left(\boldsymbol{\sigma}\boldsymbol{\tau}_t + \tilde{\boldsymbol{F}}_t\right) \\
&\qquad + \beta\,\boldsymbol{\varphi}_M^T\,(\boldsymbol{\Phi}\boldsymbol{M}_t + \tilde{\boldsymbol{e}}_{\boldsymbol{t}}) + \beta\,\boldsymbol{\varphi}_{R,t+1}^T\,\left(\boldsymbol{R}_t - \boldsymbol{E}_t^{d*}\right) + \beta\,\varphi_{t+1}.
\end{aligned}$$

Using equation (3.22) and arranging terms with respect to their states yields

$$\begin{aligned}
\varphi_k\,k_t + \boldsymbol{\varphi}_\tau^T\,\boldsymbol{\tau}_t + \boldsymbol{\varphi}_M^T\,\boldsymbol{M}_t + \boldsymbol{\varphi}_{R,t}^T\,\boldsymbol{R}_t + \varphi_t &= \Big[(1+\beta\,\varphi_k)\kappa\Big]k_t + \Big[\beta\,\boldsymbol{\varphi}_\tau^T\,\boldsymbol{\sigma} - (1+\beta\,\varphi_k)\xi_0\,\boldsymbol{e}_1^T\Big]\boldsymbol{\tau}_t \\
&+ \Big[\beta\,\boldsymbol{\Phi}\boldsymbol{\varphi}_M^T + \Big((\beta\,\varphi_{\tau 1}\,\sigma_{\text{forc}})\left(f_1 + f_2\,z - f_3\,z^{1-n}\right) - (1+\beta\,\varphi_k)(a+d\,z)\Big)M_{pre}^{-1}\,\boldsymbol{e}_1^T\Big]\boldsymbol{M}_t \\
&+ \Big[\beta\,\boldsymbol{\varphi}_{R,t+1}^T\Big]R_t + \log\,x_t^* + \beta\,\varphi_k\,\log(1-x_t^*) + (1+\beta\,\varphi_k)\log\mathcal{F}(\boldsymbol{A}_t,\,\boldsymbol{\mathcal{K}}_t^*,\,\boldsymbol{N}_t^*,\,\boldsymbol{E}_t^*) \\
&+ (1+\beta\,\varphi_k)(\xi_0 + a) + \beta\,\varphi_{\tau 1}\,\sigma_{\text{forc}}\,f_0 + \beta\,\boldsymbol{\varphi}_M^T\,\tilde{\boldsymbol{e}}_t - \beta\,\boldsymbol{\varphi}_{R,t+1}^T\,\boldsymbol{E}_t^{d*} + \beta\,\varphi_{t+1}.
\end{aligned} \tag{3.23}$$

The system is linear in all states. Deriving both sides of the equation with respect to capital, k_t, yields

$$\varphi_k = (1+\beta\,\varphi_k)\kappa \quad \Leftrightarrow \quad \varphi_k = \frac{\kappa}{1-\beta\,\kappa} \tag{3.24}$$

Inserting φ_k into equation (3.21) yield the optimal consumption rate $x_t^* = 1 - \beta\,\kappa$.

Coefficient matching with respect to transformed temperatures delivers

$$\boldsymbol{\varphi}_{\tau}^{T} = -\xi_0 \, (1 + \beta \, \varphi_k) \, \boldsymbol{e}_1^T \, (\mathbf{1} - \beta \, \boldsymbol{\sigma})^{-1},$$

where we denote the entries of the inverted matrix as follows

$$(\mathbf{1} - \beta \, \boldsymbol{\sigma})^{-1} \equiv \begin{pmatrix} \tilde{\sigma}_{11} & \tilde{\sigma}_{12} \\ \tilde{\sigma}_{21} & \tilde{\sigma}_{22} \end{pmatrix},$$

yielding

$$\tilde{\sigma}_{11} = \left[(\mathbf{1} - \beta \, \boldsymbol{\sigma})^{-1} \right]_{1,1}, \tag{3.25}$$

where $[\cdot]_{1,1}$ denotes the first element of the inverted matrix in square brackets. Hence,

$$\varphi_{\tau 1} = -\xi_0 \, (1 + \beta \, \varphi_k) \tilde{\sigma}_{11}. \tag{3.26}$$

Coefficient matching with respect to the atmospheric carbon stock leads to

$$\varphi_M^T = \Big((\beta \, \varphi_{\tau 1} \, \sigma_{\text{forc}}) \left(f_1 + f_2 \, z - f_3 \, z^{1-n} \right) - (1 + \beta \, \varphi_k)(a + d \, z) \Big) M_{pre}^{-1} \, \boldsymbol{e}_1^T \, (\mathbf{1} - \beta \, \boldsymbol{\Phi})^{-1}. \tag{3.27}$$

We define

$$\tilde{\phi}_{ij} = \left[(\mathbf{1} - \beta \, \boldsymbol{\Phi})^{-1} \right]_{ij} \text{ for } i, j \in \{1, 2, 3\},$$

yielding

$$\tilde{\phi}_{11} = \left[(\mathbf{1} - \beta \, \boldsymbol{\Phi})^{-1} \right]_{1,1}, \tag{3.28}$$

where again $[\cdot]_{1,1}$ denotes the first element of the inverted matrix in square brackets. Coefficient matching with respect to the resource stock yields

$$\boldsymbol{\varphi}_{R,t}^T = \beta \, \boldsymbol{\varphi}_{R,t+1}^T \quad \Leftrightarrow \quad \boldsymbol{\varphi}_{R,t} = \beta^{-t} \boldsymbol{\varphi}_{R,0} \quad \text{(Hotelling's rule)}.$$

The initial resource values $\boldsymbol{\varphi}_{R,0}^T$ depend on the set up of the economy, including assumptions about production and the energy sector. Given the coefficients and the optimal rate of consumption equation (3.23) turns to the following condition:

$$\begin{aligned}\varphi_t - \beta\,\varphi_{t+1} = \log\, x_t^* + \beta\,\varphi_k\,\log(1-x_t^*) + (1+\beta\,\varphi_k)\log\mathcal{F}(\boldsymbol{A}_t,\,\boldsymbol{\mathcal{K}}_t^*,\,\boldsymbol{N}_t^*,\,\boldsymbol{E}_t^*)\\ + (1+\beta\,\varphi_k)(\xi_0+a) + \beta\,\varphi_{\tau 1}\,\sigma_{\text{forc}}\,f_0 + \beta\,\boldsymbol{\varphi}_M^T\,\tilde{\boldsymbol{e}}_t - \beta\,\boldsymbol{\varphi}_{R,t+1}^T\,\boldsymbol{E}_t^{d*}\end{aligned}$$

This condition will be satisfied by picking the sequence $\varphi_0, \varphi_1, \varphi_2, ...$ The additional condition $\lim_{t\to\infty}\beta^t V(\cdot) = 0 \Rightarrow \lim_{t\to\infty}\beta^t\varphi_t = 0$ pins down this initial value φ_0.

B.1.2.2 Proof of Proposition 3.1

In Appendix B.1.2.1 we have shown that the optimal level of sulfur is given by

$$S_t^* = \left(\frac{(n-1)\beta\,\varphi_{\tau 1}\,\sigma_{\text{forc}}\,f_3}{(1+\beta\,\varphi_k)\,d - \beta\,\varphi_{\tau 1}\,\sigma_{\text{forc}}\,f_2}\right)^{\frac{1}{n}} m_t. \tag{3.29}$$

The endogenous shadow value of capital $\varphi_k > 0$ is positive (see 3.24), while the endogenous shadow value of (transformed) temperature is negative $\varphi_{\tau 1} < 0$ (a bad, see 3.26). Therefore, both numerator and denominator are positive. The optimal level of sulfur deployment increases in the absolute value of the shadow price of atmospheric temperature.

Inserting (3.24) and (3.26) for the shadow values $\varphi_{\tau 1}$ and φ_k into our expression for sulfur deployment (3.29) and using the definition $\gamma = \beta\,\xi_0\,\tilde{\sigma}_{11}\,\sigma_{\text{forc}}$ delivers

$$S_t^* = \left(\frac{(1-n)\,\gamma\,f_3}{d+\gamma\,f_2}\right)^{\frac{1}{n}} m_t.$$

The parameter $\tilde{\sigma}_{11}$ is defined as in (3.25). Inserting the fit parameters from Table 3.2, and using the parameter values from the baseline calibration of the ACE model (re-calibrated for 2 temperature layers) from Table 3.5 leads to

$$S_t^* = \left(\frac{1.65}{16\% + 10^3 d}\right)^{1.45} m_t.$$

B.1.2.3 Proof of Proposition 3.2

Inserting equation (3.26) for the shadow value $\varphi_{\tau 1}$ (using $\gamma = \beta\,\xi_0\,\tilde{\sigma}_{11}\,\sigma_{\text{forc}}$) into equation (3.27) for the shadow value of the atmospheric carbon stock delivers

$$\varphi_{M1} = -(1+\beta\,\varphi_k)\Big(\gamma\left(f_1 + f_2\,z - f_3\,z^{1-n}\right) + a + d\,z\Big)M_{pre}^{-1}\,\tilde{\phi}_{11}.$$

Inserting equation (3.24) for the shadow value φ_k leads to

$$\varphi_{M1} = -\frac{1}{1-\beta\,\kappa}\Big(\gamma\left(f_1 + f_2\,z - f_3\,z^{1-n}\right) + a + d\,z\Big)M_{pre}^{-1}\,\tilde{\phi}_{11}.$$

The SCC is the negative of the shadow value of atmospheric carbon stock expressed in money-measured consumption units,

$$\begin{aligned} SCC &= -(1-\beta\,\kappa)Y_t^{net}\,\varphi_{M1} \\ &= \frac{Y_t^{net}}{M_{pre}}\left[a + f_1\,\gamma - \left(\left(\frac{f_3}{z^n} - f_2\right)\gamma - d\right)z\right]\tilde{\phi}_{11}. \end{aligned}$$

B.1.3 Regional model

B.1.3.1 Climate dynamics and geoengineering damages

Climate dynamics. Analogously to the global case, we let $T_{0,t}^A = \frac{s}{\eta}\,F_t^A$ and $T_{0,t}^B = \frac{s}{\eta}\,F_t^B$. The two regional atmospheric temperatures and the ocean temperature evolve as

$$\begin{aligned} T_{1,t+1}^A &= \frac{1}{\xi_1}\,\log\big((\sigma_A^A)\,\exp\left(\xi_1\,T_{1,t}^A\right) + \sigma_{\text{forc}}\,\exp\left(\xi_1\,T_{0,t}^A\right) + \sigma_B^A\,\exp\left(\xi_1\,T_{1,t}^B\right) \\ &\quad + \sigma_O^A\,\exp\left(\xi_1\,T_{2,t}\right)\big) \\ T_{1,t+1}^B &= \frac{1}{\xi_1}\,\log\big((\sigma_B^B)\,\exp\left(\xi_1\,T_{1,t}^B\right) + \sigma_{\text{forc}}\,\exp\left(\xi_1\,T_{0,t}^B\right) + \sigma_A^B\,\exp\left(\xi_1\,T_{1,t}^A\right) \\ &\quad + \sigma_O^B\,\exp\left(\xi_1\,T_{2,t}\right)\big) \\ T_{2,t+1} &= \frac{1}{\xi_1}\,\log\big((\sigma_O^O)\,\exp\left(\xi_1\,T_{2,t}\right) + \sigma_A^O\exp\left(\xi_1\,T_{1,t}^A\right) + \sigma_B^O\exp\left(\xi_1\,T_{1,t}^B\right)\big), \end{aligned}$$

where $\sigma_A^A = 1 - \sigma_B^A - \sigma_O^A - \sigma_{\text{forc}}$, $\sigma_B^B = 1 - \sigma_A^B - \sigma_O^B - \sigma_{\text{forc}}$, and $\sigma_O^O = 1 - \sigma_A^O - \sigma_B^O$.

We define transformed temperatures $\tau_{1,t}^A = \exp(\xi_1\,T_{1,t}^A)$, $\tau_{1,t}^B = \exp(\xi_1\,T_{1,t}^B)$, and $\tau_{2,t} = \exp(\xi_1\,T_{2,t})$ with $\xi_1 = \frac{\log 2}{s}$.

Transformed temperatures develop according to

$$\begin{pmatrix} \tau^A_{1,t+1} \\ \tau^B_{1,t+1} \\ \tau_{2,t+1} \end{pmatrix} = \underbrace{\begin{pmatrix} \sigma^A_A & \sigma^A_B & \sigma^A_O \\ \sigma^B_A & \sigma^B_B & \sigma^B_O \\ \sigma^O_A & \sigma^O_B & \sigma^O_O \end{pmatrix}}_{\equiv \boldsymbol{\sigma}^A} \begin{pmatrix} \tau^A_{1,t} \\ \tau^B_{1,t} \\ \tau_{2,t} \end{pmatrix} + \begin{pmatrix} \sigma_{\text{forc}} \exp(\frac{\log 2}{\eta} F^A_t) \\ \sigma_{\text{forc}} \exp(\frac{\log 2}{\eta} F^B_t) \\ 0 \end{pmatrix},$$

or equivalently

$$\boldsymbol{\tau}_{t+1} = \boldsymbol{\sigma}^A \boldsymbol{\tau}_t + \tilde{\boldsymbol{F}}_t, \quad \text{with} \quad \tilde{\boldsymbol{F}}_t = \begin{pmatrix} \tilde{F}^A_t = \sigma_{\text{forc}} \exp(\frac{\log 2}{\eta} F^A_t) \\ \tilde{F}^B_t = \sigma_{\text{forc}} \exp(\frac{\log 2}{\eta} F^B_t) \\ 0 \end{pmatrix}. \tag{3.30}$$

To preserve symmetry in notation, we define the corresponding matrix $\boldsymbol{\sigma}^B$ by swapping the first and second rows and columns, characterizing the identical dynamics from the perspective of region B.

The dynamics of the carbon reservoirs are given by

$$\begin{pmatrix} M_{1,t+1} \\ M_{2,t+1} \\ M_{3,t+1} \end{pmatrix} = \begin{pmatrix} \phi_{11} & \phi_{21} & \phi_{31} \\ \phi_{12} & \phi_{22} & \phi_{32} \\ \phi_{13} & \phi_{23} & \phi_{33} \end{pmatrix} \begin{pmatrix} M_{1,t} \\ M_{2,t} \\ M_{3,t} \end{pmatrix} + \begin{pmatrix} \sum_{i=1}^{I^d} E_{A,i,t} + \sum_{i=1}^{I^d} E_{B,i,t} + \sum_{i=1}^{I^d} E_{W,i,t} + E^{\text{exo}}_t \\ 0 \\ 0 \end{pmatrix}$$

or equivalently

$$\boldsymbol{M}_{t+1} = \boldsymbol{\Phi} \boldsymbol{M}_t + \tilde{\boldsymbol{e}}_{\boldsymbol{t}}.$$

Regional geoengineering damages. We define damages for region B symmetric to damages for region A, which we defined in the main part of the paper. Thus, damages for region B as a fraction of output are given by

$$D^B_t(\tau^B_{1,t}, S_t, m_t) = 1 - \exp\left[\xi^B_0 \left(1 - \tau^B_{1,t}\right) - \left(d_{BB} S^B_t + d_{AB}\, \alpha_A\, S^A_t\right) - a^B \left(m_t - 1\right)\right],$$

with

$$d_{BB} = \begin{cases} d^g_{BB} + \epsilon^g_B & \text{for} \quad S^B_t > 0 \\ d^c_{BB} - \epsilon^c_B & \text{for} \quad S^B_t < 0 \\ 0 & \text{for} \quad S^B_t = 0 \end{cases}$$

and

$$d_{AB} = \begin{cases} d^g_{AB} & \text{for} \quad \alpha_A \, S^A_t > 0 \\ d^c_{AB} & \text{for} \quad \alpha_A \, S^A_t < 0 \\ 0 & \text{for} \quad \alpha_A \, S^A_t = 0. \end{cases}$$

B.1.3.2 Bellman equation and Markov strategies

We show the existence of linear Markov strategies $S^i_t(m_t) = s^i_t m_t$ for $i \in \{A, B\}$ forming a subgame perfect Nash equilibrium of the dynamic game.

Bellman equation. In the following we show for region A that the system is linear in states and that the affine value function

$$V(k_t, \boldsymbol{\tau}_t, \boldsymbol{M}_t, \boldsymbol{R}_t, t) = \varphi^A_k \, k_t + \boldsymbol{\varphi}^T_{\tau A} \, \boldsymbol{\tau}_t + \boldsymbol{\varphi}^T_{MA} \, \boldsymbol{M}_t + \boldsymbol{\varphi}^T_{R,t} \, \boldsymbol{R}_t + \varphi_t,$$

with

$$\boldsymbol{\varphi}^T_{\tau A} = \begin{pmatrix} \varphi^{AA}_{\tau 1} \\ \varphi^{BA}_{\tau 1} \\ \varphi^{A}_{\tau 2} \end{pmatrix} \qquad \text{and} \qquad \boldsymbol{\varphi}^T_{MA} = \begin{pmatrix} \varphi^{A}_{M1} \\ \varphi^{A}_{M2} \\ \varphi^{A}_{M3} \end{pmatrix},$$

solves the system. The proof for region B is analogous. We suppress regional indices when there is no ambiguity to ease notation. We denote region A's shadow price for temperature $\tau^A_{1,t}$ in region A by $\varphi^{AA}_{\tau 1}$, and region A's shadow price for temperature $\tau^B_{1,t}$ in region B by $\varphi^{BA}_{\tau 1}$. Inserting the trial solution and the equations of motion for the next period's states

into the Bellman equation delivers

$$
\begin{aligned}
&\varphi_k^A\, k_t + \boldsymbol{\varphi}_{\tau A}^T\, \boldsymbol{\tau}_t + \boldsymbol{\varphi}_{MA}^T\, \boldsymbol{M}_t + \boldsymbol{\varphi}_{R,t}^T\, \boldsymbol{R}_t + \varphi_t \\
&= \max_{x_t, \boldsymbol{N}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{E}_t, S_t^A} \Bigg\{ \log x_t + \kappa\, k_t + \log \mathcal{F}(\boldsymbol{A}_t,\, \boldsymbol{\mathcal{K}}_t,\, \boldsymbol{N}_t,\, \boldsymbol{E}_t) + \xi_0^A\, \left(1 - \tau_{1,t}^A\right) - d_{AA} S_t^A(m_t) \\
&\quad + d_{BA}\alpha_B\, S_t^B(m_t) - a^A\,(m_t - 1) + \lambda_t^{\mathcal{K}}(1 - \sum_{i=1}^{I_K} \mathcal{K}_{i,t}) + \lambda_t^N(1 - \sum_{i=1}^{I_N} N_{i,t}) + \beta^A\, \varphi_k^A \Big(\kappa\, k_t \\
&+ \log \mathcal{F}(\boldsymbol{A}_t,\, \boldsymbol{\mathcal{K}}_t,\, \boldsymbol{N}_t,\, \boldsymbol{E}_t) + \log(1 - x_t) + \xi_0^A\, \left(1 - \tau_{1,t}^A\right) - d_{AA} S_t^A + d_{BA}\alpha_B\, S_t^B - a^A\,(m_t - 1) \Big) \\
&+ \beta^A\, \boldsymbol{\varphi}_{\tau A}^T \Big(\boldsymbol{\sigma}^A \boldsymbol{\tau}_t + \tilde{\boldsymbol{F}}_t(S_t^A(m_t), S_t^B(m_t)) \Big) + \beta^A\, \boldsymbol{\varphi}_{MA}^T\, (\boldsymbol{\Phi} \boldsymbol{M}_t + \tilde{\boldsymbol{e}}_{\boldsymbol{t}}) + \beta^A\, \boldsymbol{\varphi}_{R,t+1}^T\, \Big(\boldsymbol{R}_t - \boldsymbol{E}_t^d \Big) \\
&\qquad\qquad + \beta^A\, \varphi_{t+1} \Bigg\}. \qquad (3.31)
\end{aligned}
$$

First order conditions (apart from geoengineering). The first order conditions for $x_t, \boldsymbol{N}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{E}_t$ are structurally the same as in the global model.

Optimal response functions: collecting terms. The optimal (counter-) geoengineering deployment has to be compatible with the assumed strategies of both regions. Region A takes region B's strategy as given while maximizing its welfare over its own sulfur deployment (or, for $S_t^A < 0$ counter-geoengineering agent). The part of the r.h.s. Bellman equation depending on sulfur is

$$
\begin{aligned}
B_{nc}^A(m_t, S_t^A) \equiv\; & \beta^A\, \varphi_{\tau 1}^{AA}\, \tilde{F}_t^A(S_t^A, S_t^B(m_t)) + \beta^A\, \varphi_{\tau A}^{BA}\, \tilde{F}_t^B(S_t^A, S_t^B(m_t)) \\
& -(1 + \beta^A\, \varphi_k^A)[d_{AA} S_t^A + d_{BA}\, \alpha_B\, S_t^B(m_t)] \\
=\; & \underbrace{\left[\beta^A\, \varphi_{\tau 1}^{AA}\, \sigma_{\text{forc}}\, f_2 + \beta^A\, \varphi_{\tau 1}^{BA}\, \sigma_{\text{forc}}\, f_2\, \alpha_A - (1 + \beta^A\, \varphi_k^A) d_{AA}\right]}_{\equiv\, -\tilde{\delta}_A} S_t^A \\
& + \left[\beta^A\, \varphi_{\tau 1}^{AA}\, \sigma_{\text{forc}}\, f_2\, \alpha_B + \beta^A\, \varphi_{\tau 1}^{BA}\, \sigma_{\text{forc}}\, f_2 - (1 + \beta^A\, \varphi_k^A) \alpha_B d_{BA}\right] S_t^B(m_t) \\
& -\beta^A\, \varphi_{\tau 1}^{AA}\, \sigma_{\text{forc}}\, f_3\, m_t^n (S_t^A + \alpha_B\, S_t^B(m_t))^{1-n} \\
& -\beta^A\, \varphi_{\tau 1}^{BA}\, \sigma_{\text{forc}}\, f_3\, m_t^n (S_t^B(m_t) + \alpha_A\, S_t^A)^{1-n}. \qquad (3.32)
\end{aligned}
$$

We note that only the term $\tilde{\delta}_A$ depends on the damage term d_{AA}, which discretely switches sign and magnitude as the region changes action between counter-geoengineering, no action, and geoengineering at $S_t^A = 0$. All other terms are continuous. Given d_{AA} multiplies

S_t^A, also the term $\tilde{\delta}_A S_t^A$ remains continuous. The shadow prices of a temperature increase $\varphi_{\tau 1}^i < 0$, $i \in \{AA, BA\}$ are negative, therefore, $\tilde{\delta}_A > 0$.

First order condition. Differentiate (3.32) with respect to S_t^A delivers the equation

$$\frac{\partial B_{nc}^A(m_t, S_t^A)}{\partial S_t^A} = \underbrace{(n-1)\,\beta^A\,\varphi_{\tau 1}^{AA}\,\sigma_{\text{forc}}\,f_3}_{\equiv \tilde{a}_A} m_t^n (S_t^A + \alpha_B\, S_t^B(m_t))^{-n} + \underbrace{(n-1)\beta^A\,\varphi_{\tau 1}^{BA}\,\sigma_{\text{forc}}\,f_3\,\alpha_A}_{\equiv \tilde{b}_A} m_t^n (S_t^B(m_t) + \alpha_A\, S_t^A)^{-n} - \tilde{\delta}_A.$$

We note that $\tilde{a}_A, \tilde{b}_A > 0$ because $n < 1$ and the shadow prices of a temperature increase are negative. Moreover recall that region A takes region B's strategy $S_t^B(m_t) = s_t^B m_t$ as given. Defining $s^A \equiv \frac{S_t^A}{m_t}$ we rewrite the derivative as

$$\begin{aligned}\frac{\partial B_{nc}^A(m_t, S_t^A)}{\partial S_t^A} &= \tilde{a}_A m_t^n (S_t^A + \alpha_B s_t^B m_t)^{-n} + \tilde{b}_A m_t^n (m_t s_t^B + \alpha_A S_t^A)^{-n} - \tilde{\delta}_A \\ &= \tilde{a}_A (s^A + \alpha_B s_t^B)^{-n} + \tilde{b}_A (s_t^B + \alpha_A s^A)^{-n} - \tilde{\delta}_A. \end{aligned} \tag{3.33}$$

Strict concavity. The second order derivative in S_t^A is strictly negative so that the function $B_{nc}^A(m_t, S_t^A)$ is strictly concave at all points of continuity. We still have to check the discontinuity at $S_t^A = 0$ ($\Leftrightarrow s^A = 0$). The left and right limits of the objective function's slope at $s^A = 0$ are

$$\lim_{s^a \to -0} \frac{\partial B_{nc}^A(m_t, S_t^A)}{\partial S_t^A} = (\tilde{a}_A \alpha_B^{-n} + \tilde{b}_A) s_t^{B^{-n}} - \tilde{\delta}_A^c \text{ and}$$

$$\lim_{s^a \to +0} \frac{\partial B_{nc}^A(m_t, S_t^A)}{\partial S_t^A} = (\tilde{a}_A \alpha_B^{-n} + \tilde{b}_A) s_t^{B^{-n}} - \tilde{\delta}_A^g$$

where $\tilde{\delta}_A \in \{\tilde{\delta}_A^g, \tilde{\delta}_A^c\}$ was defined in equation (3.32) and depends on damages d_{AA}. The superindex on $\tilde{\delta}_A^c$ refers to the case of counter-geoengineering where $S_t^A < 0$ and $d_{AA} \equiv d_{AA}^c - \epsilon_A^c$, and $\tilde{\delta}_A^g$ refers to the case of (sulfur-based) geoengineering where $S_t^A > 0$ and $d_{AA} \equiv d_{AA}^g + \epsilon_A^g$. Because $d_{AA}^c \leq d_{AA}^g$ by Assumption 3.2 and because operational cost of geoengineering ϵ_A^g are positive, we have $\tilde{\delta}_A^c < \tilde{\delta}_A^g$ and

$$\lim_{s^a \to -0} \frac{\partial B_{nc}^A(m_t, S_t^A)}{\partial S_t^A} > \lim_{s^a \to +0} \frac{\partial B_{nc}^A(m_t, S_t^A)}{\partial S_t^A}.$$

Therefore, the function $B_{nc}^A(m_t, S_t^A)$ has a concave kink at $S_t^A = 0$.

Three qualitatively distinct response functions $S_t^A = m_t\, s_t^A$. We have the following cases for region A's optimal sulfur deployment:

$$1. \quad \text{if } \lim_{s^a \to -0} \frac{\partial B_{nc}^A(m_t, S_t^A)}{\partial S_t^A} = (\tilde{a}_A \alpha_B^{-n} + \tilde{b}_A) s_t^{B-n} - \tilde{\delta}_A^c < 0 \text{ then } s^A, S_t^A < 0,$$

the interior optimum lies to the left of the kink, and the region engages in counter-geoengineering.

$$2. \quad \text{if } \lim_{s^a \to -0} \frac{\partial B_{nc}^A(m_t, S_t^A)}{\partial S_t^A} > 0 > \lim_{s^a \to +0} \frac{\partial B_{nc}^A(m_t, S_t^A)}{\partial S_t^A}, \text{ then } S_t^A = s^A = 0,$$

$B_{nc}^A(m_t, S_t^A)$ is maximal at the kink, and the region remains inactive.

$$3. \quad \text{if } \lim_{s^a \to +0} \frac{\partial B_{nc}^A(m_t, S_t^A)}{\partial S_t^A} = (\tilde{a}_A \alpha_B^{-n} + \tilde{b}_A) s_t^{B-n} - \tilde{\delta}_A^g > 0 \text{ then } s^A, S_t^A > 0,$$

an interior optimum exists to the right of the kink, and the region engages in geoengineering.

Thus, region A's strategy is $S_t^A = s^A m_t$, consistent with our assumption that both regions follow a climate engineering strategy proportional to the CO_2 concentration m_t. We obtain the same result for region B by exchanging region labels.

Equilibrium strategies. We now solve for the proportionality constants s^i, $i \in \{A, B\}$, characterized by the optimality conditions above, such that the regions' strategies are mutually best responses.

(i) Let $S_t^A \neq 0$ and $S_t^B \neq 0$. Then we have shown that the optimal responses follow from the interior solution to the first order condition of equation (3.33)

$$\underbrace{(n-1)\,\beta^A\,\varphi_{\tau 1}^{AA}\,\sigma_{\text{forc}}\,f_3}_{\equiv \tilde{a}_A > 0}\,(s^A + \alpha_B\, s^B)^{-n} + \underbrace{(n-1)\beta^A\,\varphi_{\tau 1}^{BA}\,\sigma_{\text{forc}}\;f_3\,\alpha_A}_{\equiv \tilde{b}_A > 0}(s^B + \alpha_A\, s^A)^{-n} - \tilde{\delta}_A = 0 \tag{3.34}$$

$$\underbrace{(n-1)\,\beta^B\,\varphi_{\tau 1}^{BB}\,\sigma_{\text{forc}}\,f_3}_{\equiv \tilde{a}_B > 0}\,(s^B + \alpha_A\, s^A)^{-n} + \underbrace{(n-1)\beta^B\,\varphi_{\tau 1}^{AB}\,\sigma_{\text{forc}}\;f_3\,\alpha_B}_{\equiv \tilde{b}_B > 0}(s^A + \alpha_B\, s^B)^{-n} - \tilde{\delta}_B = 0 \tag{3.35}$$

Similarly as for region A, we denoted the shadow price of region B for temperature $\tau_{1,t}^{B}$ by $\varphi_{\tau 1}^{BB}$, and the shadow price of region B for temperature $\tau_{1,t}^{A}$ by $\varphi_{\tau 1}^{AB}$. Rearranging (3.35) leads to

$$(s^A + \alpha_B\, s^B)^{-n} = \frac{\tilde{\delta}_B - \tilde{a}_B (s^B + \alpha_A\, s^A)^{-n}}{\tilde{b}_B}$$

Using this result in (3.34) yields

$$(s^B + \alpha_A\, s^A)^{-n} = \frac{\tilde{\delta}_A\, \tilde{b}_B - \tilde{a}_A\, \tilde{\delta}_B}{\tilde{b}_A\, \tilde{b}_B - \tilde{a}_A\, \tilde{a}_B}.$$

From this we get

$$s^B = \left(\frac{\tilde{\delta}_A\, \tilde{b}_B - \tilde{a}_A\, \tilde{\delta}_B}{\tilde{b}_A\, \tilde{b}_B - \tilde{a}_A\, \tilde{a}_B} \right)^{-\frac{1}{n}} - \alpha_A\, s^A$$

and

$$s^A = \left(\frac{\tilde{\delta}_B\, \tilde{b}_A - \tilde{a}_B\, \tilde{\delta}_A}{\tilde{b}_B\, \tilde{b}_A - \tilde{a}_B\, \tilde{a}_A} \right)^{-\frac{1}{n}} - \alpha_B\, s^B$$

and hence

$$s^B = \frac{1}{1 - \alpha_A\, \alpha_B} \left[\underbrace{\left(\frac{\tilde{\delta}_A\, \tilde{b}_B - \tilde{a}_A\, \tilde{\delta}_B}{\tilde{b}_A\, \tilde{b}_B - \tilde{a}_A\, \tilde{a}_B} \right)^{-\frac{1}{n}}}_{\equiv z_B} - \alpha_A \underbrace{\left(\frac{\tilde{\delta}_B\, \tilde{b}_A - \tilde{a}_B\, \tilde{\delta}_A}{\tilde{b}_B\, \tilde{b}_A - \tilde{a}_B\, \tilde{a}_A} \right)^{-\frac{1}{n}}}_{\equiv z_A} \right]$$

and analogously

$$s^A = \frac{1}{1 - \alpha_A\, \alpha_B} \left[\underbrace{\left(\frac{\tilde{\delta}_B\, \tilde{b}_A - \tilde{a}_B\, \tilde{\delta}_A}{\tilde{b}_B\, \tilde{b}_A - \tilde{a}_B\, \tilde{a}_A} \right)^{-\frac{1}{n}}}_{\equiv z_A} - \alpha_B \underbrace{\left(\frac{\tilde{\delta}_A\, \tilde{b}_B - \tilde{a}_A\, \tilde{\delta}_B}{\tilde{b}_A\, \tilde{b}_B - \tilde{a}_A\, \tilde{a}_B} \right)^{-\frac{1}{n}}}_{\equiv z_B} \right].$$

We further define

$$z_A^g = \left(\frac{\tilde{\delta}_B\, \tilde{b}_A - \tilde{a}_B\, \tilde{\delta}_A^g}{\tilde{b}_B\, \tilde{b}_A - \tilde{a}_B\, \tilde{a}_A} \right)^{-\frac{1}{n}} \qquad \text{for} \qquad S_t^A > 0$$

$$z_A^c = \left(\frac{\tilde{\delta}_B\, \tilde{b}_A - \tilde{a}_B\, \tilde{\delta}_A^c}{\tilde{b}_B\, \tilde{b}_A - \tilde{a}_B\, \tilde{a}_A} \right)^{-\frac{1}{n}} \qquad \text{for} \qquad S_t^A < 0$$

This gives us

$$S_t^A = \frac{m_t}{1 - \alpha_A \, \alpha_B} \left(z_A^g - \alpha_B \, z_B \right) \qquad \text{for} \qquad S_t^A > 0$$

and

$$S_t^A = \frac{m_t}{1 - \alpha_A \, \alpha_B} \left(z_A^c - \alpha_B \, z_B \right) \qquad \text{for} \qquad S_t^A < 0$$

(ii) In the second case where $S_t^A > 0$ and $S_t^B = 0$, the first order condition for region A simplifies to

$$\tilde{a}_A \, (s^A)^{-n} + \tilde{b}_A (\alpha_A \, s^A)^{-n} - \tilde{\delta}_A^g = 0$$

which is equivalent to

$$s^A = \left(\frac{\tilde{\delta}_A^g}{\tilde{b}_A \, \alpha_A^{-n} + \tilde{a}_A} \right)^{-\frac{1}{n}}$$

and gives us

$$S_t^A = z_A^g \, m_t$$

with

$$z_A^g = \left(\frac{\tilde{\delta}_A^g}{\tilde{b}_A \, \alpha_A^{-n} + \tilde{a}_A} \right)^{-\frac{1}{n}}.$$

(iii) The last possible case is $S_t^A = 0$ and $S_t^B > 0$, which is symmetric to case (ii).

Summary of strategies. In conclusion, the following reaction functions characterize a Nash equilibrium of the dynamic game: If $S_t^B = 0$, region A chooses $S_t^A = z_A^g m_t$ and if $S_t^B \neq 0$ region A chooses

$$\begin{aligned}
S_t^A &= \frac{m_t}{1 - \alpha_A \, \alpha_B} \left(z_A^g - \alpha_B \, z_B \right) \quad \text{for} \quad S_t^A > 0 \\
S_t^A &= \frac{m_t}{1 - \alpha_A \, \alpha_B} \left(z_A^c - \alpha_B \, z_B \right) \quad \text{for} \quad S_t^A < 0 \\
S_t^A &= 0 \quad \text{otherwise.}
\end{aligned}$$

Swapping country indices characterizes region B's strategies.

B.1.3.3 Verifying solution to the Bellman equation.

Inserting the optimal control rules $\boldsymbol{N}_t^*(\boldsymbol{A}_t, \varphi_k^A, \boldsymbol{\varphi}_{MA}, \boldsymbol{\varphi}_{R,t+1})$, $\boldsymbol{\mathcal{K}}_t^*(\boldsymbol{A}_t, \varphi_k^A, \boldsymbol{\varphi}_{MA}, \boldsymbol{\varphi}_{R,t+1})$, $\boldsymbol{E}_t^*(\boldsymbol{A}_t, \varphi_k^A, \boldsymbol{\varphi}_{MA}, \boldsymbol{\varphi}_{R,t+1})$, and $S_t^{A^*}(m_t)$, which are analogous to the global solutions characterized in equations (3.21) to (3.22), into the maximized Bellman equation (3.31) yields

$$
\begin{aligned}
&\varphi_k^A\, k_t + \boldsymbol{\varphi}_{\tau A}^T\, \boldsymbol{\tau}_t + \boldsymbol{\varphi}_{MA}^T\, \boldsymbol{M}_t + \boldsymbol{\varphi}_{R,t}^T\, \boldsymbol{R}_t + \varphi_t \\
&= \log x_t^* + \kappa\, k_t + \log \mathcal{F}(\boldsymbol{A}_t,\, \boldsymbol{\mathcal{K}}_t^*,\, \boldsymbol{N}_t^*,\, \boldsymbol{E}_t^*) + \xi_0^A\, \left(1 - \tau_{1,t}^A\right) - d_{AA} S^{A^*}{}_t(m_t) + d_{BA}\, \alpha_B\, S_t^B(m_t) \\
&\quad - a^A\, (m_t - 1) + \beta^A\, \varphi_k^A \Big(\kappa\, k_t + \log \mathcal{F}(\boldsymbol{A}_t,\, \boldsymbol{\mathcal{K}}_t^*,\, \boldsymbol{N}_t^*,\, \boldsymbol{E}_t^*) + \log(1 - x_t^*) + \xi_0^A\, \left(1 - \tau_{1,t}^A\right) \\
&- d_{AA} S_t^{A^*}(m_t) + d_{BA}\, \alpha_B\, S_t^B(m_t) - a^A\, (m_t - 1) \Big) + \beta^A\, \boldsymbol{\varphi}_{\tau A}^T \left(\boldsymbol{\sigma}^A \boldsymbol{\tau}_t + \tilde{\boldsymbol{F}}_t(S_t^{A^*}(m_t), S_t^B(m_t)) \right) \\
&\quad + \beta^A\, \boldsymbol{\varphi}_{MA}^T\, \left(\boldsymbol{\Phi} \boldsymbol{M}_t + \tilde{\boldsymbol{e}}_{\boldsymbol{t}} \right) + \beta^A\, \boldsymbol{\varphi}_{R,t+1}^T\, \left(\boldsymbol{R}_t - \boldsymbol{E}_t^{d^*} \right) + \beta^A\, \varphi_{t+1} \qquad (3.36)
\end{aligned}
$$

where $\tilde{\boldsymbol{F}}_t$ is the forcing vector defined in equation (3.30) making its sulfur dependencies explicit. Arranging terms with respect to states for the different Nash equilibria yields

(i): $S_t^A \neq 0, S_t^B \neq 0$

$$
\begin{aligned}
&\varphi_k^A\, k_t + \boldsymbol{\varphi}_{\tau A}^T\, \boldsymbol{\tau}_t + \boldsymbol{\varphi}_{MA}^T\, \boldsymbol{M}_t + \boldsymbol{\varphi}_{R,t}^T\, \boldsymbol{R}_t + \varphi_t = \Big[(1 + \beta^A\, \varphi_k^A) \kappa \Big] k_t \\
&+ \Big[\beta^A\, \boldsymbol{\varphi}_{\tau A}^T\, \boldsymbol{\sigma}^A - (1 + \beta^A\, \varphi_k^A) \xi_0^A\, \boldsymbol{e}_1^T \Big] \boldsymbol{\tau}_t + \Big[\beta^A\, \boldsymbol{\varphi}_{MA}^T\, \boldsymbol{\Phi} + \Big(\beta^A\, \varphi_{\tau 1}^{AA}\, \sigma_{\text{forc}} (f_1 + f_2\, z_A - f_3\, z_A^{1-n}) \\
&+ \beta^A\, \varphi_{\tau 1}^{BA}\, \sigma_{\text{forc}}\, (f_1 + f_2\, z_B - f_3\, z_B^{1-n}) - (1 + \beta^A\, \varphi_k^A) \Big(d_{AA}\, \frac{1}{1 - \alpha_A\, \alpha_B} (z_A - \alpha_B\, z_B) \\
&+ d_{BA}\, \frac{\alpha_B}{1 - \alpha_A\, \alpha_B} (z_B - \alpha_A\, z_A) + a^A \Big) \Big) M_{pre}^{-1}\, \boldsymbol{e}_1^T \Big] \boldsymbol{M}_t + \Big[\beta^A\, \boldsymbol{\varphi}_{R,t+1}^T \Big] R_t + \log\, x_t^* \\
&+ \beta^A\, \varphi_k^A\, \log(1 - x_t^*) + (1 + \beta^A\, \varphi_k^A) \log \mathcal{F}(\boldsymbol{A}_t,\, \boldsymbol{\mathcal{K}}_t^*,\, \boldsymbol{N}_t^*,\, \boldsymbol{E}_t^*) + (1 + \beta^A\, \varphi_k^A)(\xi_0^A + a^A) \\
&+ (\beta^A\, \varphi_{\tau 1}^{AA}\, \sigma_{\text{forc}} + \beta^A\, \varphi_{\tau 1}^{BA}\, \sigma_{\text{forc}})\, f_0 + \beta^A\, \boldsymbol{\varphi}_{MA}^T\, \tilde{\boldsymbol{e}}_{\boldsymbol{t}} - \beta^A\, \boldsymbol{\varphi}_{R,t+1}^T\, \boldsymbol{E}_t^{d^*} + \beta^A\, \varphi_{t+1}, \qquad (3.37)
\end{aligned}
$$

(ii): $S_t^A > 0, S_t^B = 0$

$$\varphi_k^A\, k_t + \boldsymbol{\varphi}_{\tau A}^T\, \boldsymbol{\tau}_t + \boldsymbol{\varphi}_{MA}^T\, \boldsymbol{M}_t + \boldsymbol{\varphi}_{R,t}^T\, \boldsymbol{R}_t + \varphi_t = \Big[(1+\beta^A\,\varphi_k^A)\kappa\Big] k_t$$
$$+\Big[\beta^A\, \boldsymbol{\varphi}_{\tau A}^T\, \boldsymbol{\sigma}^A - (1+\beta^A\,\varphi_k^A)\xi_0^A\, \boldsymbol{e}_1^T\Big]\boldsymbol{\tau}_t + \Big[\beta^A\, \boldsymbol{\varphi}_{MA}^T\, \boldsymbol{\Phi} + \Big(\beta^A\, \varphi_{\tau 1}^{AA}\, \sigma_{\text{forc}}(f_1 + f_2\, z_A^g$$
$$-f_3\,(z_A^g)^{1-n}) + \beta^A\, \varphi_{\tau 1}^{BA}\, \sigma_{\text{forc}}(f_1 + f_2\, \alpha_A\, z_A^g - f_3\,(\alpha_A\, z_A^g)^{1-n})$$
$$-(1+\beta^A\,\varphi_k^A)(d_{AA}\, z_A^g + a^A)\Big) M_{pre}^{-1}\, \boldsymbol{e}_1^T\Big]\boldsymbol{M}_t + \Big[\beta^A\, \boldsymbol{\varphi}_{R,t+1}^T\Big] R_t + \log\, x_t^*$$
$$+\beta^A\, \varphi_k^A\, \log(1-x_t^*) + (1+\beta^A\,\varphi_k^A)\log \mathcal{F}(\boldsymbol{A}_t,\, \boldsymbol{\mathcal{K}}_t^*,\, \boldsymbol{N}_t^*,\, \boldsymbol{E}_t^*) + (1+\beta^A\,\varphi_k^A)(\xi_0^A + a^A)$$
$$+(\beta^A\, \varphi_{\tau 1}^{AA}\, \sigma_{\text{forc}} + \beta^A\, \varphi_{\tau 1}^{BA}\, \sigma_{\text{forc}})\, f_0 + \beta^A\, \boldsymbol{\varphi}_{MA}^T\, \tilde{\boldsymbol{e}}_{\boldsymbol{t}} - \beta^A\, \boldsymbol{\varphi}_{R,t+1}^T\, \boldsymbol{E}_t^{d*} + \beta^A\, \varphi_{t+1}, \qquad (3.38)$$

(iii): $S_t^A = 0, S_t^B > 0$

$$\varphi_k^A\, k_t + \boldsymbol{\varphi}_{\tau A}^T\, \boldsymbol{\tau}_t + \boldsymbol{\varphi}_{MA}^T\, \boldsymbol{M}_t + \boldsymbol{\varphi}_{R,t}^T\, \boldsymbol{R}_t + \varphi_t = \Big[(1+\beta^A\,\varphi_k^A)\kappa\Big] k_t$$
$$+\Big[\beta^A\, \boldsymbol{\varphi}_{\tau A}^T\, \boldsymbol{\sigma}^A - (1+\beta^A\,\varphi_k^A)\xi_0^A\, \boldsymbol{e}_1^T\Big]\boldsymbol{\tau}_t + \Big[\beta^A\, \boldsymbol{\varphi}_{MA}^T\, \boldsymbol{\Phi} + \Big(\beta^A\, \varphi_{\tau 1}^{AA}\, \sigma_{\text{forc}}(f_1 + f_2\, \alpha_B\, z_B^g$$
$$-f_3\,(\alpha_B\, z_B^g)^{1-n}) + \beta^A\, \varphi_{\tau 1}^{BA}\, \sigma_{\text{forc}}(f_1 + f_2\, z_B^g - f_3\,(z_B^g)^{1-n})$$
$$-(1+\beta^A\,\varphi_k^A)(d_{BA}\, \alpha_B\, z_B^g + a^A)\Big) M_{pre}^{-1}\, \boldsymbol{e}_1^T\Big]\boldsymbol{M}_t + \Big[\beta^A\, \boldsymbol{\varphi}_{R,t+1}^T\Big] R_t + \log\, x_t^*$$
$$+\beta^A\, \varphi_k^A\, \log(1-x_t^*) + (1+\beta^A\,\varphi_k^A)\log \mathcal{F}(\boldsymbol{A}_t,\, \boldsymbol{\mathcal{K}}_t^*,\, \boldsymbol{N}_t^*,\, \boldsymbol{E}_t^*) + (1+\beta^A\,\varphi_k^A)(\xi_0^A + a^A)$$
$$+(\beta^A\, \varphi_{\tau 1}^{AA}\, \sigma_{\text{forc}} + \beta^A\, \varphi_{\tau 1}^{BA}\, \sigma_{\text{forc}})\, f_0 + \beta^A\, \boldsymbol{\varphi}_{MA}^T\, \tilde{\boldsymbol{e}}_{\boldsymbol{t}} - \beta^A\, \boldsymbol{\varphi}_{R,t+1}^T\, \boldsymbol{E}_t^{d*} + \beta^A\, \varphi_{t+1}. \qquad (3.39)$$

Hence, for all Nash equilibria the controlled dynamics remain linear in states.

B.1.3.4 Shadow values of the states.

Coefficient matching with respect to capital k_t yields

$$\varphi_k^A = (1+\beta^A\,\varphi_k^A)\kappa \quad \Leftrightarrow \quad \varphi_k^A = \frac{\kappa}{1-\beta^A\,\kappa}.$$

Structurally similar to the global model we get the consumption rate $x_t^* = 1-\beta^A\,\kappa$.

Coefficient matching with respect to transformed temperatures delivers

$$\boldsymbol{\varphi}_{\tau A}^T = -\xi_0^A\,(1+\beta^A\,\varphi_k^A)\,\boldsymbol{e}_1^T\,[\mathbf{1}-\beta^A\,\boldsymbol{\sigma}^A]^{-1}.$$

We define

$$\tilde{\sigma}^A_{ij} = \left[(\mathbf{1} - \beta^A \boldsymbol{\sigma^A})^{-1}\right]_{ij} \text{ and } \tilde{\sigma}^B_{ij} = \left[(\mathbf{1} - \beta^B \boldsymbol{\sigma^B})^{-1}\right]_{ij} \text{ for } i, j \in \{1, 2\}.$$

Note that to preserve symmetry in notation, we defined the corresponding matrix $\boldsymbol{\sigma}^B$ by swapping the first and second rows and columns of the matrix $\boldsymbol{\sigma^A}$.

Region A's shadow values of atmospheric temperature in regions A and B are therefore

$$\varphi^{AA}_{\tau 1} = -\xi^A_0 \,(1 + \beta^A \,\varphi^A_k)\tilde{\sigma}^A_{11}, \quad \text{and} \quad \varphi^{BA}_{\tau 1} = -\xi^A_0 \,(1 + \beta^A \,\varphi^A_k)\tilde{\sigma}^A_{12}.$$

The temperature shadow values for region B follow by switching region indices. We now define $\gamma_A \equiv \beta^A \,\xi^A_0 \,\tilde{\sigma}^A_{11} \,\sigma_{\text{forc}}$, and $\gamma^{heat}_A \equiv \beta^A \,\xi^A_0 \,\tilde{\sigma}^A_{12} \,\sigma_{\text{forc}}$ (and analogously γ_B, and γ^{heat}_B). This gives us $\tilde{\delta}_A$, $\tilde{a}_A$, and $\tilde{b}_A$ (which we defined in equations 3.32 to 3.35) as a function of γ_A, γ^{heat}_A, γ_B, and γ^{heat}_B:

$$-\tilde{\delta}_A = (1 + \beta^A \,\varphi^A_k)(-\gamma_A \,f_2 - \alpha_A \,\gamma^{heat}_A \,f_2 - d_{AA})$$

$$\tilde{a}_A = (1 + \beta^A \,\varphi^A_k)(1 - n)\gamma_A \,f_3$$

$$\tilde{b}_A = (1 + \beta^A \,\varphi^A_k)(1 - n)\alpha_A \,\gamma^{heat}_A \,f_3$$

By switching region indices we gain $\tilde{\delta}_B$, $\tilde{a}_B$, and $\tilde{b}_B$. Thus, z_A and z_B are also functions of γ_A, γ^{heat}_A, γ_B, and γ^{heat}_B:

$$z_A(\gamma_A, \gamma^{heat}_A, \gamma_B, \gamma^{heat}_B) = \Big(\frac{\delta_B(\gamma_B, \gamma^{heat}_B)\, b_A(\gamma^{heat}_A) - a_B(\gamma_B)\, \delta_A(\gamma_A, \gamma^{heat}_A)}{b_B(\gamma^{heat}_B)\, b_A(\gamma^{heat}_A) - a_B(\gamma_B)\, a_A(\gamma_A)}\Big)^{-\frac{1}{n}}$$

$$z_B(\gamma_B, \gamma^{heat}_B, \gamma_A, \gamma^{heat}_A) = \Big(\frac{\delta_A(\gamma_A, \gamma^{heat}_A)\, b_B(\gamma^{heat}_B) - a_A(\gamma_A)\, \delta_B(\gamma_B, \gamma^{heat}_B)}{b_A(\gamma^{heat}_A)\, b_B(\gamma^{heat}_B) - a_A(\gamma_A)\, a_B(\gamma_B)}\Big)^{-\frac{1}{n}}$$

We note that the terms $(1 + \beta^A \,\varphi^A_k)$, and $(1 + \beta^B \,\varphi^B_k)$ cancel out and thus

$$\delta_A = \gamma_A \,f_2 + \alpha_A \,\gamma^{heat}_A \,f_2 + d_{AA}$$

$$a_A = (1-n)\gamma_A\, f_3$$

$$b_A = (1-n)\alpha_A\, \gamma_A^{heat}\, f_3.$$

The terms δ_B, a_B, and b_B are defined analogously.

Coefficient matching with respect to the carbon stocks, and using γ_A, and γ_A^{heat} yields

(i): $S_t^A \neq 0$ and $S_t^B \neq 0$

$$\boldsymbol{\varphi}_{MA}^T = (1+\beta^A\,\varphi_k^A)\Bigg(-f_2\,(\gamma_A\,z_A+\gamma_A^{heat}\,z_B)-\frac{d_{AA}\,(z_A-\alpha_B\,z_B)+\alpha_B\,d_{BA}(z_B-\alpha_A\,z_A)}{1-\alpha_A\,\alpha_B}$$
$$+\,f_3\,(\gamma_A\,z_A^{1-n}+\gamma_A^{heat}\,z_B^{1-n})-a^A-f_1\,(\gamma_A+\gamma_A^{heat})\Bigg)M_{pre}^{-1}\,\boldsymbol{e}_1^T\,(\mathbf{1}-\beta^A\,\boldsymbol{\Phi})^{-1}, \quad (3.40)$$

(ii): $S_t^A > 0$ and $S_t^B = 0$

$$\boldsymbol{\varphi}_{MA}^T = (1+\beta^A\,\varphi_k^A)\Bigg(-f_2\,(\gamma_A\,z_A^g+\gamma_A^{heat}\,\alpha_A\,z_A^g)-d_{AA}\,z_A^g+f_3\Big(\gamma_A\,(z_A^g)^{1-n}$$
$$+\,\gamma_A^{heat}\,(\alpha_A\,z_A^g)^{1-n}\Big)-a^A-f_1\,(\gamma_A+\gamma_A^{heat})\Bigg)M_{pre}^{-1}\,\boldsymbol{e}_1^T\,(\mathbf{1}-\beta^A\,\boldsymbol{\Phi})^{-1}, \quad (3.41)$$

(iii): $S_t^A = 0$ and $S_t^B > 0$

$$\boldsymbol{\varphi}_{MA}^T = (1+\beta^A\,\varphi_k^A)\Bigg(-f_2\,(\gamma_A\,\alpha_B\,z_B^g+\gamma_A^{heat}\,z_B^g)-\,d_{BA}\,\alpha_B\,z_B^g+f_3\Big(\gamma_A\,(\alpha_B\,z_B^g)^{1-n}$$
$$+\,\gamma_A^{heat}\,(z_B^g)^{1-n}\Big)-a^A-f_1\,(\gamma_A+\gamma_A^{heat})\Bigg)M_{pre}^{-1}\,\boldsymbol{e}_1^T\,(\mathbf{1}-\beta^A\,\boldsymbol{\Phi})^{-1}. \quad (3.42)$$

We define

$$\tilde{\phi}_{ij}^A = \big[(\mathbf{1}-\beta^A\,\boldsymbol{\Phi})^{-1}\big]_{ij} \text{ for } i,j \in \{1,2,3\},$$

yielding

$$\tilde{\phi}_{11}^A = \big[(\mathbf{1}-\beta^A\,\boldsymbol{\Phi})^{-1}\big]_{1,1}. \quad (3.43)$$

Switching region indices delivers the symmetric result for region B.

Coefficient matching with respect to the resource stock leads to

$$\boldsymbol{\varphi}_{R,t}^{T} = \beta\, \boldsymbol{\varphi}_{R,t+1}^{T} \quad \Leftrightarrow \quad \boldsymbol{\varphi}_{R,t} = \beta^{-t} \boldsymbol{\varphi}_{R,0} \quad \text{(Hotelling's rule)}.$$

The initial resource values $\boldsymbol{\varphi}_{R,0}^{T}$ depend on the set up of the economy, including assumptions about production and the energy sector. Given the coefficients and the optimal rate of consumption equations (3.37), (3.38), and (3.39) turn to the following condition:

$$\begin{aligned} \varphi_t - \beta^A\, \varphi_{t+1} &= \log\, x_t^* + \beta^A\, \varphi_k^A\, \log(1 - x_t^*) + (1 + \beta^A\, \varphi_k^A) \log \mathcal{F}(\boldsymbol{A}_t,\, \boldsymbol{\mathcal{K}}_t^*,\, \boldsymbol{N}_t^*,\, \boldsymbol{E}_t^*) \\ &+ (1 + \beta^A\, \varphi_k^A)(\xi_0^A + a^A) + (\beta^A\, \varphi_{\tau 1}^{AA}\, \sigma_{\text{forc}} + \beta^A\, \varphi_{\tau 1}^{BA}\, \sigma_{\text{forc}})\, f_0 + \beta\, \boldsymbol{\varphi}_{MA}^{T}\, \tilde{\boldsymbol{e}}_{\boldsymbol{t}} - \beta^A\, \boldsymbol{\varphi}_{R,t+1}^{T}\, \boldsymbol{E}_t^{d*} \end{aligned} \tag{3.44}$$

This condition will be satisfied by picking the sequence $\varphi_0, \varphi_1, \varphi_2, ...$ The additional condition $\lim_{t\to\infty}(\beta^A)^t V(\cdot) = 0 \Rightarrow \lim_{t\to\infty}(\beta^A)^t \varphi_t = 0$ pins down this initial value φ_0.

B.1.3.5 Rest of the world

In the following we show for the rest of the world that the system is linear in states and that the affine value function

$$V(k_t, \boldsymbol{\tau}_t, \boldsymbol{M}_t, \boldsymbol{R}_t, t) = \varphi_k^W\, k_t + \boldsymbol{\varphi}_{\tau W}^{T}\, \boldsymbol{\tau}_t + \boldsymbol{\varphi}_{MW}^{T}\, \boldsymbol{M}_t + \boldsymbol{\varphi}_{R,t}^{T}\, \boldsymbol{R}_t + \varphi_t, \tag{3.45}$$

where

$$\boldsymbol{\varphi}_{\tau W}^{T} = \begin{pmatrix} \varphi_{\tau 1}^{AW} \\ \varphi_{\tau 1}^{BW} \\ \varphi_{\tau 2}^{W} \end{pmatrix} \qquad \text{and} \qquad \boldsymbol{\varphi}_{MW}^{T} = \begin{pmatrix} \varphi_{M1}^{W} \\ \varphi_{M2}^{W} \\ \varphi_{M3}^{W} \end{pmatrix},$$

solves the system. We suppress regional indices when there is no ambiguity to ease notation. Inserting the trial solution and the next periods states into the Bellman equation

delivers

$$
\begin{aligned}
&\varphi_k^W\, k_t + \boldsymbol{\varphi}_{\tau W}^T\, \boldsymbol{\tau}_t + \boldsymbol{\varphi}_{MW}^T\, \boldsymbol{M}_t + \boldsymbol{\varphi}_{R,t}^T\, \boldsymbol{R}_t + \varphi_t \\
&= \max_{x_t, \boldsymbol{N}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{E}_t} \Bigg\{ \log x_t + \kappa\, k_t + \log \mathcal{F}(\boldsymbol{A}_t,\, \boldsymbol{\mathcal{K}}_t,\, \boldsymbol{N}_t,\, \boldsymbol{E}_t) + \xi_0^W\, \left(1 - \tau_{1,t}^B\right) \\
&-(d_{BW} S_t^B + d_{AW}\, \alpha_A\, S_t^A) - a^W\, (m_t - 1) + \lambda_t^{\mathcal{K}}(1 - \sum_{i=1}^{I_K} \mathcal{K}_{i,t}) + \lambda_t^N (1 - \sum_{i=1}^{I_N} N_{i,t}) \\
&+\beta^W\, \varphi_k^W \left(\kappa\, k_t + \log \mathcal{F}(\boldsymbol{A}_t,\, \boldsymbol{\mathcal{K}}_t,\, \boldsymbol{N}_t,\, \boldsymbol{E}_t) + \log(1 - x_t) + \xi_0^W\, \left(1 - \tau_{1,t}^B\right)\right. \\
&\left.-(d_{BW} S_t^B + d_{AW}\, \alpha_A\, S_t^A) - a^W\, (m_t - 1)\right) + \beta^W\, \boldsymbol{\varphi}_{\tau W}^T \left(\boldsymbol{\sigma}^A \boldsymbol{\tau}_t + \tilde{\boldsymbol{F}}_t\right) \\
&+\beta^W\, \boldsymbol{\varphi}_{MW}^T\, (\boldsymbol{\Phi} \boldsymbol{M}_t + \tilde{\boldsymbol{e}}_{\boldsymbol{t}}) + \beta^W\, \boldsymbol{\varphi}_{R,t+1}^T \left(\boldsymbol{R}_t - \boldsymbol{E}_t^d\right) + \beta^W\, \varphi_{t+1} \Bigg\}.
\end{aligned}
$$

First order conditions. The first order conditions for $x_t, \boldsymbol{N}_t, \boldsymbol{\mathcal{K}}_t, \boldsymbol{E}_t$ are structurally the same as in the global model.

Verifying solution to the Bellman equation. Inserting the optimal control rules $\boldsymbol{N}_t^*(\boldsymbol{A}_t, \varphi_k^W, \boldsymbol{\varphi}_{MW}, \boldsymbol{\varphi}_{R,t+1})$, $\boldsymbol{\mathcal{K}}_t^*(\boldsymbol{A}_t, \varphi_k^W, \boldsymbol{\varphi}_{MW}, \boldsymbol{\varphi}_{R,t+1})$, and $\boldsymbol{E}_t^*(\boldsymbol{A}_t, \varphi_k^W, \boldsymbol{\varphi}_{MW}, \boldsymbol{\varphi}_{R,t+1})$ into the maximized Bellman equation gives us

$$
\begin{aligned}
&\varphi_k^W\, k_t + \boldsymbol{\varphi}_{\tau W}^T\, \boldsymbol{\tau}_t + \boldsymbol{\varphi}_{MW}^T\, \boldsymbol{M}_t + \boldsymbol{\varphi}_{R,t}^T\, \boldsymbol{R}_t + \varphi_t = \log x_t^* + \kappa\, k_t + \log F(\boldsymbol{A}_t,\, \boldsymbol{\mathcal{K}}_t^*,\, \boldsymbol{N}_t^*,\, \boldsymbol{E}_t^*) \\
&+\xi_0^W\, \left(1 - \tau_{1,t}^B\right) - (d_{BW} S_t^B + d_{AW}\, \alpha_A\, S_t^A) - a^W\, (m_t - 1) + \beta^W\, \varphi_k^W \left(\kappa\, k_t + \log F(\boldsymbol{A}_t,\, \boldsymbol{\mathcal{K}}_t^*,\, \boldsymbol{N}_t^*,\, \boldsymbol{E}_t^*)\right. \\
&\left.+\log(1 - x_t^*) + \xi_0^W\, \left(1 - \tau_{1,t}^B\right) - (d_{BW} S_t^B + d_{AW}\, \alpha_A\, S_t^A) - a^W\, (m_t - 1)\right) + \beta^W\, \boldsymbol{\varphi}_{\tau W}^T \left(\boldsymbol{\sigma}^A \boldsymbol{\tau}_t + \tilde{\boldsymbol{F}}_t\right) \\
&+ \beta^W\, \boldsymbol{\varphi}_{MW}^T\, (\boldsymbol{\Phi} \boldsymbol{M}_t + \tilde{\boldsymbol{e}}_{\boldsymbol{t}}) + \beta^W\, \boldsymbol{\varphi}_{R,t+1}^T \left(\boldsymbol{R}_t - \boldsymbol{E}_t^{d*}\right) + \beta^W\, \varphi_{t+1} \qquad (3.46)
\end{aligned}
$$

Arranging terms by states for the different Nash equilibria yields

(i): $S_t^A \neq 0, S_t^B \neq 0$

$$\varphi_k^W\, k_t + \boldsymbol{\varphi}_{\tau W}^T\, \boldsymbol{\tau}_t + \boldsymbol{\varphi}_{MW}^T\, \boldsymbol{M}_t + \boldsymbol{\varphi}_{R,t}^T\, \boldsymbol{R}_t + \varphi_t = \Big[(1+\beta^W\,\varphi_k^W)\kappa\Big] k_t$$
$$+\Big[\beta^W\, \boldsymbol{\varphi}_{\tau W}^T\, \boldsymbol{\sigma}^A - (1+\beta^W\,\varphi_k^W)\xi_0^W\, \boldsymbol{e}_1^T\Big]\boldsymbol{\tau}_t + \Big[\beta^W\, \boldsymbol{\varphi}_{MW}^T\, \boldsymbol{\Phi}$$
$$+\Big(\beta^W\, \varphi_{\tau 1}^{AW}\, \sigma_{\text{forc}}(f_1 + f_2\, z_A - f_3\, z_A^{1-n}) + \beta^W\, \varphi_{\tau 1}^{BW}\, \sigma_{\text{forc}}(f_1 + f_2\, z_B - f_3\, z_B^{1-n})$$
$$-(1+\beta^W\,\varphi_k^W)\left(d_{AW}\,\frac{\alpha_A}{1-\alpha_A\,\alpha_B}(z_A - \alpha_B\, z_B) + d_{BW}\,\frac{1}{1-\alpha_A\,\alpha_B}(z_B - \alpha_A\, z_A)\right.$$
$$\left.+a^W\right)\Big) M_{pre}^{-1}\, \boldsymbol{e}_1^T\Big]\boldsymbol{M}_t + \Big[\beta^W\, \boldsymbol{\varphi}_{R,t+1}^T\Big] R_t + \log\, x_t^* + \beta^W\, \varphi_k^W\, \log(1-x_t^*)$$
$$+(1+\beta^W\,\varphi_k^W)\log F(\boldsymbol{A}_t,\, \boldsymbol{\mathcal{K}}_t^*,\, \boldsymbol{N}_t^*,\, \boldsymbol{E}_t^*) + (1+\beta^W\,\varphi_k^W)(\xi_0^W + a^W)$$
$$+f_0\,\beta^W(\varphi_{\tau 1}^{AW}\,\sigma_{\text{forc}} + \varphi_{\tau 1}^{BW}\,\sigma_{\text{forc}}) + \beta^W\, \boldsymbol{\varphi}_{MW}^T\, \tilde{\boldsymbol{e}}_{\boldsymbol{t}} - \beta^W\, \boldsymbol{\varphi}_{R,t+1}^T\, \boldsymbol{E}_t^{d*} + \beta^W\, \varphi_{t+1}.$$

with $z_A \in \{z_A^c, z_A^g\}$, $z_B \in \{z_B^c, z_B^g\}$, $d_{AA} \in \{d_{AA}^c, d_{AA}^g\}$, and $d_{BB} \in \{d_{BB}^c, d_{BB}^g\}$ depending on whether region A and B engage in counter-geoengineering or geoengineering.

(ii): $S_t^A > 0, S_t^B = 0$

$$\varphi_k^W\, k_t + \boldsymbol{\varphi}_{\tau W}^T\, \boldsymbol{\tau}_t + \boldsymbol{\varphi}_{MW}^T\, \boldsymbol{M}_t + \boldsymbol{\varphi}_{R,t}^T\, \boldsymbol{R}_t + \varphi_t = \Big[(1+\beta^W\,\varphi_k^W)\kappa\Big] k_t$$
$$+\Big[\beta^W\, \boldsymbol{\varphi}_{\tau W}^T\, \boldsymbol{\sigma}^A - (1+\beta^W\,\varphi_k^W)\xi_0^W\, \boldsymbol{e}_1^T\Big]\boldsymbol{\tau}_t + \Big[\beta^W\, \boldsymbol{\varphi}_{MW}^T\, \boldsymbol{\Phi}$$
$$+\Big(\beta^W \sigma_{\text{forc}}\, \varphi_{\tau 1}^{AW}(f_1 + f_2\, z_A^g - f_3\,(z_A^g)^{1-n}) + \beta^W\, \sigma_{\text{forc}}\, \varphi_{\tau 1}^{BW}(f_1 + f_2\,\alpha_A\, z_A^g - f_3\,(\alpha_A\, z_A^g)^{1-n})$$
$$-(1+\beta^W\,\varphi_k^W)\,(d_{AW}\,\alpha_A\, z_A^g + a^W)\Big) M_{pre}^{-1}\, \boldsymbol{e}_1^T\Big]\boldsymbol{M}_t + \Big[\beta^W\, \boldsymbol{\varphi}_{R,t+1}^T\Big] R_t + \log\, x_t^*$$
$$+\beta^W\, \varphi_k^W\, \log(1-x_t^*) + (1+\beta^W\,\varphi_k^W)\log F(\boldsymbol{A}_t,\, \boldsymbol{\mathcal{K}}_t^*,\, \boldsymbol{N}_t^*,\, \boldsymbol{E}_t^*) + (1+\beta^W\,\varphi_k^W)(\xi_0^W + a^W)$$
$$+f_0\,\beta^W(\varphi_{\tau 1}^{AW}\,\sigma_{\text{forc}} + \varphi_{\tau 1}^{BW}\,\sigma_{\text{forc}}) + \beta^W\, \boldsymbol{\varphi}_{MW}^T\, \tilde{\boldsymbol{e}}_{\boldsymbol{t}} - \beta^W\, \boldsymbol{\varphi}_{R,t+1}^T\, \boldsymbol{E}_t^{d*} + \beta^W\, \varphi_{t+1}.$$

(iii): $S_t^A = 0, S_t^B > 0$

$$\varphi_k^W\, k_t + \boldsymbol{\varphi}_{\tau W}^T\, \boldsymbol{\tau}_t + \boldsymbol{\varphi}_{MW}^T\, \boldsymbol{M}_t + \boldsymbol{\varphi}_{R,t}^T\, \boldsymbol{R}_t + \varphi_t = \Big[(1+\beta^W\,\varphi_k^W)\kappa\Big] k_t$$
$$+\Big[\beta^W\,\boldsymbol{\varphi}_{\tau W}^T\,\boldsymbol{\sigma}^A - (1+\beta^W\,\varphi_k^W)\xi_0^W\,\boldsymbol{e}_1^T\Big]\boldsymbol{\tau}_t + \Big[\beta^W\,\boldsymbol{\varphi}_{MW}^T\,\boldsymbol{\Phi}$$
$$+\Big(\beta^W\,\sigma_{\text{forc}}\,\varphi_{\tau 1}^{AW}(f_1 + f_2\,z_B^g - f_3\,(\alpha_B\,z_B^g)^{1-n}) + \beta^W\,\sigma_{\text{forc}}\,\varphi_{\tau 1}^{BW}(f_1 + f_2\,\alpha_A\,z_B^g - f_3\,(z_B^g)^{1-n})$$
$$-(1+\beta^W\,\varphi_k^W)\,(d_{BW}\,z_B^g + a^W)\Big) M_{pre}^{-1}\,\boldsymbol{e}_1^T\Big]\boldsymbol{M}_t + \Big[\beta^W\,\boldsymbol{\varphi}_{R,t+1}^T\Big] R_t + \log\, x_t^*$$
$$+\beta^W\,\varphi_k^W\,\log(1-x_t^*) + (1+\beta^W\,\varphi_k^W)\log F(\boldsymbol{A}_t,\,\boldsymbol{\mathcal{K}}_t^*,\,\boldsymbol{N}_t^*,\,\boldsymbol{E}_t^*) + (1+\beta^W\,\varphi_k^W)(\xi_0^W + a^W)$$
$$+f_0\,\beta^W(\varphi_{\tau 1}^{AW}\,\sigma_{\text{forc}} + \varphi_{\tau 1}^{BW}\,\sigma_{\text{forc}}) + \beta^W\,\boldsymbol{\varphi}_{MW}^T\,\tilde{\boldsymbol{e}}_{\boldsymbol{t}} - \beta^W\,\boldsymbol{\varphi}_{R,t+1}^T\,\boldsymbol{E}_t^{d*} + \beta^W\,\varphi_{t+1}.$$

Hence, in all Nash equilibria the system is linear in states.

Shadow values of the states. Coefficient matching with respect to capital k_t yields

$$\varphi_k^W = (1+\beta^W\,\varphi_k^W)\kappa \quad \Leftrightarrow \quad \varphi_k^W = \frac{\kappa}{1-\beta^W\,\kappa}$$

Structurally similar to the global model we get the consumption rate $x_t^* = 1 - \beta^W\,\kappa$.

Coefficient matching with respect to transformed temperatures delivers

$$\boldsymbol{\varphi}_{\tau W}^T = -\xi_0^W\,(1+\beta^W\,\varphi_k^W)\,\boldsymbol{e}_1^T\,(\mathbf{1} - \beta^W\,\boldsymbol{\sigma}^A)^{-1}.$$

Since the rest of the world is part of climate zone B, we define

$$\tilde{\sigma}_{ij}^W = \big[(\mathbf{1} - \beta^W\,\boldsymbol{\sigma}^{\boldsymbol{B}})^{-1}\big]_{ij} \text{ for } i,j \in \{1,2\}.$$

Note that to preserve symmetry in notation, we defined the corresponding matrix $\boldsymbol{\sigma}^B$ by swapping the first and second rows and columns of the matrix $\boldsymbol{\sigma}^{\boldsymbol{A}}$.

The rest of the world's shadow values of atmospheric temperature in regions A and B are therefore

$$\varphi_{\tau 1}^{AW} = -\xi_0^W\,(1+\beta^W\,\varphi_k^W)\tilde{\sigma}_{12}^W, \quad \text{and} \quad \varphi_{\tau 1}^{BW} = -\xi_0^W\,(1+\beta^W\,\varphi_k^W)\tilde{\sigma}_{11}^W.$$

Further, we define $\gamma_W^{heat} \equiv \beta^W\,\xi_0^W\,\tilde{\sigma}_{12}^W\,\sigma_{\text{forc}}$ and $\gamma_W \equiv \beta^W\,\xi_0^W\,\tilde{\sigma}_{11}^W\,\sigma_{\text{forc}}$.

Coefficient matching with respect to carbon stocks and using γ_W, and γ_W^{heat} yields

(i): $S_t^A \neq 0, S_t^B \neq 0$

$$\varphi_{M1}^W = (1+\beta^W \varphi_k^W)\Bigg(- f_2(\gamma_W^{heat} z_A + \gamma_W z_B) - d_{AW} \frac{\alpha_A}{1-\alpha_A \alpha_B}(z_A - \alpha_B z_B) - d_{BW} \frac{1}{1-\alpha_A \alpha_B}(z_B - \alpha_A z_A) + f_3 \left(\gamma_W^{heat} z_A^{1-n} + \gamma_W z_B^{1-n}\right) - a^W - f_1(\gamma_W^{heat} + \gamma_W)\Bigg) M_{pre}^{-1} \boldsymbol{e}_1^T (\mathbf{1} - \beta^W \boldsymbol{\Phi})^{-1} \quad (3.47)$$

(ii): $S_t^A > 0, S_t^B = 0$

$$\varphi_{M1}^W = (1+\beta^W \varphi_k^W)\Bigg(- f_2(\gamma_W^{heat} z_A^g + \gamma_W \alpha_A z_A^g) - d_{AW} \alpha_A z_A^g + f_3\Big(\gamma_W^{heat} (z_A^g)^{1-n} + \gamma_W (\alpha_A z_A^g)^{1-n}\Big) - a^W - f_1(\gamma_W^{heat} + \gamma_W)\Bigg) M_{pre}^{-1} \boldsymbol{e}_1^T (\mathbf{1} - \beta^W \boldsymbol{\Phi})^{-1} \quad (3.48)$$

(iii): $S_t^A = 0, S_t^B > 0$

$$\varphi_{M1}^W = (1+\beta^W \varphi_k^W)\Bigg(- f_2(\gamma_W^{heat} \alpha_B z_B^g + \gamma_W z_B^g) - d_{BW} z_B^g + f_3\Big(\gamma_W^{heat} (\alpha_B z_B^g)^{1-n} + \gamma_W (z_B^g)^{1-n}\Big) - a^W - f_1(\gamma_W^{heat} + \gamma_W)\Bigg) M_{pre}^{-1} \boldsymbol{e}_1^T (\mathbf{1} - \beta^W \boldsymbol{\Phi})^{-1}. \quad (3.49)$$

We define

$$\tilde{\phi}_{ij}^W = \left[(\mathbf{1} - \beta^W \boldsymbol{\Phi})^{-1}\right]_{ij} \text{ for } i,j \in \{1,2,3\},$$

yielding

$$\tilde{\phi}_{11}^W = \left[(\mathbf{1} - \beta^W \boldsymbol{\Phi})^{-1}\right]_{1,1}. \quad (3.50)$$

From coefficient matching with respect to the resource stock we have

$$\boldsymbol{\varphi}_{R,t}^T = \beta^W \boldsymbol{\varphi}_{R,t+1}^T \quad \Leftrightarrow \quad \boldsymbol{\varphi}_{R,t} = (\beta^W)^{-t} \boldsymbol{\varphi}_{R,0} \quad \text{(Hotelling's rule)}.$$

The initial resource values $\boldsymbol{\varphi}_{R,0}^T$ depend on the set up of the economy, including assump-

tions about production and the energy sector. Given the coefficients and the optimal rate of consumption equations (3.37), (3.38), and (3.39) turn to the following condition:

$$\varphi_t - \beta^W \varphi_{t+1} = \log x_t^* + \beta^W \varphi_k^W \log(1 - x_t^*) + (1 + \beta^W \varphi_k^W) \log F(\boldsymbol{A}_t, \boldsymbol{\mathcal{K}}_t^*, \boldsymbol{N}_t^*, \boldsymbol{E}_t^*)$$
$$+(1+\beta^W \varphi_k^W)(\xi_0^W + a^W) + \beta^W (\varphi_{\tau 1}^{AW} \sigma_{\text{forc}} + \varphi_{\tau 1}^{BW} \sigma_{\text{forc}}) f_0 + \beta^W \boldsymbol{\varphi}_{MW}^T \tilde{\boldsymbol{e}}_{\boldsymbol{t}} - \beta^W \boldsymbol{\varphi}_{R,t+1}^T \boldsymbol{E}_t^{d*} \tag{3.51}$$

This condition will be satisfied by picking the sequence $\varphi_0, \varphi_1, \varphi_2, ...$ The additional condition $\lim_{t\to\infty}(\beta^W)^t V(\cdot) = 0 \Rightarrow \lim_{t\to\infty}(\beta^W)^t \varphi_t = 0$ pins down this initial value φ_0.

B.1.3.6 Proofs of Propositions 3.3 & 3.7

The proofs make use of section B.1.3.2, where we derive the Markov strategies, section B.1.3.3, where we show that the strategies are consistent with the assumed linear form of the trial solution for the Bellman equation, and section B.1.3.4, where we derive the solutions for the shadow values.

General model (Proposition 3.7). In section B.1.3.2 we show that the following reaction functions characterize a Nash equilibrium of the dynamic game: If (i) $S_t^B = 0$, region A chooses $S_t^A = z_A^g m_t$ and if (ii) $S_t^B \neq 0$ region A chooses

$$S_t^A = \frac{m_t}{1 - \alpha_A \alpha_B} \left(z_A^g - \alpha_B z_B \right) \quad \text{for} \quad S_t^A > 0$$
$$S_t^A = \frac{m_t}{1 - \alpha_A \alpha_B} \left(z_A^c - \alpha_B z_B \right) \quad \text{for} \quad S_t^A < 0$$
$$S_t^A = 0 \quad \text{otherwise.}$$

If (i) $S_t^B = 0$, then

$$z_A^g = \left(\frac{\tilde{a}_A + \tilde{b}_A \alpha_A^{-n}}{\tilde{\delta}_A^g} \right)^{\frac{1}{n}},$$

and if (ii) $S_t^B \neq 0$, then

$$z_A^g = \left(\frac{\tilde{a}_A - \tilde{b}_A \frac{\tilde{b}_B}{\tilde{a}_B}}{\tilde{\delta}_A^g - \tilde{b}_A \frac{\tilde{\delta}_B}{\tilde{a}_B}} \right)^{\frac{1}{n}}, \quad z_A^c = \left(\frac{\tilde{a}_A - \tilde{b}_A \frac{\tilde{b}_B}{\tilde{a}_B}}{\tilde{\delta}_A^c - \tilde{b}_A \frac{\tilde{\delta}_B}{\tilde{a}_B}} \right)^{\frac{1}{n}},$$

where

$$-\tilde{\delta}_A^g = \beta^A \varphi_{\tau 1}^{AA} \sigma_{\text{forc}} f_2 + \beta^A \varphi_{\tau 1}^{BA} \sigma_{\text{forc}} f_2 \alpha_A - (1 + \beta^A \varphi_k^A) d_{AA}^g$$

$$-\tilde{\delta}_A^c = \beta^A \varphi_{\tau 1}^{AA} \sigma_{\text{forc}} f_2 + \beta^A \varphi_{\tau 1}^{BA} \sigma_{\text{forc}} f_2 \alpha_A - (1 + \beta^A \varphi_k^A) d_{AA}^c$$

$$\tilde{a}_A = (n-1)\, \beta^A \varphi_{\tau 1}^{AA} \sigma_{\text{forc}} f_3$$

$$\tilde{b}_A = (n-1) \beta^A \varphi_{\tau 1}^{BA} \sigma_{\text{forc}}\ f_3 \alpha_A.$$

Swapping region indices characterizes region B's strategies.

In section B.1.3.4 we show that region A's shadow values of atmospheric temperatures in regions A and B are

$$\varphi_{\tau 1}^{AA} = -\xi_0^A (1 + \beta^A \varphi_k^A) \tilde{\sigma}_{11}^A, \quad \text{and} \quad \varphi_{\tau 1}^{BA} = -\xi_0^A (1 + \beta^A \varphi_k^A) \tilde{\sigma}_{12}^A.$$

The temperature shadow values for region B follow by switching region indices.

Defining $\gamma_A \equiv \beta^A \xi_0^A \tilde{\sigma}_{11}^A \sigma_{\text{forc}}$, and $\gamma_A^{heat} \equiv \beta^A \xi_0^A \tilde{\sigma}_{12}^A \sigma_{\text{forc}}$, gives us δ_A, a_A, and b_A as a function of γ_A, and γ_A^{heat}:

$$\begin{aligned}
a_A &\equiv \frac{\tilde{a}_A}{(1 + \beta^A \varphi_k^A)} = (1-n)\gamma_A f_3 \\
b_A &\equiv \frac{\tilde{b}_A}{(1 + \beta^A \varphi_k^A)} = (1-n)\alpha_A \gamma_A^{heat} f_3 \\
\delta_A^g &\equiv \frac{\tilde{\delta}_A^g}{(1 + \beta^A \varphi_k^A)} = (d_{AA}^g + \epsilon_A^g) + \gamma_A f_2 + \alpha_A \gamma_A^{heat} f_2 \\
\delta_A^c &\equiv \frac{\tilde{\delta}_A^c}{(1 + \beta^A \varphi_k^A)} = (d_{AA}^c - \epsilon_A^c) + \gamma_A f_2 + \alpha_A \gamma_A^{heat} f_2.
\end{aligned}$$

The solutions for δ_B, a_B, and b_B follow by switching region indices. Thus, if (i) $S_t^B = 0$

$$z_A^g = \left(\frac{a_A + b_A \alpha_A^{-n}}{\delta_A^g} \right)^{\frac{1}{n}},$$

and if (ii) $S_t^B \neq 0$, then

$$z_A^g = \left(\frac{a_A - b_A \frac{b_B}{a_B}}{\delta_A^g - b_A \frac{\delta_B}{a_B}} \right)^{\frac{1}{n}}, \quad z_A^c = \left(\frac{a_A - b_A \frac{b_B}{a_B}}{\delta_A^c - b_A \frac{\delta_B}{a_B}} \right)^{\frac{1}{n}}. \tag{3.52}$$

Swapping region indices characterizes region B's propensity and reluctance.

Base model (Proposition 3.3). No heat flows, i.e. setting $\sigma_B^A = \sigma_A^B = \sigma_A^O = \sigma_B^O = 0$, simplifies the shadow values of temperature such that $\varphi_{\tau 1}^{AA} = -\xi_0^A \left(1 + \beta^A \varphi_k^A\right) \tilde{\sigma}_{11}^A$ and $\varphi_{\tau 1}^{BA} = 0$ since $\tilde{\sigma}_{12}^A = 0$. For the base model, we use a slightly simpler notation and define $\varphi_{\tau 1}^A \equiv \varphi_{\tau 1}^{AA}$.

No heat flows imply that $b_A = b_B = 0$, and therefore equations (3.52) turn to

$$z_A^g = \left(\frac{\delta_A^g}{a_A}\right)^{-\frac{1}{n}} = \left(\frac{(n-1)\beta^A \varphi_{\tau 1}^A \sigma_{\text{forc}} f_3}{(1+\beta^A \varphi_k) d_{AA}^g - \beta^A \varphi_{\tau 1}^A \sigma_{\text{forc}} f_2}\right)^{\frac{1}{n}}$$

$$z_A^c = \left(\frac{\delta_A^c}{a_A}\right)^{-\frac{1}{n}} = \left(\frac{(n-1)\beta^A \varphi_{\tau 1}^A \sigma_{\text{forc}} f_3}{(1+\beta^A \varphi_k) d_{AA}^c - \beta^A \varphi_{\tau 1}^A \sigma_{\text{forc}} f_2}\right)^{\frac{1}{n}}$$

Switching region indices gives us the analogous result for region B. Using $\gamma_A = \beta^A \xi_0^A \tilde{\sigma}_{11}^A \sigma_{\text{forc}}$, and noting that $\gamma_A^{heat} = 0$ leads to

$$z_A^g = \left(\frac{(1-n) f_3 \gamma_A}{f_2 \gamma_A + (d_{AA}^g + \epsilon_A^g)}\right)^{\frac{1}{n}} \quad \text{and} \quad z_A^c = \left(\frac{(1-n) f_3 \gamma_A}{f_2 \gamma_A + (d_{AA}^c - \epsilon_A^c)}\right)^{\frac{1}{n}}. \tag{3.53}$$

B.1.3.7 Proof of Proposition 3.4

To derive the Nash equilibria we use the reaction functions for region A and B from section B.1.3.2. We note that Assumption 3.1 excludes the case that both regions engage in counter-engineering.

i.a) In the case where both regions are cooling ($S_t^A > 0, S_t^B > 0$) we obtain

$$S_t^A(m_t) = \frac{m_t}{1 - \alpha_A \alpha_B}\left(z_A^g - \alpha_B z_B^g\right) > 0 \quad \Rightarrow \quad z_A^g > \alpha_B z_B^g$$

and

$$S_t^B(m_t) = \frac{m_t}{1 - \alpha_A \alpha_B}\left(z_B^g - \alpha_A z_A^g\right) > 0 \quad \Rightarrow \quad z_B^g > \alpha_A z_A^g. \tag{3.54}$$

Together, the two equations imply

$$\alpha_B < \underbrace{\frac{z_A^g}{z_B^g}}_{\equiv H} < \alpha_A^{-1}. \tag{3.55}$$

Note that this condition defines a non-empty range of parameter values unless $\alpha_A^n = \alpha_B^n =$

1. Thus, condition (3.55) states the range of parameter values α_B, α_A, d^g_{AA}, d^g_{BB}, ϵ^g_A, ϵ^g_B, and f_2 for which there exists an equilibrium in which both regions are cooling the world given γ_A, γ^{heat}_A, γ_B and γ^{heat}_B.

i.b) In the case where region A is cooling ($S^A_t > 0$) and region B is warming ($S^B_t < 0$) we obtain

$$S^A_t(m_t) = \frac{m_t}{1-\alpha_A\,\alpha_B}\left(z^g_A - \alpha_B z^c_B\right) > 0 \quad \Leftrightarrow \quad z^g_A > \alpha_B z^c_B$$

$$S^B_t(m_t) = \frac{m_t}{1-\alpha_A\,\alpha_B}\left(z^c_B - \alpha_A z^g_A\right) < 0 \quad \Leftrightarrow \quad z^c_B < \alpha_A z^g_A. \tag{3.56}$$

Therefore, the parameter range in which this case defines the Nash equilibrium is characterized by

$$\frac{z^g_A}{z^c_B} > \max\{\alpha_B, \alpha_A^{-1}\} \quad \Leftrightarrow \quad h \equiv \frac{z^g_A}{z^c_B} > \alpha_A^{-1} = \max\{\alpha_B, \alpha_A^{-1}\}.$$

i.c) The case where region A is warming ($S^A_t < 0$) and region B is cooling ($S^B_t > 0$) follows by symmetry (switching the region indices)

$$\frac{z^g_B}{z^c_A} > \alpha_B^{-1} = \max\{\alpha_A, \alpha_B^{-1}\} \quad \Rightarrow \quad \hat{H} \equiv \frac{z^c_A}{z^g_B} < \alpha_B.$$

ii) In the case where region A is cooling ($S^A_t > 0$) and region B is not acting ($S^B_t = 0$) it has to be optimal for region B to neither engage in cooling, nor in counter-engineering. Given region A is taking the same actions as in scenarios i and ii, region B's reaction function can neither satisfy equation (3.54) nor (3.56). Therefore, it must be that $H \geq \alpha_A^{-1}$ and $h \leq \alpha_A^{-1}$. In addition, region A's reaction function becomes

$$S^A_t(m_t) = z^g_A\, m_t > 0,$$

which will always be satisfied. The reaction function of region B is obviously $S^B_t = 0$.

iii) Finally, the symmetric reasoning for region B cooling ($S^B_t > 0$) and region A not

acting ($S_t^A = 0$) delivers $\frac{1}{H} \geq \alpha_B^{-1}$ and $\frac{1}{\hat{H}} \leq \alpha_B^{-1}$ or

$$\frac{1}{\hat{H}} \leq \alpha_B^{-1} \leq \frac{1}{H} \quad \Leftrightarrow \quad \hat{H} \geq \alpha_B \geq H.$$

Here the reaction functions are

$$S_t^B(m_t) = z_B^g \, m_t > 0$$

and locally $S_t^A = 0$. These 5 cases are mutually exclusive and cover the full parameter domain.

B.1.3.8 Proofs of Propositions 3.5 & 3.8

The proofs make use of the solutions for the shadow values of the atmospheric carbon stock from section B.1.3.4.

General model (Proposition 3.8). Inserting φ_k^A into (3.40), (3.41), and (3.42) delivers:

(i): $S_t^A \neq 0$ and $S_t^B \neq 0$

$$\varphi_{M1}^A = \frac{1}{1-\beta^A \kappa} \Bigg(- f_2 \left(\gamma_A \, z_A + \gamma_A^{heat} \, z_B\right) - \frac{d_{AA} \left(z_A - \alpha_B \, z_B\right) + \alpha_B \, d_{BA} (z_B - \alpha_A \, z_A)}{1 - \alpha_A \, \alpha_B} + f_3 \left(\gamma_A \, z_A^{1-n} + \gamma_A^{heat} \, z_B^{1-n}\right) - a^A - f_1 \left(\gamma_A + \gamma_A^{heat}\right) \Bigg) M_{pre}^{-1} \, \tilde{\phi}_{11}^A,$$

(ii): $S_t^A > 0$ and $S_t^B = 0$

$$\varphi_{M1}^A = \frac{1}{1-\beta^A \kappa} \Bigg(- f_2 \left(\gamma_A \, z_A^g + \gamma_A^{heat} \, \alpha_A \, z_A^g\right) - d_{AA} \, z_A^g + f_3 \Big(\gamma_A \left(z_A^g\right)^{1-n} + \gamma_A^{heat} \left(\alpha_A \, z_A^g\right)^{1-n}\Big) - a^A - f_1 \left(\gamma_A + \gamma_A^{heat}\right) \Bigg) M_{pre}^{-1} \, \tilde{\phi}_{11}^A,$$

(iii): $S_t^A = 0$ and $S_t^B > 0$

$$\varphi_{M1}^A = \frac{1}{1-\beta^A\,\kappa}\Bigg(- f_2\,(\gamma_A\,\alpha_B\,z_B^g + \gamma_A^{heat}\,z_B^g) - d_{BA}\,\alpha_B\,z_B^g + f_3\Big(\gamma_A\,(\alpha_B\,z_B^g)^{1-n} + \gamma_A^{heat}\,(z_B^g)^{1-n}\Big) - a^A - f_1\,(\gamma_A + \gamma_A^{heat})\Bigg) M_{pre}^{-1}\,\tilde{\phi}_{11}^A.$$

The regional SCC is the negative of the regional shadow value of atmospheric carbon expressed in money-measured consumption units. Thus,

(i): $S_t^A \neq 0$ and $S_t^B \neq 0$

$$\begin{aligned} SCC^A &= -(1-\beta^A\,\kappa)Y_{A,t}^{net}\,\varphi_{M1}^A \\ &= \frac{Y_{A,t}^{net}}{M_{pre}}\Bigg[a^A + f_1\,(\gamma_A + \gamma_A^{heat}) - \Bigg(\Big(\frac{f_3}{z_A^n} - f_2\Big)\gamma_A - d_{AA}\Bigg)z_A - \Big(\frac{f_3}{z_B^n} - f_2\Big)\gamma_A^{heat}\,z_B \\ &\quad - \frac{\alpha_B(z_B - \alpha_A\,z_A)(d_{AA} - d_{BA})}{1-\alpha_A\,\alpha_B}\Bigg]\tilde{\phi}_{11}^A, \end{aligned} \tag{3.57}$$

with $z_A \in \{z_A^g, z_A^c\}$, $z_B \in \{z_B^g, z_B^c\}$, $d_{AA} \in \{d_{AA}^g + \epsilon_A^g, d_{AA}^c - \epsilon_A^c\}$, and $d_{BA} \in \{d_{BA}^g, d_{BA}^c\}$ depending on whether the corresponding region engages in geoengineering (g) or counter-geoengineering (c).

(ii): $S_t^A > 0$ and $S_t^B = 0$

$$\begin{aligned} SCC^A &= -(1-\beta^A\,\kappa)Y_{A,t}^{net}\,\varphi_{M1}^A \\ &= \frac{Y_{A,t}^{net}}{M_{pre}}\Bigg[a^A + f_1\,(\gamma_A + \gamma_A^{heat}) - \Bigg(\Big(\frac{f_3}{(z_A^g)^n} - f_2\Big)\gamma_A - (d_{AA}^g + \epsilon_A^g)\Bigg)z_A^g \\ &\quad - \Big(\frac{f_3}{(\alpha_A\,z_A^g)^n} - f_2\Big)\gamma_A^{heat}\,\alpha_A\,z_A^g\Bigg]\tilde{\phi}_{11}^A, \end{aligned} \tag{3.58}$$

(iii): $S_t^A = 0$ and $S_t^B > 0$

$$\begin{aligned} SCC^A &= -(1-\beta^A\,\kappa)Y_{A,t}^{net}\,\varphi_{M1}^A \\ &= \frac{Y_{A,t}^{net}}{M_{pre}}\Bigg[a^A + f_1\,(\gamma_A + \gamma_A^{heat}) - \Bigg(\Big(\frac{f_3}{(\alpha_B\,z_B^g)^n} - f_2\Big)\gamma_A - d_{BA}^g\Bigg)\alpha_B\,z_B^g \\ &\quad - \Big(\frac{f_3}{(z_B^g)^n} - f_2\Big)\gamma_A^{heat}\,z_B^g\Bigg]\tilde{\phi}_{11}^A. \end{aligned} \tag{3.59}$$

Summarizing all terms in the SCC that do not depend on γ_A^{heat} in the term $SCC_{w/o}^A$ leads

to Proposition 3.8. Swapping region indices characterizes region B's SCC.

Base model (Proposition 3.5). In the special case, where σ_B^A, σ_A^B, σ_A^O, σ_B^O are equal to zero and thus also $\gamma_A^{heat} = 0$, the general equations (3.57), (3.58), and (3.59) for the regional SCC turn to

(i): $S_t^A \neq 0$ and $S_t^B \neq 0$

$$SCC^A = \frac{Y_{A,t}^{net}}{M_{pre}} \left[a^A + f_1\,\gamma_A - \left(\left(\frac{f_3}{z_A^n} - f_2 \right) \gamma_A - d_{AA} \right) z_A - \frac{\alpha_B (z_B - \alpha_A\, z_A)(d_{AA} - d_{BA})}{1 - \alpha_A\,\alpha_B} \right] \tilde{\phi}_{11}^A,$$

with $z_A \in \{z_A^c, z_A^g\}$, $z_B \in \{z_B^c, z_B^g\}$, $d_{AA} \in \{d_{AA}^c + \epsilon^c, d_{AA}^g + \epsilon^g\}$, and $d_{BB} \in \{d_{BB}^c + \epsilon^c, d_{BB}^g + \epsilon^g\}$.

(ii): $S_t^A > 0$ and $S_t^B = 0$

$$SCC^A = \frac{Y_{A,t}^{net}}{M_{pre}} \left[a^A + f_1\,\gamma_A - \left(\left(\frac{f_3}{(z_A^g)^n} - f_2 \right) \gamma_A - d_{AA} \right) z_A^g \right] \tilde{\phi}_{11}^A,$$

(iii): $S_t^A = 0$ and $S_t^B > 0$

$$SCC^A = \frac{Y_{A,t}^{net}}{M_{pre}} \left[a^A + f_1\,\gamma_A - \left(\left(\frac{f_3}{(\alpha_B\, z_B^g)^n} - f_2 \right) \gamma_A - d_{BA} \right) \alpha_B\, z_B^g \right] \tilde{\phi}_{11}^A.$$

B.1.3.9 Proof of Proposition 3.6

The proof makes use of the solutions for the shadow values of the atmospheric carbon stock from section B.1.3.5.

General model. Inserting φ_k^W into (3.47), (3.48), and (3.49) delivers

(i): $S_t^A \neq 0, S_t^B \neq 0$

$$\varphi_{M1}^W = \frac{1}{1 - \beta^W\,\kappa} \Bigg(- f_2(\gamma_W^{heat}\, z_A + \gamma_W\, z_B) - d_{AW}\, \frac{\alpha_A}{1 - \alpha_A\,\alpha_B} (z_A - \alpha_B\, z_B)$$
$$- d_{BW}\, \frac{1}{1 - \alpha_A\,\alpha_B} (z_B - \alpha_A\, z_A) + f_3 \left(\gamma_W^{heat}\, z_A^{1-n} + \gamma_W\, z_B^{1-n} \right) - a^W - f_1(\gamma_W^{heat} + \gamma_W) \Bigg) M_{pre}^{-1}\, \tilde{\phi}_{11}^W,$$

(ii): $S_t^A > 0, S_t^B = 0$

$$\varphi_{M1}^W = \frac{1}{1-\beta^W \kappa}\Bigg(- f_2(\gamma_W^{heat} z_A^g + \gamma_W \alpha_A z_A^g) - d_{AW} \alpha_A z_A^g + f_3\Big(\gamma_W^{heat} (z_A^g)^{1-n} + \gamma_W (\alpha_A z_A^g)^{1-n}\Big) - a^W - f_1(\gamma_W^{heat} + \gamma_W)\Bigg) M_{pre}^{-1} \tilde{\phi}_{11}^W,$$

(iii): $S_t^A = 0, S_t^B > 0$

$$\varphi_{M1}^W = \frac{1}{1-\beta^W \kappa}\Bigg(- f_2(\gamma_W^{heat} \alpha_B z_B^g + \gamma_W \ z_B^g) - d_{BW} z_B^g + f_3\Big(\gamma_W^{heat} (\alpha_B z_B^g)^{1-n} + \gamma_W (z_B^g)^{1-n}\Big) - a^W - f_1(\gamma_W^{heat} + \gamma_W)\Bigg) M_{pre}^{-1} \tilde{\phi}_{11}^W.$$

The regional SCC is the negative of the regional shadow value of atmospheric carbon expressed in money-measured consumption units. Thus,

(i): $S_t^A \neq 0$ and $S_t^B \neq 0$

$$\begin{aligned} SCC^W &= -(1-\beta^W \kappa) Y_{W,t}^{net} \varphi_{M1}^W \\ &= \frac{Y_{W,t}^{net}}{M_{pre}}\Bigg[a^W + f_1(\gamma_W^{heat} + \gamma_W) - \Big(\frac{f_3}{z_A^n} - f_2\Big)\gamma_W^{heat} z_A - \Big(\Big(\frac{f_3}{z_B^n} - f_2\Big)\gamma_W - d_{BW}\Big) z_B \\ &\quad - \frac{\alpha_A (z_A - \alpha_B z_B)(d_{BW} - d_{AW})}{1 - \alpha_A \alpha_B}\Bigg] \tilde{\phi}_{11}^W, \end{aligned} \tag{3.60}$$

with $z_A \in \{z_A^c, z_A^g\}$, $z_B \in \{z_B^c, z_B^g\}$, $d_{AW} \in \{d_{AW}^c, d_{AW}^g\}$, and $d_{BW} \in \{d_{BW}^c, d_{BW}^g\}$.

(ii): $S_t^A > 0$ and $S_t^B = 0$

$$\begin{aligned} SCC^W &= -(1-\beta^W \kappa) Y_{W,t}^{net} \varphi_{M1}^W \\ &= \frac{Y_{W,t}^{net}}{M_{pre}}\Bigg[a^W + f_1(\gamma_W^{heat} + \gamma_W) - \Big(\Big(\frac{f_3}{(z_A^g)^n} - f_2\Big)\gamma_W^{heat}\Big) z_A^g \\ &\quad - \Big(\Big(\frac{f_3}{(\alpha_A z_A^g)^n} - f_2\Big)\gamma_W - d_{AW}^g\Big)\alpha_A z_A^g\Bigg] \tilde{\phi}_{11}^W, \end{aligned} \tag{3.61}$$

(iii): $S_t^A = 0$ and $S_t^B > 0$

$$\begin{aligned} SCC^W &= -(1-\beta^W\,\kappa)Y_{W,t}^{net}\,\varphi_{M1}^W \\ &= \frac{Y_{W,t}^{net}}{M_{pre}}\Bigg[a^W + f_1\,(\gamma_W^{heat}+\gamma_W) - \Bigg(\Bigg(\frac{f_3}{(\alpha_B\,z_B^g)^n} - f_2\Bigg)\gamma_W^{heat}\Bigg)\alpha_B\,z_B^g \\ &\quad -\Bigg(\Bigg(\frac{f_3}{(z_B^g)^n} - f_2\Bigg)\gamma_W - d_{BW}^g\Bigg)z_B^g\Bigg]\,\tilde{\phi}_{11}^W. \end{aligned} \tag{3.62}$$

Base model (Proposition 3.6). In the special case, where σ_B^A, σ_A^B, σ_A^O, σ_B^O are equal to zero, $\gamma_W^{heat} = 0$. The general equations (3.60), (3.61), and (3.62) for the SCC in the rest of the world turn to

(i): $S_t^A \neq 0$ and $S_t^B \neq 0$

$$SCC^W = \frac{Y_{W,t}^{net}}{M_{pre}}\Bigg[a^W + f_1\,\gamma_W - \Bigg(\Bigg(\frac{f_3}{z_B^n} - f_2\Bigg)\gamma_W - d_{BW}\Bigg)z_B - \frac{\alpha_A(z_A - \alpha_B\,z_B)(d_{BW} - d_{AW})}{1-\alpha_A\,\alpha_B}\Bigg]\,\tilde{\phi}_{11}^W,$$

with $z_A \in \{z_A^c, z_A^g\}$, $z_B \in \{z_B^c, z_B^g\}$, $d_{AW} \in \{d_{AW}^c, d_{AW}^g\}$, and $d_{BW} \in \{d_{BW}^c, d_{BW}^g\}$.

(ii): $S_t^A > 0$ and $S_t^B = 0$

$$SCC^W = \frac{Y_{W,t}^{net}}{M_{pre}}\Bigg[a^W + f_1\,\gamma_W - \Bigg(\Bigg(\frac{f_3}{(\alpha_A\,z_A^g)^n} - f_2\Bigg)\gamma_W - d_{AW}\Bigg)\alpha_A\,z_A^g\Bigg]\,\tilde{\phi}_{11}^W,$$

(iii): $S_t^A = 0$ and $S_t^B > 0$

$$SCC^W = \frac{Y_{W,t}^{net}}{M_{pre}}\Bigg[a^W + f_1\,\gamma_W - \Bigg(\Bigg(\frac{f_3}{(z_B^g)^n} - f_2\Bigg)\gamma_W - d_{BW}\Bigg)z_B^g\Bigg]\,\tilde{\phi}_{11}^W.$$

4 Carbon dioxide removal in a global analytic climate economy

Felix D. Meier[1,2]

[1] *Kiel Institute for the World Economy*

[2] *German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig*

Abstract: This paper investigates the option of carbon dioxide removal (CDR) and storage in different reservoir types in an analytic climate-economy model, and derives implications for optimal mitigation efforts, and CDR deployment. I show that the introduction of CDR lowers net energy input and net emissions over the entire time path. Furthermore, CDR affects the Social Cost of Carbon (SCC) via changes in total economic output but leaves the analytic structure of the SCC unchanged. In the first years after CDR becomes available the SCC is lower and in later years it is higher compared to a standard climate-economy model. Carbon dioxide emissions are first higher and then lower relative to a world without CDR. The quantitative analysis shows that the effect of CDR on the SCC is minor and only accounts for an increase of 3 USD/tCO_2 by the year 2100.

Keywords: carbon dioxide removal, climate change, integrated assessment, social cost of carbon, optimal carbon tax

4.1 Introduction

In line with the Paris Agreement to limit global warming to well below 2°C many countries have declared their intention to transition towards a net-zero emissions economy by the second half of this century (Tanaka and O'Neill, 2018). To accomplish this goal technologies that remove carbon from the atmosphere or capture carbon directly from large stationary emission sources, such as power plants, have been proposed. Subsequently, both technologies require the storage of carbon in secure geological reservoirs or the ocean (Anderson and Newell, 2004). The process of capturing, transporting and storing carbon however consumes additional energy and thus potentially leads to new emissions (IPCC, 2005). Moreover, the availability of carbon dioxide removal and storage as an 'end of pipe' mitigation technology may be perceived as a substitute for conventional emission mitigation, which might lead to rebound effects (e.g. Geden et al., 2019).

I explore these trade-offs in an analytic integrated assessment model where Carbon Dioxide Removal (CDR) technologies can be used to reduce atmospheric carbon concentrations at a cost measured in energy units. I derive the optimal level of CDR deployment and analyze how emissions, energy input, and the SCC (optimal carbon tax) are affected by the introduction of CDR. Although the model focuses on CDR, it is general enough to also consider the potential of Carbon Capture and Storage (CCS) technologies.

Atmospheric carbon dioxide only represents a small fraction of the total carbon stock in the Earth System, namely 829 gigatons (Gt) out of a total of more than 45,696 gigatons (IPCC, 2013). The rest of it is bound in other reservoirs such as the oceans that have served as an important carbon sink over the past 200 years (Sabine et al., 2004). Due to the large storage capacity, the ocean has been suggested to serve as a reservoir for the intentional injection of carbon dioxide via ships or pipelines (Rickels and Lontzek, 2012).

Whether a geological reservoir, such as an exploited oil field, is well suited for CDR is mainly determined by the rate at which carbon leaks back to the atmosphere (van der Zwaan and Gerlagh, 2009). A similar problem arises, if carbon is stored in the ocean. Due to feedback and saturation effects in the carbon cycle some of the carbon that is injected into the oceans will eventually still end up in the atmosphere. Rickels et al. (2018) investigate how well these effects are captured in currently used Integrated Assessment Models (IAMs). Rickels and Lontzek (2012) explore the economic implications of the ocean's imperfect storage property. They show that optimally each ton of carbon sequestered to the ocean is taxed at a rate lower than the optimal carbon tax for atmospheric carbon emission. In this paper, I derive the SCC, which quantifies the optimal tax on carbon emissions, for different reservoir types and analyze how the optimal carbon tax is affected by the introduction of CDR technologies by comparing the results of model specifications with and without the availability of CDR.

The paper is based on the recently emerging literature on analytic IAMs which have the feature that the SCC can be written as a constant fraction of total economic output (e.g. Traeger, 2018; Gerlagh and Lsiki, 2018; Golosov et al., 2014). This result arises from specific assumptions on utility and climate change damages which ensure that the climate-economy model is linear in the model's state variables, in particular human-made capital and the stocks of carbon in the different reservoirs (Karp, 2017; Traeger, 2018). I show that due to the linear-in-states property of analytic IAMs the deployment of CDR technologies has no effect on the analytic structure of the SCC. However, CDR alters the time path of total economic output and therefore influences the level of the SCC. The quantitative analysis shows that this effect is minor and only increases the SCC by 3 USD/tCO_2 by the year 2100, where the absolute level of the SCC is around 800 USD/tCO_2.

The paper is structured as follows. The next section introduces the option of CDR in an analytic climate-economy model. Section 4.3 presents the theoretical

results on optimal emissions, CDR deployment, and the SCC, and compares them to the outcome of a standard climate-economy model without the option of CDR. The last section provides a numerical simulation for calibrated versions of both model types.

4.2 Analytic climate-economy model

This section introduces the option of CDR and the storage of carbon in different reservoirs types in an analytic integrated assessment model of climate change. The model is based on Traeger (2018) and Golosov et al. (2014). I consider a global economy where gross output Y_t is a function of technology $A_{0,t}$, capital K_t, labor $N_{0,t}$, and net energy input I_t,

$$Y_t = A_{0,t}\, K_t^{\kappa}\, N_{0,t}^{1-\kappa-v}\, I_t^{v} \qquad \text{with } K_0 > 0 \text{ given.} \tag{4.1}$$

The subscript zero denotes that technology and labor are prescribed by time dependent exogenous processes. Subscript t denotes the point in time. I distinguish between gross energy E_t and net energy I_t, whereby only the latter enters the gross output function.

Energy E_t is derived from an exhaustible resource R_t which comprises all fossil fuels (coal, oil, and natural gas), and is measured in terms of its carbon content (in GtC). Absent a CCS technology, E_t can also be interpreted as carbon emissions. The resource stock R_t develops over time according to

$$R_{t+1} = R_t - E_t, \qquad \text{with } R_0 > 0 \text{ given.} \tag{4.2}$$

I follow Traeger (2018) and use the carbon cycle from the DICE model (Nordhaus and Sztorc, 2013). There are three carbon reservoirs, the atmosphere (M_1), the upper ocean (M_2), and the deep ocean (M_3). According to IPCC (2013), the current

amounts are $M_1 = 830$ GtC, $M_2 = 900$ GtC, $M_3 = 37,255$ GtC. Additionally, I include a geological reservoir (M_4). The dynamics of the carbon stocks are given by

$$\begin{pmatrix} M_{1,t+1} \\ M_{2,t+1} \\ M_{3,t+1} \\ M_{4,t+1} \end{pmatrix} = \begin{pmatrix} \phi_{11} & \phi_{21} & \phi_{31} & \phi_{41} \\ \phi_{12} & \phi_{22} & \phi_{32} & \phi_{42} \\ \phi_{13} & \phi_{23} & \phi_{33} & \phi_{43} \\ \phi_{14} & \phi_{24} & \phi_{34} & \phi_{44} \end{pmatrix} \begin{pmatrix} M_{1,t} \\ M_{2,t} \\ M_{3,t} \\ M_{4,t} \end{pmatrix} + \begin{pmatrix} E_t^{net} + E_t^{\text{exo}} \\ G_{2,t} \\ G_{3,t} \\ G_{4,t} \end{pmatrix}, \tag{4.3}$$

or in matrix notation

$$\boldsymbol{M}_{t+1} = \boldsymbol{\Phi} \boldsymbol{M}_t + \boldsymbol{E_t}.$$

The transition matrix $\boldsymbol{\Phi}$ shows the rates of carbon flows between the reservoir types. I assume that the geological reservoir only exchanges with the atmosphere. I set the carbon flow coefficients ϕ_{14}, ϕ_{24}, ϕ_{34}, ϕ_{42}, ϕ_{43} equal to zero. The decay coefficient ϕ_{41} shows how much carbon leaks from the geological reservoir to the atmosphere. The persistence of carbon in reservoir M_4 is expressed by ϕ_{44}, and $\phi_{41} + \phi_{44} = 1$. Different to Lafforgue et al. (2008), I assume that there exist no capacity constraints for the geological reservoir.

CDR technologies allow to remove carbon from the atmosphere M_1, and to store it in another reservoir type M_i with i $\in \{2, 3, 4\}$. The amount of additional carbon that is stored in each reservoir is measured in GtC and denoted by $G_{i,t}$. Net emissions are given by the difference between the carbon emitted during the production process and the carbon that is removed from the atmosphere,

$$E_t^{net} = E_t - G_{1,t}. \tag{4.4}$$

In each time period, the sum of carbon injected in all reservoir types must be equal to the amount of carbon that is subtracted from the atmospheric reservoir, thus

$G_{1,t} = \sum_{i=2}^{r=4} G_{i,t}$. If $E_t^{net} > 0$, CDR can also be interpreted as carbon capture and storage. Net negative emissions are only present if $E_t^{net} < 0$.

The total amount of carbon that enters or leaves the atmosphere is given by the sum of net emissions E_t^{net}, and emissions from exogenous processes including land use change and forestry, which are collected in E_t^{exo} and also measured in GtC.

CDR consumes energy. It is therefore convenient to measure its operational costs $f_i(G_{i,t})$ in energy units. Thus, net energy input is the result of fossil energy net the energy used for CDR,

$$I_t = E_t - \sum_{i=2}^{r=4} f_i(G_{i,t}). \tag{4.5}$$

If a storage reservoir is not used, the corresponding costs of CDR are zero, $f_i(0) = 0$. Marginal costs are assumed to be positive and increasing for all storage units, $f_i'(G_{i,t}) > 0$ and $f_i''(G_{i,t}) > 0$. A reduction in carbon emissions into the atmosphere can either be achieved by reducing energy input E_t directly (mitigation) or by using CDR ($G_{i,t}$). Since the cost for mitigation and CDR deployment can both be measured in energy units (in GtC), reservoir i will only be used if its cost (and marginal cost) is lower than the cost (and marginal cost) of mitigation, thus $f_i(G_{i,t}) \leq G_{i,t}$, and $f_i'(G_{i,t}) \leq 1$. Note that without the option of CDR net energy input, emissions, and net emissions are equivalent, $I_t = E_t = E_t^{net}$.

A common assumption of analytic IAMs is that climate change damages have an exponential impact on global output via increases in atmospheric temperature (above the pre-industrial level). Following Golosov et al. (2014), I assume a direct mapping of climate change damages from the atmospheric carbon stock $M_{1,t}$. The damage function, that shows climate damage as a fraction of gross output, is given by

$$D_t(M_{1,t}) = 1 - \exp\left[-\xi_0 \left(M_{1,t} - M_1^{pre}\right)\right], \tag{4.6}$$

where M_1^{pre} denotes the pre-industrial atmospheric carbon concentration, and $\xi_0 > 0$ the climate change damage parameter that scales the marginal climate damage of atmospheric carbon. Output net climate change damages is therefore given by $Y_t^{net} = Y_t \left[1 - D_t \left(M_{1,t}\right)\right]$. The model does not include any impacts from increasing carbon concentrations in the ocean (e.g. from ocean acidification).

As most analytic IAMs, I assume full depreciation of capital over the course of 10 years, the model's time step. Thus, the economy's capital stock in the next period is given as the difference between net output Y_t^{net}, and consumption C_t,

$$\begin{aligned} K_{t+1} &= Y_t \left[1 - D_t \left(M_{1,t}\right)\right] - C_t \\ &= Y_t \exp\left[-\xi_0 \left(M_{1,t} - M_1^{pre}\right)\right] - C_t. \end{aligned} \tag{4.7}$$

The consumption rate is defined as $x_t = \frac{C_t}{Y_t^{net}}$, such that $1 - x_t$ is the savings rate.

I solve the model for a social planner who maximizes the infinite stream of consumption flows by choosing the consumption rate, emissions, and CDR deployment,

$$\max_{x_t, E_t, G_{i,t}} \sum_{t=0}^{\infty} \beta^t \log(C_t), \tag{4.8}$$

subject to the constraints imposed by the economy and the climate system, equations (4.1) to (4.7). The parameter β denotes the utility discount factor.

4.3 Theoretical results

This section presents the results of the climate-economy model, and compares them to the outcome of an alternative model specification without the option of CDR.

4.3.1 Carbon dioxide removal

Appendix C.1.1 solves the intertemporal optimization problem. It shows that the optimal rate of consumption is constant over time, $x_t^* = 1 - \beta\,\kappa$, and that the shadow

value of the fossil resource stock, denoted by $\varphi_{R,t}$, monotonically grows over time according to Hotelling's (1931) rule, $\varphi_{R,t} = \beta^{-t}\varphi_{R,0}$. In the following, I summarize the results on optimal CDR deployment.

Proposition 4.1. *The optimal level of CDR deployment for reservoir i is given by*

$$G^*_{i,t} = f_i'^{-1}\left(\frac{\beta\,\xi_0\left[(\mathbf{1}-\beta\,\mathbf{\Phi})^{-1}\right]_{1,1} - \beta\,\xi_0\left[(\mathbf{1}-\beta\,\mathbf{\Phi})^{-1}\right]_{1,i}}{\beta\,\xi_0\,\left[(\mathbf{1}-\beta\,\mathbf{\Phi})^{-1}\right]_{1,1} + (1-\beta\,\kappa)\beta^{-t}\varphi_{R,0}}\right), \tag{4.9}$$

where $[\cdot]_{1,1}$ denotes the first, and $[\cdot]_{1,i}$ denotes the i^{th} element of the first column of the inverted matrix in square brackets. Note that the inverse of the marginal cost function is expressed by $f_i'^{-1}$ and that $\left[(\mathbf{1}-\beta\,\mathbf{\Phi})^{-1}\right]_{1,1} > \left[(\mathbf{1}-\beta\,\mathbf{\Phi})^{-1}\right]_{1,i}$.

Proof. See Appendix C.1.2. □

Optimal CDR deployment is a function of constant model parameters, and the endogenously determined shadow value of the resource stock, which monotonically grows over time. Since $f_i'(G_{i,t})$ is an increasing function, also its inverse $f_i'^{-1}$ is an increasing function. Thus, optimal CDR immediately starts with its maximum level and then monotonically declines over time.

The interpretation of the carbon dynamics contributions follows Traeger (2018): The term $\left[(\mathbf{1}-\beta\,\mathbf{\Phi})^{-1}\right]_{1,1}$ characterizes the discounted sum of carbon persisting in and returning to the atmospheric carbon stock in all future periods. The term $\left[(\mathbf{1}-\beta\,\mathbf{\Phi})^{-1}\right]_{1,i}$ characterizes the long-term contribution to the atmospheric carbon reservoir from carbon that is currently stored in reservoir i.

The numerator in equation (4.9) shows the marginal benefit of the new technology. CDR reduces the marginal damage of emissions as it allows to remove carbon from the atmosphere and store it in a less damaging reservoir i. The denominator shows the marginal cost of fossil energy. It captures the opportunity cost of the resource and the marginal damage that it creates.

The magnitude of the benefit from CDR is determined by the difference in the carbon dynamics contributions of the atmosphere and reservoir i. A decrease in the

carbon persistence of reservoir i increases its carbon dynamics contribution as more carbon eventually finds its way into the atmosphere. This decreases the marginal benefit of CDR, and hence G_t^* declines. In contrast, an increase in the climate change damage parameter ξ_0 or an increase in the atmospheric carbon dynamics contribution $[(\mathbf{1} - \beta\,\mathbf{\Phi})^{-1}]_{1,1}$ raises the marginal damage of emissions and makes CDR technologies more attractive.

4.3.2 Emissions and energy input

Using the solution for CDR deployment allows to derive the optimal levels for emissions, and net energy input.

Proposition 4.2. *Optimal carbon emissions into the atmosphere are given by*

$$E_t^* = \frac{v}{\beta\,\xi_0\,\left[(\mathbf{1} - \beta\,\mathbf{\Phi})^{-1}\right]_{1,1} + (1 - \beta\,\kappa)\beta^{-t}\varphi_{R,0}} + \sum_{i=2}^{r=4} f_i\left(G_{i,t}^*\right), \tag{4.10}$$

with optimal CDR deployment $G_{i,t}^*$ as defined in equation (4.9).

Proof. See Appendix C.1.3. □

Optimal emissions are given by the sum of two terms. The first term captures the marginal benefit (numerator) and the marginal cost (denominator) from fossil energy. The term monotonically declines over time as the shadow value of the fossil resource increases. The second term shows the total cost of CDR deployment (measured in energy units). According to Proposition 4.1 optimal deployment monotonically declines over time, and thus optimal emissions decline over time as well.

An increase in $\varphi_{R,0}$ makes the fossil resource a more expensive input for production, and decreases both terms in equation (4.10). An increase in the carbon dynamics contribution $[(\mathbf{1} - \beta\,\mathbf{\Phi})^{-1}]_{1,i}$ increases the marginal damage from reservoir i. As a result CDR deployment declines, and thus optimal emissions are lower.

The outcome of an increase in $[(\mathbf{1} - \beta\,\mathbf{\Phi})^{-1}]_{1,1}$ and ξ_0 is ambiguous as there are two opposing effects. It decreases the first term in equation in (4.10) but leads to a higher level of CDR which increases the second term.

Using the solutions $G^*_{i,t}$ and E^*_t allows to solve for optimal net energy input I^*_t,

$$I^*_t = \frac{v}{\beta\,\xi_0\,[(\mathbf{1} - \beta\,\mathbf{\Phi})^{-1}]_{1,1} + (1 - \beta\,\kappa)\beta^{-t}\varphi_{R,0}}. \tag{4.11}$$

Net energy is defined as the difference between fossil energy and the energy spent on CDR. It is therefore equivalent to the first term in equation (4.10). Net energy input increases in the energy share v, and decreases in climate change damages ξ_0, the initial resource shadow value $\varphi_{R,0}$, and the carbon dynamics contribution $[(\mathbf{1} - \beta\,\mathbf{\Phi})^{-1}]_{1,1}$.

In order to gain insides on what changes due to the introduction of CDR, I specify an alternative scenario by removing the option of CDR from the climate-economy model in section 4.2. I denote the variables of the alternative model without CDR by a tilde. In the following, I show that the introduction of CDR influences the initial shadow value of the nonrenewable resource, and analyze how this affects net energy input, and net emissions. I discuss the implications of CDR for E^*_t in the subsequent section.

Proposition 4.3. *CDR increases the shadow value of the fossil resource, and decreases net energy input and net emissions.*

Proof. See Appendix C.1.4. □

CDR increases the value of the fossil fuel resource as it creates an additional option to mitigate the negative effects from carbon emissions, and thus reduces the social costs of using fossil fuels. Due to the linear-in-states property of the model, there is no direct effect of CDR on the marginal damage of carbon emissions. As a result, the net effect of CDR on the cost of the fossil resource is positive, and thus net energy input declines, $\Delta\, I^*_t \equiv I^*_t - \tilde{I}^*_t < 0$.

Next, I compare how net emissions differ between both model types. The difference is given by

$$\begin{aligned}\Delta E_t^{net^*} &\equiv E_t^{net^*} - \tilde{E}_t^{net^*} \\ &= I_t^* - \tilde{I}_t^* + \sum_{i=2}^{r=4} \left(f_i(G_{i,t}^*) - G_{i,t}^* \right) < 0,\end{aligned}$$

since $\tilde{I}_t^* > I_t^*$ and $f_i(G_{i,t}^*) \leq G_{i,t}^*$. CDR leads to lower net emission over the entire time path. This result is driven by two effects. First, as already shown CDR lowers net energy input, and second, the cost of CDR is lower than the cost of mitigation (both measured in energy units).

4.3.3 Social cost of carbon

This section derives the SCC for all reservoir types and explores how CDR influences the optimal carbon tax. Due to the linear-in-states property of the model the marginal damage for each reservoir type is independent of its stock size. This leads to the following result.

Proposition 4.4. *CDR leaves the structure of the atmospheric SCC (optimal carbon tax) unchanged. The SCC for reservoir i is proportional to net output,*

$$SCC_{Mi} = Y_t^{net}\, \xi_0 \, \left[(\mathbf{1} - \beta\, \mathbf{\Phi})^{-1} \right]_{1,i}. \tag{4.12}$$

As defined above, $[\cdot]_{1,i}$ denotes the i^{th} element of the first column of the inverted matrix in square brackets.

Proof. See Appendix C.1.5. □

The persistence of carbon differs between reservoir types such that each reservoir has its own SCC. For the DICE carbon cycle, the carbon dynamics contribution of the deep ocean is smaller than the carbon dynamics contribution of the shallow

ocean. This leads to the following ordering: $SCC_{M1} > SCC_{M2} > SCC_{M3}$. The carbon dynamics contribution of reservoir $M_{4,t}$ (the geological storage unit) depends on the decay rate ϕ_{41}. If it is a secure deposit and the decay rate is zero, then $SCC_{M4} = 0$.

Deriving the atmospheric SCC (optimal carbon tax) for the alternative model specification without CDR leads to the same result as in equation (4.12). The availability of CDR leaves the analytic structure of the atmospheric SCC unchanged. This result is driven by two crucial assumptions of analytic IAMs. First, utility is a logarithmic function of consumption, and second, climate change damages have an exponential impact on output. These two assumptions ensure that the climate-economy model is linear-in-states and can be solved by a linear affine value function (Karp, 2017). The linear-in-states property implies that the marginal damage from an additional unit of carbon in the atmosphere is constant and does not depend on the atmospheric carbon concentration. Hence, removing a unit of carbon from the atmosphere has no effect on the marginal damage, and the atmospheric SCC. This is different for other geoengineering measures such as stratospheric aerosol injections (see Chapter 3).

Next, I analyze how the level of the SCC is affected by the availability of CDR, compared to the situation without CDR. According to Proposition 4.3 CDR leads to a lower net energy input over the entire time path. As a result initial output declines. Since the initial atmospheric carbon concentration and initial climate change damages are equivalent for both model types, initial net output decreases as well. This lowers the initial level of the atmospheric SCC, and therefore rises the level of emissions in the beginning. However, as CDR is an option it must increase net output eventually as otherwise it would no be used. Thus, there must exist a period in the future in which the SCC is higher compared to the model without CDR. The interpretation of this result is straight forward. As climate change damages are measured in percent of output, an increase in Y_t also increases

the money-measured welfare loss from global warming. In other words: The better off the economy is, the more the economy loses from climate change. The next section quantifies this effect.

4.4 Quantitative analysis

This section illustrates the previous theoretical findings. It provides a calibration of the climate-economy model for a high and low-cost scenario of oceanic CDR, and compares the results to the alternative model specification without CDR.

4.4.1 Climate-economy model without CDR

The simulation starts in t = 2010 and ends in t = 2200 with one period representing ten years, which is a standard in the literature. Economic growth is driven by increasing total factor productivity $A_{0,t}$, which develops exogenously over time according to

$$A_{0,t} = A_0 \, (1 + w)^t. \tag{4.13}$$

The growth rate of total factor productivity is assumed to be 2 percent per year, $w = 0.02$. The initial population is set to 6.9 billion and assumed to grow logistically over time to a maximum of 11 billion in 2200 as in Gerlagh and Lsiki (2018). Output for the initial decade is set to 700 trillion (tn) USD. I use the same shares of capital, $\alpha = 0.3$, and net energy, $v = 0.04$, as in Golosov et al. (2014). The utility discount rate is set to 1.4 percent per year (Traeger, 2018). The given parameter set implies an optimal constant savings rate of $s \approx 0.25$. The initial capital stock is assumed to be 135 trillion USD, approximately the output of two years, and fully depreciates over the course of a decade. I use the carbon cycle from DICE 2013 (Nordhaus and Sztorc, 2013), and the climate change damage parameter $\xi_0 = 5.3 \, \text{x} \, 10^{-5}$ from Golosov et al. (2014). The pre-industrial carbon stock is set to 600 GtC. The carbon concentration for the first decade is set to 830.4 GtC yielding initial climate change

damages of $D_0 = 1.2$ percent.

Assuming emissions of 86.7 GtC for the first decade (Gerlagh and Lsiki, 2018) allows to solve for the initial level of total factor productivity, and delivers $A_0 = 38$. I then calibrate the initial resource stock such that it matches the initial level of emissions. This implies an initial fossil fuel stock size of 793.25 GtC. Table 4.1 summarizes the model parameters and initial stock values.

Table 4.1: Parameter values

K_0	N_0	R_0	κ	v	β	w	A_0	ξ_0
135	6.9	793.25	0.3	0.04	0.986	0.02	38.02	5.3×10^{-5}
trillion USD	billion	GtC			1/year	1/year		1/GtC

Figure 4.1 shows the outcome of the standard model without CDR. The fossil resource is scarce and used up over the time horizon. Emissions start at 86 GtC per decade and monotonically decline over time as the shadow price of the resource increases. Damages start at 1.2 percent of global output and increase up to around 3 percent by the year 2100. Afterwards, damages start to decline as less energy is used and more carbon is taken up by the oceans. Relative net production (GDP) rises over time due to the growth of total factor productivity. The atmospheric SSC starts at around 45 USD/tCO_2 and increases up to around 800 USD/tCO_2 by the year 2100. All these results are very much in line with results of common IAMs (e.g. Golosov et al., 2014).

4.4.2 Climate-economy model with oceanic CDR

This section introduces the option of oceanic CDR and explores how it affects the outcome of the standard climate-economy model. Cost estimates for the storage of carbon in the oceans are still uncertain and vary widely. IPCC (2005) estimates the cost for oceanic storage between 22 and 114 USD/tC. Rickels et al. (2018) consider a convex cost function with a broad parameter range for the quadratic cost term to account for uncertainty about the cost of large-scale deployment.

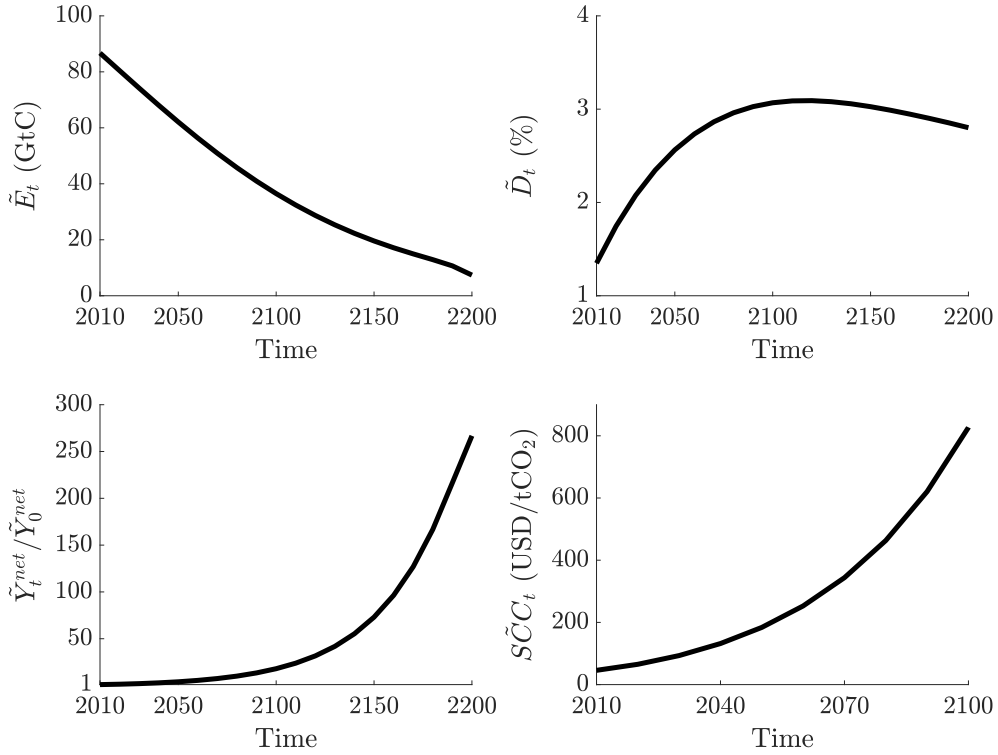

Figure 4.1: The graph shows emissions per decade ($\tilde{E}_t$), damages ($\tilde{D}_t$), relative net output ($\tilde{Y}_t^{net}/\tilde{Y}_0^{net}$), and the social cost of carbon ($S\tilde{C}C_t$) for the calibrated standard climate-economy model without CDR.

To capture the cost uncertainty for oceanic CDR I consider a low and high-cost scenario. For the low-cost case, the cost function for CDR is given by

$$f_l(G_t) = g_l\, G_t^2, \tag{4.14}$$

with parameter g_l to be calibrated. As a point of reference, I use the linear quadratic cost function from Rickels et al. (2018) and combine it with the lower bound cost estimate for oceanic storage of 22 USD/tC from IPCC (2005), which leads to

$$F(G_t) = 0.022\, G_t + 0.01833\, G_t^2. \tag{4.15}$$

CDR deployment G_t is measured in GtC and $F(G_t)$ shows the costs in trillion USD (tn USD). I calibrate the cost function $f_l(G_t)$ to equation (4.15) for the initial time

period. Minimizing the squared difference over the interval $G_t \in (0, 18.5)$ yields $g_l = 0.056$. I choose this interval since for $G_t \geq 18.5$ the cost of CDR is higher than the cost of mitigation. Figure 4.2 shows the quality of the fit, and the cost of mitigation in trillion USD. For the high-cost scenario, I consider the upper bound

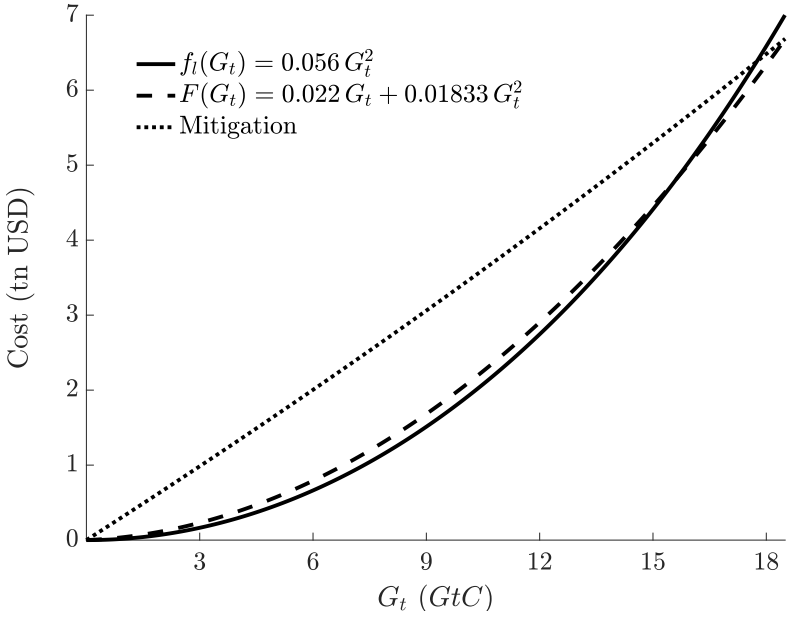

Figure 4.2: The graph shows the calibrated cost function $f_l(G_t)$ (solid line), the cost estimate based on Rickels et al. (2018) and IPCC (2005) (dashed line), and the cost of mitigation (dotted line).

of previous estimates. As the upper bound cost estimate is expected to surpass the lower bound cost estimate by a factor of five (IPCC, 2005), I assume $g_h = 5 \times g_l$. Due to the assumption of a quadratic cost function the level of CDR will still be positive but considerably lower than in the low-cost scenario. Figures 4.3 and 4.4 show how the results change due to the introduction of CDR. The black solid lines show the results for the low-cost scenario and the dotted green lines show the outcome for the high-cost scenario.

The simulation illustrates the analytic results from the previous section. In the first decade, in the low-cost case around 4.5 GtC are removed from the atmosphere and stored in the deep ocean. In the high-cost scenario, CDR deployment is considerably lower with only 1 GtC in the first decade. As described in Propositions 4.1

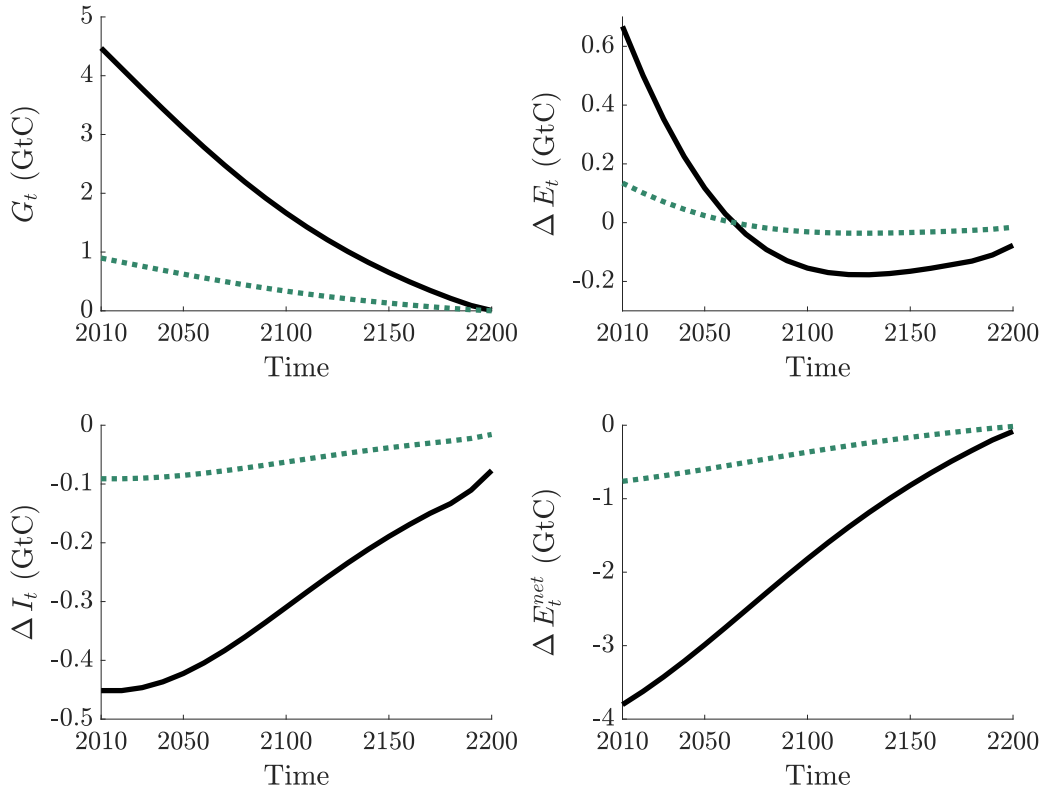

Figure 4.3: The graph shows the optimal deployment of oceanic CDR (G_t) per decade and the difference in emissions (ΔE_t), net energy input (ΔI_t), and net emissions (ΔE_t^{net}) compared to the outcome of the standard model without CDR for the low cost (black solid lines), and high cost scenario (green dotted lines).

and 4.2 CDR deployment and emissions monotonically decline over time. In line with Proposition 4.3, net emissions and net energy input is lower over the entire time horizon compared to the model without CDR. In both scenarios emissions are first higher and then lower than in the model without CDR. In the low cost scenario the difference is more pronounced.

In the low-cost case, CDR reduces the atmospheric carbon concentration by 20 GtC in 2125 and damages are lower by around 0.1 percentage points of output. Towards the end, the negative effect on the atmospheric carbon concentration and climate damages wears off as CDR deployment goes to zero and more and more carbon has cycled back from the oceans. In the high-cost case, the negative effect on atmospheric carbon is minor and only decreases damages by around 0.01 percent.

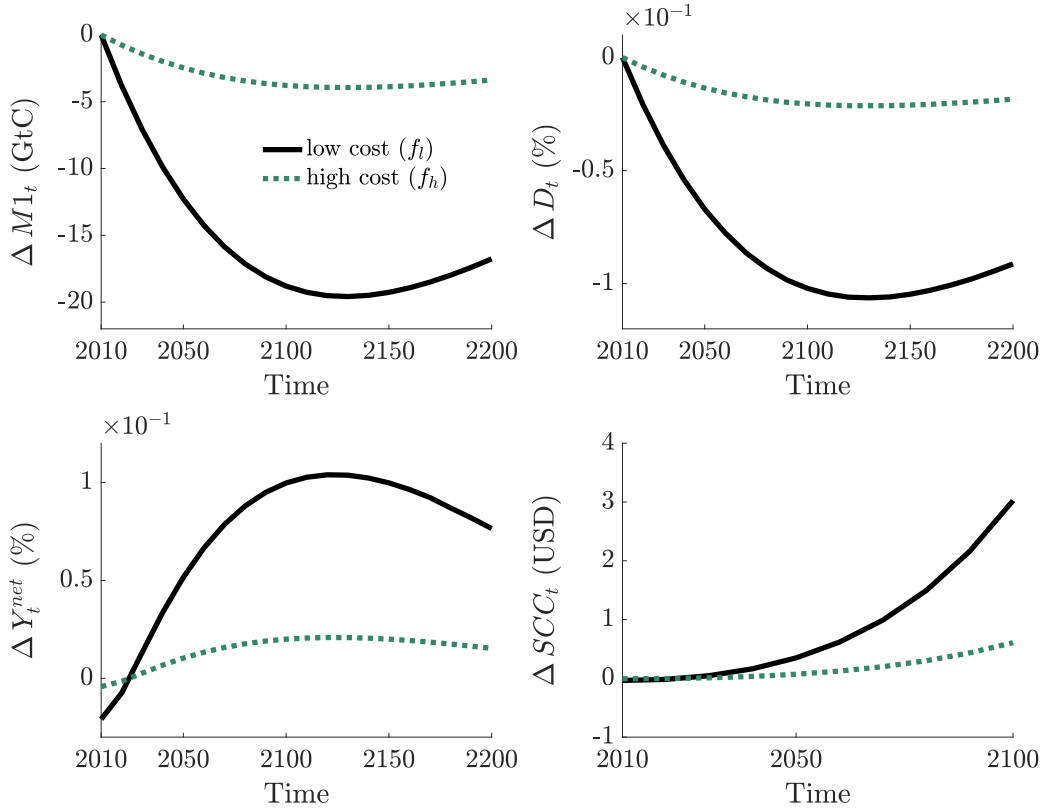

Figure 4.4: The graph shows the difference in atmospheric carbon concentration ($\Delta M_{1,t}$), climate change damages (ΔD_t), net output (ΔY_t^{net}), and the social cost of carbon (ΔSCC_t) compared to the outcome of the standard model without CDR for the low cost (black solid lines), and high cost scenario (green dotted lines).

The numerical simulation also allows to assess how strongly net output and the atmospheric SCC (optimal carbon tax) are affected by the introduction of CDR. As already discussed in the theoretical part of the paper initial net output declines as CDR becomes available. Figure 4.4 shows that this effect is rather small. Net output declines by 0.025 percent in the low-cost scenario. Afterwards, the effect on net output becomes positive and grows until 2125 to around 0.11 percent. Similar to net output, the SCC is first lower and then higher. The economy first emits more and then less. The simulation shows that the effect of CDR on the SCC is minor. By 2100 the SCC is only higher by 3 USD/tCO_2 compared to the model without CDR.

4.5 Summary and conclusions

The paper introduces the option of carbon dioxide removal (CDR) and storage in different reservoir types into an analytic climate-economy model and compares the results to a model variant without CDR. The analytic model shows that the availability of CDR alters the level of the SCC. However, the quantitative analysis suggests that this effect is negligible. In the low-cost scenario, CDR increases initial emissions by around 0.6 GtC, which is equivalent to around 0.7 percent of total carbon emissions. Thus, with an optimal policy in place the introduction of CDR has hardly any effect on mitigation incentives. The model suggests that CDR is needed on top of traditional mitigation efforts.

Furthermore, the paper provides basic implications for the optimal implementation of CDR technologies. One option that has been proposed in the literature is the introduction of a differentiated carbon tax (Rickels and Lontzek, 2012). This paper presents a simple formula for the reservoir specific carbon tax, and characterizes its components. Another suggestion for the optimal implementation of CDR is the introduction of carbon credits (Chomitz and Lecocq, 2004; Sedjo and Marland, 2003), for which this paper also offers a simple way to calculate it. Finally, I note that the current setting does not include renewable energies, which would be an interesting extension to explore.

References

Anderson, Soren and Newell, Richard. Prospects for carbon capture and storage technologies. Annu. Rev. Environ. Resour., 29:109–142, 2004.

Chomitz, Kenneth M and Lecocq, Franck. Temporary sequestration credits: an instrument for carbon bears. Climate Policy, 4(1):65–74, 2004.

Nordhaus, William and Sztorc, Paul. DICE2013R: Introduction and User's Manual. 2013. URL `http://www.econ.yale.edu/~nordhaus/homepage/homepage/documents/DICE_Manual_100413r1.pdf`.

Geden, Oliver; Peters, Glen P, and Scott, Vivian. Targeting carbon dioxide removal in the European Union. Climate Policy, 19(4):487–494, 2019.

Gerlagh, Reyer and Liski, Matti. Consistent climate policies. Journal of the European Economic Association, 16(1):1–44, 2018.

Golosov, Mikhail; Hassler, John; Krusell, Per, and Tsyvinski, Aleh. Optimal Taxes on Fossil Fuel in General Equilibrium. Econometrica, 82(1):41–88, 2014.

Hotelling, Harold. The economics of exhaustible resources. Journal of Political Economy, 39(2):137–175, 1931.

IPCC, The Intergovernmental Panel on Climate Change. Special Report on Carbon Dioxide Capture and Storage. Prepared by Working Group III of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2005.

IPCC, The Intergovernmental Panel on Climate Change. Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2013.

Karp, Larry. Provision of a public good with multiple dynasties. The Economic Journal, 127(607):2641–2664, 2017.

Lafforgue, Gilles; Magné, Bertrand, and Moreaux, Michel. Energy substitutions, climate change and carbon sinks. Ecological Economics, 67(4):589–597, 2008.

Rickels, Wilfried; Reith, Fabian; Keller, David; Oschlies, Andreas, and Quaas, Martin F. Integrated assessment of carbon dioxide removal. Earth's Future, 6(3):565–582, 2018.

Rickels, Wilfried and Lontzek, Thomas S. Optimal global carbon management with ocean sequestration. Oxford Economic Papers, 64(2):323–349, 2012.

Sabine, Christopher L; Feely, Richard A; Gruber, Nicolas; Key, Robert M; Lee, Kitack; Bullister, John L; Wanninkhof, Rik; Wong, CS; Wallace, Douglas WR; Tilbrook, Bronte, and others. The oceanic sink for anthropogenic CO2. Science, 305:367–371, 2004.

Sedjo, Roger A and Marland, Gregg. Inter-trading permanent emissions credits and rented temporary carbon emissions offsets: some issues and alternatives. Climate Policy, 3(4):435–444, 2003.

Tanaka, Katsumasa and O'Neill, Brian C. The Paris Agreement zero-emissions goal is not always consistent with the 1.5 °C and 2 °C temperature targets. Nature Climate Change, 8:319–324, 2018.

Traeger, Christian P. ACE - Analytic Climate Economy (with Temperature and Uncertainty). 2018. URL `https://ssrn.com/abstract=3307622`.

van der Zwaan, Bob and Gerlagh, Reyer. Economics of geological CO2 storage and leakage. Climatic change, 93:285–309, 2009.

C.1 Appendix to Chapter 4

C.1.1 Solving the linear-in-states model

For the proof of the linear-in-states property I follow Traeger (2018). The consumption rate can be written as

$$x_t = \frac{C_t}{Y_t \left[1 - D_t \left(M_{1,t}\right)\right]},$$

such that

$$\log C_t = \log x_t + \log A_{0,t} + \kappa \log K_t + (1 - \kappa - v) \log N_{0,t} + v \log I_t - \xi_0 \left(M_{1,t} - M_1^{pre}\right).$$

I transform the optimization problem into its dynamic programming form (Bellman equation)

$$V(k_t, \boldsymbol{M}_t, R_t, t) = \max_{x_t, E_t, G_t} \Bigg\{ \log x_t + \log A_{0,t} + \kappa \log K_t + (1 - \kappa - v) \log N_{0,t} + v \log I_t(E_t, G_{i,t}) - \xi_0 \left(M_{1,t} - M_1^{pre}\right) + \beta \, V(k_{t+1}, \boldsymbol{M}_{t+1}, R_{t+1}, t+1) \Bigg\},$$

where $k_t = \log K_t$ with the equation of motion

$$k_{t+1} = \log A_{0,t} + \kappa \log K_t + (1 - \kappa - v) \log N_{0,t} + v \log I_t - \xi_0 \left(M_{1,t} - M_1^{pre}\right) + \log(1 - x_t). \tag{4.16}$$

To solve the intertemporal optimization problem, I use the following guess for the value function

$$V(k_t, \boldsymbol{M}_t, R_t, t) = \varphi_k \, k_t + \boldsymbol{\varphi}_M^T \, \boldsymbol{M}_t + \varphi_{R,t} \, R_t + \varphi_t, \tag{4.17}$$

where φ is used to denote the shadow values for the different states, and ^{T} denotes the transpose of a vector of shadow values.

Inserting the trial solution and the next periods states (equations 4.2, 4.3, and 4.16)

into the Bellman equation delivers

$$
\begin{aligned}
&\varphi_k\, k_t + \boldsymbol{\varphi}_M^T\, \boldsymbol{M}_t + \varphi_{R,t}\, R_t + \varphi_t \\
&= \max_{x_t, E_t, G_{i,t}} \Bigg\{ \log x_t + \log A_{0,t} + \kappa\, k_t + (1-\kappa-v) \log N_{0,t} + v \log I_t(E_t, G_{i,t}) - \xi_0\, (M_{1,t} - M_1^{pre}) \\
&+ \beta\, \varphi_k \left(\log A_{0,t} + \kappa\, k_t + (1-\kappa-v) \log N_{0,t} + v \log I_t(E_t, G_{i,t}) - \xi_0\, (M_{1,t} - M_1^{pre}) + \log(1-x_t) \right) \\
&+ \beta\, \boldsymbol{\varphi}_M^T\, (\boldsymbol{\Phi} \boldsymbol{M}_t + \boldsymbol{E}_t) + \beta\, \varphi_{R,t+1}\, (R_t - E_t) + \beta\, \varphi_{t+1} \Bigg\}. \quad (4.18)
\end{aligned}
$$

First order conditions. Maximizing the right hand side over x_t yields

$$
\frac{1}{x_t} - \beta\, \varphi_k \frac{1}{1-x_t} = 0 \quad \Longrightarrow \quad x_t^* = \frac{1}{1+\beta\, \varphi_k}. \quad (4.19)
$$

Next, I find the first order condition for CDR deployment for reservoir i

$$
-v(1+\beta\, \varphi_k)\, \frac{f_i'(G_{i,t})}{I_t} = \beta(\varphi_{M1} - \varphi_{Mi}), \quad (4.20)
$$

and the first order condition for emissions

$$
v(1+\beta\, \varphi_k)\, \frac{1}{I_t} = \beta(\varphi_{R,t+1} - \varphi_{M1}). \quad (4.21)
$$

Inserting (4.21) into (4.20) and solving for $G_{i,t}$ leads to

$$
G_{i,t}^* = f_i'^{-1} \left(\frac{\varphi_{M1} - \varphi_{Mi}}{\varphi_{M1} - \varphi_{R,t+1}} \right), \quad (4.22)
$$

where the inverse of the marginal cost function is denoted by $f_i'^{-1}$. Summing up CDR deployment over all reservoir types yields

$$
G_{1,t}^* = \sum_{i=2}^{r=4} f_i'^{-1} \left(\frac{\varphi_{M1} - \varphi_{Mi}}{\varphi_{M1} - \varphi_{R,t+1}} \right)
$$

Using (4.22) and solving for optimal emissions yields

$$E_t^* = \frac{v(1+\beta\,\varphi_k)}{\beta(\varphi_{R,t+1}-\varphi_{M1})} + \sum_{i=2}^{r=4} f_i\left(f_i'^{-1}\left(\frac{\varphi_{M1}-\varphi_{Mi}}{\varphi_{M1}-\varphi_{R,t+1}}\right)\right). \tag{4.23}$$

First order conditions deliver optimal controls x_t^*, E_t^*, and $G_{i,t}^*$ which are independent of the states.

Using E_t^* and $G_{i,t}^*$ one can solve for the optimal net energy input I_t^*.

$$I_t^* = E_t^* - \sum_{i=2}^{r=4} f_i(G_{i,t}^*) = \frac{v(1+\beta\,\varphi_k)}{\beta(\varphi_{R,t+1}-\varphi_{M1})}. \tag{4.24}$$

Inserting the optimal controls into (4.18) and arranging terms with respect to their states yields

$$\begin{aligned}\varphi_k\,k_t + \boldsymbol{\varphi}_M^T\,\boldsymbol{M}_t + \varphi_{R,t}\,R_t + \varphi_t = \Big[(1+\beta\,\varphi_k)\kappa\Big]k_t + \Big[\beta\,\boldsymbol{\Phi}\boldsymbol{\varphi}_M^T - (1+\beta\,\varphi_k)\xi_0\,\boldsymbol{e}_1^T\Big]\boldsymbol{M}_t \\ + \Big[\beta\,\varphi_{R,t+1}\Big]R_t + \log\,x_t^* + \beta\,\varphi_k\,\log(1-x_t^*) + (1+\beta\,\varphi_k)\log A_{0,t} + (1+\beta\,\varphi_k)(1-\kappa-v)\log N_{0,t} \\ + (1+\beta\,\varphi_k)v\log I_t^* + (1+\beta\,\varphi_k)\xi_0\,M_1^{pre} + \beta\,\varphi_{M1}(E_t^* + E_t^{\text{exo}} - G_{1,t}^*) + \beta\,\varphi_{M2}\,G_{2,t}^* + \beta\,\varphi_{M3}\,G_{3,t}^* \\ + \beta\,\varphi_{M4}\,G_{4,t}^* - \beta\,\varphi_{R,t+1}\,E_t^* + \beta\,\varphi_{t+1}.\end{aligned} \tag{4.25}$$

Given the optimal controls the maximized Bellman equation is linear in all states.

Shadow values. Coefficient matching with respect to capital, k_t, yields

$$\varphi_k = (1+\beta\,\varphi_k)\kappa \quad \Leftrightarrow \quad \varphi_k = \frac{\kappa}{1-\beta\,\kappa} \tag{4.26}$$

Inserting φ_k into equation (4.19) yield the optimal consumption rate $x_t^* = 1-\beta\,\kappa$.

I match the coefficients of each state from both sides of the equation, which leads to

$$\boldsymbol{\varphi}_M^T = -\xi_0\,(1+\beta\,\varphi_k)\,\boldsymbol{e}_1^T\,[\mathbf{1}-\beta\,\boldsymbol{\Phi}]^{-1}$$

Using (4.26) the vector of shadow prices turns to

$$\boldsymbol{\varphi}_M^T = -\xi_0\,\frac{1}{1-\beta\,\kappa}\,\boldsymbol{e}_1^T\,[\mathbf{1}-\beta\,\boldsymbol{\Phi}]^{-1} \tag{4.27}$$

Coefficient matching with respect to the resource stock yields

$$\varphi_{R,t} = \beta\,\varphi_{R,t+1} \quad \Leftrightarrow \quad \varphi_{R,t} = \beta^{-t}\varphi_{R,0} \quad \text{(Hotelling's rule)}. \tag{4.28}$$

The initial resource values $\varphi_{R,0}$ depend on the set up of the economy, including assumptions about production and the energy sector. Given the coefficients and the optimal rate of consumption equation (4.25) turns to the following condition:

$$\begin{aligned}\varphi_t - \beta\,\varphi_{t+1} = \log\, x_t^* + \beta\,\varphi_k\, \log(1-x_t^*) + (1+\beta\,\varphi_k)\log A_{0,t} + (1+\beta\,\varphi_k)(1-\kappa-v)\log N_{0,t} \\ + (1+\beta\,\varphi_k)v\log I_t^* + (1+\beta\,\varphi_k)\xi_0\, M_1^{pre} + \beta\,\boldsymbol{\varphi}_M^T\,\boldsymbol{E}_t^* - \beta\,\varphi_{R,t+1}\,E_t^*\end{aligned}$$

This condition will be satisfied by picking the sequence $\varphi_0, \varphi_1, \varphi_2, ...$ The additional condition $\lim_{t\to\infty}\beta^t V(\cdot) = 0 \Rightarrow \lim_{t\to\infty}\beta^t\varphi_t = 0$ pins down this initial value φ_0.

C.1.2 Proof of Proposition 4.1

Inserting the solutions for the shadow values, equations (4.26) to (4.28), into (4.22) yields

$$G_{i,t}^* = f_i'^{-1}\left(\frac{\beta\,\xi_0\left[(\mathbf{1}-\beta\,\mathbf{\Phi})^{-1}\right]_{1,1} - \beta\,\xi_0\left[(\mathbf{1}-\beta\,\mathbf{\Phi})^{-1}\right]_{1,i}}{\beta\,\xi_0\,\left[(\mathbf{1}-\beta\,\mathbf{\Phi})^{-1}\right]_{1,1} + (1-\beta\,\kappa)\beta^{-t}\varphi_{R,0}}\right), \tag{4.29}$$

where $[\cdot]_{1,1}$ denotes the first, and $[\cdot]_{1,i}$ denotes the i^{th} element of the first column of the inverted matrix in square brackets. Note that $\left[(\mathbf{1}-\beta\,\mathbf{\Phi})^{-1}\right]_{1,1} > \left[(\mathbf{1}-\beta\,\mathbf{\Phi})^{-1}\right]_{1,i}$.

C.1.3 Proof of Proposition 4.2

Inserting the solutions for the shadow values, equations (4.26) to (4.28), into (4.23) yields

$$E_t^* = \frac{v}{\beta\,\xi_0\,\left[(\mathbf{1}-\beta\,\mathbf{\Phi})^{-1}\right]_{1,1} + (1-\beta\,\kappa)\beta^{-t}\varphi_{R,0}} + \sum_{i=2}^{r=4} f_i\left(G_{i,t}^*\right), \tag{4.30}$$

where

$$G_{i,t}^* = f_i'^{-1}\left(\frac{\beta\,\xi_0\left[(\mathbf{1}-\beta\,\mathbf{\Phi})^{-1}\right]_{1,1} - \beta\,\xi_0\left[(\mathbf{1}-\beta\,\mathbf{\Phi})^{-1}\right]_{1,i}}{\beta\,\xi_0\,\left[(\mathbf{1}-\beta\,\mathbf{\Phi})^{-1}\right]_{1,1} + (1-\beta\,\kappa)\beta^{-t}\varphi_{R,0}}\right).$$

C.1.4 Proof of Proposition 4.3

Consider the climate-economy model from section 4.2 without the option of CDR, and let the variables of this model specification be denoted by a tilde.

From the first order condition (4.20) it follows that optimal emissions without the option of CDR are given by

$$\tilde{E}_t^* = \frac{v}{\beta\,\xi_0\,\left[(\mathbf{1}-\beta\,\mathbf{\Phi})^{-1}\right]_{1,1} + (1-\beta\,\kappa)\beta^{-t}\tilde{\varphi}_{R,0}}. \tag{4.31}$$

The only endogenous term in equation (4.31) is the initial shadow value of the resource stock, which is denoted by $\tilde{\varphi}_{R,0}$. In both model specifications, the size of the resource stock is the same and will be used up eventually. Therefore,

$$R_0 = \sum_{t=0}^{\infty} E_t^* = \sum_{t=0}^{\infty} \tilde{E}_t^*.$$

Using equations (4.30) and (4.31), and rearranging leads to

$$\sum_{t=0}^{\infty}\sum_{i=2}^{r=4} f_i(G_{i,t}^*) = \sum_{t=0}^{\infty}\left(\frac{v}{\beta\,\xi_0\,\left[(\mathbf{1}-\beta\,\mathbf{\Phi})^{-1}\right]_{1,1} + (1-\beta\,\kappa)\beta^{-t}\tilde{\varphi}_{R,0}} - \frac{v}{\beta\,\xi_0\,\left[(\mathbf{1}-\beta\,\mathbf{\Phi})^{-1}\right]_{1,1} + (1-\beta\,\kappa)\beta^{-t}\varphi_{R,0}}\right).$$

If there exists at least one point in time where $\sum_{i=2}^{r=4} f_i(G_{i,t}^*) > 0$, the left term of the equation is positive, and thus $\tilde{\varphi}_{R,0} < \varphi_{R,0}$. From this it directly follows that $\Delta\, I_t^* \equiv I_t^* - \tilde{I}_t^* < 0$. Comparing net emissions with and without the option of CDR yields

$$\begin{aligned}\Delta E_t^{net^*} &\equiv E_t^{net^*} - \tilde{E}_t^{net^*} \\ &= I_t^* - \tilde{I}_t^* + \sum_{i=2}^{r=4}\left(f_i(G_{i,t}^*) - G_{i,t}^*\right) < 0,\end{aligned}$$

since $\tilde{I}_t^* > I_t^*$ and $f_i(G_{i,t}) \leq G_{i,t}^*$.

C.1.5 Proof of Proposition 4.4

The SCC is the negative of the shadow value of carbon reservoir i expressed in money-measured consumption units,

$$\begin{aligned} SCC_{Mi} &= -(1-\beta\,\kappa)Y_t^{net}\,\varphi_{Mi} \\ &= Y_t^{net}\,\xi_0\,\left[(\mathbf{1}-\beta\,\mathbf{\Phi})^{-1}\right]_{1,i}, \end{aligned}$$

where again $[\cdot]_{1,i}$ denotes the i^{th} element of the first column of the inverted matrix in square brackets.

Eidesstattliche Erklärung

Ich erkläre hiermit, dass ich meine Doktorarbeit

Economics of using fossil fuels and tackling climate change

selbstständig und ohne fremde Hilfe angefertigt habe und dass ich als Koautor maßgeblich zu den weiteren Fachartikeln beigetragen habe. Alle von anderen Autoren wörtlich übernommenen Stellen, wie auch die sich an die Gedanken anderer Autoren eng anlehnenden Ausführungen der aufgeführten Beiträge wurden besonders gekennzeichnet und die Quellen nach den mir angegebenen Richtlinien zitiert.

Unterschrift:

Datum: