An Analysis of Emission Reduction Strategies in Upstream and Midstream Oil Operations

Zur Erlangung des akademischen Grades eines

Doktors der Wirtschaftswissenschaften

(Dr. rer. pol.)

von der KIT-Fakultät für Wirtschaftswissenschaften
des Karlsruher Instituts für Technologie (KIT)

genehmigte

Dissertation

von

Giacomo Benini, PhD

Tag der mündlichen Prüfung:	26.03.2025
Erster Gutachter:	Prof. Dr. Wolf Fichtner
Zweiter Gutachter:	Prof. Dr. Mohammad Masnadi

Karlsruhe (2025)

Acknowledgments

Completing a dissertation is not a solitary endeavour; it is the result of the combined efforts, guidance, and support of many people. I would like to express my gratitude to those who have helped me along this journey.

My first thanks go to my Doktorvater, Prof. Dr. Wolf Fichtner. Accepting an external doctoral student is not an easy task. The lack of physical presence complicates the development of a professional relationship. However, Prof. Dr. Fichtner never made me feel like an outsider. I always felt like a member of the KIT student body - perhaps a distant one, but a part nonetheless. Thanks to his feedback, I have improved the quality of my work, giving it a more coherent structure. His contribution was crucial in developing the analysis of the electricity market in the third research paper. In particular, his knowledge of the power-gas network was critical in developing the optimal taxation structure. I am particularly grateful for the enormous amount of freedom he gave me. When the thesis deviated from its initial design, instead of hindering its development and forcing it toward a more secure outcome, he was very supportive in exploring new possibilities. I will always be thankful for the flexibility he afforded me.

The quality of my work would not have been the same without the unwavering support of my co-supervisor, Prof. Dr. Mohammad Masnadi. Throughout my PhD journey, Prof. Masnadi provided invaluable guidance and encouragement. He was always available to discuss my research, offering critical insights that significantly enhanced my understanding and approach. His expertise in energy systems and policy greatly enriched my work, particularly in refining the methodologies and ensuring the robustness of my results. Prof. Masnadi's dedication to fostering my academic growth was evident in his detailed reviews and constructive feedback, which were instrumental in shaping the final form of this dissertation. His idea to connect the global petroleum merit base curve with the cumulative emission curve was instrumental in developing the first two research papers. I am deeply grateful for his mentorship and the positive impact he had on my doctoral experience.

I would also like to thank Dr. Patrick Jochem, who served as my supervisor during the early stages of my doctoral journey. Dr. Jochem's expertise in sustainable mobility and energy systems laid the foundation for the initial development of my research. His thoughtful guidance and support were critical in helping me shape the original direction of the thesis. Although he later transitioned to a new position at DLR, his early contributions continue to resonate throughout the final work, and for that, I am grateful.

During my time as an external PhD student at KIT, I was employed by Saudi Aramco (one year), Stanford University (four years), and the Norwegian School of Economics (two years). Over this period, I was fortunate enough to collaborate with great scientists who helped me develop a global view of the petroleum sector. At Aramco, Dr. Hassan El-Houjeiri introduced me to advanced petroleum economics, while helping me link abstract concepts to accounting data. During the same period, I enriched my vision thanks to Dr. Steve Przesmitzki, Prof. Dr. Xin He, and Dr. Lang Sui, who helped me see the production-emission problem from the perspective of the automotive sector. At Stanford, Prof. Dr. Adam R. Brandt was my supervisor. He is the most multidisciplinary person I have ever met - a true Renaissance man operating at the intersection of energy engineering, numerical optimization, physics, earth science, and economics. This thesis would not have been possible without his enthusiastic support and patience. I would also like to thank all the PhD students of the Stanford Energy Resource Engineering group for their useful comments. In particular, I acknowledge PhD candidate Zhan Zhang, who helped me construct various databases. Over the seven years, the person who helped me the most was my co-author and friend, Prof. Dr. Valerio Dotti. He framed my confused ideas into well-posed mathematical problems and helped me understand the economic implications of the equilibria presented in the second and third research papers.

My appreciation also extends to those who provided me with the technical tools to conceive this dissertation. In this regard, the biggest thanks go to Prof. Dr. Stefan A. Sperlich. With immense patience, he taught me statistics and econometrics over four intense years. The treatment of co-integration, endogeneity, and heterogeneity across the entire thesis would not have been possible without his previous teachings.

Finally, I would like to acknowledge the contribution of my family and friends. The biggest thanks go to my wife, Alice. I remember when I told her that I was going to try to complete a doctoral program as an external candidate. At the time, we had just started our relationship. We were on track three of the Neuchâtel train station; it was raining, and she said, “Do not do it; we are moving to the US. You should look forward, not backward.” After some disagreements, she became my strongest supporter. She is not only of the co-authors of the thesis and the person who double-checked virtually all the thesis’s coding but also the love of my life. My family, including my parents Monica and Marco and my two sisters Marcella and Virginia, was incredibly supportive, enduring reduced time with me during Christmas breaks to allow me to attend the Doktoranden Seminars in Bad Herrenalb. Among my friends, I would like to thank Davide, Valentina, Paul, Edoardo, Francesco, and Carlo. Thank you so much.

Giacomo Benini, 6th January 2025, Bergen, Norway.

Abstract

Global pledges to mitigate climate change continue to gain momentum. The 26th Conference of the Parties to the United Nations Framework Convention on Climate Change and the third meeting of the Parties to the Paris Agreement have committed the international agenda to a sharp reduction in greenhouse gas emissions.

The oil and gas industry is a key player in the quest to reduce greenhouse gas emissions. According to the World Energy Outlook by the International Energy Agency, of the 48.9 billion metric tonnes of carbon dioxide equivalent emitted each year, 18.82 billion tonnes come from the oil and gas sector. Of these, 73% are due to the combustion of refined products like jet fuel, fuel oil, naphtha, gasoline, diesel, dry natural gas, and mixed liquefied petroleum gas. The remaining 27% occur before combustion, during the exploration, extraction, transportation, and processing of petroleum liquids and raw natural gas.

These two types of emissions differ significantly. While combustion emissions are relatively predictable regardless of the engine's efficiency, emissions from extracting and processing oil can vary greatly. For example, extracting high-viscosity oil from an oilfield that is well-connected to the natural gas pipeline system and sends its crude to a nearby refinery powered by solar energy is vastly different from mining bitumen, upgrading it on-site, and sending it to a refinery that does not capture the methane mixed with the crude.

The aim of this PhD thesis is to study the origin of up- and midstream emissions, link them to the economic decisions of different producers, and explore how to minimize them.

The PhD contains three studies. The first one evaluates the carbon intensity of crude oil production from marginal fields, which are smaller, less productive, and often economically challenged. Utilizing life cycle assessment methodologies, it highlights the disproportionately high greenhouse gas emissions associated with these fields due to inefficiencies in extraction and processing. The research underscores the need for targeted policies and technologies to mitigate emissions from marginal oil production, emphasizing the potential benefits of adopting best practices and advanced technologies.

The second study expands on the life cycle analysis framework by integrating economic and environmental assessments within the context of the Hotelling theorem's applicability to the oil industry. It assesses how the discovery-depletion process alters emissions profiles across producers, challenging the traditional assumption that more reserves imply lower emissions. The article provides a nuanced perspective, creating a spectrum that shows how natural pressure variances affect the relationship between extraction costs and emissions of an oilfield.

The first two papers suggest that the oil and gas industry is both similar to and different from other sectors that need to decarbonize. On one hand, its emissions are heavily linked to the quantity of energy used on-site to extract oil - the more energy-intensive the procedure, the higher the carbon intensity of the extraction process. On the other hand, there is an aspect specific to the oil industry: some oilfields consume a small quantity of energy but have a high carbon footprint due to an associated commodity channel. Hydrocarbon reservoirs contain a mixture of oil and natural gas. In many regions, the cost of capturing, purifying, and transporting natural gas exceeds its market price or the opportunity cost of reusing it on-site. Therefore, the natural gas is often either flared, releasing carbon dioxide (and methane), or vented directly into the atmosphere, releasing methane.

The third study describes a fiscal policy designed to eliminate flaring and venting in order to make the emissions of the oil and gas industry no different from those of any other energy-intensive sector. This research proposes a novel, revenue-neutral tax reform aimed at eliminating routine flaring and venting at zero net societal cost. By adjusting the tax on oil production in proportion to the reservoir gas-oil ratio and reducing the tax on natural gas sales, the proposed policy ensures that firms have no financial incentive to flare or vent the co-extracted natural gas. Instead, it encourages the capture and utilization of associated gas, thereby reducing greenhouse gas emissions without compromising the profitability of oil operations.

The combined insights from the first two studies provide a comprehensive framework for identifying who emits what and how profitable these emissions are. The third article proposes a solution to the type of emissions

that make the oil and gas industry different from any other. The proposed interventions not only reduce the carbon intensity of crude oil production but also enhance the industry's overall contribution to climate change mitigation. Policymakers, industry stakeholders, and researchers can leverage these findings to develop and implement strategies that promote cleaner, more efficient oil and gas production practices, ultimately supporting global efforts to meet climate targets.

Contents

Part I

Introduction

Chapter 1

Overview

1.1 Motivation

The global energy system is transforming to reduce greenhouse gas emissions per unit of physical work (heat) transferred. Central to this transition is replacing the three primary fossil fuels — coal, oil, and natural gas — with energy sources that do not emit carbon dioxide, methane, or nitrous oxide during conversion from chemical to kinetic energy.

Most techno-economic studies analyse this transformation through one of two approaches. The first begins with a predefined greenhouse gas reduction target and calculates the corresponding decrease in fossil fuel consumption needed to meet this target [McGlade and Ekins, 2015]. The second approach models a state-of-the-world scenario, incorporating projections on demographics, technological advancements, and policy assumptions, which then determine potential reductions in coal, oil, and natural gas use [Brandt et al., 2018].

However, both approaches generally assume the substitution of a representative quantity of fossil fuels. This simplification works well for the downstream segment of the supply chain, where refined products — such as lignite, jet fuel, fuel oil, gasoline, diesel, and pipeline-quality natural gas — produce relatively consistent emissions profiles upon combustion. Emissions variability in downstream phases is minor, driven only by fuel heterogeneity and combustion efficiency, making emissions per unit of energy relatively predictable and generalizable [Chang et al., 2004].

In contrast, this representative fuel assumption is less effective for up- and midstream emissions. The emissions profiles of upstream and midstream stages — including exploration, extraction, initial processing, transportation, and refining — vary widely due to factors like geological differences, extraction methods, field-specific characteristics, management practices, and refinery configurations [Masnadi et al., 2018]. For example, oil sands and heavy crude oils generally produce higher emissions during extraction than lighter crudes due to energy-intensive extraction processes. Similarly, hydraulic fracturing combined with horizontal drilling can lead to high emissions when co-extracted natural gas is mismanaged. Ignoring these variabilities can lead to underestimations or overestimations in emissions, thereby compromising the accuracy of environmental assessments in techno-economic models.

Addressing the variability in up- and midstream emissions is essential for developing effective policies and advancing a sustainable energy transition. By accounting for these variations, techno-economic models can provide a more accurate assessment of the environmental impacts across the oil supply chain. This thesis is driven by the need to enhance these models, bridging gaps in current approaches and improving our understanding of the diverse emissions profiles at each stage.

1.2 Objectives and Research Questions

The thesis pursues two intertwined objectives. The first objective is to comprehensively study the relationship between the profitability of the oil industry and its up- and midstream carbon intensity. This involves understanding how carbon emissions, particularly methane, impact the financial performance of oil and gas firms and identifying the key factors that influence this relationship. The second objective is to design and evaluate policy measures that can effectively reduce the carbon intensity of the oil industry without compromising its profitability. This involves developing innovative fiscal frameworks that address the specific challenges of the industry carbon and methane management.

To achieve its objectives, the thesis must answer the following research questions:

1. What is the relationship between the profitability of oil and gas firms and their carbon intensity?
   - This question seeks to explore the direct and indirect impacts of carbon emissions on the financial performance of oil and gas companies. It aims to identify the key drivers of carbon intensity in the industry and how these factors influence profitability.
2. Given the link between profitability and carbon intensity, what are the key challenges to implement effective carbon management policies in the oil and gas industry?
   - This question aims to identify the main obstacles to the successful implementation of carbon management policies in the oil and gas sector. By understanding these challenges, the research provides practical recommendations for policymakers and industry stakeholders to enhance the effectiveness of carbon management initiatives.
3. Considering the key challenges in implementing effective carbon management policies, is it possible to design a policy that reduces the carbon intensity of the oil and gas industry?
   - This question focuses on the development and evaluation of policy frameworks that can mitigate carbon emissions, particularly methane, in the oil and gas industry. It seeks to identify innovative approaches to carbon management that are both economically viable and environmentally effective.
4. What are the potential economic and environmental benefits of implementing these policies?
   - This question seeks to quantify the potential environmental and economic benefits of the proposed carbon management policies evaluating their implications on greenhouse gas emissions and financial performances.

The four research questions shape the structure of the thesis. The first step is to understand the connection between profitability and carbon intensity, followed by an analysis of the strength of this link. Next, policies are designed to break this link, enabling emissions reduction without compromising profitability. Finally, the economic and environmental impacts of these policies are measured.

1.3 Structure of the Thesis

The thesis is cumulative, divided into three main parts. **Part I** introduces the work in three chapters. Chapter 1 provides an overview of the motivation, scientific goals, and structure of the thesis. Chapter 2 reviews historical and current research in oil economics, focusing on how economists have modelled the industry. Chapter 3 outlines the economic and environmental tools used to capture heterogeneity in the up- and midstream sectors of the industry.

Part II comprises three research papers:

PAPER 1 "*Carbon Implications of Marginal Oils from Market-Derived Demand Shocks*" explores the variability in carbon intensity across global oil production and its sensitivity to demand changes. Traditional life cycle assessments (LCAs) often assume average crude displacement, overlooking the emissions diversity of specific crudes. This study links an econometric model of production profitability with data on the carbon intensity of nearly 1,933 oilfields (around 90% of 2015 world supply) to assess responses to demand reductions. Findings indicate that smaller demand shocks (2.5% reduction) typically displace heavier crudes with carbon intensities 25-54% above the global average, though this bias lessens with larger shocks, especially when producers with market power coordinate their responses. Emissions reduction benefits are closely tied to both demand magnitude and market structure.

PAPER 2 "*The Economic and Environmental Consequences of the Petroleum Industry Extensive Margin*" analyses the economic and environmental impacts of extraction costs based on microeconomic foundations. Unlike Paper 1, this research models extraction costs by viewing each oilfield as a continuum of wells, defined by three factors: natural pressure, depletion rate, and operational status. Resulting costs are influenced by volumes extracted, reserves, and geological factors affecting the substitutability between natural and artificial pressure. The study reveals that understanding the economic and environmental impacts of oil production requires recognizing extraction cost variability due to differences in discoveries and depletion — a factor often underrepresented in existing literature.

PAPER 3 "*A Zero-Cost Policy to Eliminate Methane Emissions from the Oil and Gas Industry*" examines the most challenging emissions issue in the oil and gas sector — the one resulting from the mismanagement of co-extracted natural gas through flaring and venting. The paper argues that traditional policies are inadequate for addressing this specific type of emission because they create two unintended substitution effects. First, flaring becomes relatively more expensive than venting. Second, the increased cost of voluntary gas disposal diminishes companies' returns on investments in equipment maintenance and leakage detection. Both channels increase the industry methane footprint. The paper proposes a tax reform, which solves both problems incentivizing companies to capture and utilize natural gas rather than wasting it. By adjusting taxes on oil production and natural gas sales, the paper demonstrates that it is possible to eliminate flaring and venting without adversely affecting industry profits, consumer incomes, or government revenue. The paper estimates that in the United States, raising the average oil tax by $12.37 per barrel while simultaneously reducing the natural gas tax by $9.60 per barrel of oil equivalent could have eliminated 70% of methane emissions from the oil and gas industry between 2005 and 2020.

In summary, Paper 1 explores emissions reductions through demand variations, while Paper 2 employs a micro-level model to analyse the impact of geophysical and operational factors on extraction costs and environmental impact. Together, they provide a comprehensive understanding of the industry's emissions, highlighting both the energy intensity of extraction and the unique emissions challenges posed by natural gas co-production. Paper 3 offers a practical solution to the second challenge through a policy reform, encouraging the capture and use of co-extracted natural gas and effectively reducing flaring and venting emissions without financial drawbacks.

The following list provides the full citations for each paper:

PAPER 1 Masnadi, Mohammad S.; Benini, Giacomo; El-Houjeiri, Hassan M.; Milivinti, Alice; Anderson, James E.; Wallington, Timothy J.; De Kleine, Robert; Dotti, Valerio; Jochem, Patrick; Brandt, Adam R. (2021): Carbon implications of marginal oils from market-derived demand shocks. In Nature 599 (7883), pp. 80–86. DOI: 10.1038/s41586-021-03932-2.

PAPER 2 Benini, Giacomo; Brandt, Adam; Dotti, Valerio; El-Houjeiri, Hassan (2023): The Economic and Environmental Consequences of the Petroleum Industry Extensive Margin. Department of Economics, Ca' Foscari University of Venice, Working Paper Series No. 14/WP/2023. DOI: 10.2139/ssrn.4556513.

PAPER 3 Benini, Giacomo; Dotti, Valerio; Berentsen, Geir Drage; Otneim, Håkon; Jahnke, Eric; Schuhmacher, Johannes; El-Houjeiri, Hassan M.; Ardone, Armin; Fichtner, Wolf; Jochem, Patrick; Gordon, Deborah; Brandt, Adam R.; Masnadi, Mohammad S. (2025): A Zero-Cost Policy for Eliminating Methane Emissions in the Oil and Gas Industry. To be submitted to a scientific journal.

Part III presents the thesis conclusions, discussing outcomes, limitations, and proposing a research agenda for unexplored areas.

Chapter 2

Background

2.1 Global Oil Value Chains, Emissions, and Mitigation Strategies

2.1.1 New Crudes, Old Powers: The Changing Dynamics of the Oil Market

Oil has long been the backbone of the global energy system [Yergin, 1991]. Its dominance as a primary energy carrier stems from its versatility, high energy density, and the relatively low cost of production and transportation compared to other energy sources [Smith and Taylor, 2018]. Currently, oil accounts for over 30% of global primary energy consumption, surpassing coal, natural gas, and renewables [BP Energy, 2022]. Its influence is particularly evident in the transportation, industrial, and petrochemical sectors [IEA, 2021b].

Historically, oil's rise as the world's leading energy carrier began in the mid-19th century with the development of kerosene as a safer, cleaner alternative to whale oil for lighting [Yergin, 1991]. The first commercial oil well, drilled in 1859 in Titusville, Pennsylvania, marked the beginning of the modern oil industry [Williamson and Daum, 1959]. Early oil production was largely focused on meeting lighting demands, but this changed dramatically with the advent of the internal combustion engine in the late 19th century. The mass production of automobiles created a voracious demand for gasoline, a byproduct of crude oil refining. By the early 20th century, oil had overtaken coal as the primary fuel for transportation, cementing its role as a cornerstone of industrial progress and mobility [BP Energy, 2022].

World War I and World War II underscored the strategic importance of oil, as nations with access to reliable oil supplies gained critical advantages in fueling military vehicles, ships, and aircraft [Black, 1998]. The interwar period saw the establishment of major oil companies, often referred to as the "Seven Sisters" (Standard Oil of New Jersey, Royal Dutch Shell, Anglo-Iranian Oil Company, Standard Oil of New York, Standard Oil of California, Gulf Oil, and Texaco), which dominated global production and trade [Samuel, 1988].

The post-war economic boom further solidified oil's dominance as nations rebuilt infrastructure and expanded industrial output. The rapid growth of the automotive industry, suburbanization, and the construction of highways in North America, Western Europe, and Japan led to a surge in oil consumption [Yergin, 1991]. At the same time, the discovery of vast reserves in the Middle East transformed the region into the epicenter of global oil production [Mahdi, 2012]. Countries like Saudi Arabia, Iraq, and Iran became critical suppliers, with their low-cost production giving them a competitive edge. The establishment of the Organization of Petroleum Exporting Countries (OPEC) in 1960 marked a turning point, as oil-producing nations began to assert greater control over pricing and production, challenging the dominance of the Seven Sisters [OPEC, 2005].

The oil shocks of the 1970s, fueled by geopolitical tensions and OPEC embargoes, underscored the vulnerabilities of a global economy heavily dependent on affordable oil from the Middle East [Hamilton, 2013]. These disruptions revealed an urgent need for secure and diversified energy sources, prompting a global rethinking of energy strategies [Adelman, 1993]. In an effort to reduce future supply risks and address the limitations of traditional energy sources, the industry developed innovative techniques that made it possible to exploit resources once deemed uneconomical or inaccessible [Maugeri, 2012]. This era of innovation paved the way for the extraction of unconventional resources, such as shale formations [Krauss, 2012], oil sands [Choquette-Levy et al., 2005], and ultra-deepwater reservoirs [Campbell, 2011], fundamentally transforming the energy landscape.

These "new crudes" differ from traditional oil in their physical and chemical characteristics, presenting challenges as well as opportunities. Traditional oil, primarily light crude with high API gravity and low sulfur content, is relatively easy to extract and refine, yielding premium products like gasoline and diesel with minimal processing. In contrast, "unconventional sources," such as shale formations, oil sands, and ultra-deepwater reserves, are often heavier, with lower API gravity and higher sulfur content, or involve additional complexities in extraction and processing. For instance, shale oil, although typically light, may contain variable impurities depending on the geological formation, complicating its extraction and refining processes [Krauss, 2012]. Oil sands produce bitumen, an extremely heavy crude that requires extensive upgrading and refining before it can be processed [Choquette-Levy et al., 2005]. Similarly, ultra-deepwater reserves often yield medium to heavy crudes, extracted from remote and challenging locations that require advanced technological and logistical solutions [Campbell, 2011]. Each of these unconventional resources required a distinct technological revolution, with three key advancements standing out: the development of hydraulic fracturing and horizontal drilling for shale oil and gas, the implementation of Steam-Assisted Gravity Drainage (SAGD) for extracting bitumen and extra-heavy crudes, and the adoption of Enhanced Oil Recovery (EOR) techniques to maximize output from mature reservoirs. These tailored innovations were designed to overcome the unique extraction and processing challenges posed by each resource, collectively driving transformative progress across the oil industry.

The first revolution, known as the shale revolution, was driven by advancements in horizontal drilling and hydraulic fracturing, which transformed access to hydrocarbons trapped in low-permeability shale formations. Horizontal drilling, a technique where the wellbore is drilled vertically to the desired depth and then deviated to run parallel to the shale layer, significantly increases the contact area with the reservoir, enhancing production potential [Montgomery and Smith, 2010]. Hydraulic fracturing, or "fracking," involves injecting a high-pressure fluid mixture of water, proppant (usually sand), and chemical additives into the formation to create and maintain fractures in the rock [King, 2010]. These fractures bypass the low-permeability matrix, providing pathways for hydrocarbons to flow toward the wellbore. The process is carefully designed, with proppants ensuring that fractures remain open under subsurface pressure, while chemical additives are used to optimize viscosity, reduce friction, and prevent damage to the formation [Veatch, 1983]. Initial breakthroughs in Texas's Barnett Shale demonstrated the synergy between these technologies, showing the economic feasibility of extracting light oil from impermeable rocks [Mayerhofer et al., 2010]. This success catalyzed widespread adoption across North America, transforming major shale basins like Bakken, Eagle Ford, and Marcellus into prolific producers [Wang et al., 2014].

The second revolution was Steam-Assisted Gravity Drainage (SAGD), a thermal recovery technique that revolutionized the extraction of extra-heavy crude oil and bitumen, particularly in Canada's oil sands. The process involves drilling two horizontal wells: an upper injection well and a lower production well, typically spaced 5–10 meters apart [Butler, 1991]. Steam is injected into the upper well, heating the surrounding bitumen to reduce its viscosity. This creates a steam chamber, allowing the bitumen to flow via gravity drainage into the lower production well, from where it is pumped to the surface [Butler, 1994]. Additionally, the low API gravity and high sulfur content of bitumen require upgrading processes like hydrocracking and coking to produce synthetic crude oil, further increasing energy consumption and costs [Brandt, 2008]. Despite these challenges, SAGD has unlocked vast bitumen reserves and continues to evolve with innovations such as co-injection of solvents with steam to improve its efficiency [Nasr and Ayodele, 2002].

The third revolution, Enhanced Oil Recovery (EOR), comprises a suite of advanced methods designed to extract additional oil from reservoirs that are otherwise difficult or uneconomical to exploit, often achieving recovery rates of 60% or more. These techniques alter the physical or chemical properties of the oil or reservoir to improve flow and displacement efficiency [Butler, 1991, Brandt, 2008]. EOR methods are categorized into thermal, chemical, and gas-based approaches. Thermal techniques, such as steam injection and in-situ combustion, are widely used in heavy oil fields and involve injecting heat to reduce oil viscosity [Jordaan, 2009]. Chemical EOR employs polymers, surfactants, or alkaline solutions to modify reservoir wettability and reduce interfacial tension, enhancing oil displacement [Nasr and Ayodele, 2002]. Gas-based methods, including carbon dioxide injection, increase reservoir pressure and improve oil mobility, with carbon dioxide injection offering the additional benefit of sequestering greenhouse gases [Brandt, 2008]. Hybrid approaches, such as steam-alternating-gas (SAG), combine techniques to maximize recovery efficiency. These innovations have extended the productive lifespan of oil fields while addressing the challenges of unconventional reservoirs like heavy and extra-heavy crude. Despite their success, EOR remains energy-intensive, and ongoing research aims to improve efficiency, ensuring its role as a cornerstone of modern petroleum engineering [Mayerhofer et al., 2010].

Figure 2.1 illustrates the connection between the three technological revolutions — shale, SAGD, and EOR — and the pivotal events that laid the groundwork for their emergence. While early milestones, such as the drilling of the first oil well in 1859 and the widespread adoption of kerosene in the 1860s, highlighted the industry's dependence on innovation to meet growing energy demands, it was the two oil shocks of the 1970s that fundamentally reshaped the global energy landscape. These crises, driven by geopolitical tensions and OPEC embargoes, underscored the vulnerabilities of relying on conventional oil sources and spurred the search for alternatives. The dramatic increase in oil prices provided the economic incentives and urgency to develop technologies capable of accessing previously uneconomical or inaccessible resources. As a result, the 1970s set the stage for the shale, SAGD, and EOR revolutions, which transformed the industry's ability to meet energy demands while navigating a more complex and constrained global market.

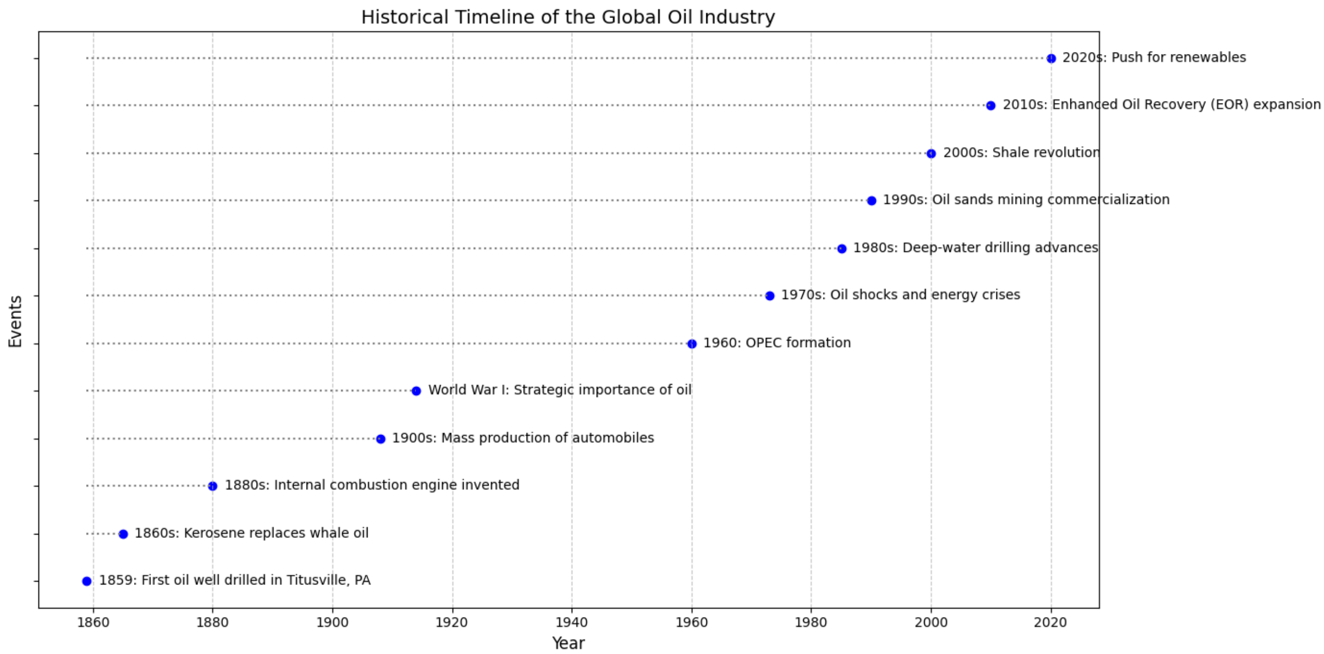

Figure 2.1: Time line of major events in the oil industry [Yergin, 1991].

As a result, non-OPEC countries significantly increased their production and capacity, fundamentally altering the global energy landscape. These innovations allowed the exploitation of resources that were previously uneconomical or technically inaccessible, shifting a substantial share of global production away from OPEC's traditional strongholds. This redistribution of capacity posed a direct challenge to OPEC's market dominance, prompting its leading members, particularly Saudi Arabia, to adapt their strategy. In 2016, Saudi Arabia spearheaded the creation of the OPEC+ alliance, incorporating Russia and other key non-OPEC producers into a coordinated framework, as shown in Figure 2.2. The OPEC+ alliance currently includes the 13 OPEC members —

Algeria, Angola, Congo, Equatorial Guinea, Gabon, Iran, Iraq, Kuwait, Libya, Nigeria, Saudi Arabia, the United Arab Emirates, and Venezuela — and 10 additional non-OPEC members: Russia, Azerbaijan, Bahrain, Brunei, Kazakhstan, Malaysia, Mexico, Oman, South Sudan, and Sudan [OPEC, 2021b]. Together, these countries account for over 50% of global oil production and nearly 90% of proven oil reserves, making them a dominant force in the energy market [OPEC, 2021a]. Through coordinated production quotas and joint decision-making, OPEC+ members aim to balance supply and demand, prevent extreme price fluctuations, and ensure long-term market stability. By adjusting output levels in response to global economic trends, geopolitical events, and shifts in energy demand, the alliance plays a critical role in influencing global oil prices and maintaining the profitability of oil-producing nations.

The interplay between the three technological revolutions and the strategic decision to create OPEC+ has created a distinctly divided global oil market. This market is broadly categorized into four macro-groups: 1) “new oils” from unconventional reserves in non-OPEC+ countries, 2) “old oils” from conventional reserves in non-OPEC+ countries, 3) “new oils” from unconventional reserves in OPEC+ countries, and 4) “old oils” from conventional reserves in OPEC+ countries. Among these, the most dominant are unconventional oil from non-OPEC+ countries and conventional oil from OPEC+ countries. The profitability of the former, despite its higher production costs, hinges on both technological advancements and OPEC+’s deliberate production strategies. By intentionally restricting the output of their abundant, low-cost conventional reserves, OPEC+ nations sustain higher global oil prices than would prevail in a perfectly competitive market. This calculated underproduction indirectly supports the economic viability of unconventional oil producers. The interdependence between OPEC+ underproduction and profitability of unconventional oils underscores the delicate balance that defines the modern oil market. OPEC+ leverages its reserve dominance to influence global prices, while its restraint allows unconventional oil producers to compete effectively. Together, these forces illustrate how technological innovation and strategic policymaking interact to shape global energy supply and demand, driving the evolution of the modern energy landscape. As the market continues to diversify, the nature of the crude oil being produced — whether from “old oils” or “new oils” — further challenges downstream operations, particularly in refining, where the attributes of the crude oil dictate the complexity and energy intensity of the processes involved.

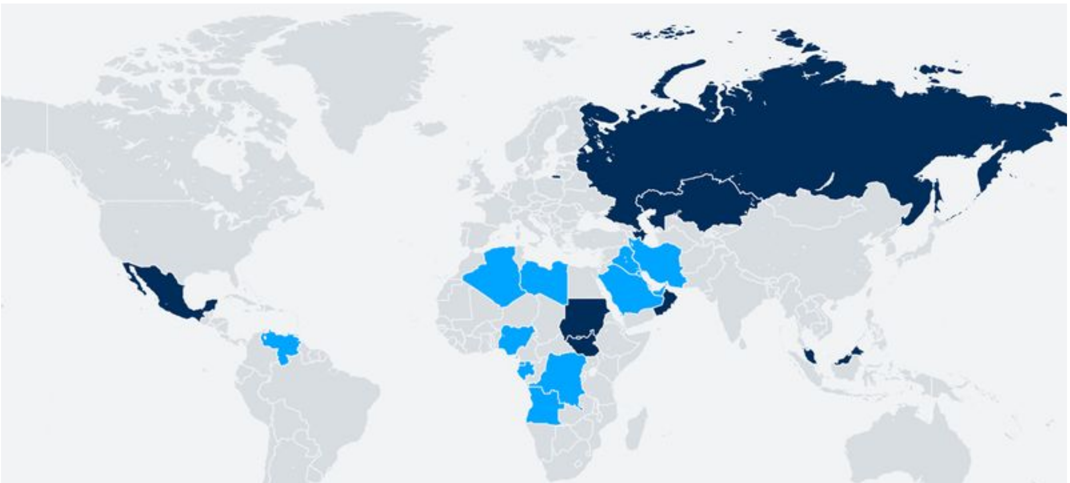

Figure 2.2: OPEC (light blue) and OPEC+ (dark blue) members [DW and IEA, 2022].

2.1.2 Petroleum Refining

Oil refining transforms crude oil, a complex mixture of hydrocarbons, into usable products such as gasoline, diesel, jet fuel, and petrochemical feedstocks [Jones and Patel, 2019]. The transformation process is profoundly influenced by the characteristics of the crude oil entering the refinery; whether derived from traditional sources ("old oils") or unconventional reserves ("new oils"), these attributes significantly determine the complexity and energy requirements of refining operations [McCarthy and Singh, 2016].

Traditional crude oils are typically light with high API gravity and low sulfur content. These "old oils" are easier and less energy-intensive to refine, yielding higher proportions of premium products like gasoline and diesel with minimal processing [Meyers, 2004]. In contrast, "new oils" from unconventional sources, including oil sands, shale oil, and ultra-deepwater reserves, often deviate significantly from these characteristics [Liang and Chen, 2017]. Heavy crudes like bitumen from oil sands have low API gravity and high sulfur content, necessitating extensive upgrading and more energy-intensive refining processes [Gray, 2002]. Similarly, shale oil, though typically lighter, may contain impurities that complicate refining. These differences in feedstock characteristics require distinct approaches within refineries to optimize product yields and maintain environmental compliance [Meyers, 2004].

Adapting to these challenges has led to the evolution of sophisticated refining processes that account for the variability in crude properties. Regardless of whether the feedstock originates from "old oils" or "new oils," the transformation of crude oil into refined products relies on three primary stages, with refiners continuously innovating to maximize efficiency and address the complexities posed by unconventional crudes:

- Separation (fractional distillation): The first step in refining is the separation of crude oil into its constituent components through fractional distillation [Watkinson and Wilson, 2012]. This process leverages the differing boiling points of hydrocarbons to separate them into fractions. The crude oil is first heated in a furnace to temperatures as high as 350°C, causing it to vaporize [Parkash, 2003]. The vaporized oil is then introduced into a distillation tower, where lighter components, such as propane and butane, rise to the top while heavier components, like diesel and residual fuel oils, condense at lower levels of the tower. At the end of the distillation process, a variety of fractions are produced, ranging from gases at the top to heavy residues at the bottom [Treese and Treese, 2015]. These fractions form the basis for further refining and upgrading steps.

- Conversion (upgrading to valuable products): After separation, the heavier and less valuable fractions undergo conversion processes to transform them into lighter, more commercially valuable products [Speight, 2011]. This stage is particularly important for processing heavier crude oils that contain a larger proportion of heavy hydrocarbons [Gary et al., 2007]. One of the most critical conversion processes is cracking, which breaks down large hydrocarbon molecules into smaller ones [Weissman and Klein, 1985]. Thermal cracking relies on heat and pressure to achieve this, while catalytic cracking uses catalysts to reduce energy requirements and improve efficiency [Sadeghbeigi, 2012]. Another key process is hydrocracking, where heavy hydrocarbons are combined with hydrogen to produce lighter products. For the heaviest residues, coking is often employed, resulting in solid petroleum coke and lighter hydrocarbon fractions [Weissman and Klein, 1985, Sadeghbeigi, 2012]. These conversion techniques ensure that refineries maximize the yield of desirable products like gasoline and diesel, even from heavy or unconventional crude oils.

- Treatment (ensuring quality and environmental compliance): The final stage of refining focuses on removing impurities and blending products to meet market specifications and environmental regulations. Sulfur, a common impurity in crude oil, is removed through desulfurization processes to minimize sulfur dioxide emissions during combustion [Lefebvre and Deschenes, 1997]. Similarly, nitrogen and metal impurities are extracted to improve product quality and protect refinery equipment. Once the impurities have been removed, it is possible to blend different fractions to create finished products that meet specific standards, such as the octane rating for gasoline or the cetane number for diesel [Speight, 2004]. This final step ensures that the refined products are not only functional but also compliant with safety and performance requirements.

The complexity and adaptability of refining processes allow them to transform crude oil, regardless of its origin, into a wide array of refined products that power modern life and industry. These products, ranging from light hydrocarbons to heavy residues, are characterized by their varying carbon chain lengths, which influence their properties, applications, and production methods. The lightest products are Liquefied Petroleum Gas (LPG) and ethane, primarily composed of light hydrocarbons such as propane, butane, and ethane. They are extracted during fractional distillation and natural gas processing. LPG serves as a widely used cooking fuel and a key feedstock for petrochemical production, while ethane is predominantly utilized in ethylene production through cracking. Gasoline, with hydrocarbons typically in the C5–C12 range, is a cornerstone of global energy demand, primarily powering passenger vehicles. Catalytic cracking and reforming processes enhance its octane rating, improving combustion efficiency and reducing engine knocking. Additives such as oxygenates and detergents further optimize performance and minimize emissions, ensuring gasoline remains a dominant fuel in global markets. Naphtha, another fraction in the C5–C12 range, plays a crucial role as a precursor in petrochemical production. It undergoes processes like steam cracking to yield ethylene, propylene, and aromatics, which are foundational for the synthesis of plastics, synthetic rubber, and various industrial chemicals. Naphtha exemplifies the refinery's dual role in producing both energy carriers and raw materials for industry. Jet fuel/kerosene, a middle distillate fraction with a carbon chain length of approximately C10–C16, is valued for its high energy density and thermal stability. These properties make it indispensable for aviation, where consistent combustion and resistance to extreme temperatures are critical [Treese and Treese, 2015]. The production of jet fuel involves hydrocracking and hydrotreating to ensure low sulfur content and compliance with stringent aviation standards [Clark, 2004]. Diesel/gasoil, another middle distillate, contains hydrocarbons in the C12–C20 range. Known for its high energy density and efficiency in compression-ignition engines, diesel is integral to freight transportation, agriculture, and industrial machinery. Hydrotreating removes sulfur and improves cetane numbers, ensuring cleaner combustion and regulatory compliance [Ho and Smith, 2019]. Fuel oil, derived from heavier fractions with hydrocarbon chains of C20 and above, is traditionally used in marine transportation and industrial heating. Its high viscosity requires preheating for efficient combustion [Farina, 2011]. Modern desulfurization technologies produce low-sulfur fuel oil (LSFO) to meet strict international maritime standards [IMO, 2020]. Lastly, specialized products such as lubricants, asphalt, and waxes are derived from the heaviest fractions of crude oil. Lubricants undergo hydroprocessing to enhance purity and thermal stability, making them indispensable for reducing friction and wear in machinery. Asphalt, valued for its high viscosity and adhesive properties, is used extensively in road construction and roofing, demonstrating the versatility of refining operations to meet diverse industrial needs.

The diverse range of refined products, from light hydrocarbons like LPG to heavy residues like asphalt, highlights the complexity of modern refineries and their ability to meet the demands of global markets. However, the journey of these products does not end at the refinery gate. Ensuring that crude oil reaches refineries and that refined products are delivered to final consumers requires an extensive and efficient transportation network.

2.1.3 Transportation

Oil transportation, both from oilfields to refineries and from refineries to final consumers, is a vital component of the global energy supply chain [Yergin, 1991]. The primary modes of transportation include pipeline networks, railcars, trucks, and tanker vessels [Smith and Brown, 2018]. Pipelines are the most efficient and widely used method for transporting large volumes of crude oil over land. These extensive networks connect oilfields to refineries, often spanning thousands of kilometers, as seen in regions like North America, where the Keystone Pipeline plays a crucial role [Smith and Taylor, 2021]. However, when pipelines are unavailable or insufficient, railcars provide a flexible alternative, allowing oil producers to transport crude across regions [Harris, 2019]. While rail is less cost-effective and poses safety risks, it offers access to remote areas and smaller refineries not connected to pipeline infrastructure. Trucks are typically used for short-distance transport from smaller oilfields to nearby storage facilities or pipeline terminals, offering flexibility but at higher costs and lower capacities. Tanker vessels, on the other hand, dominate the transportation of crude oil across oceans, connecting major oil-producing regions like the Middle East to global markets [Lee and Martin, 2020]. These massive ships, such as Very Large

Crude Carriers (VLCCs), can carry millions of barrels of oil, making them an essential link in international trade.

Once refined into products such as gasoline, diesel, and jet fuel, transportation networks ensure these products reach final consumers. Pipelines remain critical for moving large volumes of refined products to distribution hubs, where they are transferred to railcars, trucks, or barges for further distribution [Smith and Johnson, 2020]. Trucks play a pivotal role in delivering fuel to gas stations, airports, and industrial consumers, offering the flexibility needed for last-mile delivery [Harris and Taylor, 2019]. Rail is also used to transport refined products over long distances where pipelines are not feasible, such as in regions with dispersed markets or challenging terrain. For international distribution, tanker vessels transport refined products to distant markets, utilizing specialized ships for specific products, such as LPG carriers or clean product tankers [Lee and Martin, 2021]. Together, these transportation modes form an interconnected network, enabling the seamless flow of crude oil and refined products from production sites to end-users, ensuring a stable supply in the global energy market [Cruz and Peterson, 2013].

2.1.4 Up, Mid, and Downstream Emissions

The life-cycle emissions of the oil value chains described in the previous section can be broadly divided into three stages: upstream, midstream, and downstream. Upstream emissions are generated during the extraction and initial processing of crude oil. These activities include drilling, pumping, and handling associated natural gas. These methods demand substantial energy inputs, often derived from fossil fuels, resulting in high greenhouse gas emissions. Midstream emissions encompass both the transportation and refining of crude oil. Transportation emissions arise from activities such as pipeline operations, shipping, and terminal handling, with energy-intensive modes like marine shipping contributing significantly to the carbon footprint. Refining emissions occur during the transformation of crude oil into usable products, such as gasoline, diesel, and jet fuel. This stage involves energy-intensive processes like distillation, catalytic cracking, and hydrotreating, which vary in intensity depending on the type of crude processed. Finally, downstream emissions are associated with the final combustion of refined products during end use, such as in vehicles, aircraft, and industrial machinery. As the largest contributor to life-cycle emissions, downstream activities highlight the critical role of fuel combustion in determining the carbon footprint of oil value chains. Understanding the distinct characteristics and contributions of each of these emissions is essential for addressing the environmental impact of the oil value chain, as each phase presents unique challenges for mitigation:

- Upstream Emissions: These emissions are mainly driven by on-site energy demand and the management of the co-extracted natural gas.
  - On-site Energy Demand: The advent of the three technological revolutions previously described — shale, SAGD, and EOR — has significantly diversified the demand for on-site energy. Traditional extraction methods for conventional oil (light and medium crudes) initially required relatively low energy inputs. Early in a reservoir's production life, natural pressure drives oil to the surface with minimal external energy. However, as reservoirs age and pressure declines, producers increasingly rely on secondary recovery techniques, such as water injection, to sustain production levels. This process involves injecting large volumes of water into the reservoir to displace oil and maintain flow rates. Over time, the energy intensity of these operations rises due to an increasing water cut — the proportion of water to oil in the produced fluid. The management of this produced water, including separation, reinjection, or disposal, further contributes to the energy demands and emissions associated with conventional oil production. The shale revolution, underpinned by horizontal drilling and hydraulic fracturing, has drastically increased global oil production from low-permeability formations. Horizontal drilling enables operators to access extensive reservoir areas by creating wells that turn horizontally through the formation. Hydraulic fracturing, or "fracking," involves injecting high-pressure water, sand, and chemicals into the formation to create fractures that allow hydrocarbons to flow to the

wellbore. While highly effective in unlocking unconventional reserves, these techniques are energy-intensive and often release significant volumes of associated natural gas, exacerbating greenhouse gas emissions. Similarly, SAGD involves drilling two horizontal wells — one for steam injection and the other for bitumen production — enabling the recovery of bitumen with minimal surface disturbance. The first of the two wells is highly energy-intensive since it requires substantial volumes of steam to reduce the viscosity of bitumen. This steam is predominantly generated through on-site natural gas combustion, resulting in substantial greenhouse gas emissions [Jordaan, 2009]. Finally, EOR techniques, including thermal, gas-based, and chemical methods, have revolutionized the recovery of oil from mature or declining reservoirs. Thermal EOR, such as steam injection and in-situ combustion, is particularly energy-intensive, as it relies on heat to mobilize trapped oil. Gas-based EOR, such as carbon dioxide injection, enhances recovery by increasing reservoir pressure or altering fluid properties while also offering the benefit of sequestering greenhouse gases underground. Chemical EOR uses surfactants or polymers to reduce interfacial tension between oil and water, improving displacement efficiency. While EOR extends the productive lifespan of reservoirs, it requires a careful balance between economic feasibility and the emissions associated with its energy-intensive processes [Keith, 2020].

  - Management of Co-Extracted Natural Gas: A significant contributor to upstream emissions across all extraction methods is the management of associated natural gas, which is often released as a by-product of oil extraction. In remote or offshore locations where capturing and utilizing this gas is technically challenging or economically unfeasible, it is typically flared or vented. Flaring, which burns the gas to convert methane into carbon dioxide, is less harmful than venting but still contributes significantly to greenhouse gas emissions, releasing over 400 million tons of carbon dioxide in 2021 alone [Allen, 2020]. On the other hand, venting, which directly emits methane into the atmosphere, poses an even greater environmental concern due to methane's global warming potential, which is 25 times higher than that of carbon dioxide over a 100-year period [IPCC, 2021]. Additionally, methane leakage during processes like hydraulic fracturing further exacerbates its environmental impact, highlighting the need for effective management strategies [Brandt, 2013].

  - Interplay between On-site Energy Use and Management of the Co-Extracted Natural Gas: The interplay between on-site energy use and the management of co-extracted natural gas determines almost the entirety of the upstream emissions. In some cases, this dynamic aligns, as seen in shale formations, where energy-intensive extraction methods are often accompanied by inadequate management of natural gas. This results in significant on-site demand for energy as well as high flaring and methane emissions. Conversely, in techniques like SAGD and EOR, the relationship between energy use and natural gas management diverges. While these methods are highly energy-intensive due to their reliance on steam generation, they primarily extract oil with minimal associated natural gas, leading to comparatively lower levels of flaring and methane emissions.

- Midstream Emissions: From a refining perspective, the relative carbon intensity of refined products is influenced by the energy requirements and chemical transformations involved in their production.

  - On-site Energy Demand: Different types of emissions arise at various stages of refining processes, depending on the nature of the crude and the desired products. For instance, producing lighter fractions like gasoline and jet fuel involves catalytic cracking and hydrocracking, processes that emit significant levels of carbon dioxide due to their reliance on high-temperature and high-pressure environments. Conversely, heavier products like fuel oil and asphalt generate more emissions during their refining because of their complex molecular structures and the need for extended thermal cracking or coking. Advanced refining techniques, such as desulfurization and hydroprocessing, aim to reduce impurities like sulfur and nitrogen, which not only improve product quality but also ensure compliance with stringent environmental regulations. However, these processes themselves are energy-intensive, leading to indirect greenhouse gas emissions due to the combustion of fossil fuels used to generate the required heat and hydrogen.

  – Management of Natural Gas: Flaring in petroleum refineries serves as a critical safety mechanism and operational necessity, used to burn off excess hydrocarbons released during various processes. Flare systems are designed to manage pressure build-ups, vent gases during maintenance, or safely dispose of unburned hydrocarbons in emergencies. The process involves combusting these gases in a controlled environment, converting volatile hydrocarbons into carbon dioxide and water. While flaring ensures safe operations by preventing the release of harmful gases like methane, it is also a significant source of greenhouse gas emissions. Refinery flaring systems are typically divided into two categories: elevated flares and ground flares. Elevated flares are more common and consist of a stack equipped with a flare tip where gases are burned at high altitudes to disperse emissions. Ground flares, on the other hand, use burners at ground level and are often enclosed to minimize visual and noise impacts. Both systems include components like flare headers to transport the gases, knockout drums to remove liquids, and steam or air assist systems to ensure complete combustion. Technological advancements, such as flare gas recovery systems, aim to mitigate these effects by capturing and reusing waste gases. These systems compress and process flare gases, converting them into valuable resources like fuel gas or feedstock for petrochemical processes, thereby reducing emissions and improving energy efficiency. Modern refineries are increasingly adopting these technologies to align with sustainability goals and regulatory requirements, highlighting the importance of innovation in addressing the environmental challenges of flaring.

- Downstream Emissions: Downstream emissions, arising from the combustion of refined petroleum products, vary significantly depending on the type of fuel and its specific use. Diesel and gasoil are among the highest contributors to emissions due to their widespread use in freight transportation, industrial processes, and marine applications. These fuels are prized for their energy density and efficiency in compression-ignition engines, but their combustion releases substantial amounts of carbon dioxide. Similarly, fuel oil, often used for heating and marine transport, is another major emitter because of its high carbon content and the large quantities typically consumed. Gasoline, the dominant fuel for passenger vehicles, also contributes significantly to downstream emissions. Its widespread use globally amplifies its impact, even though its carbon intensity is slightly lower than diesel. Jet fuel and kerosene, critical for aviation, are notable for their emissions as they provide the high energy density and thermal stability required for aircraft operation. LPG and ethane, on the other hand, are relatively lower-emission fuels. Widely used for cooking, heating, and as feedstock in the petrochemical industry, their cleaner combustion characteristics make them a less carbon-intensive option compared to heavier products. Naphtha, primarily utilized as a petrochemical feedstock, has relatively moderate emissions, as it is often not combusted directly for energy. Finally, refined products not intended for combustion, such as lubricants, asphalt, and waxes, constitute a diverse category with varied environmental implications. Although these products are not directly burned for energy, their production and application processes can still indirectly contribute to downstream emissions through energy-intensive manufacturing and usage practices.

The previous discussion on up, mid, and downstream emissions highlights how the shale, SAGD, and EOR revolutions have shaped the global oil market not only from an economic perspective but also from an environmental one. From an upstream perspective, “new oils” from unconventional reserves in non-OPEC countries are characterized by high on-site energy demand (as seen with SAGD and EOR) and relatively low flaring and venting emissions. In contrast, shale formations exhibit medium to high on-site energy demand but often suffer from poor management of co-extracted natural gas, leading to significant flaring and methane emissions. “Old oils” from conventional reserves, whether in non-OPEC+ or OPEC+ countries, display emissions profiles that depend heavily on the maturity of the oilfield. Younger fields with light and medium crudes tend to have low energy demand and, in some cases, benefit from substantial investments in natural gas infrastructure that mitigate flaring and venting. Countries like Saudi Arabia, Kuwait, and the United Arab Emirates exemplify this trend, having significantly reduced emissions through robust gas utilization strategies. Conversely, fields in countries such as Iraq and Iran often face high flaring emissions due to inadequate natural gas management infrastructure. Midstream emissions, primarily associated with refining, further amplify these distinctions. The refining of unconventional crudes such as those extracted via SAGD or EOR, and extra-heavy crudes like bitumen, is highly energy-intensive. These crudes often require extensive upgrading and additional refining processes, such as coking, hydrocracking, and hydrotreating, to produce marketable fuels. Such processes significantly increase greenhouse gas emissions compared to refining lighter, sweeter crudes. Shale oil, while lighter and less complex to refine, may still pose challenges due to impurities that necessitate additional processing steps, such as stabilization and desulfurization, further contributing to midstream emissions. Traditional light and medium crudes, particularly those from OPEC+ countries with advanced refining infrastructure, generally incur lower midstream emissions due to their simpler refining requirements. However, variations in refinery configurations and product slates mean that even conventional oils can have differing midstream carbon footprints.

2.1.5 Mitigation Strategies

The previous two sections explored the economic and environmental distinctions among different types of oil and oil producers, which underpin the four value chains previously described. Each segment of these value chains — upstream, midstream, and downstream — faces unique challenges and opportunities for emissions reduction. While upstream and midstream emissions can be mitigated through technological advancements, process optimization, and well-designed policies, downstream emissions, primarily resulting from the combustion of refined products, are intrinsically tied to oil consumption levels. Addressing downstream emissions would require a broader transition to alternative energy sources and a reduced reliance on oil — objectives beyond the scope of this thesis. Instead, the focus here is on actionable strategies to minimize emissions within the upstream and midstream segments while balancing environmental objectives with economic feasibility.

Achieving meaningful emissions reductions in these segments depends on two complementary approaches. First, technological advancements can enhance efficiency and lower carbon intensity within extraction, processing, and transport operations. Second, policy measures play a critical role in shaping industry behavior by creating incentives and regulatory frameworks that drive sustainable practices. Together, these strategies form the foundation for a practical and effective pathway to reducing emissions within the oil industry.

- Technological Solutions
  - Upstream One of the most significant contributors to emissions is the energy-intensive nature of oil extraction, particularly for unconventional resources such as shale formations, oil sands, and reservoirs utilizing Enhanced Oil Recovery (EOR) techniques. Implementing more efficient extraction technologies, such as electric submersible pumps and advanced steam injection systems, can substantially lower energy use and the associated carbon footprint. Additionally, shifting drilling operations to renewable or low-carbon electricity presents an opportunity for further reductions. Carbon Capture and Storage (CCS) technologies at production sites also offer a viable pathway to mitigating emissions, particularly in heavy and unconventional crude extraction, where the carbon intensity is highest. Another critical upstream challenge lies in the management of co-extracted natural gas. In many cases, technical and economic constraints lead to flaring or venting, both of which contribute significantly to greenhouse gas emissions. While flaring converts methane into carbon dioxide, reducing its immediate warming potential, it remains a major emissions source, whereas venting releases methane

directly, with an even greater climate impact. Advances in gas capture and utilization, such as improved reinjection techniques and micro-LNG systems, provide technological solutions to mitigate these emissions.

  - Midstream Oil refining is an inherently energy-intensive process that accounts for a substantial portion of industry-wide emissions. Optimizing refinery operations through enhanced heat recovery systems, advanced process control technologies, and cogeneration systems can reduce energy consumption and improve efficiency. Additionally, transitioning to low-carbon hydrogen for desulfurization and other refining processes further lowers emissions, while upgrading refineries to process lighter, less carbon-intensive crude oils can significantly reduce overall energy demand. At the same time, refinery flaring remains a significant emissions source, often used as a safety measure or for disposing of excess hydrocarbons. The adoption of flare gas recovery technologies, which allow excess gases to be repurposed as fuel or feedstock for petrochemical processes, offers a practical emissions reduction strategy while improving resource utilization.
  - Transportation Pipelines, the most efficient mode of transport, can achieve emission reductions through network optimization and the use of renewable-powered electric pumps. Rail and truck transport can transition to lower-carbon fuels such as biofuels and electricity, while marine transport emissions can be reduced through retrofitting tanker vessels with energy-efficient designs, adopting wind-assist technologies, and shifting to LNG as a transitional fuel. Short-distance deliveries, particularly those made by trucks, can achieve near-zero emissions through electrification. Collectively, integrating low-carbon technologies across transportation modes has the potential to reduce emissions by 20–30%, with further improvements anticipated as hydrogen fuel cells and electric vehicles become more widely adopted.

- Policy Solutions
  - Upstream The key policy challenge in upstream operations is regulating flaring and venting, which are major sources of emissions but difficult to control due to asymmetric information. Regulators struggle to accurately monitor firms' emissions and distinguish between flaring and venting, leading to unintended consequences when imposing restrictions. For example, strict flaring limits may push companies toward venting, which has an even higher climate impact. Moreover, firms may underreport emissions or adjust operations in ways that minimize compliance costs rather than actual emissions. Designing an effective policy to eliminate flaring and venting requires overcoming these informational barriers while ensuring firms have the right incentives to capture and utilize co-extracted gas instead of wasting it.
  - Midstream Unlike in upstream operations, asymmetric information is not a major concern in the midstream sector, making emissions regulation relatively straightforward. The key policy objective here is to eliminate refinery flaring, which can be effectively addressed through Pigouvian taxation. By imposing a tax based on the volume and carbon content of flared gas, refineries would face direct financial incentives to invest in flare gas recovery technologies and optimize their operations. Because emissions sources in refining are easier to monitor and measure than in upstream production, this taxation approach could achieve significant reductions without the risk of unintended substitution effects.
  - Transportation Financial incentives for biofuel infrastructure development, rail and truck electrification, and tanker vessel retrofits can accelerate the shift away from high-emission transport fuels. At the same time, investments in renewable-powered pipeline infrastructure and hydrogen-powered shipping can create long-term pathways toward emissions reductions.

Building on the insights from the previous sections, this thesis aims to develop a comprehensive and flexible modelling framework capable of simultaneously capturing the economic and environmental characteristics of the upstream, midstream, and transportation segments of the oil value chains. By focusing on actionable strategies for emissions reduction within these segments, the framework seeks to balance economic feasibility with environmental objectives. To lay the foundation for this effort, the thesis conducts a historical review of how the oil industry and its emissions have been modelled, providing critical insights into the strengths and limitations of existing approaches. This review not only informs the proposed framework but also ensures that it addresses the complex interplay between production processes, emissions profiles, and policy interventions.

2.2 Economic Modelling of the Oil Industry

The economic modelling of the oil and gas industry has significantly evolved since its early theoretical foundations. One of the earliest contributions can be traced back to 1914 when Gray published "Rent under the Assumption of Exhaustibility" in the *Quarterly Journal of Economics* [Gray, 1914]. This pioneering work aimed to model the production process of nonrenewable resources by calculating the economic rent obtainable from a finite resource. Gray's approach laid the groundwork for understanding how resource depletion impacts resource value over time.

In 1931, Hotelling further advanced the field with his seminal paper, "The Economics of Exhaustible Resources", which introduced a formal framework for analysing the optimal extraction of nonrenewable resources [Hotelling, 1931]. Hotelling's model defined the net price path of a resource as a function of time, aiming to maximize economic rent throughout the resource's depletion. The concept of Hotelling rent, or scarcity rent, became central to understand how the value of a resource should increase over time due to its growing scarcity. According to Hotelling, the efficient extraction of a resource requires the rate of increase in the net price to match the discount rate, thereby optimizing the present value of the resource over its extraction period. Consequently, resource owners will extract the resource only if it yields a higher return compared to other standard financial instruments; otherwise, the optimal choice is to leave the resource in the ground. This result is based on four key assumptions: 1) a finite stock (the resource is available in a fixed, limited quantity), 2) profit maximization (resource owners aim to maximize the present value of their profit from extraction), 3) perfect foresight (resource owners can accurately predict future prices and costs), and 4) no externalities (environmental impacts are not considered in the decision-making process). Within this framework, the nonrenewable resource should be allocated to equalize the marginal net benefit of extraction across different periods, adjusted for the rate of interest.

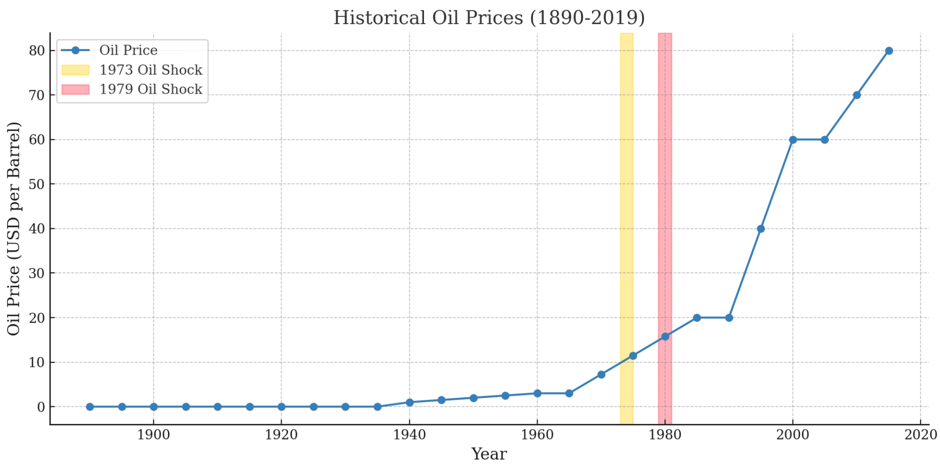

Figure 2.3: Time series of global oil prices [BP Energy, 2023].

Despite its theoretical robustness, Gray and Hotelling contributions remained largely an intellectual curiosity until the early 1970s. This period marked a turning point in the modelling of nonrenewable resources. During this decade, the world experienced two major oil shocks that underscored the vulnerability of the global energy system, see Figure 2.3. The first oil shock occurred in 1973, triggered by an oil embargo imposed by the Organization of Arab Petroleum Exporting Countries (OAPEC) in response to the Yom Kippur War. The embargo led to a dramatic increase in oil prices, causing economic turmoil in many Western countries and highlighting their reliance on fossil fuels for energy needs [Yergin, 1991]. The second oil shock occurred in 1979, following the Iranian Revolution, which disrupted global oil supplies and caused another surge in oil prices. In the same period, the Club of Rome published "The Limits to Growth", a report that used computer simulations to model the consequences of economic and population growth in a world with finite resources, see Figure 2.4 for oil projections. This publication sparked widespread debate and highlighted the potential for ecological collapse if growth trends continued at a rate comparable to the ones of the previous decades [Meadows et al., 1972].

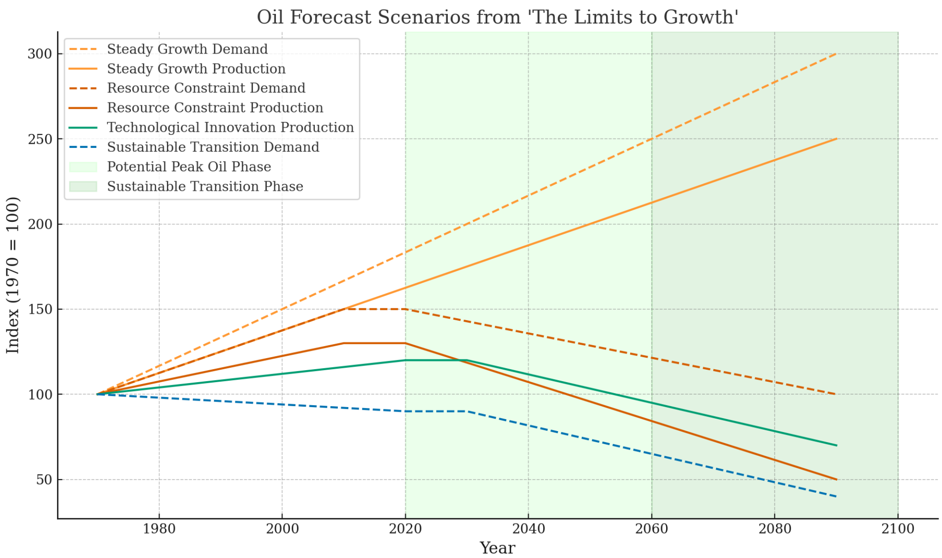

Figure 2.4: Forecasts about future oil demand according to the World3 computer model [Meadows et al., 1972].

Together, the oil shocks of the 1970s and the "Limits to Growth" report played a crucial role in pushing modellers to overcome the limitations of the Hotelling framework. The geopolitical turmoil caused by the shocks underscored the need for a deeper understanding of the inter-temporal optimization problem and the market power wielded by the Organization of Petroleum Exporting Countries (OPEC). The simulations made in the report highlighted the importance of including environmental externalities.

2.2.1 Inter-Temporal Behaviour and OPEC Market Power

Inter-Temporal Behaviour An accurate description of an oilfield inter-temporal optimization requires accounting for economic and geological factors, largely overlooked in the Hotelling framework. Economically, it is important to account for the non-quantifiable uncertainties surrounding future oil and natural gas prices as well as for the discovery of new deposits. Hotelling's model assumes a deterministic environment with no uncertainty about future prices or discoveries; it only considers a finite amount of resources to be extracted along an optimal path. Geologically, it is crucial to include the role of the reservoir pressure and composition (oil, natural gas, and water) as drivers of marginal extraction costs. Hotelling's model, which was initially developed for mining, does not account for either[1].

[1] It is interesting to notice that, even for the fraction of the oil and gas industry that can be modelled within a mining framework (i.e., oil sands), the Hotelling framework is unable to describe the mines' inter-temporal behaviour because it ignores the decisions at the extensive margin (opening *vs* close), focusing instead on the intensive margin (how much oil to extract). In other words, virtually all applications of the Hotelling model assume that mine-level depletion can be aggregated to generate a representative mine, and use its depletion to fit the optimal production path of a particular commodity. However, since, in most cases, the extensive margin decisions are not permanent but rather determined by the interaction between the commodity's aggregate demand and the peculiar geological conditions of a mine, the resulting

Starting in the 1970s several attempts were made to address these two points. In the book chapter "The Econometrics of Exhaustible Resource Supply: A Theory and an Application" [Epple, 2014], Epple explores the econometric modelling of exhaustible resource supply, focusing on the theoretical underpinnings and practical applications of these models. The chapter delves into the economic principles governing the supply of resources like oil and gas, examining how factors such as price expectations, technological advancements, and resource scarcity influence supply decisions. Epple discusses the challenges of modelling exhaustible resource supply, including the uncertainty of future prices and the depletion of resources. In the book "Competition in the Market for an Exhaustible Resource" [Farzin, 1986], Farzin examines the competitive dynamics in markets for exhaustible resources including oil and gas. Farzin analyses how competition among firms affects the extraction and pricing of these resources, taking into account factors like market structure, cost conditions, and strategic interactions. The book explores the implications of different competitive scenarios, including perfect competition and oligopoly, for resource depletion and economic efficiency. Farzin also considers the impact of technological change and government policies on market outcomes. Both the work of Epple and Farzin ignore the impact of new discoveries on prices and extraction costs.

The article "Supply and Costs in the U.S. Petroleum Industry: Two Econometric Studies" by Fisher investigates the supply and cost dynamics of the U.S. petroleum industry through two econometric analyses [Fisher, 1974]. Fisher focuses on developing models to understand the supply response and cost structures within the industry, emphasizing the role of econometric methods in capturing the complexities of the petroleum supply. The study addresses factors such as production costs, market conditions, and regulatory influences, aiming to provide insights into how these factors interact to affect the petroleum supply. The findings highlight the variability in costs and supply elasticity. Three subsequent studies extent Fisher's analysis to the natural gas market. Erickson and Spann article "Supply Response in a Regulated Industry: The Case of Natural Gas" explored how regulation impacts the supply behaviour of natural gas producers, using an econometric approach to model the industry's supply response to changes in price and regulation [Erickson and Spann, 1971]. The study finds that regulatory policies significantly influence supply dynamics, often leading to inefficiencies and distortions in the market. The authors argue that understanding these effects is crucial for designing policies that can effectively balance regulation with market forces. Khazzoom's article "The FPC Staff's Econometric Model of Natural Gas Supply in the United States" provides a detailed analysis of the econometric model developed by the Federal Power Commission (FPC) staff to understand natural gas supply in the U.S. [Khazzoom, 1971]. The model aims to forecast supply by considering factors such as production costs, technological changes, and regulatory impacts. Khazzoom critically evaluates the model's assumptions and methodologies, highlighting its strengths and limitations. The article emphasizes the importance of accurate modelling for policy-making and regulatory decisions, and suggests improvements for better capturing the complexities of the natural gas market. Finally, the technical report "The Economics of the Natural Gas Shortage (1960-1980)" of MacAvoy and Pindyck investigates the causes and economic implications of the natural gas shortage in the United States between 1960 and 1980 [MacAvoy and Pindyck, 1974]. The authors analyse the impact of regulatory policies on supply and demand dynamics, arguing that regulation contributed to the shortage by distorting market signals and investment incentives. The report uses econometric models to assess the effects of price controls and other regulatory measures, highlighting the need for policy reforms to alleviate the shortage and promote a more efficient natural gas market. The study offers recommendations for improving regulatory frameworks to better align with market realities. Both the Fisher article on oil and the subsequent ones on natural gas are econometrics studies in "reduced form" (i.e. neither the exploration nor the extraction equations are derived from an economic model, which describes how the different agents behave). Furthermore, all four studies lack a geological base.

In 1990, Pesaran was the first to address inter-temporal uncertainty in a micro-founded manner, integrating rigorous geological analysis to develop theory-consistent extraction and exploration equations as solutions to a

production pattern does not equalize the marginal net benefit of extraction across different periods [Aguirregabiria and Luengo, 2016].

unified optimization problem [Pesaran, 1990]. Solving the optimization problem within this framework leads to four key consequences. First, the shadow prices of both discovered and undiscovered oil - represented by the Lagrangian multipliers - are not predetermined but are endogenous variables that must be estimated along with other model parameters. Second, these values are influenced by past decisions, thus endogenizing today's optimal levels of exploration and extraction based on yesterday's decisions. Third, the model allows for the incorporation of engineering data on pressure dynamics and reservoir composition into the cost and supply functions. Fourth, all quantities are computed as expected values, accounting for uncertainties related to future oil prices and discoveries. For this groundbreaking work, Pesaran received the Royal Economic Society Prize for the best paper published in the *Economic Journal* during the 1990-1991 period. Since then, most inter-temporal problems in the oil industry have been modelled using similar frameworks [Crémer and Salehi-Isfahani, 2013].

Throughout the thesis, every inter-temporal problem is modelled à la Pesaran. In the simplest framework, every oilfield i is a risk-neutral firm, which exerts no market power. The firm decides in period t its production and investment plan for all subsequent periods. Its intra-temporal profits,

$$\text{Profits}_t^i = \text{Oil \& Gas Revenues}_t^i - \text{Extraction Costs}_t^i - \text{Exploration Costs}_t^i \ ,$$

are the difference between revenues, extraction costs, and exploration costs. The revenues are the product of the oil and gas prices and the quantity of oil and gas extracted. The extraction costs are a function of the quantity of oil extracted and of the quantity of reserves available when the production starts and of the geological specificities of the deposit. The exploration costs are the expenses incurred to discover new oil located in field i. Every oilfield faces two physical constraints. The first one makes the reserves at time t equal to the reserves at time $t-1$, plus new discoveries, minus the volumes of oil extracted at time t[2]. The second constraint ensures that the cumulative discoveries at time t are equal to the ones obtained till time $t-1$ plus the ones obtained at time t. Within this framework, the firm decides the volumes of production and of investment by maximizing the expected discounted future stream of profits. The resulting shadow prices equate inter-temporal marginal revenues with inter-temporal marginal costs.

The thesis exploits the inter-temporal equality between marginal revenues and marginal costs in two different way. In the first two research articles, it calculates the shadow price of discovered oil,

$$\text{Shadow Price of Discovered Oil}_t^i = \text{Marginal Revenues}_t^i - \text{Marginal Extraction Costs}_t^i \ ,$$

as the difference between marginal revenues and marginal extraction cost. In the third article, it finds formula for the shadow price of the reservoir oil and gas composition, the oilfield capacity, the oilfield capacity law-of-motion, and the input-output transformation function as well as the marginal technical rate of substitution between flaring and venting, and the marginal flaring costs. All these variables are used to estimate the flaring and venting supply function. In other words, the first two articles estimate the shadow price of discovered oil, while the third article substitute for three shadow prices to estimate two supply functions. All three exercise adhere to the Pesaran formulation of the inter-temporal problem.

OPEC Market Power Although Hotelling's work included sections on monopolistic and duopolistic behaviour, it did not adequately address the dynamics observed during the oil crises of the 1970s, which underscored the need for a deeper analysis of OPEC's market influence. In 1983, Smith's paper, "OPEC and the World Oil Market: The Dynamics of Price Determination", was instrumental in examining how OPEC's pricing strategies and market behaviour impacted global oil prices [Smith, 1983]. Smith's analysis highlighted the interplay between OPEC's production decisions and market volatility, establishing a foundational understanding of the organization's role

[2]As an alternative, it is possible to make the extraction costs a function of the discovery and of the depletion rate and substitute this constraint with one that imposes that the cumulative depletion of the oilfield exerted until $t-1$, plus the production at time t, equals the cumulative depletion at time t.

in price determination. Hull's 1986 study, "The Economics of OPEC: A Game Theoretic Approach", applied game theory to analyse OPEC's strategic interactions within the oil market, offering a theoretical framework for understanding the competitive dynamics and decision-making processes of member countries [Hull, 1986].

The 1990s marked a period of increasing scrutiny of OPEC's market influence, as researchers sought to understand the organization's adaptation to a rapidly changing global oil landscape. Verleger's 1991 research, "OPEC Behavior and World Oil Prices: A Study of the 1990s", examined how OPEC adjusted its strategies in response to fluctuating global oil prices and rising production from non-OPEC countries [Verleger, 1991]. This study illustrated the complexities of OPEC's market power and its ability to influence prices amidst a shifting supply and demand environment. Additionally, Wills' 1995 paper, "The Role of OPEC in the Global Oil Market: Evidence from the 1990s", provided further insights into OPEC's strategies and their effectiveness in shaping market outcomes during a decade characterized by significant changes in oil production and consumption patterns [Wills, 1995].

Entering the 2000s, research increasingly focused on quantitative assessments of OPEC's market influence. Beattie's 2000 article, "OPEC and Non-OPEC Oil Supply: The Role of OPEC in the New Millennium", analysed the impact of OPEC's policies in the context of rising non-OPEC oil supply, emphasizing the challenges the organization faced in maintaining its market power [Beattie, 2000]. Yergin's 2004 study, "The Dynamics of Oil Prices and OPEC's Market Power", further explored the relationship between oil prices and OPEC's market power, providing a comprehensive analysis of how the organization navigated market fluctuations [Yergin, 2004]. By the late 2000s, McMillan's 2008 paper, "OPEC's Influence on Global Oil Prices: A Quantitative Analysis", quantified OPEC's impact on oil prices and demonstrated how the organization's strategies were adapting to the global market's evolving conditions [McMillan, 2008]. These studies collectively offer a nuanced understanding of OPEC's role and influence.

During the 2010s, research continued to focus on quantitative assessments of OPEC's power, particularly its role in influencing global oil prices, production strategies, and market stability. For example, the paper "OPEC in a Shale Oil World: Where to Next?" by Fattouh examined OPEC's strategies in response to the rise of shale oil production in the United States and its implications for OPEC's market power [Fattouh, 2014]. Similarly, the study "The Impact of OPEC's Production Announcements on Oil Prices: Does OPEC Still Matter?" by Kilian and Zhou investigated the effectiveness of OPEC's production announcements on oil prices during this period, questioning whether the organization still held significant influence over the market [Kilian and Zhou, 2018].

Virtually all these contributions can be integrated into the inter-temporal problem previously described by assuming that each oilfield i is managed by a risk-neutral oil and gas firm k that owns $n(k)$ oilfields, potentially exerting no market power. The firm's intra-temporal profits are calculated as

$$\text{Profits}_t^k = \sum_{i=1}^{n(k)} \text{Profits}_t^{i,k} = \sum_{i=1}^{n(k)} \Big(\text{Oil \& Gas Revenues}_t^{i,k} - \text{Extraction Costs}_t^{i,k} - \text{Exploration Costs}_t^{i,k} \Big) \ ,$$

the difference between revenues, extraction costs, and discovery costs, aggregated across all controlled fields. Within this framework, oil firms internalize the market-clearing condition so that in each period, the equilibrium price ensures that the demand for oil from field i equals its supply. Under relatively mild assumptions, the effect of an increase in the quantity produced by field i in period t equals the oil price plus the oil price rescaled by the firm's market power. Thus, the shadow price of discovered oil in oilfield i owned by firm k becomes

$$\text{Shadow Price}_t^{i,k} = (1 + \text{Market Power}_t^k) \times \text{Oil Price}_t^i - \text{Marginal Extraction Costs}_t^i \ ,$$

where the market power equals the market share enjoyed by firm (or group of firms) k, which is a pure number defined between zero and one, divided by the price elasticity of global oil demand. In other words, the market power correction term divides the capacity of firms to influence the global reference price by the extent to which the demand side of the oil market responds to changes in aggregate supply. This way of incorporating market power can be applied in all three articles. In the first two, it is possible to directly compute the market power and rescale the perfect competition version of the shadow price. In the third article, the shadow price of the oilfield capacity and the law-of-motion for that capacity can be made functions of the firm's market power. This adjustment allows for the derivation of flaring and venting supply functions that are corrected by the degree of market power exerted by the firm owning the oilfield.

A further step to integrate OPEC and other forms of market power into the analysis would require endogenizing market power, making it, for example, a function of the degree of coordination among OPEC members,

$$\text{Market Power}_t^k = f\left(\text{Degree of Coordination among OPEC Members}\right).$$

This step goes beyond the aim of this thesis and is described in the Conclusions as one of the main limitations of the present work.

2.2.2 Environmental Externalities

Prior to the publication of "The Limits to Growth" there was no mathematical link between the Hotelling model and the concept of environmental externalities. In the 1970s, significant advancements were made in the integration of environmental considerations into economic models of resource extraction. Baumol and Oates's seminal 1971 paper, "Environmental Pollution and the Theory of the Second Best", provided a foundational framework for understanding how policies to manage environmental externalities could be integrated into economic theory [Baumol and Oates, 1971]. Their work, published in *Public Economics*, emphasized the importance of considering environmental costs alongside traditional resource management concerns, laying the groundwork for subsequent research on the economic implications of environmental policies in resource extraction. Building on this foundation, Arrow and Fisher published in 1974 the study "Natural Resource Extraction with Irreversible Environmental Effects" addressing the impact of irreversible environmental damages on resource extraction [Arrow and Fisher, 1974]. Their paper, published in the *Journal of Economic Theory*, explored how the potential for irreversible harm from resource extraction necessitates a different approach to managing these resources, incorporating the concept of irreversibility into economic models. This work highlighted the need for strategies that account for long-term environmental impacts, influencing future research on sustainable resource management. Solow's 1974 contribution, "The Economics of Exhaustible Resources", furthered the discussion by focusing on sustainability and environmental impacts within the resource management framework. His paper examined the conditions under which resource extraction could be balanced with environmental preservation, de facto introducing the concept of sustainability in the context of exhaustible resources by integrating environmental constraints into resource management models [Solow, 1974].

The mid-1970s also saw the work of Koopmans, whose 1974 paper, "Optimal Growth in a Nonrenewable Resource Using Economy", explored optimal growth strategies considering nonrenewable resources and environmental constraints. Koopmans' research integrated environmental considerations into the Hotelling framework of resource depletion, offering a model that accounted for both economic growth and environmental sustainability [Koopmans, 1974]. Building on these ideas, Neher's 1976 paper, "Optimal Depletion with Resource Augmenting Technical Progress", introduced the role of technological progress in mitigating environmental damage [Neher, 1976]. While Koopmans framed the environmental challenge within a context of fixed technology and capital accumulation as the sole dynamic force, Neher's model integrated technological change itself as a factor influenced by the use of nonrenewable resources. This approach marked a shift towards understanding how technological advancements could alter resource depletion strategies and address environmental impacts more effectively. In the late 1970s several other studies incorporated environmental externalities and technological change into extraction models [Conrad, 1979, Dasgupta and Heal, 1979].

All these works were strictly theoretical. In different forms, they constructed an Hotelling-augmented framework, in which there were environmental externalities.

2.3 Field-Level Analysis

2.3.1 Inter-Temporal Behaviour and OPEC Market Power

Discussions about the Hotelling model's limitations in accurately representing the global oil and gas market post-1970s were either theoretical or conducted using aggregate data, typically at the national level. These studies implicitly assume that depletion effects observed at individual oilfields can be aggregated to reflect similar effects across the entire industry. However, the conditions necessary for this "representative model" to be valid are quite restrictive and generally do not hold true. This limitation is particularly evident when considering the heterogeneity across oilfields and producers in terms of control variables (such as extraction and exploration) and state variables (including reserve volumes, geological characteristics, etc.). Consequently, using aggregate data to test the Hotelling rule or analyse OPEC's behaviour can be misleading. More importantly, estimating models at the industry level can introduce significant biases in assessing short- and long-run responses to demand and supply shocks and in evaluating the effects of public policies.

Recent advancements in business intelligence have overcome these data limitations, significantly enhancing researchers' ability to analyse different behavioural assumptions. Two detailed data repository provide field-level information:

- WoodMac Upstream Data Tool is an analytical platform designed to optimize the management of oil and gas exploration and production data. WoodMac integrates geological, geophysical, and production data, offering real-time monitoring of exploration and extraction decisions.

- Rystad Energy's UCube (Upstream Database Cube) is a dynamic data tool providing detailed insights into the global oil and gas upstream sector. UCube offers an extensive database covering over 65,000 fields and licenses, with critical information on reserves, production, costs, and financials. By integrating geological, technical, and economic data, UCube provides a comprehensive view of the upstream sector.

The advent of these and other micro-level data repositories has opened new avenues for studying the global oil and gas market. In particular, the increased granularity of available information has enhanced researchers' ability to study field and firm-level behaviour jointly with the market structure. This data-rich environment has reopened debates on inter-temporal behaviour and OPEC's market power.

Inter-Temporal Behaviour with Micro-Data The paper "Hotelling Under Pressure" by Anderson, Kellogg, and Salant challenges traditional models that use a unified optimization problem to determine the exploration-extraction equilibrium by introducing reservoir pressure as a key constraint [Anderson et al., 2018]. Unlike the Hotelling model, or Pesaran formulation, which assume a direct response of production to oil prices, empirical evidence suggests that existing wells' production rates are less sensitive to price changes, see Figure 2.5. Instead, drilling activity correlates strongly with oil prices. In other words, oil production does not respond at the intensive margin (more or less production from existing wells) but only at the extensive margin (more wells drilled when the oil price increases, less wells drilled when the oil price deceases). Within this framework, it is possible to write a new Hotelling model, where firms decide on drilling new wells while facing constraints from declining reservoir pressure in existing wells. This model better explains observed patterns in oil prices, drilling activity, and production. It suggests that adjustments to the traditional Hotelling rule are necessary, incorporating costs and constraints related to reservoir pressure into the extraction strategy.

The second article of the thesis borrows from Anderson, Kellogg, and Salant constructing a model where each well output depends solely upon its capacity, which is itself a function of its depletion. This implies that oil firms

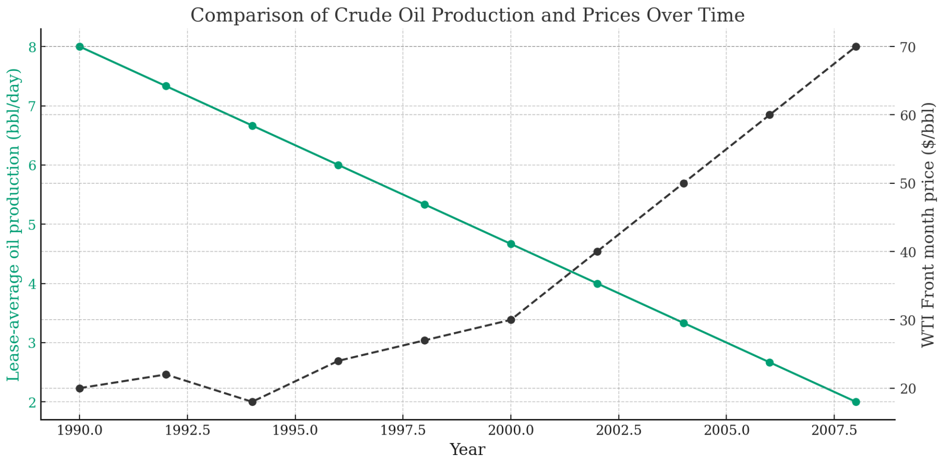

Figure 2.5: Reaction of production from existing wells to changes in prices as presented in Anderson, Kellogg, and Salant (2018).

can respond to long-term anticipated changes in oil prices by increasing the overall capacity of the oilfield at the extensive margin (i.e., by drilling new extraction wells). However, the model presented in the thesis differs from their because it gives the possibility to oil firms to respond to short- and medium-term market shocks by boosting the natural pressure of a well through the injection of liquids and/or gases. Injections are performed through existing or newly drilled injection wells and using specific inputs (steam, water, electricity, chemicals, etc.), which are purchased by the firm at market prices and contribute to boosting extraction costs. This addition captures the key features of the oil extraction process as described in the previous sections while retaining most of the empirically relevant features of the analysis of Anderson, Kellogg, and Salant.

OPEC Market Power with Micro-Data The paper "(Mis)Allocation, Market Power, and Global Oil Extraction" by Asker, Collard-Wexler, and De Loecker presents a pioneering analysis of inefficiencies in global oil extraction by examining the extent of misallocation in the industry [Asker et al., 2019]. The authors achieve this by comparing actual industry cost curves, derived from field-level costs, with hypothetical scenarios that assume no distortions. Unlike studies that model OPEC's behaviour, this article treats OPEC's influence as a given, focusing instead on the economic outcomes observed in the market versus those that would result from a perfectly competitive environment, where production is ordered from the cheapest to the most expensive oilfield. The study identifies significant misallocation, estimating welfare losses of $744 billion, with a substantial portion attributed to the market power wielded by industry players. This research marks a significant first step in quantifying the economic costs imposed by OPEC's coordination efforts on the global economy, demonstrating the potential of micro-data to illuminate the impacts of market distortions in the oil industry.

Asker, Collard-Wexler, and De Loecker's approach does not use micro-level data to test different hypotheses about OPEC's behaviour but rather to compare two distinct market scenarios. In the first scenario, "OPEC does whatever OPEC does," resulting in the current production allocation. In the second scenario, a perfectly competitive market generates a price vector that optimally allocates production based on costs. By comparing these allocations, the authors assess the economic distortion caused by market power. The thesis adopts a similar methodology, incorporating market power without specifically modeling OPEC's behavior. Instead, market power is quantified by calculating two quantities: one derived from micro-level production data that aggregates the output from all oilfields owned by a particular firm or group, and another that serves as an aggregate measure of global demand elasticity. Although micro-data are used differently, both approaches treat OPEC's behaviour as an external factor rather than modelling it directly. This consistent approach underscores the focus on measuring the consequences of market power and allocation inefficiencies without exploring the intricacies of OPEC's strategic decisions.

2.3.2 Environmental Externalities

All the studies that introduced environmental externalities into the Hotelling model required to calculate the marginal emissions of oil and gas operations in order to determine their associated social costs. However, this approach is fundamentally challenged by the fact that such a uniform measure of emissions “does not exist”. Each barrel of oil extracted from every oilfield, transported to a refinery, and eventually combusted in its final form produces a unique quantity of greenhouse gas emissions. The idea of circumventing this complexity by assuming the emissions of a representative barrel - and consequently its displacement - rather than leveraging micro-level information is fraught with issues.

Upstream As highlighted by Masnadi et al. in their study “Global Carbon Intensity of Crude Oil Production”, published in *Science*, there is substantial variability in carbon intensity across different oilfields [Masnadi et al., 2018]. This variability arises from diverse factors such as geographic location, reservoir characteristics, and the specific technologies used in extraction and processing. Their analysis, based on data from over 8,000 oilfields representing about 98% of global oil production, underscores the critical need to account the fact that emissions can vary widely depending on specific field-level operations, such as flaring and venting practices, or the energy intensity of different extraction techniques like enhanced oil recovery (EOR). Relying on the concept of a generalized representative barrel can obscure these nuances, leading to inaccurate assessments of greenhouse gas emissions and potentially misinforming upstream-focused policy.

Midstream Beyond the variability in emissions during extraction, the carbon intensity associated with crude oil refining also varies significantly. Jing et al., in their paper “Carbon Intensity of Global Crude Oil Refining and Mitigation Potential”, published in *Nature Climate Change*, highlight how the carbon intensity of crude oil refining is influenced by factors such as refinery configuration, operational practices, and the quality of the crude oil being processed [Jing et al., 2020]. The study emphasizes the potential for mitigation through optimization of refining operations and the adoption of more efficient technologies. The findings suggest that emissions from the midstream sector can vary widely, depending on the specific characteristics of the refining process, similar to the variability seen in the upstream extraction phase. Ignoring these distinctions and assuming a representative barrel in the refining process risks underestimating the true environmental impact of oil production, thereby leading to misguided decisions in midstream emissions management.

To the best of my knowledge, this thesis represents the first comprehensive attempt to model the inter-temporal problem using Pesaran’s framework, while simultaneously incorporating and expanding the micro-level critique of Anderson, Kellogg, and Salant, and accounting for OPEC’s market power as analysed by Asker, Collard-Wexler, and De Loecker, all while coupling this disaggregated framework with the micro-level analysis of environmental externalities pioneered by the Adam Brandt research group [El-Houjeiri et al., 2017].

Chapter 3

Methodology

3.1 Economic Analysis

The present dissertation adopts an Applied Industrial Organization (AIO) approach, a framework that combines economic theory with empirical methods to analyse demand and supply decisions within an industry. This approach aims to understand and investigate the behaviour of firms, consumers, and policymakers, providing a structured lens through which to examine strategic interactions within specific market structures and regulatory contexts. By integrating theoretical insights with real-world data, the AIO approach facilitates a nuanced understanding of the competitive dynamics and policy implications that shape industry outcomes.

Demand Side In the context of the oil and gas industry, demand-side modelling involves understanding how various factors influence consumer and industrial demand for oil and gas derived products. Demand can be influenced by factors such as price changes, income levels, technological advancements, and substitution effects from alternative energy sources. Standard econometric models estimate the demand function incorporating all these variables.

This thesis does not primarily focus on modelling the demand side of the oil and gas industry. Instead, it treats demand in aggregate form, utilizing elasticity estimates from prior empirical studies to calculate the market power correction term in the first two articles [Kilian, 2022]. In the third article, an aggregate estimate of the cross-price elasticity between coal and natural gas, as well as between natural gas and nuclear power, is derived using a seemingly unrelated regression model [Zellner, 1962]. This approach allows the thesis to focus more intensely on the supply side and market structure - key sources of the upstream and midstream heterogeneity that underpin the central motivations of the research. From an economic perspective, the decision to model demand at a macro level while focusing on micro-level supply-side modelling is addressed in the conclusions as one of the main limitations of the thesis. This approach may not capture the full range of interactions between supply and demand, particularly in how micro-level supply decisions can influence macro-level demand outcomes and vice versa. Recognizing this discrepancy highlights an area for future research, suggesting that further studies could integrate more detailed demand-side analysis to complement the supply-side focus of the thesis[1].

Supply Side In the context of the oil and gas industry, supply-side modelling involves understanding the strategic decisions of firms regarding exploration, extraction, production, and pricing. Within the standard supply-side framework the thesis focuses on estimating two quantities: 1) the marginal extraction costs, and 2) the supply function of unsold natural gas.

[1] Note that, from an environmental perspective, the decision to model demand at a macro level while assuming homothetic preferences for consumers seems less significant. A change in the composition of demand for oil-derived products, given a change in demand level, would have minor effects on emissions due to the combustion of refined products.

Marginal Extraction Costs The first two papers of this thesis develop an econometric framework for quantifying the value of one extra barrel of oil located in a particular oilfield at a specific point in time. Contrary to previous research, the thesis is not interested in fitting the exploration and/or the oil and gas supply function but rather in quantifying the opportunity cost of adding one barrel of additional proven reserves in a particular oilfield. As shown in section 2.2, in order to compute this quantity it is necessary to collect four pieces of information. First, the elasticity of global oil demand. As stated in the previous paragraph, this is not the main focus of the thesis, and these numbers are taken from previous publications [Kilian, 2009]. Second, the market power of the firm owning that particular oilfield. This quantity is computed by aggregating production for oilfields owned by International Oil Companies (IOC) and aggregating the production of all the oilfields owned by OPEC members for National Oil Companies (NOC) members of the cartel[2]. Third, the price at which a particular oilfield sells its output. This is computed using a pricing equation where the price of a specific oilfield's output deviates from the global average depending on the API gravity of the sold crude and its sulphur content[3]. Finally, the marginal extraction costs. This last quantity is (almost always) computed in AIO starting from the production function. For example, if the quantity of oil extracted by oilfield i in year t, Q_t^i, is a function of labour L_t^i, capital K_t^i, and the remaining stock of the resource R_t^i,

$$Q_t^i = f(L_t^i, K_t^i, R_t^i) \ ,$$

and $f(.)$ is a Cobb-Douglas production function $Q_t^i = \mathcal{A}{L_t^i}^{\alpha}{K_t^i}^{\beta}{R_t^i}^{\gamma}$, where $\mathcal{A}$ is a constant representing total factor productivity, and α, β, and γ are the output elasticities of labour, capital, and the remaining stock of the resource, then it is possible to estimate the marginal costs using the following procedure. First, write a cost function, which represents the minimum cost of producing a given level of output Q for a given vector of input prices. Given the production function, the cost minimization problem can be set up as follows:

$$\text{Minimize } C_t^i = w_t L_t^i + r_t K_t^i + c_t R_t^i$$

$$\text{subject to } Q_t^i = \mathcal{A}{L_t^i}^{\alpha}{K_t^i}^{\beta}{R_t^i}^{\gamma}$$

where w_t is the price of labour, r_t for capital, and c_t for the cost related to the remaining stock of the resource. To solve this minimization problem, we can use the method of Lagrange multipliers. The Lagrangian function writes

$$\mathcal{L}_t^i = w_t L_t^i + r_t K_t^i + c_t R_t^i + \lambda_t^i[Q_t^i - \mathcal{A}{L_t^i}^{\alpha}{K_t^i}^{\beta}{R_t^i}^{\gamma}] \ .$$

Using the first-order conditions,

$$\begin{aligned}
\frac{\partial \mathcal{L}_t^i}{\partial L_t^i} &= w_t - \lambda_t^i \alpha \mathcal{A}{L_t^i}^{\alpha-1}{K_t^i}^{\beta}{R_t^i}^{\gamma} = 0 \\
\frac{\partial \mathcal{L}_t^i}{\partial K_t^i} &= r_t - \lambda_t^i \beta \mathcal{A}{L_t^i}^{\alpha}{K_t^i}^{\beta-1}{R_t^i}^{\gamma} = 0 \\
\frac{\partial \mathcal{L}_t^i}{\partial R_t^i} &= c_t - \lambda_t^i \gamma \mathcal{A}{L_t^i}^{\alpha}{K_t^i}^{\beta}{R_t^i}^{\gamma-1} = 0 \\
\frac{\partial \mathcal{L}_t^i}{\partial \lambda_t^i} &= Q_t^i - \mathcal{A}{L_t^i}^{\alpha}{K_t^i}^{\beta}{R_t^i}^{\gamma} = 0 \ ,
\end{aligned}$$

[2]This assumption of perfect coordination among cartel members can be relaxed by multiplying the market share by a coefficient measuring the "degree-of-coordination" within the cartel; see the second research article for a detailed discussion.

[3]Note that this type of pricing equation is consistent with the aggregate demand model from which the elasticity is derived.

it is possible to solve for λ_t^i,

$$\lambda_t^i = \frac{w_t}{\alpha \mathcal{A}(L_t^i)^{\alpha-1}(K_t^i)^{\beta}(R_t^i)^{\gamma}} = \frac{r_t}{\beta \mathcal{A}(L_t^i)^{\alpha}(K_t^i)^{\beta-1}(R_t^i)^{\gamma}} = \frac{c_t}{\gamma \mathcal{A}(L_t^i)^{\alpha}(K_t^i)^{\beta}(R_t^i)^{\gamma-1}} .$$

Rearranging the previous three equations, it is possible to obtain the input demand functions

$$\begin{aligned} L_t^i &= \left(\frac{\alpha Q_t^i}{\mathcal{A} w_t}\right)^{\frac{1}{\alpha}} \left(\frac{K_t^i}{L_t^i}\right)^{\frac{\beta}{\alpha}} \left(\frac{R_t^i}{L_t^i}\right)^{\frac{\gamma}{\alpha}} \\ K_t^i &= \left(\frac{\beta Q_t^i}{\mathcal{A} r_t}\right)^{\frac{1}{\beta}} \left(\frac{L_t^i}{K_t^i}\right)^{\frac{\alpha}{\beta}} \left(\frac{R_t^i}{K_t^i}\right)^{\frac{\gamma}{\beta}} \\ R_t^i &= \left(\frac{\gamma Q_t^i}{\mathcal{A} c_t}\right)^{\frac{1}{\gamma}} \left(\frac{L_t^i}{R_t^i}\right)^{\frac{\alpha}{\gamma}} \left(\frac{K_t^i}{R_t^i}\right)^{\frac{\beta}{\gamma}} , \end{aligned}$$

for the three factors of production. Plugging the input demand functions (i.e., the optimal values of L, K, and R) into the cost equation $C(Q_t^i) = w_t L_t^{i\,*} + r_t K_t^{i\,*} + c_t R_t^{i\,*}$, it is possible to write the cost function,

$$C(Q_t^i) = w_t \left(\frac{\alpha Q_t^i}{\mathcal{A} w_t}\right)^{\frac{1}{\alpha}} \left(\frac{K_t^i}{L_t^i}\right)^{\frac{\beta}{\alpha}} \left(\frac{R_t^i}{L_t^i}\right)^{\frac{\gamma}{\alpha}} + r_t \left(\frac{\beta Q_t^i}{\mathcal{A} r_t}\right)^{\frac{1}{\beta}} \left(\frac{L_t^i}{K_t^i}\right)^{\frac{\alpha}{\beta}} \left(\frac{R_t^i}{K_t^i}\right)^{\frac{\gamma}{\beta}} + c_t \left(\frac{\gamma Q_t^i}{\mathcal{A} c_t}\right)^{\frac{1}{\gamma}} \left(\frac{L_t^i}{R_t^i}\right)^{\frac{\alpha}{\gamma}} \left(\frac{K_t^i}{R_t^i}\right)^{\frac{\beta}{\gamma}}$$

Rearranging for Q_t^i, we obtain the cost function

$$C(Q_t^i) = \mathcal{A}^{-1} {Q_t^i}^{\frac{1}{\alpha+\beta+\gamma}} (w_t^{\alpha} + r_t^{\beta} + c_t^{\gamma}) .$$

This function provides the total cost of producing a given level of output Q based on the prices of labour, capital, and resource stock.

In order to obtain the structural parameters $(\hat{\alpha}, \hat{\beta}, \hat{\gamma})$, it is necessary to estimate the production function without incurring in the transmission bias [Eberhardt and Helmers, 2010]. The two most widely used methods are those proposed by Olley and Pakes [Olley and Pakes, 1992] and Levinsohn and Petrin [Levinsohn and Petrin, 2003]. The Olley and Pakes approach is based on five key assumptions:

1. Capital Adjustment. Firms can adjust their capital stock smoothly over time.
2. Exogenous Inputs. Inputs other than capital (such as labour, energy, materials, and services) are exogenous to productivity shocks.
3. Perfect Foresight. Firms have perfect foresight and make optimal dynamic decisions regarding investment.
4. Monotonic Relationship. The model assumes a strict monotonic relationship between investment and unobserved productivity shocks.
5. Firm Exit. Exit decisions are based solely on productivity shocks and are used as a source of identification.

While the first two assumptions might not be too restrictive, the remaining three are clearly violated in the context of the oil and gas industry. The third assumption is violated due to the largely non-quantifiable uncertainty that surrounds the future movements of oil prices and the discoveries of new oilfields as discussed in section 2.2. The fourth hypothesis is violated because investment in future capacity can become part of an optimal strategy for OPEC members who wish to remain or become the principal in OPEC's coordination game. Finally, the fifth hypothesis is violated because the decision to stop producing from a particular oilfield might not be related to its productivity but rather to market conditions (e.g., a decline in demand) or, more importantly, strategic interactions (e.g., the outcome of the game in quantities played by IOCs or a coordination effort made by the NOCs member of OPEC). Similarly, Levinsohn and Petrin assume a monotonic relationship between intermediate inputs and productivity shocks. This assumption is clearly violated due to the role pressure plays in determining marginal extraction costs. The most productive oilfields are those with the highest (average) natural pressure, which is a substitute for intermediate inputs. As pressure declines the intermediate inputs become a substitute for it. However, this does not always happen in a "monotonic-way", but it rather depends upon the geological characteristics of a particular oilfield. Furthermore, it is almost impossible to find reliable data on intermediate inputs (such as materials) at a field level.

Therefore, the thesis, contrary to mainstream AIO, uses a *cost data approach*. In other words the first two articles of the thesis estimate the marginal extraction costs fitting the cost function directly. This method has the advantage of using available accounting cost data. Contrary to the indirect approach of first fitting the production function and then get the marginal costs, the direct fitting relies on two assumptions. First, firms are cost-minimizers. Second, in a panel context,

$$C(Q_t^i) = \text{Fix Costs}_t^i + \text{Variable Costs}_t^i + \epsilon_t^i \ ,$$

if the error term ϵ_t^i contains a field-specific effect and a random noise, $\epsilon_t^i = \mu^i + \eta_t^i$. Then, by differentiating period t and $t-1$, $\eta_t^i - \eta_{t-1}^i$ is uncorrelated with the deltas in fix and variable costs, solving the reverse causality problem. This way of fitting the marginal extraction costs presents two main limitations. First, the regression does not exploit the panel natural of the data loosing the opportunity of quantifying unobserved field-specific heterogeneity. Second, if the accounting costs are not the real economic costs (i.e., accounting costs do not capture all the costs firms face.), then the dependent variable might contain a measurement error. The latter is more likely to survive a first-difference then a within-transformation.

Supply Functions of Unsold Natural Gas The third paper connects the identification condition of the theoretical model with the estimation of two supply functions. The first is the flaring supply function. In the model, the optimal quantity of flaring can be negative, but flaring is only observable when it is zero or positive. Consequently, the supply regression is specified as follows:

$$\text{Flare}_t^{ik} = \begin{cases} \alpha_0^{ik} + \alpha_1 P_t^{\text{Gas}} + \eta_t^{ik} & \text{if } \eta_t^{ik} > -\alpha_0^{ik} - \alpha_1 P_t^{\text{Gas}} \\ 0 & \text{otherwise .} \end{cases}$$

Given that the observed data on flaring is censored - visible only when it is non-negative - the appropriate method of estimation is a panel Tobit model. This model is an extension of standard regression techniques, frequently used in panel data analysis when the dependent variable is subject to censoring or truncation. The panel Tobit model accommodates both the cross-sectional and temporal dimensions of the data, capturing the dynamics of variables with limited observability. It is particularly advantageous when dealing with datasets where a substantial portion of the outcome variable's observations is clustered at a boundary, like in the case of flaring. By incorporating random effects, this model accounts for unobserved heterogeneity across individual oilfields. The estimation process typically involves a non-standard maximum likelihood approach, see section 6.4.1 for a detailed description. These residuals are subsequently used as explanatory variables in the second supply function, which quantifies the amount of natural gas intentionally vented by a firm using a standard linear regression model.

3.2 Environmental Analysis

The thesis employs two engineering-based life cycle assessment (LCA) tools designed to measure greenhouse gas (GHG) emissions across the stages of production, processing, transport, and refining of crude petroleum. By combining the system boundaries of these tools, the LCA analysis spans from initial exploration to the refinery exit gate. Both tools incorporate essential LCA concepts, including goal and scope definition, system boundaries, functional units, and allocation methods, in accordance with the principles outlined in the International Standard ISO 14041 [ISO, 1998]. ISO 14041, part of the ISO 14040 series, establishes guidelines for defining the purpose and boundaries of an LCA study and conducting inventory analysis to quantify inputs and outputs associated with a product system. Though ISO 14041 has since been integrated into the broader standards of ISO 14040 and ISO 14044, its foundational principles remain central to contemporary LCA practices.

3.2.1 Upstream Emissions

The Oil Production Greenhouse Gas Emissions Estimator (OPGEE) model is a comprehensive tool designed to compute the greenhouse gas (GHG) intensity of crude oil production from different oilfields [El-Houjeiri et al., 2017]. Developed at Stanford University, OPGEE provides a detailed and systematic methodology to evaluate the carbon intensity associated with each stage of oil production, from extraction to the entry gate of a refinery. The methodology implemented by OPGEE involves detailed modelling of the field's characteristics, the production method, energy use, and GHG emissions.

Overview of OPGEE Methodology OPGEE's methodology involves six key stages to calculate the carbon intensity of oilfields:

1. Data Collection and Input Requirements
2. Process Emissions Calculation
3. Energy Consumption and Fuel Use
4. Embodied Emissions
5. Land Use Change Impacts
6. GHG Intensity Calculation

Each of these stages includes specific steps and considerations to account for the diverse and complex nature of oil production operations.

1. Data Collection and Input Requirements The first step in OPGEE's methodology involves gathering data on various parameters related to oil production. These parameters include: 1) field characteristics (depth, age, and production rates of the oilfields), 2) production method (types of recovery methods: primary, secondary, tertiary), and 3) infrastructure and equipment (types and quantities of materials used in well construction, surface processing, and transportation infrastructure).

2. Process Emissions Calculation OPGEE calculates process emissions by evaluating the emissions associated with each stage of oil production. This includes: 1) drilling and completion (emissions from the use of drilling rigs, completion equipment, and the consumption of materials such as cement and steel), 2) production operations (emissions from the operation of pumps, compressors, and other equipment used to extract oil), and 3) separation and treatment (emissions from separating oil, gas, and water, and treating the produced fluids). The model considers both direct emissions (e.g., combustion of fuels) and indirect emissions (e.g., emissions from electricity consumption) of each of the oil production stages.

3. Energy Consumption and Fuel Use Energy consumption and fuel use are critical components of the GHG emissions calculation. OPGEE accounts for: 1) fuel use (types and quantities of fuels used in various stages of oil production, such as diesel for drilling rigs and natural gas for processing facilities), and 2) electricity use (consumption of electricity for operations and its associated emissions based on regional grid mixes). OPGEE uses default values for energy efficiencies and emissions factors, which can be adjusted based on specific field data.

4. Embodied Emissions Embodied emissions refer to the GHG emissions associated with the production, processing, and transportation of materials used in oilfields. OPGEE includes embodied emissions for: 1) wellbore steel and cement (emissions from the production of steel and cement used in well construction), 2) surface piping and equipment (emissions from manufacturing and transporting surface infrastructure components), 3) fracturing sand and water (emissions from procuring and transporting sand and water for hydraulic fracturing operations).

5. Land Use Change Impacts Land use change impacts are included in OPGEE to account for emissions resulting from the disturbance of land during oilfield development. This includes: 1) soil carbon oxidation (emissions from the oxidation of soil carbon when land is cleared), 2) biomass carbon oxidation (emissions from the disturbance and oxidation of biomass carbon), 3) foregone sequestration (reduced carbon sequestration capacity due to land clearing). OPGEE provides options for analyzing land use impacts over different timeframes (e.g., 30 *vs* 150 years).

6. GHG Intensity Calculation Finally, OPGEE integrates all the emissions from the various stages and processes to compute the GHG intensity of crude oil production. This includes: 1) combustion emissions (CO_2, CH_4, and N_2O emissions from the combustion of fuels used in operation), 2) non-combustion emissions from routine flaring, non-routine flaring, venting, and leakages during the production process, and 3) embodied emissions (summed emissions from the production and transport of materials used).

3.2.2 Midstream Emissions

The Petroleum Refinery Life-cycle Inventory Model (PRELIM) is a tool designed to estimate the carbon intensity and environmental impacts of different petroleum refineries [Abella et al., 2015]. Developed at Calgary University, PRELIM provides a detailed and systematic methodology to evaluate the carbon intensity associated with each stage of oil refining, from the entry gate of a refinery to the exit gate of a refinery. The methodology implemented by PRELIM involves detailed modelling of refinery configurations, process units, energy use, and GHG emissions.

Overview of PRELIM Methodology PRELIM's methodology involves four key stages to calculate the carbon intensity of a petroleum refinery:

1. Data Collection and Input Requirements
2. Refinery Configuration and Process Emissions Calculation
3. Energy Consumption and Fuel Use
4. GHG Intensity Calculation

1. Data Collection and Input Requirements PRELIM starts by using crude oil quality information, typically available as a crude oil assay. The assay data is converted into a PRELIM-specific format that includes five key parameters: crude distillation curve, sulphur content, API gravity, carbon residue content, and hydrogen content.

These parameters are specified for nine fractions of the crude oil, each associated with a specific cut-temperature. This detailed breakdown allows PRELIM to accurately model the processing characteristics of different crude oils.

2. Refinery Configuration and Process Emissions Calculation PRELIM simulates seven different refinery configurations, each including various process units such as crude distillation, naphtha catalytic reforming, isomerization, steam methane reforming, and pollution control measures like flue gas treatment and sulphur recovery. The configurations range from simple hydroskimming setups to complex deep conversion refineries.

3. Energy Consumption and Fuel Use For each process unit, PRELIM calculates the energy required and predicts the quantity and type of products produced. The model uses data on process unit energy requirements, intermediate product yields, and crude assay information to determine the overall energy use and GHG emissions. Emissions associated with electricity and natural gas are also considered. PRELIM employs an allocation method to trace energy use and emissions through the refinery. Energy and emissions are allocated to the process unit throughputs, and this allocation is carried through the entire refinery to the final products. The model allows for flexibility in choosing which products are assigned emissions and on what basis (e.g., hydrogen content, mass, market value, or energy content).

4. GHG Intensity Calculation PRELIM calculates combustion emissions resulting from the generation of heat or steam from fuel combustion and assigns these to refinery products based on their aggregate heat demand. Process emissions from refinery units are assigned to downstream products using an input-output approach, reflecting the assay and model parameters selected. Fugitive emissions, though an order of magnitude lower than combustion emissions, are also included as support services emissions. To ensure accuracy, PRELIM verifies the energy and material balance in the system by comparing the overall energy requirements calculated by summing the embedded energy in all refinery final products with the total energy requirements from all process units. This step helps validate the model's predictions and ensures consistency in the results.

3.2.3 Downstream Emissions

Throughout the thesis, the demand for oil-derived products is modelled under the assumption of homothetic preferences. In a homothetic demand framework, reductions in total demand for oil are distributed proportionally across all oil-derived products. For instance, if there is a 1% decrease in overall oil demand, and jet fuel represents 25% of the refined products, then the demand for jet fuel would also decrease by 1% of its share, resulting in a 0.25% reduction in jet fuel demand. This proportional reduction applies consistently across all products derived from crude oil, maintaining a fixed relationship between the consumption levels of each product. This approach makes it essential to include combustion-related emissions for each product category to accurately assess the entire LCA impact under varying demand scenarios.

The Center for Corporate Climate Leadership of the Environmental Protection Agency (EPA) offers comprehensive guidance for calculating direct GHG emissions from stationary combustion sources. This guidance, outlined in the "Greenhouse Gas Inventory Guidance: Direct Emissions from Stationary Combustion Sources," is based on the GHG Protocol developed by the World Resources Institute (WRI) and the World Business Council for Sustainable Development (WBCSD) [WRI, 2004, WBCSD, 2005]. It extends these principles to incorporate EPA-specific calculation methodologies and emission factors [EPA, 2022b]. The structured methodology presented in the inventory guidance aids organizations in accurately assessing their direct emissions, ensuring alignment with regulatory requirements and promoting transparency in GHG reporting, which is crucial for accurately representing the full scope of emissions in scenarios involving changes in global oil demand. This structure positions the EPA's guidance as a foundation for capturing the combustion emissions necessary for a complete LCA in demand-modelling scenarios [EPA, 2018b, 2022a].

The primary GHGs produced from stationary combustion are CO_2, CH_4, and N_2O. There are two main methods for estimating GHG emissions from stationary combustion sources:

Continuous Emissions Monitoring System (CEMS) Method[4] CEMS is used for the continuous measurement of pollutants emitted from combustion processes. It measures CO_2 emissions through either a direct monitor of CO_2 concentration and flow rate or an O_2 concentration monitor combined with theoretical calculations based on fuel characteristics. Continuous emissions monitoring is the continuous measurement of pollutants emitted into the atmosphere in exhaust gases from combustion or industrial processes. There are two approaches to determine CO_2 emissions using CEMS:

1. A monitor measuring hourly average CO_2 (or O_2) concentration percent by volume of flue gas and a flow monitoring system measuring the volumetric flow rate of flue gas can be used to determine CO_2 mass emissions.
2. A monitor measuring O_2 concentration percent by volume of flue gas combined with theoretical CO_2 and flue gas production based on fuel characteristics can be used to determine CO_2 flue gas emissions and CO_2 mass emissions.

Fuel Analysis Method This method involves calculating emissions based on the carbon content of the fuel combusted. Three equations are provided, with the choice of equation depending on the available data about the fuel (mass/volume, heat content, or carbon content). The fuel analysis method to calculate CO_2 emissions involves determining the carbon content of fuel combusted using either fuel-specific information or default emission factors and applying that carbon content to the amount of fuel burned during the reporting year. Three equations can be used to calculate the CO_2 emissions from each type of fuel combusted. The first one,

$$\text{Emissions} = \text{Fuel} \cdot \text{EF}_1 \ ,$$

calculates the mass of the mass of CO_2, CH_4, or N_2O emitted multiplying the mass or volume of fuel combusted by the emissions factor of CO_2, CH_4, or N_2O per mass or volume unit. The second one,

$$\text{Emissions} = \text{Fuel} \cdot \text{HHV} \cdot \text{EF}_2 \ ,$$

calculates the mass of the mass of CO_2, CH_4, or N_2O emitted multiplying the mass or volume of fuel combusted by the fuel heat content (higher heating value) measured in units of energy per mass or volume of fuel and by the CO_2, CH_4, or N_2O emissions factor per energy unit. The third one,

$$\text{Emissions} = \text{Fuel} \cdot \text{CC} \cdot \frac{44}{12} \ ,$$

calculates the mass of the mass of CO_2, CH_4, or N_2O emitted multiplying the mass or volume of fuel combusted by the fuel carbon content measured in units of mass of carbon per mass or volume of fuel and by 44/12 (i.e. the ratio of molecular weights of CO_2 and carbon).

Two of these equations can also be used to calculate CH_4 and N_2O emissions using appropriate emission factors. The appropriate equation to use depends on what is known about the characteristics of the fuel being consumed. The first equation is recommended when fuel consumption is known only in mass or volume units, and no information is available about the fuel heat content or carbon content. This equation is the least preferred. It has the most uncertainty because its emission factors are based on default fuel heat content, rather than actual heat content. The second equation is recommended when the actual fuel heat content is provided by the fuel supplier or is otherwise known. It is also recommended when the fuel use is provided in energy units (e.g., therms of natural gas). In such cases, the fuel use in energy units can be multiplied directly by the emission factor (EF_2). The second equation is a preferable approach over the first one because it uses emission factors that are based on energy units as opposed to mass or volume units. Emission factors based on energy units are less variable than factors per mass or volume units because the carbon content of a fuel is more closely related to the heat content of the fuel than to the physical quantity of fuel. The third equation is recommended to calculate CO_2 emissions when the actual carbon content of the fuel is known. Carbon content is typically expressed as a percentage by mass, which requires fuel use data in mass units. This equation is most preferred for CO_2 calculations because CO_2 emissions are directly related to the fuel's carbon content. Because the last equation is only applicable to CO_2 emissions, the first or the second should be used in conjunction to calculate CH_4 and N_2O emissions.

3.2.4 Integrating Up, Mid, and Downstream Life Cycle Analysis

The outcomes of OPGEE and PRELIM are integrated with traditional repositories of stationary fossil fuel combustion in different ways throughout the thesis.

The first two articles utilize the three emissions calculators in decreasing order of granularity. Initially, virtually all global oilfields are evaluated using OPGEE to determine their upstream carbon intensity. Ideally, the system boundary approach would suggest linking each field with the refineries to which the crude is sold and then running them through PRELIM to obtain the well-to-refinery exit gate environmental footprint of the global oil supply. However, due to the confidentiality of seller-buyer contracts at the oilfield-refinery level, it is not possible to access this data. Therefore, the two articles decrease the level of granularity. Assuming that a fraction of the global oil demand changes, it is possible to know how the chemical characteristics of the global crude would change. Specifically, since most of the displaced oil is heavy (API gravity < 22) and sour (sulphur content > 1%), the global traded crude would become lighter and sweeter. At this point, it is possible to run the new stream of crude through an imaginary refinery whose returns to API gravity and sulphur content comes from averaging the returns of hundreds of refineries previously analysed via PRELIM. Once the change in midstream emissions is calculated, the associated change in downstream emissions is determined based on the decline in demand for various refinery products, including transportation fuels (gasoline, kerosene, diesel), heavy fuel oil, hydrogen, refinery fuel gas, and coke. This change is computed assuming homothetic preferences among consumers; in other words, if global oil demand declines by 1%, gasoline consumption declines by the same fraction.

The third article extensively uses OPGEE to calculate the volumes of natural gas used on-site to generate heat and electricity or (re-)injected. These quantities are then employed as explanatory variables in the second stage of the regression model.

In conclusion, by leveraging the granular data and methodologies provided by OPGEE and PRELIM, the thesis provides a comprehensive analysis of the environmental impact of the oil supply chain, from initial exploration to final product consumption.

Part II

Research Papers

Chapter 4

Carbon implications of marginal oils from market-derived demand shock

This chapter is entirely based on Masnadi, Mohammad S.; Benini, Giacomo; El-Houjeiri, Hassan M.; Milivinti, Alice; Anderson, James E.; Wallington, Timothy J.; De Kleine, Robert; Dotti, Valerio; Jochem, Patrick; Brandt, Adam R. (2021): Carbon implications of marginal oils from market-derived demand shocks. In Nature 599 (7883), pp. 80–86.

The energy sector is in a state of rapid change. Several countries announced a variety of "green" policies to recover from the 2020 COVID-19 downturn. Many of these policies could have a long lasting effect on the oil and gas industry [Kaihan Mintz-Woo and Schinko, 2020, Barbosa et al., 2020]. The industry could enter into an era of declining demand, technology-led supply response, intense competition, investors' scepticism, and increasing public and government pressure regarding impacts of the oil sector on the environment [Barbosa et al., 2020].

Environmental impacts of oil are commonly measured using LCA methods. The life-cycle carbon footprint, or CI, of oil-derived transportation fuels (for example, gasoline) includes the greenhouse gas (GHG) emissions resulting from the combustion of fuels themselves as well as emissions from production and refining of petroleum products. So-called "upstream" emissions from exploration, extraction and transportation of crude oil differ widely between oilfields ($\sim$20–300 kg of CO_2 equivalent per barrel $kgCO_2ebbl^{-1}$ of oil) due to diverse sub-surface geological properties of the deposit, physical and thermodynamic properties of the hydrocarbons, and production and resource management practices [Masnadi et al., 2018]. Similarly, "midstream" emissions from refining vary widely ($\sim$10–60 $kgCO_2ebbl^{-1}$ oil) due to the quality of stream of processed crude and the refining technologies applied [Jing et al., 2020]. These emissions contribute to variability in the life-cycle CI of different crude oil supply chains.

The profitability of crude oil production somewhat mirrors the heterogeneity in GHG emissions. The cost-effectiveness of the upstream sector varies due to the properties of the crude extracted, the marginal production costs, the capacity of the producers to affect the global oil price, and the global oil demand elasticity. Thus, some fields are very profitable, while others barely break even. Recent studies have separately analysed the heterogeneity in the GHG emissions [Masnadi et al., 2018, Jing et al., 2020] and the economics of the oil market [Benini et al., 2023]. However, the interaction between the two remains poorly understood. As a result, the characteristics of marginally economic oilfields are not systematically available. This interaction is important because it affects the magnitude of emissions mitigation potential as less profitable oil producers are displaced when demand declines. The demand drops can be due to socio-economic effects (for example, recessions or the ongoing COVID-19 pandemic), substitution effects (for example, more extensive use of alternative fuels/vehicles), and technological change within the transportation sector (for example, greater fuel efficiency).

In the past decade, development of "consequential" LCA aimed to incorporate numerous economic factors into previously static engineering-based analysis [Wallington et al., 2017, Earles and Halog, 2011, Plevin et al., 2014, Yang and Heijungs, 2018]. These analyses attempt to model income and substitution effects of introducing alternatives, instead of simply assuming that a new product directly displaces an old product. To date, this consequential LCA paradigm has not reached crude oil LCA, and studies of alternatives to crude oil (for example, electric vehicles (EVs)) nearly always assume that an alternative simply displaces average crude oil. The merging of CI and profitability allow us to conduct the first consequential LCA study of the global oil supply (to the best of our knowledge).

The present work connects the upstream CI of 1,933 oilfields (~90% of 2015 global crude production) with their profitability. The CI of fields is calculated using a well-to-refinery estimation tool, which assesses the emissions due to the production of an additional barrel of crude from a particular oilfield. The profitability is calculated using a microeconomic model, which determines how much money a company is willing to pay to manage an additional barrel of crude located in a particular oilfield (see section 4.2). The integration of field-specific CI and profitability allows us to identify the emissions of fields close to the break-even point (extensive margin of the industry). In other words, we isolate the emissions of those fields where the management choice hangs in the balance between "how much should I produce?" and "should I keep producing or cease business operations?" Our results suggest that an environmental policy designed around non-market informed LCA results could ignore first-order effects. In addition, the structure of the global oil market systematically affects the life-cycle benefits from a decline in oil demand. These results could serve public (for example, the US Department of Energy National Energy Modeling System) and private energy system models to better assess the benefits of technological change within the transportation sector.

4.1 Results and Discussion

Country-level Figure 4.1 presents the global map of national volume-weighted average (VWA) marginal production costs (MC) in 2015. The numbers below the name of each country in the map are the corresponding upstream VWA CIs (in $kgCO_2ebbl^{-1}$). The global average MC estimate - shown by the horizontal dashed line in Fig. 4.1 - is ~US\$5.9bbl^{-1} crude oil, with country-level MCs ranging from 2.8 (Iraq (IRQ)) to US\$21.5bbl^{-1}(Columbia (COL)). Fields with the lowest production costs are mainly conventional resources located in the Middle East and North Africa. There is a wide range of production emissions associated with these regions, with routine flaring as the major driver of high CI due to lack of investment/infrastructure for gas handling (see Supplementary Material). Among large producers, Venezuelan, Mexican and Canadian oils are the most expensive and tend to have high production CIs. The US oil industry stands near the global average in terms of GHG emissions and has a high MC (~US\$7.3bbl^{-1}). Note that the dynamics of the emissions presented in Fig. 4.1 can vary over time [Masnadi and Brandt, 2017a,b]. However, due to the fact that substantial change in production strategies takes time, the relative magnitude of the presented emissions can be expected to hold for a short-term outlook of <5–10 years. See Supplementary Material for production economics time-series dissection.

Crude type Table 1 groups field-level results into summary statistics of a set of global crude classes. Heavy fields (most commonly located in Venezuela) and extra heavy fields (mostly located in Canada) are the least profitable fields with relatively high MC and low selling price (due to low API gravity). Oil sands have the lowest selling oil price due to low API gravity and high sulfur content. However, their MCs have substantially decreased in recent years [Bloomberg, 2019, Markit, 2020] making them more competitive vis-à-vis heavy and extra heavy crudes. Contrary to all the other types of crude, oil sands are all located in a single country (Canada). Therefore, they are particularly sensitive to national-specific shocks and transport logistics issues. Shale and tight oil resources are somewhat more competitive, with relatively higher profit margins in all three economic cases, lower emissions, and lighter density crude (higher selling price and lower refining emissions). Conventional light and medium fields are the largest and cheapest to extract crude oil with high selling price, low MC, and relatively low

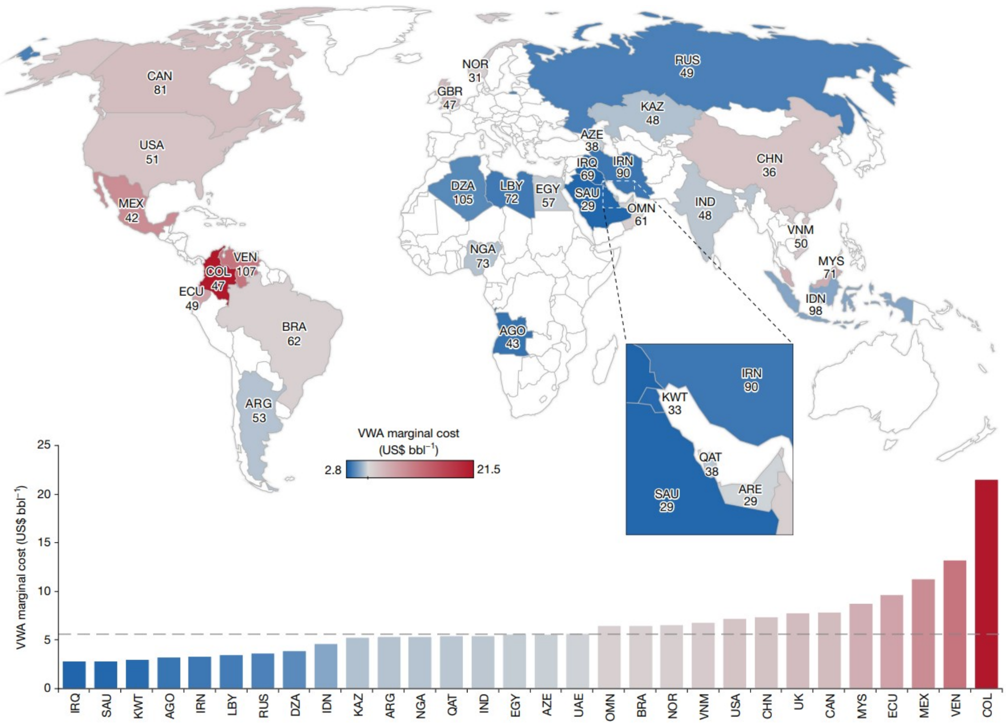

Figure 4.1: Estimated global crude oil upstream marginal cost of production (2015). National volume weighted average (VWA) upstream marginal cost of production in US$bbl^{-1} crude oil produced. Map shows national VWA upstream marginal CI below each country name (in kgCO$_2$ebbl^{-1} crude oil delivered to refinery). The global VWA MC estimate is shown by the horizontal dashed line (∼US$5.6bbl^{-1}). Reference year is 2015. Top 30 global producers are mapped (see Supplementary Data 1 for full list). Countries are named based on their ISO3 code. Colour scheme reflects national VWA MC: dark blue for lowest MC, dark red for highest MC.

Table 1 | 2015 global oilfields characteristics based on crude type

Crude type	Share in global production[a] (%)	Total no. fields	CI[b] (kgCO$_2$ebbl^{-1})	Oil price[b] (US$bbl^{-1})	MC[b] (US$bbl^{-1})	SP-PC[b] (US$bbl^{-1})	SP-oligopoly[b,c] (US$bbl^{-1})	SP-cartel[b,c] (US$bbl^{-1})	API gravity[b] (°API)	Flare-to-oil ratio[b] (scfbbl^{-1})
Light and medium	77.5%	1,259	49.0	51.1	4.2	46.9	42.1	36.0	33.7	154.8
Heavy	8.6%	157	61.0	47.2	16.6	30.6	28.7	25.8	17.2	122.4
Shale and tight oil	7.7%	314	53.4	52.2	6.8	45.5	44.9	44.6	29.8	193.7
Oil sands	2.0%	21	129.2[e]	45.3	6.3	39.0	38.3	38.3	19.3	2.0
Extra heavy	0.5%	9	60.3	50.1	20.2	29.9	28.9	28.9	13.7	32.3
Other oil[d]	3.6%	173	42.3	52.4	5.9	46.6	45.0	45.0	28.2	84.4
Global average	100.0%	1,933	51.9	50.7	5.6	45.1	41.1	36.1	31.4	148.7

[a]Global production covered in this work: 71.0 MMbbld^{-1}.
[b]Volume-weighted average based on field-level share of production. scf, standard cubic foot.
[c]The oil demand elasticity of $\eta=-0.35$ is used for both oligopolistic and cartel competition cases.
[d]Fields that are not characterized in the used dataset and are excluded in the paper discussions.
[e]Research is in progress to re-evaluate the oil sands upstream CI based on regional and real operation emission data (for example, see ref. [61] Sleep).

CI. The average profitability (shadow price (SP)) of different crude types changes accordingly to the assumption on the market structure. In perfect competition (PC), every field is an independent firm, which exerts no market power. In oligopolistic competition (oligopoly), a limited number of firms owns many fields. In cartel competition (cartel), a limited number of firms coordinate their production decisions via a syndicate (for example, the Organization of the Petroleum Exporting Countries (OPEC) - see Methods). In PC, the oilfield SP is the difference between the price at which it sells its output and the MC. The conventional light and medium fields are the most profitable producing units (see SP-PC in Table 1). In oligopoly, the volume of production of the firm, which owns the field, affects its SP. In cartel, the volume of production of OPEC affects the SP of the field's member of the union. As a result, in oligopoly and cartel cases, shutting down or reducing production from individually profitable oilfields is rational, since the firm/cartel will sell less output but at a higher price. As a result, in oligopoly and in cartel cases many light and medium fields owned by large international or national oil companies shift to a least profitable position. Irrespectively of the underlying market structure, heavy fields tend to remain the least profitable formation. Thus, these are the crudes most likely to be displaced by an oil demand reduction. Carbon taxation would also significantly affect their profitability due to their high production CI. Gas management (that is, routine flaring and methane venting and fugitives) is the major CI contributor for light and medium, and shale and tight oil crudes. The profitability of these fields is therefore exposed to gas management regulations (for example, production restriction as imposed in eastern Canada [Canada-Newfoundland and Labrador Offshore Petroleum Board (C-NLOPB), 2017]).

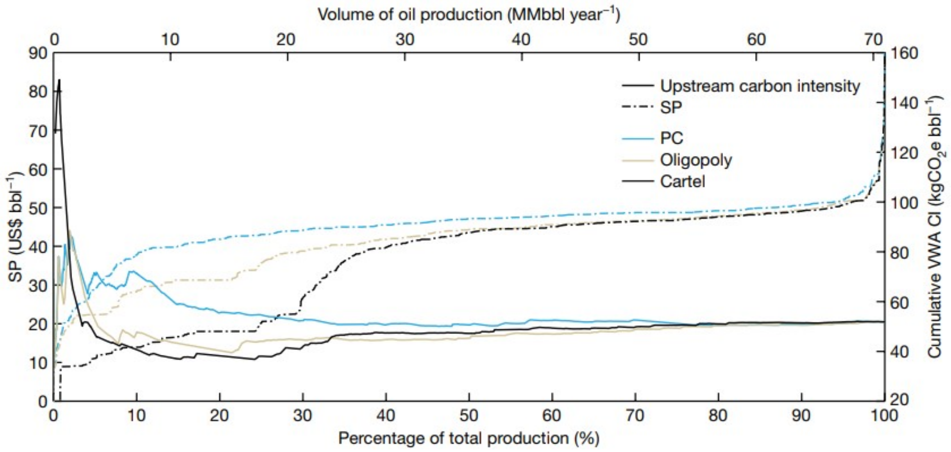

Figure 4.2: Upstream cumulative volume-weighted average CIs (right axis) and sorted SPs (left axis) of global oilfields for PC, oligopoly and cartel economic cases versus the percentage of total oil production in 2015. The oil demand elasticity of $\eta = -0.35$ is used for both oligopoly and cartel competition cases. See Supplementary Material for results variation based on different oil demand elasticities and further discussions.

Field-level To estimate field-level CI (see Methods), we separate the GHG emissions due to the production of the next barrel from the emissions due to the exploration, and drilling and development. The former identifies the environmental footprint linked to the SP of discovered oil (that is, the one identified in the econometric analysis). The latter - exploration and development emissions - are smaller in most cases and are coupled to the SP of undiscovered oil (not included in this work). Next, the computed SPs of discovered oilfields are sorted from smallest to largest (that is, low to high profitability). As a result, we obtain a merit base curve, which links profitability to production CI for the three market structures (PC, oligopoly and cartel - see Methods). Figure 4.2 combines the upstream cumulative VWA CIs (right axis) and the sorted SPs (left axis), against the percentage of total oil production covered in this work. Analogous to the upstream CI, the presented wide range of SPs illustrates heterogeneity of production costs due to diverse operational, physical, chemical, and geological properties of different oilfields. Fields in the highest fifth percentile ($\sim$US$53bbl^{-1}) make over 17% more marginal profit per barrel than the median field ($\sim$US$46bbl^{-1}) for all economic cases. Each local peak along the CI curve in Fig. 4.2 indicates an addition of a field with relatively high CI and production rate compared to the preceding covered fields. For example, the early sharp peaks by using the cartel model (black curve in Fig. 4.2) correspond to Venezuelan heavy fields. Large peaks in cumulative CIs at the beginning (0–20% of total production) imply that many less-economic fields with relatively low SPs also emit high GHG emissions (few exceptions are unprofitable depleted conventional fields with low SP and low emissions). These marginal oilfields are consequently more vulnerable to any future carbon taxation/regulation regime and more likely to be displaced by a demand shift. In all economic cases, the cumulative CI curve trends downward due to covering fewer emitting fields. This trend continues for the PC case until reaching 51.9kgCO_2ebbl^{-1}(at 100% production coverage), which is the global

VWA marginal CI (see Table 1). However, for the other two cases (that is, oligopoly and cartel), including the global oil demand elasticity and market power correction in computing the SP results in less profitability of several low-emitting conventional producers. Thus, for oligopoly and cartel models after few high peaks, the cumulative CI curve first trends descending (conventional with low CI) and later trends ascending (remaining fields with higher CI than conventional closer to margin). See Supplementary Material for additional field-level data.

Displacement implications Many reports estimate near and long-term volume of oil that is going to be displaced and/or stranded by technological developments and/or policy measures [Eberhart, 2018, Kah, 2019, Sharma, 2019, IEA, 2019, Bloomberg, 2020]. These estimates depend on numerous scenario assumptions (for example, growth rate of EVs, global income growth and the way these factors interact) and their conclusions differ markedly. Instead of selecting any one scenario, we create abstract round number shocks to identify the environmental effect resulting from the displacement of the extensive margin of the oil industry. Such shocks might stem from policies to counter climate change, economic slowdowns, geopolitical conflict, or (as the case in 2020) global diseases like COVID-19. We first consider an oil demand reduction of 2.5% relative to the baseline ($\sim$1.8 million barrels per day (MMbbld^{-1})), which we call small shock scenario. Then, we consider a reduction of $\sim$5% ($\sim$3.6MMbbld^{-1}), which we call COVID-19 pandemic scenario due to its resemblance with the contraction in oil demand observed during the 2020 pandemic. Finally, we consider a reduction of $\sim$10% ($\sim$7.1MMbbld^{-1}), which we call medium shock. The latter could result from a vigorous adoption of alternatives or major macroeconomic downturns like a global financial recession. Note that in medium and large demand reduction scenarios (that is, roughly $>$5–10%), only the PC-SP is informative. In the other two cases (that is, oligopoly and cartel), the estimated SP is likely to become uninformative, since the market power correction term would change due to a transformation of the market structure (for example, countries leave/join OPEC, different propensity of countries to respect OPEC quotas, or different outcomes of the game-in-quantities played among oligopolists).

Table 2 | Characteristics of small and COVID-19 shock scenarios for crude oil demand reduction using different economic models

	Perfect competition (PC)			Oligopolistic competition (oligopoly)[b]			Cartel competition (cartel)[b]			Total (reference)
Shock scenario[d]	Small shock, 2.5% reduction	COVID-19 shock, 5% reduction	Medium shock, 10% reduction	Small shock, 2.5% reduction	COVID-19 shock, 5% reduction	Medium shock, 10% reduction	Small shock, 2.5% reduction	COVID-19 shock, 5% reduction	Medium shock, 10% reduction	100% reduction
Displaced volume (MMbbld^{-1})	1.8	3.6	7.1	1.8	3.6	7.1	1.8	3.6	7.1	71.0
Median production (bbld^{-1})	10,000	8027	5452	12,493	21,342	19,370	38,507	68,986	50,000	7,507
MC of production[a] (US\$bbl^{-1})	26.6	21.9	18.1	26.6	15.9	14.1	11.5	10.4	7.0	5.6
SP[a] (US\$bbl^{-1})	18.6	22.5	28.3	16.9	19.7	22.3	2.3	6.5	9.9	45.1/41.1/36.1[c]
CI[a] (kgCO_2ebbl^{-1})	79.7	70.6	71.2	80.3	50	47.8	65	45.8	40.1	51.9
Total no. fields	37	68	276	33	42	79	16	25	39	1,933
Light and medium (vol%)	1	2	15	1	52	49	59	75	87	77.5
Oil sands (vol%)	0	8	11	0	0	4	0	0	0	2
Heavy (vol%)	96	85	64	96	47	44	37	23	11	9
Extra heavy (vol%)	0	3	5	0	0	2	0	0	0	1
Shale and tight oil (vol%)	0	0	2	0	0	0	3	1	1	8
Unconventional total (vol%)	96	96	82	96	47	50	40	24	12	19
Annual upstream mitigation potential (MtCO_2e)	54	92	184	53	69	124	39	61	109	1,343

[a]Volume-weighted average based on field-level share of production.
[b]The oil demand elasticity of $\eta = -0.35$ is used for both oligopolistic and cartel competition cases.
[c]SP-PC, SP-oligopoly, and SP-cartel, respectively.
[d]The accuracy of estimate is likely to decrease for larger demand reduction shocks (for example, 10% reduction) of oligopoly and cartel model results. See the corresponding text for further discussion.

Table 2 characterizes the small, COVID-19, and medium shock scenarios using the different market structures. In PC, the marginal fields are mostly small producers, with median production of $\sim$8,000–10,000bbld^{-1}. The oligopoly and cartel cases shift few large conventional producers close to the industry margins (Table 2 and Supplementary Material). The shift occurs for the same reasons explained above. Namely, oligopolists and members of the cartel adjust production from profitable fields to maximize their total profit. In all three economic models, the VWA MC of the marginal fields is much higher (25–375%) than the global average MC of ∼US$5.6bbl^{-1}. Shifting low carbon intensive light and medium conventional fields towards the margin lowers the average CI of the displaced oil, but several heavy fields stay at the margin. We conclude that oil demand shocks result in non-linear carbon emissions reduction. In all three economic cases, the CI of the crudes displaced by the small shock is ∼25–54% larger than the global average of 51.9kgCO_2ebbl^{-1}. The CI of the displaced crudes by COVID-19 shock is ∼35% larger than the global average for PC, but is close to the global average CI using oligopolistic and cartel competition. The PC model still provides accurate estimates for 10% reduction shock where the CI of the displaced crudes is ∼37% larger than the global average. However, the oligopoly and cartel models might not capture the market behaviour for such a large shock. The average CIs for these two economic cases due to 10% demand reduction are lower than global average with large volume share of light and medium crudes being displaced due to market power considerations. The demand reduction magnitude affects the average CI of displaced crudes. Heavy oilfields with high CI (mostly located in Venezuela) have consistent contribution in all demand reduction scenarios and across all economic models. The total share of unconventional crudes (by volume) generally decreases by including market power corrections in the economic model, as it becomes more viable for large national oil companies to exert market power by reducing production from productive conventional fields. Our results show that given the proposed three economic cases, the small, COVID-19 and 10% reduction shocks in the global oil demand would result in the elimination of 39–54, 61–92 and 109–184MtCO_2e per year of upstream emissions, respectively. The Supplementary Material shows a full range of annual carbon mitigation potential versus the amount of oil displaced using the three economic models. Larger reductions of GHG emissions associated with refining of oil and the final combustion of corresponding products would also occur, but are not included in these calculations. See Supplementary Material for well-to-wheel mitigation potential estimate ranges and further discussions on demand sector GHG emissions. In this work, we only included the production economics and identified the extensive margin of the oil industry. However, various other dynamic forces such as production agreements, region-specific fiscal regimes, regulations (for example, fuel standard policies), geopolitics (for example, sanctions, trade wars), technical advances and incidental events could move a particular oilfield toward or away from the margin. Further analysis of these factors is beyond the scope of this work, but could be pursued in future research (see Supplementary Material).

4.2 Methods

Research scope This work covers upstream emissions (including production and transport of crude oil to refinery gate) and costs. Due to lack of access to refinery cost data, we cannot generate a fully market-informed (consequential) well-to-wheel emissions analysis that goes all the way to refined fuels. Nevertheless, we provide a discussion on how upstream displacement could affect the emissions of the demand side (see Supplementary Material).

Carbon intensity model The field-level CI is estimated using the Oil Production Greenhouse Gas Emissions Estimator (OPGEE version 2.0) [El-Houjeiri et al., 2017, California Air Resources Board, 2017, Group, 2017]. OPGEE is an open-source, peer-reviewed [Masnadi and Brandt, 2017a,b, El-Houjeiri et al., 2017, Vafi and Brandt, 2016, El-Houjeiri et al., 2013, Masnadi et al., 2018, Brandt et al., 2018, 2015, Brandt, 2013, Brandt et al., 2014, Tripathi and Brandt, 2017], bottom-up, engineering-based model. The OPGEE system boundary is "well-to-refinery" (WTR, that is, exploration, drilling and development, production and extraction, surface processing, maintenance, waste disposal, and crude transport to the refinery). Reported emissions are measured in gCO_2e emitted per 1MJ lower heating value (LHV) of crude petroleum delivered to the refinery entrance gate. All GHGs are converted to gCO_2e using AR5 GWP100 conversion factors (without carbon feedback) [Myhre et al., 2013]. See the OPGEE user guide [El-Houjeiri et al., 2017] for more details of each process stage. OPGEE estimates

CI using up to 50 parameters as input data for each modelled oilfield. If input data are not available for some parameters (common), OPGEE supplies defaults based on statistical analysis of petroleum engineering literature and commercial data sources (for example Oil Gas Journal (O&GJ) [PenWell Corporation, 2015]) enabling the software to estimate a field's CI without complete data [El-Houjeiri et al., 2017, PenWell Corporation, 2015]. In this work, field exploration, and drilling and development emissions are excluded from CIs reported in prior work citepMasnadi2018 to estimate GHG emissions associated with production of the next barrel of crude oil (that is, marginal upstream CIs). These two sectors hold a very low share of the total upstream GHG emissions (see supplementary Fig. S20 of ref. [Masnadi et al., 2018]).

Global oilfields In the previous work1, CIs were estimated for 8,966 global active oilfields (so-called child fields) supplying 78.9 million barrels per day (MMbbld1), and capturing ∼98% of 2015 global crude oil and condensate production [EIA, 2017]. A combination of government reported data (Norway [NPD, 2017, Gass, 2015], Canada [Alberta Energy Regulator, 2015, Canada-Newfoundland and Labrador Offshore Petroleum Board (C-NLOPB), 2021, , NEB, Natural Resources Canada (NRCan), 2017], Denmark [Danish Energy Agency, 2016], UK [UK Government, 2017], Nigeria [Corporation, 2015], and US California [State of California Department of Conservation, Division of Oil, Gas, & Geothermal Resources, 2015], US Alaska [Alaska Department of Administration, 2017] and US shale oils [EIA, 2021a]), public literature (total of nearly 800 sources) and proprietary/commercial data sources (O&GJ 2015 survey [PenWell Corporation, 2015] and Wood Mackenzie (WM) oilfield datasets [Mackenzie, 2018]) were used as input data [Masnadi et al., 2018]. Government and public literature data were collected and used for 1,009 global fields, accounting for about 64.3% of global crude oil production. Commercial data are utilized for the remainder (mostly small fields). We select 2015 as the reference year due to lags in some data sources. See our previous study supplementary materials document for further details.

Economic model We frame our economic model as a profit-maximization problem. We study three different cases described here heuristically with mathematical details presented in the Supplementary Material. In the perfect competition (PC) case, every field is an independent firm which exerts no market power. In this context, the field management solves the profit-maximization problem taking the oil price as given. In the oligopolistic competition (oligopoly) case, a limited number of firms own many fields. In this context, the field management solves the profit-maximization problem knowing that the quantity of oil produced by the firm who owns the field as well as its competitors influences the oil price. In the cartel competition (cartel) case, a limited number of firms coordinate their production decisions via a syndicate. In this context, the field management solves the profit-maximization problem knowing that the quantity of oil produced by the members of the syndicate influences the oil price. Said differently, in the oligopoly case a small number of oligopolistic competitors play a game-in-quantities. In the cartel case, a few firms work together to coordinate their production decisions around a union. In our model, the members of the cartel are the national oil companies associated with OPEC. Due to the complexities in modelling the realities of cartel dynamics, our cartel case assumes that the cartel operates in unison. The effect of a cartel with imperfect coordination would fall somewhere between individual company market power (oligopoly case) and the perfect cartel (cartel case). In the PC case, field profits are the difference between field revenues and field costs, which we divide into two macro-classes: (1) costs to extract the oil (extraction costs) and (2) costs to discover new oil (exploration costs),

$$\text{Profits} = (\text{oil price} \cdot \text{volumes of oil extracted}) - \text{extraction costs} - \text{exploration costs}\,.$$

In the oligopoly and cartel cases, the field profits are the same. However, in the oligopoly case the management takes into consideration the effect of the volumes of oil produced by the firm who owns the field on the oil price, while in the cartel case the management takes into consideration the effect of the volumes of oil produce by the cartel on the oil price.

In all three cases, the decision choices are: what volume of oil to extract and how much money to spend in exploration [Devarajan and Fisher, 1982, Pindyck, 1978]. While making these decisions the management faces two physical constraints. First, the quantity of reserves available at time t equals the reserves at time $t-1$ minus the volumes of oil extracted at time t plus the quantity of oil discovered at time t. Second, the cumulative discoveries at time t equals the cumulative discoveries until time $t-1$ plus the discoveries at time t. The first-order condition of the optimization problem with respect to the volumes of oil extracted identifies how much money a producer is willing to spend to manage one extra barrel of oil. This value is called shadow price (of discovered oil),

$$\text{Shadow price} = \text{oil price} - \text{marginal extraction cost} + \text{market power correction term} .$$

The shadow price (SP) equals the difference between the oil price and the marginal extraction cost (MC) [Pesaran, 1990] (that is, the cost of extracting the next barrel; this quantity is obtained by taking the first-order derivative of the extraction costs with respect to volumes of oil extracted) readjusted by a market power correction term. If every field is an independent firm with no capacity to influence price (that is, PC), the market power correction term shrinks to zero and the SP becomes the difference between the oil price and the MC. For example, if a field sells its output at US\$50bbl^{-1}and its MC is US\$40bbl^{-1}, the owner of the field is willing to spend (up to) US\$10 to manage one more barrel located in that particular deposit. In the case of oligopolistic competition/perfect collusion behaviour, the SP takes into account the capacity of the firm/cartel to influence the global (average) oil price rescaled by the propensity of consumers to decrease the quantity of oil consumed due to an increase in oil price. Section 1 of the Supplementary Material provides the mathematical details of the economic framework linking the concept of SP to standard oil economics. As the SP of a field approaches zero, the management problem shifts from 'how much should I produce?' (intensive margin choice) to 'should I produce or not?' (extensive margin choice). In other words, the fields with a SP close to zero identify the extensive margin of the oil industry. The emissions of this portion of the industry are the most sensitive to a drop in oil price caused by a reduction in the transportation fuel demand. Note that estimating field-level gross profit was the main aim of this work, not the net profit. The gross profit is a better representative of fields' geological and physical characteristics and production practices, whereas the net profit includes additional fiscal regimes (that is, royalties, severance taxes, income taxes, production sharing and so on), which are complex, country/region-specific and subject to change. Incorporating these fiscal terms is out of the capacity and the scope of the presented work.

Econometric analysis All three variables making up SP are unobserved. To estimate them, we face three econometric problems: (1) the non-stationary nature of oil prices, (2) the endogenous link between costs, quantities and reserves, and (3) the uncertainty about the magnitude of the oil demand elasticity. We do not know the price at which a field sells its output because we do not have access to commercial agreements between oil producers and oil refiners. However, we know the prices of publicly traded oil classes. More precisely, we know the landed costs of imported crudes in the United States from 1979 to 2018 [EIA, 2021c], as well as some key physical and chemical characteristics of every traded class [of Canada , PSA] (see Supplementary Fig. 1 and Supplementary Table 1. In the same way, we know the average price at which US refineries buy imported crudes [EIA, 2021c] and the average physical and chemical characteristics of crudes imported in the United States [EIA, 2021b]. The physical and chemical characteristics most important to refineries are the crude density (measured as API gravity) and the sulfur content (measured as wt% sulfur). We regress the difference between the price of a particular oil class and the average price at which refineries buy imported crudes on the differences between the API gravity of the oil class and the average API gravity of imported crudes as well as on the difference between the sulfur content of the oil class and the average sulfur content of imported crudes [Kilian and Murphy, 2012, Fattouh, 2010, Bacon and Tordo, 2005]. In doing so, we solve the non-stationarity problem while assuming that the difference between the price of a particular oil class and the average one is a linear function of the oil's characteristics. We allow these linear deviations to be time-specific to adjust for changing in demand of transportation fuels as well as for technological change within the refinery sector. For example, in 2015 the average oil price was US$50.39bbl^{-1}, its API gravity 31.46, and its sulfur content 1.40%. In 2015 increasing the API gravity by one degree increased the value of a crude stream by US$0.13bbl^{-1}, while increasing sulfur content by 1% lowered the value of a crude stream by US$2.86bbl^{-1}. In 2016, these two quantities were +US$0.03bbl^{-1}and US$0.85bbl^{-1}. This change could be due to (1) a modification in the composition of the demand for transportation fuels (for example, more demand for gasoline, less demand for diesel), (2) a change in the technologies employed by US refineries, and (3) a combination of (1) and (2). Our econometric model is flexible enough to take into account all three possibilities (see paragraph Firm expected price of section 4.3). We can use the two structural coefficients, which weight the impact of API gravity and sulfur content, to estimate field-level selling prices (see Supplementary Material section 2.1, equation 8). Using the API gravity and sulfur content reported in the 2018 WM dataset47, we estimate the selling price of 1,933 'parent project' fields over the decade 2009–2018, thereby obtaining 1,933×10=19,330 simulated selling prices (see Supplementary Material section 2.1 for a detailed discussion on the results). See Supplementary Figs. 2 and 3 for a cross-sectional snapshot.

Next, we estimate the MC. An accurate measurement of MC is complicated because it is difficult to determine which factors of production are fixed and which are variable. However, the use of detailed accounting data, combined with standard econometric techniques, allows us to have a good first-order approximation of the MC of different types of fields. We use the WM dataset to obtain yearly cost data for the same 1,933 fields over the time interval 2009–2018. Then, we obtain the extraction costs summing the operational expenditures (OPEX), which include consumable inputs, labour, maintenance, repairs, accounting costs, license fees, office expenses, utilities and insurance. We also include capital expenditures not linked to exploration activities (non-exploration CAPEX, which include installation, acquisition, upgrading and restoring of the physical assets used to extract the oil). After computing the extraction costs, we regress them against the volumes of oil extracted while controlling for the depletion level of the field, the geological characteristics of the field, and technological trends in the broader oil industry. We estimate the structural coefficients of the cost function re-expressing the regression in first differences. The combination of the longitudinal structure of the dataset with the first-difference estimation method allows us to attenuate (or in the best case scenario to solve) eventual endogeneity problems. The first-order derivative of the fit returns the estimated MC. Section 4.3 provides all the econometric details. Finally, for two of the three cases analysed, we compute the market power correction term. Its expression is the same in both cases. Namely, the capacity of the firm/cartel to influence the average oil price rescaled by the capacity of consumers to lower their demand for oil-derived products when their prices increase. Said differently, the market power correction term adjusts the SP of every field by capturing the effect of a unit increase in the production of a specific field on the equilibrium oil price and, in turn, on the firm's profits. Higher market power — corresponding to larger firm size — implies, ceteris paribus, a lower shadow price, because the effect of a fall in price due to the production of an extra unit of crude on the firm total revenues is proportional to the total production. For instance, if the production of an extra barrel causes the oil price to fall by 0.01 cents, then the firm must trade-off the profits generated by selling that extra barrel and a loss of 0.01 cents per barrel times the total number of barrels produced by the firm. Thus, accounting for market power makes the marginal unit produced by each firm/cartel less valuable, resulting in lower SPs. This effect is increasing in the firm/cartel size. Since the magnitude of the oil demand elasticity is object of econometric debate, we validate our results using different point estimates within the interval 0.20 up to 0.35.

Data matching and coverage The previous work on the CI of global oilfields is provided at a child field level. Child fields are individual discoveries that are part of a parent project. Parent fields are combinations of geologic deposits collected for the purposes of a combined valuation. The linkage with the economic data, available only at parent level, requires us to match the child field CIs to parent fields. The majority of the child non-technical oil-fields from WM datasets (accessed 2018) — whose corresponding parent fields are available — directly matched with the OPGEE global dataset. We paired the remaining with smart string search and string distance matching using R as well as manual matching for the countries with poor total production coverage. Finally, we conduct an additional treatment on two important global producers (Canada and United States) based on the available data (see Supplementary Material section 2.2). After the matching process is completed, we examine the representativeness of our techno-economic dataset. In total, we matched 1,933 parent fields located in 77 countries. Their combined production is ∼71MMbbld1 and it captures ∼90% of the 2015 global crude oil and condensate production. Supplementary Table 5 returns the coverage summary of the top 20 largest global producers, and Supplementary Fig. 4 zooms in on the geographic location and the CI of the mapped fields.

4.3 Supplementary Material

Theoretical economic model

We assume K oil firms (or cartels of firms) competing in quantities. Each firm $k = 1, 2, ..., K$ controls I^K oil fields and maximizes the present discounted value of the sum of present and future profits. Each firm k anticipates the effect of its production choices on the equilibrium market prices. The target oil price $\overline{P}_{t+s}$ is determined by the global demand for oil, which has the isoelastic form $q^W_{t+s} = \left(\overline{P}_{t+s}\right)^\eta \times h_{t+s}\left(Y_{t+s}\right)$, where η is the price elasticity of demand, Y_{t+s} is a vector of observables affecting aggregate demand (e.g., World GDP), and h_{t+s} is a time-specific function. Let Q^W_{t+s} denote the global oil supply in period $t + s$. We require the oil market to clear in each period; i.e., $q^W_{t+s} = Q^W_{t+s}$ for all $s = 0, 1, 2, ...$ Note that for $|\eta| \to \infty$ the equilibrium price is constant in q^W_{t+s} and all firms behave as price-takers; i.e., the oil market is perfectly competitive in that case. Conversely, for finite values of η firms enjoy some market power. The firm decides in period t its production and investment plan for all periods $t + s$ with $s = 0, 1, 2, ...$ In other words, the firm commits in period t to its future production and investment plans. While this assumption is admittedly unrealistic, it is often imposed in this class of models because it eases the derivation and interpretation of the results while having negligible consequences on the implications on the analysis.

The firm k's infra-temporal profits in each period $t + s$,

$$\Pi^k_{t+s} = \sum_{i=1}^{I^k} \Pi^{k,i}_{t+s} = \sum_{i=1}^{I^k} P^{k,i}_{t+s} Q^{k,i}_{t+s} - C^{k,i}_{t+s}(Q^{k,i}_{t+s}, R^{k,i}_{t+s-1}, \epsilon^{k,i}_{t+s}) - W^{k,i}_{t+s} \,, \tag{4.1}$$

are the difference between the firm's total revenues and the firm's total costs. The field-level profits are the product between the price at which field i sells its output $P^{k,i}_{t+s}$ and the quantity of the output produced $Q^{k,i}_{t+s}$. The field-level costs are divided into two macro-classes. The first ones are production costs. The second ones are discovery costs. Production costs are function of the quantity of oil extracted and of the amount of reserves available when the production decision starts. The latter are equivalent to the initial size of the deposit $R^{k,i}$ plus the discoveries occurred after the initial assessment of the field $L^{k,i}_{t+s-1} = L^{k,i} + \sum_{r=1}^{t+s-1} D^{k,i}_r$, where $D^{k,i}_r$ are the new discoveries in period r, minus the sum of extracted liquids $B^{k,i}_{t+s-1} = B^{k,i} + \sum_{r=1}^{t+s-1} Q^{k,i}_r$, such that $R^{k,i}_{t+s-1} = R^{k,i} + L^{k,i}_{t+s-1} - B^{k,i}_{t+s-1}$. Finally, the extraction costs are function of an idiosyncratic shock $\epsilon^{k,i}_{t+s}$, which can randomly increase or decrease the extraction costs due to unexpected events. The exploration costs, $W^{k,i}_{t+s}$, are the expenses incurred to discover new oil located in field i. They equal the product between the number of spud wells and the per-well cost.

Every firm maximizes (4.1) while facing two physical constraints for each controlled field $i = 1, 2, ..., I^k$. The first one,

$$L^{k,i}_{t+s} = L^{k,i}_{t+s-1} + D^{k,i}_{t+s}(W^{k,i}_{t+s}, L^{k,i}_{t+s-1}, \xi^{k,i}_{t+s}) \,, \tag{4.2}$$

makes the cumulative discoveries obtained till time $t+s$ equal to the cumulative discoveries obtained till time $t+s-1$ plus the ones obtained at time $t+s$, denoted by $D^{k,i}_{t+s}(.)$. The amount of new discoveries is a function of the exploration costs, of the cumulative amount of past discoveries, and of an idiosyncratic shock $\xi^{k,i}_{t+s}$, which can randomly increase or decrease the volumes of discovered reserves.

The second constraint ensures that the reserves available at time $t+s$ equal the ones at time $t+s-1$, plus the new discoveries, minus the quantity of oil extracted, plus eventual idiosyncratic revision-extensions of previously discovered reserves, $\nu^{k,i}_{t+s}$, such that

$$R^{k,i}_{t+s} = R^{k,i}_{t+s-1} + D^{k,i}_{t+s} - Q^{k,i}_{t+s} + \nu^{k,i}_{t+s} \,. \tag{4.3}$$

The firm k solves the optimization problem consulting an information set, Ω^k_{t+s-1}, which includes previous prices, quantities and shocks,

$$\Omega^k_{t+s-1} = \left\{ \left[P^{k,i}_s\right]^{t-1}_{s=0},\ \left[Q^{k,i}_s, W^{k,i}_s, R^{k,i}_s, L^{k,i}_s\right]^{t-1}_{s=0},\ \left[\epsilon^{k,i}_s, \xi^{k,i}_s, \nu^{k,i}_s\right]^{t-1}_{s=0} \right\}^{I^k}_{i=1} \,.$$

Cost and discovery function

We make the extraction costs function of the volumes of oil produced and of the volumes of reserves available when the production decision is taken.

The link between extraction costs and volumes of production reflects the convex nature of the field's costs. In other words, fields increase their costs as they increase production. Furthermore, the more the production level is closed to the peak capacity, the more extracting the next barrel becomes costly[1]. The link between extraction costs and volumes of reserves reflects the role played by the reservoir pressure in the production process. If a reservoir contains low viscosity oil, trapped in impermeable rocks (a.k.a. Shale & Tight Oil), the wells fracture the oil-containing rocks to allow the natural pressure of the reservoir to lift above ground a mixture of oil, water and stones. In the same way, if a reservoir contains low (a.k.a. Light & Medium Oil) or high (a.k.a. Heavy & Extra Heavy Oil) viscosity oil, trapped in permeable rocks, wells drill vertically to reach the deposit. Once the wells reach the petroleum liquids, the natural pressure lifts the oil above ground. When the pressure declines, the management needs to inject increasing amounts of water and/or of steam to keep the volumes of production constant. In other words, there is an inverse relation between the costs of production and the reservoir pressure in all types of oil formation, with the exception of oil Sands. Capturing this reality would require to collect information about the volumes of water injected, the volumes of steam injected, the injection pressure, the water-oil-ratio, the steam-oil-ratio, and the water-injection ratio, possibly starting from well-level data. To the best of our knowledge no such data are available on a global scale. Therefore, we follow a long-standing micro-econometric tradition and use the volumes of available reserves as a first-order approximation of all the mentioned variables under the general assumption that more reserves equal more pressure and those lower marginal costs.

We capture the two previous intuitions rewriting the cost function used in [Pesaran, 1990] in panel data form,

$$C^{k,i}_t = \theta^{Geo}_0 + \theta^{Geo}_1 Q^{k,i}_t + \frac{\theta^{Geo}_2}{2} Q^{k,i\ 2}_t + \frac{\theta^{Geo}_3}{2} \frac{Q^{k,i\ 2}_t}{R^{k,i}_{t-1}} + \epsilon^{k,i}_t \,. \tag{4.4}$$

In equation (5.17), the dependent variable $C^{k,i}_t$ equals the sum of Operating (OPEX) and of Capital Expenditures not linked to exploration (Non Exp CAPEX)[2] measured in Million US Dollars (MM \$) spend per Year. The explanatory variable $Q^{k,i}_t$ equals the amount of output produced, measured in Million Barrels of Oil Equivalent

[1] If this assumption is not respected, the optimal production level is the peak capacity of the field. An explanatory analysis of our dataset suggests that this is never the case for the analyzed sample, with a median distance between the actual level of production and the peak production capacity of 11,510 barrels per day and an average distance of 46,849 barrels per day.

[2] OPEX includes expenditures like accounting, license fees, maintenance, repairs, office expenses, utilities and insurance, while the CAPEX expenditures (not linked to the discovery process) comprehend the installation, acquisition, repairing, upgrading and restoring of the physical assets used to extract oil.

(MM BOE) extracted per Year[3], while $R_{t-1}^{k,i}$ is the volume of recoverable reserves, measured in MM BOE. The idiosyncratic shock is normally distributed with finite homoskedastic variance, $\epsilon_t^{k,i} \overset{iid}{\sim} \mathcal{N}(0, \sigma_\epsilon^2)$. Finally, θ_0^{Geo} identifies the fixed costs, while $(\theta_1^{Geo}, \theta_2^{Geo}, \theta_3^{Geo})$ identify the variable production costs, rescaled by the volumes of available reserves. All the coefficients are expected to be positive. θ_0^{Geo} is expected to be positive because fields experience positive costs even when the production level is zero such as license fees and insurances. θ_1^{Geo} is expected to be positive because increasing production increases costs by increasing expenses such as utilities. In the same way, $(\theta_2^{Geo}, \theta_3^{Geo})$ are expected to be positive because increasing production is marginally more expensive, irrespective of the volumes of reserves.

We make the volumes of discoveries function a quadratic function of current exploration expenditures $W_t^{k,i}$ and of cumulated past discoveries $L_{t-1}^{k,i}$, and of the idiosyncratic shock $\xi_t^{k,i}$,

$$D_t^{k,i} = \gamma_1 W_t^{k,i} + \gamma_2 W_t^{k,i\,2} + \gamma_3 L_{t-1}^{k,i} + \gamma_4 L_{t-1}^{k,i\,2} + \xi_t^{k,i} \,. \tag{4.5}$$

In equation (4.5), $D_t^{k,i}$ is measured in MM BOE discovered in one Year, $W_t^{k,i}$ is the Exploration CAPEX measured in MM \$ spent per Year, while $L_{t-1}^{k,i}$ equals the sum of past findings measured in MM BOE. γ_1 and γ_2 identify the link between exploration expenditures and amount of discoveries. γ_1 is expect to be positive, since the more a field invests in exploration the more it discovers new oil. γ_2 is expected to be negative since marginal discoveries are declining in exploration CAPEX. γ_3 is expected to be negative since the more oil has been discovered in a field the less likely is to find new one. γ_4 is expected to be negative since marginal discoveries are decreasing in cumulative past levels of discoveries.

Optimization

The firm's optimization problem,

$$\max_{\substack{\{\{Q_{t+s}^{k,i}, W_{t+s}^{k,i}, R_{t+s}^{k,i},\\ L_{t+s}^{k,i}\}_{i=1}^{I^k}\}_{s=0}^{\infty} \in X}} \mathbb{E}_{t-1}\left\{ \sum_{i=1}^{I^k} \sum_{s=0}^{\infty} \kappa^s \Pi_{t+s}^{k,i}(P_{t+s}^{k,i}, Q_{t+s}^{k,i}, W_{t+s}^{k,i}, R_{t+s-1}^{k,i}, L_{t+s-1}^{k,i}, \epsilon_{t+s}^{k,i}) \big| \Omega_{t-1}^k \right\},$$

$$s.t. \left\{ \left\{ \begin{array}{c} D_{t+s}^{k,i} - Q_{t+s}^{k,i} + \nu_{t+s}^{k,i} - R_{t+s}^{k,i} + R_{t+s-1}^{k,i} = 0 \\ L_{t+s}^{k,i} - L_{t+s-1}^{k,i} - D_{t+s}^{k,i} = 0 \\ Q_{t+s}^{k,i} \geq 0;\ W_{t+s}^{k,i} \geq 0 \\ R_{t+s}^{k,i} \geq 0;\ L_{t+s}^{k,i} \geq 0 \end{array} \right\}_{i=1}^{I^k} \right\}_{s=0}^{\infty},$$

where $0 \leq \kappa < 1$ is the inter-temporal discount factor, can be solved using standard methods[4]. The functional form of the cost and of the discovery function are those in (5.17) and (4.5), respectively.

We need to show that the (typically unique) solution to the FOCs of the firm optimization problem is a global maximum under the restrictions on parameters $\gamma_2 \leq 0$, $\gamma_4 \leq 0$, $\theta_2^{Geo} \geq 0$, $\theta_3^{Geo} \geq 0$, and $|\eta| \geq 0.2$. First, we perform variable change by substituting $R_{t+s}^{i,k}$ with the previously defined variable $B_{t+s}^{i,k}$ using the formula $B_{t+s}^{k,i} = R^{k,i} + L_{t+s}^{k,i} - R_{t+s}^{k,i}$. The cost function becomes

$$\widetilde{C}_t^{k,i} = \theta_0^{Geo} + \theta_1^{Geo} Q_t^{k,i} + \frac{\theta_2^{Geo}}{2} Q_t^{k,i} + \frac{\theta_3^{Geo}}{2} \frac{Q_t^{k,i}}{R^{k,i} - B_{t-1}^{k,i} + L_{t-1}^{k,i}} + \epsilon_t^{k,i}$$

[3] The decision to use BOE, rather than the traditional Barrel (B), allows to sum the production of condensate, gas, natural gas liquids (NGL) and oil, so to compare the marginal costs of fields with a different composition of the output. For example, the BOE allows to confront the marginal costs of Sands formations which produce almost only oil with the one of Shale & Tight accumulations which produce considerable quantities of NGL and associated gases.

[4] Note that we are not explicitly accounting the presence of a natural capacity limit for each oil field. This assumption is mostly innocuous if production costs are sufficiently convex, such that marginal production costs become large in the proximity of the capacity limit and the optimal production level is always lower than its natural upper bound as shown in Footnote (1).

and the problem gains the constraints $R^{k,i} + L^{k,i}_{t+s} - B^{k,i}_{t+s} \geq 0$ for $s = 0, 1, 2, ...$ (instead of $R^{k,i}_{t+s} \geq 0$). Thus, the firm's optimization problem becomes:

$$\max_{\{\{Q^{k,i}_{t+s}, W^{k,i}_{t+s}, B^{k,i}_{t+s}, L^{k,i}_{t+s}\}_{i=1}^{I^k}\}_{s=0}^{\infty}} \mathbb{E}_{t-1}\left\{ \sum_{i=1}^{I^k}\sum_{s=0}^{\infty} \kappa^s \widetilde{\Pi}^{k,i}_{t+s}(P^{k,i}_{t+s}, Q^{k,i}_{t+s}, W^{k,i}_{t+s}, B^{k,i}_{t+s-1}, L^{k,i}_{t+s-1}, \epsilon^{k,i}_{t+s}) \middle| \Omega^k_{t-1} \right\}$$

$$s.t. \left\{ \left\{ \begin{array}{c} -B^{k,i}_{t+s} + B^{k,i}_{t+s-1} + Q^{k,i}_{t+s} \leq 0 \\ L^{k,i}_{t+s} - L^{k,i}_{t+s-1} - D^{k,i}_{t+s} \leq 0 \\ Q^{k,i}_{t+s} \geq 0 \\ W^{k,i}_{t+s} \geq 0 \\ R^{k,i} - B^{k,i}_{t+s} + L^{k,i}_{t+s} \geq 0 \\ L^{k,i}_{t+s} \geq 0 \end{array} \right\}_{i=1}^{I^k} \right\}_{s=0}^{\infty}.$$

where the function $\widetilde{\Pi}^{k,i}_{t+s}(P^{k,i}_{t+s}, Q^{k,i}_{t+s}, W^{k,i}_{t+s}, B^{k,i}_{t+s-1}, L^{k,i}_{t+s-1}, \epsilon^{k,i}_{t+s})$ is simply equal to $\Pi^{k,i}_{t+s}(P^{k,i}_{t+s}, Q^{k,i}_{t+s}, W^{k,i}_{t+s}, (R^{k,i} + L^{k,i}_{t+s-1} - B^{k,i}_{t+s-1}), L^{k,i}_{t+s-1}, \epsilon^{k,i}_{t+s})$. The Lagrangian of the optimization problem is

$$\begin{aligned} \mathcal{L}^k_t = \mathbb{E}_{t-1}\Bigg\{ \sum_{i=1}^{I^k}\sum_{s=0}^{\infty} \kappa^s \widetilde{\Pi}^{k,i}_{t+s} &- \lambda^{k,i}_{t+s}\left[B^{k,i}_{t+s-1} - B^{k,i}_{t+s} + Q^{k,i}_{t+s} - \nu^{k,i}_{t+s}\right] + \\ &- \mu^{k,i}_{t+s}\left[L^{k,i}_{t+s} - D^{k,i}_{t+s}\left(W^{k,i}_{t+s}, L^{k,i}_{t+s-1}, \xi^{k,i}_{t+s}\right) - L^{k,i}_{t+s-1}\right] + \\ &- \phi^{k,i}_{t+s}\left[-Q^{k,i}_{t+s}\right] - \chi^{k,i}_{t+s}\left[-W^{k,i}_{t+s}\right] + \\ &- \psi^{k,i}_{t+s}\left[-R^{k,i} + B^{k,i}_{t+s} - L^{k,i}_{t+s}\right] - \zeta^{k,i}_{t+s}\left[-L^{k,i}_{t+s}\right] \left|\Omega^k_{t-1}\right.\Bigg\}. \end{aligned}$$

Define the vector of constraints

$$\mathbf{g}^k_t\left(\left\{\left\{Q^{k,i}_{t+s}, W^{k,i}_{t+s}, L^{k,i}_{t+s-1}, B^{k,i}_{t+s-1}\right\}_{i=1}^{I^k}\right\}_{s=0}^{\infty}\right) =$$

$$\begin{pmatrix} g^{k,1}_{t,1}(W^{k,1}_t, L^{k,1}_t, L^{k,1}_{t-1}, \xi^{k,1}_t) \\ g^{k,1}_{t,2}(Q^{k,1}_t, B^{k,1}_t, B^{k,1}_{t-1}, W^{k,1}_t, L^{k,1}_{t-1}, \xi^{k,1}_t, \nu^{k,1}_t) \\ g^i_{t+1,3}(Q^i_{t+1}) \\ g^i_{t+1,4}(W^i_{t+1}) \\ g^i_{t+1,5}(B^i_{t+1}, L_{t+1}) \\ g^{k,1}_{t,6}(L^{k,1}_t) \\ g^{k,2}_{t,1}(W^{k,2}_t, L^{k,2}_t, L^{k,2}_{t-1}, \xi^{k,2}_t) \\ \vdots \\ g^{k,I^k}_{t,2}(Q^{k,1}_t, B^{k,1}_t, B^{k,1}_{t-1}, W^{k,1}_t, L^{k,1}_{t-1}, \xi^{k,1}_t, \nu^{k,1}_t) \\ \vdots \\ g^{k,1}_{t+1,1}(W^i_{t+1}, L^i_{t+1}, L^i_t, \xi^i_{t+1}) \\ \vdots \\ g^{k,I^k}_{t+s,6}(L^{k,I^k}_{t+s}) \\ \vdots \end{pmatrix} \leq \mathbf{0}$$

where each element is defined as:

$$\begin{aligned} g^{k,i}_{t+s,1}(W^{k,i}_{t+s}, L^{k,i}_{t+s}, L^{k,i}_{t+s-1}, \xi^{k,i}_{t+s}) &= -D^{k,i}_{t+s}(W^{k,i}_{t+s}, L^{k,i}_{t+s-1}, \xi^{k,i}_{t+s}) + L^{k,i}_{t+s} - L^{k,i}_{t+s-1} \\ g^{k,i}_{t+s,2}(Q^{k,i}_{t+s}, B^{k,i}_{t+s}, B^{k,i}_{t+s-1}, \nu^{k,i}_{t+s}) &= -B^{k,i}_{t+s} + B^{k,i}_{t+s-1} + Q^{k,i}_{t+s} - \nu^{k,i}_{t+s} \\ g^{k,i}_{t+s,3}(Q^{k,i}_{t+s}) &= -Q^{k,i}_{t+s} \\ g^{k,i}_{t+s,4}(W^{k,i}_{t+s}) &= -W^{k,i}_{t+s} \\ g^{k,i}_{t+s,5}(B^{k,i}_{t+s}, L^{k,i}_{t+s}) &= -R^{k,i} - L^{k,i}_{t+s} + B^{k,i}_{t+s} \\ g^{k,i}_{t+s,6}(L^{k,i}_{t+s}) &= -L^{k,i}_{t+s} \end{aligned}$$

for all $s = 0, 1, 2, 3, ...$ and $i = 1, 2, ..., I^k$. The most common approach is that of checking the second-order sufficient (necessary) conditions. This consists in verifying whether the bordered Hessian

$$\begin{bmatrix} 0 & D\mathbf{g}^k_t \\ D\mathbf{g}^k_t & D^2\mathcal{L}^k_t \end{bmatrix}$$

is positive definite (positive semidefinite). Due to the complexity of checking such conditions, we follow a different approach. Namely, we show that this is a convex optimization problem under the stated restrictions on parameters. We follow a two steps procedure:

1. We prove that the objective function $OF^k_t = \mathbb{E}_{t-1}\left\{ \sum_{i=1}^{I^k} \sum_{s=0}^{\infty} \kappa^s \widetilde{\Pi}^{k,i}_{t+s} | \Omega^k_{t-1} \right\}$ is concave;
2. We prove that the feasible set $\left\{ x \in X \mid \mathbf{g}^k_t \left(\left\{ \left\{ Q^{k,i}_{t+s}, W^{k,i}_{t+s}, L^{k,i}_{t+s-1}, B^{k,i}_{t+s-1} \right\}_{i=1}^{I^k} \right\}_{s=0}^{\infty} \right) \leq \mathbf{0} \right\}$ is a convex set.

This approach has the further advantage of ensuring that the solution to the FOCs is globally optimal.

Step 1. Define the functions TR^k_{t+s} and $CF^{k,i}_t$ as follows: $TR^k_{t+s} = \mathbb{E}_{t-1}\left[\sum_{i=1}^{I^k} P^{k,i}_{t+s} Q^{k,i}_{t+s} | \Omega^k_{t-1}\right]$ and $CF^{k,i}_t = -\mathbb{E}_{t-1}\left[\sum_{s=0}^{\infty} \widetilde{C}^{k,i}_{t+s}(Q^{k,i}_{t+s}, B^{k,i}_{t+s-1}, \epsilon^{k,i}_{t+s}) + W^{k,i}_{t+s} | \Omega^k_{t-1}\right]$, and note that $OF^k_t = \sum_{s=0}^{\infty} TR^k_{t+s} + \sum_{i=1}^{I^k} CF^{k,i}_t$. Because the sum of concave functions is itself concave, for the function OF^k_t to be concave in the choice variables

$$\left\{ \left\{ Q^{k,i}_{t+s}, W^{k,i}_{t+s}, L^{k,i}_{t+s-1}, B^{k,i}_{t+s-1} \right\}_{i=1}^{I^k} \right\}_{s=0}^{\infty}$$

it is sufficient to show that each function in the collection

$$\left\{ \left\{ TR^k_{t+s} \right\}_{s=0}^{\infty} ; \left\{ CF^{k,i}_t \right\}_{i=1}^{I^k} \right\}$$

is concave in the same variables.

Step 1a. Concavity of $TR^k_{t+s} = \mathbb{E}_{t-1}\left[\sum_{i=1}^{I^k} P^{k,i}_{t+s} Q^{k,i}_{t+s} | \Omega^k_{t-1}\right]$ for each $i = 1, 2, ..., I^K$ and $s = 0, 1, 2, 3, ...$. Let $a^k_{t+s}(x^{k,j}_{t+r}, y^{k,l}_{t+m})$ denote the cross derivative of the function TR^k_{t+s} w.r.t. any two choice variables $x^{k,j}_{t+q}, y^{k,l}_{t+r}$. Given the inverse demand function $\bar{P}_{t+s} = \left(q^W_{t+s}\right)^{1/\eta} h_{t+s}(Y_{t+s})^{-1}$, the field-level price formula $P^{k,i}_{t+s} = \alpha^{k,i} + \bar{P}_{t+s} + err^{k,i}_{t+s}$ (details in section 2.1 of this appendix), and the market equilibrium condition $q^W_{t+s} = Q^W_{t+s}$, it is easy to show that $a^k_{t+s}(Q^{k,j}_{t+q}, Q^{k,l}_{t+r}) = 0$ if either $q \neq s$ or $r \neq s$ (or both). Moreover, we can show that for all $i = 1, 2, ..., I^k$ and $l = 1, 2, ..., I^k$ $a^k_{t+s}(Q^{k,j}_{t+q}, Q^{k,l}_{t+r})$ has value:

$$a^k_{t+s}(Q^{k,i}_{t+s}, Q^{k,l}_{t+s}) = \frac{1}{\eta}\mathbb{E}_{t-1}\left\{ \frac{\overline{P}_{t+s}}{Q^W_{t+s}} \left[2 + \frac{1-\eta}{\eta} \frac{Q^k_{t+s}}{Q^W_{t+s}} \right] | \Omega^k_{t-1} \right\} = \bar{a}^k_{t+s}$$

where $Q^k_{t+s} \equiv \sum_{i=1}^{I^k} Q^{k,i}_{t+s}$ is the aggregate output of firm k and Q^W_{t+s} is the global oil supply. Note that

$a^k_{t+s}(Q^{k,i}_{t+s}, Q^{k,l}_{t+s})$ has the same value $\bar{a}^k_{t+s}$ for all fields controlled by firm k and is negative if $\frac{Q^k_{t+s}}{Q^W_{t+s}} < 2\left|\frac{\eta}{1-\eta}\right|$. This condition is satisfied for all empirically relevant values of η — i.e., $|\eta| \geq 0.2$ — as long as no firm possesses sufficient aggregate productive capacity to cover more than 50% of the global oil supply. Because no firm possesses such large productive capacity (even if one considers the entire OPEC as a perfect cartel acting as a single firm), we use the result $\bar{a}^k_{t+s} < 0$ throughout the paper and we omit the implicit constraints $Q^k_{t+s}/Q^W_{t+s} \leq 0.5$ for $s = 0, 1, 2, ...$ because they are never binding.

In order to prove that the objective function is concave we must show that the matrix

$$H1^k_{t+s} = \begin{bmatrix} a^k_{t+s}(Q^{k,1}_{t+s}, Q^{k,1}_{t+s}) & a^k_{t+s}(Q^{k,1}_{t+s}, Q^{k,2}_{t+s}) & \cdots & a^k_{t+s}(Q^{k,1}_{t+s}, Q^{k,I^k}_{t+s}) & 0 & 0 & \cdots \\ a^k_{t+s}(Q^{k,2}_{t+s}, Q^{k,1}_{t+s}) & a^k_{t+s}(Q^{k,2}_{t+s}, Q^{k,2}_{t+s}) & \cdots & a^k_{t+s}(Q^{k,2}_{t+s}, Q^{k,I^k}_{t+s}) & 0 & 0 & \cdots \\ \vdots & \vdots & \ddots & \vdots & 0 & 0 & \cdots \\ a^k_{t+s}(Q^{k,I^k}_{t+s}, Q^{k,1}_{t+s}) & a^k_{t+s}(Q^{k,I^k}_{t+s}, Q^{k,2}_{t+s}) & \cdots & a^k_{t+s}(Q^{k,I^k}_{t+s}, Q^{k,2}_{t+s}) & 0 & 0 & \cdots \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix},$$

where the zeros represents the cross derivatives w.r.t. all choice variables

$$\left\{\left\{Q^{k,i}_{t+s}, W^{k,i}_{t+s}, L^{k,i}_{t+s-1}, B^{k,i}_{t+s-1}\right\}_{i=1}^{I^k}\right\}_{s=0}^{\infty}$$

other than those in $\left\{Q^{k,i}_{t+s}\right\}_{i=1}^{I^k}$, is negative semidefinite. Because $a^k_{t+s}(Q^{k,i}_{t+s}, Q^{k,l}_{t+s}) = \bar{a}^k_{t+s}$ for all $i, j = 1, 2, ...I^k$, for any non-zero vector $\mathbf{v}$ we get $\mathbf{v}^T H^k_{t+s}\mathbf{v} = \left(\sum_{j=1}^{I^k} v^j\right)^2 \times \bar{a}^k_{t+s}$. Thus, $\mathbf{v}^T H^k_{t+s}\mathbf{v}$ is weakly negative for any non-zero vector $\mathbf{v}$ because $\bar{a}^k_{t+s} < 0$ given the assumption $|\eta| \geq 0.2$, as shown in the previous paragraph. In turn, each H^k_{t+s} is negative semidefinite, implying that TR^k_{t+s} is a concave function for any value of η that satisfies $|\eta| \geq 0.2$.

Step 1b. Let $d^{k,i}(x^{k,j}_{t+s}, y^{k,l}_{t+r})$ denote the cross derivative of the function $CF^{k,i}_t = -\mathbb{E}_{t-1}\left[\sum_{s=0}^{\infty} \widetilde{C}^{k,i}_{t+s}(Q^{k,i}_{t+s}, B^{k,i}_{t+s-1}, \epsilon^{k,i}_{t+s}) + W\right.$ w.r.t. any two choice variables $x^{k,j}_{t+s}, y^{k,l}_{t+r}$. In order to prove that the objective function is concave we must show that the matrix

$$H2^{k,i}_t = \begin{bmatrix} d^{k,i}(Q^{k,i}_t, Q^{k,i}_t) & d^{k,i}\left(Q^{k,i}_t, L^{k,i}_t\right) & d^{k,i}\left(Q^{k,i}_t, B^{k,i}_t\right) & d^{k,i}\left(Q^{k,i}_t, W^{k,i}_t\right) & ... \\ d^{k,i}(L^{k,i}_t, Q^{k,i}_t) & d^{k,i}\left(L^{k,i}_t, L^{k,i}_t\right) & d^{k,i}\left(L^{k,i}_t, B^{k,i}_t\right) & d^{k,i}\left(L^{k,i}_t, W^{k,i}_t\right) & ... \\ d^{k,i}(B^{k,i}_t, Q^{k,i}_t) & d^{k,i}\left(B^{k,i}_t, L^{k,i}_t\right) & d^{k,i}\left(B^{k,i}_t, B^{k,i}_t\right) & d^{k,i}\left(B^{k,i}_t, W^{k,i}_t\right) & ... \\ d^{k,i}(W^{k,i}_t, Q^{k,i}_t) & d^{k,i}\left(W^{k,i}_t, L^{k,i}_t\right) & d^{k,i}\left(W^{k,i}_t, B^{k,i}_t\right) & d^{k,i}\left(W^{k,i}_t, W^{k,i}_t\right) & ... \\ d^{k,i}(Q^{k,i}_{t+1}, Q^{k,i}_t) & d^{k,i}\left(Q^{k,i}_{t+1}, L^{k,i}_t\right) & d^{k,i}\left(Q^{k,i}_{t+1}, B^{k,i}_t\right) & d^{k,i}\left(Q^{k,i}_{t+1}, W^{k,i}_t\right) & ... \\ d^{k,i}(L^{k,i}_{t+1}, Q^{k,i}_t) & d^{k,i}\left(L^{k,i}_{t+1}, L^{k,i}_t\right) & d^{k,i}\left(L^{k,i}_{t+1}, B^{k,i}_t\right) & d^{k,i}\left(L^{k,i}_{t+1}, W^{k,i}_t\right) & ... \\ d^{k,i}(B^{k,i}_{t+1}, Q^{k,i}_t) & d^{k,i}\left(B^{k,i}_{t+1}, L^{k,i}_t\right) & d^{k,i}\left(B^{k,i}_{t+1}, B^{k,i}_t\right) & d^{k,i}\left(B^{k,i}_{t+1}, W^{k,i}_t\right) & ... \\ d^{k,i}(W^{k,i}_{t+1}, Q^{k,i}_t) & d^{k,i}\left(W^{k,i}_{t+1}, L^{k,i}_t\right) & d^{k,i}\left(W^{k,i}_{t+1}, B^{k,i}_t\right) & d^{k,i}\left(W^{k,i}_{t+1}, W^{k,i}_t\right) & ... \\ ... & ... & ... & ... & ... \end{bmatrix}$$

$$\begin{matrix} ... & d^{k,i}\left(Q_t^{k,i},Q_{t+1}^{k,i}\right) & d^{k,i}\left(Q_t^{k,i},L_{t+1}^{k,i}\right) & d^{k,i}\left(Q_t^{k,i},B_{t+1}^{k,i}\right) & d^{k,i}\left(Q_t^{k,i},W_{t+1}^{k,i}\right) & ... \\ ... & d^{k,i}\left(L_t^{k,i},Q_{t+1}^{k,i}\right) & d^{k,i}(L_t^{k,i},L_{t+1}^{k,i}) & d^{k,i}(L_t^{k,i},B_{t+1}^{k,i}) & d^{k,i}\left(L_t^{k,i},W_{t+1}^{k,i}\right) & ... \\ ... & d^{k,i}\left(B_t^{k,i},Q_{t+1}^{k,i}\right) & d^{k,i}(B_t^{k,i},L_{t+1}^{k,i}) & d^{k,i}(B_t^{k,i},B_{t+1}^{k,i}) & d^{k,i}\left(B_t^{k,i},W_{t+1}^{k,i}\right) & ... \\ ... & d^{k,i}\left(W_t^{k,i},Q_{t+1}^{k,i}\right) & d^{k,i}(W_t^{k,i},L_{t+1}^{k,i}) & d^{k,i}(W_t^{k,i},B_{t+1}^{k,i}) & d^{k,i}\left(W_t^{k,i},W_{t+1}^{k,i}\right) & ... \\ ... & d^{k,i}\left(Q_{t+1}^{k,i},Q_{t+1}^{k,i}\right) & d^{k,i}(Q_{t+1}^{k,i},L_{t+1}^{k,i}) & d^{k,i}(Q_{t+1}^{k,i},B_{t+1}^{k,i}) & d^{k,i}\left(Q_{t+1}^{k,i},W_{t+1}^{k,i}\right) & ... \\ ... & d^{k,i}\left(L_{t+1}^{k,i},Q_{t+1}^{k,i}\right) & d^{k,i}(L_{t+1}^{k,i},L_{t+1}^{k,i}) & d^{k,i}(L_{t+1}^{k,i},B_{t+1}^{k,i}) & d^{k,i}\left(L_{t+1}^{k,i},W_{t+1}^{k,i}\right) & ... \\ ... & d^{k,i}\left(B_{t+1}^{k,i},Q_{t+1}^{k,i}\right) & d^{k,i}(B_{t+1}^{k,i},L_{t+1}^{k,i}) & d^{k,i}(B_{t+1}^{k,i},B_{t+1}^{k,i}) & d^{k,i}\left(B_{t+1}^{k,i},W_{t+1}^{k,i}\right) & ... \\ ... & d^{k,i}\left(W_{t+1}^{k,i},Q_{t+1}^{k,i}\right) & d^{k,i}(W_{t+1}^{k,i},L_{t+1}^{k,i}) & d^{k,i}(W_{t+1}^{k,i},B_{t+1}^{k,i}) & d^{k,i}\left(W_{t+1}^{k,i},W_{t+1}^{k,i}\right) & ... \\ ... & ... & ... & ... & ... & ... \end{matrix}\Bigg]$$

is negative semidefinite for all values of $\left\{\left\{Q_{t+s}^{k,i},W_{t+s}^{k,i},L_{t+s-1}^{k,i},B_{t+s-1}^{k,i}\right\}_{i=1}^{I^k}\right\}_{s=0}^{\infty}$ in X. Since all the cross-derivatives $d^{k,i}(x_{t+s}^{k,j},y_{t+r}^{k,l})$ for $x_{t+s}^{k,j}\neq y_{t+r}^{k,l}$ are all equal to zero, i.e. $\frac{d^2CF_t^{k,i}}{dL_{t+s}^{k,i}dB_{t+r}^{k,i}}=0$; $\frac{d^2CF_t^{k,i}}{dL_{t+s}^{k,i}dQ_{t+r}^{k,i}}=0$; $\frac{d^2CF_t^{k,i}}{dB_{t+s}^{k,i}dW_{t+r}^{k,i}}=0$; $\frac{d^2CF_t^{k,i}}{dB_{t+s}^{k,i}dL_{t+r}^{k,i}}=0$; $\frac{d^2CF_t^{k,i}}{dL_{t+s}^{k,i}dW_{t+r}^{k,i}}=0$; $\frac{d^2CF_t^{k,i}}{dQ_{t+s}^{k,i}dW_{t+r}^{k,i}}=0$ for all $s=1,2,3,...$, $r=1,2,3,....$, and $i=1,2,...,I^k$; except for the following:

$$d^{k,i}\left(Q_{t+s}^{k,i},B_{t+s-1}^{k,i}\right)=\frac{d^2CF_t^{k,i}}{dQ_{t+s}^{k,i}dB_{t+s-1}^{k,i}}=-\kappa^s\frac{\theta_3^{Geo}Q_{t+s}^{k,i}}{R_{t+s-1}^{i}{}^2},$$

$$d^{k,i}\left(Q_{t+s}^{k,i},L_{t+s-1}^{k,i}\right)=\frac{d^2CF_t^{k,i}}{dQ_{t+s}^{i}dL_{t+s-1}^{i}}=\kappa^s\frac{\theta_3^{Geo}Q_{t+s}^{k,i}}{R_{t+s-1}^{i}{}^2},$$

$$d^{k,i}\left(B_{t+s-1}^{k,i},L_{t+s-1}^{k,i}\right)=\frac{d^2CF_t^{k,i}}{dB_{t+s-1}^{k,i}dL_{t+s-1}^{k,i}}=\kappa^s\frac{\theta_3^{Geo}Q_{t+s}^{k,i\,2}}{R_{t+s-1}^{k,i}{}^3},$$

the matrix of cross derivatives can be rearranged in the form of a diagonal matrix:

$$H2_t^{k,i}=\begin{bmatrix} A_1^{k,i} & \mathbf{0}_{12} & \mathbf{0}_{13} & ... \\ \mathbf{0}_{21} & A_2^{k,i} & \mathbf{0}_{23} & ... \\ \mathbf{0}_{31} & \mathbf{0}_{32} & A_3^{k,i} & ... \\ ... & ... & ... & ... \end{bmatrix}$$

where each $(n_j\times n_j)$ submatrix $A_j^{k,i}$ is either diagonal, or it contains all the cross derivatives that are non-zero with respect to a specific choice variable, and where $\mathbf{0}_{jm}$ is a $(n_m\times n_j)$ null matrix. Then, the matrix $H2_t^{k,i}$ is negative semidefinite if all the non-zero submatrices $A_1^{k,i}$, $A_2^{k,i}$, $A_3^{k,i}$, ... are negative semidefinite. First, notice that the own second derivatives are

$$d^{k,i}(Q_{t+s}^{k,i},Q_{t+s}^{k,i})=-\kappa^s(\theta_2^{Geo}+\frac{\theta_3^{Geo}}{R_{t+s-1}^{k,i}})\qquad d^{k,i}(L_{t+s}^{k,i},L_{t+s}^{k,i})=-\kappa^{s+1}\theta_3^{Geo}Q_{t+s+1}^{k,i}{}^2R_{t+s}^{k,i\;-3}$$

$$d^{k,i}(B_{t+s}^{k,i},B_{t+s}^{k,i})=-\kappa^{s+1}\theta_3^{Geo}Q_{t+s+1}^{k,i\;2}R_{t+s}^{k,i\;-3}\qquad d^{k,i}(W_{t+s}^{k,i},W_{t+s}^{k,i})=0\,,$$

for all $s=0,1,2,....$ where $R_{t+s-1}^{k,i}=R^{k,i}+B_{t+s-1}^{k,i}+L_{t+s-1}^{k,i}$. Since all the own second derivatives satisfy $d^{k,i}(x_{t+s}^{k,i},x_{t+s}^{k,i})\leq 0$, then any diagonal submatrix $A_j^{k,i}$ is negative semidefinite. Thus, for the purposes of this prove, it is sufficient to show that each non-diagonal submatrix $\tilde{A}_j^{k,i}$ is also negative semidefinite. Each such non-diagonal submatrix includes only elements $d^{k,i}(x^{k,j}_{t+s}, y^{k,l}_{t+r})$ such that $s = r$ and $j = l = i$ and has form:

$$\tilde{A}^{k,i}_j = \begin{bmatrix} d^{k,i}\left(Q^{k,i}_{t+s}, Q^{k,i}_{t+s}\right) & d^{k,i}\left(Q^{k,i}_{t+s}, B^{k,i}_{t+s-1}\right) & d^{k,i}\left(Q^{k,i}_{t+s}, L^{k,i}_{t+s-1}\right) & d^{k,i}\left(Q^{k,i}_{t+s}, W^{k,i}_{t+s}\right) \\ d^{k,i}\left(B^{k,i}_{t+s-1}, Q^{k,i}_{t+s}\right) & d^{k,i}\left(B^{k,i}_{t+s-1}, B^{k,i}_{t+s-1}\right) & d^{k,i}\left(B^{k,i}_{t+s-1}, L^{k,i}_{t+s-1}\right) & d^{k,i}\left(B^{k,i}_{t+s-1}, W^{k,i}_{t+s}\right) \\ d^{k,i}\left(L^{k,i}_{t+s-1}, Q^{k,i}_{t+s}\right) & d^{k,i}\left(L^{k,i}_{t+s-1}, B^{k,i}_{t+s-1}\right) & d^{k,i}\left(L^{k,i}_{t+s-1}, L^{k,i}_{t+s-1}\right) & d^{k,i}\left(L^{k,i}_{t+s-1}, W^{k,i}_{t+s}\right) \\ d^{k,i}\left(W^{k,i}_{t+s}, Q^{k,i}_{t+s}\right) & d^{k,i}\left(W^{k,i}_{t+s}, B^{k,i}_{t+s-1}\right) & d^{k,i}\left(W^{k,i}_{t+s}, L^{k,i}_{t+s-1}\right) & d^{k,i}\left(W^{k,i}_{t+s}, W^{k,i}_{t+s}\right) \end{bmatrix} =$$

$$= \kappa^s \begin{bmatrix} -\theta^{Geo}_2 - \frac{\theta^{Geo}_3}{R^i_{t+s-1}} & -\frac{\theta^{Geo}_3 Q^i_{t+s}}{{R^i_{t+s-1}}^2} & \frac{\theta^{Geo}_3 Q^i_{t+s}}{{R^i_{t+s-1}}^2} & 0 \\ -\frac{\theta^{Geo}_3 Q^i_{t+s}}{{R^i_{t+s-1}}^2} & -\frac{\theta^{Geo}_3 {Q^i_{t+s}}^2}{{R^i_{t+s-1}}^3} & \frac{\theta^{Geo}_3 {Q^i_{t+s}}^2}{{R^i_{t+s-1}}^3} & 0 \\ \frac{\theta^{Geo}_3 Q^i_{t+s}}{{R^i_{t+s-1}}^2} & \frac{\theta^{Geo}_3 {Q^i_{t+s}}^2}{{R^i_{t+s-1}}^3} & -\frac{\theta^{Geo}_3 {Q^i_{t+s}}^2}{{R^i_{t+s-1}}^3} & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

for $s = 1, 2, ...$ and is negative semidefinite for $\theta^{Geo}_2 \geq 0$, $\theta^{Geo}_3 \geq 0$. Lastly, notice that if all submatrices $\tilde{A}^{k,i}_j$ are negative semidefinite, then the matrix $H2^{k,i}_t$ is negative semidefinite, implying that the function $CF^{k,i}_t$ is concave. Since $Q^{k,i}_{t+s}$ and $R^{k,i}_{t+s-1}$ are weakly positive scalars, the matrix $\tilde{A}^{k,i}_j$ is negative semidefinite if the following conditions hold true: $\theta^{Geo}_2 \geq 0, \theta^{Geo}_3 \geq 0$. Thus, the function $CF^{k,i}_t = -\mathbb{E}_{t-1}\left[\sum_{s=0}^{\infty} \widetilde{C}^{k,i}_{t+s}(Q^{k,i}_{t+s}, B^{k,i}_{t+s-1}, \epsilon^{k,i}_{t+s}) + W^{k,i}_{t+s}|\Omega^k_{t-1}\right]$ is weakly concave given the assumptions $\theta^{Geo}_2 \geq 0$, $\theta^{Geo}_3 \geq 0$.

*Step 1*c. The objective function $OF^k_t = \mathbb{E}_{t-1}\left\{\sum_{i=1}^{I^k}\sum_{s=0}^{\infty}\kappa^s\widetilde{\Pi}^{k,i}_{t+s}|\Omega^k_{t-1}\right\}$ is the sum of concave functions:

$$OF^k_t = \sum_{s=0}^{\infty} TR^k_{t+s} + \sum_{i=1}^{I^k} CF^{k,i}_t$$

and as such, OF^k_t is concave.

Step 2. We show that the set defined by

$$\left\{x \in X \mid \mathbf{g}^k_t\left(\left\{\left\{Q^{k,i}_{t+s}, W^{k,i}_{t+s}, L^{k,i}_{t+s-1}, B^{k,i}_{t+s-1}\right\}_{i=1}^{I^k}\right\}_{s=0}^{\infty}\right) \leq \mathbf{0}\right\}$$

is a convex set. First, recall that—because X is a convex subset of $\mathbb{R}^{\infty}_{+}$—the set $\left\{x \in X \mid g^{k,i}_{t+s,j}(...) \leq 0\right\}$ is convex if $g^{k,i}_{t+s,j}$ is a convex function of the choice variables. Second, recall that intersections of convex sets are convex sets. Thus, in order to prove the result it is sufficient to show that each set $\left\{x \in X \mid g^{k,i}_{t+s,j}(...) \leq 0\right\}$ for $s = 1, 2, ...,$ $i = 1, 2, ..., I^k$, and $j = 1, 2, 3, 4, 5, 6$ is a convex set. First, (i) each function $g^{k,i}_{t+s,j}$ for $j = 2, 3, 4, 5, 6$ is trivially weakly convex because is either linear or constant in each choice variable. Thus, it is sufficient to show that each function $g^{k,i}_{t+s,1}(W^{k,i}_{t+s}, L^{k,i}_{t+s}, L^{k,i}_{t+s-1}, \xi^{k,i}_{t+s})$ for $s = 0, 1, 2, ...$ and $i = 1, 2, ..., I^k$ is convex. Then we have:

$$g^{k,i}_{t+s,1}(W^{k,i}_{t+s}, L^{k,i}_{t+s}, L^{k,i}_{t+s-1}, \xi^{k,i}_{t+s}) = L^{k,i}_{t+s} - D^{k,i}_{t+s}\left(W^{k,i}_{t+s}, L^{k,i}_{t+s-1}, \xi^{k,i}_{t+s}\right) - L^{k,i}_{t+s-1}$$

Let $c^{k,i}_{t+s}\left(x^{k,j}_{t+q}, y^{k,l}_{t+r}\right) \equiv \frac{\partial^2 g^{k,i}_{t+s,1}}{\partial x^{k,j}_{t+q}\partial y^{k,l}_{t+r}}$. Note that $c^{k,i}_{t+s}\left(x^{k,j}_{t+q}, y^{k,l}_{t+r}\right) = 0$ if either $q \neq s$, or $r \neq s$, or $j \neq i$, or $l \neq i$

(or more than one such condition). Thus, in order to prove that $g^{k,i}_{t+s,1}$ is convex, we need to show that the matrix:

$$\begin{bmatrix} c^{k,i}_{t+s}\left(W^{k,i}_{t+s},W^{k,i}_{t+s}\right) & c^{k,i}_{t+s}\left(W^{k,i}_{t+s},L^{k,i}_{t+s-1}\right) & c^{k,i}_{t+s}\left(W^{k,i}_{t+s},L^{k,i}_{t+s}\right) & 0 & 0 & \cdots \\ c^{k,i}_{t+s}\left(L^{k,i}_{t+s-1},W^{k,i}_{t+s}\right) & c^{k,i}_{t+s}\left(L^{k,i}_{t+s-1},L^{k,i}_{t+s-1}\right) & c^{k,i}_{t+s}\left(L^{k,i}_{t+s-1},L^{k,i}_{t+s}\right) & 0 & 0 & \cdots \\ c^{k,i}_{t+s}\left(L^{k,i}_{t+s},W^{k,i}_{t+s}\right) & c^{k,i}_{t+s}\left(L^{k,i}_{t+s},L^{k,i}_{t+s-1}\right) & c^{k,i}_{t+s}\left(L^{k,i}_{t+s},L^{k,i}_{t+s}\right) & 0 & 0 & \cdots \\ 0 & 0 & 0 & 0 & 0 & \cdots \\ 0 & 0 & 0 & 0 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix} =$$

$$= \begin{bmatrix} -2\gamma_2 & 0 & 0 & 0 & 0 & \cdots \\ 0 & -2\gamma_4 & 0 & 0 & 0 & \cdots \\ 0 & 0 & 0 & 0 & 0 & \cdots \\ 0 & 0 & 0 & 0 & 0 & \cdots \\ 0 & 0 & 0 & 0 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

is positive semidefinite, which holds true for $\gamma_2 \leq 0$ and $\gamma_4 \leq 0$. Thus, (ii) $g^{k,i}_{t+s,1}(W^{k,i}_{t+s}, L^{k,i}_{t+s}, L^{k,i}_{t+s-1}, \xi^{k,i}_{t+s})$ is convex given the assumptions $\gamma_2 \leq 0$ and $\gamma_4 \leq 0$. In turn, the two results (i) and (ii) together imply that the set

$$\left\{ x \in X \mid \mathbf{g}^k_t \left(\left\{ \left\{ Q^{k,i}_{t+s}, W^{k,i}_{t+s}, L^{k,i}_{t+s-1}, B^{k,i}_{t+s-1} \right\}_{i=1}^{I^k} \right\}_{s=0}^{\infty} \right) \leq \mathbf{0} \right\}$$

is a convex set given the restrictions $\gamma_2 \leq 0$, $\gamma_4 \leq 0$.

Lastly, the results of Step 1 and 2 together imply that the firm maximization is a convex optimization problem given the restrictions on the parameters $\gamma_2 \leq 0$, $\gamma_4 \leq 0$, $\theta^{Geo}_2 \geq 0$, $\theta^{Geo}_3 \geq 0$, and $|\eta| \geq 0.2$. This implies that the solution to the FOCs is a global maximum. Q.E.D.

First-Order Necessary Conditions for a Global Maximum. The FOCs for each $s = 0, 1, 2, 3, ..$ write:

$$[Q^{k,i}_{t+s}] : \mathbb{E}_{t-1} \left\{ P^{k,i}_{t+s} + \frac{1}{\eta} \overline{P}_{t+s} \frac{Q^k_{t+s}}{Q^W_{t+s}} - \frac{\partial C^{k,i}_{t+s}(.)}{\partial Q^{k,i}_{t+s}} - \lambda^{k,i}_{t+s} \middle| \Omega^k_{t-1} \right\} = 0$$

$$[W^{k,i}_{t+s}] : \mathbb{E}_{t-1} \left\{ -1 + \mu^{k,i}_{t+s} \frac{\partial D^{k,i}_{t+s}(.)}{\partial W^{k,i}_{t+s}} \middle| \Omega^k_{t-1} \right\} = 0$$

$$[L^{k,i}_{t+s}] : \mathbb{E}_{t-1} \left\{ -\mu^{k,i}_{t+s} - \kappa \frac{\partial C^{k,i}_{t+s+1}(.)}{\partial R^i_{t+s}} + \kappa \mu^{k,i}_{t+s+1} \frac{\partial D^{k,i}_{t+s+1}(.)}{\partial L^{k,i}_{t+s}} + \kappa \mu^{k,i}_{t+s+1} \middle| \Omega^k_{t-1} \right\} = 0$$

$$[B^{k,i}_{t+s}] : \mathbb{E}_{t-1} \left\{ \lambda^{k,i}_{t+s} - \kappa \frac{\partial C^{k,i}_{t+s+1}(.)}{\partial R^{k,i}_{t+s}} - \kappa \lambda^{k,i}_{t+s+1} \middle| \Omega^k_{t-1} \right\} = 0$$

$$[\mu^{k,i}_{t+s}] : \mathbb{E}_{t-1} \left\{ -D^{k,i}_{t+s}(W^{k,i}_{t+s}, L^{k,i}_{t+s-1}, \xi^{k,i}_{t+s}) + L^{k,i}_{t+s} - L^{k,i}_{t+s-1} \middle| \Omega^k_{t-1} \right\} \leq 0$$

plus the standard primal feasibility, dual feasibility and complementary slackness conditions.

Interior Solution. From the first part we know that the solution is a global maximum. The solution must be interior for $(Q^{k,i}_{t+s}, W^{k,i}_{t+s}, L^{k,i}_{t+s}, B^{k,i}_{t+s})$—i.e., in field i and in period $t+s$—if $(Q^{k,i}_{t+s}, W^{k,i}_{t+s}, L^{k,i}_{t+s}, R^{i,k} - B^{k,i}_{t+s} + L^{k,i}_{t+s})^* \gg 0$. Thus, whenever the observed values of production, investment, discoveries and reserves satisfy $(Q^{k,i}_{t+s}, W^{k,i}_{t+s}, L^{k,i}_{t+s}, R^{k,i}_{t+s})^* \gg 0$, the FOCs w.r.t. $Q^{k,i}_{t+s}, W^{k,i}_{t+s}, L^{k,i}_{t+s}, B^{k,i}_{t+s}, \lambda^{k,i}_{t+s}, \mu^{k,i}_{t+s}$ must be binding, and therefore they can be used to derive the following equilibrium conditions. Q.E.D.

Equilibrium Conditions (Shadow-Prices). The results in the previous sections lead to the following equilibrium conditions for all fields that satisfy $(Q_t^{k,i}, W_t^{k,i}, L_t^{k,i}, R_t^{k,i})^* \gg 0$:

1. Shadow-price of discovered oil:

$$\mathbb{E}_{t-1}\left[\lambda_t^{k,i}\left|\Omega_{t-1}^k\right.\right] = \mathbb{E}_{t-1}\left[P_t^{k,i}\middle|\,\Omega_{t-1}^k\right] + \mathbb{E}_{t-1}\left[\frac{1}{\eta}\overline{P}_t\frac{Q_t^k}{Q_t^W}\middle|\,\Omega_{t-1}^k\right] - \mathbb{E}_{t-1}\left[\frac{\partial C_t^i(.)}{\partial Q_t^i}\middle|\,\Omega_{t-1}^k\right]$$

2. Expected shadow-price of undiscovered oil:

$$\mathbb{E}_{t-1}\left[\mu_t^{k,i}\left|\Omega_{t-1}^k\right.\right] = \mathbb{E}_{t-1}\left[\left(\frac{\partial D_t^{k,i}(.)}{\partial W_t^{k,i}}\right)^{-1}\middle|\,\Omega_{t-1}^k\right]$$

3. Law of motion of the shadow-price of discovered oil:

$$\mathbb{E}_{t-1}\left[\lambda_{t+1}^{k,i}\left|\Omega_{t-1}^k\right.\right] = \mathbb{E}_{t-1}\left[\frac{\lambda_t^{k,i}}{\kappa} - \frac{\partial C_{t+1}^{k,i}(.)}{\partial R_t^{k,i}}\middle|\,\Omega_{t-1}^k\right]$$

4. Law of motion of the expected shadow-price of undiscovered oil:

$$\mathbb{E}_{t-1}\left[\mu_{t+1}^{k,i}\left|\Omega_{t-1}^k\right.\right] = \mathbb{E}_{t-1}\left[\left(\frac{\mu_t^{k,i}}{\kappa} + \frac{\partial C_{t+1}^{k,i}(.)}{\partial L_t^{k,i}}\right)\left(\frac{\partial D_{t+1}^{k,i}(.)}{\partial L_t^{k,i}} + 1\right)^{-1}\middle|\,\Omega_{t-1}^k\right],$$

Note that $\lambda_t^{k,i}$ is fully known at time t by the firm – i.e., $\mathbb{E}_{t-1}\left[\lambda_t^{k,i}\left|\Omega_{t-1}^i\right.\right] = \hat{\lambda}_t^{k,i}$ – but it is a random variable for the econometrician, because the exact realizations of the random coefficients are unknown. All these results are in line with standard natural resource economics. Lastly, the component $\mathbb{E}_{t-1}\left[\frac{1}{\eta}\overline{P}_t\frac{Q_t^k}{Q_t^W}\middle|\,\Omega_{t-1}^k\right] \leq 0$ in the formula for the shadow-price $\mathbb{E}_{t-1}\left[\lambda_t^{k,i}\left|\Omega_{t-1}^k\right.\right]$ captures the effect of firm k's market power due to imperfect competition. If $\eta \to -\infty$ this component vanishes and all the firm behaves as price-takers. Conversely, for finite values of η the shadow-price is, ceteris paribus, decreasing in the firm's size Q_t^k/Q_t^W because a marginal increase in the output of field i causes a fall in the equilibrium price, which negatively affects the firm's revenues. This effect is proportional to the total oil production of the firm. To see why, note that a fall in oil price of $\$x$ due to a marginal increase in production generates a revenue loss for the firm equal to $\$x \times Q_{t+s}^k$, which is proportional to the firm's aggregate production. As a consequence, the value of the marginal barrel for the firm - and therefore the shadow-price of oil - is decreasing in the firm's size.

Empirical economic model

The empirical analysis associates the shadow price of a field,

$$\mathbb{E}_{t-1}[\lambda_t^{k,i}|\Omega_{t-1}^k] = \mathbb{E}_{t-1}[P_t^{k,i}|\Omega_{t-1}^k] - \mathbb{E}_{t-1}\left[\frac{\partial C_t^i(.)}{\partial Q_t^i}\middle|\Omega_{t-1}^k\right] + \frac{1}{\eta}\mathbb{E}_{t-1}\left[\bar{P}_t\,\frac{Q_t^k}{Q_t^w}\middle|\Omega_{t-1}^k\right], \tag{4.6}$$

to its Carbon Intensity (CI). Its realization relies on four distinct datasets.

Environmental The OPGEE global CI dataset contains information about the emissions of 8,966 children commercial fields in 2015.

Price The Energy Information Administration dataset on Landed Costs of Imported Crude for Selected Crude Streams contains the future prices for twenty-three oil classes over the time interval 1979-2018 [EIA, 2020a]. Complementary information about the relative chemical composition of the Selected Crude Streams is imported from the PSA Management and Services BV database.

Costs The WoodMac Upstream Data Tool contains information about costs, production and reserves across two levels: 30,235 children and standalone commercial oil fields and 1,916 parent fields over the time interval 1965-2018.

The granularity of the data sources is heterogeneous. The environmental information is available at children level, the price at oil class level, while the costs both at the children or standalone and at the parent level. Nevertheless, since the children's level costs presents a considerable number of missing values, we decided to preserve the largest amount of information by focusing on the parent and standalone level. Therefore, we harmonize the granularity of all the other data sources to the standalone and parent level. First, we refer to the oil category prices, enhanced by the information of the chemical composition of the category, to approximate the price at which parent and standalone producers sell their oil (see Section Firm Expected Prices). Then, we aggregate the CI at the parent level by computing a weighted average of the 8,817 fields resulting from the harmonizing the OPGEE and the WoodMac IDs (see Section Data Harmonization and IDs Match).

Firm expected prices

The prices at which fields sell their output respond to the timely interaction between demand and supply. In addition, they depend upon some of the crude chemical characteristics, such as the gravity and the sulphur content [Lanza et al., 2005, Fattouh, 2010][5]. Consequently, the field-level price,

$$\mathbb{E}_{t-1}[P_t^{k,i}|\Omega_{t-1}^i] = \bar{P}_t + \beta_{1t}(API^i - \overline{API}_t) + \beta_{2t}(S^i - \bar{S}_t)\ , \tag{4.7}$$

can be thought as a time moving value around the average oil price $(\bar{P}_t)$. The latter is the average real price at which the United States refineries import crude streams[6]. The unit of account of $\bar{P}_t$ is \$ /BOE. The same applies to β_{1t} and β_{2t}, since API and sulphur are dimensionless quantities. The data on $(\bar{P}_t, \overline{API}_t, \bar{S}_t)$ are obtained using the Energy Information Administration dataset [EIA, 2020a,c].

The variations around the average price depend upon the delta between the gravity of field i (API^i) and the average gravity of imported crudes $(\overline{API}_t)$ at time t as well as the delta between the sulphur content of field i (S^i) and the average sulphur content of imported crudes $(\bar{S}_t)$ at time t. We allow the magnitude of these deviations to be time-dependent. The time change could be due to: 1) the composition of the demand for oil derived products, 2) the technology employed by the refineries, 3) a combination of 1) and 2). For example, if the demand for light products, like gasoline and jet fuel increases, while the demand for heavy products, like liquid heavy ends and naphtha, declines, the impact of $(API^i - \overline{API}_t)$ on $\mathbb{E}_{t-1}[P_t^i|\Omega_{t-1}^i]$ might increase. In the same way, if an improvement in technologies allows refineries to produce lighter products using heavier oils without increasing their operational costs, the impact of $(API^i - \overline{API}_t)$ on $\mathbb{E}_{t-1}[P_t^i|\Omega_{t-1}^i]$ might decrease. In the same way, it is possible to imagine an interplay between these two effects.

$\mathbb{E}_{t-1}[P_t^i|\Omega_{t-1}^i]$ is unobservable to the researcher, but it can be proxied by the prices of several mixtures of crudes, which group oils coming from different fields into a tradable oil class. Using the Energy Information Administration dataset and the PSA Management and Services BV database, we collect, respectively, the yearly future prices and the chemical characteristics of twenty-three oil classes over the time interval 1978-2018[7] [EIA, 2021b, of Canada , PSA]. Table 5.1 provides the summary statistics of the future prices, together with the gravity

[5] A variation of oil prices due to their quality can be thought as the outcome of a Bertrand competition between two or more buyers (typically refineries or intermediaries), which evaluate the crude depending upon its qualities.

[6] The average prices are available only in nominal terms. To obtain the real ones, we use the WoodMac Consumer Price Index (CPI). This adjustment guarantees the harmonization of all the variables of equation (4.6).

[7] We take advantage of the full time span of the series to improve the model's fitting.

(API) and sulphur content (S). Figure (4.3) shows how the future prices (FP) of twenty-three oil classes z homogeneously respond to a general trend, whereas Figure (4.4) portrays the positive relation between the future price of a category of crude ($FP(z)$) with its API and the negative relation between $FP(z)$ with its sulphur content in a single year (2015).

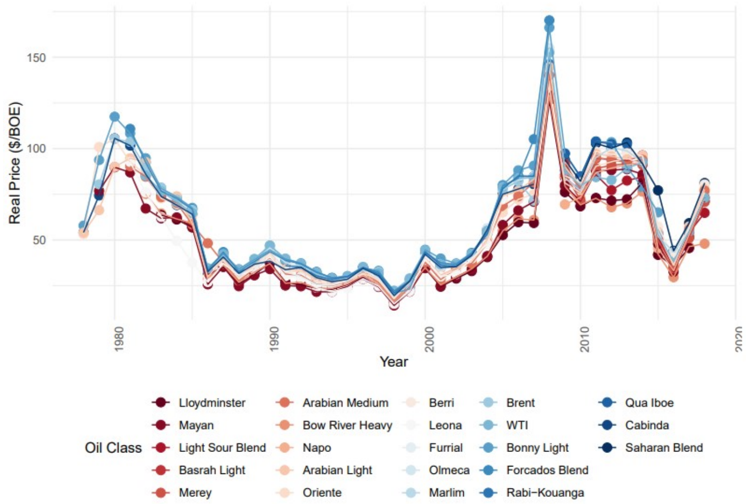

Figure 4.3: Future prices of twenty-three oil class over the time interval 1978-2018. Colors reflect the sulphur content, where dark red represents high content whereas dark blue low percentages.

Table 4.1: Summary statistics of $FP_t(z)$ for twenty-three oil classes.

Oil Class (z)	Country of Origin	Mean	SD	Min	Max	API	S
Arabian Light	Saudi Arabia	40.39	29.36	12.36	109.43	32.8	1.97
Arabian Medium	Saudi Arabia	40.70	29.34	10.86	107.12	30.2	2.59
Basrah Light	Iraq	76.11	25.70	39.90	106.93	30.5	2.90
Berri	Saudi Arabia	78.82	25.79	45.62	110.77	38.5	1.50
Bonny Light	Nigeria	42.21	30.84	13.62	117.70	33.4	0.16
Bow River Heavy	Canada	33.96	22.82	10.41	84.29	24.7	2.10
Brent Crude	United Kingdom	28.10	13.30	13.94	64.60	38.3	0.37
Cabinda	Angola	26.90	13.92	12.69	69.17	32.4	0.13
Forcados Blend	Nigeria	32.34	22.95	14.35	111.07	30.8	0.16
Furrial	Venezuela	18.27	4.26	12.24	28.23	30.0	1.06
Leona	Venezuela	20.98	9.36	9.79	51.55	24.0	1.50
Light Sour Blend	Canada	69.09	20.51	40.04	96.52	64.0	3.00
Lloydminster	Canada	33.88	23.95	10.15	82.50	20.9	3.50
Marlim	Brazil	78.42	27.83	47.77	114.32	19.6	0.67
Mayan	Mexico	36.01	27.09	9.21	100.29	21.8	3.33
Merey	Venezuela	72.31	24.94	38.97	103.28	15.0	2.70
Napo	Ecuador	70.78	25.76	37.46	101.53	19.0	2.00
Olmeca	Mexico	31.82	22.98	13.58	101.14	37.3	0.84
Oriente	Ecuador	39.10	27.57	11.55	105.50	24.1	1.51
Qua Iboe	Nigeria	99.73	22.16	68.26	117.02	36.3	0.14
Rabi-Kouanga	Gabon	33.79	23.38	13.65	95.46	37.7	0.15
Saharan Blend	Algeria	83.16	24.68	49.82	115.82	45.0	0.09
WTI	United States	42.30	27.68	14.34	99.56	39.6	0.24

Sources:[EIA, 2021b, of Canada , PSA].

In order to justify the use of such proxy, we assume that the oil class prices equals the weighted average of the field prices belonging to the respective class. In this context, if all the information is publicly available, $\Omega^k_{t-1} = \Omega^{pub}_{t-1}$, the time-varying weights $\{w^{k,i}_t\}^{N(z)}_{i=1}$ identify the relative importance of a field belonging to class

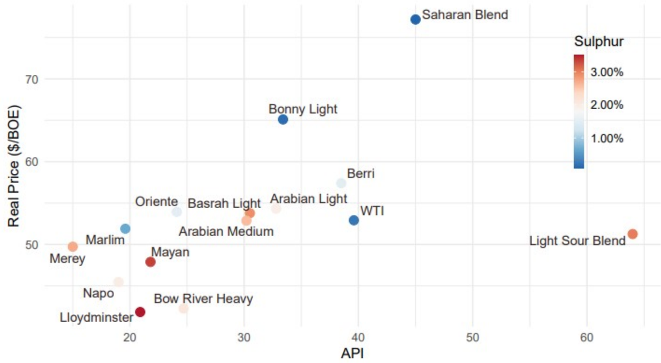

Figure 4.4: 2015 prices of fourteen oil classes based on their API and sulphur content.

z in period t, such that

$$FP_t(z) = \mathbb{E}_{t-1}\left[\sum_{i\in z} w_t^{k,i} P_t^{k,i} | \Omega_{t-1}^{pub}\right] = \mathbb{E}_{t-1}\left[\sum_{i\in z} w_t^{k,i} P_t^{k,i} | \Omega_{t-1}^{k}\right] . \quad (4.8)$$

Thanks to the assumption entailed in equation 4.8[8], we can stack the different time series of the oil classes future price into a longitudinal dataset and we can estimate regression

$$FP_t(z) = \bar{P}_t + \beta_{1t}(API(z) - \overline{API}_t) + \beta_{2t}(S(z) - \bar{S}_t) + \varsigma_t , \quad (4.9)$$

where

$$\mathbb{E}_t[\varsigma_t | API(z), S(z)] = \mathbb{E}_t\left[\sum_{i\in z}(w_t^{k,i} - w^{k,i})(\beta_{1t} API^{k,i} + \beta_{2t} S^{k,i}) | API(z), S(z)\right] = 0$$

is the random error component.

Among the twenty-three oil classes analyzed, ten (Cabinda, Bow River Heavy, Lloydminster, Oriente, Mayan, Bonny Light, Forcados Blend, Arabian Light, Arabian Medium and WTI) have time series long enough to allow us to perform unit roots tests. We run a Panel Covariate-Augmented Dickey-Fuller Test for unit roots, which also controls for unobserved cross-sectional correlation [Pesaran, 2007, Costantini and Lupi, 2013]. We performed different variations of the test: 1) with linear trend without Pesaran cross-sectional correlation, 2) with linear trend with Pesaran cross-sectional correlation, 3) with drift without Pesaran cross-sectional correlation, 4) with drift with Pesaran cross-sectional correlation. All the results fail to reject the null hypothesis of presence of a unit root with the p-values ranging from ∼0.88 to ∼0.41. In the same way, an Augmented Dickey-Fuller test for unit root of the time series $\bar{P}_t$ fails to reject the null hypothesis of presence of unit root (p-value = ∼0.30) [Said and Dickey, 1984, Banerjee et al., 1993]. The lack of stationarity of both the dependent and the explanatory variable requires the use of an unit root compatible strategy to estimate equation (5.8). Following Bacon and Tordo [2005], we regress the differential between the class prices and the average one on the differences in API gravity and sulphur content,

$$FP_t(z) - \bar{P}_t = \beta_{1t}(API(z) - \overline{API}) + \beta_{2t}(S(z) - \bar{S}) + \varsigma_t . \quad (4.10)$$

Being the difference between two variables cointegrated of order one, the new dependent is stationary. More precisely, we perform the four previously mentioned tests and obtain p-values ranging from 4.8e-06 to 1.4e-08.

Table 4.2 reports the results using a Pooled Ordinary Least Square (POLS) ($\beta_{1t} = \beta_1$ and $\beta_{2t} = \beta_2$) and a Random Coefficient Model (RCM) where β_{1t} and β_{2t} are normally distributed [Swamy, 1970, Bates, 2005, De Boeck et al., 2011].

[8] Please notice that equation 4.8 only serves as a theoretical justification of our next steps and it is not estimated.

According to the POLS estimates a unit increase in the ΔAPI augments the value of the crude by 0.13 \$/BOE, while a 1% increase in ΔS content decreases the value of the oil by 2.52 \$/BOE. The adjusted R^2 is 0.32. The conditional modes of the RCM are presented in Table 5.3[9]. The delta in API ranges from -0.04 \$/BOE in 2012 to a maximum of 0.13 \$/BOE in 2015, with an average value of 0.04 \$/BOE, and a median one 0.03 \$/BOE. Similarly, the delta in sulphur ranges from a minimum of -8.25 \$/BOE in 2008 to a maximum of -0.03 in 1986, with a average value of -2.55 \$/BOE, and a median one of -2.20 \$/BOE. This second model increases the adjusted R^2 computed according to Nakagawa and Schielzeth [2013] to 0.40. In other words, a model which allows for time variations in the returns on deltas explains almost half of the variance of the delta between the future price of a particular oil class and the average oil price suggesting that 7% of the variance of $FP_t(z) - \bar{P}_t$ is due to a combination of changes in the composition of the demand for oil derived products and of technological changes in the refinery sector.

Table 4.2: Pooled ordinary least square (POLS) results.

	Dependent variable: ΔPrice			
	Value	Std. Err.	C.I.2.5%	C.I.97.5%
ΔAPI	0.13^{***}	0.03	0.07	0.20
ΔS	-2.52^{***}	0.23	-2.97	-2.07
Observations	484			
Adjusted R^2	0.33			
Residual Std. Error	5.26 (df = 482)			
F Statistic	119.05^{***} (df = 2; 482)			
Note:		*p<0.1; **p<0.05; ***p<0.01		

We use the estimates in Table 4.3 to obtain the field-level prices. For example, in 2015, when $\bar{P}_t = 50.39$ \$/BOE, $\overline{API}_t = 31.46$, $\bar{S}_t = 1.4\%$, $\hat{\beta}_{1,2015} = 0.13$ \$/BOE and $\hat{\beta}_{2,2015} = -2.87$ \$ / BOE, the estimated price at which field i containing oil with $API^i = 55$ and $S^i = 3\%$ would be,

$$\mathbb{E}_{t-1}[\hat{P}_t^i|\Omega_{t-1}^i] = 50.39 + \underset{(0.08)}{0.13} * (55 - 31.46) - \underset{(1.01)}{2.87} * (0.03 - 0.014) = \frac{53.40\$}{\text{BOE}} .$$

Notice that under the assumption stated in this section $\hat{P}_t^i$ is an unbiased estimator of $\mathbb{E}_{t-1}[P_t^i|\Omega_{t-1}^i]$ as long as $\hat{\beta}_{1t}, \hat{\beta}_{2t}$ are unbiased estimators of β_{1t}, β_{2t}. Figure (4.5) shows the resulting field-level mapping for the year 2015.

Data harmonization and IDs match

This subsection describes the harmonization of the OPGEE and the WoodMac IDs necessary to link the shadow prices to their emissions. The WoodMac Upstream Data Tool contains historical data for 20,522 children and standalone commercial oil & gas fields [Mackenzie, 2018]. The OPGEE global carbon intensity models the emissions of 8,966 commercial children fields. The OPGEE field IDs were harmonized to the Wood Mac ones with the following steps:

- Harmonization of the apostrophes and of the countries' names (`Sudan` vs `South Sudan`, `Republic of Congo` vs `Congo`, `Brunei` vs `Brunei Darussalam`)
- Partition of the `Sacha` field into `Sacha (pre-2016)` and `Sacha (post-2016)`
- Aggregation of fields `Garzad A`, `Garzan B` and `Garzan C` into a unique field labeled `Garzan`
- Removal of the field `Columba BD` since it was overlapping with two separate fields `Columba B` and `Columba D`

[9] The RCM is estimated using the R-package `lme4` [Bates et al., 2015] and the conditional modes are extracted with the command `ranef`.

Table 4.3: Conditional modes of the random coefficient model.

	ΔAPI				ΔS			
Year	Value	Std. Dev.	C.I.2.5%	C.I.97.5%	Value	Std. Dev.	C.I.2.5%	C.I.97.5%
1985	0.07	0.11	-0.14	0.28	-2.75	1.10	-4.90	-0.59
1986	0.05	0.11	-0.16	0.26	-0.03	1.11	-2.21	2.15
1987	0.02	0.11	-0.19	0.23	-1.76	1.12	-3.96	0.43
1988	0.02	0.11	-0.19	0.23	-2.06	1.13	-4.28	0.15
1989	0.03	0.11	-0.18	0.24	-1.94	1.14	-4.17	0.28
1990	0.05	0.11	-0.16	0.26	-1.88	1.14	-4.12	0.36
1991	0.06	0.11	-0.15	0.27	-3.09	1.13	-5.30	-0.88
1992	0.04	0.11	-0.17	0.25	-2.81	1.13	-5.03	-0.60
1993	0.04	0.11	-0.17	0.25	-2.22	1.13	-4.43	-0.01
1994	0.03	0.11	-0.18	0.24	-1.55	1.13	-3.76	0.66
1995	0.03	0.11	-0.18	0.24	-1.18	1.13	-3.39	1.02
1996	0.03	0.11	-0.18	0.24	-1.37	1.13	-3.58	0.84
1997	0.03	0.11	-0.18	0.24	-2.03	1.14	-4.26	0.20
1998	0.03	0.11	-0.18	0.24	-1.67	1.14	-3.90	0.56
1999	0.03	0.11	-0.18	0.24	-0.74	1.14	-2.97	1.48
2000	0.03	0.11	-0.18	0.24	-2.41	1.14	-4.63	-0.18
2001	0.03	0.11	-0.18	0.24	-3.52	1.14	-5.74	-1.29
2002	0.03	0.11	-0.18	0.24	-1.70	1.14	-3.93	0.52
2003	0.02	0.11	-0.20	0.23	-2.47	1.13	-4.69	-0.25
2004	0.06	0.11	-0.15	0.27	-3.65	1.14	-5.87	-1.42
2005	0.09	0.11	-0.12	0.30	-5.88	1.14	-8.11	-3.65
2006	0.09	0.11	-0.13	0.30	-5.45	1.14	-7.68	-3.22
2007	0.03	0.11	-0.18	0.25	-5.58	1.14	-7.80	-3.35
2008	0.06	0.11	-0.15	0.27	-8.26	1.21	-10.62	-5.89
2009	0.12	0.08	-0.03	0.28	-2.35	0.98	-4.27	-0.43
2010	0.10	0.08	-0.05	0.25	-2.17	0.98	-4.09	-0.26
2011	0.07	0.08	-0.08	0.22	-3.66	0.98	-5.57	-1.74
2012	-0.04	0.08	-0.19	0.11	-3.98	0.98	-5.91	-2.05
2013	0.01	0.08	-0.14	0.17	-3.31	0.99	-5.25	-1.38
2014	0.00	0.08	-0.16	0.15	-0.37	1.02	-2.37	1.64
2015	0.13	0.08	-0.02	0.29	-2.87	1.01	-4.85	-0.89
2016	0.03	0.08	-0.12	0.19	-0.85	1.05	-2.91	1.21
2017	0.06	0.08	-0.10	0.21	-1.01	1.05	-3.06	1.05
2018	0.02	0.08	-0.13	0.18	-0.22	1.17	-2.52	2.08

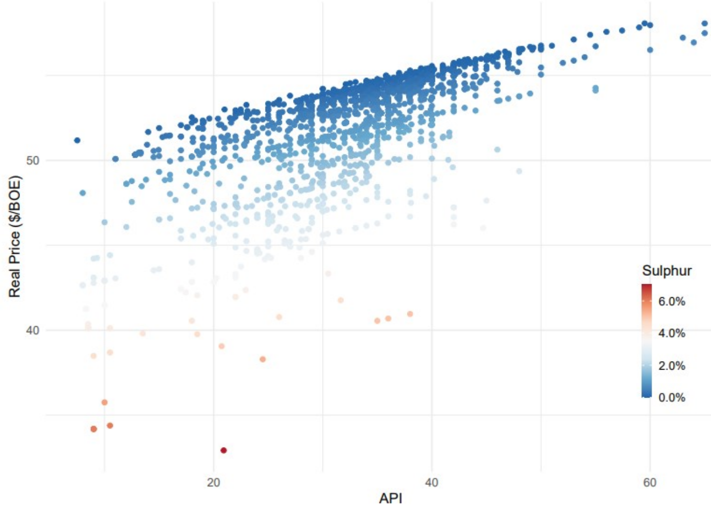

Figure 4.5: 2015 estimated prices of 1,931 oil fields.

- Aggregation of the fields `G-Ain Dar`, `G-Haradh`, `G-Hawiyah`, `G-Shedgum`, `G-Uthmaniyah` into a unique field named `Ghawar`.

The first attempt to join the two datasets, using only the fields ID, matches 8,632 out of 8,966 fields. Next, we perform a smart string search using `R` commands `str_dectect` and `stringdist_left_joint` to recover 24 out of the 325 unmatched fields. Then, we match by-hand 29 fields checking the field names spelling and their geographical coordinates. Finally, we assign to unmatched Canadian fields the CI measure for the representative Canadian field `Other heavy (conv. production)` to all the unmatched Heavy and Extra Heavy Canadian fields. In the same way, we assign the CI measure for the representative Canadian field `Other conventional` to all the unmatched Light & Medium and Other Oil Canadian fields. At the end, we were able to match 8,817 fields out of the initial 8,966.

The WoodMac Upstream Data Tool almost exclusively provides balance sheet data at the parent or standalone level. In order to pair the granularity of the two datasets, we compute the volumes weighted average CI, measured in kg CO_2eq. per BOE,

$$CI^i = \frac{\sum_{j=1}^{J} Q_t^j CI^j}{\sum_{j=1}^{J} Q_t^j} \quad \forall j \in i \, ,$$

of 6,841 children fields[10]. The result is a techno-economic dataset containing the accounting and the environmental characteristics of 995 parent fields and 2,062 standalone fields.

After restricting our sample to fields which have production bigger than zero and provide information about their OPEX and CAPEX expenditures, the final dataset follows 2,062 fields, across 77 countries, over the decade 2009-2018. Our panel contains 17,494 data points[11] and covers an average of 72.57 percent of the global oil demand as estimated by the Organization for Economic Co-operation and Development (OECD). The dataset does not only cover a large fraction of the global oil demand, but it is also representative of the supply in terms of field location (On Shore, Shallow Water, Deepwater, Ultra-Deepwater) and of the chemical and geological peculiarities. Namely, it contains information about low viscosity oil trapped in impermeable rocks (Shale & Tight Oil), low viscosity oil trapped in permeable rocks (Light & Medium Oil), high viscosity oil trapped in permeable rocks (a.k.a. Heavy & Extra Heavy Oil), and oil sands, see Table 4.4.

[10]Please notice that no action as been taken for the remaining standalone fields.

[11]Note that a balanced dataset would have had $N \times T = 2,062 * 10 = 20,620$ data points. The discrepancy originates because the 15.16 percent of the fields are not observed at every point in time because they either stop or start operations during the studied period.

Table 4.4: Absolute frequency of different geological formations.

	On Shore	Shallow Water	Deepwater	Ultra-Deepwater
Light & Medium	682	480	118	31
Heavy	102	47	10	5
Shale & Tight	339	7	0	0
Sands	27	0	0	0
Extra Heavy	9	0	0	0
Other Oil	195	6	4	1

Additional treatments for North America Many U.S. oil fields (shale & tight oils in particular) are not named consistently and systematically. These fields are typically labeled based on their basin, the producer company name, or a combination of both. Therefore, it is very difficult to match these fields between different datasets. In order to improve the U.S. coverage in this work, after a rigorous field-by-field manual matching, two generic oil fields - *US tight oil_generic* and *US heavy_generic* - are created, and volume-weighted-average CI and MC of other U.S. matched tight oil and heavy fields are attributed to these two generic fields. The volumetric production of these two generic fields are estimated based on the missing production volume from the total production of the corresponding crude types using EIA U.S. tight oil production [EIA, 2020a] and crude oil and lease condensate production by API gravity [EIA, 2020b] statistics in 2015, respectively.

Data coverage The production coverage summary of the top 20 largest global crude oil producers in 2015 shown in Table 4.5) confirms that the dataset used in this study covers different types of crudes from different countries and fairly represents the global oil production market. See the supplementary Excel sheet for list of all countries coverage.

Table 4.5: Production coverage summary of top 20 largest global crude oil producers in year 2015.

Country	Count of parent fields	2015 total production*, $mmbbl/d$	Coverage in this work, %
Russian Federation	151	10.3	96
Saudi Arabia	15	10.2	100
United States	490	9.4	80
China	65	4.3	97
Iraq	27	4.0	70
Canada	151	3.7	90
Iran	23	3.3	92
United Arab Emirates	8	3.1	44
Kuwait	4	2.8	100
Venezuela	44	2.5	69
Brazil	34	2.4	94
Mexico	19	2.3	99
Nigeria	73	2.2	91
Angola	32	1.8	95
Kazakhstan	35	1.7	84
Norway	56	1.6	100
Qatar	9	1.5	51
Algeria	20	1.4	83
Colombia	59	1.0	91
Oman	10	1.0	100

**Crude oil including lease condensate. Source:*[EIA, 2017].

The 2015 carbon intensity map of global oil fields is illustrated in Figure (4.6).

Marginal extraction cost

Extraction costs are the sum of all the expenditures faced to get the oil out from the ground. Exploration costs are the sum of all the expenditures faced to find new oil. WoodMac classifies costs into twenty-three categories, see

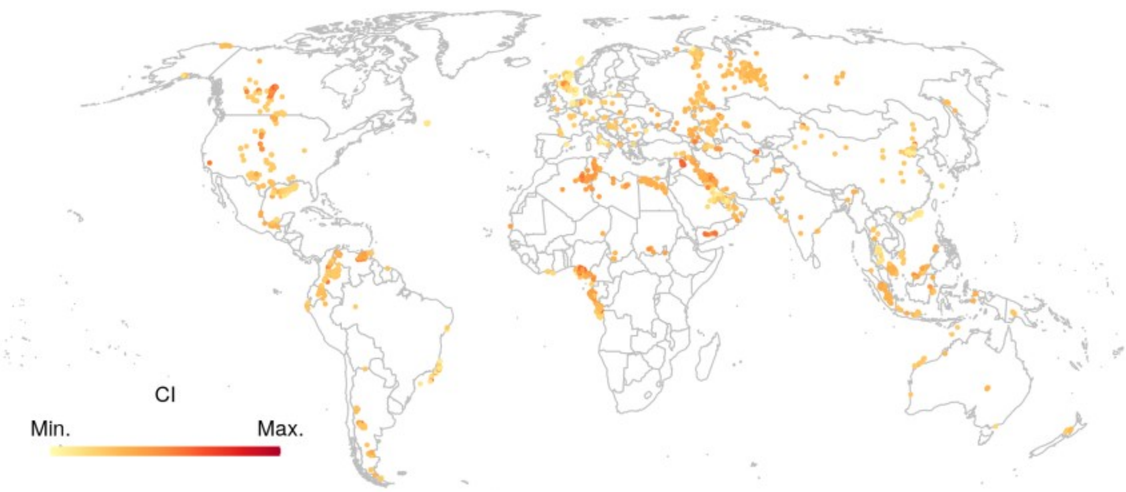

Figure 4.6: Geographical location and the carbon intensity of covered global oil fields in 2015.

Table 4.6 . We sum the first twenty-one of them to obtain the extraction costs,

$$C_t^i = \text{Abandonment Costs}_t^i + \text{Capital Receipts}_t^i + ... + \text{Terminal}_t^i \ ,$$

and the last two to obtain the exploration costs,

$$W_t^i = \text{Development Drilling}_t^i + \text{Exploration and Appraisal}_t^i \ .$$

Table 4.6: Summary statistics of the twenty-three types of cost listed in WoodMac.

Cost Type	Number of Observations	Mean	Std. Dev.	Min	Max
Extraction Costs					
Abandonment Costs	14,196	1.16	9.77	-8.87	378.72
Capital Receipts	401	6.22	59.26	0.00	1,044.79
Country Specific CAPEX	3,859	7.82	60.81	0.00	1,586.93
Country Specific OPEX	1,772	3.96	12.66	0.00	317.39
Field Fixed Costs	18,056	73.58	237.96	0.00	5,175.10
Field Variable Costs	17,701	45.78	117.94	0.00	2,928.38
General and Administrative	2,549	6.39	12.90	0.00	186.45
Insurance	40	0.07	0.41	0.00	2.62
Non Tariff Transport	2,055	20.05	70.27	0.00	931.03
Offshore Loading	854	3.93	20.70	0.00	264.49
Other CAPEX	10,461	19.14	77.25	-235.08	1,805.82
Other Costs	776	52.39	94.01	0.00	1,246.24
Other OPEX	726	9.08	31.83	0.00	212.17
Pipeline	14,491	6.30	32.50	-17.14	1,060.07
Processing Equipment	13,287	30.31	138.37	-32.60	3,191.38
Production Facilities	16,670	40.38	188.19	-40.14	6,260.00
Subsea	3,701	25.61	84.89	0.00	1,378.67
Tariff Gas	6,987	11.53	60.26	0.00	1,678.20
Tariff Oil	11,731	30.49	134.68	0.00	3,378.91
Tariff Receipts	1,604	11.47	25.17	0.00	239.46
Terminal	759	3.40	17.45	0.00	269.04
Exploration Costs					
Development Drilling	17,90	82.59	206.94	0.00	4,495.88
Exploration and Appraisal	123	3.46	12.33	0.00	82.87

Sources:[Mackenzie, 2018].

Figure 4.7 shows the relative importance of the different cost categories, where fixed costs ($\sim$ 30%) and

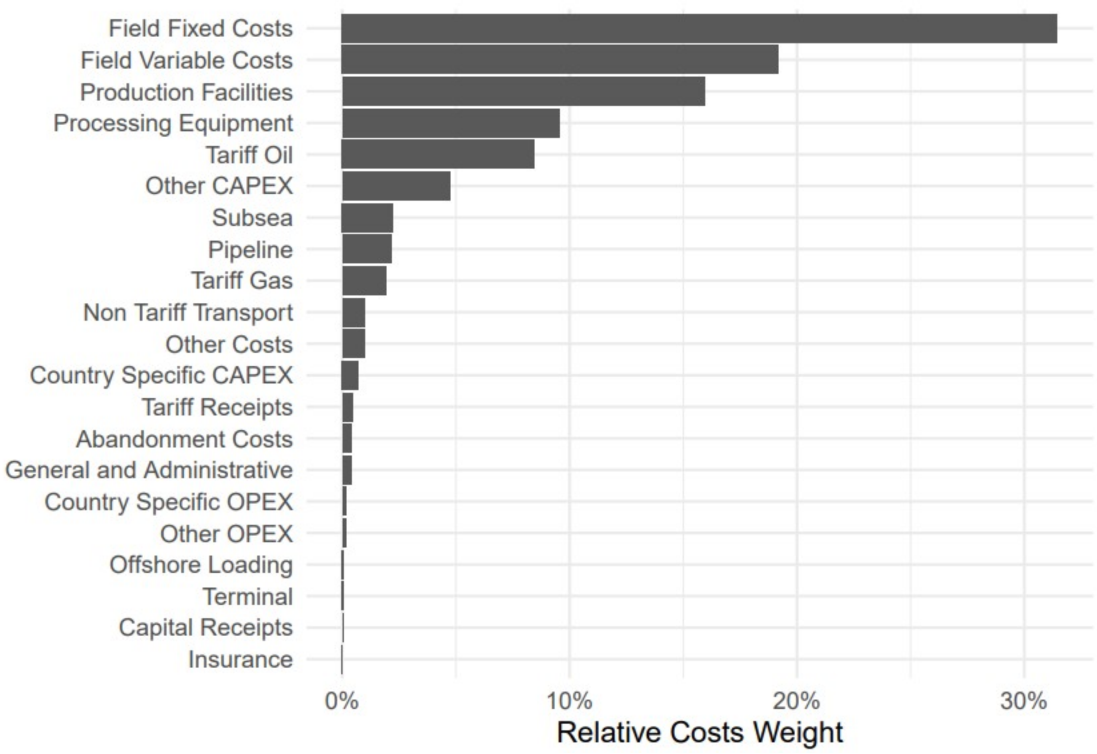

Figure 4.7: Relative weight of the different cost categories.

variable costs ($\sim$ 20%) already account for a half of the total costs a field faces. Figure 4.7 reflects equation (4.4), where θ_0^{Geo} captures the field fixed costs and the joint behavior of $\theta_1^{Geo}, \theta_2^{Geo}$ and θ_3^{Geo} the sum of all the other extraction costs. An explanatory analysis of C_t^i confirms the positive relation of the costs with both the production (see Figure 4.8) and the reserves (see Figure 4.9). However, the relations vary across different geological formations. Light & Medium, Heavy and Shale & Tight Oil depict a concave pattern as the quantity produced increases, whereas Extra Heavy, Sands and Other Oil depict a convex one. The scenario is a little different when visualizing the relation between the costs and the reserves. Heavy, Shale & Tight and Extra Heavy report a concave pattern, Light & Medium a linear pattern, Other Oil a convex pattern, while Sands a non-linear one.

In equation 4.4 the dependent variable is always positive defined. In order to avoid the use of generalized linear model, we estimate it in first differences. More precisely, we estimate five versions of equation 4.4. We start using a Linear Mixed Model, column (1) in Table 4.7, where all the random coefficients are normally distributed. Then, we add a time effect, such that $\epsilon_t^i = \theta_4^{Geo} Q_t^i \, t + \varepsilon_t^i$, to isolate the impact of technological change, column (2), and a quadratic time effect, such that $\epsilon_t^i = \theta_4^{Geo} Q_t^i \, t + \theta_5^{Geo} Q_t^i \, t^2 + \varepsilon_t^i$[12], to allow the technological improvement to have a (linear) trend. Finally, column (5) combines the three fixed effects, such that $\epsilon_t^i = \varphi_t + Geo^i + \eta^i + \varepsilon_t^i$. In all four cases, we assume $\varepsilon_t^i \overset{iid}{\sim} \mathcal{N}(0, \sigma_\varepsilon^2)$.

In all cases the coefficients have the expected sign. Namely, both $\hat{\theta}_1^{Geo}$ and $\hat{\theta}_2^{Geo} + \hat{\theta}_3^{geo}/R_{t-1}^i$ are bigger than zero. The introduction of a constant technological effect does not substantially change the magnitude of the coefficients suggesting the absence of a homogenous shock across fields belonging to a same geological class. The introduction of a linear technological trend diminishes the magnitude of $\hat{\theta}_1^{Geo}$, while increasing the overall convexity of the problem while rising the adjusted R^2 from 17% to 23%. Note that, since the regression is estimated in first differences, these goodness-of-fit measures are in line with standard econometric results[13].

[12] Geo^i is a categorical variable which takes values Light & Medium, Heavy, Extra Heavy, Shale & Tight, Sands and Other Oil.
[13] Further diagnostic analysis are avilable upon request of the reader.

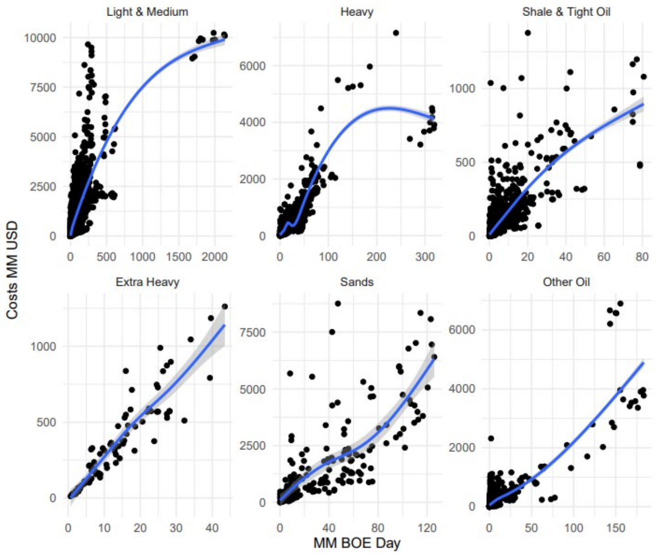

Figure 4.8: Relation between costs and production for different geological formations.

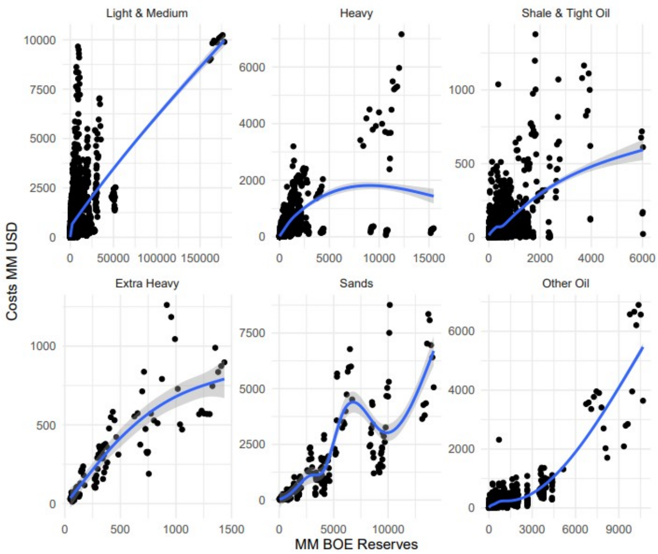

Figure 4.9: Relation between costs and reserves for different geological formations.

Shadow prices

In this section we bridge the theoretical with the empirical model. Namely, we conjugate the results obtained in section Firm Expected Prices and Marginal Extraction Costs with equation (4.6) to obtain

$$\hat{\lambda}_t^i = \underbrace{\bar{P}_t + \hat{\beta}_{1t}(API^i - \overline{API}_t) + \hat{\beta}_{2t}(S^i - \bar{S}_t)}_{\text{Expected Field Future Price}} - \underbrace{\left(\hat{\theta}_1^{Geo} + \hat{\theta}_2^{Geo} Q_t^i + \hat{\theta}_3^{Geo} \frac{Q_t^i}{R_{t-1}^i} + \hat{\theta}_4^{Geo} t + \hat{\theta}_5^{Geo} t^2\right)}_{\text{Marginal Extraction Cost}} . \quad (4.11)$$

The previous equation identifies the profitability of a field at a point in time. In other words, when $\hat{\lambda}_t^i$ is strictly positive the field is lucrative, when it is equal to zero the field breaks-even, and when it is negative the field looses money.

The magnitude of $\hat{\lambda}$ is a versatile measure of the extensive margin of the industry. On a disaggregated level it shows the least profitable fields. According to our estimates all the field are profitable except for one (Kucavo-

Table 4.7: Cost function regression results.

	Dependent variable: Delta Total Costs C_t^i MM USD		
	(1)	(2)	(3)
Quantity	11.090***	11.062***	10.106***
	(0.162)	(0.161)	(0.157)
Quantity squared	−0.004***	−0.004***	−0.004***
	(0.0001)	(0.0001)	(0.0001)
Reserves squared	127.168***	128.797***	167.830***
	(7.157)	(7.135)	(6.946)
Constant Technology	No	Yes	Yes
Linear Technological Trend	No	No	Yes
Observations	17,503	17,503	17,503
Adjusted R^2	0.16	0.17	0.23
Note:			*p<0.1; **p<0.05; ***p<0.01

Arreza (Albania)), which has a negative shadow price. The least profitable field is West Esh El Mallaha A-1 (Egypt), an On Shore Heavy deposit hosting petroleum with an API fo 20.92 and 7% sulphur content. According to our estimates its shadow price in perfect competition is

$$\hat{\lambda}_t^i = 50.39 + \underset{(0.08)}{0.13} * (20.92 - 31.46) - \underset{(1.01)}{2.87} * (7.00 - 1.4) - 28.97 = \frac{3.95\$}{\text{BOE}}\ .$$

Either a decline in selling price or a rise in marginal extraction costs of 4 $ / BOE would make the field unprofitable. In other words, this field identifies the extensive margin of the industry in 2015.

Time trends

Our analysis gives a single-year snapshot between carbon intensities (CIs) and the shadow prices (λ). Namely, we used 2015, the latest year for which a comprehensive data set was available, to pair the two quantities. However, both variables can vary over time.

CIs change over time. This paper cannot provide insights about the dynamics of emissions, since time-series upstream operation data are generally missing on a global scale. However, prior regional works have shown that the energy intensity of crude oil production tends to increase with depletion [Masnadi and Brandt, 2017a, Brandt, 2011]. The increase is mostly due to increased work of fluid lifting and increased fluid injections during secondary and tertiary recovery schemes. This trend, coupled with the shift to unconventional resources, suggests that the CIs of average and marginal petroleum resources is likely to slowly increase over time [Wallington et al., 2017].

Similarly, shadow prices change over time due to variations in oil prices, MCs, and market power. Furthermore, our estimates of the λ rely on eight structural coefficients (two for the field-level oil price, five for the MCs, and one for the market power correction term). With the exception of the two pricing coefficients, which are time varying; all the other coefficients are time-invariant. Thus, major economic variations could render them uninformative. For example, we obtain the MC regressing the extraction costs on the volumes of oil extracted, the depletion level of the field, the geological characteristics of the field, and the technological trend within the oil industry. This structure of the cost function allows the MCs to change over time. In particular, the presence of a (quadratic) technological trend, which interacts with the geological peculiarities of the fields, allows us to isolate the rate of technological change for different types of crude. As time passes the cost of producing one more barrel of different crude types might change because of technological change modelled within our cost function. However, the differentiation between fixed and variable factors of production remains fix. In presence of major technological changes, this could render the structure of costs different from the one modelled by our cost function.

Figure (4.10) plots the estimated marginal cost of production in different years of the WoodMac (WM) dataset from oilfields, sorted in descending order. It is evident that the marginal cost of production in different years for some unconventional crudes with relatively high MC reduces over time. This decrease in MC is mostly associated with shale oil and oil sand fields due to technological improvements captured in our economic model. Reports [Bloomberg, 2019, Markit, 2020, Ihejirika, 2019, Morgan, 2020, Crooks, 2016] in recent years describe relentless production cost-reduction by unconventional producers (e.g., oil sands of Canada and U.S. shale) confirming this dynamic adjustment, in good agreement with Figure (4.10) trend. For example, the cost to construct a new oil sands project was estimated to be between 25% to 33% cheaper in 2018 than in 2014 [Bloomberg, 2019, Markit, 2020]. Deflation in capital costs was a factor, but reengineering efforts such as simplifying project designs, and speeding construction played a role in the reductions. Operating costs fell by more than 40% on average from 2014 to 2018 due to reducing facility downtime and increasing fluid throughput through facilities [Bloomberg, 2019, Markit, 2020]. Conventional light and medium crude producers with lower average MC (see Table 1 of the main text) are less sensitive to technological enhancements in different years.

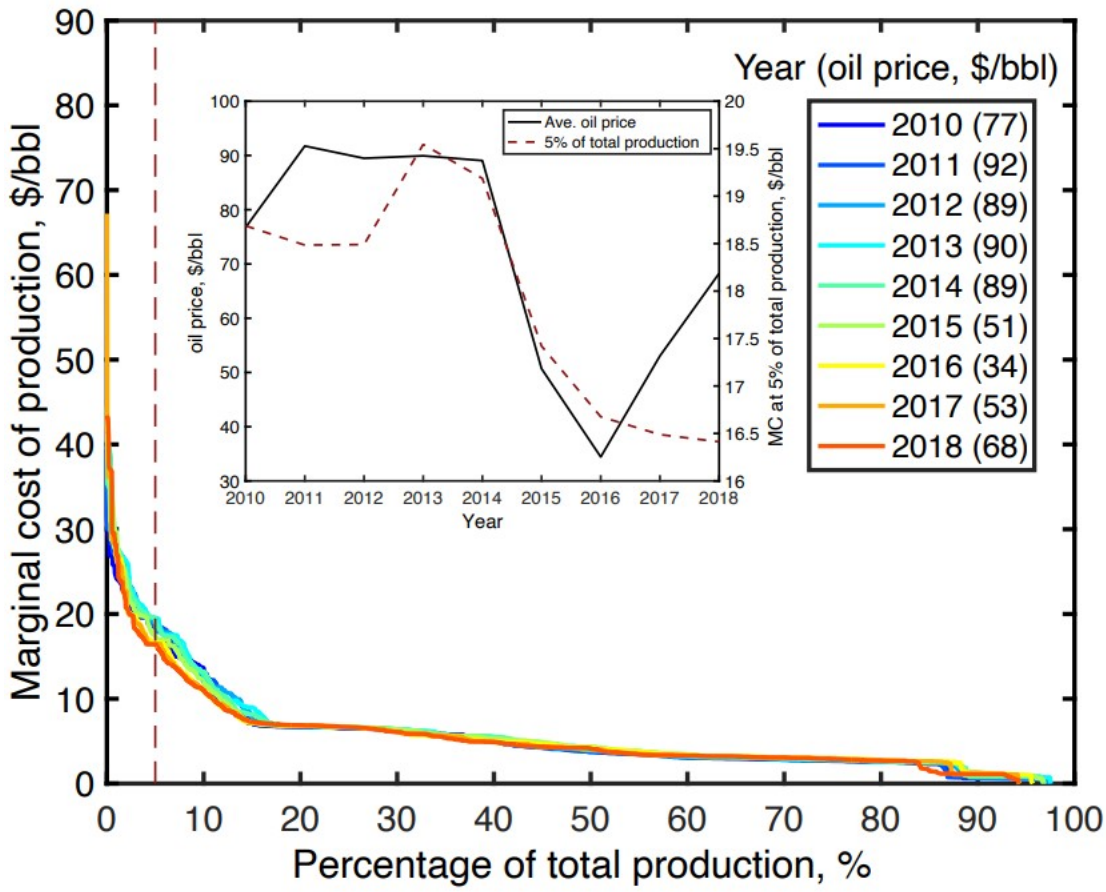

Figure 4.10: Year 2010-2018 descending temporal marginal cost of production of global oilfields versus cumulative volumetric oil production. The oil prices listed in the legend are annual VWA prices in the corresponding year. The inset graph illustrates the average oil price and marginal costs by displacing 5% of total production in different years. See Figure (4.11) for 2.5-15% displacing range of total production in different years.

The MCs are influenced by the status of the global oil market. As the oil price increases, producers increase their output. Since MCs are convex in quantities, the cost per BOE increases. In the same way, as the oil price drops, producers reduce their output and those their MCs. For example, the pronounced oil price drop after 2014 forced the marginal cost downward due to lowering of production and technological improvement. However, our model assumes a fixed production technology. In other words, we assume that the input mix of the production function is incapable to adapt to changes in the relative prices of the different inputs as well as of the output.

Therefore, if technological change within the oil industry renders one of the inputs of the production function no longer fixed but variable, the technological trend (which we modelled in the cost function) is incapable to capture this structural change, since the true (unobserved) new cost function would identify a different minimum than the one we fitted with historical time series.

The un-modelled capacity of the oil industry to change its input mix may offset the tendency of alternative vehicles and fuels to drive oil out of the market. It is generally thought that as alternatives such as EVs become cheaper and consequently start to penetrate in the global transport market, the demand for crude oil will decline, resulting in a drop in the oil price and shutting in of oilfields [Wallington et al., 2017]. However, this effect might be muted by the capacity of the oil industry to change its production set which could hinder penetration of these oil-displacing alternatives.

Figure (4.11) expands Figure (4.10) inset graph and illustrates the average oil price and marginal costs by displacing 2.5-15% of total production in different years.

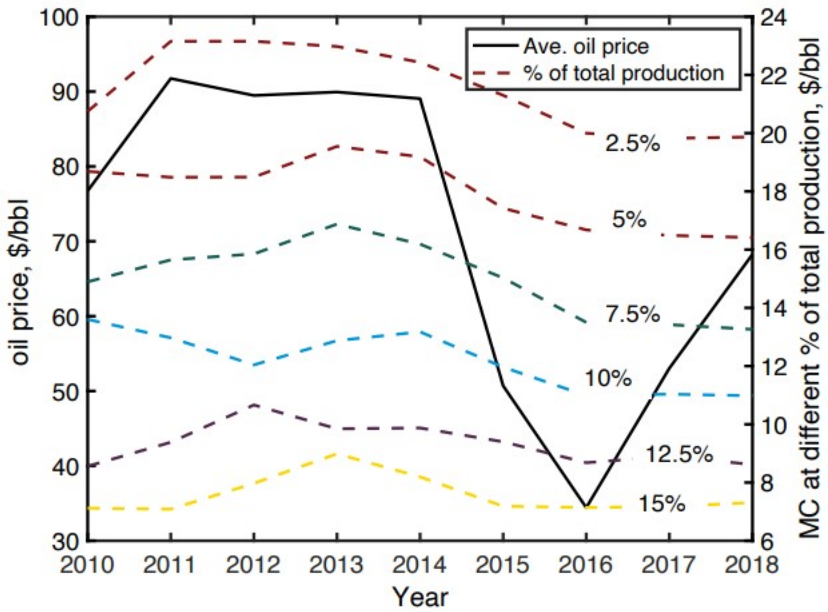

Figure 4.11: Average oil price and marginal costs by displacing 2.5-15% of total production in different years.

Effect of oil demand elasticity

Following the main text Figure 2 discussions, Figure (4.12) and Figure (4.13) show upstream cumulative volume weighted average CIs (right-axis) and sorted shadow prices (left-axis) of global oilfields for oligopolistic and cartel competition model using different oil demand elasticities versus the percentage of total oil production in 2015, respectively. The less elastic consumers are, the more the market power impacts the SP.

The cartel shadow price in the first percentages of total production equals zero (see Figure (4.13) and Figure 2 of the main text). For example, why Venezuela, a country pertaining to the cartel, would accept a situation in which its profitability is null? As in every production model, allocating all factors of production is the optimal choice. Namely, between two choices of 1) do not produce, 2) produce at a profit equal to zero; producers usually choose to keep producing at zero profit for a period of time. Between the choice of not using the factors of production and using them as efficiently as the producer can but without making any profit, it is optimal to use them.

Furthermore, note that Venezuela,s profits would be smaller under the decisions to exit the cartel. Namely, profits would be smaller than zero. Therefore, the optimal decision is to stay in the cartel and produce at a profit equal to zero. In addition, a country like Venezuela runs a national oil company and economic profitably is only one decision factor among many others (national security, national supply, government subsidies, etc.) to keep or terminating the production. As we stated in the paper, we are not able to capture many other factors in the economic model.

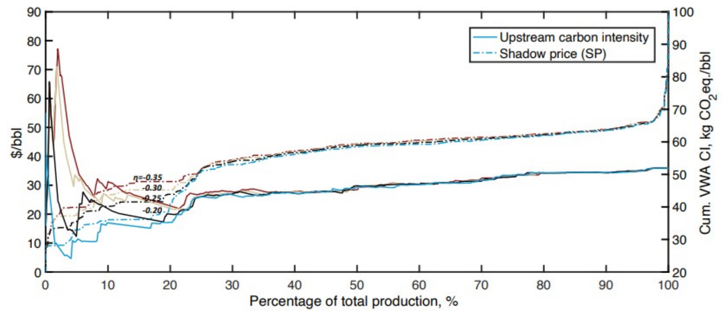

Figure 4.12: Upstream cumulative volume weighted average CIs (right-axis) and sorted shadow prices (left-axis) of global oilfields for oligopolistic competition model using different oil demand elasticities versus the percentage of total oil production in 2015.

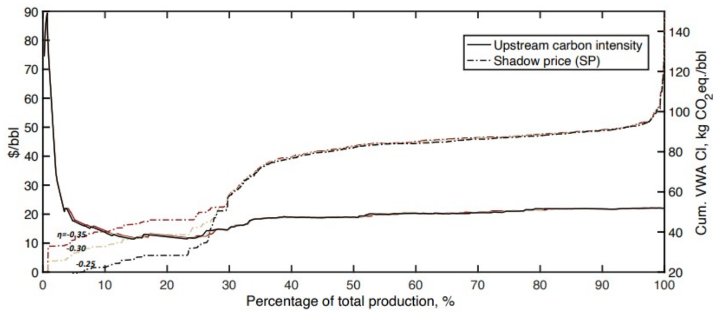

Figure 4.13: Upstream cumulative volume weighted average CIs (right-axis) and sorted shadow prices (left-axis) of global oilfields for cartel competition model using different oil demand elasticities versus the percentage of total oil production in 2015.

Global field-level merit base curves

Following the main text Figure 2 discussions, Figure (4.14) shows the global field-level perfect competition shadow-price (SP-PC) merit base curve, identifying fields using three different types of descriptors. Each bar width reflects the oil production of a particular field in 2015. Figures (4.15) and (4.16) present the corresponding results for Oligopoly and Cartel economic cases. Figure (4.14)(a) (and Figures (4.15)(a) and (4.16)(a)) characterizes each field based on its corresponding CI percentile. For both PC and Oligopoly cases, the least profitable 10% of global production volume includes several fields with high CI (dark green) mainly due to crudes high density and/or high gas flaring rates (see Figure (4.14)(b)). Approximately 82% (~ 5.8 mmbbl/d) and 50% (~ 3.5 mmbbl/d) of these marginal crudes correspond to heavy, extra heavy, and oil sands unconventional fields based on PC and Oligopoly cases, respectively (see Figure (4.14)). Incorporating a market power correction term (especially the Cartel case) in computing the SP shifts several conventional producers with low CI towards the margin. However, several high CI heavy crudes, mostly located in Venezuela, remain at margin for both oligopolistic and cartel competition models.

In the current world where upstream GHG emissions are not regulated or priced globally, a high CI is not always associated with low profitability. Figures (4.14), (4.15), and (4.16) illustrate that many fields in the highest 10% of global profitability (90-100% of x-axis) have high shadow prices but release high GHG emissions mainly due to routine flaring of large amounts of gas. These economically productive fields are mostly light medium onshore fields.

Light and medium deepwater offshore fields are consistently among the most profitable fields in all economic models (see the orange bars in Figures (4.14), (4.15), and (4.16)). The oilfields in the dataset are ~ 25 years old in average with deepwater fields as the youngest geological group. In particular, the ones containing light medium crude are in average ~ 9 years old. A younger field tends to correlate with stronger reservoir drive, less water production, and thus larger per-well productivity, reducing the extraction energy intensity and marginal costs of production. The competitiveness of these formations could change over time as they become more depleted.

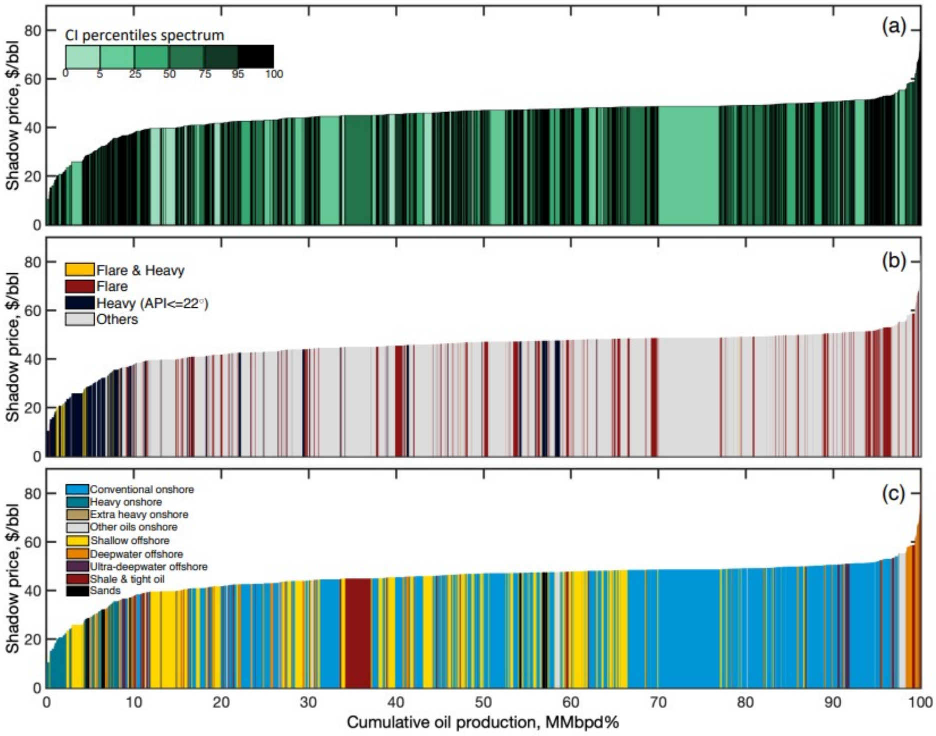

Figure 4.14: Shadow price supply curve of 1,933 parent global oilfields versus the cumulative volumetric oil production using perfect competition economic case. Bar width reflects the oil production of a particular field in 2015. The bars are colored based on: (a) CI percentiles (b) contribution of high flaring (Flare with FOR $>$ 75th percentile of all fields) and oil density (Heavy with API gravity $<= 22°$)(c) oilfield geology.

Extensive margin versus stranding

Several papers [Kilian, 2008, 2009, Ramcharran, 2002, Güntner, 2014] studied the effect of exogenous demand shocks and price changes on oil producers and quantified the relative importance of demand and supply shocks in the global crude oil market and economy in general. These studies typically utilize historical empirical data to perform economic analysis without incorporating environmental indices. Other research groups [McGlade and Ekins, 2014, Bauer et al., 2016, Mercure et al., 2018] utilized economics and integrated assessment models (IAM) to explore the impacts of global environmental policies (e.g., based on limiting global warming to 2°C) and technological trajectories on fossil fuel markets, and to analyze the characteristics of stranded fossil fuel assets. Unlike the present work, these studies include forthcoming fossil fuel resources in order to model long-term future scenarios. They tend to include broad resource classes and lack specific data on individual operating oilfields. In contrast, the more granular, near-term results presented in this work could inform such macroeconomic studies by providing estimates of the heterogeneity in marginal cost of production and corresponding CI of different crude types as well as shedding light on impacts of short- and mid-term demand reductions.

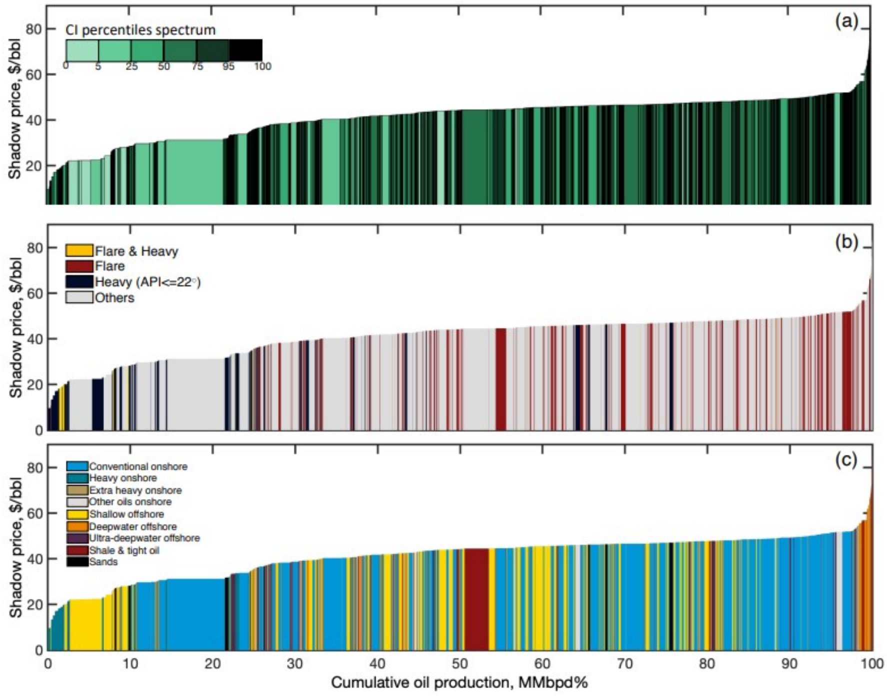

Figure 4.15: Shadow price supply curve of 1933 parent global oilfields versus the cumulative volumetric oil production using oligopolistic competition economic model (oil demand elasticity=-0.35). Bar width reflects the oil production of a particular field in 2015. The bars are colored based on: (a) CI percentiles (b) contribution of high flaring (Flare with FOR > 75th percentile of all fields) and oil density (Heavy with API gravity <= 22°)(c) oilfield geology.

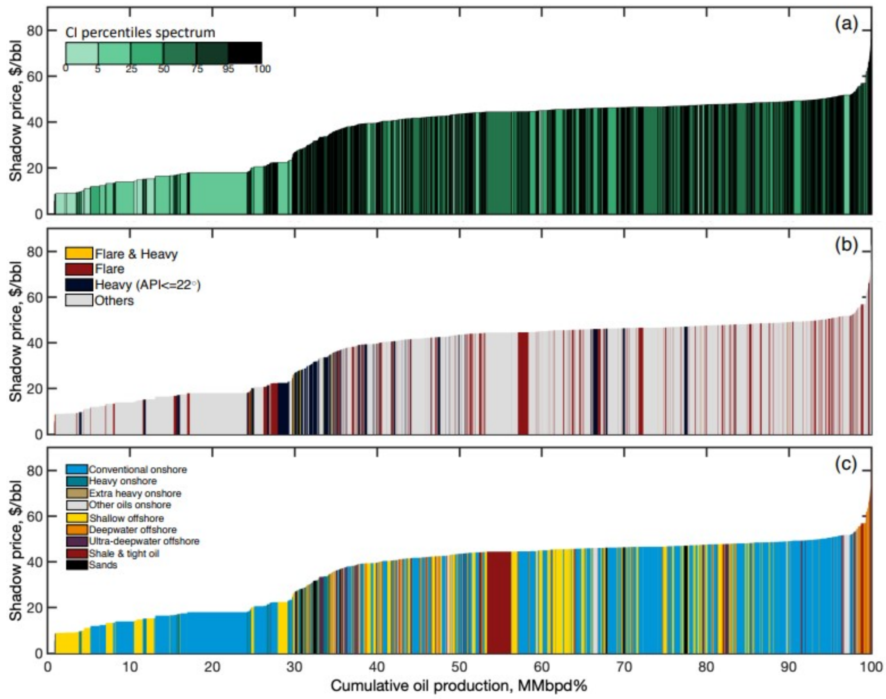

Figure 4.16: Shadow price supply curve of 1933 parent global oilfields versus the cumulative volumetric oil production using cartel competition economic model (oil demand elasticity=-0.35). Bar width reflects the oil production of a particular field in 2015. The bars are colored based on: (a) CI percentiles (b) contribution of high flaring (Flare with FOR > 75th percentile of all fields) and oil density (Heavy with API gravity <= 22°)(c) oilfield geology.

Life cycle analysis

Upstream mitigation estimate

Following Table 2 of the main text, Figure (4.17) shows annual upstream carbon mitigation potential versus the amount of oil displaced using perfect competition, oligopolistic competition, and cartel competition. For most of the volume displaced, PC case results in the highest mitigation potential as more unconventional fields with high CI are at margin in PC. However, incorporating a market power correction term in Oligo and Cartel cases in computing the SP shifts several conventional producers with low CI towards the margin resulting in less GHG emissions mitigation potential than PC case.

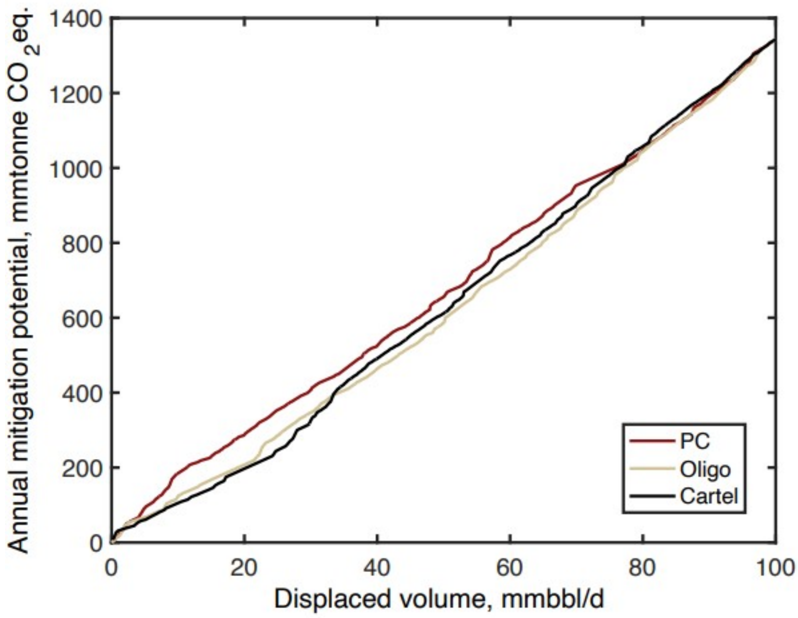

Figure 4.17: Annual GHG emissions mitigation potential vs the amount of oil displaced.

Well-to-wheel mitigation estimate

Table 2 of the main text presents the upstream mitigation potential of different demand reduction scenarios. In order to estimate the well-to-wheel mitigation potential, we need to couple them with mid (i.e., refining) and downstream (i.e., end-use emissions where oil products are combusted) emissions reduction. Estimating them requires granular data, not dissimilar from the one used for upstream part of the model [Oil Climate Index (OCI), 2020]. Unfortunately, we do not have access to refinery-level data and/or consumption data of finite products at a global level.

However, we know that the upstream emissions reduction reverberates on refining emissions across two dimensions. First, if most displaced crude is heavy and extra heavy, then the average traded crude will have a higher API gravity. This change will have economic and environmental consequences. From an economic perspective, the displacement of heavy crude is equivalent to a general improvement in the quality of the input used by petroleum refineries. From an environmental perspective, the displacement of heavy crude corresponds to a reduction - all else equal - of refinery emissions. This second result is due to the inverse correlation between API gravity and carbon intensity during the refining process [Jing et al., 2020]. For example, in the linear regression presented in Jing et al. [2020],

$$CI^r = \varphi_0 + \varphi_1 API^r + \varsigma^r \ ,$$

where CI^r is the carbon intensity of refinery r, API^r is the average gravity of the processed crude, and ς^r is a normally distributed idiosyncratic shock with finite homoskedastic variance $\varsigma_t^r \overset{iid}{\sim} \mathcal{N}(0, \sigma_\varsigma^2)$, the estimate coefficient $\hat{\varphi}_1$ is negative and statistically significant, see Figure (4.18). Using the estimated coefficients $(\hat{\varphi}_0, \hat{\varphi}_1)$, as presented in Jing et al. [2020], we compute a possible range of GHG emissions mitigation.

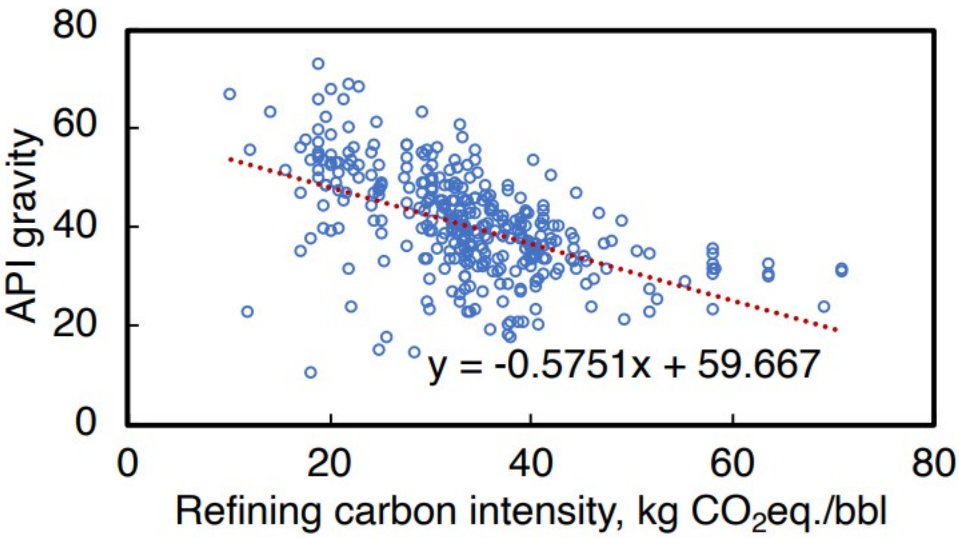

Figure 4.18: Estimated crude oil refining carbon intensity as a function of API gravity [Jing et al., 2020].

The effect of an oil demand reduction on downstream emissions are hard to calculate because, contrary to up and midstream emissions, they are sensitive to the type of demand shock experienced. For example, a symmetric global shock in income, like the 2008-2009 financial crisis, would reduce the demand of finite product in the same proportion, as long as consumer preferences are (intra-temporally) homothetic. As a result, if demand for oil decreases by 5%, due to a decrease in global income, and gasoline represents the 25% of the global oil product demand, we should observe a decline of 1.25% of gasoline combustion emissions. Therefore, the emission reduction would equal 1.25% multiply by the gasoline emission factor. To the contrary, an asymmetric shock in the world income, like the COVID-19 pandemic, would cause a change in the composition of global oil derived product demand. For instance, the demand of jet fuel will decline more than the demand for gasoline or ultra-low sulfur diesel because the pandemic impacted the flying industry more than the transportation one. The same applies to the diffusion of specific technologies (e.g. EVs). In other words, the idea to use abstract round number shocks frees us from stringent hypothesis on policy or technology scenarios and it does not impact our capacity to compute reduction in up and midstream sector. However, it is incapable to be precise about displacement of downstream emissions. This is a limitation of our approach, which could be solved using disaggregated consumption data to micro-found the oil demand elasticity. We try to overcome this limitation using Oil Climate Index (OCI) [2020] emission factors of 333-552 kg CO_2eq./bbl for the end-use emissions.

Table (4.8) presents our annual estimates on the well-to-wheel reduction potential of small (2.5%), COVID-19 (5%), and medium (10%) demand shocks.

Table 4.8: Estimated well-to-wheel GHG emissions mitigation potential measured in mmtonne CO_2eq. according to three economic models for three different demand shocks.

	Perfect Competition			Oligopoly			Cartel		
	2.5%	5%	10%	2.5%	5%	10%	2.5%	5%	10%
Upstream	54	92	184	53	69	124	39	61	109
Midstream	53	100	187	52	93	169	40	82	150
Downstream	219-363	438-725	863-1431	219-363	438-725	863-1431	219-363	438-725	863-1431
Total	326-470	630-917	1234-1802	324-468	600-887	1156-1723	298-442	581-868	1122-1689

Sector mitigation measured in mmtonne CO_2eq.

Limitations and future research

We develop a techno-economic analysis of the global oil supply. Our main goal is to connect the GHG emissions of different oilfields with their economic characteristics. The link between these two quantities allows us to identify the volume of CO_2eq. emissions associated with the extensive margin of the oil industry.

According to our estimates, several high carbon intensive heavy fields are consistently at margin and are likely to be displaced by demand reductions in all three economic cases. Our economic cases also show that despite the fact that the average crude oil production cost of some of the major global producers (i.e., OPEC member countries) is relatively low, their role to regulate the global oil market sacrifices the field-level profitability in the short term.

As mentioned in the main text, we only included the production economics and identified the extensive margin of the oil industry. However, various other dynamic forces such as production agreements, region-specific fiscal regimes, regulations, geopolitics, technical advances, and incidental events could move a particular oilfield toward or away from the margin. Two pertinent examples: 1) governments have announced a variety of policy responses designed at lessening the consequences of the COVID-19 crisis [IMF, 2020], and several governments have announced their intentions to earmark portions of their packages for a "green recovery". In the post COVID-19 context, this means that, beyond fiscal stimuli, recovery packages should include incentives, policies, taxes or rebates, mandates and penalties, and other supportive regulations to facilitate the achievement of long-term climate goals. However most governments have not yet signaled how they intend to spend their money. Recovery packages that include unconditional oil & gas company bailouts may keep the marginal producers in business in the short run, while increasing the number of assets that will someday be stranded [IMF, 2020, Brandt et al., 2021]. 2) fields with high routine flaring and a medium shadow price would be significantly affected by a carbon tax and may become marginal even if they are currently far from the margin.

While granular at the upstream level, our analysis lacks a detailed investigation of the techno-economics peculiarities of the mid and downstream sector. Future research could focus on them. Such work would need to model midstream emissions as well as the corresponding refining economics. For example, it would be interesting to couple our findings with midstream models where refineries solve a two-stage problem. First, they choose how much capacity to install for different transportation fuels. Then, they compete on prices with other refineries to buy raw crude oil. Expanding our research to the refinery level would shade light on the carbon intensity of oil networks (from field A to refinery B versus from field C to refinery D) and discover the shadow price of different global connections. Sad differently, the modeling of global oil networks would identify which routes are resilient to demand-driven shocks and what is their environmental footprint.

Furthermore, it would be important to couple oilfields-refineries networks with a micro-foundation of the elasticity of the global oil demand. In our current framework, different values of η come from different estimation strategies. However, they all rely on macro-econometric regressions, which use aggregate quantities and prices. As a result, they cannot distinguish across the elasticities of different oil-derived products. Using microdata on transportation, it would be possible to estimate the demand elasticity of products like gasoline, ultra low sulfur diesel, and jet fuel. The latter could be use to correct the shadow price of different oilfields-refineries networks by their market power conditional on which final product they sell.

Finally, it is unclear how the time frame for petroleum supply-demand considerations interact with the long-term time frame for climate change. Do near-term reductions in petroleum demand simply delay, but not prevent, the consumption of marginal petroleum over the relevant time frame? Or can such a delay buy time for eventual changes in technology or climate policy, such that less carbon intensive petroleum is extracted?

Chapter 5

The economic and environmental consequences of the petroleum industry extensive margin

This chapter is entirely based on Benini, Giacomo; Brandt, Adam; Dotti, Valerio; El-Houjeiri, Hassan (2023): The Economic and Environmental Consequences of the Petroleum Industry Extensive Margin. Department of Economics, Ca' Foscari University of Venice, Working Paper Series No. 14/WP/2023.

5.1 Introduction

The global oil market is an immense network that links thousands of oilfields with billions of consumers, relying on vast infrastructures, including over two million kilometers of pipelines and extensive merchant shipping capacity [Cruz, 2013]. This network has facilitated the transformation of regional markets into a unified global system, where opportunities for price arbitrage are increasingly scarce due to logistical and financial standardization [Adelman, 1984, Nordhaus, 2009]. Over time, collaboration between logistics and financial sectors has streamlined oil demand globally, reducing inefficiencies and reinforcing a cohesive market [Milonas and Henker, 2001]. At the same time, technological advancements in extraction have expanded the variety and complexity of global oil supplies, allowing access to new, challenging reservoirs like shale, oil sands, and deepwater deposits, which has diversified extraction costs across regions [Maugeri, 2012]. Consequently, oilfields now vary significantly in profitability, and cost dynamics are shaped by specific geological and technological factors, emphasizing the sector's shift from a homogenous market to one marked by diverse extraction profiles and economic variability [Roberts et al., 2019, Masnadi et al., 2021].

The diversification stems from the heterogeneous role that reservoir pressure plays in defining the geometry of the cost function. When technology allowed producers to extract only high-viscosity oil trapped in highly permeable rocks (Light & Medium oil), extraction costs were modelled as a function of the volume of oil extracted *and* the amount of recoverable reserves [Livernois and Uhler, 1987, Pesaran, 1990, Favero, 1992]. These two quantities capture the entire extraction process, which becomes a proxy for the logical sequence: more reserves → more pressure → fewer inputs needed to extract a given quantity of oil. In this framework, discovering one barrel should compensate, in terms of marginal costs, for the extraction of one barrel. As a result, the oil sector was no different from any other exhaustible resource industry [Solow and Wan, 1976], like coal or copper mining, where, once controlled for ore grade, only the size of the mine impacts the marginal extraction costs [Zimmerman, 1977, Aguirregabiria and Luengo, 2016].

The ability to extract low-viscosity oil trapped in high-permeability rocks (Heavy & Extra Heavy oil), high-viscosity oil trapped in low-permeability rocks (Shale & Tight oil), and mine for bitumen (Oil Sands) has created a spectrum of the importance of natural pressure in determining the variable costs of an oilfield. Moving from Sands $\rightarrow$ Heavy & Extra Heavy $\rightarrow$ Light & Medium $\rightarrow$ Shale & Tight, we observe an increasing role of natural pressure in determining the field variable costs. Conversely, moving in the opposite direction, we see an increasing role of inputs such as steam, water, electricity, and labour, which substitute for natural pressure in the production process. In each of these formations, extraction costs emerge as the interaction between the geological characteristics of the oilfield and the endogenous production decisions of the firm's management.

To capture this fundamental change, we construct an extraction-exploration model where oilfields are risk-neutral firms facing heterogeneous revenues and costs. We quantify the firm's mark-up by estimating their selling prices and marginal extraction costs. We exploit the theoretical framework to derive a field-level pricing equation, which separates global demand-driven shocks from the impact of crude-specific characteristics on the value of the extracted oil. Then, we use the future prices of publicly traded oil classes and the cost of imported crude oils in the United States to estimate it.

In a similar way, we derive a micro-founded field-level cost function, which disentangles the impact of natural pressure from the one of all the other factors of production. To achieve this goal, we build on the formulation proposed by Anderson et al. [2018] allowing for the possibility that discoveries and depletion affect the firm's marginal costs through separate channels with potentially different magnitudes. This increased flexibility is particularly useful to model the marginal costs of non-conventional formations. Then, we use the WoodMac Upstream Data Tool [Mackenzie, 2018] to estimate the cost function.

Subtracting the marginal extraction costs from the marginal revenues, we obtain the shadow price of discovered oil. In other words, we approximate how much money a firm is willing to pay in order to manage an extra barrel of oil located in a particular field at a specific point in time. Ordering these values, from the smallest to the largest, we construct a global *merit order curve*, which identifies how profitable the different segments of the supply are.

Once identified the oil industry extensive margin, we analyse the economic and the environmental consequences of its displacement. From an economic prospective, we multiply the amount of technically recoverable reserves currently available, which would become unprofitable by a marginal change in the market conditions, by their estimated selling price to obtain the monetary value of the stranded resources. According to our estimate, decreasing the global oil demand by 1% would result in stranding circa 15.56 billion barrels of oil with a commercial value of approximately 578.65 billion US Dollars. From an environmental prospective, we multiply the production volumes, which would become unprofitable by a marginal change in the market conditions, by their upstream carbon intensity to obtain the volume of greenhouse gas savings. To compute the entire well-to-wheel emission reduction, we estimate and add the mid- and downstream savings. According to our results, decreasing the global oil demand by 1% would result in reducing upstream emissions by 24.95 $MMtCO_2e$ per year. This quantity is approximately equal to the annual carbon footprint of 5.4 million cars (i.e., more than 5.3% of all the privately-owned cars registered in the United States). The corresponding well-to-wheel emission savings equal approximately 123.66 $MMtCO_2e$. We show that these results are robust to an imperfect competition scenario where large International Oil Companies play a game in quantities, while National Oil Companies member of the Organization of the Petroleum Exporting Countries (OPEC) behave as members of a cartel.

To the best of our knowledge, the present paper is the first attempt to derive how oil quality, global demand trends, reservoir pressure, market power, and upstream emissions are intertwined starting from a rigorous theoretical formulation, which begins from the behaviour of a single well and ends with the impact of OPEC on the global average oil price.

5.2 The Oil Shadow Price

We study a general equilibrium economy featuring four types of players: oil firms, refineries, consumers, and the government. In this section we focus solely on the key actors, the oil firms. Within this framework, every oilfield is owned by an international or a national oil company, which may exert market power [Golombek et al., 2018, Asker et al., 2019]. International oil companies play a game in quantities, while the national oil companies member of the Organization of Petroleum Exporting Countries (OPEC) collude in the form of a cartel[1]. In this framework, K risk-neutral oil firms (or groups of colluding firms) compete à la Cournot. Each firm $k = 1, 2, ..., K$ controls $n(k)$ oilfields and maximizes the present discounted value of the sum of present and future profits while anticipating the effect of its production choices on the equilibrium prices. Firm k decides in period t its production and investment plan for all periods $t + s$ with $s = 0, 1, 2, ...$ Its intra-temporal profits,

$$\Pi^k_{t+s} = \sum_{i=1}^{n(k)} \Pi^i_{t+s} = \sum_{i=1}^{n(k)} \left[P^i_{t+s} Q^i_{t+s} - C^i_{t+s}(Q^i_{t+s}, L^i_{t+s-1}, M^i_{t+s-1}, Geo^i, \epsilon^i_{t+s}) - W^i_{t+s}\right] , \tag{5.1}$$

are the difference between revenues, extraction costs, and discovery costs, aggregated across all the controlled fields. The field revenues are the product between the oil price P^i_{t+s} and the quantity of oil produced Q^i_{t+s}. The oil price is a function of the chemical characteristics of the crude produced by field i. The functional form of the cost function is obtained as the outcome of a cost minimization problem solved by the firm in each period, whose details are outlined in section 5.3.2. The resulting field-level extraction costs C^i_{t+s} are a function of the quantity of oil extracted and of the quantity of reserves available when the production starts. Let D^i_r denote the new discoveries in period r. Available reserves are equivalent to the initial size of the deposit R^i plus the discoveries occurred after the initial assessment of the field $L^i_{t+s-1} = R^i + \sum_{r=1}^{t+s-1} D^i_r$ and minus the sum of extracted liquids $M^i_{t+s-1} = \sum_{r=1}^{t+s-1} Q^i_r$. Finally, the costs are function of the peculiar geology of the field Geo^i and of an idiosyncratic shock ϵ^i_{t+s}. The exploration costs, W^i_{t+s}, are the expenses incurred to discover new oil located in field i.

Every field faces two physical constraints. The first one,

$$L^i_{t+s} \leq L^i_{t+s-1} + D^i_{t+s}(W^i_{t+s}, L^i_{t+s-1}, \xi^i_{t+s}) , \tag{5.2}$$

restrains the cumulative amount of discoveries at time $t + s$ to be lower or equal to the one obtained till time $t + s - 1$ plus the new ones $D^i_{t+s}(.)$[2], which are function of the exploration expenditures W^i_{t+s}, of the quantity of discoveries made in the past L^i_{t+s-1}, and of an idiosyncratic shock ξ^i_{t+s}.

The second constraint ensures that the cumulative depletion exerted until $t + s - 1$, denoted by M^i_{t+s-1}, plus the production at time $t + s$, equals or exceeds[3] the cumulative depletion at time $t + s$,

$$M^i_{t+s} \geq M^i_{t+s-1} + Q^i_{t+s} . \tag{5.3}$$

Each firm in period t decides the volumes of production, $Q^i_t, Q^i_{t+1}, ...$ and the rates of investment in exploratory $W^i_t, W^i_{t+1}, ...$ by maximizing the expected discounted future stream of profits. The decision is conditioned by the available information set Ω^k_{t+s-1} which includes previous prices, quantities, and shocks,

$$\Omega^k_{t+s-1} = \left\{ \left[P^i_s\right]^{t-1}_{s=0}, \ \left[Q^i_s, W^i_s, M^i_s, L^i_s\right]^{t-1}_{s=0}, \ \left[\epsilon^i_s, \xi^i_s\right]^{t-1}_{s=0} \right\}^{n(k)}_{i=1} .$$

[1] This assumption is rather extreme and not meant to provide a realistic description of OPEC decision-making process. Conversely, it defines a benchmark case that accounts for the maximum possible degree of market power given the current structure of the oil industry.

[2] The inequality captures the implicit assumption that the firm is free to ignore/disregard newly discovered oil in its assessment of total available reserves.

[3] The inequality captures the implicit assumption that the firm can dispose of extracted oil for free.

The resulting inter-temporal profit maximization problem can be solved using standard methods. The Lagrangian writes

$$\mathcal{L}_t^k = \mathbb{E}_{t-1}\Bigg\{ \sum_{i=1}^{n(k)} \sum_{s=0}^{\infty} \kappa^s \left[\Pi_{t+s}^i + \lambda_{t+s}^i [M_{t+s}^i - M_{t+s-1}^i - Q_{t+s}^i] + \right. \\ \left. + \mu_{t+s}^i \left[L_{t+s-1}^i + D_{t+s}^i - L_{t+s}^i\right]\right] |\Omega_{t+s-1}^k \Bigg\} ,$$

where $0 \leq \kappa < 1$ is the inter-temporal discount factor. To obtain a tractable formula for the shadow price, we exploit the demand side of the economy (i.e., the behaviour of oil refineries). Oil firms internalize the market-clearing condition so that in each period the equilibrium price must be such that the demand of oil from field i equals its supply,

$$Q_{t+s}^i = AD_{t+s}^i \quad \forall i = 1, 2, 3, ...,$$

where AD_{t+s}^i is the aggregate demand for oil produced by field i. Under relatively mild assumptions, we show that the effect of an increase in the quantity produced by field i in period $t + s$,

$$\frac{\partial}{\partial Q_{t+s}^i} \left\{ \mathbb{E}\left[\sum_{j=1}^{n(k)} P_{t+s}^j Q_{t+s}^j |\Omega_{t-1}^k \right] \right\} = \mathbb{E}\left[P_{t+s}^i + \sum_{j=1}^{n(k)} \frac{\partial P_{t+s}^j}{\partial AD_{t+s}^i} Q_{t+s}^j |\Omega_{t+s-1}^k \right] , \tag{5.4}$$

equals $(1 + MS_{t+s}^k / EL_P)\mathbb{E}_{t-1}[P_t^i|\Omega_{t-1}^k]$, where $MS_t^k = \sum_{j=1}^{n(k)} P_t^j Q_t^j / \sum_{i=1}^{n} P_t^i Q_t^i$ is the market share of firm k in period $t + s$ and EL_P is the price elasticity of the global oil demand. As a result, for finite values of EL_P, firm k exerts a positive degree of market power.

The shadow price of discovered oil in field i in period t,

$$\mathbb{E}_{t-1}[\lambda_t^i|\Omega_{t-1}^k] = \left(1 + \frac{MS_t^k}{EL_P}\right) \mathbb{E}_{t-1}[P_t^i|\Omega_{t-1}^k] - \mathbb{E}_{t-1}\left[\frac{\partial C_t^i(.)}{\partial Q_t^i} \middle| \Omega_{t-1}^k \right] , \tag{5.5}$$

equals the perfect competition one only re-scaled by the market power correction term identified by equation 5.4. The latter is the expected market share enjoyed by firm (or group of firms) k, which is a pure number defined between zero and one, divided by the price elasticity of global oil demand. In other words, the market power correction term divides the capacity of firms to influence the global reference price by the extent to which the demand side of the oil market responds to changes in the aggregate supply.

The closer its magnitude is to zero the more the decision of the field management shifts from "*how much should the field produce?*" (intensive margin) to "*should the field keep producing or cease business operations?*" (extensive margin). While our analysis allows for both types of decisions, our empirical exercise focuses exclusively on extensive-margin choices.

5.3 Empirical Analysis

5.3.1 Expected Prices

To the best of our knowledge, business intelligence companies do not provide data about the selling prices of individual oilfields. However, we know that such prices depends upon the global price and the chemical characteristics of the crude oil [Lanza et al., 2005, Fattouh, 2010]. Making use of the optimality conditions of the demand side of the theoretical model, we structurally derive a field-level expected price equation,

$$\mathbb{E}_{t-1}[P_t^i|\Omega_{t-1}^i] = \bar{P}_t + \beta_{1t}(API^i - \overline{API}_t) + \beta_{2t}(S^i - \bar{S}_t) , \tag{5.6}$$

where the price at which field i expects to sell its output equals a global reference price $\bar{P}_t$ adjusted by the delta between the gravity of field i and the average gravity of the oil traded at time t $(API^i - \overline{API}_t)$ and by the delta between the sulfur content of field i and the average sulfur content of the oil traded at time t $(S^i - \bar{S}_t)$. We take as reference price the average price at which United States refineries import different streams of crude [EIA, 2020c] and as $(\overline{API}_t, \bar{S}_t)$ the average gravity and sulfur content of crude imported in the United States [EIA, 2020a]. $\bar{P}_t$ is measured in United States dollars per Barrel of Oil Equivalent ($/BOE), while $(API^i - \overline{API}_t)$ and $(S^i - \bar{S}_t)$ are dimensionless quantities expressed as pure numbers. As a result, (β_{1t}, β_{2t}) are measured in $/BOE. Both these coefficients might vary over time. The variations could be due to a change in the composition of the demand for oil derived products, in the technology employed by the refineries, or in a combination of the two. For example, an increase in the relative demand for light products, like gasoline and jet fuel increases, could boost the impact of $(API^i - \overline{API}_t)$ on $\mathbb{E}_{t-1}[P_t^i|\Omega_{t-1}^i]$. Conversely, if technological progress allows refineries to produce lighter products using heavier oils without facing higher operational costs, the impact of $(API^i - \overline{API}_t)$ on $\mathbb{E}_{t-1}[P_t^i|\Omega_{t-1}^i]$ could decrease. Lastly, an interplay between these two effects may occur.

We do not observe $\mathbb{E}_{t-1}[P_t^i|\Omega_{t-1}^i]$. However, we observe the prices of several crude mixtures, which group oils coming from different producers into a tradable class. Using the Energy Information Administration dataset and the PSA Management and Services BV database, we collect the yearly future prices[4] and the chemical characteristics of twenty-three oil classes over the time interval 1978-2018 [EIA, 2021b, of Canada , PSA]. Table 5.1 provides the summary statistics of the future prices, the gravity, and the sulfur content of every class. Figure 5.1 shows how the future prices (FP) respond to the interaction between demand and supply. To the contrary, Figure 5.2 shows that, for any given average real price, the lighter and sweeter the crude stream is (high API - low S), the more valuable it becomes.

Table 5.1: Summary Statistics of $FP_t(z)$.

Oil Class (z)	Country of Origin	Mean	SD	Min	Max	API	S
Arabian Light	Saudi Arabia	40.39	29.36	12.36	109.43	32.8	1.97
Arabian Medium	Saudi Arabia	40.70	29.34	10.86	107.12	30.2	2.59
Basrah Light	Iraq	76.11	25.70	39.90	106.93	30.5	2.90
Berri	Saudi Arabia	78.82	25.79	45.62	110.77	38.5	1.50
Bonny Light	Nigeria	42.21	30.84	13.62	117.70	33.4	0.16
Bow River Heavy	Canada	33.96	22.82	10.41	84.29	24.7	2.10
Brent Crude	United Kingdom	28.10	13.30	13.94	64.60	38.3	0.37
Cabinda	Angola	26.90	13.92	12.69	69.17	32.4	0.13
Forcados Blend	Nigeria	32.34	22.95	14.35	111.07	30.8	0.16
Furrial	Venezuela	18.27	4.26	12.24	28.23	30.0	1.06
Leona	Venezuela	20.98	9.36	9.79	51.55	24.0	1.50
Light Sour Blend	Canada	69.09	20.51	40.04	96.52	64.0	3.00
Lloydminster	Canada	33.88	23.95	10.15	82.50	20.9	3.50
Marlim	Brazil	78.42	27.83	47.77	114.32	19.6	0.67
Mayan	Mexico	36.01	27.09	9.21	100.29	21.8	3.33
Merey	Venezuela	72.31	24.94	38.97	103.28	15.0	2.70
Napo	Ecuador	70.78	25.76	37.46	101.53	19.0	2.00
Olmeca	Mexico	31.82	22.98	13.58	101.14	37.3	0.84
Oriente	Ecuador	39.10	27.57	11.55	105.50	24.1	1.51
Qua Iboe	Nigeria	99.73	22.16	68.26	117.02	36.3	0.14
Rabi-Kouanga	Gabon	33.79	23.38	13.65	95.46	37.7	0.15
Saharan Blend	Algeria	83.16	24.68	49.82	115.82	45.0	0.09
WTI	United States	42.30	27.68	14.34	99.56	39.6	0.24

Sources: [EIA, 2021b, of Canada , PSA].

Our theoretical framework implies that the oil class prices equal the weighted average of the field prices belonging to that class. Using this result, and assuming that all private information is publicly available ($\Omega_{t-1}^k =$

[4]The Energy Information Administration dataset provides only nominal prices. In order to make them comparable with the marginal extraction costs, we download WoodMac costs in nominal and in real terms. Using both values, we compute the Consumer Price Index (CPI) used by WoodMac to transform nominal costs into real ones. Rescaling the nominal future prices and $\bar{P}_t$ using the WoodMac CPI ensures that field level expected prices and marginal extraction costs are comparable quantities.

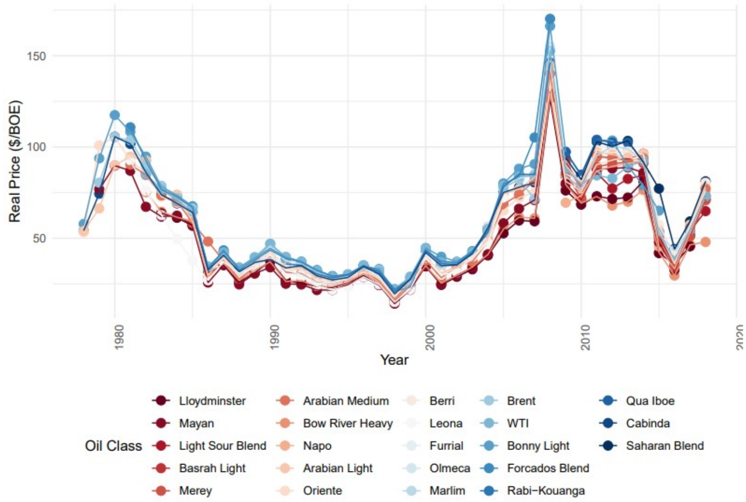

Figure 5.1: Future prices of twenty-three oil class over the time interval 1978-2018. Colors reflect the sulfur content, where dark red represents high content and dark blue low percentages.

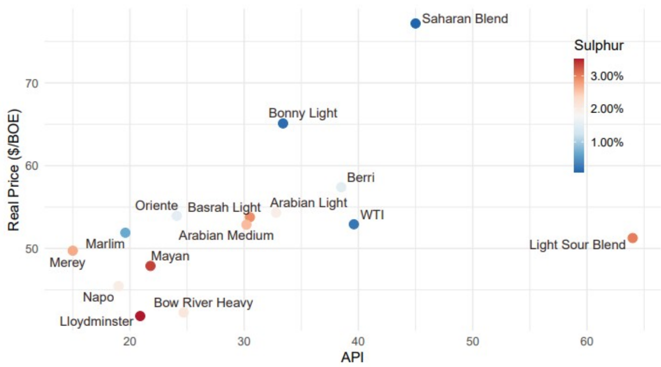

Figure 5.2: Spot prices of fourteen oil classes in 2015. The position on the horizontal axis reflects lightness. Colors reflect the sulfur content, where dark red represents high content and dark blue low percentages.

Ω_{t-1}^{pub}), we derive a formula for the future price of oil of class z in period t,

$$FP_t(z) = \mathbb{E}_{t-1}\left[\sum_{i\in z} w_t^i P_t^i | \Omega_{t-1}^{pub}\right] = \mathbb{E}_{t-1}\left[\sum_{i\in z} w_t^i P_t^i | \Omega_{t-1}^{k}\right] , \tag{5.7}$$

where $\{w_t^i\}_{i=1}^{N(z)}$ are time-varying weights identifying the relative importance of a field belonging to class z in period t. Then, we substitute the structural equation (5.6) for the expected price of oil from field i into the formula (5.7). As a result, we can glue the twenty-three future prices time series into a panel structure and run the regression

$$FP_t(z) = \bar{P}_t + \beta_{1t}(API(z) - \overline{API}_t) + \beta_{2t}(S(z) - \bar{S}_t) + \varsigma_t , \tag{5.8}$$

where we assume $\mathbb{E}_t[\varsigma_t|API(z), S(z)] = 0$, and use the estimated $(\hat{\beta}_{1t}, \hat{\beta}_{2t})$ to predict the unobserved response $\mathbb{E}_{t-1}[\hat{P}_t^i|\Omega_{t-1}^i]$.

Before running equation (5.8) and use its structural coefficient to reverse-engineer $\mathbb{E}_{t-1}[P_t^i|\Omega_{t-1}^k]$, we check if the future and reference prices are stationary. Not all oil classes have time series long enough to perform unit root tests. However, a subgroup of ten does[5]. Following Pesaran [2007] and Costantini and Lupi [2013],

[5]The subgroup is made of Cabinda, Bow River Heavy, Lloydminster, Oriente, Mayan, Bonny Light, Forcados Blend, Arabian Light,

we perform four types of Panel Covariate-Augmented Dickey-Fuller Test[6]: 1) with linear trend without Pesaran cross-sectional correlation, 2) with linear trend with Pesaran cross-sectional correlation, 3) with drift without Pesaran cross-sectional correlation, and 4) with drift with Pesaran cross-sectional correlation. We fail to reject the null hypothesis of presence of a unit in all cases obtaining p-values, which range from ∼0.88 to ∼0.41. In a similar way, we fail to reject the presence of a unit root for the time series of $\bar{P}_t$ (p-value = ∼0.30) using an Augmented Dickey-Fuller test [Said and Dickey, 1984, Banerjee et al., 1993]. The lack of stationarity on both sides of equation (5.8) allows us to move $\bar{P}_t$ to the lefthandside of the future price equation,

$$FP_t(z) - \bar{P}_t = \beta_{1t}(API(z) - \overline{API}_t) + \beta_{2t}(S(z) - \bar{S}_t) + \varsigma_t \; , \tag{5.9}$$

and run a regression in which the dependent variable, being the difference between two variables cointegrated of order one, is stationary[7] [Hamilton, 2020].

Therefore, we can use standard panel techniques to estimate β_{1t} and β_{2t}. We start running three Pooled Ordinary Least Square (POLS) regressions, which do not allow the coefficient to be time specific ($\beta_{1t} \equiv \beta_1$, $\beta_{2t} \equiv \beta_2$). Column (1) of Table 5.2, which uses $API(z) - \overline{API}_t$ as the only explanatory variable, suggests that a unit increase in $API(z) - \overline{API}_t$ increases the price of a BOE by 0.30 \$. Column (2), which uses $S(z) - \bar{S}_t$ as the only explanatory variable, indicates that a unit increase in $S(z) - \bar{S}_t$ lowers the price of a BOE by 2.98 \$. Comparing the results of these first two regressions suggests that sulfur explains a larger fraction of the variance of $FP_t(z) - \bar{P}_t$ than gravity. Column (3) includes both variables, as in equation (5.9). In this case, the magnitude of $\hat{\beta}_1$ shifts from 0.30 to 0.13, while the one of $\hat{\beta}_2$ from -2.98 to -2.52. Furthermore, the Adjusted R^2 is the largest among the POLS estimates suggesting that the combined presence of gravity and sulfur can explain circa one third of the variance of the delta between the future price of an oil class and a global reference price. The last three columns of Table 5.2 report the results obtained using a Random Coefficient Model (RCM) where β_{1t} and β_{2t} are normally distributed and vary across time[8] [Swamy, 1970, Bates, 2005, De Boeck et al., 2011]. The numbers reported in columns (4)-(5)-(6) are the average of the obtained $(\hat{\beta}_{1t}, \hat{\beta}_{2t})$. Column (4) returns an average impact of $API(z) - \overline{API}_t$ of $\sum_{t=1}^{T} \hat{\beta}_{1t}/T = 0.39$ \$/BOE, while column (5) returns an average impact of $S(z) - \bar{S}_t$ of $\sum_{t=1}^{T} \hat{\beta}_{2t}/T = -2.70$ \$/BOE. Contrary to the POLS estimates, in the RCM using only $API(z) - \overline{API}_t$ or only $S(z) - \bar{S}_t$ roughly explains the same portions of the variance of $FP_t(z) - \bar{P}_t$. Using both explanatory variables rescales the average impact of $\hat{\beta}_{1t}$ to 0.04 \$/BOE and the one of $\hat{\beta}_{2t}$ to -2.55 \$/BOE. We decompose the average estimates presented in column (6) $(\sum_{t=1}^{T} \hat{\beta}_{1t}/T, \sum_{t=1}^{T} \hat{\beta}_{2t}/T)$ in Table 5.3. The delta in API ranges from -0.04 \$/BOE in 2012 to a maximum of 0.13 \$/BOE in 2015, with an average value of 0.04 \$/BOE, and a median one 0.03 \$/BOE. Similarly, the delta in sulfur ranges from a minimum of -8.25 \$/BOE in 2008 to a maximum of -0.03 in 1986, with a average value of -2.55 \$/BOE, and a median one of -2.20 \$/BOE. This last model increases the (Nakagawa) adjusted R^2 to 0.40 [Nakagawa and Schielzeth, 2013]. In other words, a model which allows for time variations in the returns on deltas explains 40% of the variance of the delta between the future price of a particular oil class and the average oil price suggesting that 7% of the variance of $FP_t(z) - \bar{P}_t$ is due to a combination of changes in the composition of the demand for oil derived products and of the technological changes in the refinery sector.

Using the estimates portrayed in Table 5.3, we obtain the field-level expected prices. For example, in 2015, when $\bar{P}_t = 50.39$ \$/BOE, $\overline{API}_t = 31.46$, $\bar{S}_t = 1.4\%$, $\hat{\beta}_{1,2015} = 0.13$ \$/BOE and $\hat{\beta}_{2,2015} = -2.87$ \$/BOE, a hypothetical field i with an $API^i = 55$ and $S^i = 3\%$ would sell its output at

$$\widehat{\mathbb{E}_{t-1}[P_t^i|\Omega_{t-1}^i]} = 50.39 + \underset{(0.08)}{0.13} \cdot (55 - 31.46) - \underset{(1.01)}{2.87} \cdot (0.03 - 0.014) = \frac{53.40\$}{\text{BOE}} \; .$$

Arabian Medium, and WTI.

[6]The tests are performed using the R-package CADFtest running the command pCADFtest [Lupi, 2010].

[7]This theoretical result is confirmed by the four previously mentioned tests, which reject the hypothesis of non stationary of $FP_t(z) - \bar{P}_t$, with p-values ranging from 4.8e-06 to 1.4e-08.

[8]The RCM is estimated using the R-package lme4 [Bates et al., 2015] and the conditional modes are extracted with the command ranef.

Table 5.2: Future Price Regressions

	Dependent variable: $FP_t(z) - \bar{P}_t$ ($/BOE)					
	(1)	(2)	(3)	(4)	(5)	(6)
$API(z) - \overline{API}_t$	0.30***		0.13***	0.39***		0.04***
	(0.03)		(0.03)	(0.30)		(0.03)
$S(z) - \bar{S}_t$		-2.98***	-2.52***		-2.70***	-2.55***
		(0.20)	(0.23)		(1.80)	(1.75)
Observations	484	484	484	484	484	484
Adj. R^2	0.16	0.30	0.33	0.40	0.39	0.40
Note:						*p<0.1; **p<0.05; ***p<0.01

In that same year a different hypothetical field j containing oil with $API^j = 25$ and $S^j = 4\%$ would sell its output at

$$\mathbb{E}_{t-1}\widehat{[P_t^j|\Omega_{t-1}^j]} = 50.39 + \underset{(0.08)}{0.13} \cdot (25.00 - 31.46) - \underset{(1.01)}{2.87} \cdot (0.03 - 0.04) = \frac{49.58\$}{\text{BOE}} .$$

Notice that under the assumption stated in this section $\mathbb{E}_{t-1}\widehat{[P_t^i|\Omega_{t-1}^i]}$ is an unbiased estimator of $\mathbb{E}_{t-1}[P_t^i|\Omega_{t-1}^i]$ as long as $\hat{\beta}_{1t}, \hat{\beta}_{2t}$ are unbiased estimators of β_{1t}, β_{2t}.

Table 5.3: Year Specific β_{1t} and β_{2t} of Table 2 Column (6).

	$API(z) - \overline{API}_t$				$S(z) - \bar{S}_t$			
Year	Value	Std. Dev.	C.I.2.5%	C.I.97.5%	Value	Std. Dev.	C.I.2.5%	C.I.97.5%
1985	0.07	0.11	-0.14	0.28	-2.75	1.10	-4.90	-0.59
1986	0.05	0.11	-0.16	0.26	-0.03	1.11	-2.21	2.15
1987	0.02	0.11	-0.19	0.23	-1.76	1.12	-3.96	0.43
1988	0.02	0.11	-0.19	0.23	-2.06	1.13	-4.28	0.15
1989	0.03	0.11	-0.18	0.24	-1.94	1.14	-4.17	0.28
1990	0.05	0.11	-0.16	0.26	-1.88	1.14	-4.12	0.36
1991	0.06	0.11	-0.15	0.27	-3.09	1.13	-5.30	-0.88
1992	0.04	0.11	-0.17	0.25	-2.81	1.13	-5.03	-0.60
1993	0.04	0.11	-0.17	0.25	-2.22	1.13	-4.43	-0.01
1994	0.03	0.11	-0.18	0.24	-1.55	1.13	-3.76	0.66
1995	0.03	0.11	-0.18	0.24	-1.18	1.13	-3.39	1.02
1996	0.03	0.11	-0.18	0.24	-1.37	1.13	-3.58	0.84
1997	0.03	0.11	-0.18	0.24	-2.03	1.14	-4.26	0.20
1998	0.03	0.11	-0.18	0.24	-1.67	1.14	-3.90	0.56
1999	0.03	0.11	-0.18	0.24	-0.74	1.14	-2.97	1.48
2000	0.03	0.11	-0.18	0.24	-2.41	1.14	-4.63	-0.18
2001	0.03	0.11	-0.18	0.24	-3.52	1.14	-5.74	-1.29
2002	0.03	0.11	-0.18	0.24	-1.70	1.14	-3.93	0.52
2003	0.02	0.11	-0.20	0.23	-2.47	1.13	-4.69	-0.25
2004	0.06	0.11	-0.15	0.27	-3.65	1.14	-5.87	-1.42
2005	0.09	0.11	-0.12	0.30	-5.88	1.14	-8.11	-3.65
2006	0.09	0.11	-0.13	0.30	-5.45	1.14	-7.68	-3.22
2007	0.03	0.11	-0.18	0.25	-5.58	1.14	-7.80	-3.35
2008	0.06	0.11	-0.15	0.27	-8.26	1.21	-10.62	-5.89
2009	0.12	0.08	-0.03	0.28	-2.35	0.98	-4.27	-0.43
2010	0.10	0.08	-0.05	0.25	-2.17	0.98	-4.09	-0.26
2011	0.07	0.08	-0.08	0.22	-3.66	0.98	-5.57	-1.74
2012	-0.04	0.08	-0.19	0.11	-3.98	0.98	-5.91	-2.05
2013	0.01	0.08	-0.14	0.17	-3.31	0.99	-5.25	-1.38
2014	0.00	0.08	-0.16	0.15	-0.37	1.02	-2.37	1.64
2015	0.13	0.08	-0.02	0.29	-2.87	1.01	-4.85	-0.89
2016	0.03	0.08	-0.12	0.19	-0.85	1.05	-2.91	1.21
2017	0.06	0.08	-0.10	0.21	-1.01	1.05	-3.06	1.05
2018	0.02	0.08	-0.13	0.18	-0.22	1.17	-2.52	2.08

5.3.2 Marginal Costs

Several business intelligence companies collect data about oilfields revenues and costs. Among others, WoodMac classifies capital and operational expenditures of (parent) oilfields into twenty-three categories. Table 5.4 provides the summary statistics of the different classes over the twenty years time interval 1999-2018. We sum the first twenty-one of them to obtain the expenditures faced to “get the oil out from the ground”,

$$C_t^i = \text{Abandonment Costs}_t^i + \text{Capital Receipts}_t^i + ... + \text{Terminal}_t^i \ , \tag{5.10}$$

and the last two to “find new oil”,

$$W_t^i = \text{Development Drilling}_t^i + \text{Exploration and Appraisal}_t^i \ . \tag{5.11}$$

Table 5.4: Summary Statistics of the Twenty-Three Types of Cost.

Cost Type	Number of Observations	Mean	Std. Dev.	Min	Max
Extraction Costs					
Abandonment Costs	14,196	1.16	9.77	-8.87	378.72
Capital Receipts	401	6.22	59.26	0.00	1,044.79
Country Specific CAPEX	3,859	7.82	60.81	0.00	1,586.93
Country Specific OPEX	1,772	3.96	12.66	0.00	317.39
Field Fixed Costs	18,056	73.58	237.96	0.00	5,175.10
Field Variable Costs	17,701	45.78	117.94	0.00	2,928.38
General and Administrative	2,549	6.39	12.90	0.00	186.45
Insurance	40	0.07	0.41	0.00	2.62
Non Tariff Transport	2,055	20.05	70.27	0.00	931.03
Offshore Loading	854	3.93	20.70	0.00	264.49
Other CAPEX	10,461	19.14	77.25	-235.08	1,805.82
Other Costs	776	52.39	94.01	0.00	1,246.24
Other OPEX	726	9.08	31.83	0.00	212.17
Pipeline	14,491	6.30	32.50	-17.14	1,060.07
Processing Equipment	13,287	30.31	138.37	-32.60	3,191.38
Production Facilities	16,670	40.38	188.19	-40.14	6,260.00
Subsea	3,701	25.61	84.89	0.00	1,378.67
Tariff Gas	6,987	11.53	60.26	0.00	1,678.20
Tariff Oil	11,731	30.49	134.68	0.00	3,378.91
Tariff Receipts	1,604	11.47	25.17	0.00	239.46
Terminal	759	3.40	17.45	0.00	269.04
Exploration Costs					
Development Drilling	17,90	82.59	206.94	0.00	4,495.88
Exploration and Appraisal	123	3.46	12.33	0.00	82.87

Sources:[Mackenzie, 2018].

Figure 5.3 shows the relative importance of the different cost categories. Excluding fixed costs, the most relevant ones are the variable costs linked to the production process. For the purpose of informing the constructing the firm’s cost function, we briefly illustrate their origin.

When the mineral extraction rights are assigned, a team of geologists and engineers assesses the production potential of a field drilling a number of exploration wells, see left panel of Figure 5.4. The number of exploration wells multiplied by the per-well cost is a good proxy of W_t^i. Once the productive potential of a certain region is assessed, the area is divided into different sub-areas using a point pattern system. For illustrative purposes, we show a five-spot patterns method, see central panel of Figure 5.4. This technique divides the initial area into regular squares. Then, at the vertices of the squares are bored wells. The wells placed at the vertices of the square can be opened and transformed into producing wells, see right panel of Figure 5.4.

Once a well is opened, oil free-flows due to the natural pressure of the reservoir. Over time the natural pressure of the reservoir declines causing a decline in output. At this point, the management can artificially increase the

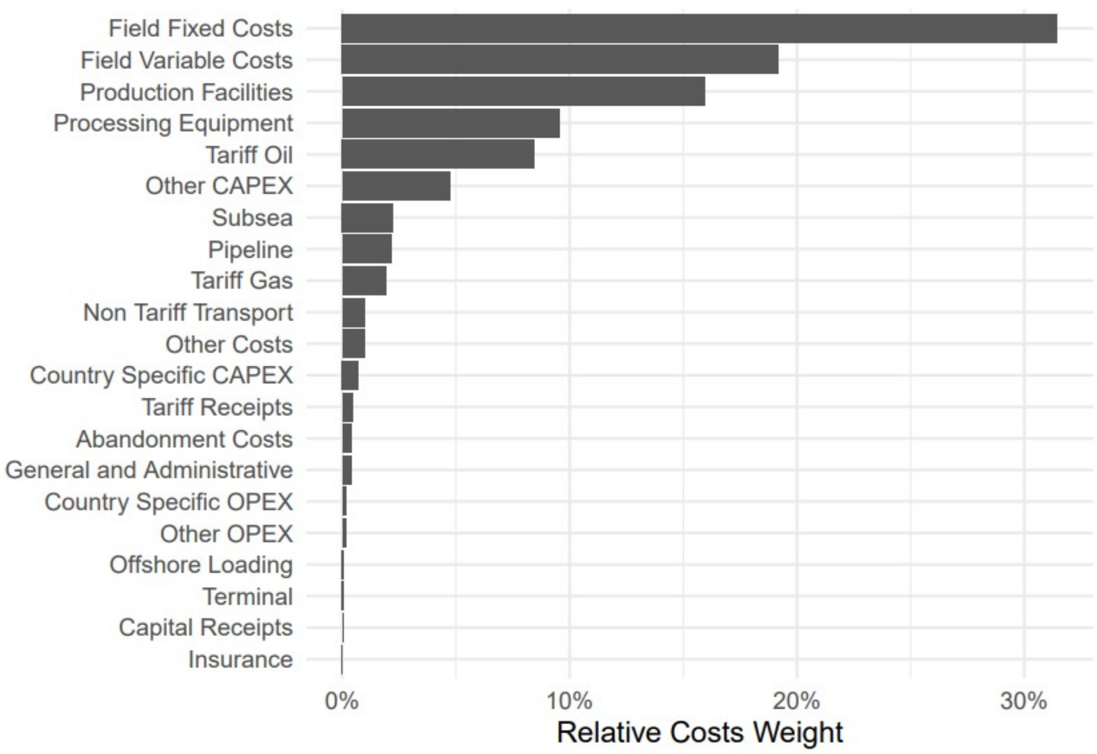

Figure 5.3: Relative weight of the different cost categories.

pressure in the reservoir by drilling a second type of well called injection wells. The later pump water, steam, chemicals and/or natural gas to keep the production process fluid, increase the reservoir temperature, and its pressure, see Figure 5.5.

The possibility to use injection wells depends upon the type of oil hosted in the deposit. If the reservoir contains low viscosity oil trapped in impermeable rocks (a.k.a. Shale & Tight Oil), untapped wells drill vertically. Then, once the deposit is reached, untapped wells drill horizontally through the oil-containing rocks. The horizontal section of the well is then fractured by opening fissures in the rocks. When the fracturing is completed, the well is tapped. The natural pressure of the reservoir lifts a mixture of oil, water, and stones above the ground (a.k.a. primary production phase). During this phase, it is impossible to increase production injecting water, steam, or natural gas. To the contrary, if the reservoir contains low viscosity oil trapped in permeable rocks (a.k.a. Light &

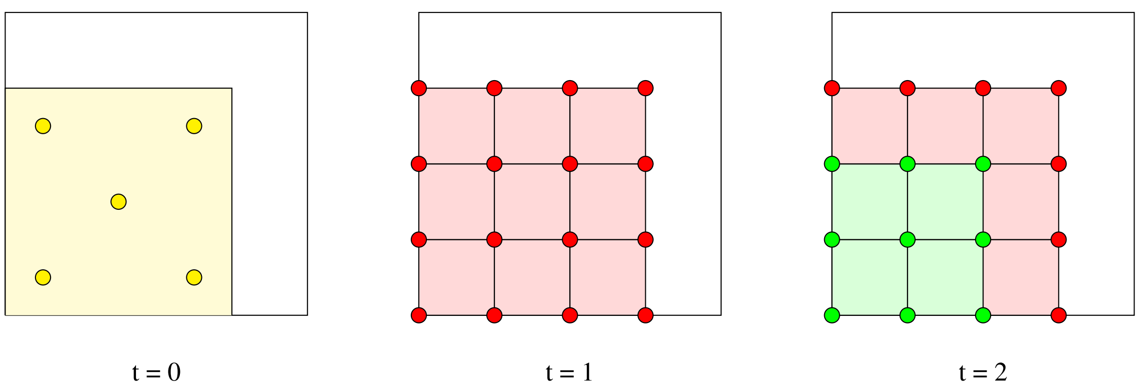

Figure 5.4: Exploration Wells ●, Explored Area, Untapped Wells ●, Potentially Producing Area, Tapped Wells ●, Producing Area.

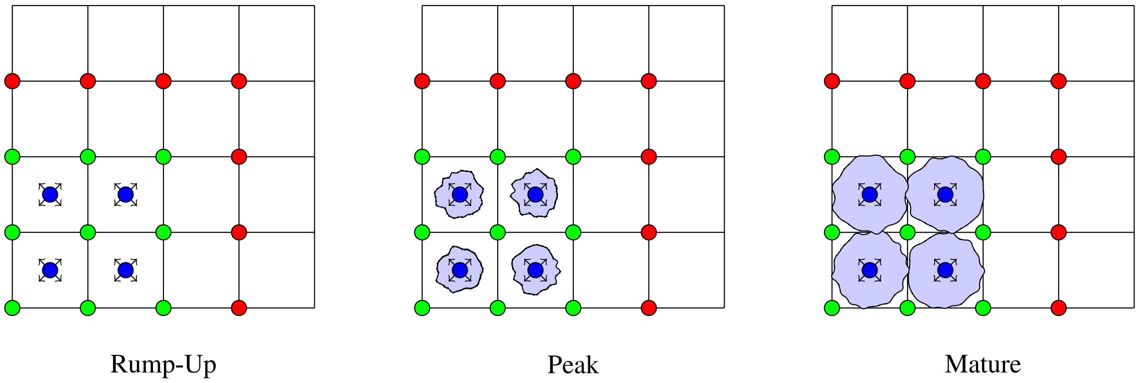

Figure 5.5: Untapped Wells •, Tapped Wells •, Injection Wells •, Injected Water/Steam

Medium Oil), untapped wells drill vertically. Then, once the deposit is reached and the wells tapped, the natural pressure of the reservoir lifts the oil above ground. As the field gets depleted, the pressure declines and with it the free-flow of liquids. However, in both the aforementioned cases, it is possible to re-bust the reservoir pressure injecting water in the deposit (a.k.a. secondary production phase). Finally, if the reservoir contains high viscosity oil trapped in permeable rocks (a.k.a. Heavy & Extra Heavy Oil), untapped wells drill vertically. Then, once the deposit is reached and the wells are tapped, the natural pressure of the reservoir lifts the oil above the ground. Contrary to the previous two cases, the pressure is short lived and, after few years, it sharply declines. However, unlike for Light & Medium oil, it is not possible to increase production by injecting water, since the oil would not flow due to its complex molecular structure[9]. Therefore, it is necessary to heat the water, transform it into steam, and inject an aerosol mixture in the reservoir (a.k.a. tertiary production phase or enhanced recovery method), see Table 5.5. In the case of Light & Medium and Heavy & Extra Heavy, it is possible to substitute and/or complement the injection of water with the one of natural gas. The latter can push the oil through the pores and the cracks of the matrix block guiding it toward the production well and increasing the reservoirs' recovery factor[10]. Finally, there are oil sands. They are loose or partially consolidated stones, which contain oil. The stones are generally saturated with Extra Heavy oil (bitumen). In this case, the natural pressure of the reservoir does not play a role in the production process since the oil is mined. The traditional method is to mine the sands and subsequently upgrade the resulting Extra Heavy oil in order to make the final product lighter [Shah et al., 2010]. More recently, it has become possible to heat the sands in-situ and avoid the upgrading.

Table 5.5: Production Phases by Typology of Oil

Production Phase	Mean of Production	Light & Medium	Heavy & Extra Heavy	Shale & Tight
Primary	Natural Pressure	✓	✓	✓
Secondary	Water Injection	✓	✗	✗
Tertiary	Steam Injection	TD	✓	✗

The previous discussion allows us to construct a categorical variable Geo^i, which connects pressure and input costs. Namely, if Geo^i classifies the different fields according to: 1) the porosity of the rocks (high vs low permeability), and 2) the oil consistence (high vs low viscosity), which orders the importance of natural pressure in determining the field variable costs, then moving from Sands → Heavy & Extra Heavy → Light & Medium → Shale & Tight, we observe an increasing role of natural pressure in determine the field variable costs. Moving in the opposite direction, we observe an increasing role of inputs, which substitute natural pressure, in the production

[9]Note that the opposite is not true. It would be possible to recover Light & Medium Oil injecting steam in the reservoir. However, this procedure is Technologically Dominated (TD) by the possibility to increase pressure by injecting water, which allows oil companies to obtain the same result at a lower cost.

[10]Natural gas injection could be used to increase the reservoir pressure. However, this practice is not common due to its high costs.

process like steam, water, electricity, and labour. In each of these formations the extraction costs emerge as the interaction between the geological characteristics of the oilfield and the endogenous production decision of the firm management. This interplay has been largely ignored by the existing literature, which usually makes the extraction costs function of the volumes of oil extracted *and* of the amount of recoverable reserves. This last quantity should captures the entire extraction process becoming the proxy for the logical sequence: more reserves $\rightarrow$ more pressure $\rightarrow$ less inputs. In this framework discovering one barrel should compensate, in terms of marginal costs, the extraction of one.

Extraction Costs

We assume that each firm i faces six types of costs, each corresponding to one or more cost classes listed in Table 5.4: extraction costs $ExtrCosts_t^i$ (classes 4, 6, 13, 16), transportation costs $TransCosts_t^i$ (classes 9, 10, 14, 21), fiscal costs $Taxes_t^i$ (classes 18, 19, 20), maintenance costs $MantCosts_t^i$ (classes 4, 5, 15, 17), disruption costs $DisrCosts_t^i$ (classes 1, 5) and other costs $OtherCosts_t^i$ (classes 1, 2, 3, 5, 7, 8, 11, 12). The firm's objective is to choose an input mix (labor, water, steam, electricity, etc), which minimizes its cost structure[11],

$$\begin{aligned} C_t^i = & \underbrace{ExtrCosts_t^i + TransCosts_t^i + Taxes_t^i}_{\text{Variable Costs}} + \\ & + \underbrace{DisrCosts_t^i + MaintCosts_t^i + OtherCosts_t^i}_{\text{Fixed Costs}} . \end{aligned} \tag{5.12}$$

Following [Anderson et al., 2018], we model each oilfield i as a continuum of wells. The field's expected size (denoted by $Size^i$) is equal to the amount of initial reserves R^i times a constant S capturing the potential for further discoveries in that field. Each well is characterized by a three-dimensional vector $\boldsymbol{\eta} \in T$ with $T := (0, +\infty) \times [0,1] \times \{0,1\}$, where the first element η_1 is a random variable capturing the initial natural pressure of a well of type $\boldsymbol{\eta}$ (measured at the time of tapping) and the second element η_2 is the depletion rate of the well. Lastly, η_3 is an indicator that equals 1 if the well is tapped and 0 otherwise. The variables η_1, η_2, η_3 are jointly distributed with conditional probability density function $f^i(\boldsymbol{\eta} \mid h_t^i)$, where h_t^i denotes the history of the field up to period $t-1$, in particular each wells pressure at the time of discovery, its depletion rate and whether it is tapped or not in each period $0, 1, ..., t-1$. Oil extraction is performed using a combination of n productive inputs in each well. The firm purchases inputs of type j at unit price p^j. Let $PInputs_t^{i,j}(\boldsymbol{\eta})$ denote the amount of inputs of type $j \in J$ used in a well of type $\boldsymbol{\eta}$. Firm i's extraction costs in period t are equal to the total cost of the productive inputs purchased by the firm during that period,

$$ExtrCosts_t^i = \sum_{j=1}^{n} \int_T p^j \; PInputs_t^j(\boldsymbol{\eta}) \, f^i\left(\boldsymbol{\eta} \mid h_t^i\right) d^3\boldsymbol{\eta} \, . \tag{5.13}$$

The quantity of each specific input used in each well is the outcome of an endogenous choice made by the firm's management. Specifically, we assume that a firm aiming to achieve a given production level Q_t^i chooses its input bundle seeking to minimize the total cost of producing such amount of output, and that the efficiency of a given input mix depends upon the firm's production technology. Input mix, technology, and geological characteristics together shape the output of the oilfield, which equals the aggregate capacity of its wells. As a result, the capacity of a well of type $\boldsymbol{\eta}$ is

$$WellCapacity_t^i\left(\boldsymbol{\eta}, h_t^i\right) = F_t^i\left(\left\{PInputs_t^{i,j}(\boldsymbol{\eta})^*\right\}_{j=1,\boldsymbol{\eta}\in T}^{n}, \boldsymbol{\eta} \,\middle|\, h_t^i\right) , \tag{5.14}$$

[11] Note that the inter-temporal profit maximization problem described in section 5.2 and the within-period cost minimization problem outlined in section 5.3.2 are consistent with each other, as in any standard multiple inputs-single output production theory framework. This equivalence allows us to incorporate the specific characteristics of the oil extraction technology within a standard microeconomic framework.

where $PInputs_t^{i,j}(\boldsymbol{\eta})^*$ is the (optimally chosen) amount of input j used by firm i in well η and the technology embedded in F_t^i is assumed to be smooth and exhibit constant returns to scale. Equation (5.14) is flexible enough to accommodate the characteristics of the oil extraction technology outlined in the previous section. In particular, the capacity of each well depends upon its natural pressure at discovery and the extent to which such natural pressure declines with depletion. Moreover, $WellCapacity_t^i$ varies together with the average pressure of other wells in the field and with the share of tapped wells, both captured by the field history h_t^i. For instance, the extraction of large quantities of crude from a given well may affect the natural pressure of all the other wells in the same field.

The reader may appreciate how the way we model the production of each well borrows from Anderson et al. [2018], in particular in assuming that the well's output depends solely upon its capacity, which is itself a function of its depletion. This implies that oil firms can respond to long-term anticipated changes in oil prices by increasing the overall capacity of the oilfield at the extensive margin (i.e., by drilling new extraction wells). However, our framework crucially differs from theirs because equation (5.14) embeds the possibility that the oil firm can respond to short- and medium-term market shocks by boosting the natural pressure of a well through the injection of liquids and/or gases. Injections are performed through existing or newly drilled injection wells and using specific inputs (steam, water, electricity, chemicals, etc.), which are purchased by the firm at market prices and contribute to boosting extraction costs, as illustrated in equation (5.13). This addition captures the key features of the oil extraction process we described in the previous section while retaining most of the empirically relevant features of the analysis of Anderson et al. [2018]. Moreover, it allows for non-trivial intensive-margin production choices[12] by the firm, which is a necessary feature to obtain a fully specified functional relationship between the firm's expected mark-up and the shadow-price of its oil reserves.

Next, we connect the capacity of individual wells with the overall capacity of the oilfield. Specifically, we assume that the normalized capacity of the oilfield in period t is given by the Dixit-Stiglitz aggregator of the capacity of its wells,

$$FieldCapacity_t^i = \left\{ \int_T \left[WellCapacity_t^i\left(\boldsymbol{\eta}, h_t^i\right)\right]^\delta f^i\left(\boldsymbol{\eta} \mid h_t\right) d^3\boldsymbol{\eta} \right\}^{\frac{1}{\delta}}, \tag{5.15}$$

for $\delta < 0$. The idea underpinning the use of this aggregator is that there is potentially some degree of complementarity or substitutability between productive inputs used across different wells[13]. Lastly, we assume that the total quantity of oil produced by every field is a strictly increasing function of its aggregate capacity, which equals the expected size of the field multiplied by the normalized field capacity, with formula:

$$Q_t^i = (Size^i \cdot FieldCapacity_t^i)^\xi ,$$

where the value of the parameter capturing the returns to scale of the firm's technology $\xi < 1$ ensures that the firm's optimal output choice problem presented in section 5.2 is well-behaved.[14] Using this setup, we show that at the optimal bundle of productive inputs the formula for extraction costs writes:

$$ExtrCosts_t^i = \left[\tilde{\theta}_2 + \theta_3^{Geo}\frac{L_{t-1}^i}{R^i} + \theta_3^{Geo}\frac{M_{t-1}^i}{R^i}\right] Q_t^{i\ 2}, \tag{5.16}$$

where $\tilde{\theta}_2 > 0$ and $\theta_3^{Geo}, \theta_4^{Geo}$ are geology-specific scalars.

[12] Anderson et al. [2018] implicitly rule out the possibility of manipulating the well pressure through injections by assuming that well-level marginal production costs are equal to zero. As a result, each well production level is always equal to its maximum predetermined capacity and firms only choose production levels at the extensive margin.

[13] Note that as $\delta \to -1$ the above equals the harmonic mean of the field's well capacities.

[14] Conversely, the cost minimization problem described in this section is well-behaved regardless of the value of ξ. Thus, the theoretical structure underpinning the derivation of the cost function does not rely on the assumption of diminishing return to scale.

We derive the maintenance costs using a similar procedure, where the cost of the optimal bundle of maintenance inputs is endogenously shaped by the field's characteristics. The other components of the firm's cost structure are modelled as a linear-quadratic function of oil output, augmented by firm- and time-specific factors and a stochastic component. The resulting cost function,

$$C_t^i = \theta_1 Q_t^i + \left(\theta_2 + \theta_3^{Geo}\frac{L_{t-1}^i}{R^i} + \theta_4^{Geo}\frac{M_{t-1}^i}{R^i}\right) Q_t^{i\,2} + \theta_5^{Geo}\left(\frac{M_{t-1}^i}{R^i}\right)^2 + \epsilon_t^i\,, \tag{5.17}$$

disentangles the impact of production choices and average reservoir pressure on marginal production costs for different types of geological formations. More precisely, in equation (5.17) the dependent variable C_t^i is the sum of Operating (OPEX) and of Capital Expenditures not linked to exploration (Non Exp CAPEX), as identified by equation (5.10), measured in Million US Dollars (MM \$) spend per Year. The independent variable Q_t^i equals the amount of output produced, measured in Million Barrels of Oil Equivalent (MM BOE) extracted per Year[15]. L_{t-1}^i/R^i and M_{t-1}^i/R^i measure the impact of discoveries and depletion, both rescaled by the initial volumes of reserves. These two quantities are pure numbers. The error term ϵ_t^i contains a field-specific effect, a time-specific effect, and an idiosyncratic cost shock $\epsilon_t^i = \theta_0^i + \theta_{0t} + \varepsilon_t^i$. The latter is normally distributed with finite homoskedastic variance, $\varepsilon_t^i \overset{iid}{\sim} \mathcal{N}(0, \sigma_\varepsilon^2)$.

Equation (5.17) differs from cost functions in which the extraction costs are function of the volumes of oil extracted and of the amount of recoverable reserves [Pesaran, 1990, Masnadi et al., 2021] because it allows for the magnitude of the effect of discovered reserves on the marginal production cost to differ from that of the depletion rate. Intuitively, this distinction captures the fact that depletion affects the field's capacity solely via its effect on well pressure, whereas the discovery of new reserves may also result in an increase in the field's installed capacity and, in turn, in the number of tapped wells.

Estimation

The estimation of equation (5.17) faces three main econometric issues. First, the empirical probability density function of C_t^i is virtually always positive[16] and over-dispersed ($\mathbb{V}[C_t^i] >> \mathbb{E}[C_t^i] > 0$). Second, C_t^i might be co-integrated. Third, Q_t^i, L_{t-1}^i, and M_{t-1}^i might be endogenous due to reverse causality in cost-production choices [Marschak and Andrews, 1944, Wooldridge, 2010]. Estimating (5.17) in first differences effectively tackles the first two problems, while attenuating the third one.

The empirical probability density function of $\Delta C_t^i = C_t^i - C_{t-1}^i$ is not always positive defined. Furthermore, a Wilcoxon signed-rank test rejects the hypothesis that ΔC_t^i and a simulated variable $\Delta C_t^{sim} \overset{iid}{\sim} \mathcal{N}(1.54, 19, 237.96)$ (i.e. a normal distribution with mean and variance equal to the ones of ΔC_t^i) are drawn from two statistically different distributions [Taheri and Hesamian, 2013]. Therefore, we do not need to estimate the model using a generalized linear models, which would return non-constant marginal extraction costs [Nelder and Wedderburn, 1972, Liang and Zeger, 1986]. Furthermore, if we run on C_t^i the four Panel Covariate-Augmented Dickey-Fuller Tests presented in section 5.3.1, we obtain discordant results. Namely, the two tests with drift reject the null hypothesis of presence of a unit root, while the two tests with linear trends (one with one without cross-sectional correlation) fail to reject the null hypothesis of presence of a unit root. This second results holds true even if we add as an explanatory variable the quantity of oil extracted (p-value = ∼0.06) and/or the development and depletion (p-value = ∼0.06). To the contrary, the same tests on ΔC_t^i regressed on drifts, trends, and/or the other explanatory variables evaluated in delta always reject the presence of a unit root. Finally, if we assume that the idiosyncratic cost shock is the sum of a field-specific unobserved fixed effect ϖ^i, possibly

[15] The decision to use BOE, rather than the traditional Barrel (BBL), allows us to sum the production of condensate, gas, natural gas liquids (NGL) and oil, so to compare the marginal costs of fields with a different composition of the output. For example, the BOE allows us to confront the marginal costs of Sands formations, which produce almost only oil, with the one of Shale & Tight formations, which produce considerable quantities of associated gases.

[16] 43 out of 28,924 observations present negative costs (0.14% of the sample). This are mostly North American fields, which, due to fiscal reasons, had more rebates than expenditures during the first or second year of production, such that $Taxes_t^i < 0$.

correlated with all the explanatory variables, and a random noise $\chi_t^i \overset{iid}{\sim} \mathcal{N}(0, \sigma_\chi^2)$, which is independent from all the explanatory variables; then the θs estimated in first differences do not suffer from reverse causality bias [McElroy, 1987]. Namely, since $\chi_t^i - \chi_{t-1}^i$ is uncorrelated to $Q_t^i - Q_{t-1}^i$, $L_{t-1}^i - L_{t-2}^i$, and $M_{t-1}^i - M_{t-2}^i$, the θs should be unbiased. While this is a standard solution in empirical industrial organization [Kawaguchi, 2021], reducing endogeneity by evaluating the model in first differences presents two problems. Firstly, it restricts our ability to capture cross-sectional heterogeneity since we renounce to estimate field-specific effects [Bai, 2009], those forgoing to exploit the panel nature of the dataset. Secondly, if the dependent and/or the explanatory variables contain significant measurement errors, then a first difference estimation might generate higher biases than a POLS or a fixed effect one, since measurement errors are more likely to survive a first difference than a within transformation. We tackle the first problem by fitting the model using a Random Coefficient Model, similar to the one used for fitting the pricing equation, which allows the coefficients to vary across geological groups while keeping them uncorrelated with the explanatory variables. While we cannot specifically address the second problem, we are confident it should have limited consequences for our analysis given that WoodMac is possibly the most reliable data provider of the petroleum industry.

The global production is dominated by Light & Medium deposits, which are responsible for 84.51% of all the extracted oil. The average Light & Medium field produces as much oil as the average Heavy & Extra Heavy one ($\sim$17 MM BOE per Year). To the contrary, the average Shale & Tight field produces 23.91% of the average Light & Medium oilfield, while the average sand mine produced 175.27% the one of the average Light & Medium oilfield. While on average sand mines produced more oil than any other type of formation, the largest oilfields are responsible for volumes of production unmatched by sand mines. For example, the largest Light & Medium fields extract 2,317,88 MM BOE in one year, the largest sand mine 131.01. The same difference is not reflected in the costs, where the maximum costs faced by Light & Medium fields were only 23% higher than the maximum costs registered for sand formations. Table 5.6 summarizes the costs and the volumes of production for the different formations.

Table 5.6: Summary Statistics of the Extraction Costs and the Production Volumes.

Variable	Number of Observations	Mean	Std. Dev.	Min	Max
Extraction Costs					
Light & Medium	23,066	225.21	596.01	-400.63	10,806.95
Heavy & Extra Heavy	2,649	322.80	801.70	0.00	11,013.02
Shale & Tight	2,882	66.76	123.58	-420.57	1,377.38
Sands	327	1,305.68	1,771.70	2.14	8,761.26
Production Volumes					
Light & Medium	23,066	17.23	72.12	0.01	2,317.88
Heavy & Extra Heavy	2,649	17.08	47.60	0.01	785.48
Shale & Tight	2,882	4.12	8.55	0.01	109.71
Sands	327	30.20	34.79	0.05	131.01

Sources:[Mackenzie, 2018].

Over the time interval 1999-2018, the average discovery rate D_t^i/R^i is 0.23% per year. This rate of development implies that, if the deposit does not extract any oil, in little over three hundred years, it can double its size. However slow this rate might appear, it still significantly bigger than the median one, which is zero, since approximately half of the assets analyzed did not made any additional discovery after the initial assessment of the field size. The largest discoveries are done in shallow, deep, and ultradeep waters and in Shale & Tight deposits where few outliers increase their original assessment up to 20% on a year-by-year base. The average depletion rate Q_t^i/R^i is 1.70%. Contrary to the discovery rate, the average depletion rate is close to the median one (1.70% vs 1.18%). Figure 5.6 shows the difference between the discovery and the cumulative depletion rate for different types of geology.

We run four regressions. Column (1) of Table 5.7 assumes that discoveries and depletion play no role in

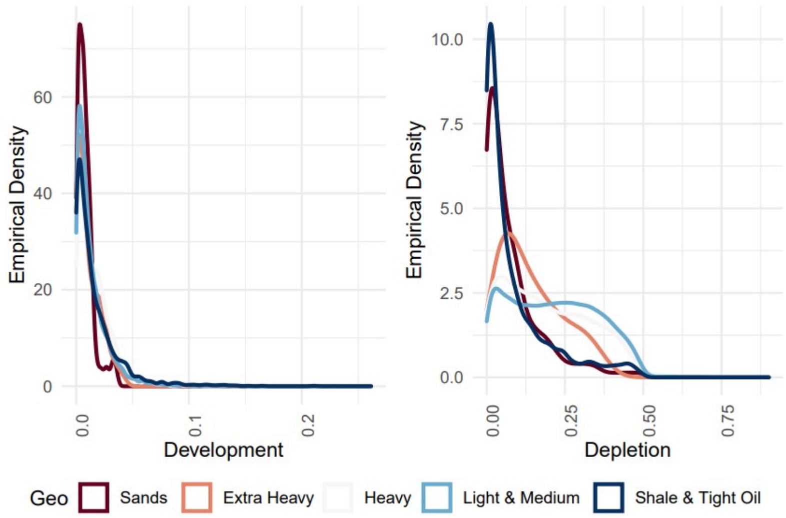

Figure 5.6: Empirical Density of the Discovery rate (D_t^i/R^i) and the Cumulative Depletion Rate (M_t^i/R^i).

Table 5.7: Marginal Costs Regressions

		Dependent variable: $C_t^i - C_{t-1}^i$ *(MM \$)*			
	Geo Class	(1)	(2)	(3)	(4)
Q_t^i		4.13***	5.28***	4.99***	4.98***
		(0.13)	(0.14)	(0.15)	(0.15)
$Q_t^{i\,2}$		−0.001***	0.06**	0.09***	0.08***
		(0.00)	(0.02)	(0.02)	(0.02)
$Q_t^{i\,2} \cdot L_{t-1}^i/R^i$			−0.06**		
			(0.02)		
$Q_t^{i\,2} \cdot M_{t-1}^i/R^i$			0.01***		
			(0.00)		
$Q_t^{i\,2} \cdot L_{t-1}^i/R^i$	Light & Medium			−0.095*	-0.007
				(0.017)	
$Q_t^{i\,2} \cdot L_{t-1}^i/R^i$	Heavy & Extra Heavy			−0.089**	-0.079
				(0.02)	
$Q_t^{i\,2} \cdot L_{t-1}^i/R^i$	Shale & Tight			−0.12***	-0.104
				(0.02)	
$Q_t^{i\,2} \cdot L_{t-1}^i/R^i$	Sands			−0.17*	-0.128
				(0.02)	
$Q_t^{i\,2} \cdot M_{t-1}^i/R^i$	Light & Medium			0.008*	0.008
				(0.003)	
$Q_t^{i\,2} \cdot M_{t-1}^i/R^i$	Heavy & Extra Heavy			0.002	0.002
				(0.003)	
$Q_t^{i\,2} \cdot M_{t-1}^i/R^i$	Shale & Tight			0.07	0.009
				(0.09)	
$Q_t^{i\,2} \cdot M_{t-1}^i/R^i$	Sands			0.38*	0.0106
				(0.18)	
Num. obs.		27,616	27,616	27,616	27,616
Adj. R^2		0.03	0.04	0.05	0.08

Note:
*p<0.1; **p<0.05; ***p<0.01

determining marginal extraction costs. Column (2) includes the impact discoveries and depletion without differentiating across geological formations. Column (3) includes geology as a categorical variable, which interacts with (L_{t-1}^i/R^i, M_{t-1}^i/R^i). Finally, column (4) reports the results of a RCM, which exactly matches the theoretically derived cost structure of equation (5.17), and where ($\theta_3^{Geo}, \theta_4^{Geo}, \theta_5^{Geo}$) are normally distributed and vary across geological classes, like in the pricing equation. The first regression estimates a cost of extracting the first barrel of 4.13 MM USD. The problem seems to be non-convex in quantities. Every barrel after the fist one would decrease marginal extraction costs by 1,000 USD. In the second regression extracting the first barrel costs

5.28 MM USD. Including discoveries and depletion makes the problem convex since every other barrel increases extraction costs by 6,000 USD. However, for an increase of 1% of the size of the deposit this increase in costs is perfectly offset. Conversely, decreasing the size of the deposit by 1% increases extraction costs by 1,000 USD per barrel. Combing the impact of $(\hat{\theta}_2, \hat{\theta}_3, \hat{\theta}_4)$, we conclude that discoveries offset the convexity of the cost function and the only element, which make cost grow is the field depletion. The last two columns introduce geology. In both cases, all discoveries decrease the marginal costs and all depletion increase them. While both the last one increase the capacity of explaining the variance of ΔC_t^i by 38% shifting the adjusted R^2 from 5 to 8%. Using the outcome of column (4), we obtain the expected marginal extraction costs,

$$\mathbb{E}_{t-1}\left[\widehat{\frac{\partial C_t^i(.)}{\partial Q_t^i}}\middle|\Omega_{t-1}^i\right] = \underset{(0.15)}{4.98} + 2\left(\underset{(0.02)}{0.08} + \underset{(97.33)}{\hat{\theta}_3^{Geo}}\frac{L_{t-1}^i}{R^i} + \underset{(64.04)}{\hat{\theta}_4^{Geo}}\frac{M_{t-1}^i}{R^i}\right)Q_t^i\,, \tag{5.18}$$

which can be subtracted to the estimated price equation to find the shadow price of oil in each field, which we use in the next section to derive the main results of the paper.

5.4 Economic & Environmental Effects of a Marginal Displacement

In order to analyze the economic and environmental footprint of the petroleum industry extensive margin, we merge the Oil Production Greenhouse Gas Emissions Estimator (OPGEE) global carbon intensity dataset with the estimated shadow prices[17]. The former contains the upstream emissions of 8,966 (children) oilfields. The latter the production and cost of 20,522 (children and standalone) oil & gas fields. Limiting our analysis to fields for which we have all the required information, we are able to match the shadow price of 2,017 fields covering circa 80% of the 2015 global oil supply[18]. Rank-ordering the obtained shadow prices from the lowest to the higher, we obtain the global merit order curve. Superimposing to every shadow price, its upstream emissions allows us to identify the environmental footprint of the global petroleum industry, as illustrated in Figure (5.7). We use this empirical tool to estimate the economic and environmental effects of an exogenous shocks on the global oil demand, focusing on the endogenous supply-side responses by extraction firms and refineries[19].

Upstream We analyse the effects of an oil demand reduction of 1%. According to our estimates, the least profitable 1% of the global production is made out of eight Heavy & Extra Heavy and five Sand formations. Both types of fields extract low-viscosity oils. The Heavy & Extra Heavy formations need to inject large quantities of steam as soon as the natural pressure declines. The oil sands need to add heat or inject fluids ‘in situ’ to reduce the bitumen’s viscosity. Both procedures increase the upstream emissions making them significantly bigger than the one of the standard Light & Medium formation, especially if the latter is well connected to the natural gas infrastructure and avoids large volumes of natural gas flaring and venting. The volume-weighted average carbon intensity of this fraction of the global oil supply is 114.61 $KgCO_2e$ (113.04 Heavy & Extra Heavy; 116.08 Sands) versus a sample average of 54.35 $KgCO_2e$. In other words, the 1% of least profitable fields emits more than double than the average global producer. This implies that a fall in the global oil demand by 1% translates in a reduction of upstream emissions equal to 24.95 $MMtCO_2e$ per year, approximately equal to more than 5.3% of all the privately-owned cars registered in the US[20].

[17] For a detailed description on the merging of the cross-sectional dimension of the production and cost information available in the WoodMac Upstream Data Tool with the emissions estimated by OPGEE, see the Appendix of [Masnadi et al., 2021].

[18] We choose 2015 as the reference year, since OPGEE emissions have been calculated for 2015.

[19] Thanks to the general equilibrium framework, we can analytically derive the formulas for macroeconomic shocks, which affect the equilibrium prices of crude oil and fossil fuels. For instance, we can analyse the effect of an exogenous shock on global GDP, which is equal to the effect of a shock of global oil demand of the same proportion times the equilibrium income elasticity of global oil demand. In this framework, the equilibrium elasticity differs from the standard notion of income elasticity of demand because it includes the effect of the endogenous adjustment of equilibrium prices due to the shock. While not shown in the paper, the results of a global income reduction (or of a change in taste for fossil fuels) have consequences virtually identical to the one of a decline of oil demand.

[20] The quantification is based on an average emission of 4.6 metric tons of carbon dioxide per year per vehicle as quantified by the United Stated Environmental Protection Agency [EPA, 2018a]. The number of privately-owned automobiles for private and commercial use in the

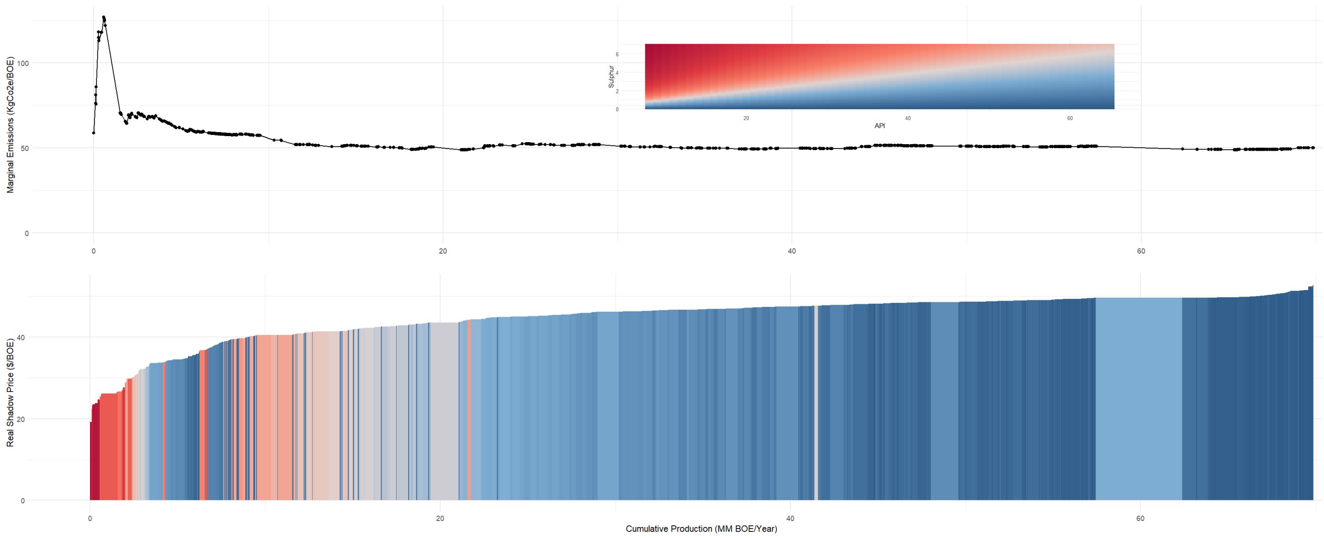

Figure 5.7: 2015 Global Merit Base Curve coupled with Upstream Carbon Intensity. Colors reflect the API gravity and the sulfur content, where dark red represents high sulfur content and low gravity and dark blue represents low sulfur content and high gravity.

Similarly, the least profitable 2.5% of the global production is made of fourteen Heavy & Extra Heavy formations, six Sand formations, and four Light & Medium deposits. The volume-weighted average carbon intensity of this fraction of the global oil supply is 78.75 $KgCO_2e$ (70.90 Heavy & Extra Heavy; 115.84 Sands, 36.47 Light & Medium). In other words, on average the 2.5% of least profitable fields is 31% less carbon intensive than the 1% and generates emissions equal to 51.68 $MMtCO_2e$ per year. Finally, the least profitable 5% is made of thirty-three Heavy & Extra Heavy, seventeen Sands, twenty-three Light & Medium, and two Shale & Tight formation. Their volume-weighted average carbon intensity is 70.92 $KgCO_2e$ (69.11 Heavy & Extra Heavy; 97.67 Sands, 54.11 Light & Medium, 50.50 Shale & Tight) and is responsible for 93.02 $MMtCO_2e$ of emissions per year. After passing the 5% least profitable production, the carbon intensity starts converging to the global average, see Figure (5.7). Table 5.8 summarizes all the results.

Table 5.8: Estimated Upstream Impact of a Marginal Decline in Oil Demand

Scenario	Carbon Intensity $KgCO_2e$ / BOE	Demand Decline MM BOE / Day	Carbon Savings $MMtCO_2e$ / Year
1%	114.61	0.72	24.95
2.5%	78.75	1.80	51.68
5%	70.92	3.60	93.02

From an economic prospective, the least profitable 1% of the global production manages a quantity of reserves equal to 15.55 billions BOE equal to 0.75% of the global pool. Similarly, the 2.5% of the least profitable fields manages 1.72% of the global reserves, while the 5% the 5.27%. In other words, the volume of production displaced by a demand shock are similar to the volumes of reserves stranded. To the contrary, the value of the displaced oil is smaller than the volume-weighted global average of 50.66 dollars per BOE. The 1% extensive margin sells its output at a volume-weighted price of 36.30 dollars per BOE. In a similar way, the 2.5% sells its oil at 40.31 dollar per barrel and the 5% at 41.50 dollars per BOE. Stranding the 1% less profitable oilfields would keep 1500 $MMtCO_2e$ underground. Cutting off the 2.5% and the 5% less profitable formations would increase the carbon savings to 3270 and 8440 $MMtCO_2e$, respectively.

US in 2021 was equal to 101,601,344 [DOT, 2022].

Midstream Petroleum refineries use as input blend of multiple streams of crude feedstocks. The first step of the refining process requires to separate the natural gas from the liquids. Then, the different gas-free streams are allocated to different sub-units depending upon the boiling point of their molecules. In each of the processing units a set of chemical and thermal processes fragments and rearranges the carbon and the hydrogen bonds of the input in order to increase the hydrogen-carbon ratio of the output, while eliminating the sulfur and the nitrogen. The heavier and sourer the crude stream, the more energy intensive the process becomes.

Since the least profitable oilfields extract heavier and sourer oil than the global average, their displacement impacts the midstream emissions. In order to quantify this effect, we run a linear carbon intensity equation,

$$\frac{\widehat{\mathrm{KgCO_2e}}^{\,r}}{\mathrm{BOE}} = \underset{(2.37)}{62.03} - \underset{(0.06)}{0.62}\, API^r - \underset{(0.64)}{0.95}\, S^r \;, \tag{5.19}$$

on the gravity and sulfur content of the processed crude using data from 343 refineries $r = 1, 2, ..., 343$, located in 83 countries[21] as elaborated by Cooney et al. [2017] and Jing et al. [2020]. The dependent variable is the refinery carbon intensity computed by the Petroleum Refinery Life Cycle Inventory Model 1.4 (PRELIM) [Abella and Bergerson, 2012]. The API^r is the average gravity of the processed crude and S^r is the average sulfur content. The carbon intensity is measured in kilograms of carbon dioxide equivalent emitted per barrel of oil equivalent refined $\mathrm{KgCO_2e/BOE}$, while (API^r, S^r) are dimensionless quantities expressed as pure numbers. As a result, the two coefficients are measured in $\mathrm{KgCO_2e/BOE}$. Their magnitude is obtained using Ordinary Least Squared (OLS) estimates. According to our results, the unconditional emissions equal 62.03 $\mathrm{KgCO_2e/BOE}$. Every increase in gravity makes the emissions decline by 0.62 $\mathrm{KgCO_2e/BOE}$. The impact of a gravity change is highly statistically significant. To the contrary, sulfur content does not play a statistically significant role[22].

We do not know which oilfields sell to which refineries. Therefore, we cannot examine how a decline in global oil demand would impact the trading routes between up- and midstream and calculate how this change would affect midstream emissions. However, we know from the upstream analysis that the global pool of crude would be lighter and sweeter. Therefore, using an approach similar to Masnadi et al. [2021], we use the estimates of equation (5.19) to construct a partial equilibrium counterfactual analysis where the global oil demand declines by 1%, 2.5%, and 5%. Then, we measure the volume-weighted change in gravity and sulfur content of the global pool, and assume that this new stream of crude is processed by a representative refinery, whose emissions decline by 0.62 $\mathrm{KgCO_2e/BOE}$ every time the global pool becomes lighter by one degree, and by 0.95 $\mathrm{KgCO_2e/BOE}$ every time the global pool becomes sourer by 1%.

The pre-shock global pool has a volume-weighted API gravity of 31.93 and a sulfur content of 1.27%. The 1% reduction scenario of 32.11 and of 1.24%. Therefore, the midstream carbon intensity shifts from 40.89 to 40.81 $\mathrm{KgCO_2e/BOE}$. In a similar way, in the 2.5% gravity is 32.39 and the sulfur content 1.20% reducing the carbon intensity to 40.67 $\mathrm{KgCO_2e/BOE}$. Finally, the 5% reduction scenario the gravity is 32.69 and the sulfur content of 1.14% further reducing the carbon intensity to 40.54%. Table 5.9 shows how this changes in chemical composition and consequent change in carbon intensity affect the aggregate emissions of the refineries. According to our results, the savings range from a minimum of 0.006 to a maximum of 0.025 $\mathrm{MMtCO_2e}$ / Year. Although non-negligible, these savings are significantly smaller than the upstream ones. For instance, in the 1% demand decline scenario, the former represent only 0.02% of the latter. Similar ratios emerge in the 2.5% (0.03%) and 5% (0.03%) scenario.

Downstream The effects of an oil demand reduction on downstream emissions are hard to calculate because, contrary to up- and midstream emissions, they are sensitive to the type of demand shock experienced. For example, a symmetric global shock in income, like the 2008-2009 financial crisis, would reduce the demand of different

[21] For homogeneity reason, the cross-section is taken in 2015.

[22] Despite being statistically insignificant, we use it to predict $\mathrm{KgCO_2e/BOE}$ since skipping S^r increases the mean square prediction error of the regression.

Table 5.9: Estimated Midstream Impact of a Marginal Decline in Oil Demand

Scenario	Carbon Intensity $KgCO_2e$ / BOE	Demand Decline MM BOE / Day	Carbon Savings $MMtCO_2e$ / Year
1%	40.81	0.72	0.006
2.5%	40.67	1.80	0.016
5%	40.54	3.60	0.025

products in the same proportion, as long as consumer preferences are (intra-temporally) homothetic, as in our theoretical model. To the contrary, an asymmetric shock in the world income, like the COVID-19 pandemic, would cause a change in the composition of global oil derived product demand. For instance, the demand of jet fuel will decline more than the demand for gasoline or ultra-low sulfur diesel because the pandemic impacted the flying industry more than the transportation one. These difficulties are well illustrated in Masnadi et al. [2021]. In keeping with their analysis, we focus on demand shocks that do not depend upon the consumer's relative preferences for different oil products. We depart from their methodology in the fact that we do not construct bounds on the average carbon intensity of downstream emissions based on estimates in the literature. Instead, we exploit the consumer side of the theoretical model to calculate the change in the demand for each of seven types of oil products identified in the Statistical Review of World Energy published by BP [BP Energy, 2022]. We calculate the change in demand for each product using the formula for the consumer demand from our theoretical model and assuming an average refinery hydrocarbon loss equal to 0.75%, a prudent value based on firms' best practice [Trident Consulting, 2023]. Then, we use the stationary combustion emissions values from [EPA, 2022b] as prudent measures of the marginal emissions due to the consumption of each product. Lastly, we multiply the change in the demand for each product for its carbon intensity and sum up over all the products to quantify the decline in downstream emissions. While the results of this exercise – which are summarized in Table (5.10) – do not differ significantly from those in the recent literature [Masnadi et al., 2021], our approach constitutes an improvement with respect to the methodology adopted by recent studies.

Table 5.10: Estimated Downstream Impact of a Marginal Decline in Oil Demand

	Carbon Intensity $KgCO_2e$ / BOE	Demand Decline MM BOE / Day			Carbon Savings $MMtCO_2e$ / Year		
Scenario		*1%*	*2.5%*	*5%*	*1%*	*2.5%*	*5%*
Jet/Kerosene	410.93	0.054	0.135	0.271	8.13	20.32	40.63
Fuel Oil	429.00	0.059	0.148	0.297	9.30	23.25	46.49
Naphtha	358.40	0.047	0.118	0.236	6.18	15.46	30.92
Gasoline	370.16	0.178	0.445	0.891	24.07	60.19	120.37
Diesel/gasoil	430.25	0.209	0.522	1.045	32.83	82.07	164.13
LPG/ethane	239.60	0.092	0.230	0.459	8.04	20.09	40.19
Others	373.15	0.075	0.186	0.373	10.16	25.41	50.82
Total		**0.715**	**1.786**	**3.573**	**98.71**	**246.78**	**493.57**

Well-to-Wheel We merge the results from up-, mid-, and downstream processes to obtain a comprehensive well-to-wheel quantification of the carbon emission reduction effects of a modest decline in the global oil demand. Our findings – summarized in Table (5.11) – suggest that both down- and upstream processes play a substantial role in shaping the magnitude of the effect of a demand shock on the greenhouse gas emissions of the oil sector, with the latter accounting for a share of the overall emission reductions ranging from 15% to 20% across the three scenarios analyzed in this study. Conversely, midstream emissions play a secondary role, never exceeding the 0.004% of the total effect of the demand shock. In aggregate, emission savings are substantial. A 1% fall in the global oil demand translates into a reduction in greenhouse emissions of 123.67 $MMtCO_2e$ per year, approximately equal to the total annual GHG emissions of the US State of Colorado in 2016 [EPA, 2023].

Table 5.11: Estimated Well-to-Wheel Impact of a Marginal Decline in Oil Demand

Scenario	Upstream Savings $MMtCO_2e$ / Year	Midstream Savings $MMtCO_2e$ / Year	Downstream Savings $MMtCO_2e$ / Year	Well-to-Wheel Savings $MMtCO_2e$ / Year
1%	24.95	0.006	98.71	**123.67**
2.5%	51.68	0.016	246.78	**298.48**
5%	93.02	0.025	493.57	**568.61**

5.5 Conclusions

The present paper provides a fully micro-founded empirical tool, which quantifies the impact of aggregate macroeconomic shocks on the production decisions of oil firms. Combining its output with life cycle analysis tools, we estimate the environmental consequences of a marginal oil displacement across the global supply-chain.

We start identifying an extraction-exploration equilibrium, where output, costs, and shadow prices are endogenously determined by the firm profit-maximizing behavior. Then, we estimate the magnitude of the shadow prices to quantify the response of each oilfield to an exogenous change in aggregate demand. According to our results, the marginal profitability is highly heterogeneous. The most profitable fields can be twice as profitable than the least competitive ones. Coupling the profitability with the upstream carbon intensity, we compute the impact of an exogenous demand shock on the emissions released to 'get the oil out from the ground'. According to our results, the impact of marginal displacement is substantial because the least profitable oilfields, which are the most likely to shut down in response to a fall in oil prices, are also those exhibiting the highest carbon intensity. This finding is robust to the introduction of strategic behavior among firms. In particular, Venezuelan Heavy and Extra Heavy fields are a large fraction of the industry extensive margin when firms are price-takers and when they play a game in quantities. We complement these finding with novel estimates the effect of an exogenous demand shock on both mid- and downstream emissions, and aggregate the results to quantify the overall well-to-wheel GHG emission reduction.

Our findings have two key policy implications. First, they imply that the responses to targeted taxes and subsidies on oil production and consumption are likely to be highly heterogeneous across fields with different geological characteristics. In particular, a uniform excise tax on oil production is likely to hit severely the production and investment choices of firms producing Heavy and non-conventional oil, while it would have little effect on other fields. Second, they suggest that an optimal Pigouvian tax and/or tradable permit scheme might not deliver substantially more efficient outcomes relative to some normatively inferior but easier-to-implement policy alternatives, such as excise taxes on oil production or a sales tax on fossil fuels consumption.

Far from being fully exhaustive, the present paper opens a path in the direction of a more all-inclusive approach toward an increasingly diversified oil industry and offers a new perspective on how to tackle its contribution to global warming.

Chapter 6

A zero-cost policy to eliminate methane emissions from the oil and gas industry

This chapter is entirely based on Benini, Giacomo; Dotti, Valerio; Berentsen, Geir Drage; Otneim, Håkon; Jahnke, Eric; Schuhmacher, Johannes; El-Houjeiri, Hassan M.; Ardone, Armin; Fichtner, Wolf; Jochem, Patrick; Gordon, Deborah; Brandt, Adam R.; Masnadi, Mohammad S. (2025): A Zero-Cost Policy for Eliminating Methane Emissions in the Oil and Gas Industry. To be submitted to a scientific journal.

The elimination of methane emissions is one of the most significant goals that policymakers can prioritize to meet the Paris Agreement climate targets [Lee and Romero, 2023]. The combination of a short lifespan, high global warming potential, and positive commercial value makes methane emission abatement a highly cost-effective way to limit the concentration of greenhouse gases (GHG) in the atmosphere. The oil and gas industry is a major methane emitter [Masnadi et al., 2018], releasing an estimated 79.5 megatons of methane into the atmosphere in 2023 alone [IEA, 2024]. Most of these emissions derive from an associated commodity channel. Hydrocarbon reservoirs contain a mixture of oil and natural gas. In many regions, the cost of capturing, purifying, and transporting natural gas exceeds its market price and/or the opportunity-cost of reusing it on site. Therefore, natural gas is either burnt off via flare stacks, releasing both carbon dioxide and methane or vented into the atmosphere, releasing methane [Plant et al., 2022]. Together, these two practices account for roughly 3% of global natural gas production and 7% of global energy-related GHG emissions [IEA, 2023a, Stanford Doerr School of Sustainability, 2023, EDGAR, 2023, Statista, 2024].

Given that the recovery of natural gas that is currently being wasted requires lower investment compared to most other GHG emission reduction strategies [IEA, 2021a], many policymakers consider the reduction of flaring and venting the low hanging fruit of climate change mitigation [IEA, 2023b]. In an effort to address both practices, the World Bank launched the Global Gas Flaring Reduction Partnership (GGFRP) in 2015. To date, 36 governments, 58 oil and gas companies, and 15 development banks have signed the Zero Routine Flaring by 2030 Initiative [Bank, 2023]. The main goal of this intergovernmental partnership is to eliminate flaring practices through voluntary governmental regulation while explicitly stating that venting is "not an acceptable substitute for flaring" [Bank, 2023].

Calel and Mahdavi raise concerns about the effectiveness of the GGFRP approach [Calel and Mahdavi, 2020]. Through satellite data analysis, they observe an increase in methane emissions in areas where flaring decreased, suggesting that oil and gas companies may have ramped up venting practices in an effort to comply with GGFRP flaring standards. Without rigorous monitoring, the overall impact of GGFRP-like agreements could be negligible or even counterproductive for climate change mitigation, considering that methane has a significantly higher global warming potential compared to carbon dioxide. The potential backfire arises from two unintended effects of introducing a new regulatory regime. First, when the relative cost of flaring exceeds that of venting, producers are incentivized to vent a portion of natural gas that, in the absence of anti-flaring policies, would have been flared as a routine practice. Second, the increased costs associated with voluntary gas disposal due to both anti-flaring and anti-venting policies reduce companies' return on investment in equipment maintenance and leak detection. Consequently, these negative outcomes undermine the effectiveness of policies that aim to restrict both flaring and venting [EPA, 2024]. Similarly, traditional approaches such as Pigouvian taxation, emissions trading, and incentive schemes targeting non-point externalities (e.g., agricultural runoffs) [Griffin and Bromley, 1982] are ineffective and may even exacerbate the problem. Newer solutions that allow companies to self-select into different fiscal regimes often result in underreporting of emissions data, leading to minimal emissions reductions [Cicala et al., 2023]. Moreover, both traditional and innovative emissions pricing policies can have adverse effects that compromise their feasibility and desirability when implemented in a single country, including reduced consumer purchasing power [Sager, 2023], weakened competitiveness of domestic companies in global markets [Dorsey-Palmateer and Niu, 2020], and carbon leakage [Böhringer et al., 2022].

To address these unresolved challenges, we propose a fiscal reform that incorporates the complexities of natural gas management. Building on the work of Masnadi et al. [Masnadi et al., 2021], we develop a general equilibrium model that includes four key types of agents: 1. oil and gas firms, which are central to the model and make profit-maximizing decisions regarding oil extraction (see supplementary materials [SM], section 6.1.1); 2. refineries, natural gas processing facilities, and power plants (SM, section 6.1.2); 3. consumers (SM, section 6.1.3); and 4. governments (SM, section 6.1.4).

Within our modeling framework, the management of natural gas co-extracted with oil plays a crucial role in shaping the economic and environmental outcomes of oil production. When a certain quantity of oil is extracted, a corresponding amount of natural gas — determined by the reservoir's gas-oil ratio — is co-extracted. Firms have five main options for handling this gas: sale, reinjection, onsite use, flaring, or release into the atmosphere. The flaring of natural gas is categorized into routine and non-routine flaring. Routine flaring includes gas combusted for purposes other than essential safety and maintenance operations. Released natural gas is further classified into venting (operators opening of O&G systems) and leaking (gas escaping from supposedly closed O&G systems). Venting refers to deliberate actions, such as opening oil tanks, ignoring detected leaks, or failing to repair flare stacks to meet flaring regulations. Leaking, on the other hand, refers to both the unintentional and intentional release of gas due to ageing equipment, inadequate maintenance, or safety-related pressure relief measures. Because the connection between leaking and safety measures creates a trade-off between minimizing leaking and ensuring worker safety [Collins et al., 2022], an effective regulatory approach should focus on eliminating routine flaring and venting while discouraging, but not completely banning, leaking. However, distinguishing between venting and leaking can be challenging from a legal standpoint. To address this problem, we propose a system of endogenous incentives that motivates oil and gas firms to voluntarily adopt best practices, thereby reducing the need for strict legal distinctions.

We propose a revenue-neutral, stakeholder-friendly tax reform, which eliminates routine flaring and venting (RFV) at a net-zero societal cost. The tax reform is intended to mitigate the undesired flaring-venting substitution effect and raise the maintenance level for natural gas equipment, without impacting corporate profits, consumer income, or government revenue (see SM, section 6.2.1). The reform does not alter the equilibrium prices of oil, natural gas, and derivative products nor does it generate carbon leakage. Furthermore, it is compatible with the existing tax schedule on fossil fuels in many countries, including the United States (see SM, section 6.5.3).

At the core of the proposed reform is a strategic adjustment of two key taxes. First, the specific tax on oil production should increase proportionally to the reservoir's gas-oil ratio. This ensures that firms extracting oil with a higher gas content pay a tax that reflects the environmental risks associated with the potential mismanagement of the co-extracted gas. Second, the specific tax on natural gas sold by oil producers should decrease by an amount equivalent to the maximum net cost of selling unprofitable natural gas (see SM, section 6.1.1 for guidance on distinguishing between oil and gas fields for tax purposes). These two tax adjustments are designed to work in opposite directions: while increasing the tax on oil production addresses the environmental cost, the reduction in tax on natural gas sales compensates firms for reducing potential social costs by selling the gas instead of wasting it. Under this scheme, a firm that completely eliminates voluntary natural gas waste would receive full compensation. As a result, the reform does not negatively impact a firm's profitability or the optimal level of oil production. Additionally, the reform includes a slight adjustment to the specific tax on natural gas sales by gas-only fields, ensuring that the policy has no effect on all equilibrium prices.

As a case study, we examine what would have happened if the United States — the largest oil and gas producer in the world — had implemented our proposed reform from 2005 to 2020. This period was specifically chosen because it provides the only timeframe with a consistent treatment of all the micro-data necessary for our research. Utilizing a comprehensive dataset that includes production decisions from 556 onshore oilfields, we estimate the policy's outcome. The first step in evaluating the impact of the proposed tax reform is to disentangle venting from leaking. To separate these two quantities, we use the firm's equilibrium conditions to identify a venting lower bound. This identification strategy minimizes the risk of overestimating the impact of our policy proposal (see SM, section 6.3). Then, we quantify the environmental savings using a two-step estimation strategy. Using a panel Tobit I model, the first step calculates the expected change in volumes of routine flaring resulting from changes in the price of natural gas (see SM, section 6.4.1). The second step estimates a venting lower bound using a linear panel regression model, where the quantity of unsold and unflared natural gas is a function of the flaring status of oilfields, the price of natural gas, the flaring-venting substitution effect (obtained as the residual of the first step), the expected level of maintenance, the level of oil production, and the volumes of natural gas (re-)injected to maintain reservoir pressure or used onsite to generate heat or electricity (see SM, section 6.4.2).

Our calculations for the case study indicate that, on average, 2.78% of the total energy from oil and natural gas extraction was wasted, amounting to 0.197 million barrels of oil equivalent per day (MM BOE/Day). Routine flaring accounted for 3.30% (0.086 MM BOE/Day) of all wasted natural gas extracted from oilfields, while venting (0.033 MM BOE/Day), non-routine flaring (0.001 MM BOE/Day), and leaking (0.076 MM BOE/Day) contributed 1.26%, 0.04%, and 2.87%, respectively. These quantities are sensitive to changes in natural gas prices: routine flaring decreases by 1.74 BOE/Day for every $1/BOE increase in nominal gas prices, venting by 2.37 BOE/Day, and leaking by 0.68 BOE/Day. Figure 6.1 illustrates the geolocation of and estimated composition of natural gas waste for the oilfields analysed in this study. The mismanagement of the co-extracted natural gas resulted in an average emission of 0.465 million metric ton of carbon dioxide equivalent per day (MM TCO_2e/Day). According to our estimates, 35.31% of these emissions (0.164 MM TCO_2e/Day) are discretionary and thus avoidable. They are spread between routine flaring (0.032 MM TCO_2e/Day) and venting (0.132 MM TCO_2e/Day), with the latter accounting for 30.05% of all methane emissions.

Policy Outcome We study the effect of the proposed tax scheme on four different venting reduction targets: 25%, 50%, 75%, and 100%. According to our estimates, with an average West Texas Intermediate price of 67.30 dollars per barrel ($/BBL) and an average Henry Hub natural gas price of 21.07 $/BOE from 2005 to 2020, the elimination of 75% of all venting (0.025 MM BOE/Day) would have required increasing the average oil tax by 20.92 $/BBL, while simultaneously decreasing the natural gas tax by 16.23 $/BOE. The oilfields responsible for the majority of production have a very low gas-oil ratio (median 0.48 BOE/BBL). As a result, the oilfields responsible for 80% of production would experience an oil tax increase smaller than 12.79 $/BBL. This scheme would have also lowered routine flaring by 0.020 MM BOE/Day. On average, the tax reform would reduce emissions caused by the mismanagement of co-extracted natural gas by 23.41%, resulting in a decrease of 0.109 MM TCO_2e/Day. The reduction would entail decreases in both carbon dioxide (0.009 MM TCO_2e/Day) and methane (0.099 MM TCO_2e/Day) emissions. Figure 6.2 shows the economic and environmental consequences of the four venting reduction targets, displaying the changes in oil and gas taxation, as well as the targets' environmental consequences, in absolute and relative terms. Our economic and environmental estimates are prudent across three dimensions. First, we design the identification strategy to quantify the policy lower bound. Second, we assume a 98% efficiency rate for the flare stack, which is the industry standard. However, a recent study suggests that

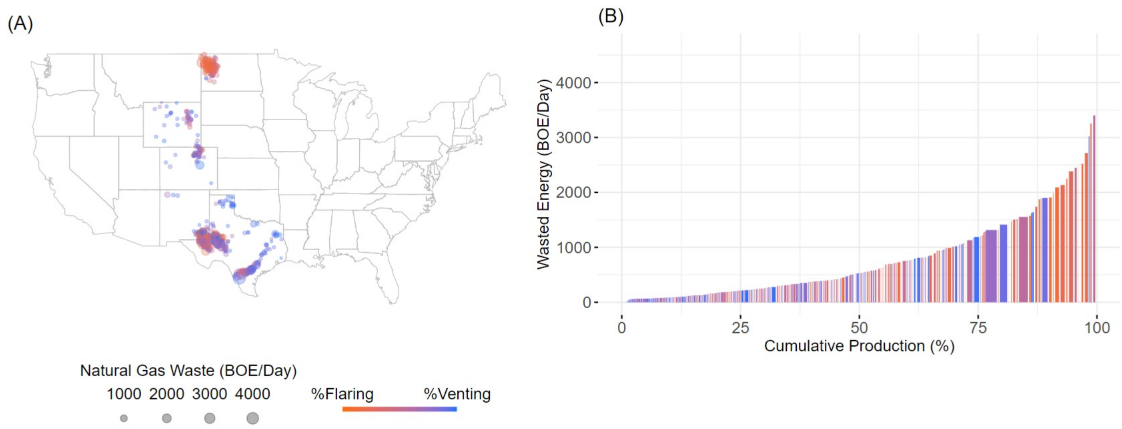

Figure 6.1: Geolocation and Composition of Energy Waste. (A) **Map of the United States** Each dot represents an oilfield. The dot size indicates the magnitude of the energy waste (flaring + venting + leaking) measured in Barrels of Oil Equivalent per Day (BOE/Day). The dot color represents the composition of the waste (Orange = Flaring, Light Blue = Venting + Leaking). (B) **Merit-Based Curve of Energetic Waste** Each bar represents an oilfield. The bar base shows the volume of oil produced, while the bar height indicates the volume of energy waste. The bar color represents the composition of the waste (Orange = Flaring, Light Blue = Venting + Leaking).

the rate of efficiency might be as low as 91% (6), in which case the flaring savings would be significantly higher due to a change in the carbon dioxide/methane composition, resulting in an additional reduction of 0.006 MM TCO_2e/Day in emissions (for the 75% savings scenario). Third, the tax reform replaces natural gas produced by gas-only fields with natural gas recovered from oilfields to ensure that the aggregate supply remains unchanged. Alternatively, regulators could incentivize the use of recovered natural gas as a coal substitute by imposing a modest tax on coal purchases by power plants (approximately 0.002 $ per million British Thermal Units) and leveraging the inter-fuel elasticity of substitution within the power sector. Had this approach been implemented, the reform could have further reduced emissions by up to an additional 0.023 MM TCO_2e/Day (see SM, sections 6.5.4 and 6.5.5).

Policy Implications The proposed reform differs from Pigouvian taxation, emissions markets, and other incentivization schemes across three dimensions. First, these policies are prone to suffer from misreporting and limited legal enforcement, because small deliberate gas emissions are hard to detect and punish. The proposed reform is immune to these issues because oil and gas sales are thoroughly documented in firms' income statements. Thus, oil and gas sales are easy to monitor, unambiguously intentional, and successfully taxed in many countries. Furthermore, the field gas-oil ratio can be accurately measured using a test separator and verified through multiple empirical formulas [Al Dhaif et al., 2021]. Second, these policies can harm firms' profits, consumer purchasing power, and government revenue. Thus, lobbying by the hydrocarbon sector, political opposition (e.g., anti-environmentalist populism), and the risk of social unrest (e.g., yellow vest-type protests) can undermine their implementation. The proposed reform is more politically feasible because it maintains the profitability of oil and gas firms leaving their overall tax burden unchanged by redistributing a portion of the tax burden from co-extracted natural gas to oil sales. As a result, firms have no incentive to lobby against such a reform and should be willing to adopt the proposed scheme in place of the current tax regime, even on a voluntary basis [Cicala et al., 2023]. For the same reason, fraudulent practices (e.g., misreporting venting amounts) are meaningless since disclosing actual methane emissions is typically incentive compatible. Moreover, the reform does not reduce either consumers' lifetime income or governmental tax revenue, resulting in unchanged purchasing power and improved welfare. As a result, voters and policymakers have no incentive to lobby against the reform. Third, Pigouvian taxation, emissions markets, and other incentivization schemes implemented by a single country may cause a loss of competitiveness for domestic firms, resulting in profit reduction, carbon leakage, and production relocation to less

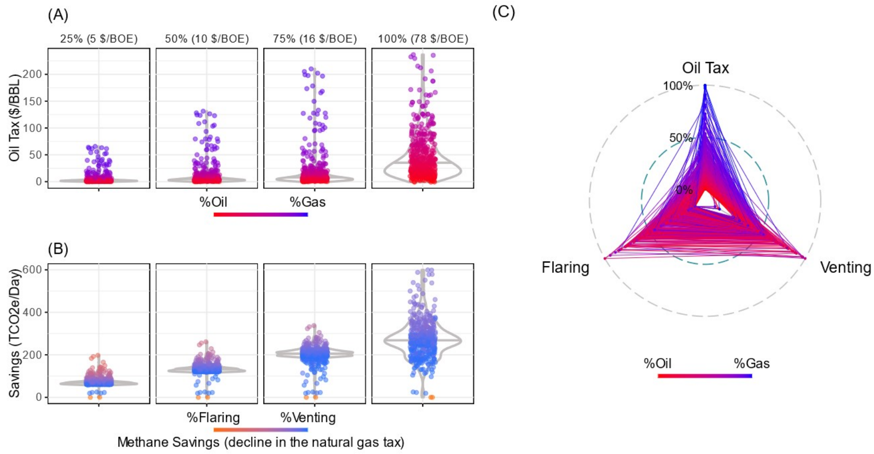

Figure 6.2: Economic and Environmental Consequences of Four Venting Reduction Targets (25%, 50%, 75%, 100%). (A) **Empirical Probability Density Function of the Increase in Oil Tax for a Given Decline in Natural Gas Tax** Each dot represents an oilfield. The distance of the dot from zero indicates the required increase in the specific oil tax for that oilfield. The color of the dots indicates the gas-oil ratio (Red = Oil, Blue = Natural Gas) measured in Barrels of Oil Equivalent per Barrel of Oil (BOE/BBL). Most oilfields should expect a very modest increase in their oil tax in the 25% (associated to a decline in the specific natural gas tax of 5 $/BOE), 50% (associated to a decline in the specific natural gas tax of 10 $/BOE), and 75% (associated to a decline in the specific natural gas tax of 16 $/BOE) scenarios. In the 100% scenario (associated to a decline in the specific natural gas tax of 78 $/BOE), approximately one-quarter should expect an increase of more than \$100/BBL. (B) **Empirical Probability Density Function of the Abatement of Routine Flaring and Venting Emissions for a Given Decline in Natural Gas Tax** Each dot represents an oilfield. The distance of the dot from zero indicates the savings in metric tons of carbon dioxide equivalent per day (TCO_2e/Day) associated with the change in taxation presented in (A). The color of the dots indicates the composition of the savings (Orange = Flaring, Light Blue = Venting). As savings increase, the relative importance of venting increases. (C) **Radar Plot of the Link Between Oil Tax, Routine Flaring, and Venting** Each triangle represents an oilfield. Most oilfields can recover 100% of their energetic waste with a relatively small percentage increase in their oil tax.

stringent jurisdictions. Thus, their viability depends upon either lengthy transnational negotiation or the introduction of a carbon border adjustment mechanism, which might violate World Trade Organization rules [Böhringer et al., 2022]. By contrast, our proposed tax reform has no impact on international trade, so even a single country can implement it in isolation without facing loss of competitiveness or carbon leakage (see SM, section 6.5.2, SWOT Table).

Although our solution is not the most theoretically efficient mechanism for tackling methane emissions, it is a cost-effective, lobby-resilient, and readily-implemented strategy to drastically reduce RFV (see SM, sections 6.5.1 and 6.5.2). We hope that public (e.g., the World Bank and the United States Department of Energy) and private (e.g., the Environmental Defense Fund) energy modelers will incorporate the following two results into their future policy proposals. First, the introduction of anti-flaring and/or anti-venting policies may have limited or adverse effects on the methane footprint of the oil and gas industry. Second, regulators might overcome the difficulty of monitoring hard-to-quantify emissions by designing incentive schemes which rely on alternative and easily measurable quantities. In the case of the oil and gas sector, the reservoir gas-oil ratio is a good proxy for the potential social cost of oil extraction.

Supplementary Material

6.1 Model Setup

We study an infinite-horizon production economy with $S \geq 2$ sovereign countries populated by four types of agents: oil & gas firms (upstream, section 6.1.1, refineries & transformation firms (midstream, section 6.1.2), consumers (downstream, section 6.1.3), and national governments (section 6.1.4). The economy is affected by climate change caused by anthropogenic greenhouse gas (GHG) emissions (section 6.1.5). In this section we present the setup of the model. Section 6.2 describes the main analytical results. Section 6.3 illustrates the identification of the structural parameters of the model.

6.1.1 Upstream: Oil&Gas Firms

In each country s there are K^s infinitely-living oil&gas profit-maximizing firms. Firm k in country s owns I^{ks} oil&gas fields denoted by $i \in \left\{1, 2, ..., I^{sk}\right\}$ and compete in a Cournot oligopoly fashion on the crude and natural gas markets.[1]

Production Technology. Let Oil_t^{iks} and Gas_t^{iks} be the Barrel of Oil Equivalent (BOE) amounts of crude and natural gas sold by field i in period t, respectively. Flare_t^{iks} denotes the BOE amount of natural gas flared and PInS_t^{iks} the amount purchased for in-situ use for electricity production or heating purposes. M_t^{iks} is the maintenance capital accumulated by the field and Z_t^{iks} the US dollar value of other net outputs, such as labor and electricity.[2] Lastly, ReInS_t^{iks} denotes the amount of extracted gas that is reused in-situ[3] and NRF_t^{iks} is the amount of flaring that field i cannot avoid to produce given the technology available in period t (minimum non-routine flaring). Following standard Microeconomic Theory, we assume that the oil&gas production technology is described by a field-specific real analytic transformation function $TF_t^{iks} : (-\infty, +\infty)^7 \to \mathbb{R}$, whose argument, the net output vector, writes $\left(\text{Oil}_t^{iks}, \text{Gas}_t^{iks}, \text{Flare}_t^{iks}, \text{PInS}_t^{iks}, \text{M}_t^{iks}, \text{Z}_t^{iks}; \text{ReInS}_t^{iks}\right)$. We assume that natural gas venting is a costless output and does not affect the production technology in any way other than, of course, through regulatory and fiscal costs, which we consider separately from the production technology in the next section. As a consequence of this assumption, we suppress Vent_t^i from the arguments of $TF_t^{iks}(\cdot)$. Thus, the production set is defined by the inequality:

$$TF_t^{iks}\left(\text{Oil}_t^{iks}, \text{Gas}_t^{iks}, \text{Flare}_t^{iks}, \text{PInS}_t^{iks}, \text{M}_t^{iks}, \text{Z}_t^{i}; \text{ReInS}_t^{iks}\right) \leq 0\ . \tag{6.1}$$

Note that $\text{PInS}_t^{iks}, \text{M}_t^{iks}, \text{Z}_t^{iks}; \text{ReInS}_t^{iks}$ are net input whose value is allowed to be positive or negative. The sign of these variables is determined endogenously and is shaped by the function TF_t^{iks}. For instance, the assumptions $\frac{\partial TF_t^{iks}(\cdot)}{\partial \text{Oil}_t^i} \leq 0$ for all $\text{Oil}_t^{iks} \leq 0$ ensure that the oil production is always weakly positive. In particular, we restrict the attention to the class of weakly separable and twice differentiable convex transformation functions in the form:

$$\begin{aligned} TF_t^{iks} = \quad & F_t^{iks}\left(TF_{1t}^{iks}\left(\text{Oil}_t^i\right) + TF_{2t}^{iks}\left(\text{Gas}_t^i\right) + TF_{3t}^{iks}\left(\text{Flare}_t^i\right)\right. \\ & \left. + TF_{4t}^{iks}\left(\text{PInS}_t^{iks}; \text{ReInS}_t^{iks}\right) + TF_{5t}^{iks}\left(\text{M}_t^{iks}\right) + TF_{6t}^{iks}\left(\text{Z}_t^{iks}\right)\right) \end{aligned} , \tag{6.2}$$

where F_t^{iks} is strictly increasing and twice differentiable. Lastly, the production of field i in period t is bounded

[1] However, oil fields usual produce modest quantities of gas, which make the effect of their productive choices on gas prices small or negligible. Because of that, we assume that oil firms are price-takers on the natural gas market.

[2] Note that the domain of Z_t^{iks} is $(-\infty, +\infty)$. However, it typically takes negative values because it includes the value of all productive inputs, such as labor, energy purchases, etc.

[3] Note that formally ReInS_t^{iks} is not a net output of the field transformation function, the quantity of gas reused in-situ affects the production technology by partially replacing other sources of energy required in the production process. Also note that the domain of each net input is the entire set of real number. However, when we setup the firm's problem, we restrict the range of feasible values by adding appropriate constraints (e.g., $\text{Oil}_t^{iks} \geq 0$)

above by its capacity K_t^i, such that the aggregate production of hydrocarbons must satisfy the following inequality:

$$\mathrm{Oil}_t^{iks} + \mathrm{TotGas}_t^{iks} \leq \mathrm{K}_t^{iks} \tag{6.3}$$

Where the variable TotGas_t^{iks} denotes the total amount of gas extracted from the field in period t. Note that the constraint in (6.3) is stated in terms of total hydrocarbons measured in BOE. Under the technological constraints we introduce in section 1.1.1, this modelling choice is equivalent to imposing a constraint on the oil production capacity. The management can sell, flare, vent, re-inject or use the extracted gas in-situ for electricity production, therefore its formula writes:

$$\mathrm{TotGas}_t^{iks} = \mathrm{Gas}_t^{iks} + \mathrm{Flare}_t^{iks} + \mathrm{TotVent}_t^{iks} + \mathrm{ReInj}_t^{iks} + \mathrm{ReInS}_t^{iks}\ , \tag{6.4}$$

where $\mathrm{TotVent}_t^{iks}$ denotes the total amount of gas released in the atmosphere in period t.

Investment, Development, and Discoveries. The field faces investment costs in field development and new discoveries, as well as in field's capacity in the form, represented by a real analytic function

$$\begin{gathered} InvCost_t^{iks}\left(\mathrm{ID}_t^{iks}, \mathrm{IM}_t^{iks}, \mathrm{ReInj}_t^{iks}, \mathrm{PInj}_t^{iks}, \mathrm{OInj}_t^{iks}; PP_t^{iks,\mathrm{Gas}}\right) = \\ \mathrm{ID}_t^{iks} + \mathrm{IM}_t^{iks} + PP_t^{iks,\mathrm{Gas}}\mathrm{PInj}_t^{iks} + IC_t^{iks}\left(\mathrm{ReInj}_t^{iks}, \mathrm{PInj}_t^{iks}\right) + C_t^{iks}\mathrm{OInj}_t^{iks} \end{gathered}\ , \tag{6.5}$$

where ID_t^{iks} is the US dollar amount of investment in field development and exploration, IM_t^{iks} is investment in field maintenance and $PP_t^{iks,\mathrm{Gas}}$ denotes the purchase price of natural gas from nearby fields.[4] The field capacity can be increased through the discovery and development of new reserves. Alternatively, the field's management can increase the pressure of the reservoir by injecting natural gas or other liquids and/or gases through injection wells. Injections ReInj_t^{iks} and PInj_t^{iks} denote the amounts injected natural gas produced by the field and purchased from nearby fields, respectively, whereas OInj_t^{iks} is the gas-equivalent amount of other types of injections, such as steam and chemicals. The increase in field capacity depends upon total injections, $\mathrm{TotInj}_{t-1}^{iks} = \mathrm{ReInj}_t^{iks} + \mathrm{PInj}_t^{iks} + \mathrm{OInj}_t^{iks}$. Lastly, the field's capacity declines with the amount of extracted hydrocarbons, capturing the fall in well pressure due to depletion. In detail, the capacity of oil field i in period $t+1$ solves the following inequality:

$$\mathrm{K}_{t+1}^{iks} \leq \mathrm{K}_t^{iks} + D_t^{iks}\left(\mathrm{ID}_t^{iks}, \mathrm{L}_{t-1}^{iks}\right) + B_t^i\left(\mathrm{TotInj}_{t-1}^{iks}\right) - \zeta\left[\mathrm{Oil}_t^{iks} + \mathrm{TotGas}_t^{iks}\right] \tag{6.6}$$

where D_t^{iks} and B_t^{iks} are real analytic functions capturing the effect on the field's capacity of investment in discoveries and injections, respectively[5], whereas L_{t-1}^{iks} denotes the cumulative investment in new discoveries up to period $t-1$ and follows a low of motion:

$$\mathrm{L}_t^{iks} = \mathrm{L}_{t-1}^{iks} + \mathrm{ID}_t^{iks}\ . \tag{6.7}$$

Before describing the firm's profit maximization problem we describe in detail the technological and regulatory constraint an oil&gas firm faces in the management of the natural gas produced by each field.

The Natural Gas Management Problem

Technological Constraints. Consider an oil&gas field i owned by firm k. In each period t, field i extracts Oil_t^{iks} and a quantity of natural gas, denoted by TotGas_t^{iks} and measured in BOE. Gas extraction may be either the outcome of a deliberate choice of the management or a byproduct of oil production. In both cases, we assume a constant field-specific gas-to-oil ratio $GOR^{iks} \in [0, +\infty)$, and impose the constraint $\mathrm{TotGas}_t^{iks} \geq GOR^{iks}\mathrm{Oil}_t^{iks}$.

[4] Note that the purchase price of gas $PP_t^{iks,\mathrm{Gas}}$ faced by field i is allowed to differ from the gross sales price $P_t^{iks,\mathrm{Gas}}$. This capture cases in which oil firms purchase gas from nearby fields that are not connected to a gas pipeline at a cheaper-than-market price.

[5] The use of an inequality constraint captures the possibility that the firm chooses to disregard some of its productive capacity, for instance by postponing the start of productive activity of some newly tapped wells.

This constraint captures the fact that a certain quantity of natural gas is extracted as a by-product of oil production and is trivially non-binding for gas-only fields. The total quantity of gas vented is divided into two macro categories. A quantity vented intentionally IVent_t^{iks} and a quantity vented unintentionally $\text{UVent}_t^{iks} = UVent_t^{iks}\left(\text{Oil}_t^{iks}, \text{M}_t^{iks}\right)$ (leaking, also known as unintentional venting),

$$\text{TotVent}_t^{iks} = \text{IVent}_t^{iks} + UVent_t^{iks}(\text{TotGas}_t^{iks}, \text{M}_t^{iks}) \tag{6.8}$$

The former is defined as the amount of gas vented as a direct and deterministic consequence of a deliberate action or omission by the firm's personnel which is not justified by true health and safety concerns. It is mostly due to the disposal of gas accumulated at the top of oil tanks. It therefore excludes leakages due to poor maintenance. The latter is for the most part caused by the cleaning, testing, and poor maintenance of the gas equipment and safety-related pressure releases.

The set of feasible net output vectors is determined by the production technology. Following standard Microeconomic Theory, we assume that the production technology of the oil field in period t is described a real analytic transformation function $TF_t^{iks} : (-\infty, +\infty)^6 \to \mathbb{R}$, whose argument – the net output vector – is described in detail in the next paragraphs. Note that TF_t^{iks} has a very large domain. However, we impose constraints to the firm's choice problem (e.g., $\text{Oil}_t^{iks} \geq 0$) in order to avoid unplausible outcomes. For the purpose of modeling the gas management problem, we impose three key restrictions on TF_t^{iks}. First, even if natural gas is often treated as a by-product of oil extraction by oil firms, its production for commercial purposes is costly for the firm; i.e., $\frac{\partial^2 TF_t^{iks}(\cdot)}{\partial \text{Gas}_t^{iks}} > 0$, because it requires energy to capture and compress the extracted gas prior to entering the market. Second, flaring is also (weakly) costly and such cost is weakly increasing and convex in the quantity of gas flared,

$$\frac{\partial TF_t^{iks}(\cdot)}{\partial \text{Flare}_t^{iks}} \geq 0 \; \frac{\partial^2 TF_t^{iks}(\cdot)}{\partial \text{Flare}_t^{iks2}} > 0 \tag{6.9}$$

for all possible values of the argument of TF_t^{iks}. This convex cost structure is motivated by technological considerations. First, the high pressure gas contained in the heater-treater can be flared at a very small marginal production cost –virtually equal to zero[6]. However, the low pressure gas contained in the oil tank cannot be flared at a marginal cost equal to zero. It is necessary to use a small compressor, an air assisted blower, or a gas assist options to get the gas out of the tank in a pressurized form and then burn it, see Figure 6.3. This operation has a positive marginal cost, which is possibly increasing in the quantity of gas flared because the energy required to get the gas out of the tank increases at a more-than-proportional rate as the amount of residual gas in the tank decreases. The magnitude of $\frac{\partial TF_t^{iks}(\cdot)}{\partial \text{Flare}_t^{i}}$ depends upon the specific configuration of every field. However, a rule-of-thumb estimation can be done multiplying the quantity of natural gas or electricity needed to re-pressurize the low pressure gas by the cost of electricity (for a detailed description see section VRU (lines 54-63) of the OPGEE 3.0 manual Brandt et al. [2020])[7]. Lastly, we assume that the amount of unintentional venting is a real analytic function of the amount of crude extracted Oil_t^{iks} and the stock of maintenance capital M_t^{iks} in the form:

$$UVent_t^{iks}\left(\text{TotGas}_t^{iks}, \text{M}_t^{iks}\right) = \vartheta^{iks}\text{TotGas}_t^{iks} - Maint_t^{iks}\left(\text{M}_t^{iks}\right) + \epsilon_t^{iks}\,, \tag{6.10}$$

where $Maint_t^{iks}(\cdot)$ is weakly increasing and concave, and ϵ_t^{iks} is an i.i.d. shock with $\mathbb{E}\left[\epsilon_t^{iks}\right] = 0$. Given the the assumption on the functional form of $UVent_t^{iks}$ stated above, the formula for TotGas_t^{iks} in (6.4) can be solved

[6]We assume that not all the gas, which goes into the flare-stack, is flared. In articular, we assume a 98% flaring rate within the flare stack, which corresponds to the best practice in the industry. In presence of strong wind and/or low tech combustors, the flaring efficiency could decline (as low as $\sim 91\%$) and a larger part of the gas in the heater-treater could be vented.

[7]Note that in world with stringent regulation on methane emissions also the flaring of high pressure gas is costly because the flare stuck must be maintained to ensure that all the gas is combusted all the time.

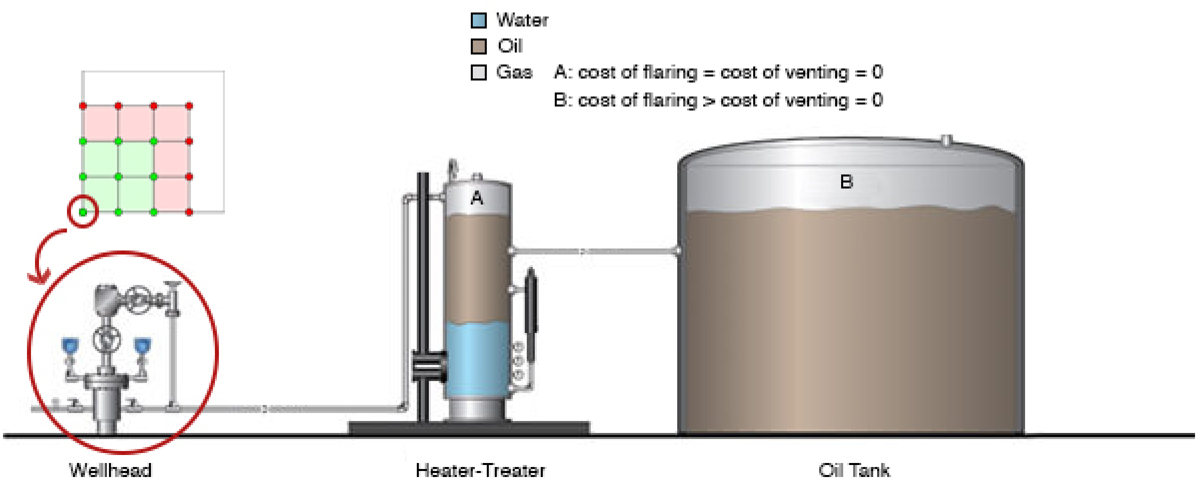

Figure 6.3: Flaring Marginal Costs according to the pressure of the co-extracted Gas.

recursively to obtain:

$$\text{TotGas}_t^{iks} = \left(1 - \vartheta^{iks}\right)^{-1} \left[\text{Gas}_t^{iks} + \text{Flare}_t^{iks} + \text{IVent}_t^{iks} - Maint_t^{iks}\left(\text{M}_t^{iks}\right) + \text{ReInj}_t^{iks} + \text{ReInS}_t^{iks}\right] \quad (6.11)$$

Lastly, the stock of maintenance capital follows a law of motion:

$$\text{M}_{t+1}^{iks} = \text{M}_t^{iks}\left(1 - \rho^i\right) + \text{IM}_t^{iks} \quad (6.12)$$

where IM_t^{iks} is the investment in the maintenance of field i made in period t.

Information. In order to correctly design the regulatory and fiscal framework in the next paragraphs, we first describe the information set of each player. Each oil&gas firm is assumed to possess full information at the moment in which production decisions are made. That is, a firm's information set Ω_t^{ks} includes the full history of prices, own and other firm's costs, own and other firm's decisions and outcomes, tax rates and regulation in place. Moreover, each firm knows the future realizations of all fields' marginal costs and all other time-variant exogenous variables, and possess perfect foresight regarding all endogenous variables, such as prices and other firms' production choices. Because of this assumption, we can omit expectations in the firm's problem and treat it as an optimization in a deterministic environment. Note that under these assumptions, a the solution of the problem of a firm choosing all it production plans in period 1 for all periods $t = 1, 2...$ is identical to that of a firm choosing the production plan for each period $t = 1, 2, ...$ at the beginning of such period.

All the information –with one piece of information being a notable exception– is assumed to be public and contractible for the government, whose information set is denoted by Ω_t^{PUB}. For instance, oil and gas sales are well-documented in the firm's balance sheets. Moreover, they are relatively easy to measure and verify for the regulatory authority, implying that substantial misreporting for these variables is very unlikely. Regarding flaring, the regulator may not rely solely on self-reported quantities, which could be distorted using under/over billing tricks. In particular, the regulator can assess the volumes of disposed gas using quantity-monitoring technologies. They can supervise flaring activities using satellite, airplane and in-person tracking. All these methods tend to be accurate. Given these considerations, we assume thatOil $_t^i$, Gas_t^i, and Flare_t^i are fully observable and contractible by all agents. However, some information is not publicly available. In particular, gas venting is deemed hard to detect, measure and attribute to a specific emitter [Allen, 2020]. In principle, the regulator can supervise venting using technologies similar to the ones adopted to monitor flaring. However, in the case of venting bottom-up as well as top- down measures tend to be inaccurate. A general lack in the understanding of the spatio-temporal heterogeneity of methane emissions renders these measures prone to commit measurement errors. Furthermore, most legislation regulate intentional venting, which is not easy to separate form unintentional venting. In other words, the regulator wants to supervise intentional venting but it is incapable to separate this quantity from unintentional venting and/or measurement difficulties. Moreover, even if some amount of venting is detected, it may be challenging for the regulatory authority to establish in legally binding terms that such venting occurred as the result of a voluntary action of the firm's personnel which was not justified by health and safety reasons. Thus, we assume that IVent_t^{iks} and UVent_t^{iks} are observable with probability equal to 1 by firm i only. The public only receives an imperfect contractible public signal $ivent_t^{iks} \in \{0,1\}$, where $ivent_t^{iks} = 1$ only if $\text{IVent}_t^{iks} > 0$, such that $Pr_t^{iks}\left(ivent_t^{iks} = 1 \mid \Omega_t^{PUB}, \text{IVent}_t^{iks} > 0\right) \in (0,1)$ and $Pr_t^{iks}\left(ivent_t^{iks} = 1 \mid \Omega_t^{PUB}, \text{IVent}_t^{iks} = 0\right) = 0$. This implies that even if TotGas_t^{iks}, Gas_t^{iks}, Flare_t^{iks}, ReInj_t^{iks}, ReInS_t^{iks} are public information, such that it is possible to obtain a reliable measure of total venting Vent_t^{i}, the intentional part IVent_t^{i} is only partially observable and contractible.

Flaring & Venting Regulation. firms in their production decisions are not solely shaped by technology. One must also account for the legal and fiscal restrictions that both flaring and venting face in most countries.

In the United States, flaring and venting are regulated at the federal and at the local level. Federal laws focus on offshore production, local requirements on onshore. Offshore fields must require flaring and venting permits to dispose of the extracted Gas. The permits are released by the Interior Bureau of Safety and Environmental Enforcement if at least one of the following criterion is met: 1) the national interest requires it (e.g. when a major hurricane could cause infrastructure damage), 2) the production from a completed well would likely be permanently lost, or 3) the short-term flaring or venting would likely yield a smaller volume of lost Gas than if the facility were to shut in and restart later (with flaring and venting necessary to restart the facility). Similarly, onshore fields must require flaring and venting permits. The latter are released by State Environmental agencies. Several states release unlimited flaring permits, while regulating/forbidding intentional venting (Alabama, Arizona, Colorado, Florida, Illinois, Indiana, Kentucky, Louisiana, Mississippi, Missouri, Nebraska, Nevada, Ohio, Oklahoma, Pennsylvania, South Dakota, Tennessee, West Virginia). Other states release a limited number of flaring permits, those capping the quantity of flaring, while regulating/forbidding intentional venting (Alaska, California, Idaho, Kansas, Michigan, Montana, New Mexico, North Dakota, New York, Texas, Utah, Virginia, Wyoming). Finally, one state (Arkansas) taxes flaring, while regulating/forbidding intentional venting. Table 6.1 provides the legal sources of the current flaring and venting regulation in the United States.

Table 6.1: Sources of Legal Regulation

Region	Source
State Regulation	
Alabama	State Oil & Gas Board of Alabama Administrative Code, Rules 400-1 - 400-7. Alabama Statute, Title 09, Chapter 17, Section 9-17-11.
Alaska	Alaska Oil & Gas Conservation Act, Section 31.05.095
Arizona	Arizona Administrative Code Title 12; Chapter 7, Section R12-7-138
Arkansas	Arkansas Code, Title 15, Section 15-72-105 and Section 15-72-208
California	California Code of Regulations, Title 17, Division 3, Chapter 1, Subchapter 10 Climate Change, Article 4, Sub-article 13
Colorado	Colorado Code of Regulations, Rule 912, Page 183
Florida	Florida Statutes and Rules, Chapter 377, 62C-25 - 30
Idaho	Idaho Administrative Rule 20.07.02, Sections 413 and 430
Illinois	Illinois Oil and Gas Act (225 ILCS 725). Illinois Hydraulic Fracturing Regulatory Act, Section 245.900 and 245.910
Indiana	Indiana Code, Title 14, Article 37, Chapter 11, Subsection 14-37-11-1. Indiana Administrative Code, Title 312, Article 29, Subsection 312 IAC 29-3-3.
Kansas	Kansas Statute 55-102b. Kansas Administrative Regulations, Sections 82-3-208, 82-3-209, 82-3-314
Kentucky	Kentucky Revised Statutes Chapter 353, Section 353.160, 353.520
Louisiana	Louisiana Administrative Code, Title 33, Part III. Title 43, Part XIX
Michigan	Natural Resources and Environmental Protection Act, 1994, Public Act 451, Part 615
Mississippi	Mississippi Statewide Rules and Regulations, Rule 62
Missouri	Revised Statues of Missouri, Chapter 259.060
Montana	Administrative Rules of Montana (ARM) 17.8.1603, 17.8.1711, 36.22.1220
Nebraska	Revised Statutes of Nebraska, Chapter 57, Section 902 and 903
Nevada	Nevada Revised Statutes, Section 522.010 and 522.039. Nevada Administrative Codes, Section 522.3
New Mexico	New Mexico Administrative Code, Chapter 15, Title 19, Subsection 18
New York	New York Codes, Rules and Regulations, Title 6, Parts 550-559, Chapter V, Subchapter B
North Dakota	North Dakota Administrative Code 33.1-15-07-02, 33.1-15-03-03.1, 33.1-15-20. North Dakota Industrial Commission Order No. 24665
Ohio	Ohio Revised Code Title 15, Chapter 1509.20. Ohio Administrative Code, Chapter 1501:9-9
Oklahoma	Oklahoma Register, Chapter 10, Subsection 3-15. Oklahoma Statues 2-5-102, et seq.. Oklahoma Administrative Code, Title 252, Chapter 100
Pennsylvania	No specific Regulation
South Dakota	Administrative Rules of South Dakota, Article 74:12, Section 74:12:05:04
Tennessee	Tennessee Code, Chapter 1, Title 60, Section 60-1-101 and Section 60-1-102
Texas	Texas Air Quality State Implementation Plan, Regulation 30 TAC 115.720-115.729. Texas Administrative Code, Title 16, Part 1, Chapter 3, Rule 3.32
Utah	Utah Administrative Code, Rule R649-3
Virginia	Code of Virginia, Title 45.1, Chapter 22.1. Virginia Administrative Code, Title 4, Agency 25, Chapter 150
West Virginia	West Virginia Code, Chapter 22, Articles 5 and 18. West Virginia Legislative Rules, 45CSR, Series 6 and 13
Wyoming	
Federal Regulation	
US Federal Offshore	At the discretion of the Department of the Interior's Bureau of Safety and Environmental Enforcement (BSEE)

In order to capture these legal and fiscal restrictions, we assume that each oil&gas field also faces expected regulatory costs $RegCost_t^{iks}$, which depend upon the fines, taxes and emission permits the firm must pay to the government (if any). The flaring regulation is in the form of a threshold $\overline{\text{Flare}}_t^{iks}$[8], such that a violation occurs whenever the amount of gas flared exceed the threshold.[9] In line with the current regulation of all US states we assume that intentional venting is illegal, such that the firm faces a fine whenever $\text{IVent}_t^{iks} > 0$ is detected.

Let VF_t^{iks} and FF_t^{iks} denote the fines for violation of the venting and flaring regulation, respectively. Because flaring is assumed to be fully observable and contractible by the regulatory authority, any violation of the flaring regulation results in a fine with probability equal to 1. Conversely, because intentional venting is not observable and/or contractible, a violation of the venting regulation results in a fine only when a signal $ivent_t^{iks} = 1$ is observed, which occurs with probability $Pr_t^{iks}\left(ivent_t^{iks} = 1 \mid \Omega_t^{ks}\right)$. For simplicity, we model the expected regulatory costs of venting as an increasing and strictly convex real analytic function of IVent_t^{iks} and Flare_t^{iks}; i.e., $VF_t^{iks} \times Pr_t^{iks}\left(ivent_t^{iks} = 1 \mid \Omega_t^{ks}\right) = VF_t^{iks}\left[PF_t^{iks}\left(\text{IVent}_t^{iks}, \text{Flare}_t^{iks}\right) + PU_t^{iks}\left(\text{M}_t^{iks}\right)\right]$, where the inclusion of Flare_t^{iks} in the arguments of PF_t^{iks} captures any possible substitutability between intentional venting and flaring for the oil&gas firm, whereas the second term in the square brackets PU_t^{iks} models the possibility that the level of maintenance affects the probability of detection of intentional venting, for instance because it may be hard for the regulatory authority to provide sufficient evidence that a given amount of venting performed by a poorly maintained field is the result of a deliberate action of the field management rather than an unintentional leakage due to damaged pipelines. Given these assumptions the formula for regulatory costs write:

$$\begin{aligned} RegCost_t^{iks}\left(\text{IVent}_t^{iks}, \text{Flare}_t^{iks}, \text{M}_t^{iks}\right) = \; & VF_t^{iks}\left[PF_t^{iks}\left(\text{IVent}_t^{iks}, \text{Flare}_t^{iks}\right) + PU_t^{iks}\left(\text{M}_t^{iks}\right)\right] \\ & + FF_t^{iks} \times \mathbf{1}\left[\text{Flare}_t^{iks} > \overline{\text{Flare}}_t^{iks}\right] \end{aligned}, \quad (6.13)$$

While these assumptions deliver an admittedly stylized picture of the highly heterogeneous regulatory framework concerning flaring and venting in US States, we believe that it represents a useful simplification that incorporates all the key features for the purpose of this analysis.

Market Structure. We assume that the markets for crude oil and unrefined natural gas are imperfectly competitive. Specifically, oil firms compete on the crude market in a global Cournot-style oligopoly and are price-taker on the gas market. There is a unique global average crude price P_t^{Oil}, but individual fields face different prices $P_t^{iks,\text{Oil}} = P_t^{\text{Oil}} + \sigma^{iks}$, where σ^{iks} captures the time-invariant quality of crude from field i. Conversely, we assume that the quality of natural gas is identical across fields, and that gas firms compete in an oligopoly in quantities on the unrefined natural gas market, whose demand side consists of a number of midstream firms (gas-processing facilities), but we allow for gas produced in different countries to be imperfect substitutes. Thus, we allow for different gas prices $P_t^{s,\text{Gas}}$ across different countries. This assumption captures the geographically segmented nature of the natural gas market at international level and accommodates for the possibility of transport costs and bottlenecks affecting specific local segments of the market. Lastly, there is a market for each consumption good. Both midstream firms and consumers are assumed to be price-taker on all markets they participate in. The price of general consumption (numéraire) is normalized to 1, whereas the prices of oil&gas goods are listed in the vector $\mathbf{p}_t^s$ and they are allowed to differ across country. The collection of all prices in period t is $\mathbf{P}_t = \left\{P_t^{\text{Oil}}, \left\{P_t^{s,\text{Gas}}, \mathbf{p}_t^s\right\}_{s=1}^{S}\right\}$, whereas $\mathbf{P} = \{\mathbf{P}_t\}_{t=1}^{\infty}$ denotes the collection of all prices in all periods. All prices are assumed to be endogenously determined in a competitive equilibrium in each period t and all agents possess perfect foresight regarding future equilibrium prices. In particular, the price of crude oil must clear the global oil market, whereas each country possesses its own market for unrefined natural gas, resulting in a country-specific

[8]The *Zero Routine Flaring by 2030 Initiative* aims to set $\overline{\text{Flare}}_t^{iks} \cong 0$. We allow for a more general regulatory scenario with$\overline{\text{Flare}}_t^{iks} \geq 0$.

[9]Current United States regulation expresses $\overline{\text{Flare}}_t^{iks}$ as a ratio between the volume of Gas flared and the total volume of Gas extracted, $\overline{\text{Flare}}_t^{iks} = \text{Flare}_t^{iks}/\text{TotGas}_t^{iks}$. Other legislation (e.g. Canada) express $\overline{\text{Flare}}_t^{iks}$ as a ratio between the volume of gas flared and the volume of oil extracted, $\overline{\text{Flare}}_t^{iks} = \text{Flare}_t^{iks}/\text{Oil}_t^{iks}$.

natural gas price that clears such market.

Fiscal Framework. Oil&gas firms in the US face a complex system of taxation, fees, and royalties collected at both Federal, State, and local level. For the sake of simplicity, we reduce this complex and heterogeneous framework to a stylized system of linear taxes. Moreover, the assumption on the information structure of the model implies that taxation can be imposed only on quantities that are observable and contractible by the government. As a result, the government can tax flaring, but taxation of intentional venting is not feasible. In principle, the regulator could impose a tax on all natural gas released into the atmosphere by the firm, irrespective of its intentional or unintentional nature. However, this regulatory approach could lead to unintended and undesirable safety consequences. By making it expensive for the firm to implement safety-related pressure relief measures, it could inadvertently raise the risk of explosions. Moreover, it would represent a monetary incentive to emission misreporting, as illustrated in section 5.2. Therefore, we propose an incentive scheme that avoids penalizing unintentional leaks. In detail, we assume a tax system on oil&gas upstream firms featuring: (i) a linear tax rate T_t^s on corporate income, defined as the firm's revenue minus total costs excluding the payment of fines; (ii) a vector of (possibly field-specific) specific linear taxes $\left(\tau_t^{iks,\text{Oil}}, \tau_t^{iks,\text{Gas}}\right)$ on oil and gas sales, which also includes any royalties on hydrocarbon extraction to be paid to the government; (iii) a (possibly field-specific) specific linear tax on flaring, which also includes the unitary cost of any flaring permits the firm may be required to purchase, denoted by $\tau_t^{iks,\text{Flare}}$ (it may be equal to zero). Lastly, firm k may be partially allowed to deduct from the taxable income generated by field i other production and/or investment cost that do not enter directly the firm's balance sheets, such as the value of extracted gas which is re-injected in the field or used in-situ for electricity production. Specifically, the firm's deductible amount is given by function $Deduct_t^{ikt}$, which has formula:

$$Deduct_t^{ikt}\left(\text{ReInj}_t^{iks}, \text{ReInS}_t^{iks}\right) = \delta_{0t}^{iks} + \delta_{1t}^{iks}\left(\text{ReInj}_t^{iks} + \text{ReInS}_t^{iks}\right) \ , \tag{6.14}$$

where δ_{0t}^{iks} denotes lump-sump deductions aiming to captures other off-balance firm-specific costs and δ_{1t}^{iks} denotes the deduction rate for gas re-injected or reused in-situ. These tax provisions are complemented by a system of taxes on other agents (midstream firms and consumers), which are described in detail in the corresponding sections of this Appendix. For a complete definition of a *tax scheme*, see Section 1.4.

Identification of Types of Upstream Fields for Tax Purposes. For tax purposes, oil&gas fields are classified in three mutually-exclusive categories (types): *oil fields*, *gas-only fields*, and *mixed oil&gas fields*. Let $MC_t^{iks,\text{Gas}} = \frac{\partial TF_t^{iks}(\cdot)}{\partial \text{Gas}_t^{iks}} \Big/ \frac{\partial TF_t^{iks}(\cdot)}{\partial \text{Z}_t^{iks}}$ and $MC_t^{iks,\text{Oil}} = \frac{\partial TF_t^{iks}(\cdot)}{\partial \text{Oil}_t^{iks}} \Big/ \frac{\partial TF_t^{iks}(\cdot)}{\partial \text{Z}_t^{iks}}$. We assume that $MC_t^{iks,\text{Gas}}$ is constant in all endogenous variables for all oil fields and impose specific restrictions that depends upon the type field considered, as described in the remainer of this paragraph. Firstly, we assume that all oil fields in each period t satisfy the following condition:

$$\begin{aligned} -\left[P_t^{s,\text{Gas}} - \tau_t^{iks,\text{Gas}} - MC_t^{iks,\text{Gas}}\right]\,\left(1+\vartheta^{iks}\right)^{-1} GOR^{iks} \\ > \left[P_t^{\text{Oil}} + \sigma_1^{iks} - \tau_t^{iks,\text{Oil}} - MC_t^{iks,\text{Oil}}\left(0\right)\right] \geq 0 \end{aligned} \tag{6.15}$$

in all periods $t = 1, 2, ...$ given the tax rates $\tau_t^{iks,\text{Gas}}, \tau_t^{iks,\text{Oil}}$ and equilibrium prices $P_t^{s,\text{Gas}}$, P_t^{Oil} that prevail under the existing tax scheme. Note that, because the tax reform we propose is such that for oil fields the change in $\tau_t^{iks,\text{Oil}}$, denoted by $\Delta\tau_t^{iks,\text{Oil}}$, satisfies $\Delta\tau_t^{iks,\text{Oil}} = -\Delta\tau_t^{s,\text{Gas}}\left(1+\vartheta^{iks}\right)^{-1} GOR^{iks}$, the condition above is unaffected by the introduction of the tax reform at constant prices, meaning that the classification of a given field in the oil category does not change with the introduction of the tax reform as long as the reform does not affect equilibrium prices. Intuitively, condition (6.15) states that oil fields are those fields for which gas production is not profitable. Thus, natural gas production (if any) is a by-product of oil production for those fields. Secondly, we assume that all oil&gas fields satisfy the following condition:

$$\begin{aligned} P_t^{\text{Oil}} + \sigma_1^{iks} - \tau_t^{iks,\text{Oil}} - MC_t^{iks,\text{Oil}}\left(\text{Oil}_t^{iks}\right)\left(1+\vartheta^{iks}\right) \leq \\ P_t^{s,\text{Gas}}\left(1-\varsigma_t^{\text{Gas}}\right) - \tau_t^{iks,\text{Gas}} - MC_t^{iks,\text{Gas}}\left(GOR^{iks}\text{Oil}_t^{iks}\right) \end{aligned} \tag{6.16}$$

for all $\text{Oil}_t^{iks} \geq 0$ in all periods $t = 1, 2, ...$ given the the tax rates $\tau_t^{iks,\text{Gas}}, \tau_t^{iks,\text{Oil}}$ and equilibrium prices $P_t^{s,\text{Gas}}$, P_t^{Oil} that prevail under the existing tax scheme. Note that if a field satisfies the condition above under the existing tax scheme, then it also satisfies the condition under the tax reform proposed in this paper at constant prices, meaning that the classification of a given field in the "other oil&gas" category does not change with the introduction of the tax reform as long as the reform does not affect equilibrium prices. Intuitively, this assumption (6.16) captures the fact that gas production is profitable for this type of fields. Lastly, gas-only fields are characterized by $MC_t^{iks,\text{Oil}}(0) = +\infty$ in all periods $t = 1, 2, ...$, implying that they never produce any positive quantity of crude. Given these assumptions, it is possible to show (see proof to Proposition 3 below) that a field is uniquely identified as an oil field in period t if its production choices satisfy $\text{Oil}_t^{iks} > 0$ and $\text{TotGas}_t^{iks} \leq \text{Oil}_t^{iks}\, GOR^{iks}$, as an other oil&gas field if $\text{Oil}_t^{iks} > 0$ and $\text{TotGas}_t^{iks} > \text{Oil}_t^{iks}\, GOR^{iks}$ and as a gas-only field if $\text{Oil}_t^{iks} = 0$. Moreover, we can show that the identification of a field's category is not affected by the introduction of the tax reform.

Technical Assumptions. Each field is endowed with an initial condition $\text{M}_0^{iks}, \text{K}_0^{iks}, \text{L}_{-1}^{iks}$. We impose a lower bound on ReInS_t^{iks} such that $\text{ReInS}_t^{iks} \geq RI_t^{iks}$ where for technical reasons we allow for RI_t^{iks} to be negative and arbitrarily large in magnitude. This restriction is mostly innocuous because we impose conditions on $TF_{4t}^{iks}\left(\text{PInS}_t^{iks}; \text{ReInS}_t^{iks}\right)$ that ensures that $\text{ReInS}_t^{iks} \geq 0$ at all optimal choices. Specifically, we impose that all oil fields and other oil&gas fields satisfy $\frac{\partial TF_t^{iks}(\cdot)}{\partial \text{ReInS}_t^{iks}} \Big/ \frac{\partial TF_t^{iks}(\cdot)}{\partial \text{Z}_t^{iks}} \geq \max\left\{P_t^{s,\text{Gas}} - \tau_t^{iks,\text{Gas}} - MC_t^{iks,\text{Gas}}, 0\right\}$ for all feasible $\text{Oil}_t^{iks}, \text{Gas}_t^{iks}, \text{Flare}_t^{iks}, \text{PInS}_t^{iks}, \text{M}_t^{iks}, \text{Z}_t^{iks}$ at any value $\text{ReInS}_t^{iks} < 0$. For gas fields, we allow for occasional negative values for ReInS_t^{iks}, although negative values do not generally occur at the optimal choice. This assumption is rather strong, but it is used solely for two specific purposes: (1) in section 1.1.3 to ensure that the firm's optimization problem satisfies the Slater's condition; and (2) in the proof to Proposition 3 Part (vi) to show that the tax scheme proposed increases consumer's welfare.

Oil&gas Firm's Problem

Each oil&gas field i owned by firm k generates revenues Rev_t^{iks} from selling net outputs $\left(\text{Oil}_t^{iks}, \text{Gas}_t^{iks}, \text{Flare}_t^{iks}, \text{PInS}_t^{iks}, \text{Z}_t^{iks}\right)$ at net prices $\mathbf{np}_t^{iks} = \left(P_t^{iks,\text{Oil}} - \tau_t^{iks,\text{Oil}}, P_t^{s,\text{Gas}} - \tau_t^{iks,\text{Gas}}, -\tau_t^{iks,\text{Flare}}, -PP_t^{iks,\text{Gas}}, 1\right)$ expressed in $\$/BOE$ for the first four arguments. The value of the last argument equals one because it corresponds to the normalized price of the numéraire. Under the assumptions stated in the previous paragraphs, the gross revenue from field i has formula:

$$\begin{aligned} Rev_t^{iks}(\cdot) = \quad & \left(P_t^{iks,\text{Oil}} - \tau_t^{iks,\text{Oil}}\right)\text{Oil}_t^{iks} + \left(P_t^{s,\text{Gas}} - \tau_t^{iks,\text{Gas}}\right)\text{Gas}_t^{iks} \\ & -\tau_t^{iks,\text{Flare}}\text{Flare}_t^{iks} - PP_t^{iks,\text{Gas}}\text{PInS}_t^{iks} - \text{IM}_t^{iks} + \text{Z}_t^{iks} \end{aligned} . \tag{6.17}$$

Such that the intra-temporal profits $\Pi_t^{iks}\left(\text{Oil}_t^{iks}, \text{Gas}_t^{iks}, \text{Flare}_t^{iks}, \text{ReInj}_t^{iks}, \text{PInj}_t^{iks}, \text{PInS}_t^{iks}, \text{Z}_t^{iks}, \text{ID}_t^{iks}, \text{M}_t^{iks}, \text{IM}_t^{iks}, \text{OInj}_t^{iks}, \text{IVent}_t^{iks}\right)$ generated by field i have formula:

$$\begin{aligned} \Pi_t^{iks}(\cdot) = \quad & \left(1 - T_t^{ks}\right)\Big[Rev_t^{iks}\left(\text{Oil}_t^{iks}, \text{Gas}_t^{iks}, \text{Flare}_t^{iks}, \text{PInj}_t^{iks}, \text{PInS}_t^{iks}, \text{Z}_t^{iks}\right) \\ & -InvCost_t^{iks}\left(\text{ID}_t^{iks}, \text{IM}_t^{iks}, \text{ReInj}_t^{iks}, \text{PInj}_t^{iks}, \text{OInj}_t^{iks}; PP_t^{iks,\text{Gas}}\right) \\ & +Deduct_t^{ikt}\left(\text{ReInj}_t^{iks}, \text{ReInS}_t^{iks}\right)\Big] - RegCost_t^{iks}\left(\text{IVent}_t^{iks}, \text{Flare}_t^{iks}\right) \end{aligned} \tag{6.18}$$

Therefore, after substituting (6.11) and (6.8) and the formula for $\text{TotInj}_{t-1}^{iks}$ into the inequalities (6.3), (6.6), and (6.7) and using such inequalities plus the inequality in (6.1) as constraints to the firm's choice, we can construct the firm k's profit maximization problem in period 1, which writes:

$$
\max_{\substack{\left\{\mathrm{Oil}_t^{iks},\mathrm{Gas}_t^{iks},\mathrm{Flare}_t^{iks},\mathrm{IVent}_t^{iks},\\ \mathrm{ID}_t^{iks},\mathrm{ReInj}_t^{iks},\mathrm{OInj}_t^{iks}\mathrm{ReInS}_t^{iks},\mathrm{PInj}_t^{iks},\\ \mathrm{PInS}_t^{iks},\mathrm{M}_{t+1}^{iks},\mathrm{IM}_t^{iks},\mathrm{L}_t^{iks},\mathrm{Z}_t^{iks},\mathrm{K}_{t+1}^{iks}\right\}_{i=1}^{I_k}\in X_u^s}} \sum_{t=1}^{\infty}\sum_{k=1}^{K^s}\sum_{i=1}^{I^{ks}}\beta^{t-1}\,\Pi_t^{iks}(\cdot)
$$

$$
s.t.\ \left\{\begin{array}{c}
TF_t^{iks}\left(\mathrm{Oil}_t^{iks},\mathrm{Gas}_t^{iks},\mathrm{Flare}_t^{iks},\mathrm{ReInS}_t^{iks},\mathrm{PInS}_t^{iks},\mathrm{M}_t^{iks},\mathrm{Z}_t^{iks}\right)\le 0\\[2ex]
GOR^{iks}\mathrm{Oil}_t^{iks}-\left(1-\vartheta^{iks}\right)^{-1}\Big[\mathrm{Gas}_t^{iks}+\mathrm{Flare}_t^{iks}+\mathrm{IVent}_t^{iks}\\
-Maint_t^{iks}\left(\mathrm{M}_t^{iks}\right)+\mathrm{ReInj}_t^{iks}+\mathrm{ReInS}_t^{iks}-\mathrm{K}_t^{iks}\Big]\le 0\\[2ex]
\mathrm{Oil}_t^{iks}+\left(1-\vartheta^{iks}\right)^{-1}\Big[\mathrm{Gas}_t^{iks}+\mathrm{Flare}_t^{iks}+\mathrm{IVent}_t^{iks}\\
-Maint_t^{iks}\left(\mathrm{M}_t^{iks}\right)+\mathrm{ReInj}_t^{iks}+\mathrm{ReInS}_t^{iks}-\mathrm{K}_t^{iks}\Big]\le 0\\[2ex]
\mathrm{K}_{t+1}^{iks}-D_t^{iks}\left(\mathrm{ID}_t^{iks},\mathrm{L}_{t-1}^{iks}\right)-B_t^{iks}\left(\mathrm{ReInj}_t^{iks}+\mathrm{PInj}_t^{iks}+\mathrm{OInj}_t^{iks}\right)\\
-\mathrm{K}_t^{iks}+\zeta\Big\{\mathrm{Oil}_t^{iks}+\left(1-\vartheta^{iks}\right)^{-1}\Big[\mathrm{Gas}_t^{iks}+\mathrm{Flare}_t^{iks}+\mathrm{IVent}_t^{iks}+\\
-Maint_t^{iks}\left(\mathrm{M}_t^{iks}\right)+\mathrm{ReInj}_t^{iks}+\mathrm{ReInS}_t^{iks}\Big]\Big\}\le 0\\[2ex]
\mathrm{Oil}_t^{iks}\ge 0,\mathrm{Flare}_t^{iks}-\mathrm{NRF}_t^{iks}\ge 0,\mathrm{IVent}_t^{iks}\ge 0,\mathrm{ReInj}_t^{iks}\ge 0,\\
\mathrm{ReInS}_t^{iks}-RI_t^{iks}\ge 0,\mathrm{M}_{t+1}^{iks}\ge 0,\mathrm{L}_t^{iks}\ge 0,\mathrm{K}_{t+1}^{iks}\ge 0\\[2ex]
\mathrm{L}_t^{iks}-\mathrm{L}_{t-1}^{iks}-\mathrm{ID}_t^{iks}=0\\[2ex]
\mathrm{M}_{t+1}^{iks}-\mathrm{M}_t^{iks}\left(1-\rho^{iks}\right)-\mathrm{IM}_t^{iks}=0
\end{array}\right\}\ \begin{array}{l} t=1,2,...,\\ i=1,...,I_k\end{array}
\tag{6.19}
$$

where $X_u^s=\left\{\left\{X_{ut}^{iks}\right\}_{i=1}^{I^{ks}}\right\}_{t=1}^{\infty}$ with $X_{ut}^{iks}=(-\infty,+\infty)$[15]. That is, the firm k solves a constrained maximization problem with 14 inequality constraints and 2 linear equality constraints. Note that we are not constraining Gas_t^{iks}, ID_t^{iks} and IM_t^{iks} to be positive. This captures the fact that, in principle, oil field may purchase natural gas for injection or in-situ use. However, the extent of which Gas_t^{iks} can be negative is limited by the other constraints, in particular $GOR^{iks}\mathrm{Oil}_t^{iks}-\mathrm{TotGas}_t^{iks}\le 0$ and $\mathrm{Oil}_t^{iks}\ge 0$. Similarly, negative investment is allowed in our framework, but the extent of negative investment in maintenance and/or discoveries is bounded by the constraints on the values of L_t^{iks} and M_{t+1}^{iks}. Given these assumptions, the Lagrangian of the firm's problem writes:

$$
\begin{aligned}
\mathcal{L}_u^{ks} = \sum_{t=1}^{\infty}\sum_{i=1}^{I^{ks}} \Big\{ \Big[& \beta^{t-1}\Pi_t^{iks}\Big(\mathrm{Oil}_t^{iks}, \mathrm{Gas}_t^{iks}, \mathrm{Flare}_t^{iks}, \mathrm{ReInj}_t^{iks}, \\
& \mathrm{PInj}_t^{iks}, \mathrm{PInS}_t^{iks}, \mathrm{Z}_t^{iks}, \mathrm{ID}_t^{iks}, \mathrm{M}_t^{iks}, \mathrm{IM}_t^{iks}, \mathrm{OInj}_t^{iks}, \mathrm{IVent}_t^{iks}\Big)\Big] + \\
- \phi_{1t}^{iks} & \Big[TF_t^{iks}\Big(\mathrm{Oil}_t^{iks}, \mathrm{Gas}_t^{iks}, \mathrm{Flare}_t^{iks}, \mathrm{ReInS}_t^{iks}, \mathrm{PInS}_t^{iks}, \mathrm{M}_t^{iks}, \mathrm{Z}_t^{iks}\Big)\Big] + \\
- \phi_{2t}^{iks} & \Big\{GOR^{iks}\mathrm{Oil}_t^{iks} - \left(1-\vartheta^{iks}\right)^{-1}\Big[\mathrm{Gas}_t^{iks} + \mathrm{Flare}_t^{iks} + \mathrm{IVent}_t^{iks} + \\
& -Maint_t^{iks}\left(\mathrm{M}_t^{i}\right) + \mathrm{ReInj}_t^{iks} + \mathrm{ReInS}_t^{iks}\Big]\Big\} \\
- \phi_{3t}^{iks} & \Big\{\mathrm{Oil}_t^{iks} + \left(1-\vartheta^{iks}\right)^{-1}\Big[\mathrm{Gas}_t^{iks} + \mathrm{Flare}_t^{iks} + \mathrm{IVent}_t^{iks} + \\
& -Maint_t^{iks}\left(\mathrm{M}_t^{i}\right) + \mathrm{ReInj}_t^{iks} + \mathrm{ReInS}_t^{iks}\Big] - \mathrm{K}_t^{iks}\Big\} + \\
- \phi_{4t}^{i} & \Big\{\mathrm{K}_{t+1}^{iks} - \mathrm{K}_t^{iks} - D_t^{iks}\left(\mathrm{ID}_t^{iks}, \mathrm{L}_{t-1}^{iks}\right) - B_t^{iks}\Big(\mathrm{ReInj}_t^{iks} + \mathrm{PInj}_t^{iks} + \mathrm{OInj}_t^{iks}\Big) + \\
& + \zeta\Big[\mathrm{Oil}_t^{iks} + \left(1-\vartheta^{iks}\right)^{-1}\Big[\mathrm{Gas}_t^{iks} + \mathrm{Flare}_t^{iks} + \mathrm{IVent}_t^{iks} + \\
& -Maint_t^{iks}\left(\mathrm{M}_t^{i}\right) + \mathrm{ReInj}_t^{iks} + \mathrm{ReInS}_t^{iks}\Big]\Big\} \\
+ \phi_{5t}^{iks} & \mathrm{Oil}_t^{iks} + \phi_{6t}^{iks}\Big(\mathrm{Flare}_t^{iks} - \mathrm{NRF}_t^{iks}\Big) + \phi_{7t}^{iks}\mathrm{IVent}_t^{i} + \phi_{8t}^{iks}\mathrm{ReInj}_t^{iks} \\
+ \phi_{9t}^{iks} & \Big(\mathrm{ReInS}_t^{iks} - RI_t^{iks}\Big) + \phi_{10t}^{iks}\mathrm{M}_{t+1}^{iks} + \phi_{11t}^{iks}\mathrm{L}_t^{i} + \phi_{12t}^{iks}\mathrm{K}_{t+1}^{iks} \\
- \lambda_{1t}^{iks} & \left[\mathrm{L}_t^{iks} - \mathrm{L}_{t-1}^{iks} - \mathrm{ID}_t^{iks}\right] - \lambda_{2t}^{iks}\left[\mathrm{M}_{t+1}^{iks} - \mathrm{M}_t^{iks}\left(1-\rho^{iks}\right) - \mathrm{IM}_t^{iks}\right] \Big\}
\end{aligned}
$$

For ease of notation, we define the vector

$$
\mathbf{x}_t^{iks} = \Big(\mathrm{Oil}_t^{iks}, \mathrm{Gas}_t^{iks}, \mathrm{Flare}_t^{iks}, \mathrm{ReInj}_t^{iks}, \mathrm{PInj}_t^{iks}, \mathrm{PInS}_t^{iks}, \mathrm{Z}_t^{iks}, \mathrm{ID}_t^{iks}, \mathrm{M}_t^{iks}, \mathrm{IM}_t^{iks}, \mathrm{OInj}_t^{iks}, \mathrm{IVent}_t^{iks}\Big) \quad (6.20)
$$

with $\mathbf{x}_t^{iks} \in X_{ut}^{iks}$ and the corresponding profit function $\mathit{\Pi}_t^{iks}\left(\mathbf{x}_t^{iks}; \mathbf{T}\right)$, which we use in the proofs in section 1.1.3, where $\mathbf{T}$ denotes the tax scheme as defined in section 1.4, and the corresponding collection $\mathbf{x}^{iks} = \left\{\mathbf{x}_t^{iks}\right\}_{t=1}^{\infty}$. Given this newly defined notation, we restate the firm's problem in the following parsimonious form.

$$
\begin{aligned}
\max_{\mathbf{x}_t^{iks} \in X_u^s} \quad & \textstyle\sum_{t=1}^{\infty}\sum_{k=1}^{K^s}\sum_{i=1}^{I^{ks}} \beta^{t-1}\, \mathit{\Pi}_t^{iks}\left(\mathbf{x}_t^{iks}; \mathbf{T}\right) \\
s.t. \quad & \left\{ \begin{array}{c} \left\{g_{w,ut}^{iks}\left(\mathbf{x}_t^{iks}\right) \le 0\right\}_{w=1}^{14} \\ \\ \left\{e_{z,ut}^{iks}\left(\mathbf{x}_t^{iks}\right) \le 0\right\}_{z=1}^{2} \end{array} \right\} \quad \begin{array}{l} t = 1, 2, ..., \\ i = 1, ..., I_k \end{array}
\end{aligned} \quad (6.21)
$$

where $g_{w,ut}^{iks}\left(\mathbf{x}_t^{iks}\right)$ corresponds to the wth inequality constraint and $e_{z,ut}^{iks}\left(\mathbf{x}_t^{iks}\right)$ to the zth equality constraint of the original firm's problem in (6.19).

Optimality Conditions

First, we establish that the firm's problem in (6.21) is a convex maximization problem. The objective function in (6.21) is a concave function given the assumptions on its functional form. The set X_u^s is convex. Moreover, each inequality constraint $g_{w,ut}^{iks}(x) \le 0$ is such that $g_{w,ut}^{iks}\left(\mathbf{x}^{iks}\right)$ is a weakly convex function, and all equality constraints $e_{z,ut}^{iks}\left(\mathbf{x}^{iks}\right) = 0$ are linear, implying in turn that each set $CS_{ut}^{iks} \equiv \left\{\mathbf{x}^{iks} \in X_u^s \mid \mathbf{g}_{ut}^{iks}\left(\mathbf{x}^{iks}\right) \le \mathbf{0}, \mathbf{e}_{ut}^{iks}\left(\mathbf{x}^{iks}\right) = \mathbf{0}\right\}$ for $i = 1, 2, ..., I^k$ and $t = 1, 2, ...$ is a convex set. Lastly, the set $\left[\bigcap_{i=1}^{I^k}\bigcap_{t=1}^{\infty} CS_{ut}^{iks}\right] \cap X_u^s$ is the intersection of convex sets, which is a convex set. Thus, the firm's problem is convex. Second, note that because the domain of the variables Slater's condition is satisfied. To prove this result we must prove that there exists a feasible choice vector that all the inequality constraints are satisfied with strict inequality. Consider the following choice vector. In each period t and for each field i, set $\text{Oil}_t^{iks}, \text{Gas}_t^{iks}, \text{Flare}_t^{iks}, \text{IVent}_t^{iks}, \text{ReInj}_t^{iks}, \text{M}_{t+1}^{iks}$ equal to an arbitrarily small strictly positive number ϵ_t^{iks}, $\text{L}_t^{iks} = \text{L}_{t-1}^{iks} > 0$, $\text{K}_{t+1}^{iks} = \text{K}_t^{iks} > 0$ $\text{ReInS}_t^{iks} \leq 0$ and set Oil_t^{iks} equal to a strictly positive number that satisfies $\text{Oil}_t^{iks} = \epsilon 2_t^{iks} < \frac{5\epsilon_t^{iks}}{\overline{G}}$, where $\overline{G}$ is the upper bound of GOR^{iks}. This ensures the proposed element satisfies seven constraints in (6.21), which correspond to constraints 2-7-8-9-10-11-12 in (6.19), with strict inequality. Then select $\text{PInj}_t^{iks} = \text{OInj}_t^{iks} = \epsilon_t^{iks}$, and $\text{ReInS}_t^{iks} = -\max\left\{\epsilon 2_t^{iks} + 4\epsilon_t^{iks} + UVent_t^i\left(\epsilon 2_t^{iks}, \epsilon_t^{iks}\right), -D_t^i\left(\epsilon_t^{iks}, \text{L}_0^{iks}\right)\right.$ $\left.-B_t^{iks}\left(\epsilon_t^{iks}\right) + \zeta\left[\epsilon 2_t^{iks} + 4\epsilon_t^{iks} + UVent_t^{iks}\left(\epsilon 2_t^{iks}, \epsilon_t^{iks}\right)\right]\right\} - \epsilon_t^{iks}$ (i.e., just sufficiently large in magnitude to ensure that constraints 3 and 4 are satisfied with strict inequality). Lastly, set Z_t^{iks} sufficiently small such that $TF_t^{iks}(\text{Oil}_t^{iks}, \epsilon_t^{iks}, \epsilon_t^{iks}, \text{ReInS}_t^{iks}, 0, \epsilon_t^{iks}, \text{Z}_t^{iks}) < 0$, for which, given that TF_t^{iks} is weakly decreasing in ReInS_t^{iks} for negative values of ReInS_t^{iks} (see the technical assumptions section), it is sufficient that $TF_t^{iks}(\epsilon_t^{iks}, \epsilon_t^{iks}, \epsilon_t^{iks}, 0, 0, \epsilon_t^{iks}, \text{Z}$ 0. Thus, because ϵ_t^{iks} is arbitrarily small and TF_t^{iks} has finite first derivatives and strictly positive first derivative w.r.t. Z_t^{iks}, there exists $\text{Z}_t^{iks} < 0$ such that all the inequalities are satisfied with strict inequality. Thus, the Slater's condition is satisfied and the Karush–Kuhn–Tucker theorem implies the global maximizer of the constrained optimization problem (if it exists) has to satisfy the KKT conditions (First-order necessary conditions - FOCs). Let ς_t^{Oil} ($\varsigma_t^{s,\text{Gas}}$) denote the elasticity of oil (natural gas) demand, and $MS_t^{ks,\text{Oil}}$ ($MS_t^{ks,\text{Gas}}$) be the share of the oil (natural gas) market that is controlled by firm k. We derive the FOCs for a global maximum, which for each $t = 1, 2, ...$ and each $i = 1, 2, .., I^k$ write:

$$\frac{\partial \mathcal{L}_u^{ks}}{\partial \text{Oil}_t^{iks}} = \left[P_t^{\text{Oil}}\left(1-\varsigma_t^{\text{Oil}} MS_t^{ks,\text{Oil}}\right)+\sigma_1^{iks}-\tau_t^{iks,\text{Oil}}\right](1-T_t^s)-\phi_{1t}^{iks}\frac{\partial TF_{1t}^{iks}(.)}{\partial \text{Oil}_t^{iks}}$$
$$-\phi_{2t}^{iks} GOR^{iks}-\phi_{3t}^{iks}-\phi_{4t}^{iks}\zeta+\phi_{5t}^{iks}=0$$
$$\frac{\partial \mathcal{L}_u^{ks}}{\partial \text{Gas}_t^{iks}} = \left[P_t^{s,\text{Gas}}\left(1-\varsigma_t^{s,\text{Gas}} MS_t^{ks,\text{Gas}}\mathbf{1}\left[\text{Oil}_t^{isk}=0\right]\right)-\tau_t^{isk,\text{Gas}}\right](1-T_t^s)$$
$$-\phi_{1t}^{iks}\frac{\partial TF_{1t}^{iks}(.)}{\partial \text{Gas}_t^{iks}}+\left(\phi_{2t}^{iks}-\phi_{3t}^{iks}-\phi_{4t}^{iks}\zeta\right)\left(1-\vartheta^{iks}\right)^{-1}=0$$
$$\frac{\partial \mathcal{L}_u^{ks}}{\partial \text{Flare}_t^{iks}} = -\tau_t^{iks,\text{Flare}}(1-T_t^s)-VF_t^{iks}\frac{PF_t^{iks}(\cdot)}{\partial \text{Flare}_t^{iks}}-\phi_{1t}^{iks}\frac{\partial TF_{3t}^{iks}(.)}{\partial \text{Flare}_t^{iks}}$$
$$+\left(\phi_{2t}^{iks}-\phi_{3t}^{iks}-\phi_{4t}^{iks}\zeta\right)\left(1-\vartheta^{iks}\right)^{-1}+\phi_{6t}^{iks}=0$$
$$\frac{\partial \mathcal{L}_u^{ks}}{\partial \text{IVent}_t^{iks}} = -VF_t^{iks}\frac{PF_t^{iks}(\cdot)}{\partial \text{IVent}_t^{i}}+\left(\phi_{2t}^{iks}-\phi_{3t}^{iks}-\phi_{4t}^{iks}\zeta\right)\left(1-\vartheta^{iks}\right)^{-1}+\phi_{9t}^{iks}=0$$
$$\frac{\partial \mathcal{L}_u^{ks}}{\partial \text{PInj}_t^{iks}} = -\frac{IC_t^{iks}\left(\text{ReInj}_t^{iks},\text{PInj}_t^{iks}\right)}{\partial \text{PInj}_t^{iks}}(1-T_t^s)-PP_t^{iks,\text{Gas}}(1-T_t^s)+\phi_{4t}^{iks}\frac{\partial B_t^{iks}(\cdot)}{\partial \text{PInj}_t^{iks}}=0$$
$$\frac{\partial \mathcal{L}_u^{ks}}{\partial \text{ReInj}_t^{iks}} = \left[-\frac{IC_t^{iks}\left(\text{ReInj}_t^{iks},\text{PInj}_t^{iks}\right)}{\partial \text{ReInj}_t^{iks}}+\delta_{1t}^{iks}\right](1-T_t^s)$$
$$+\left(\phi_{2t}^{iks}-\phi_{3t}^{iks}-\phi_{4t}^{iks}\zeta\right)\left(1-\vartheta^{iks}\right)^{-1}+\phi_{4t}^{iks}\frac{\partial B_t^{iks}(\cdot)}{\partial \text{ReInj}_t^{iks}}+\phi_{8t}^{iks}=0$$
$$\frac{\partial \mathcal{L}_u^{ks}}{\partial \text{OInj}_t^{iks}} = -C_t^{iks}(1-T_t^s)+\phi_{4t}^{iks}\frac{\partial B_t^{iks}(\cdot)}{\partial \text{OInj}_t^{iks}}=0$$
$$\frac{\partial \mathcal{L}_u^{ks}}{\partial \text{ReInS}_t^{iks}} = -\phi_{1t}^{iks}\frac{\partial TF_{4t}^{iks}(.)}{\partial \text{ReInS}_t^{iks}}+\delta_{1t}^{iks}+\left(\phi_{2t}^{iks}-\phi_{3t}^{iks}-\phi_{4t}^{iks}\zeta\right)\left(1-\vartheta^{iks}\right)^{-1}+\phi_{9t}^{iks}=0$$
$$\frac{\partial \mathcal{L}_u^{ks}}{\partial \text{PInS}_t^{iks}} = -PP_t^{iks,\text{Gas}}(1-T_t^s)+\phi_{1t}^{iks}\frac{\partial TF_{4t}^{iks}(.)}{\partial \text{PInS}_t^{iks}}=0$$
$$\frac{\partial \mathcal{L}_u^{ks}}{\partial \text{Z}_t^{iks}} = (1-T_t^s)-\phi_{1t}^{iks}\frac{\partial TF_t^{iks}(.)}{\partial \text{Z}_t^{iks}}=0$$
$$\frac{\partial \mathcal{L}_u^{ks}}{\partial \text{K}_{t+1}^{iks}} = -\phi_{4t}^{iks}+\beta\phi_{3t+1}^{iks}+\beta\phi_{4t+1}^{iks}+\phi_{11t}^{iks}=0$$
$$\frac{\partial \mathcal{L}_u^{ks}}{\partial \text{ID}_t^{iks}} = -(1-T_t^s)+\phi_{4t}^{iks}\frac{\partial D_t^{iks}\left(\text{ID}_t^{iks},\text{L}_{t-1}^{iks}\right)}{\partial \text{ID}_t^{i}}-\lambda_{1t}^{iks}=0$$
$$\frac{\partial \mathcal{L}_u^{ks}}{\partial \text{L}_t^{iks}} = \lambda_{1t}^{iks}+\beta\phi_{4t+1}^{iks}\frac{\partial D_{t+1}^{iks}\left(\text{ID}_{t+1}^{iks},\text{L}_t^{iks}\right)}{\partial \text{L}_t^{iks}}-\beta\lambda_{1t+1}^{iks}=0$$
$$\frac{\partial \mathcal{L}_u^{ks}}{\partial \text{IM}_t^{iks}} = -(1-T_t^s)+\lambda_{2t}^{iks}=0$$
$$\frac{\partial \mathcal{L}_u^{ks}}{\partial \text{M}_{t+1}^{iks}} = -\beta\phi_{1t+1}^{iks}\frac{\partial TF_{5t+1}^{iks}(.)}{\partial \text{M}_{t+1}^{iks}}-\lambda_{2t}^{iks}+\beta\lambda_{2t+1}^{iks}\left(1-\rho^{iks}\right)+\phi_{10t}^{iks}-\beta VF_t^{iks}\frac{PF_{t+1}^{iks}(\cdot)}{\partial \text{M}_{t+1}^{iks}}$$
$$-\beta\left(1-\vartheta^{iks}\right)^{-1}\left[\phi_{2t+1}^{iks}-\phi_{3t+1}^{iks}-\phi_{4t+1}^{iks}\zeta\right]\frac{\partial Maint_{t+1}^{iks}(.)}{\partial \text{M}_{t+1}^{iks}}=0$$

plus the standard primal feasibility conditions $\left\{\left\{\left\{g_{w,ut}^{iks}(\cdot)\le 0\right\}_{w=1}^{12},\left\{e_{z,ut}^{iks}(\cdot)=0\right\}_{z=1}^{2}\right\}_{i=1}^{I^k}\right\}_{t=1}^{\infty}$, the dual feasibility conditions $\left\{\left\{\left\{\phi_{wt}^{iks}\ge 0\right\}_{w=1}^{12}\right\}_{i=1}^{I^k}\right\}_{t=1}^{\infty}$, and the complimentary slackness conditions $\left\{\left\{\left\{\phi_{wt}^{iks}g_{w,ut}^{iks}(\cdot)=0\right\}_{w=1}^{12}\right\}_{i=1}^{I^k}\right\}$

6.1.2 Midstream: Refinery&Transformation Firms

We assume there are J^s price-taker and infinitely-living midstream firms in each country s and each period t. This category includes oil refineries, gas-processing facilities, and transformation firms which use crude O_t^{js} and natural gas G_t^{js} (measured in BOE) and other commodities with US dollar value MZ_t^{js} as inputs (negative net outputs) and produce a B-dimensional vector of oil&gas products $\mathbf{y}_t^{js}$ (positive net outputs) and a weakly positive amount of flaring F_t^{js}. Transformation firms include all type of firms that use crude oil and/or unrefined natural gas as inputs, such as power plants and petrolchemical firms. Oil&gas products include, among other, gasoline, heavy fuels, LPG, natural gas, and plastic materials. Lastly, midstream firms may recover some of the natural gas produced as a byproduct of crude processing and use it for electricity production in-situ in quantity MInS_t^{js}.

Technological Constraints. The firm's technology is represented by the real analytic transformation function $MTF_t^{js} : X_t^m \to \mathbb{R}$ where $X_t^m = (-\infty, +\infty)^{B+5}$. As a consequence, the production set of the R&F firm is described by the inequality:

$$MTF_t^j \left(\mathbf{y}_t^j, \mathrm{O}_t^j, \mathrm{G}_t^j, \mathrm{F}_t^j, \mathrm{MInS}_t^j, \mathrm{MZ}_t^j \right) \leq 0 \tag{6.22}$$

For tractability, we assume that MTF_t^{js} possess the additively separable form:

$$\begin{aligned} MTF_t^{js} \left(\mathbf{y}_t^{js}, \mathrm{O}_t^{js}, \mathrm{G}_t^{js}, \mathrm{F}_t^{js}, \mathrm{MInS}_t^{js}, \mathrm{MZ}_t^{js} \right) = \quad & MTF_{1t}^j \left(\mathbf{y}_t^j \right) + MTF_{2t}^j \left(\mathrm{O}_t^j \right) \\ & + MTF_{3t}^j \left(\mathrm{G}_t^j \right) + MTF_{4t}^j \left(\mathrm{F}_t^j \right) \\ & + MTF_{5t}^j \left(\mathrm{MInS}_t^j \right) + MTF_{6t}^j \left(\mathrm{MZ}_t^j \right) \end{aligned} \tag{6.23}$$

The technology of any midstream firm satisfies the constraints $\mathrm{O}_t^{js} \leq 0$ and $\mathrm{G}_t^{js} \leq 0$, $\mathrm{F}_t^{js} \geq RF_t^{js}$; i.e., crude and unrefined natural gas are net inputs for midstream firms, and some of them –specifically, oil refineries and gas processing facilities– must produce a certain amount of flaring (minimum non-routine flaring). First, oil refineries are defined as midstream firms such that $\frac{\partial MTF_2^{js}(\mathrm{O}_t^{js})}{\partial \mathrm{O}_t^{js}} > 0$ for all O_t^j, $\frac{\partial MTF_4^{js}(\mathrm{F}_t^{js})}{\partial \mathrm{F}_t^{js}} \neq +\infty$, and $\frac{\partial MTF_3^{js}(\mathrm{G}_t^{js})}{\partial \mathrm{G}_t^j} = 0$ for all G_t^j. Under these assumptions, the constraint $\mathrm{G}_t^{js} \leq 0$ is always binding, implying that oil refineries optimally only use crude oil and MZ_t^j as inputs and they may produce positive flaring and natural gas as by-products. Second, in a similar way, a gas processing facility features $\frac{\partial MTF_2^{js}(\mathrm{G}_t^{js})}{\partial \mathrm{G}_t^{js}} > 0$ for all G_t^{js}, $\frac{\partial MTF_4^{js}(\mathrm{F}_t^{js})}{\partial \mathrm{F}_t^{js}} \neq +\infty$, and $\frac{\partial MTF_3^{js}(\mathrm{O}_t^{js})}{\partial \mathrm{O}_t^{js}} = 0$ for all O_t^{js}. Third, midstream firms other than oil refineries and gas processing facilities ("transformation firms") feature $\frac{\partial MTF_4^{js}(\mathrm{F}_t^{js})}{\partial \mathrm{F}_t^{js}} = +\infty$ for all $\mathrm{F}_t^{js} \geq 0$ and $RF_t^{js} = 0$; i.e., they do not perform any gas flaring. Fourth, we do not restrict the elements of vector $\mathbf{y}_t^{js}$ to be weakly positive in order to capture the fact that some goods that are net output for some midstream firms may be net inputs for other firms in the same class. For instance, refined natural gas is a net output for gas processing facilities and a net input for gas-operated steel factories. Lastly, all midstream firms face the following constraint:

$$-\mathrm{MInS}_t^{js} - GOR^{js}\mathrm{O}_t^{js} - \mathrm{F}_t^{js} \leq 0$$

where GOR^{js} represents the amount of natural gas that is produced as byproduct of crude processing per unit of crude. This constraint captures the fact that such byproduct gas can be either flared or used in-situ for the production of electricity. Note that this constraint is not binding for gas processing facilities that feature $\mathrm{O}_t^{js} = 0$ at the optimal choice, and for other midstream firms that are characterized by $GOR^{js} = 0$; i.e., they do not produce any natural gas as byproduct.

Information. In terms of information structure, we assume that midstream firms operate under full information. Differently from upstream firms, however, the information set of midstream firms corresponds to the public one; i.e., $\Omega_t^j = \Omega_t^{PUB}$ for all j. This assumption captures the fact that midstream firms do not engage in venting of natural gas of any type because of safety concerns, therefore all their endogenous choices are observable and contractible for the government.

Flaring & Venting Regulation. Midstream firms (including oil refineries and gas processing facilities) cannot vent any amount of natural gas because of safety concerns. For the same reason, they are also typically allowed to flare gas resulting from their operations. Therefore, assume no venting and/or flaring regulation applies to such firms.

Fiscal Framework. We assume a tax system on midstream firms featuring: (i) a linear tax rate T_t^s on corporate income, defined as the firm's revenue minus total costs;(ii) a (possibly firm-specific) specific linear tax on flaring, which also includes the unitary cost of any flaring permits the firm may be required to purchase, denoted by $\tau_t^{js,\mathrm{F}}$ (it may be equal to zero). (iii) a vector of (possibly field-specific) specific linear sales taxes $\mathbf{a}_t^{js}$ on oil&gas products; (iv) a (possibly field-specific) specific linear sales tax b_t^{js} on natural gas. Note that taxes (iii) and (iv) are defined on net supplies. Thus, if the corresponding net outputs are negatives, their values should be interpreted as subsidy rates rather than tax rates.

Midstream Firm's Problem

We define $X_m = \{X_{mt}\}_{t=1}^{\infty}$. The within-period profits of a midstream firm write:

$$\Pi_t^{js}(\cdot) = (1 - T^s)\left[\left(\mathbf{p}_t^s - \mathbf{a}_t^{js}\right)'\mathbf{y}_t^{js} + P_t^{\mathrm{Oil}}\mathrm{O}_t^{js} + \left(P_t^{s,\mathrm{Gas}} - b_t^{js}\right)\mathrm{G}_t^{js} + \mathrm{MZ}_t^{js} - \mathrm{F}_t^{js}\tau_t^{js,\mathrm{F}}\right] \tag{6.24}$$

Given the assumptions stated in the previous section, the problem of a midstream firm writes:

$$\begin{array}{cc}
\underset{\substack{\left\{\mathbf{y}_t^{js},\mathrm{O}_t^{js},\mathrm{G}_t^{js},\mathrm{F}_t^{js},\right.\\ \left.\mathrm{MInS}_t^{js},\mathrm{MZ}_t^{js}\right\}_{t=1}^{\infty}\in X_m}}{\max} & \sum_{t=1}^{\infty}\beta^{t-1}\Pi_t^{js}\left(\mathbf{y}_t^{js},\mathrm{O}_t^{js},\mathrm{G}_t^{js},\mathrm{F}_t^{js},\mathrm{MInS}_t^{js},\mathrm{MZ}_t^{js}\right) \\
 & MTF_t^{js}\left(\mathbf{y}_t^{js},\mathrm{O}_t^{js},\mathrm{G}_t^{js},\mathrm{F}_t^{js},\mathrm{MInS}_t^{js},\mathrm{MZ}_t^{js}\right)\leq 0 \\
s.t. & -\mathrm{MInS}_t^{js} - GOR^{js}\mathrm{O}_t^{js} - \mathrm{F}_t^{js}\leq 0 \\
 & \mathrm{O}_t^{js}\leq 0,\ \mathrm{G}_t^{js}\leq 0,\ \mathrm{F}_t^{js} - RF_t^{js}\geq 0
\end{array} \tag{6.25}$$

The Lagrangian of this problem writes:

$$\begin{aligned}
\mathcal{L}_m^{js} = \quad & \sum_{t=1}^{\infty}\beta^{t-1}\left\{(1 - T_t^s)\left[\left(\mathbf{p}_t^s - \mathbf{a}_t^{js}\right)'\mathbf{y}_t^{js} + P_t^{\mathrm{Oil}}\mathrm{O}_t^{js} + \left(P_t^{s,\mathrm{Gas}} - b_t^{js}\right)\mathrm{G}_t^{js} - \mathrm{F}_t^{js}\tau_t^{js,\mathrm{F}} + \mathrm{MZ}_t^{js}\right]\right. \\
& -\psi_{1t}^{js}\left[MTF_t^{js}\left(\mathbf{y}_t^{js},\mathrm{O}_t^{js},\mathrm{G}_t^{js},\mathrm{F}_t^{js},\mathrm{MInS}_t^{js},\mathrm{MZ}_t^{js}\right)\right] \\
& -\psi_{2t}^{js}\left[-\mathrm{MInS}_t^{js} - GOR^{js}\mathrm{O}_t^{js} - \mathrm{F}_t^{js}\right] \\
& -\psi_{3t}^{js}\mathrm{O}_t^{js} - \psi_{4t}^{j}\mathrm{G}_t^{js} + \psi_{5t}^{j}\left[\mathrm{F}_t^{js} - RF_t^{js}\right]
\end{aligned} \tag{6.26}$$

For ease of notation, we define the vector

$$\mathbf{z}_t^{js} = \left(\mathrm{y}_{1t}^{js},\mathrm{y}_{2t}^{js},...,\mathrm{y}_{Bt}^{js},\mathrm{O}_t^{js},\mathrm{G}_t^{js},\mathrm{F}_t^{js},\mathrm{MInS}_t^{js},\mathrm{MZ}_t^{js}\right) \tag{6.27}$$

with $\mathbf{z}_t^{js}\in X_{mt}$ and the corresponding profit function $\mathit{\Pi}_t^{js}\left(\mathbf{z}_t^{js};\mathbf{T}\right)$, which we use in the proofs in section 1.1.3, where $\mathbf{T}$ denotes the tax scheme as defined in section 1.4, and the corresponding collection $\mathbf{z}^{js} = \left\{\mathbf{z}_t^{js}\right\}_{t=1}^{\infty}$. Given this newly defined notation, we restate the firm's problem in the following parsimonious form.

$$
\begin{aligned}
\max_{\mathbf{z}_t^{js} \in X_m} \quad & \sum_{t=1}^{\infty} \beta^{t-1} \, \Pi_t^{js}\left(\mathbf{z}_t^{js}; \mathbf{T}\right) \\
s.t. \quad & \left\{ \left\{ g_{w,mt}^{js}\left(\mathbf{z}_t^{js}\right) \leq 0 \right\}_{w=1}^{5} \right\}_{t=1,2,...,}
\end{aligned}
\tag{6.28}
$$

where $g_{w,mt}^{js}\left(\mathbf{z}_t^{js}\right)$ corresponds to the wth inequality constraint of the original firm's problem in (6.25).

Optimality Conditions

First, we establish that the firm's problem in (6.28) is a convex maximization problem. The objective function in (6.28) is a concave function because it is linear. The set X_m is convex. Moreover, each inequality constraint $g_{w,mt}^{js}\left(\mathbf{z}_t^{js}\right) \leq 0$ is such that $g_{w,mt}^{js}\left(\mathbf{z}_t^{js}\right)$ is a weakly convex function, implying in turn that each set $CS_{w,mt}^{js} \equiv \left\{\mathbf{z}_t^{js} \in X_{mt} \mid g_{w,mt}^{js}\left(\mathbf{z}_t^{js}\right) \leq 0\right\}$ for $w = 1,2,3,4,5$ and $t = 1,2,...$ is a convex set. Lastly, the set $\left[\bigcap_{w=1}^{5} \bigcap_{t=1}^{\infty} CS_{w,mt}^{js}\right] \cap X^m$ is the intersection of convex sets, which is a convex set. Thus, the firm's problem is convex. Second, note that the Slater's condition is satisfied. To prove this result, it is sufficient to choose an arbitrarily small strictly positive ε_t^{js} and set $\mathbf{y}_t^{js} = \mathbf{0}$, $\mathrm{O}_t^{js} = -\varepsilon_t^{js}$, $\mathrm{G}_t^{js} \leq -\varepsilon_t^{js}$, $\mathrm{F}_t^{js} = RF_t^{js} + \varepsilon_t^{js}$, $\mathrm{MInS}_t^{js} = GOR^{js}\varepsilon_t^{js} - RF_t^{js}$ and set MZ_t^{js} to a value small enough to ensure that the first constraint is satisfied with strict inequality, where such value exists given that $\frac{\partial MTF_{6t}^{js}\left(\mathrm{MZ}_t^{js}\right)}{\mathrm{MZ}_t^{js}} > 0$ for all values of MZ_t^{js}. Thus, the Slater's condition is satisfied and the Karush–Kuhn–Tucker theorem implies the global maximizer of the constrained optimization problem (if it exists) has to satisfy the KKT conditions (First-order necessary conditions - FOCs). Second, we derive the First-order Necessary Conditions for a global maximum, which write:

$$
\begin{aligned}
\frac{\partial \mathcal{L}_m^{js}}{\partial \mathrm{O}_t^{js}} &= P_t^{\mathrm{Oil}}\left(1 - T_t^s\right) - \psi_{1t}^{js}\frac{\partial MTF_t^{js}(\cdot)}{\partial \mathrm{O}_t^{js}} + \psi_{2t}^{js} GOR^{js} - \psi_{3t}^{js} = 0 \\
\frac{\partial \mathcal{L}_m^{js}}{\partial \mathrm{G}_t^{js}} &= \left(P_t^{s,\mathrm{Gas}} - b_t^{js}\right)\left(1 - T_t^s\right) - \psi_{1t}^{js}\frac{\partial MTF_t^{js}(\cdot)}{\partial \mathrm{G}_t^{js}} - \psi_{4t}^{js} = 0 \\
\frac{\partial \mathcal{L}_m^{js}}{\partial \mathrm{F}_t^{js}} &= -\tau_t^{js,\mathrm{F}}\left(1 - T_t^s\right) - \psi_{1t}^{js}\frac{\partial MTF_t^{js}(\cdot)}{\partial \mathrm{F}_t^{js}} + \psi_{2t}^{js} - \psi_{5t}^{js} = 0 \\
\frac{\partial \mathcal{L}_m^{j}}{\partial \mathrm{MInS}_t^{js}} &= -\psi_{1t}^{js}\frac{\partial MTF_t^{js}(\cdot)}{\partial \mathrm{MInS}_t^{js}} + \psi_{2t}^{js} = 0 \\
\frac{\partial \mathcal{L}_m^{js}}{\partial \mathrm{MZ}_t^{js}} &= \left(1 - T_t^{js}\right) - \psi_{1t}^{js}\frac{\partial MTF_t^{js}(\cdot)}{\partial \mathrm{MZ}_t^{js}} = 0 \\
\frac{\partial \mathcal{L}_m^{js}}{\partial \mathrm{y}_{1t}^{j}} &= \left(p_{bt}^s - a_{bt}^{js}\right)\left(1 - T_t^s\right) - \psi_{1t}^{js}\frac{\partial MTF_t^{js}(\cdot)}{\partial \mathrm{y}_{bt}^{js}} = 0 \qquad \forall t = 1,2,... \\
&\vdots \\
\frac{\partial \mathcal{L}_m^{js}}{\partial \mathrm{y}_{bt}^{js}} &= \left(p_{bt}^s - a_{bt}^{js}\right)\left(1 - T_t^s\right) - \psi_{1t}^{js}\frac{\partial MTF_t^{js}(\cdot)}{\partial \mathrm{y}_{bt}^{js}} = 0 \\
&\vdots \\
\frac{\partial \mathcal{L}_m^{js}}{\partial \mathrm{y}_{Bt}^{js}} &= \left(p_{bt}^s - a_{bt}^{js}\right)\left(1 - T_t^s\right) - \psi_{1t}^{js}\frac{\partial MTF_t^{js}(\cdot)}{\partial \mathrm{y}_{Bt}^{js}} = 0
\end{aligned}
\tag{6.29}
$$

plus the standard primal feasibility $\left\{\left\{g_{w,mt}^{js}\left(\mathbf{z}_t^{js}\right) \leq 0\right\}_{w=1}^{5}\right\}_{t=1}^{\infty}$, dual feasibility $\left\{\left\{\psi_{wt}^{js} \geq 0\right\}_{w=1}^{5}\right\}_{t=1}^{\infty}$, and complimentary slackness conditions $\left\{\left\{\psi_{wt}^{js} g_{w,mt}^{js}\left(\mathbf{z}_t^{js}\right) = 0\right\}_{w=1}^{5}\right\}_{t=1}^{\infty}$.

6.1.3 Downstream: Consumers

We assume a single infinitely-living and price-taker consumer in each country $s = 1,2,...,S$, who consumes the numéraire good C_t^s (other consumption) and a B-dimensional vector of oil&gas products (gasoline, natural gas, electricity, plastic materials, etc.), with typical element $\mathbf{c}_t^s$. Each element of vector $\mathbf{c}_t^s$ is expressed in an appropriate unit (e.g., gallons, KWh, Mt, etc.), whereas other consumption C_t^s is expressed in USD.

Preferences. In each period t the consumer has preferences over C_t^s, $\mathbf{c}_t^s$, and the concentration (stock) of greenhouse gases in the atmosphere expressed in CO_2-equivalent amounts. In detail, their within-period preferences of a consumer in country s are represented by the utility function:

$$U\left(C_t^s, \mathbf{c}_t^s, ExcTCO2e_t\right) = C_t^s + u^s\left(\mathbf{c}_t^s\right) - Ext \times ExcTCO2e_t \tag{6.30}$$

where u^s is continuous and strictly concave. Ext represents the marginal social cost of one CO_2-equivalent unit of greenhouse gases in the atmosphere (i.e., the price of CO_2) and $ExtTCO2e_t = TCO2e_t - \overline{TCO2e_t}$ is the difference between the actual concentration of GHG in the atmosphere in period t and a target value $\overline{TCO2e_t}$ for such variable (e.g., the concentration observed before the industrial revolution, $\sim$ 1750 AD). Thus, the term $Ext \times ExcTCO2e_t$ in formula (6.30) represents the disutility from excess GHG concentration in the atmosphere. Provided that the average global temperature is increasing in $TCO2e_t$, this term can be interpreted as a measure of the utility cost of Global Warming faced by the consumer in a given period t. The concentration of GHGs in the atmosphere is taken as given by the consumer, such that the amount of individual emissions which contribute to increase the value of $TCO2e_t$ (see equation (6.50) in the climate change section) is excluded from the consumer's problem. This assumption is equivalent to that of an economy with a large number of identical consumers, such that the consumption choices of each consumer have a negligible impact on global emissions.

Gross Income. In each period t the consumer in country s earns exogenous income e_t^s (e.g., labor income) and capital income. Regarding the former, we assume that it is large enough to ensure positive consumption. Regarding the latter, we assume that all firms in each country are owned by domestic consumers. This assumption implies that firms' net profits from each oil&gas firm $\sum_{k=1}^{K^s} \Pi_t^{iks}$ and each midstream firm $\sum_{j=1}^{J^s} \Pi_t^{js}$ are entering the consumer's income. Consumers take firms' profits as given because they do not manage the firms' production choices and are price-takers.

Fiscal Framework. In each period t the consumer in country s pays a net lump-sum income tax $ITax_t^s$. As a result, their disposable income in period t –denoted by Y_t^s– has formula:

$$Y_t^s = e_t^s + \left[\sum_{k=1}^{K^s} \Pi_t^{iks} + \sum_{j=1}^{J^s} \Pi_t^{js}\right] - ITax_t^s \tag{6.31}$$

However, the consumer does not exert control on neither firm nor government's decisions, meaning that Y_t^s is treated by the consumer as exogenous. Moreover, the consumer also faces a vector $\mathbf{v}_t^s$ of specific linear consumption taxes on oil&gas products. Thus, in each period t the consumer in country s faces a budget constraint:

$$C_t^s + \left(\mathbf{p}_t + \mathbf{v}_t^s\right)' \mathbf{c}_t^s - Y_t^s \leq 0 \tag{6.32}$$

Because consumer's disposable income Y_t^s is treated as given, the consumer's choice set reduces to $X_c = \{X_{ct}\}_{t=1}^{\infty}$ where $X_{ct} = \left\{\left(C_t^S, \mathbf{c}_t^s\right) \mid \left(C_t^S, \mathbf{c}_t^s\right) \in (0, +\infty) \times (0, +\infty)^B\right\}$. As a result, the consumer in country s solves the following problem:

$$\begin{aligned} \max_{\left\{\left(C_t^S, \mathbf{c}_t^s\right)\right\}_{t=1}^{\infty} \in X_c} \quad & \textstyle\sum_{t=1}^{\infty} \beta^{t-1}\left[C_t^s + u^s\left(\mathbf{c}_t^s\right) - Ext \times ExcTCO2e_t\right] \\ & s.t. \ \Big\{C_t^s + \left(\mathbf{p}_t^s + \mathbf{v}_t^s\right)' \mathbf{c}_t^s - e_t^s - Y_t^s \leq 0, \\ & \qquad \{c_{bt}^s \geq 0\}_{b=1}^{B}\Big\}_{t=1}^{\infty} \end{aligned} \tag{6.33}$$

The Lagrangian of consumer s's problem writes:

$$\begin{aligned}\mathcal{L}_c^s = \quad & \sum_{t=1}^{\infty} \beta^{t-1} \{C_t^s + u(\mathbf{c}_t^s) - Ext \times ExcTCO2e_t \\ & -\theta_{0t}^s \left[C_t^s + (\mathbf{p}_t^s + \mathbf{v}_t^s)' \mathbf{c}_t^s - Y_t^s\right] \\ & \sum_{b=1}^{B} \theta_{bt}^s \mathrm{c}_{bt}^s \Big\}\end{aligned} \tag{6.34}$$

Note that we are not constraining C_t^s to be weakly positive. However, the assumption that e_t^s is large ensures positive consumption in each period t.

For ease of notation, we define the vector

$$\mathbf{b}_t^s = \left(\mathbf{C}_t^s, \mathrm{c}_{1t}^s, \mathrm{c}_{2t}^{js}, ..., \mathrm{c}_{Dt}^s\right) \tag{6.35}$$

with $\mathbf{b}_t^{js} \in X_{ct}$ and the corresponding utility function $\mathbb{U}_t^s(\mathbf{b}_t^s; \mathbf{T}, ExcTCO2e_t)$, which we use in the proofs in section 1.1.3, where $\mathbf{T}$ denotes the tax scheme as defined in section 1.4, and the corresponding collection $\mathbf{b}^s = \{\mathbf{b}_t^s\}_{t=1}^{\infty}$. Given this newly defined notation, we restate the firm's problem in the following parsimonious form.

$$\begin{aligned}\max_{\mathbf{b}_t^s \in X_c} \quad & \sum_{t=1}^{\infty} \beta^{t-1}\, \mathbb{U}_t^s(\mathbf{b}_t^s; \mathbf{T}, ExcTCO2e_t) \\ s.t. \quad & \left\{ \left\{ g_{w,ct}^s\left(\mathbf{b}_t^{js}\right) \leq 0 \right\}_{w=1}^{B+1} \right\}_{t=1,2,...,}\end{aligned} \tag{6.36}$$

where $g_{w,ct}^{js}\left(\mathbf{b}_t^{js}\right)$ corresponds to the wth inequality constraint of the original firm's problem in (6.33).

Optimality Conditions

First, we establish that the consumer's problem in (6.36) is a convex maximization problem. The objective function in (6.36) is a concave given the assumption that u^s is strictly concave. The set X_c is convex. Moreover, each inequality constraint $g_{w,ct}^s(\mathbf{b}_t^s) \leq 0$ is such that $g_{w,ct}^s(\mathbf{b}_t^s)$ is a weakly convex function, implying in turn that each set $CS_{ct}^s \equiv \left\{\mathbf{b}_t^s \in X_{ct} \mid g_{w,ct}^s(\mathbf{b}_t^s) \leq 0\right\}$ for $w = 1, 2, ..., B+1$ and $t = 1, 2, ...$ is a convex set. Lastly, the set $[\bigcap_{t=1}^{\infty} CS_{ct}^s] \cap X_c$ is the intersection of convex sets, which is a convex set. Thus, the firm's problem is convex. Moreover, it is easy to show that Slater's condition is satisfied. To prove that result, note that it is sufficient to set all c_{bt}^s equal to an arbitrarily small and strictly positive value, and such bundle satisfies all the constraints with strict inequality given that e_t^s is assumed to be large. This implies that a global maximizer exists and solves the First-order Conditions. Second, we derive the First-order Necessary Conditions for a global maximum, which write:

$$\begin{aligned}\frac{\partial \mathcal{L}_c^s}{\partial \mathrm{C}_t^s} &= 1 - \theta_{0t}^s = 0 \\ \frac{\partial \mathcal{L}_c^s}{\partial \mathrm{c}_{bt}^s} &= \frac{\partial u^s(\mathbf{c}_t^s)}{\partial \mathrm{c}_{bt}^s} - \theta_{bt}^s(p_{kt} + v_{kt}) + \theta_{bt}^s = 0\end{aligned} \quad \forall t = 1, 2, ... \tag{6.37}$$

plus the standard primal feasibility $\left\{\left\{g_{w,ct}^s(\mathbf{b}_t^s) \leq 0\right\}_{w=1}^{B+1}\right\}_{t=1}^{\infty}$, dual feasibility $\left\{\{\theta_{wt}^s \geq 0\}_{w=1}^{B+1}\right\}_{t=1}^{\infty}$, and complimentary slackness conditions $\left\{\left\{\theta_{wt}^s g_{w,ct}^s(\mathbf{b}_t^s) = 0\right\}_{w=1}^{B+1}\right\}_{t=1}^{\infty}$.

6.1.4 Government

We assume a government in each country s which spend an exogenous amount G_t^s in each period t. Public spending is a mere cost for the society[10] and is financed through solely through tax revenues (there is no sovereign debt). Tax revenues include those generated by: (i) specific taxes on the production of oil, gas, and other commodities, (ii) specific taxes on the consumption of goods, (iii) personal income taxes $ITax_t^s$, (iv) corporate income taxes, (v) flaring tax (if any), and (vi) the payments of any fine due by firms because of the violation of flaring and/or venting regulation (if any). The government is a passive player solely defined by its budget constraint, which is assumed to be balanced in every period and has the following functional form:

$$\begin{aligned} G_t^s - \textstyle\sum_{i\in K^s} \tau_t^{iks,\text{Oil}}\text{Oil}_t^{iks} + \tau_t^{iks,\text{Gas}}\text{Gas}_t^{iks} + \tau_t^{iks,\text{Flare}}\text{Flare}_t^{i} \\ +VF_t^{iks} \times \mathbf{1}\left[\text{ivent}_t^{iks}=1\right] + FF_t^{iks} \times \mathbf{1}\left[\text{Flare}_t^{iks} > \bar{F}^i\right] + \left(\mathbf{a}_t^{js}\right)'\mathbf{y}_t^{js} \\ +\tau_t^{js,\text{F}}\text{F}_t^{js} + (\mathbf{v}_t)'\,\mathbf{c}_t^s + ITax_t^s + \left[\left(\textstyle\sum_{k=1}^{K^s}\sum_{i=1}^{I^k} T_t^{ks} G\Pi_t^{iks}\right) + \left(\textstyle\sum_{j=1}^{J^s} T_t^{js} G\Pi_t^{js}\right)\right] = 0 \end{aligned} \tag{6.38}$$

where

$$\begin{aligned} G\Pi_t^{iks} = \quad & Rev_t^{iks}\left(\text{Oil}_t^{iks}, \text{Gas}_t^{iks}, \text{Flare}_t^{iks}, \text{PInj}_t^{iks}, \text{PInS}_t^{iks}, \text{Z}_t^{iks}\right) \\ & -InvCost_t^{iks}\left(\text{ID}_t^{iks}, \text{IM}_t^{iks}, \text{ReInj}_t^{iks}, \text{PInj}_t^{iks}, \text{OInj}_t^{iks}; PP_t^{iks,\text{Gas}}\right) \end{aligned} \tag{6.39}$$

and

$$G\Pi_t^{js} = \left[\left(\mathbf{p}_t^s - \mathbf{a}_t^{js}\right)'\mathbf{y}_t^{js} + P_t^{\text{Oil}}\text{O}_t^{js} + P_t^{s,\text{Gas}}\text{G}_t^{js} + \text{MZ}_t^{js} - \text{F}_t^{js}\tau_t^{js,\text{F}}\right] \tag{6.40}$$

are the gross profits (excluding fines) generated by field i owned by upstream firm k and by midstream firm j, respectively. The rates of each tax $\tau_t^{iks,\text{Oil}}, \tau_t^{iks,\text{Gas}}, \tau_t^{iks,\text{Flare}}, T_t^{ks}, \tau_t^{js,\text{F}}\mathbf{a}_t^{js}, T_t^{js}, \mathbf{v}_t^s, ITax_t^s$ are the (exogenous) policy variables that we seek to set at a desirable level. A tax scheme in country s is the collection $\mathbf{T}^s$ of all taxes imposes on all firms and consumers in each period t; i.e.,

$$\mathbf{T}^s = \left\{\left\{\tau_t^{iks,\text{Oil}}, \tau_t^{iks,\text{Gas}}, \tau_t^{iks,\text{Flare}}, T_t^{ks}\right\}_{k=1}^{K^s}, \left\{\tau_t^{js,\text{F}}\mathbf{a}_t^{js}, T_t^{js}\right\}_{j=1}^{J^s}, \mathbf{v}_t^s, ITax_t^s\right\}_{t=1}^{\infty} \tag{6.41}$$

Lastly, we denote with $\mathbf{T} = \{\mathbf{T}^s\}_{s=1}^S$ the collection of all the tax schemes adopted by each country $s = 1, 2, ..., S$.

Tax Reform

A tax reform is defined as a change in the value of some of the tax rates in $\mathbf{T}^s$ relative to their values under the existing tax scheme, that delivers a new tax scheme $\check{\mathbf{T}}^s$. We use the symbol Δ to denote a change in a given variable; for instance:

$$\Delta\tau_t^{iks,\text{Oil}} = \check{\tau}_t^{iks,\text{Oil}} - \tau_t^{iks,\text{Oil}} \tag{6.42}$$

where $\check{\tau}_t^{iks,\text{Oil}}$ is an element of $\check{\mathbf{T}}^s$. Let $type_t^{iks} \in \{oil, gas, mixed\}$ denote the type of oil&gas field, as defined in section 1.1.1. The reform proposed in this paper consists in the following adjustments:

1. A change in the tax rate on unrefined natural gas sales:

$$\Delta\tau_t^{iks,\text{Gas}} = \begin{cases} \min_{i\in\{\{1,2,...,I_k\}\}_{k=1}^K}\left\{P_t^{s,\text{Gas}} - \tau_t^{iks,\text{Gas}} - MC_t^{iks,\text{Gas}}\right\} & if \quad type_t^{iks} = oil \\ \frac{P_t^{s,\text{Gas}}\left(1-\varsigma_t^{\text{Gas}}MS_t^{ks,\text{Gas}}\right)}{\eta_{\text{G},P}^{s,GO}} \frac{\sum_{k=1}^{K^s}\sum_{i=1}^{I^k}\Delta\text{Gas}_t^{iks}\mathbf{1}\left[\text{Oil}_t^{iks}>0\right]}{\sum_{k=1}^{K^s}\sum_{i=1}^{I^k}\text{Gas}_t^{iks}\mathbf{1}\left[\text{Oil}_t^{iks}=0\right]} & if \quad type_t^{iks} = gas \\ 0 & otherwise \end{cases} \tag{6.43}$$

[10] This assumption can be easily relaxed. For instance, one could impose the alternative assumption that G_t^s enters the utility function of the consumer in country s, with no consequences for the present analysis.

2. A change in the tax rate on crude oil sales:

$$\Delta\tau_t^{iks,\text{Oil}} = \begin{cases} -\Delta\tau_t^{iks,\text{Gas}} GOR^{iks}\left(1-\vartheta^{iks}\right) & if \quad type_t^{iks} = oil \\ 0 & otherwise \end{cases} \tag{6.44}$$

3. A change in the rate of deduction of non-commercial gas use and unavoidable gas losses:

$$\Delta\delta_{1t}^{iks} = \begin{cases} -\Delta\tau_t^{iks,\text{Gas}} & if \quad type_t^{iks} = oil \\ 0 & otherwise \end{cases} \tag{6.45}$$

4. A change in the lump-sum deduction amount:

$$\Delta\delta_{0t}^{iks} = \begin{cases} -\Delta\tau_t^{iks,\text{Gas}} \widehat{Gas}_t^{iks} & if \quad type_t^{iks} = gas \\ -\Delta\tau_t^{iks,\text{Gas}} \left(\text{NRF}_t^{iks} - \widehat{Maint}_t^{iks}\right) & if \quad type_t^{iks} = oil \\ 0 & otherwise \end{cases} \tag{6.46}$$

where

$$\widehat{Gas}_t^{iks} = \frac{\sum_{l\neq k}\sum_{j=1}^{I^l} \text{Gas}_t^{jls}\mathbf{1}\left[type_t^{jls} = Gas\right]}{\left(K^s-1\right)\left(\sum_{j=1}^{I^k}\mathbf{1}\left[type_t^{jls} = Gas\right]\right)} \tag{6.47}$$

is the average natural gas production from gas-only fields owned by firms other than k, and

$$\widehat{Maint}_t^{iks} = \frac{\sum_{l\neq k}\sum_{j=1}^{I^l} \text{ResGas}_t^{jls}\mathbf{1}\left[type_t^{jls} = Oil\right]}{\left(K^s-1\right)\sum_{j=1}^{I^k}\mathbf{1}\left[type_t^{jks} = Oil\right]} \tag{6.48}$$

where residual gas ResGas_t^{jls} has formula:

$$\text{ResGas}_t^{jls} = \left(1-\vartheta^{iks}\right)^{-1}\text{TotGas}_t^{jls} - \text{ReInj}_t^{jls} - \text{ReInS}_t^{jls} - RF_t^{jls} - \text{Gas}_t^{jls} \tag{6.49}$$

and represents the average natural gas extracted from oil fields owned by firms other than k that cannot be attributed to observable variables; i.e., it equals the sum over fields j owned by firms other than k of $\text{IVent}_t^{jls} + \text{Flare}_t^{jls} - RF_t^{jls} - Maint_t^{jls}\left(\text{M}_t^{jls}\right)$. Note that both $\widehat{Gas}_t^{iks}$ and $\widehat{Maint}_t^{iks}$ are independent of changes in choice variables controlled by firm k.

A further optional tax adjustment consists in the introduction of a tax rate on flaring for midstream firms, whose description we postpone to section 2.2. Lastly, we assume that the current (pre-reform) tax scheme features no flaring fee (i.e., $FF_t^{iks} = 0$) and the same tax rate on natural gas production for all firms; i.e., $\tau_t^{iks,\text{Gas}} = \tau_t^{s,\text{Gas}}$ for all fields i in country s.

6.1.5 Climate Change

We assume that the excess concentration of greenhouse gases in the atmosphere follows a simple law of motion:

$$-\Lambda Emissions_t + ExcTCO2e_{t+1} - (1-\Gamma)\, ExcTCO2e_t = 0 \tag{6.50}$$

where $Emissions_t$ represents the global emissions of greenhouse gases in period t, Λ is a parameter that captures the effect of new GHG emissions in period t on the concentration of GHG in the atmosphere, whereas Γ represents the annual rate of natural decline of the excess concentration of GHG in the atmosphere. While admittedly stylized, these assumptions capture the key consequence of GHG emissions for the purpose of this analysis: the long-lasting effect on the global climate, which represent a cost for individuals both in the current period and in the future (see section 1.3). Note that all climate change variables have no superscript: this notation captures the fact that they are global-level variables.

6.2 Main Analytical Results

This section contains the main analytical results and their proofs.

6.2.1 Upstream

Proposition 1. *If (i)* $\frac{\partial^2 PF_t^{iks}\left(IVent_t^{iks}, Flare_t^{iks}\right)}{\partial IVent_t^{iks}\partial Flare_t^{iks}} \geq 0$ *then intentional venting* $IVent_t^{iks}$ *by field* i *is weakly increasing in the flaring tax rate* $\tau_t^{iks,Flare}$ *and weakly decreasing in the flaring ceiling* $\overline{Flare}_t^{iks}$. *Moreover, if (ii)* $\frac{\partial^2 PF_t^{iks}\left(IVent_t^{iks}, Flare_t^{iks}\right)}{\partial IVent_t^{iks}\partial Flare_t^{iks}} \left[\frac{\partial^2 PF_t^{i}\left(IVent_t^{iks}, Flare_t^{iks}\right)}{\partial \left(IVent_t^{iks}\right)^2}\right]^{-1} > \frac{CI^{Flare}}{CI^{Vent}}$ *then the total GHG emissions* $CO2e_t^{iks}$ *by field* i *is weakly increasing in the flaring tax rate* $\tau_t^{iks,Flare}$ *and weakly decreasing in the flaring ceiling* $\overline{Flare}_t^{iks}$.

Proof. Part (i): effect of an increase in $\tau_t^{iks,\text{Flare}}$ for field i. First, note that the first constraint is always binding, implying $\phi_{1t}^{iks} > 0$ at the optimal solution, because otherwise $\frac{\partial TF_{6t}^{i}(.)}{\partial Z_t^{i}} > 0$ implies that it would be possible to obtain strictly larger profits in period t with no effect on the profit gained in any other period through a marginal increase in Z_t^{iks}. Secondly, under the assumption that $MC_t^{iks,\text{Gas}} \equiv \frac{\partial TF_{2t}^{iks}(.)}{\partial \text{Gas}_t^{iks}} \Big/ \frac{\partial TF_{6t}^{iks}(.)}{\partial Z_t^{iks}}$ is constant and given that the domain of the variable Gas_t^{iks} is unbounded and we know that the solution must satisfy the FOCs, the FOC w.r.t. Gas_t^{iks} implies that $-\phi_{2t}^{iks} + \phi_{3t}^{iks} + \phi_{4t}^{iks}\zeta$ is constant in $\tau_t^{iks,\text{Flare}}$. Moreover, the assumptions that (1) $MC_t^{iks,\text{Flare}}\left(\text{Flare}_t^{iks}\right)$ is constant in Z_t^{iks} and that (2) M_t^{iks} enters the formula for $Pr_t^{i}\left(ivent_t^{iks} = 1 \mid \Omega_t^{ks}\right)$ in an additively separable fashion imply that all variables except possibly Flare_t^{iks},IVent_t^{iks}, and Gas_t^{iks} are all constant in $\tau_t^{iks,\text{Flare}}$. In order to study the effect of a marginal change in $\tau_t^{iks,\text{Flare}}$ on these three variables we must distinguish two cases. Case 1. If the FOC w.r.t. ϕ_{6t}^{iks} is binding, then the optimal level of Flare_t^{iks} is a corner solution $\text{Flare}_t^{iks} = \text{NRF}_t^{iks}$ at the baseline value of $\tau_t^{iks,\text{Flare}}$ (and in turn $\text{Flare}_t^{iks} = 0$) , then $\frac{d\text{Flare}_t^{iks}}{d\tau_t^{iks,\text{Flare}}} = 0$ and trivially the FOCs imply $\frac{d\text{IVent}_t^{iks}}{d\tau_t^{iks,\text{Flare}}} = \frac{d\text{Flare}_t^{iks}}{d\tau_t^{iks,\text{Flare}}} = \frac{d\text{Gas}_t^{iks}}{d\tau_t^{iks,\text{Flare}}} = 0$ and therefore $\frac{d\text{CO2e}_t^{iks}}{d\tau_t^{iks,\text{Flare}}} = 0$. Case 2. If the FOC w.r.t. at the baseline value of $\tau_t^{iks,\text{Flare}}$ Flare_t^{iks} are binding (and in turn $\text{Flare}_t^{iks} \geq 0$), then we must make use of the FOCs w.r.t. Flare_t^{iks} and IVent_t^{iks}. We define the marginal cost of flaring as follows: $MC_t^{iks,\text{Flare}}\left(\text{Flare}_t^{iks}\right) \equiv \frac{\partial TF_{3t}^{iks}(.)}{\partial \text{Flare}_t^{iks}} \Big/ \frac{\partial TF_{6t}^{iks}(.)}{\partial Z_t^{iks}}$. Then under the condition:

$$\begin{aligned} (1 - T_t^{s})\, MC_t^{iks,\text{Flare}}\left(\text{Flare}_t^{iks}\right) + VF_t^{iks}\, \frac{\partial PF_t^{iks}\left(\text{IVent}_t^{iks}, \text{Flare}_t^{iks}\right)}{\partial \text{Flare}_t^{iks}}\Bigg|_{\text{IVent}_t^{iks}=0} \\ \geq VF_t^{iks}\, \frac{\partial PF_t^{iks}\left(\text{IVent}_t^{iks}, \text{Flare}_t^{iks}\right)}{\partial \text{IVent}_t^{iks}}\Bigg|_{\text{IVent}_t^{iks}=0} \end{aligned} \tag{6.51}$$

for all values of Flare_t^{iks}, it must be true that if $\text{Flare}_t^{iks} > 0$ at the baseline value of $\tau_t^{iks,\text{Flare}} \geq 0$ then the primal feasibility condition $\text{IVent}_t^{iks} \geq 0$ is also not binding. We prove this result by contradiction. Suppose it is. $\text{Flare}_t^{iks} > \text{NRF}_t^{iks}$ implies $\phi_{6t}^{iks} = 0$, and in turn at the optimal solution:

$$\begin{aligned} -\tau_t^{iks,\text{Flare}}\left(1 - T_t^{ks}\right) - VF_t^{iks} \times \frac{\partial PF_t^{iks}\left(\text{IVent}_t^{iks}, \text{Flare}_t^{iks}\right)}{\partial \text{Flare}_t^{iks}}\Bigg|_{\text{IVent}_t^{iks}=0} \\ -\phi_{1t}^{iks}\frac{\partial TF_{3t}^{iks}(.)}{\partial \text{Flare}_t^{iks}} + \left(\phi_{2t}^{iks} - \phi_{3t}^{iks} - \phi_{4t}^{iks}\zeta\right)\left(1 - \vartheta^{iks}\right)^{-1} = 0 \end{aligned} \tag{6.52}$$

If the constraint $\text{IVent}_t^{iks} \geq 0$ is binding, then $\phi_{7t}^{iks} > 0$ and therefore at the optimal solution:

$$-VF_t^{iks} \times \left.\frac{\partial PF_t^{iks}\left(\text{IVent}_t^{iks}, \text{Flare}_t^{iks}\right)}{\partial \text{IVent}_t^{iks}}\right|_{\text{IVent}_t^{iks}=0} + \left(\phi_{2t}^{iks} - \phi_{3t}^{iks} - \phi_{4t}^{iks}\zeta\right)\left(1-\vartheta^{iks}\right)^{-1} < 0 \tag{6.53}$$

Combining these two inequalities, we obtain

$$\begin{array}{c}\left(1-T_t^i\right) MC_t^{iks,\text{Flare}}\left(\text{Flare}_t^{iks}\right) + VF_t^{iks} \times \left.\frac{\partial PF_t^{iks}\left(\text{IVent}_t^{iks},\text{Flare}_t^{iks}\right)}{\partial \text{Flare}_t^{iks}}\right|_{\text{IVent}_t^{iks}=0} \\ +\tau_t^{iks,\text{Flare}}\left(1-T_t^{ks}\right) - VF_t^{iks} \times \left.\frac{\partial PF_t^{i}\left(\text{IVent}_t^{iks},\text{Flare}_t^{iks}\right)}{\partial \text{IVent}_t^{iks}}\right|_{\text{IVent}_t^{iks}=0} < 0\end{array}, \tag{6.54}$$

which leads to a contradiction. Thus, the solution for IVent_t^{iks} is not a corner solution with respect to its natural boundary. Moreover, the firm's objective function is globally concave and (partially) strictly concave in IVent_t^{iks} and Flare_t^{iks}. Thus we can obtain the marginal effects on $\tau_t^{iks,\text{Flare}}$ on IVent_t^{iks} and Flare_t^{iks} by differentiating the FOC w.r.t. IVent_t^{iks} and Flare_t^{iks} evaluated at $\phi_{6t}^{iks} = 0$, $\phi_{7t}^{iks} = 0$, and solving for the derivatives of interest, to get:

$$\frac{d\text{IVent}_t^{iks}}{d\tau_t^{iks,\text{Flare}}} = \frac{\frac{\partial^2 PF_t^{iks}(\cdot)}{\partial \text{IVent}_t^{i}\partial \text{Flare}_t^{i}}\left(1-T_t^s\right)}{\frac{\partial^2 PF_t^{iks}(\cdot)}{\partial\left(\text{IVent}_t^{iks}\right)^2}\left(\frac{\partial^2 PF_t^{iks}(\cdot)}{\partial\left(\text{Flare}_t^{iks}\right)^2} + \frac{1}{VF_t^{iks}}\frac{\partial^2 MC_t^{iks,\text{Flare}}(\cdot)}{\partial\left(\text{Flare}_t^{iks}\right)^2}\right) - \left(\frac{\partial^2 PF_t^{iks}(\cdot)}{\partial \text{IVent}_t^{iks}\partial \text{Flare}_t^{iks}}\right)^2} \tag{6.55}$$

which is positive if $\frac{\partial^2 PF_t^{iks}\left(\text{IVent}_t^{iks},\text{Flare}_t^{iks}\right)}{\partial \text{IVent}_t^{iks}\partial \text{Flare}_t^{iks}} \geq 0$, and

$$\frac{d\text{Flare}_t^{iks}}{d\tau_t^{iks,\text{Flare}}} = -\frac{\frac{\partial^2 PF_t^{iks}(\cdot)}{\partial\left(\text{IVent}_t^{iks}\right)^2}\left(1-T_t^s\right)}{\frac{\partial^2 PF_t^{iks}(\cdot)}{\partial\left(\text{IVent}_t^{iks}\right)^2}\left(\frac{\partial^2 PF_t^{iks}(\cdot)}{\partial\left(\text{Flare}_t^{iks}\right)^2} + \frac{1}{VF_t^{iks}}\frac{\partial^2 MC_t^{iks,\text{Flare}}(\cdot)}{\partial\left(\text{Flare}_t^{iks}\right)^2}\right) - \left(\frac{\partial^2 PF_t^{iks}(\cdot)}{\partial \text{IVent}_t^{iks}\partial \text{Flare}_t^{iks}}\right)^2} \tag{6.56}$$

and the residual effect on Gas_t^{iks}, which is obtained using $\frac{d\text{Gas}_t^{iks}}{d\tau_t^{iks,\text{Flare}}} = -\left(\frac{d\text{IVent}_t^{iks}}{d\tau_t^{iks,\text{Flare}}} + \frac{d\text{Flare}_t^{iks}}{d\tau_t^{iks,\text{Flare}}}\right)$, writes:

$$\frac{d\text{Gas}_t^{iks}}{d\tau_t^{iks,\text{Flare}}} = \frac{\left(\frac{\partial^2 PF_t^{iks}(\cdot)}{\partial\left(\text{IVent}_t^{iks}\right)^2} - \frac{\partial^2 PF_t^{iks}(\cdot)}{\partial \text{IVent}_t^{iks}\partial \text{Flare}_t^{iks}}\right)\left(1-T_t^{ks}\right)}{\frac{\partial^2 PF_t^{iks}(\cdot)}{\partial\left(\text{IVent}_t^{iks}\right)^2}\left(\frac{\partial^2 PF_t^{iks}(\cdot)}{\partial\left(\text{Flare}_t^{iks}\right)^2} + \frac{1}{VF_t^{iks}}\frac{\partial^2 MC_t^{iks,\text{Flare}}(\cdot)}{\partial\left(\text{Flare}_t^{iks}\right)^2}\right) - \left(\frac{\partial^2 PF_t^{iks}(\cdot)}{\partial \text{IVent}_t^{iks}\partial \text{Flare}_t^{iks}}\right)^2} \tag{6.57}$$

However, if the change in the tax rate extends to all the oil&gas fields, then the change in the gas supply may affect the equilibrium price of natural gas, with effects on the levels of flaring and venting optimally chosen by each firm. The overall effect writes:

$$\begin{array}{rl}\frac{d\text{CO2e}_t^{iks}}{d\tau_t^{iks,\text{Flare}}} = & \left(\frac{d\text{IVent}_t^{iks}}{d\tau_t^{iks,\text{Flare}}} + \frac{d\text{IVent}_t^{i}}{dP_t^{s,\text{Gas}}}\frac{dP_t^{s,\text{Gas}}}{d\tau_t^{iks,\text{Flare}}}\right) CI^{\text{Vent}} + \left(\frac{d\text{Flare}_t^{iks}}{d\tau_t^{i,\text{Flare}}} + \frac{d\text{Flare}_t^{iks}}{dP_t^{s,\text{Gas}}}\frac{dP_t^{s,\text{Gas}}}{d\tau_t^{iks,\text{Flare}}}\right) CI^{\text{Flare}} \\ & + \left(\frac{d\text{Gas}_t^{iks}}{d\tau_t^{iks,\text{Flare}}} + \frac{d\text{Gas}_t^{iks}}{dP_t^{s,\text{Gas}}}\frac{dP_t^{s,\text{Gas}}}{d\tau_t^{iks,\text{Flare}}}\right) CI^{iks,\text{Gas}}\end{array} \tag{6.58}$$

Because flaring and gas production are gross substitutes, we can show that $\frac{d\text{Gas}_t^{iks}}{d\tau_t^{iks,\text{Flare}}} + \frac{d\text{Gas}_t^{iks}}{dP_t^{s,\text{Gas}}}\frac{dP_t^{s,\text{Gas}}}{d\tau_t^{iks,\text{Flare}}} \geq 0$, $\frac{dP_t^{s,\text{Gas}}}{d\tau_t^{iks,\text{Flare}}} \leq 0$, $\frac{d\text{Flare}_t^{iks}}{dP_t^{s,\text{Gas}}} \leq 0$ and $\frac{d\text{IVent}_t^{iks}}{dP_t^{s,\text{Gas}}} \leq 0$. To see how, first note that (i) $\frac{d\text{Gas}_t^{iks}}{dP_t^{s,\text{Gas}}} \geq 0$ by the law of supply. We prove that (ii) $\frac{dP_t^{s,\text{Gas}}}{d\tau_t^{iks,\text{Flare}}} \leq 0$ by contradiction. Suppose $\frac{dP_t^{s,\text{Gas}}}{d\tau_t^{iks,\text{Flare}}} > 0$. Then, because the demand side of the economy is not directly affected by $\tau_t^{iks,\text{Flare}}$, the aggregate gas supply at national level must be lower after the rise in $\tau_t^{iks,\text{Flare}}$, implying that at least one field must reduce its gas supply following a rise in $\tau_t^{iks,\text{Flare}}$. But combining the FOCs of the oil&gas firm w.r.t. Gas_t^{iks} and Flare_t^{iks}, which are both binding in this case, Gas_t^{iks} is weakly increasing in $P_t^{s,\text{Gas}}$ at constant $\tau_t^{iks,\text{Flare}}$ and in $\tau_t^{iks,\text{Flare}}$ at constant $P_t^{s,\text{Gas}}$. This implies in turn that a marginal increase in both variables translates to an increase in Gas_t^{iks} in all fields, leading to a contradiction. Second, we prove that (iii) $\frac{d\text{Gas}_t^{iks}}{d\tau_t^{iks,\text{Flare}}} + \frac{d\text{Gas}_t^{iks}}{dP_t^{s,\text{Gas}}} \frac{dP_t^{s,\text{Gas}}}{d\tau_t^{iks,\text{Flare}}} \geq 0$ by contradiction. Suppose $\frac{d\text{Gas}_t^{iks}}{d\tau_t^{iks,\text{Flare}}} + \frac{d\text{Gas}_t^{iks}}{dP_t^{s,\text{Gas}}} \frac{dP_t^{s,\text{Gas}}}{d\tau_t^{iks,\text{Flare}}} < 0$. Because the second term is positive (see point (i) and (ii)), it must be true that $\frac{d\text{Gas}_t^{iks}}{d\tau_t^{iks,\text{Flare}}} < 0$. But because the denominator of (6.57) is positive by convexity of the objective function, this is true only if $\frac{\partial^2 PF_t^{iks}\left(\text{IVent}_t^{iks},\text{Flare}_t^{iks}\right)}{\partial\left(\text{IVent}_t^{iks}\right)^2} < \frac{\partial^2 PF_t^{iks}\left(\text{IVent}_t^{iks},\text{Flare}_t^{iks}\right)}{\partial \text{IVent}_t^{iks}\partial \text{Flare}_t^{iks}}$, which contradicts the assumption that the function PF_t^{iks} is convex, because that is true only if $\frac{\partial^2 PF_t^{iks}\left(\text{IVent}_t^{iks},\text{Flare}_t^{iks}\right)}{\partial\left(\text{IVent}_t^{iks}\right)^2} \geq \frac{\partial^2 PF_t^{iks}\left(\text{IVent}_t^{iks},\text{Flare}_t^{iks}\right)}{\partial \text{IVent}_t^{iks}\partial \text{Flare}_t^{iks}}$. Lastly, $\frac{d\text{Flare}_t^{iks}}{dP_t^{s,\text{Gas}}} \leq$ and $\frac{d\text{IVent}_t^{iks}}{dP_t^{s,\text{Gas}}} \leq 0$ can be proved by substituting the FOC w.r.t. Gas_t^{iks} into the FOCs w.r.t. (a) Flare_t^{iks} and (b) IVent_t^{iks} and totally differentiating each of the two resulting equation w.r.t. Flare_t^{iks} and IVent_t^{iks}, respectively. Thus, it is possible to calculate a lower bound on the effect of interest because the following inequality holds:

$$\frac{d\text{CO2e}_t^{iks}}{d\tau_t^{iks,\text{Flare}}} \geq \frac{d\text{IVent}_t^{iks}}{d\tau_t^{iks,\text{Flare}}} CI^{\text{Vent}} + \frac{d\text{Flare}_t^{iks}}{d\tau_t^{iks,\text{Flare}}} CI^{\text{Flare}} \tag{6.59}$$

where the RHS of equation 6.59 corresponds to the effect of a marginal increase in $\tau_t^{iks,\text{Flare}}$ at constant gas prices. Substituting the formulas in 6.55 and 6.56 into 6.59 we obtain:

$$\frac{d\text{CO2e}_t^{iks}}{d\tau_t^{iks,\text{Flare}}} > \frac{\left(\frac{\partial^2 PF_t^{iks}(\cdot)}{\partial \text{IVent}_t^{iks}\partial \text{Flare}_t^{iks}} CI^{\text{Vent}} - \frac{\partial^2 PF_t^{iks}(\cdot)}{\partial\left(\text{IVent}_t^{iks}\right)^2} CI^{\text{Flare}}\right)\left(1 - T_t^s\right)}{\frac{\partial^2 PF_t^{iks}(\cdot)}{\partial\left(\text{IVent}_t^{iks}\right)^2}\left(\frac{\partial^2 PF_t^{iks}(\cdot)}{\partial\left(\text{Flare}_t^{iks}\right)^2} + \frac{1}{VF_t^{iks}} \frac{\partial^2 MC_t^{iks,\text{Flare}}(\cdot)}{\partial\left(\text{Flare}_t^{iks}\right)^2}\right) - \left(\frac{\partial^2 PF_t^{iks}(\cdot)}{\partial \text{IVent}_t^{iks}\partial \text{Flare}_t^{iks}}\right)^2} \geq 0 \tag{6.60}$$

Thus, it is sufficient to derive a condition for the second inequality above to be weakly satisfied to obtain a sufficient condition for $\frac{d\text{CO2e}_t^{iks}}{d\tau_t^{iks,\text{Flare}}} > 0$. Because the firm's objective function is (partially) strictly concave in Flare_t^{iks} and IVent_t^{iks}, the denominator of the RHS of 6.60 is positive. Thus, a sufficient condition for the inequality in 6.60 to hold true writes:

$$\frac{\partial^2 PF_t^{iks}\left(\text{IVent}_t^{iks}, \text{Flare}_t^{iks}\right)}{\partial \text{IVent}_t^{iks}\partial \text{Flare}_t^{iks}} \left[\frac{\partial^2 PF_t^{iks}\left(\text{IVent}_t^{iks}, \text{Flare}_t^{iks}\right)}{\partial\left(\text{IVent}_t^{iks}\right)^2}\right]^{-1} > \frac{CI^{\text{Flare}}}{CI^{\text{Vent}}}, \tag{6.61}$$

which corresponds to the condition stated in Proposition 1.

Part (ii): effect of a marginal decrease in $\overline{\text{Flare}}_t^{iks}$. Note that if the flaring ceiling is not binding for field i then the effect of a marginal change in the value of $\overline{\text{Flare}}_t^{iks}$ is trivially zero. If the ceiling is binding, then the problem is equivalent at the margin to that in 6.19 but with three extra conditions in each period t and each field $i = 1, 2, ..., I^k$: a constraint $\text{Flare}_t^{iks} - \overline{\text{Flare}}_t^{iks} \leq 0$ with associated KT multiplier ϕ_{13t}^i, a dual feasibility condition $\phi_{13t}^{iks} \geq 0$ and a complementary slackness condition $\phi_{13t}^{iks}\left(\text{Flare}_t^{iks} - \overline{\text{Flare}}_t^{iks}\right) = 0$. Then whenever this new constraint is binding ($\phi_{13t}^{iks} > 0$) it can be interpreted as the shadow price of relaxing the constraint. This implies that the effect of a marginal decrease in $\overline{\text{Flare}}_t^{iks}$ is equivalent to that of a marginal increase in the flaring tax multiplied by $\phi_{13t}^{iks}/\left(1 - T_t^{ks}\right) > 0$. Thus, the sign of the effect of a marginal decrease in $\overline{\text{Flare}}_t^{iks}$ is the same as that of a marginal increase in the flaring tax $\tau_t^{iks,\text{Flare}}$, which is stated in part (i) of Proposition 1. Q.E.D. Note that $\frac{\partial^2 PF_t^{iks}\left(\text{IVent}_t^{iks},\text{Flare}_t^{iks}\right)}{\partial \text{IVent}_t^{iks}\partial \text{Flare}_t^{iks}} \geq 0$ is a sufficient condition for total emission being increasing in the flaring tax rate, but that may be true even if such condition is not satisfied. Secondly, note that the RHS of inequality (6.61) is a very small number (< 0.076, calculated using the values $CI^{\text{Flare}} = 0.3018\ TCO2e/BOE$ and $CI^{\text{Vent}} = 3.9583\ TCO2e/BOE$ from Brandt et al. [2018]) implying that the condition is satisfied even for modest degrees of substitutability between flaring and intentional venting. In the empirical section, we show that this condition is satisfied in the data, and we quantify the detrimental effect of an increase in flaring taxation. Note that under the parametric specification for PF_t^{iks} adopted in the empirical section of this paper, the condition in (6.61) reduces to $\frac{\kappa_2^{iks}}{\kappa_2^{iks}} \geq \frac{CI^{\text{Flare}}}{CI^{\text{Vent}}}$, where the threshold $\frac{\kappa_2^{iks}}{\kappa_1^{iks}}$ represents the marginal effect of reducing flaring by one unit on the amount of intentional venting and is identified by the model, allowing for estimation of the effect of interest.

Proposition 2. *If $\frac{\partial PU_t^{iks}(\cdot)}{\partial M_t^{iks}} \geq 0$, then an increase in the venting fine VF_t^{iks} for an oil field i translates into (i) weakly lower maintenance M_t^{iks} and (ii) weakly larger unintentional venting $UVent_t^{iks}$ for any given quantity of extracted gas $TotGas_t^{iks}$.*

Proof. Part (i). Case (1): the seventh constraint is binding (and therefore $\phi_{7t}^{iks} > 0$ and $\text{IVent}_t^{iks} = 0$). Then given the assumption $Pr_t^{iks}\left(ivent_t^{iks} = 1 \mid \Omega_t^{PUB}, \text{IVent}_t^{iks} = 0\right) = 0$ the value of VF_{t+1}^{iks} does not enter any binding optimality condition, trivially implying $\frac{d\text{M}_{t+1}^{iks}}{dVF_{t+1}^{iks}} = 0$ and $\frac{dUVent_{t+1}^{iks}}{dVF_{t+1}^{iks}} = 0$. Case (2): the the ninth constraint is not binding (and therefore $\phi_{7t}^{iks} = 0$ and $\text{IVent}_t^{iks} \geq 0$). Let us consider the FOCs of the firm's problem w.r.t. M_{t+1}^{iks} and IM_t^{iks}:

$$\begin{aligned}\frac{\partial \mathcal{L}_u^{ks}}{\partial \text{M}_{t+1}^{iks}} = \quad & -\beta\phi_{1t+1}^{iks}\frac{\partial TF_{5t+1}^{iks}\left(\text{M}_{t+1}^{iks}\right)}{\partial \text{M}_{t+1}^{iks}} - \lambda_{2t}^{iks} + \phi_{10t}^{iks} + \beta\lambda_{2t+1}^{iks}\left(1-\rho^{iks}\right) \\ & -\beta VF_{t+1}^{iks}\frac{\partial PU_{t+1}^{iks}(\cdot)}{\partial \text{M}_{t+1}^{i}} - \beta\left[\phi_{2t+1}^{iks} - \phi_{3t+1}^{iks} - \phi_{4t+1}^{iks}\zeta\right]\left(1-\vartheta^{iks}\right)^{-1}\frac{\partial Maint_{t+1}^{iks}(\cdot)}{\partial \text{M}_{t+1}^{iks}} = 0\end{aligned} \tag{6.62}$$

and

$$\frac{\partial \mathcal{L}_u^{ks}}{\partial \text{IM}_t^{iks}} = -\lambda_{2t}^{iks} + \left(1 - T_t^s\right) = 0 \tag{6.63}$$

where the second condition is always binding given that IM_t^{iks} possesses unbounded support. Second, the sixth constraint is an equality constraint and, as such, it is always binding. Thus, we can substitute $\lambda_{2t}^{i} = \left(1 - T_t^{ks}\right)$ into the FOC w.r.t. M_{t+1}^{iks} to obtain the condition:

$$\begin{aligned}&\beta\left(1-T_{t+1}^s\right)MPE_{t+1}^{iks}\left(\text{M}_{t+1}^{iks}\right) - \left[\left(1-T_t^s\right) - \beta\left(1-\rho^i\right)\left(1-T_{t+1}^s\right)\right] + \phi_{10t}^{iks} \\ &-\beta\left[VF_{t+1}^{iks}\frac{\partial PU_{t+1}^{iks}(\cdot)}{\partial \text{M}_{t+1}^{i}} + \left(1-T_{t+1}^s\right)\left(P_{t+1}^{s,\text{Gas}} - \tau_{t+1}^{iks,\text{Gas}} - MC_{t+1}^{iks,\text{Gas}}\right)\left(1-\vartheta^{iks}\right)^{-1}\frac{\partial Maint_{t+1}^{iks}(\cdot)}{\partial \text{M}_{t+1}^{iks}}\right] = 0\end{aligned} \tag{6.64}$$

where the above follows the fact that IVent_{t+1}^{iks} is constant in M_{t+1}^{iks} and $MPE_{t+1}^{i}\left(\text{M}_{t+1}^{iks}\right) \equiv \frac{\partial TF_{5t+1}^{iks}\left(\text{M}_{t+1}^{iks}\right)}{\partial \text{M}_{t+1}^{iks}} \Big/ \frac{\partial TF_{6t+1}^{iks}\left(\text{Z}_{t+1}^{iks}\right)}{\partial \text{Z}_{t+1}^{iks}}$. If the tenth constraint is binding ($\phi_{10t}^{i} > 0$), then $\text{M}_{t+1}^{iks} = 0$ and the derivative of interest is simply $\frac{d\text{M}_{t+1}^{iks}}{dVF_{t+1}^{iks}} = 0$. If the the tenth constraint is non-binding ($\phi_{10t}^{i} = 0$), then we can totally differentiate the F.O.C. w.r.t. VF_{t+1}^{iks} to obtain:

$$\begin{aligned}&-\beta\frac{\partial PU_{t+1}^{iks}(\cdot)}{\partial \text{M}_{t+1}^{i}} + \beta\left\{\left(1-T_{t+1}^s\right)\frac{\partial MPE_{t+1}^{iks}(\cdot)}{\partial \text{M}_{t+1}^{iks}} - VF_{t+1}^{iks}\frac{\partial^2 PU_{t+1}^{iks}(\cdot)}{\partial \text{M}_{t+1}^{i\ 2}}\right. \\ &\left.-\left(1-T_{t+1}^s\right)\left(P_{t+1}^{s,\text{Gas}} - \tau_{t+1}^{iks,\text{Gas}} - MC_{t+1}^{iks,\text{Gas}}\right)\left(1-\vartheta^{iks}\right)^{-1}\frac{\partial^2 Maint_{t+1}^{i}(\cdot)}{\partial \text{M}_{t+1}^{i\ 2}}\right\}\frac{d\text{M}_{t+1}^{iks}}{dVF_{t+1}^{iks}} = 0\end{aligned} \tag{6.65}$$

Note that because the optimization problem is convex, the Second-order Necessary Conditions for a global maximum must be satisfied; i.e., the bordered Hessian matrix must be negative semi-definite. This condition implies that the second derivative of the Lagrangian w.r.t. M_{t+1}^{i} must satisfy:

$$\begin{aligned}\frac{\partial^2 \mathcal{L}_u^{ks}}{\partial \text{M}_{t+1}^{iks\ 2}} = \quad & \beta\left\{\left(1-T_{t+1}^s\right)\frac{\partial MPE_{t+1}^{iks}(\cdot)}{\partial \text{M}_{t+1}^{iks}} - VF_{t+1}^{iks}\frac{\partial^2 PU_{t+1}^{iks}(\cdot)}{\partial \text{M}_{t+1}^{i\ 2}}\right. \\ & \left.-\left(1-T_{t+1}^s\right)\left(P_{t+1}^{s,\text{Gas}} - \tau_{t+1}^{iks,\text{Gas}} - MC_{t+1}^{iks,\text{Gas}}\right)\left(1-\vartheta^{iks}\right)^{-1}\frac{\partial^2 Maint_{t+1}^{i}(\cdot)}{\partial \text{M}_{t+1}^{i\ 2}}\right\} \leq 0\end{aligned} \tag{6.66}$$

at all possible values of the choice variables in X^u. In particular, evaluating condition (6.66) at the candidate solution values that solve the FOCs, the SONCs are satisfied only if:

$$\beta\left\{\left(1-T^{s}_{t+1}\right)\frac{\partial MPE^{iks}_{t+1}(\cdot)}{\partial \mathrm{M}^{iks}_{t+1}}-VF^{iks}_{t+1}\frac{\partial^2 PU^{iks}_{t+1}(\cdot)}{\partial \mathrm{M}^{i\,2}_{t+1}}\right. \\ \left.-\left(1-T^{s}_{t+1}\right)\left(P^{s,\text{Gas}}_{t+1}-\tau^{iks,\text{Gas}}_{t+1}-MC^{iks,\text{Gas}}_{t+1}\right)\left(1-\vartheta^{iks}\right)^{-1}\frac{\partial^2 Maint^{i}_{t+1}(\cdot)}{\partial \mathrm{M}^{i\,2}_{t+1}}\right\}\leq 0 \tag{6.67}$$

Lastly, note that if the weak inequality (6.67) is satisfied with strict equality, then the FOC in (6.62) does not have a unique solution for M^{iks}_{t+1}, implying that the optimal value of M^{iks}_{t+1} is pinned down by the constraint $\mathrm{M}^{iks}_{t+1}-\mathrm{M}^{iks}_{t}\left(1-\rho^{iks}\right)-\mathrm{IM}^{iks}_{t}\leq 0$ and it is therefore constant in VF^{iks}_{t+1}. Conversely, if the inequality (6.67) is satisfied with strict inequality, then we can solve (6.65) with respect to $\frac{d\mathrm{M}^{iks}_{t+1}}{dVF^{iks}_{t+1}}$ and obtain:

$$\frac{d\mathrm{M}^{iks}_{t+1}}{dVF^{iks}_{t+1}}=\frac{\partial PU^{iks}_{t+1}(\cdot)}{\partial \mathrm{M}^{i}_{t+1}}\left\{\left(1-T^{s}_{t+1}\right)\frac{\partial MPE^{iks}_{t+1}(\cdot)}{\partial \mathrm{M}^{iks}_{t+1}}-VF^{iks}_{t+1}\frac{\partial^2 PU^{iks}_{t+1}(\cdot)}{\partial \mathrm{M}^{i\,2}_{t+1}}\right. \\ \left.\left.-\left(1-T^{s}_{t+1}\right)\left(P^{s,\text{Gas}}_{t+1}-\tau^{iks,\text{Gas}}_{t+1}-MC^{iks,\text{Gas}}_{t+1}\right)\left(1-\vartheta^{iks}\right)^{-1}\frac{\partial^2 Maint^{i}_{t+1}(\cdot)}{\partial \mathrm{M}^{i\,2}_{t+1}}\right\}\right]^{-1}, \tag{6.68}$$

where the sign of the second part at the RHS of (6.68) is strictly negative. Thus, this implies that either $\frac{d\mathrm{M}^{iks}_{t+1}}{dVF^{iks}_{t+1}}=0$ or the sign of $\frac{d\mathrm{M}^{iks}_{t+1}}{dVF^{iks}_{t+1}}$ is the same as that of:

$$-\frac{\partial PU^{iks}_{t+1}(\cdot)}{\partial \mathrm{M}^{iks}_{t+1}} \tag{6.69}$$

which is strictly negative as long as $PU^{iks}_{t+1}(\cdot)$ is increasing in M^{iks}_{t+1}, Q.E.D.

Part (ii). Recall that $UVent^{i}_{t}(\cdot)=\vartheta^{iks}\text{TotGas}^{iks}_{t+1}-Maint^{iks}_{t}\left(\mathrm{M}^{iks}_{t}\right)+\epsilon^{iks}_{t}$. Thus, a marginal increase in VF^{iks}_{t} may affect the optimal choice of TotGas^{iks}_{t}, but at any given level of gas extracted $\text{TotGas}^{iks}_{t}=\overline{\text{TotGas}}^{iks}_{t}$, we obtain:

$$\frac{\partial UVent^{iks}_{t}\left(\overline{\text{TotGas}}^{iks}_{t},\mathrm{M}^{iks}_{t}\right)}{\partial VF^{iks}_{t}}=-\frac{\partial Maint^{iks}_{t}\left(\mathrm{M}^{iks}_{t}\right)}{\partial \mathrm{M}^{iks}_{t}}\frac{d\mathrm{M}^{iks}_{t}}{dVF^{iks}_{t}}\geq 0\,, \tag{6.70}$$

which is weakly positive because $Maint^{iks}_{t}$ is increasing in M^{iks}_{t} and because if $\frac{\partial PU^{iks}_{t}(\cdot)}{\partial \mathrm{M}^{iks}_{t}}\geq 0$, then we have $\frac{d\mathrm{M}^{iks}_{t}}{dVF^{iks}_{t}}\leq 0$ from part (i) of this proof. Q.E.D.

Proposition 3. *In each period t the tax reform $\check{\mathbf{T}}^{s}$ translates into (i) zero intentional venting; (ii) zero non-routine flaring; (iii) weakly lower unintentional venting for all oil&gas firms; (iv) no effect on the equilibrium level of all prices; (v) no effect on the equilibrium demand of intermediate goods O^{js}_{t}, G^{js}_{t}, and consumption goods $\mathbf{c}^{s}_{t}$; (v) strictly lower GHG emissions;, (vi) weakly larger aggregate present-discounted corporate profits, tax revenue, and net consumer income, and (vii) strictly larger social welfare.*

Proof. We begin with proving result (iv). Then we use result (iv) to prove that the results in parts (i), (ii), (iii), (v), (vi), and (vii).

Part (iv). First, recall that the first constraint is always binding (see proof to Proposition 1). Thus, $\phi^{iks}_{1t}>0$. Second, because the optimization problem is convex and satisfy the Slater's condition, the optimal solution satisfies the FOCs. Third, consider the constraint 2, which writes:

$$GOR^{iks}\text{Oil}^{iks}_{t}-\left(1-\vartheta^{iks}\right)^{-1}\left[\text{Gas}^{iks}_{t}+\text{Flare}^{iks}_{t}+\text{IVent}^{iks}_{t}-Maint^{iks}_{t}\left(\mathrm{M}^{iks}_{t}\right)+\text{ReInj}^{iks}_{t}+\text{ReInS}^{iks}_{t}\right]\leq 0 \tag{6.71}$$

we aim to show that this constraint is always binding for all oil fields and never binding for all gas fields.

Step 1. We prove that for an oil field constraint 2 is always binding at the optimal solution. This result ensures that an oil fields remains classified as such even after the tax reform is implemented. Proof. By assumption, a field classified as an oil field in period t satisfies

$$\begin{aligned} 0 < \left[P_t^{\text{Oil}}\left(1-\varsigma_t^{\text{Oil}} MS_t^{ks,\text{Oil}}\right)+\sigma_1^{iks}-\tau_t^{iks,\text{Oil}}-MC_t^{iks,\text{Oil}}\left(\text{Oil}_t^{iks}\right)\right] < \\ -\left[P_t^{s,\text{Gas}}-\tau_t^{iks,\text{Gas}}-MC_t^{iks,\text{Gas}}\right]\left(1-\vartheta^{iks}\right) GOR^{iks} \end{aligned} \quad (6.72)$$

for all $\text{Oil}_t^{iks} \geq 0$. Note that this condition is unaffected by the introduction of the tax reform at constant prices because $\Delta\tau_t^{iks,\text{Oil}} = -\Delta\tau_t^{s,\text{Gas}}\left(1-\vartheta^{iks}\right) GOR^{iks}$ implies that the condition is identical under $\check{\mathbf{T}}^s$ and $\mathbf{T}^s$ for any feasible value of Oil_t^{iks}. Suppose the constraint (6.71) is not binding in period t, therefore $\phi_{2t}^{iks} = 0$. In such case, the field is classified as an "other oil&gas field" if $\text{Oil}_t^{iks} > 0$ and faces the original tax scheme regardless of the implementation of the reform, or as a gas-only field if $\text{Oil}_t^{iks} = 0$. Suppose the optimal choice features $\text{Oil}_t^{iks} > 0$. Combining the FOCs w.r.t. Oil_t^{iks} and Gas_t^{iks} and setting $\phi_{2t}^{iks} = 0$ the optimality condition for oil production writes:

$$\begin{aligned} \left(1-T_t^{ks}\right)\left[P_t^{\text{Oil}}\left(1-\varsigma_t^{\text{Oil}} MS_t^{ks,\text{Oil}}\right)+\sigma_1^{iks}-\tau_t^{iks,\text{Oil}}-MC_t^{iks,\text{Oil}}\left(\text{Oil}_t^{i*}\right)\right] \\ -\left[P_t^{s,\text{Gas}}-\tau_t^{iks,\text{Gas}}-MC_t^{iks,\text{Gas}}\right]\left(1-T_t^{s}\right)\left(1-\vartheta^{iks}\right)^{-1}+\phi_{5t}^{iks}=0 \end{aligned} \quad (6.73)$$

Using condition (6.72) into the optimality condition (6.73) and noticing that it implies $P_t^{s,\text{Gas}} - \tau_t^{iks,\text{Gas}} - MC_t^{iks,\text{Gas}} \leq 0$, we obtain:

$$\left[P_t^{\text{Oil}}\left(1-\varsigma_t^{\text{Oil}} MS_t^{ks,\text{Oil}}\right)-\tau_t^{iks,\text{Oil}}-MC_t^{iks,\text{Oil}}\left(\text{Oil}_t^{iks*}\right)\right]\left[1+\left(GOR^{iks}\right)^{-1}\right]\left(1-T_t^{s}\right)+\phi_{5t}^{iks}<0 \quad (6.74)$$

Thus, because $\phi_{5t}^{iks*} \geq 0$, the inequality is satisfied only if

$$\left[P_t^{\text{Oil}}\left(1-\varsigma_t^{\text{Oil}} MS_t^{ks,\text{Oil}}\right)+\sigma_1^{iks}-\tau_t^{iks,\text{Oil}}-MC_t^{iks,\text{Oil}}\left(\text{Oil}_t^{iks*}\right)\right]\left(1-T_t^{s}\right)<0 \quad (6.75)$$

However, note that the FOC w.r.t. Oil_t^{iks} at the optimal vector writes:

$$\begin{aligned} \left[P_t^{\text{Oil}}\left(1-\varsigma_t^{\text{Oil}} MS_t^{ks,\text{Oil}}\right)+\sigma_1^{iks}-\tau_t^{iks,\text{Oil}}-MC_t^{iks,\text{Oil}}\left(\text{Oil}_t^{iks*}\right)\right]\left(1-T_t^{s}\right) \\ -\left(\phi_{3t}^{iks}+\phi_{4t}^{iks}\zeta\right)+\phi_{5t}^{iks}=0 \end{aligned} \;, \quad (6.76)$$

which implies that either $\phi_{5t}^{iks*} > 0$ and therefore $\text{Oil}_t^{iks*} = 0$, implying that the field would not be classified as an oil field for the purposes of the tax scheme, or if $\phi_{5t}^{iks*} = 0$, and therefore given that the K-T multipliers must satisfy $\phi_{3t}^{iks} + \phi_{4t}^{iks}\zeta \geq 0$ we obtain:

$$\left[P_t^{\text{Oil}}\left(1-\varsigma_t^{\text{Oil}} MS_t^{ks,\text{Oil}}\right)+\sigma_1^{iks}-\tau_t^{iks,\text{Oil}}-MC_t^{iks,\text{Oil}}\left(\text{Oil}_t^{iks*}\right)\right]\left(1-T_t^{s}\right) \geq 0\;. \quad (6.77)$$

Comparing inequalities (6.75) and (6.77) we find that for Oil_t^{iks*} to be optimal the left hand side of the inequality (6.75) should be both strictly negative and weakly positive. This leads to a contradiction. Lastly, suppose $\text{Oil}_t^{iks*} = 0$ and $\text{Gas}_t^{iks*} > 0$. $P_t^{\text{Gas}} + \sigma_2^{iks} - \tau_t^{iks,\text{Gas}} - MC_t^{iks,\text{Gas}} < 0$ implies that the firm is making negative profits in period t. Moreover, $\text{Oil}_t^{iks*} = 0$ implies that the second constraint is not binding. Thus, there exists an alternative feasible choice $\tilde{\mathbf{x}}^{ks} \in X^u$ such that (i) $\tilde{\text{Gas}}_t^{iks} = 0$ and (ii) all other choice variables are unchanged, which delivers strictly larger profit in period t and the same profit as $\mathbf{x}^{ks*}$ in all other periods $r \neq t$. Thus, the choice $\mathbf{x}^{ks*}$ cannot be optimal. This leads to a contradiction. Q.E.D.

Step 2a. For a "other oil&gas field" constraint 2 is always not binding at the optimal solution. Suppose it is binding, which implies $\phi_{2t}^{iks*} > 0$. Then combining the FOCs w.r.t. Gas_t^{iks} and Oil_t^{iks} we obtain:

$$\begin{gathered} -\left(1-T_t^s\right)\left[P_t^{\text{Oil}}\left(1-\varsigma_t^{\text{Oil}} MS_t^{ks,\text{Oil}}\right)+\sigma_1^{iks}-\tau_t^{iks,\text{Oil}}-MC_t^{iks,\text{Oil}}\left(\text{Oil}_t^{iks*}\right)\right]\left(1-\vartheta^{iks}\right) \\ +\left[P_t^{s,\text{Gas}}\left(1-\varsigma_t^{\text{Gas}} MS_t^{ks,\text{Gas}}\right)+\sigma_2^{iks}-\tau_t^{iks,\text{Gas}}-MC_t^{iks,\text{Gas}}\left(GOR^{iks}\text{Oil}_t^{iks}\right)\right]\left(1-T_t^s\right)= \\ -\phi_{2t}^{iks}\left(1+GOR^{iks}\right)\left(1-\vartheta^{iks}\right)+\phi_{5t}^{iks}\left(1-\vartheta^{iks}\right) \end{gathered} \tag{6.78}$$

But by assumption all “other oil&gas field” satisfy the following condition:

$$\begin{gathered} \left[P_t^{\text{Oil}}+\sigma_1^{iks}-\tau_t^{iks,\text{Oil}}-MC_t^{iks,\text{Oil}}\left(\text{Oil}_t^{iks}\right)\right]\left(1-\vartheta^{iks}\right) \leq \\ P_t^{s,\text{Gas}}\left(1-\varsigma_t^{\text{Gas}}\right)+\sigma_2^{iks}-\tau_t^{iks,\text{Gas}}-MC_t^{iks,\text{Gas}}\left(GOR^{iks}\text{Oil}_t^{iks}\right) \end{gathered} \tag{6.79}$$

for all $\text{Oil}_t^{iks} \geq 0$ implies:

$$-\phi_{2t}^{iks}\left(1+GOR^{iks}\right)\left(1-\vartheta^{iks}\right)+\phi_{5t}^{iks}\left(1-\vartheta^{iks}\right) \geq 0 \tag{6.80}$$

Thus, either (i) $\phi_{5t}^{iks*} > 0$ with $\text{Oil}_t^{iks*} = 0$ and given that we have stated that the constraint is binding, $\text{Gas}_t^{iks*} = 0$; i.e., the field is inactive, or (ii) $\phi_{5t}^{iks*} = 0$ with $\text{Oil}_t^{iks*} > 0$, then the inequality above is not satisfied. This leads to a contradiction.

Step 2b. For a gas-only field constraint 2 is always not binding at the optimal solution. Note that the assumption $MC_t^{iks,\text{Oil}}\left(\text{Oil}_t^{iks}\right) = +\infty$ ensures that $\text{Oil}_t^{iks*} = 0$ for all gas-only fields. Thus, the constraint 2 is always trivially non-binding for any active gas field; i.e., such that $\text{Gas}_t^{iks*} > 0$.

Step 3. The constraint 4 is always binding and therefore $\phi_{4t}^{iks} > 0$. Suppose the constraint 4 is not binding. Then, because of the complimentary slackness condition at the optimal choice we must have:

$$\begin{gathered} \text{K}_{t+1}^{iks}-\text{K}_t^{iks}-D_t^{iks}\left(\text{ID}_t^{iks},\text{L}_{t-1}^{iks}\right)-B_t^{iks}\left(\text{ReInj}_t^{iks}+\text{PInj}_t^{iks}+\text{OInj}_t^{iks}\right) \\ +\zeta\left\{\text{Oil}_t^{iks}+\left(1-\vartheta^{iks}\right)^{-1}\left[\text{Gas}_t^{iks}+\text{Flare}_t^{iks}\right.\right. \\ \left.\left.+\text{IVent}_t^{iks}-Maint_t^{iks}\left(\text{M}_t^{iks}\right)+\text{ReInj}_t^{iks}+\text{ReInS}_t^{iks}\right]\right\}<0 \end{gathered} \tag{6.81}$$

Because B_t^{iks} is strictly increasing, this implies that a choice vector $\tilde{\mathbf{x}}^{ks}$ identical to the optimal one $\mathbf{x}^{ks*}$ except for $\widetilde{\text{OInj}}_t^i = \text{OInj}_t^{i*} - a$ for arbitrarily small strictly positive a is feasible. This alternative choice vector is feasible and delivers exactly the same profit from field i in all periods $s \neq t$ and strictly larger profit in period t, and it does not affect the profit generated by fields other than i. This implies that the total present discounted value of profits from choosing vector $\tilde{\mathbf{x}}^{ks}$ are strictly larger that those from $\mathbf{x}^{ks*}$. Thus, the original choice vector $\mathbf{x}^{ks*}$ cannot be optimal. This leads to a contradiction.

Step 4. For an oil field the problem can be split into two independent optimization problems. Recall from Step 1 that the second constraint is always binding for oil fields, implying that the corresponding KT multiplier satisfies $\phi_{2t}^{iks} > 0$. Moreover, from Step 4 we know that the fourth K-T multiplier satisfies $\phi_{4t}^{iks} > 0$. Using these result and combining together the FOCs w.r.t. Gas_t^{iks}, and OInj_t^{iks}, Z_t^{iks}, we obtain the following conditions:

$$\begin{gathered} GOR^{iks}\text{Oil}_t^{iks}-\left(1-\vartheta^{iks}\right)^{-1}\left[\text{Gas}_t^{iks}+\text{Flare}_t^{iks}+\text{IVent}_t^{iks}\right. \\ \left.-Maint_t^{iks}\left(\text{M}_t^{iks}\right)+\text{ReInj}_t^{iks}+\text{ReInS}_t^{iks}\right]=0 \\ \left[P_t^{s,\text{Gas}}-\tau_t^{s,\text{Gas}}-MC_t^{iks,\text{Gas}}\right]\left(1-T_t^s\right)+\left(\phi_{2t}^{iks}-\phi_{3t}^{iks}-\phi_{4t}^{iks}\zeta\right)\left(1-\vartheta^{iks}\right)^{-1}=0 \\ C_t^{iks}\left(1-T_t^s\right)=\phi_{4t}^{iks}\frac{\partial B_t^{iks}(\cdot)}{\partial \text{OInj}_t^i} \\ \left(1-T_t^s\right)=\phi_{1t}^{iks}\frac{\partial TF_t^{iks}(.)}{\partial \text{Z}_t^{iks}} \end{gathered} \tag{6.82}$$

Substitute these four conditions into the FOCs. We obtain the following equilibrium conditions, divided into three subsets, which correspond to three *reduced problems*.

Reduced Problem 1. Define the collections of endogenous variables $y1_t^{iks} = \{\text{Flare}_t^{iks}, \text{IVent}_t^{iks}, \text{ReInj}_t^{iks}, \text{PInj}_t^{iks}, \text{ReInS}_t^{iks}, \text{PIn}$ $\phi1_t^{iks} = \{\phi_{8t}^{iks}, \phi_{9t}^{iks}, \phi_{10t}^{iks}, \phi_{11t}^{iks}, \phi_{12t}^{iks}\}$, and $\lambda1_t^{iks} = \lambda_{2t}^{iks}$. Also define the sets $Y1^{ks}$ with typical element $y1^{ks} = \{\{y1_t^{iks}\}_{t=1}^{\infty}\}_{i=1}^{I^k}$, $\Psi1^{ks}$ with typical element $\psi1^{ks} = \{\{\psi1_t^{iks}\}_{t=1}^{\infty}\}_{i=1}^{I^k}$, and $\Lambda1^{ks}$ with typical element $\lambda1^{ks} = \{\{\lambda1_t^{iks}\}_{t=1}^{\infty}\}_{i=1}^{I^k}$. Note that $\mathbf{x}^{ks} \in X^{ks} \rightarrow y1^{ks} \in Y1^{ks}$. Also note that the inequality constraints 8, 9, 10, 11, 12 can be redefined as functions $g1_{jt}^{iks}$ such that $g1_{jt}^{iks}\left(y1_t^{iks}\right) \geq 0$ if and only if $g_{wt}^{j,m}\left(\mathbf{x}^{ks}\right) \leq 0$ given an appropriate choice of the index j. Consider the following subset of $24 \times I^k \times T$ conditions:

$$
\begin{aligned}
F1_{1t}^{iks} &= -VF_t^{iks} \times \frac{PF_t^{iks}(\cdot)}{\partial \text{Flare}_t^{iks}} - (1-T_t^s)\Big[P_t^{s,\text{Gas}} - \tau_t^{iks,\text{Gas}} \\
&\quad -MC_t^{iks,\text{Gas}} + \tau_t^{iks,\text{Flare}} + MC_t^{iks,\text{Flare}}\left(\text{Flare}_t^{iks}\right)\Big] + \phi_{8t}^{iks} = 0 \\
F1_{2t}^{iks} &= -VF_t^{iks} \times \frac{PF_t^{iks}(\cdot)}{\partial \text{IVent}_t^{iks}} - (1-T_t^s)\left(P_t^{s,\text{Gas}} - \tau_t^{s,\text{Gas}} - MC_t^{iks,\text{Gas}}\right) + \phi_{9t}^{iks} = 0 \\
F1_{3t}^{iks} &= -\left[\frac{IC_t^{iks}\left(\text{ReInj}_t^{iks}, \text{PInj}_t^{iks}\right)}{\partial \text{PInj}_t^{iks}} + PP_t^{iks,\text{Gas}}\right](1-T_t^s) + C_t^{iks}(1-T_t^s) = 0 \\
F1_{4t}^{iks} &= \left[-\frac{IC_t^{iks}\left(\text{ReInj}_t^{iks}, \text{PInj}_t^{iks}\right)}{\partial \text{ReInj}_t^{iks}} + \delta_{1t}^{iks}\right](1-T_t^s) \\
&\quad -(1-T_t^s)\left(P_t^{s,\text{Gas}} - \tau_t^{s,\text{Gas}} - MC_t^{iks,\text{Gas}} - C_t^{iks}\right) + \phi_{10t}^{iks} = 0 \\
F1_{5t}^{iks} &= -(1-T_t^s)\left(P_t^{s,\text{Gas}} - \tau_t^{s,\text{Gas}} - MC_t^{iks,\text{Gas}} - MP_t^{iks,\text{ReIns}}\left(\text{ReInS}_t^{iks}, \text{PInS}_t^{iks}\right) - \delta_{1t}^{iks}\right) + \phi_{11t}^{iks} = 0 \\
F1_{6t}^{iks} &= (1-T_t^s)\left[MP_t^{iks,\text{PIns}}\left(\text{ReInS}_t^{iks}, \text{PInS}_t^{iks}\right) - PP_t^{iks,\text{Gas}}\right] = 0 \\
F1_{7t}^{iks} &= -(1-T_t^s) + \lambda_{2t}^{iks} = 0 \\
F1_{8t}^{iks} &= -(1-T_t^s)\beta MP_{t+1}^{iks,\text{M}}\left(\text{M}_{t+1}^{iks}\right) - \lambda_{2t}^{iks} + \beta\lambda_{2t+1}^{iks}\left(1-\rho^{iks}\right) \\
&\quad -\beta\left[P_{t+1}^{s,\text{Gas}} - \tau_{t+1}^{s,\text{Gas}} - MC_{t+1}^{iks,\text{Gas}}\right]\frac{\partial Maint_{t+1}^{iks}\left(\text{M}_{t+1}^{iks}\right)}{\partial \text{M}_{t+1}^{iks}} + \beta\phi_{12t}^{iks} = 0 \\
E1_{1t}^{iks} &= \text{M}_{t+1}^{iks} - \text{M}_t^{iks}\left(1-\rho^{iks}\right) - \text{IM}_t^{iks} = 0 \\
I1_{1t}^{iks} &= \text{Flare}_t^{iks} - \text{NRF}_t^{iks} \geq 0 \\
I1_{2t}^{iks} &= \text{IVent}_t^{iks} \geq 0 \\
I1_{3t}^{iks} &= \text{ReInj}_t^{iks} \geq 0 \\
I1_{4t}^{iks} &= \text{ReInS}_t^{iks} \geq 0 \\
I1_{5t}^{iks} &= \text{M}_{t+1}^{iks} \geq 0 \\
M1_{jt}^{iks} &= \phi1_{jt}^{iks} \geq 0,\ j = 1,2,3,4,5 \\
S1_{jt}^{iks} &= \phi1_{jt}^{iks} g1_{jt}^{iks}\left(y1_t^{iks}\right) \geq 0\ j = 1,2,3,4,5
\end{aligned}
$$

In the system above, we have 24 conditions per period t per field i. However, making use of the complimentary slackness conditions $S1_{jt}^{iks}$ for $j = 1,2,3,4,5$ we are left with 14 binding conditions per period t per field i. We have 14 unknown variables per period t and per field i: 8 in collection $y1_t^{iks}$, 5 in $\psi1_t^{iks}$, plus $\lambda1_t^{iks}$, all independent of the realizations of the variables outside of $Y1^{ks} \times \Psi1^{ks} \times \Lambda1^{ks}$. Thus the problem can be solved independently of the optimality of the other variables in the full problem. As we have shown that a solution to the full optimization problem exists, then a solution to the reduced problem must also exist. Thus, it must be the collection $\{y1^{ks*}, \psi1^{ks}, \lambda1^{ks}\}$ that solve this system.

Reduced Problem 2. Define the variable total gas injection as follows: $\text{TotInj}_t^{iks} = \text{ReInj}_t^{iks} + \text{PInj}_t^{iks} +$

OInj_t^{iks}. Then define the collections of endogenous variables $y2_t^{iks} = \left\{\text{Oil}_t^{iks}, \text{ID}_t^{iks}, \text{TotInj}_t^{iks}, \text{L}_t^{iks}, \text{K}_{t+1}^{iks}\right\}$, $\phi 2_t^{iks} = \left\{\phi_{3t}^{iks}, \phi_{4t}^{iks}, \phi_{7t}^{iks}, \phi_{13t}^{iks}, \phi_{14t}^{iks}\right\}$, and $\lambda 2_t^{iks} = \lambda_{1t}^{iks}$.. Also define the sets $Y2^{ks}$ with typical element $y2^{ks} = \left\{\left\{y1_t^{iks}\right\}_{t=1}^{\infty}\right\}_{i=1}^{I^k}$, $\Psi 2^{ks}$ with typical element $\phi 2^{ks} = \left\{\left\{\psi 1_t^{iks}\right\}_{t=1}^{\infty}\right\}_{i=1}^{I^k}$, and $\Lambda 2^{ks}$ with typical element $\lambda 2^{ks} = \left\{\left\{\lambda 2_t^{iks}\right\}_{t=1}^{\infty}\right\}_{i=1}^{I^k}$. Note that $\mathbf{x}^{ks} \in X^{ks} \rightarrow y2^{ks} \in Y2^{ks}$. Also note that the inequality constraints 3, 4, 7, 13, 14 can be redefined as functions $g2_{jt}^{iks}$ such that $g2_{jt}^{iks}\left(y2_t^{iks}\right) \geq 0$ if and only if $g_{wt}^{j,m}\left(\mathbf{x}^k\right) \leq 0$ given an appropriate choice of the index j. Consider the following subset of $21 \times I^k \times T$ conditions:

$$
\begin{aligned}
F2_{1t}^{iks} &= \left[P_t^{\text{Oil}}\left(1-\varsigma_t^{\text{Oil}} MS_t^{ks,\text{Oil}}\right)+\sigma_1^{iks}-\tau_t^{iks,\text{Oil}}-MC_t^{iks,\text{Oil}}\left(\text{Oil}_t^{iks}\right)\right]\left(1-T_t^s\right) \\
&+\left[P_t^{s,\text{Gas}}-\tau_t^{iks,\text{Gas}}-MC_t^{iks,\text{Gas}}\right]\left(1-T_t^s\right)\ GOR^{iks}\left(1-\vartheta^{iks}\right) \\
&-\left(\phi_{3t}^{iks}+\phi_{4t}^{iks}\zeta\right)\left[1+GOR^{iks}\right]+\phi_{5t}^{iks}=0 \\
F2_{2t}^{iks} &= -\phi_{4t}^{iks}+\beta\phi_{3t+1}^{iks}+\beta\phi_{4t+1}^{iks}+\phi_{13t}^{iks}=0 \\
F2_{3t}^{iks} &= -\left(1-T_t^s\right)+\phi_{4t}^{iks}\frac{\partial D_t^{iks}\left(\text{ID}_t^{iks},\text{L}_{t-1}^{iks}\right)}{\partial \text{ID}_t^i}-\lambda_{1t}^{iks}=0 \\
F2_{4t}^{iks} &= \lambda_{1t}^{iks}+\beta\phi_{4t+1}^{iks}\frac{\partial D_{t+1}^{iks}\left(\text{ID}_{t+1}^{iks},\text{L}_t^{iks}\right)}{\partial \text{L}_t^i}-\beta\lambda_{1t+1}^{iks}=0 \\
F2_{5t}^{iks} &= C_t^{iks}\left(1-T_t^s\right)-\phi_{4t}^{iks}\frac{\partial B_t^{iks}\left(\text{TotInj}_t^{iks}\right)}{\partial \text{TotInj}_t^{iks}}=0 \\
E2_{1t}^{iks} &= \text{L}_t^{iks}-\text{L}_{t-1}^{iks}-\text{ID}_t^{iks}=0 \\
I2_{1t}^{iks} &= \text{Oil}_t^{iks}\left(1+GOR^{iks}\right)-\text{K}_t^{iks}\leq 0 \\
I2_{2t}^{iks} &= \text{K}_{t+1}^{iks}-\text{K}_t^{iks}-D_t^{iks}\left(\text{ID}_t^{iks},\text{L}_{t-1}^{iks}\right)-B_t^{iks}\left(\text{TotInj}_t^{iks}\right)+\zeta\text{Oil}_t^{iks}\left(1+GOR^{iks}\right)\leq 0 \\
I2_{3t}^{iks} &= \text{Oil}_t^{iks}\geq 0 \\
I2_{4t}^{i,k} &= \text{L}_t^{iks}\geq 0 \\
I2_{5t}^{iks} &= \text{K}_{t+1}^{iks}\geq 0 \\
M2_{jt}^{iks} &= \phi 2_{jt}^{iks}\geq 0,\ j=1,2,3,4,5 \\
S2_{jt}^{iks} &= \phi 2_{jt}^{iks} g2_{jt}^{iks}\left(y2_t^{iks}\right)\geq 0\ j=1,2,3,4,5
\end{aligned}
$$

In the system above, we have 21 unknown variables per period t and per field i. However, making use of the complimentary slackness conditions $S2_{jt}^{iks}$ for $j = 1, 2, 3, 4, 5$ we are left with 11 binding conditions per period t per field i. We have 11 unknown variables per period t and per field i: 5 in collection $y2_t^{ks}$, 5 in $\psi 2_t^{ks}$, plus $\lambda 2_t^{ks}$, all independent of the realizations of the variables outside of $Y2^{ks} \times \Psi 2^{ks} \times \Lambda 2^{ks}$. Thus the problem can be solved independently of the optimality of the other variables in the full problem. As we have shown that a solution to the full optimization problem exists, then a solution to the reduced problem must also exist. Thus, it must be the collection $\left\{y2^{ks*}, \psi 2^{ks*}, \lambda 2^{ks*}\right\}$ that solve this system. In particular, note that given that the change in the tax rate on oil production satisfies:

$$
\Delta\tau_t^{iks,\text{Oil}} = -GOR^{iks}\left(1-\vartheta^{iks}\right)\Delta\tau_t^{s,\text{Gas}} \tag{6.83}
$$

then all the equilibrium conditions of reduced problem 2 are unaffected by the tax reform. Therefore, the optimal levels of all endogenous variables in $y2_t^{ks}$ $\psi 2_t^{ks}$, and $\lambda 2_t^{ks}$ for all $t = 1, 2, ...$ – including Oil_t^{iks}– are unaffected by the policy change from $\mathbf{T}^s$ to $\check{\mathbf{T}}^s$.

Reduced problem 3. Lastly, the optimal values of Z_t^{iks}, Gas_t^{iks}, ϕ_{1t}^{iks}, ϕ_{2t}^{iks} can be pinned down using the optimal values of the other endogenous variables from $y1^{ks*}, \phi1^{ks*}$ and $y2^{ks*}, \phi2^{ks*}$, plus the four conditions:

$$\begin{array}{ll} F3_{1t}^{iks} = & \left[P_t^{s,\text{Gas}} - \tau_t^{s,\text{Gas}} - MC_t^{iks,\text{Gas}}\right](1-T_t^s) + \left(\phi_{2t}^{iks} - \phi_{3t}^{iks*} - \phi_{4t}^{iks*}\zeta\right)\left(1-\vartheta^{iks}\right)^{-1} = 0 \\ F3_{2t}^{iks} = & = (1-T_t^s) - \phi_{1t}^{iks}\frac{\partial TF_t^{iks}(.)}{\partial \text{Z}_t^{iks}} = 0 \\ I3_{1t}^{iks} = & TF_t^{iks}(\text{Oil}_t^{iks*}, \text{Gas}_t^{iks}, \text{Flare}_t^{iks*}, \text{ReInS}_t^{iks*}, \text{PInS}_t^{iks*}, \text{M}_t^{iks*}, \text{Z}_t^{iks}) \leq 0 \\ I3_{2t}^{iks} = & GOR^{iks}\text{Oil}_t^{iks*} - \left(1-\vartheta^{iks}\right)^{-1}\left[\text{Gas}_t^{iks} + \text{Flare}_t^{iks*} + \text{IVent}_t^{iks*}\right. \\ & \left. -Maint_t^{iks}\left(GOR^{iks}\text{Oil}_t^{iks*}, \text{M}_t^{iks*}\right) + \text{ReInj}_t^{iks*} + \text{ReInS}_t^{iks*} \leq 0\right. \end{array},$$

which are all binding. Specifically, we have shown that $I3_{2t}^{iks}$ is binding in Step 1 of this proof, and that $I3_{1t}^{iks}$ is binding in the proof to Proposition 1. Thus, after substituting the optimal values of $y1^{ks*}, \phi1^{ks*}$ and $y2^{ks*}, \phi2^{ks*}$ from the corresponding reduced problems into conditions $F3_{1t}^{iks}$, $F3_{2t}^{iks}$, $I3_{1t}^{iks}$ and $I3_{2t}^{iks}$ and set the conditions to hold with strict equality, we can solve for Z_t^{iks} and Gas_t^{iks} from conditions $I3_{1t}^{iks}$ and $F3_{2t}^{iks}$, and for ϕ_{2t}^{iks} from condition $F3_{1t}^{iks}$, for all fields $i = 1, 2, ..., I^k$ and all periods $t = 1, 2, ...$

Step 5. The oil firms' field-level and total optimal oil supply is unaffected by the tax reform at constant global oil price P_t^{Oil},constant gas price $P_t^{s,\text{Gas}}$, and constant aggregate oil supply from firms other than k. Suppose the aggregate oil supply from firms other than k; i.e., $\sum_{l\neq k}\sum_{i=1}^{I^l} \text{Oil}_t^{ils}$, is unaffected by the tax reform. Then recall that the formula for the tax reform implies $\Delta\tau_t^{iks,\text{Oil}} GOR^{iks}\left(1-\vartheta^{iks}\right) = -\Delta\tau_t^{s,\text{Gas}}$. Thus, the formula for the optimal positive oil production for field i in period t, which writes:

$$\begin{array}{c} \left[P_t^{\text{Oil}}\left(1-\varsigma_t^{\text{Oil}} MS_t^{k,\text{Oil}}\right) + \sigma_1^{iks} - \tau_t^{iks,\text{Oil}} - MC_t^{iks,\text{Oil}}\left(\text{Oil}_t^i\right)\right](1-T_t^s) \\ + \left[P_t^{s,\text{Gas}} - \tau_t^{s,\text{Gas}} - MC_t^{iks,\text{Gas}}\right](1-T_t^s)\ GOR^{iks}\left(1-\vartheta^{iks}\right) \\ - \left(\phi_{3t}^{iks} + \phi_{4t}^{iks}\zeta\right)\left[1 + GOR^{iks}\right] + \phi_{5t}^{iks} = 0 \end{array}$$

is unaffected by the reform. The tax reform does not enter any other condition of the reduced problem 2, implying in turn that the optimal values of $\text{Oil}_t^{iks}, \text{ID}_t^{iks}, \text{TotInj}_t^{iks}, \text{L}_t^{iks}, \text{K}_{t+1}^{iks}, \phi_{3t}^{iks}, \phi_{4t}^{iks}$ are unchanged for all $i = 1, 2, .., I^k$ and all $t = 1, 2, ...$ As a result, the aggregate oil supply from all the fields owned by firm k in period t; i.e., $\sum_{i=1}^{I^k} \text{Oil}_t^{iks}$ is also unchanged, as well as $MS_t^{ks,\text{Oil}}$, which is a function of $\sum_{i=1}^{I^k} \text{Oil}_t^{iks}$, of the aggregate oil supply from firms other than k and of the oil price P_t^{Oil}. Note that the restrictions imposed on oil fields imply that oil production is always profitable, i.e., the constraint $\text{Oil}_t^{iks} \leq 0$ is never binding. That is, an oil field hit the zero production point only if $\text{K}_t^{iks} = 0$; i.e., the field does not have any residual capacity because the net return to investments in discoveries and/or injections have become weakly negative.

Step 5/b. First, because each oil firm's optimal oil supply is unaffected by the tax reform at constant global oil price P_t^{Oil},constant gas price $P_t^{s,\text{Gas}}$, and constant aggregate oil supply from firms other than k, then the global oil supply is also unchanged at any given global oil price P_t^{Oil}. Second, the new tax on gas fields ensures that the aggregate gas supply of country s is unchanged, ensuring that the national and the global natural gas supply are also unchanged at any given national gas prices $\left\{P_t^{s,\text{Gas}}\right\}_{s=1}^S$. Third, because the demand for all goods is unaffected by the tax reform then the equilibrium prices must be unchanged too. Under the assumptions imposed on B^{iks}, if $\text{PInS}_t^{iks} = 0$ we get for gas only fields $MP_t^{iks,\text{ReInS}}\left(\text{ReInS}_t^{iks}\right) = MP_t^{iks,\text{ReInS}} > 0$. Let us consider fields such that $\text{ReInS}_t^{iks} > 0$, we obtain that the production of gas-only fields solves:

$$\begin{aligned}
\frac{\partial \mathcal{L}_u^{ks}}{\partial \text{Gas}_t^{iks}} &= \left[P_t^{s,\text{Gas}}\left(1-\varsigma_t^{\text{Gas}} MS_t^{ks,\text{Gas}}\left(\text{Gas}_t^{iks}\right)\right)-\tau_t^{iks,\text{Gas}}\right](1-T_t^s) \\
&\quad -\phi_{1t}^{iks}\frac{\partial TF_{1t}^{iks}(.)}{\partial \text{Gas}_t^{i}}-\left(\phi_{3t}^{iks}+\phi_{4t}^{iks}\zeta\right)\left(1-\vartheta^{iks}\right)^{-1}=0 \\
\frac{\partial \mathcal{L}_u^{ks}}{\partial \text{ReInS}_t^{iks}} &= \phi_{1t}^{iks}\frac{\partial TF_{4t}^{iks}(.)}{\partial \text{ReInS}_t^{iks}}-\left(\phi_{3t}^{iks}+\phi_{4t}^{iks}\zeta\right)\left(1-\vartheta^{iks}\right)^{-1}+\phi_{11t}^{iks}=0 \\
\frac{\partial \mathcal{L}_u^{ks}}{\partial \text{Z}_t^{iks}} &= (1-T_t^s)-\phi_{1t}^{iks}\frac{\partial TF_t^{iks}(.)}{\partial \text{Z}_t^{iks}}=0 \\
\frac{\partial \mathcal{L}_u^{ks}}{\partial \text{K}_{t+1}^{iks}} &= -\phi_{4t}^{i}+\beta\phi_{3t+1}^{i}+\beta\phi_{4t+1}^{i}+\phi_{14t}^{i}=0 \\
\frac{\partial \mathcal{L}_u^{ks}}{\partial \text{ID}_t^{i}} &= -(1-T_t^s)+\phi_{4t}^{iks}\frac{\partial D_t^{iks}\left(\text{ID}_t^{iks},\text{L}_{t-1}^{iks}\right)}{\partial \text{ID}_t^{i}}-\lambda_{1t}^{iks}=0 \\
\frac{\partial \mathcal{L}_u^{ks}}{\partial \text{L}_t^{i}} &= \lambda_{1t}^{iks}+\beta\phi_{4t+1}^{iks}\frac{\partial D_{t+1}^{iks}\left(\text{ID}_{t+1}^{iks},\text{L}_t^{iks}\right)}{\partial \text{L}_t^{i}}-\beta\lambda_{1t+1}^{iks}+\phi_{13t}^{iks}=0 \\
\frac{\partial \mathcal{L}_u^{ks}}{\partial \text{IM}_t^{iks}} &= -\left(1-T_t^{iks}\right)+\lambda_{2t}^{iks}=0 \\
\frac{\partial \mathcal{L}_u^{ks}}{\partial \text{M}_{t+1}^{iks}} &= -\phi_{1t+1}^{iks}\frac{\partial TF_{5t+1}^{iks}(.)}{\partial \text{M}_{t+1}^{iks}}-\lambda_{2t}^{iks}+\beta\lambda_{2t+1}^{iks}\left(1-\rho^{iks}\right)+\phi_{12t}^{iks} \\
&\quad -\beta\left[\phi_{3t+1}^{iks}+\phi_{4t+1}^{iks}\zeta+\delta_{1t+1}^{iks}\right]\frac{\partial UVent_{t+1}^{iks}(.)}{\partial \text{M}_{t+1}^{iks}}=0 \\
EG_{1t}^{iks} &= \text{L}_t^{iks}-\text{L}_{t-1}^{iks}-\text{ID}_t^{iks}=0 \\
EG_{2t}^{iks} &= \text{M}_{t+1}^{iks}-\text{M}_t^{iks}\left(1-\rho^{iks}\right)-\text{IM}_t^{iks}=0 \\
IG_{1t}^{iks} &= TF_t^{iks}(0,\text{Gas}_t^{iks},\text{NRF}_t^{iks},\text{ReInS}_t^{iks},0,\text{M}_t^{iks},\text{Z}_t^{iks})\le 0 \\
IG_{2t}^{iks} &= \left(1-\vartheta^{iks}\right)^{-1}\left[\text{Gas}_t^{iks}+\text{Flare}_t^{iks}+\text{IVent}_t^{iks}-Maint_t^{iks}\left(\text{M}_t^{iks}\right)+\text{ReInS}_t^{iks}\right]-\text{K}_t^{iks}\le 0 \\
IG_{3t}^{iks} &= \text{K}_{t+1}^{iks}-\text{K}_t^{iks}-D_t^{iks}\left(\text{ID}_t^{iks},\text{L}_{t-1}^{iks}\right) \\
&\quad +\zeta\left(1-\vartheta^{iks}\right)^{-1}\left[\text{Gas}_t^{iks}+\text{Flare}_t^{iks}+\text{IVent}_t^{iks}-Maint_t^{iks}\left(\text{M}_t^{iks}\right)+\text{ReInS}_t^{iks}\right]\le 0
\end{aligned}$$

Plus $\phi_{1t}^{iks}\ge 0,\phi_{3t}^{iks}\ge 0,\phi_{4t}^{iks}\ge 0,\phi_{9t}^{iks}\ge 0,\phi_{10t}^{iks}\ge 0,\phi_{11t}^{iks}\ge 0,\phi_{12t}^{iks}\ge 0$ and the usual complementary slackness conditions. Using these FOCs, the optimality condition for gas production writes:

$$\left[P_t^{s,\text{Gas}}\left(1-\varsigma_t^{\text{Gas}} MS_t^{ks,\text{Gas}}\left(\text{Gas}_t^{iks}\right)\right)-\tau_t^{iks,\text{Gas}}-MC_t^{iks,\text{Gas}}\left(\text{Gas}_t^{iks}\right)-MP_t^{iks,\text{ReInS}}\right](1-T_t^s)=0 \tag{6.84}$$

which can be differentiated to obtain the marginal effect of an increase in the linear tax rate on gas production:

$$\begin{aligned}
\frac{\partial \text{Gas}_t^{iks}}{\partial \tau_t^{s,\text{Gas}}} &= -\left[P_t^{s,\text{Gas}}\frac{\partial \varsigma_t^{\text{Gas}} MS_t^{ks,\text{Gas}}\left(\text{Gas}_t^{iks}\right)}{\partial \text{Gas}_t^{iks}}+\frac{\partial MC_t^{iks,\text{Gas}}\left(\text{Gas}_t^{iks}\right)}{\partial \text{Gas}_t^{iks}}\right]^{-1} \\
&= -\frac{\partial \text{Gas}_t^{iks}}{\partial P_t^{s,\text{Gas}}}\left(1-\varsigma_t^{\text{Gas}} MS_t^{ks,\text{Gas}}\right)^{-1}
\end{aligned} \tag{6.85}$$

Thus, the effect of a change in the linear tax rate on gas production on the gas supply of gas-only filed i writes:

$$\Delta \text{Gas}_t^{iks}\simeq -\min\left\{\frac{\partial \text{Gas}_t^{iks}}{\partial P_t^{s,\text{Gas}}}\left(1-\varsigma_t^{\text{Gas}} MS_t^{ks,\text{Gas}}\right)^{-1}\Delta\tau_t^{iks,\text{Gas}},\text{Gas}_t^{iks}\right\} \tag{6.86}$$

Because the tax adjustment is small, we can approximate the result by assuming that for all active gas-only fields the the corrective tax affects production only at the intensive margin. Thus, after setting $\Delta\tau_t^{iks,\text{Gas}}=$

$\Delta\tau_t^{s,\text{Gas}}\left(1-\varsigma_t^{\text{Gas}}MS_t^{ks,\text{Gas}}\right)$ we obtain a formula for the effect of the corrective tax, which writes:

$$\sum_{k=1}^{K^s}\sum_{i=1}^{I^k}\Delta\text{Gas}_t^{iks}\mathbf{1}\left[\text{Oil}_t^{iks}=0\right]\simeq-\eta_{\text{G},P}^{s,GO}\left(\sum_{k=1}^{K^s}\sum_{i=1}^{I^k}\text{Gas}_t^{iks}\mathbf{1}\left[\text{Oil}_t^{iks}=0\right]\right)\frac{\Delta\tau_t^{s,\text{Gas}}}{P_t^{s,\text{Gas}}} \tag{6.87}$$

where $\eta_{\text{G},P}^{s,GO}$ denotes the aggregate own-price elasticity of gas supply by gas-only fields. Also note that the optimal values of maintenance is unaffected, because the equilibrium condition for M_{t+1}^{iks} imply that either $\phi_{12t}^{i}>0$ and M_{t+1}^{iks}, such that the optimal level of M_{t+1}^{iks} is unaffected by marginal changes in $\tau_t^{iks,\text{Gas}}$, or $\phi_{12t}^{iks}=0$ and the optimality condition writes:

$$\left(1-T_t^s\right)\left[MP_{t+1}^{iks,\text{M}}\left(\text{M}_{t+1}^{iks}\right)-1+\beta\left(1-\rho^{iks}\right)\right]+\beta\left(1-T_t^i\right)\frac{MP_t^{i,\text{ReInS}}}{1+\vartheta^i}\frac{\partial Maint_{t+1}^{iks}(.)}{\partial\text{M}_{t+1}^{iks}}=0$$

which is constant in $\tau_t^{iks,\text{Gas}}$. Lastly, we need to ensure that the tax is set to a level such that the reduced gas supply from gas-only fields exactly offsets the increased supply by gas fields. Thus, the optimal corrective tax solves:

$$\sum_{k=1}^{K^s}\sum_{i=1}^{I^k}\Delta\text{Gas}_t^{iks}\mathbf{1}\left[\text{Oil}_t^{iks}=0\right]=-\sum_{k=1}^{K^s}\sum_{i=1}^{I^k}\Delta\text{Gas}_t^{iks}\mathbf{1}\left[\text{Oil}_t^{iks}>0\right] \tag{6.88}$$

where ΔGas_t^{iks} is the change in gas supply by field i. The above equation solves for:

$$\frac{\Delta\tau_t^{iks,\text{Gas}}}{P_t^{s,\text{Gas}}}=\frac{1-\varsigma_t^{\text{Gas}}MS_t^{ks,\text{Gas}}}{\eta_{\text{G},P}^{s,GO}}\frac{\sum_{k=1}^{K^s}\sum_{i=1}^{I^k}\Delta\text{Gas}_t^{iks}\mathbf{1}\left[\text{Oil}_t^{iks}>0\right]}{\sum_{k=1}^{K^s}\sum_{i=1}^{I^k}\text{Gas}_t^{iks}\mathbf{1}\left[\text{Oil}_t^{iks}=0\right]} \tag{6.89}$$

which represents the optimal adjustment tax,; i.e., one that ensures zero change in the aggregate country-level and global natural gas supply.

$$AS_t^{r,\text{Gas}}\left(\mathbf{P};\mathbf{T}\right)=\sum_{k=1}^{K^r}\sum_{i=1}^{I^k}\text{Gas}_t^{ikr*}\left(\mathbf{P};\mathbf{T}\right) \tag{6.90}$$

Thus, given that the tax reform is such that $\Delta\tau_t^{s,\text{Gas}}$ solves equation (6.89), the aggregate supply of gas $AS_t^{r,\text{Gas}}\left(\mathbf{P};\mathbf{T}^r\right)$ by each country $r=1,2,...,S$ – including $r=s$ – is unaffected by the tax scheme at any given price collection $\mathbf{P}=\{\mathbf{P}_t\}_{t=1}^{\infty}$; i.e.,

$$AS_t^{r,\text{Gas}}\left(\mathbf{P};\mathbf{T}\right)=AS_t^{r,\text{Gas}}\left(\mathbf{P};\check{\mathbf{T}}\right)$$

Step 6. Let $\text{Gas}_t^{ikr*}\left(\mathbf{P};\mathbf{T}^r\right)$ and $\text{Oil}_t^{ikr*}\left(\mathbf{P};\mathbf{T}^r\right)$ denote the supply of natural gas and oil by field i in period t, respectively. Under the proposed tax scheme, we get that in each period t the following results hold true:

1. The aggregate supply of gas $AS_t^{r,\text{Gas}}\left(\mathbf{P};\mathbf{T}^r\right)$ by each country $r=1,2,...,S$ – including $r=s$ – is unaffected by the tax scheme at any given price collection $\mathbf{P}=\{\mathbf{P}_t\}_{t=1}^{\infty}$ by Step 5 bis; i.e.,

$$AS_t^{r,\text{Gas}}\left(\mathbf{P};\mathbf{T}\right)=AS_t^{r,\text{Gas}}\left(\mathbf{P};\check{\mathbf{T}}\right)$$

   for all $r=1,2,...,S$ and all $t=1,2,....$

2. The supply of oil by each field i in country s is unaffected by the tax scheme at any given price collection$\mathbf{P}=\{\mathbf{P}_t\}_{t=1}^{\infty}$ by Step 5. Moreover, the supply of oil by each field i in country $r\neq s$ is also unaffected by the tax scheme implemented by country s, except possibly through price effects. Thus, at any given price collection $\mathbf{P}$. Thus, the supply of oil by each field i in country $r\neq s$ is unaffected by the tax scheme at any given price collection $\mathbf{P}$. As a result, the aggregate global supply function of crude by each country $r=1,2,...,S$ –

including $r = s$ – is unaffected by the tax scheme at any given price collection $\mathbf{P}$; i.e.,

$$AS_t^{\text{Oil}}\left(\mathbf{P};\mathbf{T}^r\right) = AS_t^{\text{Oil}}\left(\mathbf{P};\check{\mathbf{T}}\right)$$

for all $t = 1, 2, ...$

3. The aggregate supply of all consumption goods y_{bt}^{js} for $b = 1, 2, ..., B$ by each country $r = 1, 2, ..., S$ – including $r = s$ – is unaffected by the tax scheme at any given price collection $\mathbf{P}$. Thus, the aggregate supply function of a given consumption good from country r is unchanged in each country $r = 1, 2, ..., S$; i.e.,

$$AS_{bt}^{r}\left(\mathbf{P};\mathbf{T}\right) = AS_{bt}^{r}\left(\mathbf{P};\check{\mathbf{T}}\right)$$

for all $r = 1, 2, ..., S$, all $b = 1, 2, ..., B$, and all $t = 1, 2, ...$

4. The aggregate demand function in each country r of gas $AD_t^{r,\text{Gas}}\left(\mathbf{P};\mathbf{T}\right)$ and all consumption goods $AD_{bt}^{r}\left(\mathbf{P};\mathbf{T}\right)$, as well as the global demand for crude oil $AD_t^{\text{Oil}}\left(\mathbf{P};\mathbf{T}\right)$ are all unaffected by the tax collection $\check{\mathbf{T}}$ relative to $\mathbf{T}$ because no tax change affects midstream and downstream markets. Because all the aggregate demand and all the aggregate supply functions are unaffected by the tax reform, the market equilibrium in each period $t = 1, 2, ...$ under scheme $\check{\mathbf{T}}$ solves the same system of $[1 + s \times (B + 1)]$ equations per period that delivers the equilibrium price vector under scheme $\mathbf{T}$. Thus, the equilibrium price collection $\mathbf{P}$ that clears all the markets in each period t under scheme $\mathbf{T}$, also clears all the markets in each period t under the tax scheme $\check{\mathbf{T}}$. Thus, we have shown that the proposed tax reform has no effect on the equilibrium prices of oil, gas, and of all consumption goods. Q.E.D.

Part (i). Recall that the FOCs of the oil&gas firm's problem w.r.t. IVent_t^{iks}, Gas_t^{iks}, and Z_t^{iks} write:

$$-VF_t^{iks} \times \frac{PF_t^{iks}\left(\cdot\right)}{\partial \text{IVent}_t^{iks}} + \left(\phi_{2t}^{iks} - \phi_{3t}^{iks} - \phi_{4t}^{iks}\zeta\right)\left(1 - \vartheta_t^{iks}\right)^{-1} + \phi_{9t}^{iks} = 0 \,, \tag{6.91}$$

$$\begin{array}{c}\left[P_t^{s,\text{Gas}}\left(1 - \varsigma_t^{\text{Gas}} MS_t^{ks,\text{Gas}}\right) - \tau_t^{s,\text{Gas}}\right]\left(1 - T_t^s\right) \\ -\phi_{1t}^{iks}\frac{\partial TF_{2t}^{iks}(.)}{\partial \text{Gas}_t^{iks}} + \left(\phi_{2t}^{iks} - \phi_{3t}^{iks} - \phi_{4t}^{iks}\zeta\right)\left(1 - \vartheta_t^{iks}\right)^{-1}\end{array} = 0 \,, \tag{6.92}$$

and

$$\left(1 - T_t^i\right) + \phi_{1t}^{iks}\frac{\partial TF_{6t}^{iks}(.)}{\partial \text{Z}_t^{iks}} = 0 \,, \tag{6.93}$$

respectively. Recall that IVent_t^{iks} is equal to zero if the FOC in 6.91 is satisfied at $\text{IVent}_t^{iks} = 0$. Combining the three conditions in (6.91), (6.92), and (6.93) and using $\phi_{9t}^{iks} \geq 0$, one gets that IVent_t^{iks} is equal to zero if:

$$\left[P_t^{s,\text{Gas}}\left(1 - \varsigma_t^{\text{Gas}} MS_t^{ks,\text{Gas}}\right) - \tau_t^{iks,\text{Gas}} - MC_t^{iks,\text{Gas}}\right]\left(1 - T_t^s\right) \geq -VF_t^{iks}\frac{\partial PF_{t+1}^{iks}\left(\cdot\right)}{\partial \text{IVent}_t^{iks}} \tag{6.94}$$

at $\text{IVent}_t^{iks} = 0$ and for all possible values of Flare_t^{iks}. Thus, a sufficient condition for (6.94) to hold true for all oil&gas fields is to set a uniform tax rebate $\Delta\tau_t^{iks,\text{Gas}} = \Delta\tau_t^{s,\text{Gas}}$ for all i with formula:

$$\Delta\tau_t^{s,\text{Gas}} = \min_{i \in \{\{1,2,...,I^k\}\}_{k=1}^{K^s}} \left\{P_t^{s,\text{Gas}} - \tau_t^{iks,\text{Gas}} - MC_t^{iks,\text{Gas}}\right\} \tag{6.95}$$

where $\tau_t^{iks,\text{Gas}}$ denotes the marginal tax rate on gas sales faced by any field i given the current tax framework. Note that $\Delta\tau_t^{s,\text{Gas}}$ is typically negative.

Part (ii). Recall that the FOCs of the oil&gas firm's problem w.r.t. Flare_t^{iks} writes:

$$\begin{array}{c} -\tau_t^{iks,\text{Flare}}\left(1-T_t^s\right)-VF_t^{iks}\times\frac{PF_t^{iks}(\cdot)}{\partial\text{Flare}_t^{iks}}-\phi_{1t}^{iks}\frac{\partial TF_{3t}^{iks}(.)}{\partial\text{Flare}_t^{iks}} \\ +\left(\phi_{2t}^{iks}-\phi_{3t}^{iks}-\phi_{4t}^{iks}\zeta\right)\left(1-\vartheta_t^{iks}\right)^{-1}+\phi_{8t}^{iks}=0 \end{array}, \tag{6.96}$$

with $\phi_{8t}^{iks}\geq 0$. Combining the condition in (6.96) with those in (6.92) and (6.93), one gets that Flare_t^{iks} is equal to its lower bound NRF_t^{iks} if:

$$\begin{array}{c} \left[P_t^{s,\text{Gas}}\left(1-\varsigma_t^{\text{Gas}}MS_t^{ks,\text{Gas}}\right)-\tau_t^{s,\text{Gas}}-MC_t^{iks,\text{Gas}}\right]\left(1-T_t^s\right)\geq \\ \left[-MC_t^{iks,\text{Flare}}(0)-\tau_t^{iks,\text{Flare}}\right]\left(1-T_t^s\right)-VF_t^{iks}\times\frac{PF_t^{iks}(\cdot)}{\partial\text{Flare}_t^{iks}} \end{array} \tag{6.97}$$

for all possible values of IVent_t^{iks}. Thus, a sufficient condition for (6.97) to hold true for all oil&gas fields i is:

$$\left[P_t^{s,\text{Gas}}\left(1-\varsigma_t^{\text{Gas}}MS_t^{ks,\text{Gas}}\right)-\tau_t^{s,\text{Gas}}-MC_t^{iks,\text{Gas}}\right]\geq\tau_t^{iks,\text{Flare}} \tag{6.98}$$

Note that formula (6.95) implies:

$$\varDelta\tau_t^{s,\text{Gas}}\leq P_t^{s,\text{Gas}}-\tau_t^{s,\text{Gas}}-MC_t^{iks,\text{Gas}} \tag{6.99}$$

Using the definition of the tax rebate from part (i), $\varDelta\tau_t^{s,\text{Gas}}=\check{\tau}_t^{iks,\text{Gas}}-\tau_t^{iks,\text{Gas}}$, and substituting (6.99) into the sufficient condition for zero flaring in (6.98), we obtain that – given $\varDelta\tau_t^{s,\text{Gas}}$ from (6.95) for all $i=1,2,...I^k$ and all $k=1,2,...K^s$, a sufficient condition for zero flaring by each foil&gas firm $k=1,2,...K^s$ is:

$$\tau_t^{iks,\text{Flare}}\leq 0\;\forall i=1,2,...,I^k\;, \tag{6.100}$$

which implies that, as long as no flaring tax is introduced; i.e., $\tau_t^{iks,\text{Flare}}=0$ for all $i=1,2,...,I^k$, the tax rebate $\varDelta\tau_t^{s,\text{Gas}}$ ensures zero flaring by all oil&gas firms.

Part (iii). Recall that the optimality condition for M_{t+1}^{iks} for an oil&gas firm with respect to oil field i in period $t+1$ writes:

$$\begin{array}{c} \beta\left(1-T_{t+1}^s\right)MPE_{t+1}^{iks}\left(\text{M}_{t+1}^{iks}\right)-\left[\left(1-T_t^s\right)-\beta\left(1-\rho^{iks}\right)\left(1-T_{t+1}^s\right)\right]+\phi_{12t}^{iks} \\ +\beta\left[VF_{t+1}^{iks}\frac{\partial PU_{t+1}^{iks}(\cdot)}{\partial\text{M}_{t+1}^{iks}}+\left(P_{t+1}^{s,\text{Gas}}-\tau_{t+1}^{iks,\text{Gas}}-MC_{t+1}^{iks,\text{Gas}}\right)\left(1-T_{t+1}^s\right)\right]\frac{\partial Maint_{t+1}^{iks}(\cdot)}{\partial\text{M}_{t+1}^{iks}}=0 \end{array} \tag{6.101}$$

Recall that $\varDelta\delta_{1t+1}^{iks}=-\varDelta\tau_{t+1}^{s,\text{Gas}}$ for oil fields and $\varDelta\delta_{1t+1}^{iks}=0$ for gas fields. Because the optimality condition for M_{t+1}^{iks} in (6.101) is unaffected by changes in tax rates other than $\tau_{t+1}^{iks,\text{Gas}}$, it is sufficient to show that $\frac{d\text{M}_{t+1}^{iks}}{d\tau_{t+1}^{iks,\text{Gas}}}\leq 0$. If the twelfth constraint is non-binding ($\phi_{12t+1}^{iks}>0$), then $\text{M}_{t+1}^{iks}=0$ and the derivative of interest is simply $\frac{d\text{M}_{t+1}^{iks}}{d\tau_{t+1}^{iks,\text{Gas}}}=0$. If the the twelfth constraint is binding ($\phi_{12t+1}^{iks}=0$), then given the convexity of the firm's maximization problem (see proof to Proposition 2), we can totally differentiate the F.O.C. w.r.t. $\tau_{t+1}^{iks,\text{Gas}}$ to obtain:

$$\begin{array}{c} -\beta\frac{\partial Maint_{t+1}^{iks}(\cdot)}{\partial\text{M}_{t+1}^{iks}}\left(1-T_{t+1}^s\right)+\beta\left\{VF_{t+1}^{iks}\frac{\partial^2 PU_{t+1}^{iks}(\cdot)}{\partial\text{M}_t^{iks\;2}}\right. \\ +\left(1-T_{t+1}^s\right)\left[P_{t+1}^{s,\text{Gas}}-\tau_{t+1}^{iks,\text{Gas}}-MC_{t+1}^{iks,\text{Gas}}\right]\frac{\partial^2 Maint_{t+1}^{iks}(\cdot)}{\partial\text{M}_{t+1}^{iks\;2}} \\ \left.+\left(1-T_{t+1}^s\right)\frac{\partial MPE_{t+1}^{iks}(\cdot)}{\partial\text{M}_{t+1}^{iks}}\right\}\frac{d\text{M}_{t+1}^{iks}}{d\tau_{t+1}^{iks,\text{Gas}}}=0 \end{array} \tag{6.102}$$

Where the second term in equation (6.102) is weakly negative, as shown in the proof to Proposition 2. In particular, if such term is equal to zero at the optimal level of M_{t+1}^{iks}, then the FOC in (6.101) does not have a unique solution for M_{t+1}^{iks}, implying that the optimal value of M_{t+1}^{iks} is pinned down by the constraint $\text{M}_{t+1}^{iks}-\text{M}_t^{iks}\left(1-\rho^{iks}\right)-\text{IM}_t^{iks}\leq 0$ and it is therefore constant in $\tau_{t+1}^{iks,\text{Gas}}$. Instead, if the second term in equation (6.102) is strictly negative, then we can solve (6.102) with respect to $\frac{d\mathrm{M}_{t+1}^{iks}}{d\tau_{t+1}^{iks,\text{Gas}}}$ and obtain::

$$\frac{d\mathrm{M}_{t+1}^{iks}}{d\tau_{t+1}^{iks,\text{Gas}}} = \frac{\partial Maint_{t+1}^{iks}(\cdot)}{\partial \mathrm{M}_{t+1}^{iks}} \left\{ \frac{VF_{t+1}^{iks}}{1-T_{t+1}^{s}} \frac{\partial^2 PU_{t+1}^{iks}(\cdot)}{\partial \mathrm{M}_{t}^{iks\ 2}} + \left[P_{t+1}^{s,\text{Gas}} - \tau_{t+1}^{iks,\text{Gas}} - MC_{t+1}^{iks,\text{Gas}} \right] \frac{\partial^2 Maint_{t+1}^{iks}(\cdot)}{\partial \mathrm{M}_{t+1}^{iks\ 2}} - \frac{\partial MPE_{t+1}^{iks}(\cdot)}{\partial \mathrm{M}_{t+1}^{iks}} \right\}^{-1} \leq 0 \quad , \tag{6.103}$$

where the first term $\frac{\partial Maint_{t+1}^{iks}(\cdot)}{\partial \mathrm{M}_{t+1}^{iks}}$ is strictly positive by assumption and the second term is strictly negative, implying that $\frac{d\mathrm{M}_{t+1}^{iks}}{d\tau_{t+1}^{iks,\text{Gas}}} \leq 0$.

Because the optimality condition for M_{t+1}^{iks} in (6.101) is unaffected by changes in tax rates other than $\tau_{t+1}^{iks,\text{Gas}}$ and δ_{1t+1}^{iks}, it is sufficient to consider the effect of the change in these two tax variables. If the twelfth constraint is binding ($\phi_{12t+1}^{iks} > 0$), then $\mathrm{M}_{t+1}^{iks} = 0$ and the derivative of interest is simply $\frac{d\mathrm{M}_{t+1}^{iks}}{d\tau_{t+1}^{s,\text{Gas}}} = \frac{d\mathrm{M}_{t+1}^{iks}}{d\delta_{1t+1}^{iks}} = 0$. If the the twelfth constraint is not binding ($\phi_{12t+1}^{iks} = 0$), then the optimal level of M_{t+1}^{iks} solves equation (6.101) evaluated at $\phi_{12t+1}^{iks} = 0$. In this case, $\Delta\delta_{1t+1}^{iks} = -\Delta\tau_{t+1}^{s,\text{Gas}}$ implies that the equation (6.101) is unaffected by the introduction of the reform, therefore the solution to the equation is also unchanged.

The optimality conditions for ReInj_t^{iks} , PInj_t^{iks} and OInj_t^{iks} for an oil&gas firm with respect to oil field i in period t write:

$$\begin{gathered}
\left[-\frac{IC_t^{iks}\left(\mathrm{ReInj}_t^{iks},\mathrm{PInj}_t^{iks}\right)}{\partial \mathrm{ReInj}_t^{iks}} + \delta_{1t}^{iks} - \left(P_t^{s,\text{Gas}} - \tau_t^{isk,\text{Gas}} - MC_t^{iks,\text{Gas}} \right) \right] (1 - T_t^s) \\
+\phi_{4t}^{iks} \frac{\partial B_t^{iks}(\cdot)}{\partial \mathrm{ReInj}_t^{iks}} + \phi_{8t}^{iks} = 0 \\
\left[-\frac{IC_t^{iks}\left(\mathrm{ReInj}_t^{iks},\mathrm{PInj}_t^{iks}\right)}{\partial \mathrm{PInj}_t^{iks}} - PP_t^{iks,\text{Gas}} \right] (1 - T_t^s) + \phi_{4t}^{iks} \frac{\partial B_t^{iks}(\cdot)}{\partial \mathrm{PInj}_t^{iks}} = 0 \\
-C_t^{iks} (1 - T_t^s) + \phi_{4t}^{iks} \frac{\partial B_t^{iks}(\cdot)}{\partial \mathrm{OInj}_t^{iks}} = 0
\end{gathered} \tag{6.104}$$

plus $\phi_{8t}^{iks} \geq 0$ and $\mathrm{ReInj}_t^{iks} \geq 0$. First, note that $\Delta\delta_{1t}^{iks} = -\Delta\tau_t^{s,\text{Gas}}$ implies that the FOC w.r.t. ReInj_t^{iks} is unaffected by the tax reform. Because the third condition is always binding (see Step 3 of the proof to part (iv)), then it can be substituted into the other two conditions, implying that the FOC w.r.t. PInj_t^{iks} is also unaffected by the tax reform. Lastly, we know that ϕ_{4t}^{iks} is unaffected by the tax reform from Step 4 of part (iv) of this proof. Thus, the optimal values of ReInj_t^{iks}, PInj_t^{iks}, OInj_t^{iks}, and ϕ_{8t}^{iks} given the optimal value of ϕ_{4t}^{iks}, which is unaffected by the tax reform, solve a system of four equations corresponding to four binding conditions (the three equations in (6.104) plus either $\phi_{8t}^{iks} = 0$ or $\mathrm{ReInj}_t^{iks} = 0$) with four unknown. Because all the equations are unaffected by the tax reform, the solution to the system must be the same. In the same way, it is possible to show that the optimal value of ReInS_t^{iks} and PInS_t^{iks} are both unaffected by the tax reform. Q.E.D.

Part (v). Proof. All prices are unaffected by the tax reform by part (iv), and the tax reform does not directly affect the optimality conditions of any midstream firm and any consumer in each period t. Thus, any individual demand function from either midstream firms (demand for G_t^{js} and O_t^{js}) or consumers (demand for y_{bt}^{js} for $b = 1, 2, ..., B$) evaluated at the equilibrium price collection $\mathbf{P} = \{\mathbf{P}_t\}_{t=1}^{\infty}$ must deliver the same equilibrium demand levels under both scheme $\mathbf{T}^s$ and $\check{\mathbf{T}}^s$. As a consequence, each corresponding aggregate demand must also be unchanged under scheme $\check{\mathbf{T}}^s$ relative to scheme $\mathbf{T}^s$.

Part (v). Proof. Note that all midstream and downstream GHG emissions are unchanged because the optimal choice of each midstream firm and each consumer is unaffected by the tax reform. In the upstream level, however, the emissions in each period t change as follows:

$$\begin{aligned}
\Delta Emissions_t^s = \quad & \textstyle\sum_{k=1}^{K^s} \sum_{i=1}^{I^K} CI^{iks,\text{Gas}} \Delta \mathrm{Gas}_t^{iks} + CI^{\text{Flare}} \Delta \mathrm{Flare}_t^{iks} \\
& + CI^{\text{Vent}} \left(\Delta \mathrm{IVent}_t^{iks} + \Delta \mathrm{UVent}_t^{iks} \right) \\
& + CI^{iks,\text{ReInj}} \Delta \mathrm{ReInj}_t^{i} + CI^{iks,\text{ReIns}} \Delta \mathrm{ReIns}_t^{i}
\end{aligned} \tag{6.105}$$

Note that $\sum_{k=1}^{K^s}\sum_{i=1}^{I^k}\Delta\text{Gas}_t^{iks}=0$. Thus, as long as the average carbon intensity of gas extraction is the same, then $\sum_{k=1}^{K^s}\sum_{i=1}^{I^k}CI^{iks,\text{Gas}}\Delta\text{Gas}_t^{iks}=0$. All the other terms: $\Delta\text{Flare}_t^{iks}$, $\Delta\text{IVent}_t^{iks}$, $\Delta\text{UVent}_t^{iks}$, $\Delta\text{ReInj}_t^{iks}$, $\Delta\text{ReIns}_t^{iks}$ are weakly negative, with $\Delta\text{Flare}_t^{iks}$ strictly negative for at least one field i. Thus, $\Delta Emissions_t^s<0$. On the other hands, all the productive choices made by firms in countries other than s are unaffected by the tax reform, implying that the emissions in such countries are also unchanged; i.e., $\Delta Emissions_t^r=0$ for all $r\neq s$. Lastly, these two results together imply $\Delta Emissions_t=\sum_{r=1}^{S}\Delta Emissions_t^s<0$. Q.E.D.

Part (vi). Proof. First we show that firm's profits are weakly larger unchanged under $\check{\mathbf{T}}^s$ relative to $\mathbf{T}^s$. Using the finding that the first constraint $TF_t^{iks}(\cdot)\leq 0$ is always binding (see proof to part (iii)) and that Π_t^{iks} is linear in Z_t^{iks}, we can define the function at $Z_t^{iks}\left(\text{Oil}_t^{iks},\text{Gas}_t^{iks},\text{Flare}_t^{iks},\text{ReInS}_t^{iks},\text{PInS}_t^{iks},\text{M}_t^{iks}\right)$ such that $TF_t^{iks}(\cdot)=0$ at $\text{Z}_t^{iks}=Z_t^{iks}(\cdot)$, and use it to re-define the profit function as:

$$\begin{aligned}\tilde{\Pi}_t^{iks}\left(\mathbf{x}_t^{iks};\mathbf{T}^s\right)= \quad &\Pi_t^{iks}\Big(\text{Oil}_t^{iks},\text{Gas}_t^{iks},\text{Flare}_t^{iks},\text{ReInj}_t^{iks},\text{PInj}_t^{iks},\text{PInS}_t^{iks},\\ &Z_t^{iks}(\cdot),\text{ID}_t^{iks},\text{M}_t^{iks},\text{IM}_t^{iks},\text{OInj}_t^{iks},\text{IVent}_t^{i};\mathbf{T}^s\Big)\end{aligned} \tag{6.106}$$

and note that for each $\text{x}_{jt}^{iks}\neq\text{Z}_t^{iks}$ we get:

$$\frac{\partial Z_t^{iks}(\cdot)}{\partial \text{x}_{jt}^{iks}}=-\left.\frac{\partial TF_t^{iks}(.)}{\partial \text{x}_{jt}^{iks}}\right/\frac{\partial TF_t^{iks}(.)}{\partial \text{Z}_t^{iks}} \tag{6.107}$$

This implies that the derivative of $\tilde{\Pi}_t^{iks}$ with respect to x_{jt}^{iks} writes:

$$\frac{\partial\tilde{\Pi}_t^{iks}(\cdot)}{\partial \text{x}_{jt}^{iks}}=\frac{\partial\Pi_t^{iks}(\cdot)}{\partial \text{x}_{jt}^{iks}}-\left(1-T_t^s\right)\left.\frac{\partial TF_t^{iks}(.)}{\partial \text{x}_{jt}^{iks}}\right/\frac{\partial TF_t^{iks}(.)}{\partial \text{Z}_t^{iks}} \tag{6.108}$$

for each $\text{x}_{jt}^{iks}\neq\text{Z}_t^{iks}$. Recall that the change in the tax deduction has formula:

$$\Delta\delta_{1t}^{iks}=-\Delta\tau_t^{s,\text{Gas}}\mathbf{1}\left[type_t^{iks}=Oil\right] \tag{6.109}$$

for all $i=1,2,...,I^k$, and

$$\Delta\delta_{0t}^{iks}=-\Delta\tau_t^{s,\text{Gas}}\left(1-T_t^{ks}\right)\left\{\widehat{Gas}_t^{iks}\mathbf{1}\left[type_t^{iks}=Gas\right]-\widehat{Maint}_t^{iks}\mathbf{1}\left[type_t^{iks}=Oil\right]\right\} \tag{6.110}$$

where the variable $\widehat{Gas}_t^{iks}$ is defined as follows:

$$\widehat{Gas}_t^{iks}=\frac{\sum_{l\neq k}\sum_{j=1}^{I^l}\check{\text{Gas}}_t^{jls}\mathbf{1}\left[type_t^{jls}=Gas\right]}{\left(K^s-1\right)\left(\sum_{j=1}^{I^k}\mathbf{1}\left[type_t^{jls}=Gas\right]\right)} \tag{6.111}$$

and $\widehat{Maint}_t^{iks}$ has formula:

$$\widehat{Maint}_t^{iks}=\frac{\sum_{l\neq k}\sum_{j=1}^{I^l}\text{Res}\check{\text{Gas}}_t^{jls}\mathbf{1}\left[type_t^{jls}=Oil\right]}{\left(K^s-1\right)\sum_{j=1}^{I^k}\mathbf{1}\left[type_t^{jks}=Oil\right]} \tag{6.112}$$

Thus, we can calculate the profit generated by such a bundle. First, note that the linearity of the profit function in the tax rates on oil and gas implies:

$$\begin{aligned}\tilde{\Pi}_t^{iks}\left(\tilde{\mathbf{x}}_t^{iks};\check{\mathbf{T}}^s\right) = \quad & \tilde{\Pi}_t^{iks}\left(\tilde{\mathbf{x}}_t^{iks};\mathbf{T}^s\right) - \Delta\tau_t^{s,\text{Oil}}\tilde{\text{Oil}}_t^{iks} \\ & -\Delta\tau_t^{s,\text{Gas}}\tilde{\text{Gas}}_t^{iks} + \Delta\delta_{0t}^{iks} + \Delta\delta_{1t}^{iks}\left(\tilde{\text{ReInj}}_t^{iks} + \tilde{\text{ReInS}}_t^{iks}\right)\end{aligned} \tag{6.113}$$

where $\Delta\delta_{0t}^{iks} = -\Delta\tau_t^{s,\text{Gas}}\left(\text{NRF}_t^{iks} - \widehat{Maint}_t^{iks}\right)$ for oil fields and $\delta_{0t}^{iks} = -\Delta\tau_t^{s,\text{Gas}}\widehat{Gas}_t^{iks}$ for gas-only fields.

(a) Profits - Oil Fields. Note that the introduction of the tax reform $\check{\mathbf{T}}^s$ causes all prices to be unchanged relative to $\mathbf{T}^s$. Consider the following choice vector: $\tilde{\mathbf{x}}_t^{iks}$, where $\tilde{\text{x}}_{jt}^{iks} = \text{x}_{jt}^{iks}$ for all j except the following. Set $\tilde{\text{Gas}}_t^{iks} = \text{Gas}_t^{iks} + \text{Flare}_t^{iks} - \text{NRF}_t^{iks} + \text{IVent}_t^{iks} - Maint_t^{iks}\left(\check{\text{M}}_t^{iks}\right) + Maint_t^{iks}\left(\text{M}_t^{iks}\right)$, $\tilde{\text{Flare}}_t^{iks} = \text{NRF}_t^{iks}$, and $\tilde{\text{IVent}}_t^{iks} = 0$ for all fields and leave all other choice variables (other than Z_t^{iks}) unchanged; i.e., $\tilde{\text{ReInj}}_t^{iks} = \text{ReInj}_t^{iks}$, $\tilde{\text{ReInS}}_t^{iks} = \text{ReInS}_t^{iks}$, etc. Note that choice vector $\tilde{\mathbf{x}}_t^{iks}$ lies within the domain of the function $\tilde{\Pi}_t^{iks}$ and it is feasible given the assumptions of upstream fields. Lastly, note that the FOCs of the firm's problem imply: $\frac{\partial\tilde{\Pi}_t^{iks}(\cdot)}{\partial\text{Gas}_t^{iks}} = \frac{\partial\tilde{\Pi}_t^{iks}(\cdot)}{\partial\text{Flare}_t^{iks}} = \frac{\partial\tilde{\Pi}_t^{iks}(\cdot)}{\partial\text{IVent}_t^{iks}}$ whenever the constraints $\text{Flare}_t^{iks} - \text{NRF}_t^{iks} \geq 0$, $\text{IVent}_t^{iks} \geq 0$ are not binding. If any of the these two constraints is binding, then the change in the value of the corresponding variable is equal to zero at the margin. Using these two result and the differentiablity of $\tilde{\Pi}_t^{iks}$, we can obtain a linear approximation of $\tilde{\Pi}_t^{iks}\left(\mathbf{x}_t^{iks},\mathbf{T}^s\right)$ for oil fields as follows:

$$\begin{aligned}\tilde{\Pi}_t^{iks}\left(\tilde{\mathbf{x}}_t^{iks},\mathbf{T}^s\right) \simeq \quad & \left.\frac{\partial\tilde{\Pi}_t^{iks}(\cdot)}{\partial\text{Gas}_t^{i}}\right|_{\mathbf{x}_t^{iks},\mathbf{T}^s}\left(\tilde{\text{Gas}}_t^{iks} - \text{Gas}_t^{iks}\right) - \left.\frac{\partial\tilde{\Pi}_t^{iks}(\cdot)}{\partial\text{Flare}_t^{iks}}\right|_{\mathbf{x}_t^{iks},\mathbf{T}^s}\left(\text{Flare}_t^{iks} - \text{NRF}_t^{iks}\right) \\ & - \left.\frac{\partial\tilde{\Pi}_t^{iks}(\cdot)}{\partial\text{IVent}_t^{iks}}\right|_{\mathbf{x}_t^{iks},\mathbf{T}^s}\text{IVent}_t^{iks} = \\ = \quad & \left.\frac{\partial\tilde{\Pi}_t^{iks}(\cdot)}{\partial\text{Gas}_t^{i}}\right|_{\mathbf{x}_t^{iks},\mathbf{T}^s}\left[\left(\tilde{\text{Gas}}_t^{iks} - \text{Gas}_t^{iks}\right) - \left(\text{Flare}_t^{iks} - \text{NRF}_t^{iks} + \text{IVent}_t^{iks}\right)\right] \\ \simeq \tilde{\Pi}_t^{iks}\left(\mathbf{x}_t^{iks},\mathbf{T}^s\right) &\end{aligned} \tag{6.114}$$

Thus, we found that for oil fields, the following linear approximation holds true:

$$\tilde{\Pi}_t^{iks}\left(\tilde{\mathbf{x}}_t^{iks},\mathbf{T}^s\right) \simeq \tilde{\Pi}_t^{iks}\left(\mathbf{x}_t^{iks},\mathbf{T}^s\right) \tag{6.115}$$

Using result (6.115) into (6.113) we obtain an approximation of $\tilde{\Pi}_t^{iks}\left(\tilde{\mathbf{x}}_t^{iks},\check{\mathbf{T}}^s\right)$, which writes:

$$\begin{aligned}\tilde{\Pi}_t^{iks}\left(\mathbf{x}_t^{iks},\mathbf{T}^s\right) - \Delta\tau_t^{s,\text{Oil}}\check{\text{Oil}}_t^{iks} - \Delta\tau_t^{s,\text{Gas}}\tilde{\text{Gas}}_t^{iks} \\ +\Delta\delta_{0t}^{iks} + \Delta\delta_{1t}^{iks}\left(\tilde{\text{ReInj}}_t^{iks} + \tilde{\text{ReInS}}_t^{iks}\right)\end{aligned} \tag{6.116}$$

which for oil fields, given $\Delta\tau_t^{s,\text{Gas}} = -\Delta\delta_{1t}^{iks}$ and $\Delta\delta_{0t}^{iks} = -\Delta\tau_t^{s,\text{Gas}}\left(\text{NRF}_t^{iks} - \widehat{Maint}_t^{iks}\right)$, the formula above is equivalent to:

$$\begin{aligned}\tilde{\Pi}_t^{iks}\left(\tilde{\mathbf{x}}_t^{iks},\check{\mathbf{T}}^s\right) \simeq \quad & \tilde{\Pi}_t^{iks}\left(\mathbf{x}_t^{iks},\mathbf{T}^s\right) - \Delta\tau_t^{s,\text{Oil}}\check{\text{Oil}}_t^{iks} \\ & -\Delta\tau_t^{s,\text{Gas}}\left[\tilde{\text{Gas}}_t^{iks} - \widehat{Maint}_t^{iks} + \tilde{\text{ReInj}}_t^{iks} + \tilde{\text{ReInS}}_t^{iks} + \text{NRF}_t^{iks}\right]\end{aligned} \tag{6.117}$$

which using the fact that for oil fields $GOR^{iks}\text{Oil}_t^{iks} = \check{\text{TotGas}}_t^{iks}$ and $\Delta\tau_t^{iks,\text{Oil}} = -\Delta\tau_t^{s,\text{Gas}}GOR^{iks}\left(1 - \vartheta^{iks}\right)$, the formula above rewrites as follows:

$$\tilde{\Pi}_t^{iks}\left(\tilde{\mathbf{x}}_t^{iks},\check{\mathbf{T}}^s\right) \simeq \tilde{\Pi}_t^{iks}\left(\mathbf{x}_t^{iks},\mathbf{T}^s\right) + \Delta\tau_t^{s,\text{Gas}}\left[\widehat{Maint}_t^{iks} - Maint_t^{iks}\left(\check{\text{M}}_t^{iks}\right)\right] \tag{6.118}$$

(b) Profits - Gas-only fields. For gas-only fields we set $\tilde{\text{Gas}}_t^{iks} = \check{\text{Gas}}_t^{iks}$ and $\tilde{\text{ReInS}}_t^{iks} = \text{ReInS}_t^{iks} - \tilde{\text{Gas}}_t^{iks} + \text{Gas}_t^{iks}$. Then, using the same method as in the previous paragraph, it is possible to show that $\tilde{\Pi}_t^{iks}\left(\tilde{\mathbf{x}}_t^{iks},\mathbf{T}^s\right) \simeq$

$\tilde{\Pi}_t^{iks}\left(\mathbf{x}_t^{iks}, \mathbf{T}^s\right)$ holds true for gas-only fields. Second, for gas-only fields we have $\check{\text{Oil}}_t^{iks} = 0$ and $\Delta\delta_{0t}^{iks} = \Delta\tau_t^{s,\text{Gas}}\widehat{Gas}_t^{iks}$. Thus, the approximate value of $\tilde{\Pi}_t^{iks}\left(\tilde{\mathbf{x}}_t^{iks}, \check{\mathbf{T}}^s\right)$ writes:

$$\tilde{\Pi}_t^{iks}\left(\tilde{\mathbf{x}}_t^{iks}, \check{\mathbf{T}}^s\right) \simeq \tilde{\Pi}_t^{iks}\left(\mathbf{x}_t^{iks}, \mathbf{T}^s\right) - \Delta\tau_t^{s,\text{Gas}}\left[\check{\text{Gas}}_t^{iks} - \widehat{Gas}_t^{iks}\right](1 - T_t^s) \tag{6.119}$$

where the formula for $\widehat{Gas}_t^{iks}$ is provided in the previous paragraph.

(c) *Profits - Aggregate*. By optimality, because all prices are unchanged by the introduction of the tax reform and the bundle $\left\{\left\{\tilde{\mathbf{x}}_t^{iks}\right\}_{i=1}^{I^k}\right\}_{t=1}^{\infty}$ is feasible for firm k but it is not chosen, the present discounted value of all firms profits generated by $\left\{\left\{\tilde{\mathbf{x}}_t^{iks}\right\}_{i=1}^{I^k}\right\}_{t=1}^{\infty}$ cannot fall short that obtained by the optimal bundle $\left\{\left\{\check{\mathbf{x}}_t^{iks}\right\}_{i=1}^{I^k}\right\}_{t=1}^{\infty}$ given tax policy $\check{\mathbf{T}}^s$. Therefore,

$$\sum_{t=1}^{\infty}\sum_{i=1}^{I^k}\beta^{t-1}\tilde{\Pi}_t^{iks}\left(\check{\mathbf{x}}_t^{iks}, \check{\mathbf{T}}^s\right) \geq \sum_{t=1}^{\infty}\sum_{i=1}^{I^k}\beta^{t-1}\tilde{\Pi}_t^{iks}\left(\tilde{\mathbf{x}}_t^{iks}, \check{\mathbf{T}}^s\right) \tag{6.120}$$

Subtracting $\sum_{t=1}^{\infty}\sum_{i=1}^{I^k}\beta^{t-1}\tilde{\Pi}_t^{iks}\left(\mathbf{x}_t^{iks}, \mathbf{T}^s\right)$ from both side of the inequality (6.120) we obtain:

$$\begin{aligned} &\textstyle\sum_{t=1}^{\infty}\sum_{i=1}^{I^k}\beta^{t-1}\left[\tilde{\Pi}_t^{iks}\left(\check{\mathbf{x}}_t^{iks}, \check{\mathbf{T}}^s\right) - \tilde{\Pi}_t^{iks}\left(\mathbf{x}_t^{iks}, \mathbf{T}^s\right)\right] \geq \\ &\textstyle\sum_{t=1}^{\infty}\sum_{i=1}^{I^k}\beta^{t-1}\left[\tilde{\Pi}_t^{iks}\left(\tilde{\mathbf{x}}_t^{iks}, \check{\mathbf{T}}^s\right) - \tilde{\Pi}_t^{iks}\left(\mathbf{x}_t^{iks}, \mathbf{T}^s\right)\right] \end{aligned} \tag{6.121}$$

Lastly, substituting (6.118) and (6.119) into (6.121), and denoting $\Delta\tilde{\Pi}_t^{iks}\left(\mathbf{T}^s, \check{\mathbf{T}}^s\right) = \tilde{\Pi}_t^{iks}\left(\check{\mathbf{x}}_t^{iks}, \check{\mathbf{T}}^s\right) - \tilde{\Pi}_t^{iks}\left(\mathbf{x}_t^{iks}, \mathbf{T}^s\right)$, we obtain:

$$\begin{aligned} &\textstyle\sum_{t=1}^{\infty}\sum_{i=1}^{I^k}\beta^{t-1}\Delta\tilde{\Pi}_t^{iks}\left(\mathbf{T}^s, \check{\mathbf{T}}^s\right) \geq \\ &\textstyle\sum_{t=1}^{\infty}\sum_{i=1}^{I^k}\beta^{t-1}\Delta\tau_t^{s,\text{Gas}}\left\{\left[\widehat{Maint}_t^{iks} - Maint_t^{iks}\left(\check{\text{M}}_t^{iks}\right)\right]\mathbf{1}\left[type_t^{iks} = oil\right]\right. \\ &\left.+\left[\widehat{Gas}_t^{iks} - \check{\text{Gas}}_t^{iks}\right]\mathbf{1}\left[type_t^{iks} = gas\right]\right\}(1 - T_t^s) \end{aligned} \tag{6.122}$$

That is, the change in the present discounted value of profits generated by oil fields profits is weakly larger than $\sum_{t=1}^{\infty}\beta^{t-1}\Delta\tau_t^{s,\text{Gas}}\left[\widehat{Maint}_t^{iks} - Maint_t^{iks}\left(\check{\text{M}}_t^{iks}\right)\right](1 - T_t^s)$, whereas the change in those generated by gas-only firms it is weakly larger than $\sum_{t=1}^{\infty}\beta^{t-1}\Delta\tau_t^{s,\text{Gas}}\left[\widehat{Gas}_t^{iks} - \check{\text{Gas}}_t^{iks}\right](1 - T_t^s)$. Because the $\widehat{Maint}_t^{iks} - Maint_t^{iks}\left(\check{\text{M}}_t^{iks}\right)$ – which can be either positive or negative – is typically very small in magnitude, we can conclude that the effect of the tax reform on oil fields' profits is negligible. This finding rules out concerns regarding possible effects of the tax reform on the extensive margin of oil production. Lastly, note that the profits of other oil&gas fields (i.e., neither oil fields not gas-only fields) are unaffected by the tax reform, we can sum up the the formula for oil&gas firm's profits over all firms, to obtain that the aggregate value of present-discounted profits for oil&gas firms, which write:

$$\begin{aligned} &\textstyle\sum_{k=1}^{K^s}\sum_{t=1}^{\infty}\sum_{i=1}^{I^k}\beta^{t-1}\left[\tilde{\Pi}_t^{iks}\left(\check{\mathbf{x}}_t^{iks}, \check{\mathbf{T}}^s\right) - \tilde{\Pi}_t^{iks}\left(\mathbf{x}_t^{iks}, \mathbf{T}^s\right)\right] \geq \\ &\textstyle\sum_{k=1}^{K^s}\sum_{t=1}^{\infty}\sum_{i=1}^{I^k}\beta^{t-1}\Delta\tau_t^{s,\text{Gas}}\left\{\left[\widehat{Maint}_t^{iks} - Maint_t^{iks}\left(\check{\text{M}}_t^{iks}\right)\right]\mathbf{1}\left[type_t^{iks} = oil\right]\right. \\ &\left.+\left[\widehat{Gas}_t^{iks} - \check{\text{Gas}}_t^{iks}\right]\mathbf{1}\left[type_t^{iks} = gas\right]\right\}(1 - T_t^s) \end{aligned} \tag{6.123}$$

Using the fact that the tax reform implies $\check{\text{IVent}}_t^{iks} = 0$ and $\check{\text{Flare}}_t^{iks} = \text{NRF}_t^{iks}$, the formula for $\widehat{Maint}_t^{iks}$ implies:

$$\sum_{k=1}^{K^s}\sum_{i=1}^{I^k} \widehat{Maint}_t^{iks} = \sum_{k=1}^{K^s}\sum_{i=1}^{I^k} Maint_t^{iks}\left(\check{\text{M}}_t^{iks}\right) \tag{6.124}$$

and that the formula for $\widehat{Gas}_t^{iks}$ implies:

$$\sum_{k=1}^{K^s}\sum_{i=1}^{I^k} \widehat{Gas}_t^{iks} = \sum_{k=1}^{K^s}\sum_{i=1}^{I^k} \check{\text{Gas}}_t^{iks} \tag{6.125}$$

we can use the results (6.124) and (6.125) into (6.123) to obtain:

$$\sum_{k=1}^{K^s}\sum_{t=1}^{\infty}\sum_{i=1}^{I^k} \beta^{t-1}\left[\tilde{\Pi}_t^{iks}\left(\check{\mathbf{x}}_t^{iks}, \check{\mathbf{T}}^s\right) - \tilde{\Pi}_t^{iks}\left(\mathbf{x}_t^{iks}, \mathbf{T}^s\right)\right] \geq 0\,, \tag{6.126}$$

that is, the aggregate present discounted value of profits from oil&gas firms is weakly larger under the tax reform $\check{\mathbf{T}}^s$ relative to $\mathbf{T}^s$.

(d) Tax Revenue - Oil Fields. Note that for oil fields:

$$\check{\text{TotGas}}_t^{iks} = \left(1-\vartheta^{iks}\right)^{-1}\left[\check{\text{Gas}}_t^{iks} - Maint_t^{iks}\left(\check{\text{M}}_t^{iks}\right) + \check{\text{ReInj}}_t^{iks} + \text{ReInS}_t^{iks} + \text{NRF}_t^{iks}\right] \tag{6.127}$$

First, note that because the aggregate present discounted value of profits is larger under the tax reform $\check{\mathbf{T}}^s$ relative to $\mathbf{T}^s$, then aggregate present discounted value of tax revenue is also larger. Thus, for oil fields the change in tax revenue for the government writes:

$$\begin{aligned} \varDelta TRev^{s,\text{Oil}}\left(\mathbf{T}^s, \check{\mathbf{T}}^s\right) \geq \quad & \textstyle\sum_{t=1}^{\infty}\beta^{t-1}\sum_{k=1}^{K^s}\sum_{i=1}^{I^k}\Big\{\varDelta\tau_t^{s,\text{Gas}} \\ & \left[\check{\text{TotGas}}_t^{iks} - \widehat{Maint}_t^{iks} + Maint_t^{iks}\left(\check{\text{M}}_t^{iks}\right)\right] + \\ & \tau_t^{iks,\text{Gas}}\varDelta\text{Gas}_t^{iks} + \varDelta\tau_t^{i,\text{Oil}}\check{\text{Oil}}_t^{iks}\Big\}\left(1-T_t^s\right)\mathbf{1}\left[type_t^{iks} = Oil\right] \end{aligned} \tag{6.128}$$

which using $\varDelta\tau_t^{iks,\text{Oil}}\check{\text{Oil}}_t^{iks} = \varDelta\tau_t^{s,\text{Gas}}\check{\text{TotGas}}_t^{iks}$ reduces to:

$$\begin{aligned} \varDelta TRev^{s,\text{Oil}}\left(\mathbf{T}^s, \check{\mathbf{T}}^s\right) = \quad & \textstyle\sum_{t=1}^{\infty}\beta^{t-1}\sum_{k=1}^{K^s}\sum_{i=1}^{I^k}\Big\{\tau_t^{iks,\text{Gas}}\varDelta\text{Gas}_t^{iks}\left(1-T_t^s\right) \\ & + \varDelta\tau_t^{s,\text{Gas}}\left[Maint_t^{iks}\left(\check{\text{M}}_t^{iks}\right) - \widehat{Maint}_t^{iks}\right]\Big\} \\ & \left(1-T_t^s\right)\mathbf{1}\left[type_t^{iks} = Oil\right] \end{aligned} \tag{6.129}$$

Using the fact that the tax reform implies $\check{\text{IVent}}_t^{iks} = 0$ and $\check{\text{Flare}}_t^{iks} = \text{NRF}_t^{iks}$, the formula for $\widehat{Maint}_t^{iks}$ implies:

$$\sum_{k=1}^{K^s}\sum_{i=1}^{I^k} \widehat{Maint}_t^{iks} = \sum_{k=1}^{K^s}\sum_{i=1}^{I^k} Maint_t^{iks}\left(\check{\text{M}}_t^{iks}\right) \tag{6.130}$$

Using (6.130) into (6.129) the formula reduces to:

$$\varDelta TRev^{s,\text{Oil}}\left(\mathbf{T}^s, \check{\mathbf{T}}^s\right) = \sum_{t=1}^{\infty}\beta^{t-1}\sum_{k=1}^{K^s}\sum_{i=1}^{I^k}\tau_t^{iks,\text{Gas}}\varDelta\text{Gas}_t^{iks}\left(1-T_t^s\right)\left(1-T_t^s\right)\mathbf{1}\left[type_t^{iks} = Oil\right] \tag{6.131}$$

(e) Tax Revenue - Gas-only Fields. Note that for gas-only fields the change in tax revenue for the government

writes:

$$\begin{aligned}\Delta TRev^{s,\text{Gas}}\left(\mathbf{T}^s,\check{\mathbf{T}}^s\right) = & \sum_{t=1}^{\infty}\beta^{t-1}\sum_{k=1}^{K^s}\sum_{i=1}^{I^k}\Big[\Delta\tau_t^{s,\text{Gas}}\left(\check{\text{Gas}}_t^{iks}-\widehat{Gas}_t^{iks}\right) \\ & +\tau_t^{iks,\text{Gas}}\Delta\text{Gas}_t^{iks}\Big]\left(1-T_t^s\right)\mathbf{1}\left[type_t^{iks}=Gas\right]\end{aligned} \tag{6.132}$$

The formula for $\widehat{Gas}_t^{iks}$ implies:

$$\sum_{k=1}^{K^s}\sum_{i=1}^{I^k}\widehat{Gas}_t^{iks} = \sum_{k=1}^{K^s}\sum_{i=1}^{I^k}\check{\text{Gas}}_t^{iks} \tag{6.133}$$

Using (6.133) into (6.132) the formula for $\Delta Rev_t^{s,\text{Gas}}$ reduces to:

$$\Delta TRev^{s,\text{Gas}}\left(\mathbf{T}^s,\check{\mathbf{T}}^s\right) = \sum_{t=1}^{\infty}\beta^{t-1}\sum_{k=1}^{K^s}\sum_{i=1}^{I^k}\tau_t^{iks,\text{Gas}}\Delta\text{Gas}_t^{iks}\left(1-T_t^s\right)\mathbf{1}\left[type_t^{iks}=Gas\right] \tag{6.134}$$

(f) Tax Revenue - Aggregate. Lastly, we can calculate the change in aggregate revenue:

$$\begin{aligned}\Delta TRev^{s}\left(\mathbf{T}^s,\check{\mathbf{T}}^s\right) = & \Delta TRev^{s,\text{Oil}}\left(\mathbf{T}^s,\check{\mathbf{T}}^s\right)+\Delta TRev^{s,\text{Gas}}\left(\mathbf{T}^s,\check{\mathbf{T}}^s\right) \\ & =\sum_{t=1}^{\infty}\beta^{t-1}\sum_{k=1}^{K^s}\sum_{i=1}^{I^k}\tau_t^{iks,\text{Gas}}\Delta\text{Gas}_t^{iks}\left(1-T_t^s\right)\mathbf{1}\left[type_t^{iks}=Oil\right] \\ & +\tau_t^{iks,\text{Gas}}\Delta\text{Gas}_t^{iks}\left(1-T_t^s\right)\mathbf{1}\left[type_t^{iks}=Oil\right] \\ & =\sum_{t=1}^{\infty}\beta^{t-1}\tau_t^{s,\text{Gas}}\sum_{k=1}^{K^s}\sum_{i=1}^{I^k}\Delta\text{Gas}_t^{iks}\left(1-T_t^s\right)=0\end{aligned} \tag{6.135}$$

where the last equality follows from the fact that the aggregate natural gas supply is unaffected by the policy, and therefore $\sum_{k=1}^{K^s}\sum_{i=1}^{I^k}\Delta\text{Gas}_t^{iks}=0$, and that we have assumed that prior to the reform all natural gas is taxed at the same rate $\tau_t^{s,\text{Gas}}$.

(g) Net Lifetime Income. The balanced budget assumption implies $\Delta ITax_t^s = -\Delta TRev_t^s$. The change in the aggregate lifetime income ΔLY^s of the consumer in country s is given by:

$$\Delta LY^s\left(\mathbf{T}^s,\check{\mathbf{T}}^s\right)=\sum_{k=1}^{K^s}\sum_{t=1}^{\infty}\sum_{i=1}^{I^k}\beta^{t-1}\left[\Delta\tilde{\Pi}_t^{iks}\left(\mathbf{T}^s,\check{\mathbf{T}}^s\right)+\Delta ITax_t^s\left(\mathbf{T}^s,\check{\mathbf{T}}^s\right)\right] \tag{6.136}$$

Substituting the value of $\sum_{k=1}^{K^s}\sum_{t=1}^{\infty}\sum_{i=1}^{I^k}\beta^{t-1}\Delta\tilde{\Pi}_t^{iks}\left(\mathbf{T}^s,\check{\mathbf{T}}^s\right)\geq 0$ from (6.122) and of $\sum_{t=1}^{\infty}\beta^{t-1}\Delta ITax_t^s\left(\mathbf{T}^s,\check{\mathbf{T}}^s\right)=-\Delta Rev^s\left(\mathbf{T}^s,\check{\mathbf{T}}^s\right)$ from (6.135) into (6.136) we obtain:

$$\Delta LY^s\left(\mathbf{T}^s,\check{\mathbf{T}}^s\right)\geq 0 \tag{6.137}$$

i.e., the consumer in country s enjoys weakly larger net lifetime income under tax scheme $\check{\mathbf{T}}^s$ relative to tax scheme $\mathbf{T}^s$ in each period t. Q.E.D.

Part (vii). *Proof.* In order to calculate the change in utility we must calculate the change in consumer's consumption (see part (vi)) and in the utility loss due to climate change. Note that because by part (v) of Proposition 3, global GHG emissions are lower under the tax scheme $\check{\mathbf{T}}^s$ relative to tax scheme $\mathbf{T}^s$ in every period $t=1,2,...$; i.e., $Emis\check{s}ions_t < Emissions_t$, the law of motion of the excess GHGs in the atmosphere:

$$-\Lambda Emissions_t + ExcTCO_2e_{t+1}-\left(1-\Gamma\right)ExcTCO_2e_t=0 \tag{6.138}$$

implies that $Exc\check{T}CO_2e_t < ExcTCO_2e_t$ in each period $t=1,2,....$ Moreover, we have shown in part (iv) that $\check{\mathbf{c}}_t^s=\mathbf{c}_t^s$ and that all prices and consumption taxes are unaffected by the tax reform. Therefore, the consumer's budget constraint $C_t^s + (\mathbf{p}_t + \mathbf{v}_t^s)' \mathbf{c}_t^s - Y_t^s \leq 0$, which must be binding at the optimal consumer's choice, implies

$$\sum_{t=1}^{\infty} \beta^{t-1} \left(\check{C}_t^s - C_t^s \right) = \sum_{t=1}^{\infty} \beta^{t-1} \left(\check{Y}_t^s - Y_t^s \right) = \mathit{\Delta} LY^s \left(\mathbf{T}^s, \check{\mathbf{T}}^s \right) \geq 0 \tag{6.139}$$

Thus, we have shown that the present discounted value of general consumption C_t^s, weakly larger under tax scheme $\check{\mathbf{T}}^s$ relative to tax scheme $\mathbf{T}^s$. Lastly, because all prices are unaffected by tax scheme $\check{\mathbf{T}}^s$ relative to tax scheme $\mathbf{T}^s$ in every period $t = 1, 2, ...$, then $\mathbf{c}_t^r$ and C_t^r are unaffected by the change in tax policy in country r for all $r \neq s$. Thus,

$$\check{C}_t^s + u^s \left(\check{\mathbf{c}}_t^s \right) - Ext \times Exc\check{TCO_2e_t} \geq C_t^s + u^s \left(\mathbf{c}_t^s \right) - Ext \times \mathit{\Delta} TCO_2 e \, ; \tag{6.140}$$

for all $s = 1, 2, ..., S$ in each period $t = 1, 2, ...$, and therefore:

$$\sum_{t=1}^{\infty} \beta^{t-1} \left[\check{C}_t^s + u^s \left(\check{\mathbf{c}}_t^s \right) - Ext \times Exc\check{TCO_2e_t} \right] \geq \sum_{t=1}^{\infty} \beta^{t-1} \left[C_t^s + u^s \left(\mathbf{c}_t^s \right) - Ext \times \mathit{\Delta} TCO_2 e \right] \, ; \tag{6.141}$$

for all $s = 1, 2, ..., S$; i.e., the lifetime utility of a consumer in each country s is strictly larger under tax scheme $\check{\mathbf{T}}^s$ relative to tax scheme $\mathbf{T}^s$. Q.E.D.

6.2.2 Midstream

Let $MC_t^{js,\mathrm{F}} \left(\mathrm{F}_t^j \right) \equiv \frac{\partial MTF_4^{js}\left(\mathrm{F}_t^j\right)}{\partial \mathrm{F}_t^{js}} \Big/ \frac{\partial MTF_6^{js}\left(\mathrm{MZ}_t^j\right)}{\partial \mathrm{MZ}_t^{js}}$ be the marginal cost of flaring and $\check{MC}_t^{js,\mathrm{F}}$ represent the maximum value of $MC_t^{js,\mathrm{F}}$ for firm j (see section 1.2).

Proposition 4. *A specific linear tax on flaring at rate $\tau_t^{js,F} \geq \check{MC}_t^{js,F}$ implies $F_t^{js} = 0$ and no effect on O_t^{js}, G_t^{js}, $\mathbf{y}_t^{js}$ for all midstream firms $j = 1, 2, ..., J$.*

Proof. First, note that by assumption all midstream firms other than oil refineries feature $MC_t^{js,\mathrm{F}} \left(\mathrm{F}_t^{js} \right) = +\infty$, implying that for such firms the constraint $\mathrm{F}_t^{js*} \geq RF_t^{js}$ and in turn, $\mathrm{F}_t^{js*} = RF_t^{js}$. Second, refineries are defined as all midstream firms such that $\frac{\partial MTF_2^{js}\left(\mathrm{O}_t^{js}\right)}{\partial \mathrm{O}_t^{js}} > 0$, $MC_t^{js,\mathrm{F}} \left(\mathrm{F}_t^{js} \right) \neq +\infty$, and $\frac{\partial MTF_3^{js}\left(\mathrm{G}_t^{js}\right)}{\partial \mathrm{G}_t^{js}} = 0$ for $\mathrm{G}_t^{js} < 0$. Thus, using the FOCs of the midstream firms w.r.t. F_t^{js} and MZ_t^{js}, where the latter is always binding because $\mathrm{MZ}_t^{js} \in (-\infty, +\infty)$, we get:

$$-\tau_t^{js,\mathrm{F}} \left(1 - T_t^s \right) - \left(1 - T_t^s \right) \frac{\partial MTF_4^{js} \left(\mathrm{F}_t^{js} \right)}{\partial \mathrm{F}_t^{js}} \Bigg/ \frac{\partial MTF_6^{js} \left(\mathrm{MZ}_t^{js} \right)}{\partial \mathrm{MZ}_t^{js}} + \psi_{2t}^{js} \leq 0 \tag{6.142}$$

whereas using the FOCs of the midstream firms w.r.t. O_t^{js} and MZ_t^{js} we get:

$$P_t^{\mathrm{Oil}} \left(1 - T_t^s \right) - \left(1 - T_t^s \right) \frac{\partial MTF_2^{js} \left(\mathrm{F}_t^{js} \right)}{\partial \mathrm{O}_t^{js}} \Bigg/ \frac{\partial MTF_6^{js} \left(\mathrm{MZ}_t^{js} \right)}{\partial \mathrm{MZ}_t^{js}} + \psi_{2t}^{js} GOR^{js} \leq 0 \tag{6.143}$$

Let us define the marginal product of oil $MP_t^{js,\mathrm{O}} \left(\mathrm{O}_t^{js} \right) \equiv \frac{\partial MTF_2^{js}\left(\mathrm{O}_t^{js}\right)}{\partial \mathrm{O}_t^{js}} \Big/ \frac{\partial MTF_6^{js}\left(\mathrm{MZ}_t^{js}\right)}{\partial \mathrm{MZ}_t^{js}}$. Considering oil refineries which are active; i.e., they feature $\mathrm{O}_t^{js} > 0$, the condition (6.143) rewrites:

$$\left(1 - T_t^s \right) \left[MP_t^{js,\mathrm{O}} \left(\mathrm{O}_t^{js} \right) - P_t^{\mathrm{Oil}} \right] = \psi_{2t}^{js} GOR^{js} \tag{6.144}$$

and the condition (6.142) rewrites:

$$-\left[\tau_t^{js,\mathrm{F}} + MC_t^{js,\mathrm{F}}\left(\mathrm{F}_t^{js}\right)\right](1-T_t^s) + \psi_{2t}^{js} \leq 0 \tag{6.145}$$

Combining conditions (6.144) and (6.145) we obtain:

$$-\tau_t^{js,\mathrm{F}} - MC_t^{js,\mathrm{F}}\left(\mathrm{F}_t^{js}\right) - \frac{P_t^{\mathrm{Oil}} - MP_t^{js,\mathrm{O}}\left(\mathrm{O}_t^{js}\right)}{GOR^{js}} \leq 0 \tag{6.146}$$

Thus, zero flaring is ensured if:

$$\tau_t^{js,\mathrm{F}} \geq \frac{MP_t^{js,\mathrm{O}}\left(\mathrm{O}_t^{js*}\right) - P_t^{\mathrm{Oil}}}{GOR^{js}} - MC_t^{js,\mathrm{F}}(0) \tag{6.147}$$

whereas with zero flaring tax the FOCs imply:

$$0 \geq \frac{MP_t^{js,\mathrm{O}}\left(\mathrm{O}_t^{js**}\right) - P_t^{\mathrm{Oil}}}{GOR^{js}} - MC_t^{js,\mathrm{F}}\left(\mathrm{F}_t^{js**}\right) \tag{6.148}$$

However, the optimal quantity of oil with positive flare tax in (6.147) may differ from the quantity with no flaring tax in (6.148). In order to tackle this issue, let us define $MC_t^{js,\mathrm{MInS}} \equiv \frac{\partial MTF_5^{js}\left(\mathrm{MInS}_t^j\right)}{\partial \mathrm{MInS}_t^{js}} \Big/ \frac{\partial MTF_6^{js}\left(\mathrm{MZ}_t^{js}\right)}{\partial \mathrm{MZ}_t^{js}}$. Assuming that $MC_t^{js,\mathrm{MInS}}$ is constant, the FOC w.r.t. MInS_t^{js} writes:

$$(1-T_t^s)\, MC_t^{js,\mathrm{MInS}} = \psi_{2t}^{js} \tag{6.149}$$

Substituting (6.149) into (6.144) we obtain:

$$MP_t^{js,\mathrm{O}}\left(\mathrm{O}_t^j\right) = P_t^{\mathrm{Oil}} + MC_t^{js,\mathrm{MInS}} GOR^{js} \tag{6.150}$$

which implies that O_t^{js*} is independent of $MC_t^{js,\mathrm{F}}\left(\mathrm{F}_t^{j**}\right)$ at given P_t^{Oil}. Thus, we get $\mathrm{O}_t^{js*} = \mathrm{O}_t^{js**}$. As a consequence, combining inequalities (6.148) and (6.147), a sufficient condition for inequality to hold is:

$$\tau_t^{js,\mathrm{F}} \geq MC_t^{js,\mathrm{F}}\left(\mathrm{F}_t^{js**}\right) - MC_t^{js,\mathrm{F}}(0) \tag{6.151}$$

Lastly, because the marginal cost of flaring is weakly positive and increasing, a sufficient condition for the inequality (6.151) to hold true is:

$$\tau_t^{js,\mathrm{F}} \geq \check{MC}_t^{js,\mathrm{F}}, \tag{6.152}$$

which implies that if the flaring tax satisfies inequality (6.152) the optimal amount of flaring chosen by each refinery is $\mathrm{F}_t^{js*} = 0$. Lastly, note that the optimality conditions for all other endogenous variables are unaffected by the introduction of , implying that all optimal choices –except possibly MInS_t^{js} and MZ_t^{js} – are unaffected by the policy reform, implying that aggregate net production choices $\mathrm{O}_t^{js}, \mathrm{G}_t^{js}, \mathrm{y}_t^{js}$ are all unaffected. Q.E.D.

6.3 Identification

In order to quantify the aggregate reduction in GHG emissions caused by the implementation of the tax reform proposed in the previous section we need to propose a method to identify and estimate the amount of (unobservable) intentional venting produced by of each oil&gas firm. Note that because all the empirical analysis makes use of data from a single country (the US), for ease of notation in this section we suppress the country superscript s.

6.3.1 Assumptions

For the purpose of identification, we need to impose additional restrictions to our model. Specifically, we assume that

1. The expected regulatory cost of venting possesses a linear-quadratic function of IVent_t^i and Flare_t^i; i.e.,

$$\begin{aligned} VF_t^{ik} \times Pr_t^{ik}\left(ivent_t^{ik}=1 \mid \Omega_t^{ks}\right) = \quad & \kappa_0^{ik}\text{IVent}_t^{ik} + \tfrac{\kappa_1}{2}\text{IVent}_t^{ik\ 2} \\ & +\kappa_2\text{IVent}_t^{ik}\left(\text{Flare}_t^{ik} - \text{NRF}_t^{ik}\right) + \tilde{\kappa}_3\text{M}_t^{ik}\ . \end{aligned} \tag{6.153}$$

2. The marginal cost (including the specific tax on gas sales) of gas production for oil fields $MC_t^{ik,\text{Gas}} + \tau_t^{ik,\text{Gas}}$ can be decomposed as follows:

$$MC_t^{ik,\text{Gas}} + \tau_t^{ik,\text{Gas}} = \overline{MC}^{ik,\text{Gas}} + \overline{\tau}^{ik,\text{Gas}} + \nu_t^{ik}$$

   where ν_t^{ik} is independent across fields normally distributed with zero mean and standard deviation σ_ν, potentially serially correlated, and such that $E\left[\nu_t^{ik}\right] = 0$.

3. We impose the following parametric restrictions regarding the effect of maintenance on capital on the level of unintentional venting and on firm's costs: $\frac{\partial Maint_t^{ik}(\cdot)}{\partial \text{M}_t^{ik}} = \kappa_4$ and $MPE_{t+1}^{ik}\left(\text{M}_{t+1}^{ik}\right) = \kappa_5 + \kappa_6\text{M}_t^{ik}$. Moreover, we assume that the level of maintenance capital is strictly positive for all fields in our sample. Given these assumptions, the formula for the optimal level of maintenance capital writes:

$$\begin{aligned} \text{M}_t^{ik} = & -\tfrac{\kappa_5}{\kappa_6} + \tfrac{1-T_{t-1}^s}{\beta\kappa_6(1-T_t)} - \left(1-\rho^{ik}\right) + \tfrac{\kappa_4}{\kappa_6}\tfrac{VF_t^{ik}\kappa_3}{1-T_t} \\ & + \tfrac{\kappa_4}{\kappa_6}\left[P_t^{\text{Gas}} - \tau_t^{ik,\text{Gas}} - MC_t^{ik,\text{Gas}}\right] \end{aligned} \tag{6.154}$$

4. We assume zero flaring fines; i.e., $FF_t^{ik} = 0$, and time-invariant $T_t^s = T^s$, $\tau_t^{ik,\text{Gas}} = \tau^{ik,\text{Gas}}$ and $VF_t^{ik} = VF^{ik}$. As a result, the formula for regulatory costs writes:

$$\begin{aligned} RegCost_t^{ik}\left(\text{IVent}_t^{ik}, \text{Flare}_t^{ik}, \text{M}_t^{ik}\right) = \quad & const^{ik} + \kappa_0^{ik}\text{IVent}_t^{ik} + \tfrac{\kappa_1}{2}\text{IVent}_t^{ik\ 2} \\ & +\kappa_2\text{IVent}_t^{ik}\left(\text{Flare}_t^{ik} - \text{NRF}_t^{ik}\right) + \kappa_3 P_t^{\text{Gas}} + \upsilon_t^{ik} \end{aligned}\ . \tag{6.155}$$

   where $\kappa_3 = \tilde{\kappa}_3\frac{\kappa_4}{\kappa_6}$ $const^{ik} = -\kappa_3 + \frac{\tilde{\kappa}_3}{\beta\kappa_6} - \left(1-\rho^{ik}\right) + \kappa_3\frac{VF^{ik}\kappa_3}{1-T} - \kappa_3\left[\tau^{ik,\text{Gas}} + MC^{ik,\text{Gas}}\right]$ and $\upsilon_t^{ik} = -\kappa_3\nu_t^{ik}$

5. The marginal cost of flaring $MC_t^{ik,\text{Flare}}$ has linear formula:

$$MC_t^{ik,\text{Flare}}(1-T_t) = \pi_0^{ik} + \pi_1\left(\text{Flare}_t^{ik} - \text{NRF}_t^{ik}\right) \tag{6.156}$$

6. We assume $\kappa_0^{ik} \simeq \pi_0^{ik}$. That is, the marginal expected cost of flaring and intentional venting both evaluated at $\text{Flare}_t^{ik} = 0$ and $\text{IVent}_t^{ik} = 0$ are approximately equal to each other. For instance, it is plausible to assume that for very low levels of flaring and intentional venting, the marginal expected cost of both flaring and intentional venting is very close to zero. Moreover, we assume $\pi_1 > 0$, $\kappa_1 > 0$, $\kappa_1 > \kappa_2$ and $\pi_1 > \kappa_2$ to ensures that the costs are convex and, in turn, that the supply substitution matrix is positive definite. Note that the last two inequalities also imply that the own-price effects on the supply of Flare_t^{ik} and IVent_t^{ik} are greater in magnitude than the cross-price effects.

7. We assume a field-specific level of minimum routine flaring NRF_t^{ik}, which is allowed to depend upon the realization of the field's marginal cost of gas production; i.e., $\text{NRF}_t^{ik} = \zeta_0^{ik} + \zeta_1\left(MC_t^{ik,\text{Gas}} + \tau_t^{ik}\right)$.

Given this assumption, the product of parameters $\pi_1\zeta_1$ captures the extent of the co-movement between the marginal cost of gas production and the marginal cost of flaring due, for instance, to changes in electricity prices that affect both type of costs.

6.3.2 Identification: Structural Equations

We define the following variable:

$$\text{OtherGas}_t^{ik} = \text{TotGas}_t^{ik} - \text{Gas}_t^{ik} - \text{Flare}_t^{ik} \, , \tag{6.157}$$

which is observable by the econometrician because both TotGas_t^{ik}, Flare_t^{ik} and Gas_t^{ik} are. Using the formula for TotGas_t^{ik}, the definition above implies:

$$\text{OtherGas}_t^{ik} = \text{UVent}_t^{ik} + \text{IVent}_t^{ik} + \text{ReInj}_t^{ik} + \text{ReInS}_t^{ik} \tag{6.158}$$

In order to identify the key parameters of the structural model from equation (6.158), one faces several issues, which are summarized below.

1. The variable IVent_t^{ik} is not observed by the econometrician separately from UVent_t^{ik} and is censored at $\text{IVent}_t^{ik} = 0$, because a field cannot run negative intentional venting. Thus, one cannot neither directly observe whether $\text{IVent}_t^{i} > 0$ or not.

2. The realizations of $\frac{\nu_t^{ik}}{\kappa_1}$ which enter the formula for IVent_t^{ik} are not observed separately from any shock that enters UVent_t^{ik}. Thus, one cannot directly identify IVent_t^{ik} separately from UVent_t^{ik} .

3. Note that the model implies a link between $\text{Flare}_t^{ik} > \text{NRF}_t^{ik}$ and $\text{IVent}_t^{ik} > 0$. Thus, one could, in principle, use the value of Flare_t^{ik} as a selection tool to separate IVent_t^{ik} from UVent_t^{ik}. However, even if one could separate IVent_t^{ik} from UVent_t^{ik} using Flare_t^{ik}, other identification issues may occur. In particular, note that the formulas for IVent_t^{ik} and Flare_t^{ik} imply $E\left[\frac{\nu_t^{ik}}{\kappa_1} \mid \text{Flare}_t^{ik} > \text{NRF}_t^{ik}\right] \neq 0$. Because $\frac{\nu_t^{ik}}{\kappa_1}$ is part of IVent_t^{ik}, any attempt to use Flare_t^{ik} to quantify IVent_t^{ik} must account for the fact that the stochastic component of IVent_t^{ik} is not independent of Flare_t^{ik}.

4. We do not observes the values of ReInj_t^{i} and ReInS_t^{i}, UVent_t^{ik}. We only observe (estimates of) $\text{GasInj}_t^{i} = \text{ReInj}_t^{i} + \text{PInj}_t^{i}$ and $\text{GasInS}_t^{i} = \text{ReInS}_t^{i} + \text{PInS}_t^{i}$.

In the next sections we propose a method to solve all these issues. Specifically, sections 3.3 and 3.4 describe a two-step procedure to identify IVent_t^{ik} separately from UVent_t^{ik}. In section 3.5 and 3.6 we illustrate the methodology adopted to quantify ReInj_t^{i} and ReInS_t^{i}, respectively. Section 3.7 present the final empirical equation derived in sections 3.3. 3.4, 3.5 and 3.6. Lastly section 3.8 describes a procedure to identify the effect of a flaring tax, and section 3.9 provides a discussion of the main assumptions underpinning identification.

6.3.3 Identification: Intentional Venting

Using the parametric restriction from section 3.1 and the optimality conditions from section 2, we obtain the following formulas for the optimal levels of flaring and intentional venting:

$$\text{IVent}_t^{ik} = \max\left\{ -\frac{\kappa_0^{ik} + \kappa_2^{ik}\left(\text{Flare}_t^{ik} - \text{NRF}_t^{ik}\right)}{\kappa_1} + \frac{1}{\kappa_1}\left[\overline{MC}^{ik,\text{Gas}} + \tau^{ik} - P_t^{\text{Gas}} + \nu_t^{ik}\right], 0 \right\} \tag{6.159}$$

and

$$\text{Flare}_t^{ik} = \max\left\{-\frac{\pi_0^{ik} + \tau_t^{ik,\text{Flare}} + \kappa_2^{ik}\text{IVent}_t^{ik}}{\pi_1} + RF_t^{ik} + \frac{1}{\pi_1}\left[\overline{MC}^{ik,\text{Gas}} + \tau^{ik} - P_t^{\text{Gas}} + \nu_t^{ik}\right], \text{NRF}_t^{ik}\right\}, \tag{6.160}$$

respectively. This section makes use of the assumption $\kappa_0^{ik} \simeq \pi_0^{ik}$ to solve the first identification problem stated in section 3.2. Specifically, this assumption implies that if flaring is strictly positive, intentional venting is also strictly positive – although the two quantities may obviously differ from each other. In other words, observing a value for flaring $\text{Flare}_t^{ik} > \text{NRF}_t^{ik}$ is sufficient to identify $\text{IVent}_t^{ik} > 0$. This result is summarized by the following proposition.

Proposition 5. *If $\kappa_0^{ik} = \pi_0^{ik}$ and $\tau_t^{ik,Flare} = 0$ then $Flare_t^{ik} > NRF_t^{ik}$ if and only if $IVent_t^{ik} > 0$.*

Proof. Step 1. Suppose $\text{Flare}_t^{ik} > \text{NRF}_t^{ik}$ and $\text{IVent}_t^{ik} = 0$. The formula for Flare_t^{ik}:

$$\text{Flare}_t^{ik} = \max\left\{-\frac{\pi_0^{ik} + \tau_t^{ik,\text{Flare}} + \kappa_2^{ik}\text{IVent}_t^{ik}}{\pi_1} + \text{NRF}_t^{ik} + \frac{1}{\pi_1}\left[MC_t^{ik,\text{Gas}} + \tau_t^{ik,\text{Gas}} - P_t^{\text{Gas}}\right], \text{NRF}_t^{ik}\right\} \tag{6.161}$$

implies that $\text{Flare}_t^{ik} > \text{NRF}_t^{ik}$ at $\text{IVent}_t^{ik} = 0$ only if:

$$MC_t^{ik,\text{Gas}} + \tau_t^{ik,\text{Gas}} - P_t^{\text{Gas}} - \pi_0^{ik} - \tau_t^{ik,\text{Flare}} > 0 \tag{6.162}$$

From the formula of IVent_t^{ik}, at $\text{Flare}_t^{ik} > \text{NRF}_t^{ik}$ we get that $\text{IVent}_t^{ik} = 0$ only if the latent variable IVent_t^{ik*} is weakly negative, i.e.,

$$\text{IVent}_t^{ik*} = \frac{\pi_1\left(\pi_0^{ik} + \tau_t^{ik,\text{Flare}} - \kappa_0^{ik}\right)}{\kappa_1\pi_1 - \left(\kappa_2^{ik}\right)^2} + \frac{\pi_1 - \kappa_2^{ik}}{\kappa_1\pi_1 - \left(\kappa_2^{ik}\right)^2}\left[MC_t^{ik,\text{Gas}} + \tau_t^{ik,\text{Gas}} - P_t^{\text{Gas}} - \pi_0^{ik} - \tau_t^{ik,\text{Flare}}\right] \leq 0 \tag{6.163}$$

Because the second term of (6.163) is strictly positive by result (6.162), $\text{IVent}_t^{ik} = 0$ is true only if:

$$\tau_t^{ik,\text{Flare}} - \kappa_0^{ik} + \pi_0^{ik} < 0 \tag{6.164}$$

which for $\tau_t^{ik,\text{Flare}} = 0$ implies

$$\pi_0^{ik} < \kappa_0^{ik} \tag{6.165}$$

which leads to a contradiction of the assumption $\kappa_0^{ik} = \pi_0^{ik}$. Thus, if $\kappa_0^{ik} = \pi_0^{ik}$, then $\text{Flare}_t^{ik} > 0 \longrightarrow \text{IVent}_t^{ik} > 0$.

Step 2. Suppose $\text{Flare}_t^{ik} = \text{NRF}_t^{ik}$ and $\text{IVent}_t^{ik} > 0$. From the formula of IVent_t^{ik}, at $\text{Flare}_t^{ik} = \text{NRF}_t^{ik}$ we get:

$$\text{IVent}_t^{ik*} = \frac{1}{\kappa_1}\left[MC_t^{ik,\text{Gas}} + \tau_t^{ik,\text{Gas}} - P_t^{\text{Gas}} - \kappa_0^{ik}\right] > 0 \tag{6.166}$$

where the term is square brackets must be strictly positive; i.e.,

$$MC_t^{ik,\text{Gas}} + \tau_t^{ik,\text{Gas}} - P_t^{\text{Gas}} - \kappa_0^{ik} > 0 \tag{6.167}$$

The formula for Flare_t^{ik} implies that $\text{Flare}_t^{ik} = \text{NRF}_t^{ik}$ only if:

$$\text{Flare}_t^{ik*} = -\frac{\pi_0^{ik} + \tau_t^{ik,\text{Flare}} + \kappa_2^{ik}\text{IVent}_t^{ik}}{\pi_1} + \text{NRF}_t^{ik} + \frac{1}{\pi_1}\left[MC_t^{ik,\text{Gas}} + \tau_t^{ik,\text{Gas}} - P_t^{\text{Gas}}\right] \leq \text{NRF}_t^{ik} \tag{6.168}$$

Substituting the formula for IVent_t^{ik} and $\pi_0^{ik} = \kappa_0^{ik}$ we obtain

$$-\kappa_1 \frac{\pi_0^{ik} - \kappa_0^{ik} + \tau_t^{ik,\text{Flare}}}{\pi_1\kappa_1 - \left(\kappa_2^{ik}\right)^2} + \frac{\kappa_1 - \kappa_2^{ik}}{\pi_1\kappa_1 - \left(\kappa_2^{ik}\right)^2}\left[MC_t^{ik,\text{Gas}} + \tau_t^{ik,\text{Gas}} - P_t^{\text{Gas}} - \kappa_0^{ik}\right] \leq 0 \tag{6.169}$$

because the second term of (6.169) is strictly positive by result (6.167), $\text{Flare}_t^{ik} = \text{NRF}_t^{ik}$ is true only if:

$$\pi_0^{ik} - \kappa_0^{ik} + \tau_t^{ik,\text{Flare}} > 0 \tag{6.170}$$

which for $\tau_t^{ik,\text{Flare}} = 0$ implies

$$\pi_0^{ik} > \kappa_0^{ik} \tag{6.171}$$

which leads to a contradiction of the assumption $\kappa_0^{ik} = \pi_0^{ik}$. Thus, if $\kappa_0^{ik} = \pi_0^{ik}$, then $\text{Flare}_t^{ik} > \text{NRF}_t^{ik} \longleftrightarrow \text{IVent}_t^{ik} > 0$. Q.E.D.

Step 1: Selection Equation

Proposition 5 allows us to establish whether $\text{IVent}_t^{ik} > 0$ or not simply using the value of the observable variable Flare_t^{ik}, providing in turn an empirical tool to tackle the identification problem number 1 in section 3.2. Moreover, this finding also provides a route to tackle identification problem 2. Specifically, we can introduce in the structural equation (6.158) a dummy $\text{Flare}_t^{ik} > \text{NRF}_t^{ik}$ to separate the effect of a given variable on IVent_t^{ik} from the effect of the same variable on UVent_t^{ik}. However, in order to quantify IVent_t^{ik} we still need to tackle identification problem 3. We solve this problem using a two-step approach. Using the result in Proposition 5 we know that $\text{Flare}_t^{ik} > \text{NRF}_t^{ik} \longleftrightarrow \text{IVent}_t^{ik} > 0$. Thus, we can substitute the formula for IVent_t^{ik} into the formula for Flare_t^{ik} to obtain:

$$\text{Flare}_t^{ik} = \alpha_0^{ik} + \alpha_1 P_t^{\text{Gas}} + \eta_t^{ik} \tag{6.172}$$

where using $\text{NRF}_t^{ik} = \zeta_0^{ik} + \zeta_1\left(MC_t^{ik,\text{Gas}} + \tau_t^{ik,\text{Gas}}\right)$ we obtain the parameters α_0^{ik}, α_1 and the stochastic component η_t^{ik}, which have formulas:

$$\begin{array}{ll} \alpha_0^{ik} = & \frac{\kappa_1-\kappa_2}{\kappa_1\pi_1-\kappa_2^2}\left(\overline{MC}^{ik,\text{Gas}} + \overline{\tau}^{ik,\text{Gas}} - \kappa_0^{ik}\right) + \zeta_0^{ik} + \zeta_1\left(\overline{MC}^{ik,\text{Gas}} + \overline{\tau}^{ik,\text{Gas}}\right) \\ \alpha_1 = & -\frac{\kappa_1-\kappa_2}{\kappa_1\pi_1-\kappa_2^2} \\ \eta_t^{ik} = & \left(\frac{\kappa_1-\kappa_2}{\kappa_1\pi_1-\kappa_2^2} + \zeta_1\right)\nu_t^{ik} \end{array} \tag{6.173}$$

respectively. Note that the equation in (6.172) resembles that of a standard *censored regression*. Thus, it can be estimated using standard tools for censored data. In particular, we use a version of the Tobit type I model that allows for serial correlation on the η_t^{ik}'s. Having estimated the parameters α_0^{ik}, α_1^{ik} using a consistent estimator, we can calculate the residuals for each observation such that $\text{Flare}_t^{ik} > 0$ as follows:

$$\widehat{res}_t^{ik} = \text{Flare}_t^{ik} - \hat{\alpha}_0^{ik} - \hat{\alpha}_1 P_t^{\text{Gas}}$$

where $\widehat{res}_t^{ik}$ represents an estimate of $\eta_t^{ik} = \left(\frac{\kappa_1-\kappa_2}{\kappa_1\pi_1-\kappa_2^2} + \zeta_1\right)\nu_t^{ik}$ for the cases in which Flare_t^{ik}. Conversely, when $\text{Flare}_t^{ik} = 0$ we do not possess a point estimate for η_t^{ik}. However, we can construct the expected value of η_t^{ik} conditional on $\text{Flare}_t^{ik} = 0$ using the normality assumption. Specifically, the conditional expectation of η_t^{ik} has formula:

$$E\left[\eta_t^{ik} \mid \eta_t^{ik} \leq -\alpha_0^{ik} - \alpha_1 P_t^{\text{Gas}}\right] = -\hat{\sigma}_\eta \frac{\phi\left(\alpha_0^{ik} + \alpha_1 P_t^{\text{Gas}}/\sigma_\eta\right)}{\Phi\left(\alpha_0^{ik} + \alpha_1 P_t^{\text{Gas}}/\sigma_\eta\right)} \tag{6.174}$$

Substituting α_0^{ik}, α_1, and σ_η with their estimators obtained from the censored regression in (6.172) into the formula above, we can construct the estimator for η_t^{ik} as follows:

$$\hat{\eta}_t^{ik} = \begin{cases} \widehat{res}_t^{ik} & if\ \text{Flare}_t^{ik} > \text{NRF}_t^{ik}. \\ -\hat{\sigma}_\eta \frac{\phi\left(-\widehat{\text{Flare}}_t^{ik}/\hat{\sigma}_\eta\right)}{\Phi\left(-\widehat{\text{Flare}}_t^{ik}/\hat{\sigma}_\eta\right)} & otherwise \end{cases} \tag{6.175}$$

Thus, we can use this estimate in the main structural equation (6.158) in order to quantify IVent_t^{ik} accounting for the fact that $E\left[\nu_t^{ik} \mid \text{Flare}_t^{ik} > \text{NRF}_t^{ik}\right] \neq 0$. This procedure is illustrated in the next subsection.

Step 2: Formula for IVent_t^{ik} Conditional on Flare_t^{ik}

Using the result in Proposition 5, the formula for IVent_t^{ik} conditional on $\text{Flare}_t^{ik} > 0$ writes:

$$\text{IVent}_t^{ik} = -\frac{\kappa_2 \tau_t^{ik,\text{Flare}}}{\pi_1\kappa_1 - \kappa_2^2} + \frac{\pi_1 - \kappa_2}{\pi_1\kappa_1 - \kappa_2^2}\left(\overline{MC}^{ik,\text{Gas}} + \overline{\tau}^{ik,\text{Gas}} - \kappa_0^{ik}\right) - \frac{\pi_1 - \kappa_2}{\pi_1\kappa_1 - \kappa_2^2} P_t^{\text{Gas}} + \frac{\pi_1 - \kappa_2}{\kappa_1\pi_1 - \kappa_2^2}\nu_t^{ik} \tag{6.176}$$

Thus, we can rewrite the formula for IVent_t^{ik} as follows:

$$\text{IVent}_t^{ik} = \delta_1^{ik} D_t^{ik} + \delta_2 P_t^{\text{Gas}} D_t^{ik} + \delta_3 \eta_t^{ik} D_t^{ik} \tag{6.177}$$

where D_t^{ik} is a dummy variable defined as follows:

$$D_t^{ik} = \begin{cases} 1 & if\ \text{Flare}_t^{ik} > \text{NRF}_t^{ik} \\ 0 & otherwise \end{cases} \tag{6.178}$$

and where the parameters δ_1^{ik}, δ_2, δ_3 have formulas:

$$\begin{aligned} \delta_1^{ik} &= \frac{\pi_1-\kappa_2}{\pi_1\kappa_1-\kappa_2^2}\left(\overline{MC}^{ik,\text{Gas}} + \overline{\tau}^{ik,\text{Gas}} - \kappa_0^{ik}\right) - \frac{\kappa_2\tau_t^{ik,\text{Flare}}}{\pi_1\kappa_1-\kappa_2^2} \\ \delta_2 &= -\frac{\pi_1-\kappa_2}{\pi_1\kappa_1-\kappa_2^2} \\ \delta_3 &= \frac{\pi_1-\kappa_2}{\kappa_1\pi_1-\kappa_2^2}\left(\zeta_1 + \frac{\kappa_1-\kappa_2}{\kappa_1\pi_1-\kappa_2^2}\right)^{-1} \end{aligned} \tag{6.179}$$

Lastly, note that η_t^{ik} is not observable by the econometrician. However, from section 3.3 we know that we can construct an estimate of η_t^{ik}, with formula (6.175), which we can use in equation (6.177). Note that this method allows us to tackle the identification issue number 2 in section 3.2.

6.3.4 Identification: Unintentional Venting

Using the binding constraint for oil fields $\text{TotGas}_t^{ik} = GOR^{ik}\text{Oil}_t^{ik}$ and the optimality condition for M_t^{ik} in (6.154) we get the formula for UVent_t^{ik}:

$$\begin{aligned} \text{UVent}_t^{ik} = \; & \frac{\kappa_4\kappa_5}{\kappa_6} + \kappa_4\left(1-\rho^{ik}\right) - \frac{(\kappa_4)^2}{\kappa_6}\left(\frac{VF^{ik}\kappa_3}{1-T} - \overline{MC}^{ik,\text{Gas}} - \overline{\tau}^{ik,\text{Gas}}\right) \\ & -\frac{\kappa_4}{\beta\kappa_6} + \vartheta^{ik} GOR^{ik}\text{Oil}_t^{ik} - \frac{(\kappa_4)^2}{\kappa_6}P_t^{\text{Gas}} + \frac{(\kappa_4)^2}{\kappa_6}\nu_t^{ik} + \epsilon_t^{ik} \end{aligned} \tag{6.180}$$

where $E\left[\epsilon_t^{ik}\right] = 0$. Note that ν_t^{ik} is not observable by the econometrician. However, from section 3.3 we know that we can construct an estimate of $\eta_t^{ik} = \left(\frac{\kappa_1-\kappa_2}{\kappa_1\pi_1-\kappa_2^2} + \zeta_1\right)\nu_t^{ik}$, with formula (6.175), which we can use in equation (6.180) by applying simple change of variable $\nu_t^{ik} = \left(\frac{\kappa_1-\kappa_2}{\kappa_1\pi_1-\kappa_2^2} + \zeta_1\right)^{-1}\eta_t^{ik}$. Thus, the formula for UVent_t^{ik} can be written as follows:

$$\text{UVent}_t^{ik} = \delta_0^{ik} + \delta_4^{ik}\text{Oil}_t^{ik} + \delta_5 P_t^{\text{Gas}} + \delta_6\eta_t^{ik} + \epsilon_t^{ik} \tag{6.181}$$

where the parameters possess the following formulas:

$$\begin{array}{ll} \delta_0^{ik} = & \frac{\kappa_4\kappa_5}{\kappa_6} - \frac{\kappa_4}{\beta\kappa_6} + \kappa_4\left(1-\rho^{ik}\right) - \frac{(\kappa_4)^2}{\kappa_6}\left(\frac{VF^{ik}\kappa_3}{1-T} - \overline{MC}^{ik,\text{Gas}} - \overline{\tau}^{ik,\text{Gas}}\right) \\ \delta_4^{ik} = & \vartheta^{ik}GOR^{ik} \\ \delta_5 = & -\frac{(\kappa_4)^2}{\kappa_6} \\ \delta_6 = & \frac{(\kappa_4)^2}{\kappa_6}\left(\frac{\kappa_1-\kappa_2}{\kappa_1\pi_1-\kappa_2{}^2} + \zeta_1\right)^{-1} \end{array} \tag{6.182}$$

The formula in (6.181) represents the second object of the structural equation we aim to estimate.

6.3.5 Gas Injections

We assume that the cost of performing gas injections for firm i has the following functional form:

$$IC_t^{ik}\left(\text{ReInj}_t^{ik}, \text{PInj}_t^{ik}\right) = a_1^{ik}\text{ReInj}_t^{ik} + a_2^{ik}\text{PInj}_t^{ik} + A^{ik}\left(\text{ReInj}_t^{ik}, \text{PInj}_t^{ik}\right) \tag{6.183}$$

where $A^{ik} : (-\infty, +\infty)^2 \to \mathbb{R}$ is a strictly convex, homogeneous of degree $\varrho^{ik} > 0$ real analytic function that satisfies $A_1^{ik}\left(0, \text{PInj}_t^{ik}\right) + a_1^{ik} \leq 0$ for all $\text{PInj}_t^{ik} > 0$ and $A_2^{ik}\left(\text{ReInj}_t^{ik}, 0\right) + a_2^{ik} + PP_t^{ik,\text{Gas}} \leq 0$ for all $\text{ReInj}_t^{ik} > 0$, where A_j^{ik} denotes the first derivative of A^{ik} with respect to its jth argument. Using formula (6.183) into the FOCs of the oil firm problem w.r.t. $\text{ReInj}_t^{ik}, \text{PInj}_t^{ik}$, and OInj_t^{ik} we get:

$$\begin{array}{ll} \frac{\partial\mathcal{L}_u^k}{\partial\text{ReInj}_t^{ik}} = & -\left[A_1^{ik}\left(\text{ReInj}_t^{ik}, \text{PInj}_t^{ik}\right) + a_1^{ik}\right](1-T_t) \\ & +\left(-\phi_{2t}^{ik} + \phi_{3t}^{ik} + \phi_{4t}^{ik}\zeta\right)\left(1-\vartheta^{ik}\right)^{-1} + \phi_{4t}^{ik}\frac{\partial B_t^{ik}\left(\text{TotInj}_t^{ik}\right)}{\partial\text{ReInj}_t^{ik}} + \phi_{8t}^{ik} = 0 \\ \frac{\partial\mathcal{L}_u^k}{\partial\text{PInj}_t^{ik}} = & -\left[A_2^{ik}\left(\text{ReInj}_t^{ik}, \text{PInj}_t^{ik}\right) + a_2^{ik} + PP_t^{ik,\text{Gas}}\right](1-T_t) + \phi_{4t}^{i}\frac{\partial B_t^{ik}\left(\text{TotInj}_t^{ik}\right)}{\partial\text{PInj}_t^{ik}} = 0 \\ \frac{\partial\mathcal{L}_u^k}{\partial\text{OInj}_t^{ik}} = & -C_{3t}^{ik}(1-T_t) + \phi_{4t}^{i}\frac{\partial B_t^{i}\left(\text{TotInj}_t^{ik}\right)}{\partial\text{OInj}_t^{i}} = 0 \end{array} \tag{6.184}$$

Note that the assumptions on A^{ik} implies that $\text{ReInj}_t^{ik} > 0$ if $\text{PInj}_t^{ik} > 0$, implying in turn that the observable variable total gas injection $\text{GasInj}_t^{ik} > 0$ is sufficient to establish that both $\text{ReInj}_t^{ik} > 0$ and $\text{PInj}_t^{ik} > 0$ must hold true, and therefore $\phi_{8t}^{ik} = 0$. Using the optimality condition of the oil firm's problem w.r.t. Gas_t^{ik}:

$$\left[P_t^{\text{Gas}} - \tau_t^{ik,\text{Gas}} - MC_t^{ik,\text{Gas}}\right](1-T_t) + \left(-\phi_{2t}^{ik} + \phi_{3t}^{ik} + \phi_{4t}^{ik}\zeta\right)\left(1-\vartheta^{ik}\right)^{-1} = 0 \tag{6.185}$$

and recalling that the third condition is always binding (see proof to Proposition 3), the optimality conditions for any field such that $\text{GasInj}_t^{ik} > 0$ become:

$$-\left[A_1^{ik}\left(\text{ReInj}_t^{ik}, \text{PInj}_t^{ik}\right) + a_1^{ik}\right] - \left(P_t^{\text{Gas}} - \tau_t^{ik,\text{Gas}} - MC_t^{ik,\text{Gas}}\right) - C_{3t}^{ik} = 0 \tag{6.186}$$

and

$$-\left[A_2^{ik}\left(\text{ReInj}_t^{ik}, \text{PInj}_t^{ik}\right) + a_2^{ik} + PP_t^{ik,\text{Gas}}\right] - C_{3t}^{ik} = 0 \tag{6.187}$$

Combining the two conditions and using the homogeneity of A^{ik}, we get:

$$h^{ik\ -1}\left(\frac{\text{PInj}_t^{ik}}{\text{ReInj}_t^{ik}}\right) = \frac{C_{3t}^{ik} - PP_t^{ik,\text{Gas}} - a_2^{ik}}{C_{3t}^{ik} - P_t^{i,Gas} + MC_t^{ik,\text{Gas}} + \tau_t^{ik,\text{Gas}} - a_1^{ik}} \tag{6.188}$$

for some real analytic function h^{ik} with formula:

$$h^{ik\ -1}\left(\frac{\text{PInj}_t^{ik}}{\text{ReInj}_t^{ik}}\right) = \frac{A_2^{ik}\left(\text{ReInj}_t^{ik}, \text{PInj}_t^{ik}\right)}{A_1^{ik}\left(\text{ReInj}_t^{ik}, \text{PInj}_t^{ik}\right)}, \tag{6.189}$$

where $h^{ik\ -1}$ is a function of $\frac{\text{PInj}_t^{ik}}{\text{ReInj}_t^{ik}}$ because of the homogeneity of A^{ik}, and solves for:

$$\text{PInj}_t^{ik} = h^{ik}\left(\frac{C_t^{ik} - PP_t^{ik,\text{Gas}} - a_2^{ik}}{C_t^{ik} - P_t^{\text{Gas}} + MC_t^{ik,\text{Gas}} + \tau_t^{ik,\text{Gas}} - a_1^{ik}}\right)\text{ReInj}_t^{ik} \tag{6.190}$$

Using the formula for gas injections:

$$\text{GasInj}_t^{ik} = \text{ReInj}_t^{ik} + \text{PInj}_t^{ik} \tag{6.191}$$

we obtain

$$\text{GasInj}_t^{i} = \left[1 + h^{ik}\left(\frac{C_t^{ik} - PP_t^{ik,\text{Gas}} - a_2^{ik}}{C_t^{ik} - P_t^{\text{Gas}} + MC_t^{ik,\text{Gas}} + \tau_t^{ik,\text{Gas}} - a_1^{ik}}\right)\right]\text{ReInj}_t^{i} \tag{6.192}$$

Lastly we use the formulas for $PP_t^{ik,\text{Gas}}$:

$$PP_t^{ik,\text{Gas}} = P_t^{\text{Gas}} - \sigma_2^{ik}, \tag{6.193}$$

where σ_2^{ik} captures the segmentation of the local natural gas market for injections, and for $MC_t^{ik,\text{Gas}} + \tau_t^{ik,\text{Gas}}$:

$$MC_t^{ik,\text{Gas}} + \tau_t^{ik,\text{Gas}} = \overline{MC}^{ik,\text{Gas}} + \overline{\tau}^{ik,\text{Gas}} + \nu_t^{ik} \tag{6.194}$$

into equation (6.192) to obtain:

$$\text{ReInj}_t^{ik} = SR^{ik}\left(P_t^{\text{Gas}}, \nu_t^{ik}\right)\text{GasInj}_t^{ik} \tag{6.195}$$

where $SR^{ik}\left(P_t^{\text{Gas}}, \nu_t^{ik}\right)$ is a real analytic function with formula:

$$SR^{ik}\left(P_t^{\text{Gas}}, \nu_t^{ik}\right) = \left[1 + h^{ik}\left(\frac{\sigma_2^{ik} - a_2^{ik} - P_t^{\text{Gas}}}{\overline{MC}^{ik,\text{Gas}} + \overline{\tau}^{ik,\text{Gas}} - a_1^{ik} + \nu_t^{ik} - P_t^{\text{Gas}}}\right)\right]^{-1} \tag{6.196}$$

The formula in (6.195) show that, whenever $\text{GasInj}_t^{ik} > 0$, the use of own gas and purchased gas for gas injection is regulated by a share SR^{ik} which is a function of the gas price, the realized shock on the firm marginal cost and other non time-variant variables. Because h^{ik} is a real analytic function, $SR_t^{ik}\left(P_t^{\text{Gas}}, \nu_t^{ik}\right)$ also is also real analytic. Thus, it is twice differentiable. This property allows us to define all the derivatives of SR^{ik} evaluated at $P_t^{\text{Gas}} = \overline{P}_t^{\text{Gas}}, \nu_t^{ik} = \overline{\nu}_t^{ik}$ as follows:

$$SR_{jl}^{ik} = \left.\frac{\partial^{j+k} SR^{ik}\left(P_t^{\text{Gas}}, \nu_t^{ik}\right)}{\partial\left(P_t^{\text{Gas}}\right)^j \partial\left(\nu_t^{i}\right)^l}\right|_{P_t^{\text{Gas}} = \overline{P}_t^{\text{Gas}}, \nu_t^{ik} = \overline{\nu}_t^{ik}} \tag{6.197}$$

Moreover, because SR^{ik} is a real analytic function, the formula for ReInj_t^{ik} can be written as a Taylor Series:

$$\text{ReInj}_t^{ik} = \left[\sum_{j=0}^{\infty}\sum_{j=0}^{\infty}\frac{SR_{jl}^{ik}}{j!l!}\left(P_t^{\text{Gas}} - \overline{P}_t^{\text{Gas}}\right)^j\left(\nu_t^{ik} - \overline{\nu}_t^{ik}\right)^l\right]\text{GasInj}_t^{ik} \tag{6.198}$$

However, the formula above is not suitable for empirical purposes. Therefore, we rely on an approximate a formula for ReInj_t^{ik} using a Jth-order Taylor approximation:

$$\text{ReInj}_t^{ik} \simeq \left[\sum_{j=0}^{J}\sum_{l=0}^{J}\frac{SR_{jl}^{ik}}{j!l!}\left(P_t^{\text{Gas}} - \overline{P}_t^{\text{Gas}}\right)^j\left(\nu_t^{ik} - \overline{\nu}_t^{ik}\right)^l\right]\text{GasInj}_t^{ik} \tag{6.199}$$

Specifically, for practical purposes in our empirical analysis we use a first-order (i.e., linear) Taylor approximation:

$$\text{ReInj}_t^{ik} \simeq \left(SR_{00}^{ik} - \overline{P}_t^{\text{Gas}} - \overline{\nu}_t^{ik}\right)\text{GasInj}_t^{ik} + SR_{10}^{ik}P_t^{\text{Gas}}\text{GasInj}_t^{ik} + SR_{01}^{ik}\nu_t^{ik}\text{GasInj}_t^{ik} \tag{6.200}$$

Lastly, note that ν_t^{ik} is not observable by the econometrician. However, from section 3.3 we know that we can construct an estimate of $\eta_t^{ik} = \left(\frac{\kappa_1-\kappa_2}{\kappa_1\pi_1-\kappa_2^2}+\zeta_1\right)\nu_t^{ik}$, with formula (6.175), which we can use in equation (6.200) by applying simple change of variable $\nu_t^{ik} = \left(\frac{\kappa_1-\kappa_2}{\kappa_1\pi_1-\kappa_2^2}+\zeta_1\right)^{-1}\eta_t^{ik}$. Thus, the approximate formula for ReInj_t^{ik} can be written as follows:

$$\text{ReInj}_t^{ik} = \delta_7^{ik}\text{GasInj}_t^{i} + \delta_8^{ik}P_t^{Gas}\times\text{GasInj}_t^{ik} + \delta_9^{ik}\eta_t^{ik}\times\text{GasInj}_t^{ik} \tag{6.201}$$

where the coefficients $\delta_7^{ik}, \delta_8^{ik}, \delta_9^{ik}$ are field-specific coefficients and in particular $\delta_9^{ik} = SR_{01}^{ik}\left(\frac{\kappa_1-\kappa_2}{\kappa_1\pi_1-\kappa_2^2}+\zeta_1\right)^{-1}$. Note that the coefficient in front of GasInj_t^{ik} always implies that ReInj_t^{ik} is a share $\in(0,1)$ of GasInj_t^{ik} as long as $\text{GasInj}_t^{ik}>0$, and $\text{ReInj}_t^{ik}=0$ whenever $\text{GasInj}_t^{ik}=0$ then , as expected.

6.3.6 In-Situ Use

We assume that the transformation function of field i in period t satisfies the following conditions:

$$\frac{\partial TF_{4t}^{ik}\left(\text{PInS}_t^{ik};\text{ReInS}_t^{ik}\right)}{\partial\text{ReInS}_t^{ik}}\Bigg/\frac{\partial TF_{6t}^{ik}\left(\text{Z}_t^{ik}\right)}{\partial\text{Z}_t^{ik}} = \left[b_1^{ik}+B_1^{ik}\left(\text{ReInS}_t^{i},\text{PInS}_t^{i}\right)\right](1-T_t)$$

and

$$\frac{\partial TF_{4t}^{ik}\left(\text{PInS}_t^{ik};\text{ReInS}_t^{ik}\right)}{\partial\text{PInS}_t^{ik}}\Bigg/\frac{\partial TF_{6t}^{ik}\left(\text{Z}_t^{ik}\right)}{\partial\text{Z}_t^{ik}} = \left[b_2^{ik}+B_2^{ik}\left(\text{ReInS}_t^{i},\text{PInS}_t^{i}\right)\right](1-T_t)$$

where $B^{ik}:(-\infty,+\infty)^2\to\mathbb{R}$ is a strictly convex, homogeneous of degree $\varphi^{ik}>0$ real analytic function that satisfies $B_1^{ik}\left(0,\text{PInS}_t^{ik}\right)+b_1^{ik}\le 0$ for all $\text{PInS}_t^{ik}>0$ and $B_2^{ik}\left(\text{ReInS}_t^{ik},0\right)+b_2^{ik}+PP_t^{ik,\text{Gas}}\le 0$ for all $\text{ReInS}_t^{ik}>0$, where B_j^{ik} denotes the first derivative of B^{ik} with respect to its jth argument. Note that the assumptions on B^{ik} implies that $\text{ReInS}_t^{ik}>0$ if $\text{PInS}_t^{ik}>0$, implying in turn that the observable variable total gas injection $\text{GasInS}_t^{ik}>0$ is sufficient to establish that both $\text{ReInj}_t^{ik}>0$ and $\text{PInj}_t^{ik}>0$ must hold true, and therefore $\phi_{9t}^{ik}=0$. Using formula (6.183) into the FOCs of the oil firm problem w.r.t. ReInS_t^{ik} and PInS_t^{ik}, we get:

$$\begin{aligned}\frac{\partial\mathcal{L}_u^k}{\partial\text{ReInS}_t^{ik}} &= -\left[B_1^{ik}\left(\text{ReInS}_t^{ik},\text{PInS}_t^{ik}\right)+b_1^{ik}\right](1-T_t)\\ &\quad+\left(-\phi_{2t}^{ik}+\phi_{3t}^{ik}+\phi_{4t}^{ik}\zeta\right)\left(1-\vartheta^{ik}\right)^{-1}+\phi_{9t}^{iks}=0\\ \frac{\partial\mathcal{L}_u^k}{\partial\text{PInS}_t^{ik}} &= -\left[B_2^{ik}\left(\text{ReInS}_t^{ik},\text{PInS}_t^{ik}\right)+b_2^{ik}+PP_t^{ik,\text{Gas}}\right](1-T_t)=0\end{aligned} \tag{6.202}$$

Combining the conditions above with the optimality condition of the firm's problem w.r.t. Gas_t^{ik},

$$\frac{\partial\mathcal{L}_u^k}{\partial\text{Gas}_t^{ik}} = \left[P_t^{\text{Gas}}-\tau_t^{ik,\text{Gas}}+MC_t^{ik,\text{Gas}}\right](1-T_t)+\left(-\phi_{2t}^{ik}+\phi_{3t}^{ik}+\phi_{4t}^{ik}\zeta\right)\left(1-\vartheta^{ik}\right)^{-1}=0 \tag{6.203}$$

the optimality conditions for any field such that $\text{GasInS}_t^{ik}>0$ become:

$$\begin{aligned}\frac{\partial\mathcal{L}_u^k}{\partial\text{ReInS}_t^{ik}} &= -\left[B_1^{ik}\left(\text{ReInS}_t^{ik},\text{PInS}_t^{ik}\right)+b_1^{ik}\right](1-T_t)\\ &\quad-\left[P_t^{\text{Gas}}-\tau_t^{ik,\text{Gas}}+MC_t^{ik,\text{Gas}}\right](1-T_t)+\phi_{9t}^{iks}=0\\ \frac{\partial\mathcal{L}_u^k}{\partial\text{PInS}_t^{ik}} &= -\left[B_2^{ik}\left(\text{ReInS}_t^{ik},\text{PInS}_t^{ik}\right)+b_2^{ik}+PP_t^{ik,\text{Gas}}\right](1-T_t)=0\end{aligned} \tag{6.204}$$

Using the formula for gas in-situ use:

$$\text{GasInS}_t^{ik} = \text{ReInS}_t^{ik}+\text{PInS}_t^{ik} \tag{6.205}$$

and following the same procedure used in the previous section for gas injections, we obtain the following results:

$$\text{ReInS}_t^i = \left[1 + g^{ik}\left(\frac{b_2^{ik} + PP_t^{ik,\text{Gas}}}{b_1^{ik} - MC_t^{ik,\text{Gas}} - \tau_t^{ik,\text{Gas}} + P_t^{\text{Gas}}}\right)\right] \text{GasInS}_t^i \tag{6.206}$$

or some real analytic function g^{ik}. Lastly we use the formulas for $PP_t^{ik,\text{Gas}}$:

$$PP_t^{ik,\text{Gas}} = P_t^{\text{Gas}} - \sigma_2^{ik} \tag{6.207}$$

and for $MC_t^{ik,\text{Gas}} + \tau_t^{ik,\text{Gas}}$:

$$MC_t^{ik,\text{Gas}} + \tau_t^{ik,\text{Gas}} = \overline{MC}^{ik,\text{Gas}} + \overline{\tau}^{ik,\text{Gas}} + \nu_t^{ik} \tag{6.208}$$

into equation (6.206) to obtain:

$$\text{ReInj}_t^i = SI^{ik}\left(P_t^{\text{Gas}}, \nu_t^{ik}\right) \text{GasInj}_t^i \tag{6.209}$$

where

$$SI^{ik}\left(P_t^{\text{Gas}}, \nu_t^{ik}\right) = \left[1 + g^{ik}\left(\frac{-b_2^{ik} - P_t^{\text{Gas}} + \sigma_2^{ik}}{-b_1^{ik} + \overline{MC}^{ik,\text{Gas}} + \overline{\tau}^{ik,\text{Gas}} + \nu_t^{ik} - P_t^{\text{Gas}}}\right)\right]^{-1} \tag{6.210}$$

The formula in (6.210) show that, whenever $\text{GasInS}_t^{ik} > 0$, the use of own gas and purchased gas for in-situ use is regulated by a share SI^{ik} which is a function of the gas price, the realized shock on the firm marginal cost and other non time-variant variables. Because g^{ik} is a real analytic function, $SI_t^{ik}\left(P_t^{\text{Gas}}, \nu_t^{ik}\right)$ also is also real analytic. Thus, it is twice differentiable. This property allows us to define all the derivatives of SI^{ik} evaluated at $P_t^{\text{Gas}} = \overline{P}_t^{\text{Gas}}, \nu_t^{ik} = \overline{\nu}_t^{ik}$ as follows:

$$SI_{jl}^{ik} = \left.\frac{\partial^{j+k} SI^{ik}\left(P_t^{\text{Gas}}, \nu_t^{ik}\right)}{\partial\left(P_t^{\text{Gas}}\right)^j \partial\left(\nu_t^i\right)^l}\right|_{P_t^{\text{Gas}} = \overline{P}_t^{\text{Gas}}, \nu_t^{ik} = \overline{\nu}_t^{ik}} \tag{6.211}$$

Moreover, because SI^{ik} is a real analytic function, the formula for ReInS_t^{ik} can be written as a Taylor Series:

$$\text{ReInS}_t^{ik} = \left[\sum_{j=0}^{\infty}\sum_{j=0}^{\infty} \frac{SI_{jl}^{ik}}{j!l!}\left(P_t^{\text{Gas}} - \overline{P}_t^{\text{Gas}}\right)^j \left(\nu_t^{ik} - \overline{\nu}_t^{ik}\right)^l\right] \text{GasInS}_t^{ik} \tag{6.212}$$

However, the formula above is not suitable for empirical purposes. Therefore, we rely on an approximate a formula for ReInS_t^{ik} using a Jth-order Taylor approximation:

$$\text{ReInS}_t^{ik} \simeq \left[\sum_{j=0}^{J}\sum_{l=0}^{J} \frac{SI_{jl}^{ik}}{j!l!}\left(P_t^{\text{Gas}} - \overline{P}_t^{\text{Gas}}\right)^j \left(\nu_t^{ik} - \overline{\nu}_t^{ik}\right)^l\right] \text{GasInS}_t^{ik} \tag{6.213}$$

Specifically, for practical purposes in our empirical analysis we use a first-order (i.e., linear) Taylor approximation:

$$\text{ReInS}_t^{ik} \simeq \left(SI_{00}^{ik} - \overline{P}_t^{\text{Gas}} - \overline{\nu}_t^{ik}\right) \text{GasInS}_t^{ik} + SI_{10}^{ik} P_t^{\text{Gas}} \text{GasInS}_t^{ik} + SI_{01}^{ik} \nu_t^{ik} \text{GasInS}_t^{ik} \tag{6.214}$$

Lastly, note that ν_t^{ik} is not observable by the econometrician. However, from section 3.3 we know that we can construct an estimate of $\eta_t^{ik} = \left(\frac{\kappa_1 - \kappa_2}{\kappa_1 \pi_1 - \kappa_2^2} + \zeta_1\right)\nu_t^{ik}$, with formula (6.175), which we can use in equation (6.214) by applying simple change of variable $\nu_t^{ik} = \left(\frac{\kappa_1 - \kappa_2}{\kappa_1 \pi_1 - \kappa_2^2} + \zeta_1\right)^{-1} \eta_t^{ik}$. Thus, the approximate formula for ReInS_t^{ik} can be written as follows:

$$\text{ReInS}_t^{ik} = \delta_{10}^{ik} \text{GasInS}_t^i + \delta_{11}^{ik} P_t^{Gas} \times \text{GasInS}_t^{ik} + \delta_{12}^{ik} \eta_t^{ik} \times \text{GasInS}_t^{ik} \tag{6.215}$$

where the coefficients $\delta_{10}^{ik}, \delta_{11}^{ik}, \delta_{12}^{ik}$ are field-specific coefficients and in particular $\delta_{12}^{ik} = SI_{01}^{ik}\left(\frac{\kappa_1-\kappa_2}{\kappa_1\pi_1-{\kappa_2}^2}+\zeta_1\right)^{-1}$. Note that the coefficient in front of GasInS_t^{ik} always implies that ReInS_t^{ik} is a share $\in (0,1)$ of GasInS_t^{ik} as long as $\text{GasInS}_t^{ik} > 0$, and $\text{ReInS}_t^{ik} = 0$ whenever $\text{GasInS}_t^{ik} = 0$ then , as expected.

6.3.7 Identification: Main Equation

The formulas for $\text{IVent}_t^{ik}, \text{UVent}_t^{ik}, \text{ReInj}_t^{ik}, \text{ReInS}_t^{ik}$ derived in the previous sections write:

$$\begin{aligned}
\text{IVent}_t^{ik} &= \delta_1^{ik} D_t^{ik} + \delta_2 P_t^{\text{Gas}} D_t^{ik} + \delta_3 \eta_t^{ik} D_t^{ik} \\
\text{UVent}_t^{ik} &= \delta_0^{ik} + \delta_4^{ik}\text{Oil}_t^{ik} + \delta_5 P_t^{\text{Gas}} + \delta_6 \eta_t^{ik} + \epsilon_t^{ik} \\
\text{ReInj}_t^{ik} &= \delta_7^{ik}\text{GasInj}_t^{i} + \delta_8^{ik} P_t^{\text{Gas}} \times \text{GasInj}_t^{ik} + \delta_9^{ik}\eta_t^{ik} \times \text{GasInj}_t^{ik} \\
\text{ReInS}_t^{ik} &= \delta_{10}^{ik}\text{GasInS}_t^{i} + \delta_{11}^{ik} P_t^{\text{Gas}} \times \text{GasInS}_t^{ik} + \delta_{12}^{ik}\eta_t^{ik} \times \text{GasInS}_t^{ik}
\end{aligned}$$

Substitute the formulas above into the OtherGas_t^{ik} equation:

$$\text{OtherGas}_t^{ik} = \text{IVent}_t^{ik} + \text{UVent}_t^{ik} + \text{ReInj}_t^{ik} + \text{ReInS}_t^{ik} \tag{6.216}$$

to obtain the empirical structural equation:

$$\begin{aligned}
\text{OtherGas}_t^{ik} = \; & \delta_0^{ik} + \delta_1^{ik} D_t^{ik} + \delta_2 P_t^{\text{Gas}} D_t^{ik} + \delta_3 \eta_t^{ik} D_t^{ik} \\
& + \delta_4^{ik}\text{Oil}_t^{ik} + \delta_5 P_t^{\text{Gas}} + \delta_6 \eta_t^{ik} + \delta_7^{ik}\text{GasInj}_t^{i} \\
& + \delta_8^{ik} P_t^{\text{Gas}} \times \text{GasInj}_t^{ik} + \delta_9^{ik}\eta_t^{ik} \times \text{GasInj}_t^{ik} + \delta_{10}^{ik}\text{GasInS}_t^{i} \\
& + \delta_{11}^{ik} P_t^{\text{Gas}} \times \text{GasInS}_t^{ik} + \delta_{12}^{ik}\eta_t^{ik} \times \text{GasInS}_t^{ik} + \epsilon_t^{ik}
\end{aligned} \tag{6.217}$$

which can be estimated using the estimates of $\hat{\eta}_t^{ik}$ obtained from the residuals of the flaring equation (first stage) as a new explanatory variable. Then, one can use the estimated parameters $\hat{\delta}_0^{ik}, \hat{\delta}_1^{ik}, \hat{\delta}_2, \hat{\delta}_3, \hat{\delta}_4^{ik}, \hat{\delta}_5, \hat{\delta}_6, \hat{\delta}_7^{ik}, \hat{\delta}_8^{ik}, \hat{\delta}_9^{ik}, \hat{\delta}_{10}^{ik}, \hat{\delta}_{11}^{ik}, \hat{\delta}_{12}^{ik}$ together with the regression residuals $\widehat{res2}_t^{ik}$ to construct estimates of the quantities of interest, specifically:

$$\widehat{\text{IVent}}_t^{ik} = \hat{\delta}_1^{ik} D_t^{ik} + \hat{\delta}_2 P_t^{\text{Gas}} D_t^{ik} + \hat{\delta}_3 \hat{\eta}_t^{ik} D_t^{ik} \tag{6.218}$$

and

$$\widehat{\text{UVent}}_t^{ik} = \hat{\delta}_0^{ik} + \hat{\delta}_4^{ik}\text{Oil}_t^{ik} + \hat{\delta}_5 P_t^{\text{Gas}} + \hat{\delta}_6 \hat{\eta}_t^{ik} + \widehat{res2}_t^{ik} \tag{6.219}$$

6.3.8 Identifications of Bounds on the Effect of a Flaring Tax

Exploiting the relationship between the structural parameters and the empirical equation in (6.217), which is illustrated below:

$$\begin{aligned}
\delta_1^{ik} &= \frac{\pi_1-\kappa_2}{\pi_1\kappa_1-{\kappa_2}^2}\left(\overline{MC}^{ik,\text{Gas}} + \tau^{ik} - \kappa_0^{ik}\right) + \zeta_0^{ik} + \zeta_1\left(\overline{MC}^{ik,\text{Gas}} + \overline{\tau}^{ik}\right) \\
\delta_2 &= -\frac{\pi_1-\kappa_2}{\pi_1\kappa_1-\kappa_2^2} \\
\delta_3 &= \frac{\pi_1-\kappa_2}{\kappa_1\pi_1-{\kappa_2}^2}\left(\zeta_1 + \frac{\kappa_1-\kappa_2}{\kappa_1\pi_1-{\kappa_2}^2}\right)^{-1} \\
\delta_4^{ik} &= \vartheta^{ik} GOR^{ik} \\
\delta_5 &= -\frac{(\kappa_4)^2}{\kappa_6} \\
\delta_6 &= \frac{(\kappa_4)^2}{\kappa_6}\left(\frac{\kappa_1-\kappa_2}{\kappa_1\pi_1-{\kappa_2}^2} + \zeta_1\right)^{-1} \\
\alpha_0^{ik} &= \frac{\kappa_1-\kappa_2}{\kappa_1\pi_1-{\kappa_2}^2}\left(\overline{MC}^{ik} + \tau^{ik} - \kappa_0^{ik}\right) + \zeta_0^{ik} + \zeta_1\overline{MC}^{ik,\text{Gas}} \\
\alpha_1 &= -\frac{\kappa_1-\kappa_2}{\kappa_1\pi_1-{\kappa_2}^2} \\
\eta_t^{ik} &= \left(\frac{\kappa_1-\kappa_2}{\kappa_1\pi_1-{\kappa_2}^2} + \zeta_1\right)\nu_t^{ik}
\end{aligned} , \tag{6.220}$$

we can use the estimated parameters to identify bounds on the structural objects of interest. Moreover, for the purpose of this section we impose the following additional assumptions: (i) $\delta_3 \leq 0$, which can be easily tested by checking the sign of the estimated parameter $\hat{\delta}_3$; and $\frac{\partial \text{Flare}_t^{ik*}}{\partial \tau_t^{ik,\text{Flare}}} \leq \frac{\partial \text{NRF}_t^{ik}}{\partial MC_t^{ik,\text{Gas}}}$; i.e., the effect of a marginal increase in the marginal cost of natural gas production is either positive or negative but not too large in magnitude relative to the (negative) effect of a marginal increase in the flaring tax on the amount of flaring performed by each field i. Note that the assumption stated in the previous sections also imply: $\alpha_1 \leq 0$ and $\delta_2 \leq 0$, which can also be tested by checking the sign of the corresponding estimated parameters. Then we can use the parameters α_1, δ_1, and δ_3 to identify bounds on the derivative of interest, as illustrated in the remainder of this section.

Lower bound on $\frac{\partial \textbf{IVent}_t^{ik}}{\partial \tau_t^{ik,\textbf{Flare}}}$

From the structural equation for IVent_t^{ik} we know that the derivative of interest has formula:

$$\frac{\partial \text{IVent}_t^{ik}}{\partial \tau_t^{ik,\text{Flare}}} = \frac{\kappa_2}{\pi_1\kappa_1 - \kappa_2^2} \mathbf{1}\left[\text{Flare}_t^{ik} > 0\right] \tag{6.221}$$

Consider the ratio of parameters δ_2 and δ_3. Using their structural equations in (6.220) we obtain:

$$\frac{\delta_2}{\delta_3} = -\left(\zeta_1 + \frac{\kappa_1 - \kappa_2}{\kappa_1\pi_1 - \kappa_2^2}\right) \tag{6.222}$$

which solves for

$$\frac{\kappa_2}{\pi_1\kappa_1 - \kappa_2^2} = \frac{\delta_2}{\delta_3} + \left(\zeta_1 + \frac{\kappa_1}{\kappa_1\pi_1 - \kappa_2^2}\right) \tag{6.223}$$

Under the assumption $\frac{\partial \text{Flare}_t^{ik*}}{\partial \tau_t^{ik,\text{Flare}}} \leq \frac{\partial \text{NRF}_t^{ik}}{\partial MC_t^{ik,\text{Gas}}}$ it must be true that $\zeta_1 + \frac{\kappa_1}{\kappa_1\pi_1 - \kappa_2^2} \geq 0$, which used in (6.223) implies:

$$\frac{\kappa_2}{\pi_1\kappa_1 - \kappa_2^2} \geq \frac{\delta_2}{\delta_3} \tag{6.224}$$

which can be combined with (6.221) to obtain a lower bound for $\frac{\partial \text{IVent}_t^{ik}}{\partial \tau_t^{ik,\text{Flare}}}$, which writes:

$$\frac{\partial \text{IVent}_t^{ik}}{\partial \tau_t^{ik,\text{Flare}}} \geq \frac{\delta_2}{\delta_3} \mathbf{1}\left[\text{Flare}_t^{ik} > 0\right] \tag{6.225}$$

and note that the inequality in (6.225) implies that if $\frac{\delta_2}{\delta_3} \geq 0$, then $\frac{\partial \text{IVent}_t^{ik}}{\partial \tau_t^{ik,\text{Flare}}} \geq 0$, i.e., a marginal increase in the flaring tax translate in a weakly larger amount of intentional venting by field i.

In a similar way, we can use the model to derive another lower bound on the value of the derivative of interest as follows. Note that the formulas for the structural parameters in (6.182) imply:

$$\frac{\delta_5}{\delta_6} = -\left(\zeta_1 + \frac{\kappa_1 - \kappa_2}{\kappa_1\pi_1 - \kappa_2^2}\right) \tag{6.226}$$

Following the same argument illustrated in the previous paragraph and using the previously stated assumptions on ζ_1, equation (6.226) implies:

$$\frac{\kappa_2}{\pi_1\kappa_1 - \kappa_2^2} \geq \frac{\delta_5}{\delta_6} \mathbf{1}\left[\text{Flare}_t^{ik} > 0\right] \tag{6.227}$$

Because both bound must be satisfied, we can combine them to obtain a single tighter lower bound, which has formula:

$$\frac{\partial \text{IVent}_t^{ik}}{\partial \tau_t^{ik,\text{Flare}}} \geq \max\left\{\frac{\delta_2}{\delta_3}, \frac{\delta_5}{\delta_6}\right\} \mathbf{1}\left[\text{Flare}_t^{ik} > 0\right] \tag{6.228}$$

Lower bound on $\frac{\partial \textbf{Flare}_t^{ik}}{\partial \tau_t^{ik,\textbf{Flare}}}$

From the structural equation for Flare_t^{ik} we know that the derivative of interest has formula:

$$\frac{\partial \text{Flare}_t^{ik}}{\partial \tau_t^{ik,\text{Flare}}} = -\frac{\kappa_1}{\pi_1\kappa_1 - \kappa_2^2}\mathbf{1}\left[\text{Flare}_t^{ik} > 0\right] \tag{6.229}$$

Consider the parameter α_1. Using its structural equation in (6.220) we obtain:

$$\frac{\kappa_1}{\kappa_1\pi_1 - {\kappa_2}^2} = \frac{\kappa_2}{\kappa_1\pi_1 - {\kappa_2}^2} - \alpha_1. \tag{6.230}$$

Given that $\frac{\kappa_2}{\kappa_1\pi_1 - {\kappa_2}^2} \geq 0$, the equation above implies:

$$\frac{\kappa_1}{\kappa_1\pi_1 - {\kappa_2}^2} \leq -\alpha_1 \tag{6.231}$$

which can be combined with (6.229) to obtain a lower bound for $\frac{\partial \text{Flare}_t^{ik}}{\partial \tau_t^{ik,\text{Flare}}}$, which writes:

$$\frac{\partial \text{Flare}_t^{ik}}{\partial \tau_t^{ik,\text{Flare}}} \geq \alpha_1\mathbf{1}\left[\text{Flare}_t^{ik} > 0\right] \tag{6.232}$$

and note that the inequality in (6.232) implies that if $\alpha_1 \leq 0$, then $\frac{\partial \text{Flare}_t^{ik}}{\partial \tau_t^{ik,\text{Flare}}} \leq 0$, i.e., a marginal increase in the flaring tax translate in a weakly lower amount of flaring by field i, as expected.

Lower bound on $\frac{\partial \textbf{CO2e}_t^{ik}}{\partial \tau_t^{ik,\textbf{Flare}}}$

The formula for a lower bound on overall effect of a marginal increase in the flaring tax on field i's CO2-equivalent GHG emissions is given by equation (6.59) of the theory section of this appendix, and writes:

$$\frac{\partial \text{CO2e}_t^{ik}}{\partial \tau_t^{ik,\text{Flare}}} \geq \frac{\partial \text{IVent}_t^{ik}}{\partial \tau_t^{ik,\text{Flare}}}CI^{\text{Vent}} + \frac{\partial \text{Flare}_t^{ik}}{\partial \tau_t^{ik,\text{Flare}}}CI^{\text{Flare}} \tag{6.233}$$

Using results (6.232) and (6.225) into (6.233) we obtain a lower bound for the overall effect of a marginal increase in the flaring tax CO2-equivalent GHG emissions, which writes:

$$\frac{\partial \text{CO2e}_t^{ik}}{\partial \tau_t^{ik,\text{Flare}}} \geq \left(\max\left\{\frac{\delta_2}{\delta_3}, \frac{\delta_5}{\delta_6}\right\}CI^{\text{Vent}} + \alpha_1 CI^{\text{Flare}}\right)\mathbf{1}\left[\text{Flare}_t^{ik} > 0\right] \tag{6.234}$$

which can be calculated using the estimates for α_1, δ_2, δ_3, δ_5, and δ_6 obtained using the method illustrated in section 3.7, and standard values for the carbon intensity of flaring and venting from the literature. In particular, the formula above implies that $\frac{\partial \text{CO2e}_t^{ik}}{\partial \tau_t^{ik,\text{Flare}}} \geq 0$ if the following inequality holds true:

$$-\frac{1}{\alpha_1}\max\left\{\frac{\delta_2}{\delta_3}, \frac{\delta_5}{\delta_6}\right\} \geq \frac{CI^{\text{Flare}}}{CI^{\text{Vent}}} \tag{6.235}$$

For instance, if we use the standard conversion values from Brandt et al. [2018], namely $CI^{\text{Flare}} = 0.3018$ $TCO2e/BOE$ and $CI^{\text{Vent}} = 3.9583$ $TCO2e/BOE$, such that $\frac{CI^{\text{Flare}}}{CI^{\text{Vent}}} \simeq 0.07624$, we get that the condition in (6.234) rewrites:

$$\frac{\partial \text{CO2e}_t^{ik}}{\partial \tau_t^{ik,\text{Flare}}} \geq \left(3.9583\max\left\{\frac{\delta_2}{\delta_3}, \frac{\delta_5}{\delta_6}\right\} + 0.3018\alpha_1\right)\mathbf{1}\left[\text{Flare}_t^{ik} > 0\right] \tag{6.236}$$

and therefore that $\frac{\partial \text{CO2e}_t^{ik}}{\partial \tau_t^{ik,\text{Flare}}} \geq 0$ if

$$-\frac{1}{\alpha_1}\max\left\{\frac{\delta_2}{\delta_3},\frac{\delta_5}{\delta_6}\right\} \geq 0.07624 \tag{6.237}$$

which can be tested using the estimates for α_1, δ_2, δ_3, δ_5, and δ_6 obtained using the method illustrated in section 3.7. Using our estimates for these parameters, which we present in section 4, we find that the empirical value of the left-hand side of inequality 6.237 is 2.05687, meaning that the inequality is satisfied. In turn, this finding represents compelling evidence that the introduction of a flaring tax would increase rather than decrease the overall greenhouse gas emissions produced by the oil&gas fields in our sample.

6.3.9 Identification: Discussion of Identifying Assumptions

The methodology for the identification of the key structural parameters of the model illustrated in this section is based on three key assumptions. In this section we describe these assumptions in depth and discuss the possible consequences of relaxing each of them.

- Key Assumption 1: $\pi_0^{ik} \simeq \kappa_0^{ik}$. This assumption states that the marginal expected costs of flaring and intentional venting are approximately the same at $\text{IVent}_t^{ik} = 0$ and $\text{Flare}_t^{ik} = \text{NRF}_t^{ik}$. This assumption is crucial for Proposition 5 to hold true, which states that $\text{IVent}_t^{ik} > 0$ if and only if $\text{Flare}_t^{ik} > \text{NRF}_t^{ik}$. In turn, it ensures that we can use the dummy variable D_t^{ik} as a tool to separate the effect of the natural gas price and the cost shock ν_t^{ik} on IVent_t^{ik} relative to the effect of the same variables on UVent_t^{ik}. We conjecture that the expected marginal cost of intentional venting at $\text{IVent}_t^{ik} = 0$ should be extremely close to zero, because the regulatory authority would not even start an investigation for a possible voluntary leak of an extremely small amount of natural gas, which means that the probability of a fine being issued remains arbitrarily close to zero for low levels of IVent_t^{ik}. However, one may conjecture that the marginal cost of flaring is non-zero even at $\text{Flare}_t^{ik} = \text{NRF}_t^{ik}$, because the cost of the routine maintenance of the flare stack may be roughly proportional to the amount of flaring performed. If the expected marginal cost of intentional venting evaluated at $\text{IVent}_t^{ik} = 0$ and $\text{Flare}_t^{ik} = \text{NRF}_t^{ik}$ is lower than the marginal cost of flaring at $\text{IVent}_t^{ik} = 0$ and $\text{Flare}_t^{ik} = 0$; i.e., $\pi_0^{ik} \geq \kappa_0^{ik}$, then we may have cases in which $\text{IVent}_t^{ik} > 0$ and $\text{Flare}_t^{ik} = \text{NRF}_t^{ik}$. In such a case, we conjecture that this amount of intentional venting would be incorrectly accounted as part of UVent_t^{ik}, causing the latter to be overestimated and IVent_t^{ik} to be underestimated. However, because the marginal cost of flaring is deemed to be extremely small, we conjecture that such estimation bias, if it occurs at all in our sample, is likely to be very small in magnitude. Lastly, because our proposed policy reform primarily targets intentional venting, this potential bias should result in an underestimation of the policy emission-reduction effects, which is consistent with our goal of obtaining prudent estimates.

- Key Assumptions 2 and 3: $A^{ik}\left(\text{ReInj}_t^{ik},\text{PInj}_t^{ik}\right)$ and $B^{ik}\left(\text{ReInS}_t^{ik},\text{PInS}_t^{ik}\right)$ are homogeneous functions. These two assumptions ensure that the formulas for ReInj_t^{ik} and ReInS_t^{ik} write $\text{ReInj}_t^{ik} = SR^{ik}\left(P_t^{\text{Gas}},\nu_t^{ik}\right)\text{GasInj}_t^{ik}$ and $\text{ReInS}_t^{ik} = SI^{ik}\left(P_t^{\text{Gas}},\nu_t^{ik}\right)\text{GasInS}_t^{ik}$, respectively. This is crucial for identification because it implies that ReInj_t^{ik} and ReInS_t^{ik} equal zero whenever $\text{GasInj}_t^{ik} = 0$ and $\text{GasInS}_t^{ik} = 0$, respectively. This ensures that the field fixed effect δ_0^{ik} in the structural equation can be entirely attributed to UVent_t^{ik}, which is therefore separately identified from ReInj_t^{ik} and ReInS_t^{ik}. Moreover, it also implies that ReInj_t^{ik} and ReInS_t^{ik} may be functions of P_t^{Gas} and ν_t^{ik}, but only with a functional form that is non-additive in GasInj_t^{ik} and GasInS_t^{ik}, respectively. As a consequence, this assumption allows one to identify the effect of P_t^{Gas} and ν_t^{ik} on ReInj_t^{ik} and ReInS_t^{ik} separately from the effect of the same variables on UVent_t^{ik}. We cannot exclude that the second implication may fail to hold true in our analysis. We conjecture that relaxing either of these assumptions (or both), may cause part of the effect of P_t^{Gas} and ν_t^{ik} to be attributed to GasInj_t^{ik} and GasInS_t^{ik} instead of UVent_t^{ik}. However, given that the effect of P_t^{Gas} on all those variables should be negative given the assumptions of our model and is empirically larger in magnitude relative to the effect of ν_t^{ik}, we also conjecture that a potential misspecification of the function $A^{ik}$and B^{ik} may lead to an underestimation of UVent_t^{ik} and, in turn, of the emission-reduction effects of our proposd tax reform. This is, once again, consistent with our goal of obtaining prudent estimates.

Thus, the key take-home from this analysis of the key identifying assumption of this empirical model is that the estimates of IVent_t^{ik} and UVent_t^{ik} obtained under these assumptions are likely to represent lower bounds for the quantities of interest in the case in which some of these assumptions do not hold true. Similarly, the estimated emission-reduction effects of the proposed tax reform should be interpreted as a lower bound.

6.4 Economic & Environmental Consequences

To compute the economic and environmental effects of the tax package, we use the Rystadt Shale Well Database Rystad Energy [2021]. The Rystadt Shale Well Database assigns to every well a unique ID. The latter identifies one, and only one, onshore oilfield. Aggregating them, we obtain information about 1,464 oil & gas fields over a sixteen years interval (2005-2020). The combination of the cross-sectional and of the temporal dimension creates a micro-panel (cross-sectional dimension $\gg$ time dimension; 1,464 $\gg$ 16), made out of 18,909 data points[11].

Out of 1,464 oil & gas fields, Rystadt classifies 1,091 as oilfields and 373 as gas fields. Among the oilfields, 325 (i.e. 29.79% of the sample) extract "conventional" oil by recovering high viscosity liquids from permeable rocks (Light & Medium). The remaining oilfields (i.e. 51.6% of the sample) produce "unconventional" oil either extracting low viscosity liquids from permeable rocks (Heavy & Extra Heavy, 32 oilfields) or extracting high viscosity liquids from impermeable rocks (Shale & Tight, 531 oilfields). Finally, 209 oilfields (i.e. 19.16% of the sample) are hard to classify since Rystad does not directly label these formations and we do not have information about the API gravity and/or the lithology of the rocks. Therefore, we generate a fourth category (Other), which incorporates oilfields with little or no information about the API gravity of the oil and/or the lithology of the rocks containing it. Out of 373 fields, which predominantly produce gas, 330 (i.e. 88.47% of the sample) extract raw methane from low permeability rocks (Shale & Tight gas fields) and 43 (i.e. 11.94% of the sample) from coal beds (Coalbed Methane).

Among oilfields, the output composition differs across categories. Shale & Tight formations are the main oil producers. They extract circa half of the oil (43.2%, yearly average 2.55 million barrels per day BBL/Day) and one fourth of the total natural gas (26.40%, yearly average 1.45 million barrels of oil equivalent BOE/Day). Light & Medium formations are responsible for a comparable quantity of production (38.10%, yearly average 1.29 million BBL/Day). Other formations are the third most important oil producers. They extract circa 10% of the total production (9.94%, yearly average 0.38 million BBL/Day) and almost 20% of the natural gas production (19.50%, yearly average 0.93 million BOE/Day). Finally, Heavy & Extra Heavy formations are the least important oil producers. They extract less than one tenth of the oil (8.78%, yearly average 0.31 million BBL/Day) and they virtually do not extract natural gas (0.48%, yearly average 0.02 million BOE/Day). Table 6.2 presents the summary statistics for the different types of oil producers.

Table 6.2: Oil & Gas Production

	Oil (BBL/Day)					Total Natural Gas (BOE/Day)				
	Mean	Median	SD	Min	Max	Mean	Median	SD	Min	Max
Light & Medium	4371	394	18077	0	371643	8342	509	70088	0	1301162
Heavy & Extra Heavy	9798	5117	14188	0	90354	669	176	1395	0	10286
Shale & Tight	6283	1349	13800	0	162382	3569	850	7622	0	93413
Other Oil	2389	46	7213	0	96023	5785	339	15797	0	123664

[11]The panel in unbalanced. A balanced panel would have had 1,464 · 16 = 23,424 data points. The unbalanced nature of the sample emerges because 692 fields are observed in every period, while the rest (784) either star or end their production during the studied period.

Most of the output is concentrated in few regions, which contain only one type of oil. The largest fraction of the oil is extracted from the Permian basin (located between Texas and New Mexico), and the Bakken basin (located in west North Dakota). Both basins contain Shale & Tight deposits. Figure 6.4 shows the 2019 production[12].

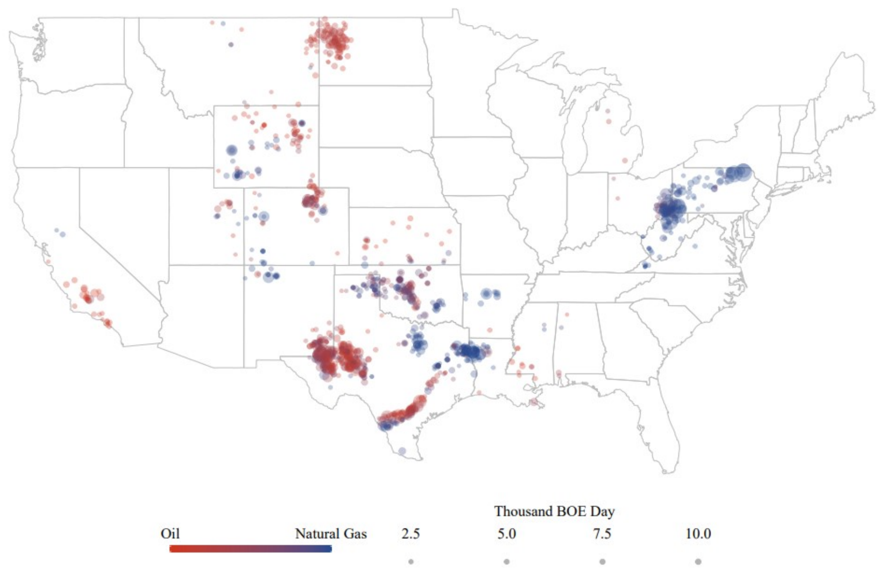

Figure 6.4: The 2019 Oil & Gas production in the United States. The colors of the dots reflect the composition of the outcome (Oil ●, Gas ●). The size of the dots reflects the aggregate volume of production.

6.4.1 First Step Estimation

The first step of the estimation process involves running a panel Tobit model,

$$\text{Flare}_t^{ik} = \begin{cases} \alpha_0^{ik} + \alpha_1 P_t^{\text{Gas}} + \eta_t^{ik} & \text{if } \eta_t^{ik} > -\alpha_0^{ik} - \alpha_1 P_t^{\text{Gas}} + \text{NRF}_t^{ik} \\ 0 & \text{otherwise} \end{cases}, \tag{6.238}$$

where the dependent variable, subject to left-censoring, represents flaring. The censoring occurs every time routine flaring is bigger than zero. Routine flaring denotes the proportion of total flaring that can be mitigated through managerial adjustments. In other words, it is the share of flaring not attributed to safety or maintenance concerns, which is instead categorized as non-routine flaring.

In order to divide routine from non-routine flaring, we study the flaring behaviour of producers who have the incentive to minimize their flaring behaviour, namely natural gas fields. These formations have as their primary source of income revenues obtained from selling natural gas. Therefore, they can be used to construct a benchmark in terms of minimum amount of flaring, which cannot be avoided. Table 6.3 presents the summary statistics for different types of natural gas producers.

Table 6.3: Flaring among Natural Gas Producers (BOE/Day)

	Mean	Median	SD	Min	Max
Coalbed Methane	11.28	1.61	27.77	0.00	202.66
Shale & Tight	56.10	5.20	169.88	0.00	1495.70

The median natural gas field has a flaring ratio, $\text{Flare}_t^{ik}/\text{TotGas}_t^{ik}$, of 0.05%, see Figure 6.5[13]. We use this quantity to identify the minimum amount of flaring necessary to guarantee the safety and the orderly maintenance

[12] Alaska's production is not displayed in the map but available upon request.

[13] Note that the average flaring ratio is bigger (0.44%) due to few outliers among Shale & Tight gas fields.

of production. In other words, we assume that it is not possible for oilfields to flare less than 0.05% of all the extracted natural gas without incurring into technical problems[14]. Therefore, we define non-routine flaring,

$$\text{NRF}_t^{ik} = \begin{cases} \text{mdn}\left(\frac{\text{Flare}_t^{ik}}{\text{TotGas}_t^{ik}}\right) \cdot \text{TotGas}_t^{ik} & \text{if } \text{mdn}\left(\frac{\text{Flare}_t^{ik}}{\text{TotGas}_t^{ik}}\right) \cdot \text{TotGas}_t^{ik} \leq \text{Flare}_t^{ik} \\ \text{Flare}_t^{ik} & \text{otherwise .} \end{cases}$$

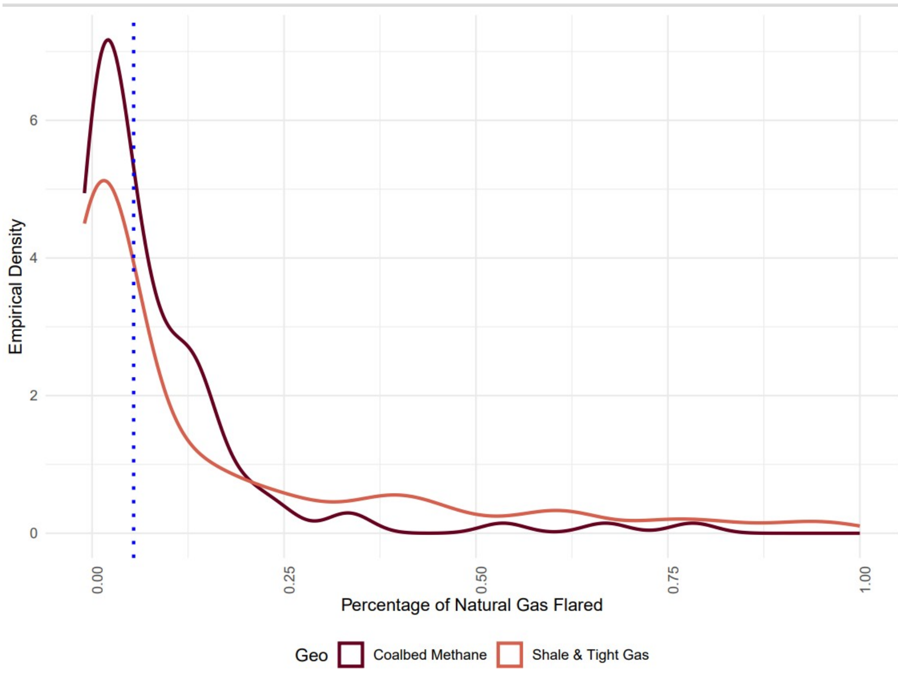

Figure 6.5: The empirical density function of the flaring rate for natural gas fields. The colors define the two types of formation (Coalbed Methane, Shale & Tight Gas). The blue dotted line indicates the median of their flaring rate (i.e. the fraction of natural gas that they flare divided by the total quantity of natural gas that they extract.). What is left-hand side of the vertical line is considered non-routine flaring.

Combining this definition with a selection procedure on the original dataset, we can construct the selection rule of equation (6.238). First, we drop all the natural gas fields reducing the dataset from 18,909 to 14,267 data points. Then, we check that all the remaining oilfields are consistent with the definition of an oilfield given by the Energy Information Administration (EIA). In other words, we check that each oilfield contained in the dataset has a gas-oil ratio smaller than 100,000 standardized cubic feet of natural gas per BBL. This second step reduces the dataset to 11,993 data points because there are 287 fields (mostly classified as Other Oil), which Rystad labels as oilfields, that do not respect the EIA definition. Then, we drop all the observations for which Flare_t^{ik} is an NA. This third step reduces the size of the dataset from 11,993 to 5,539 data points (i.e. 38.82% of the original sample). After this last step, we can construct the dependent variable of equation (6.238),

$$\text{RoutineFlare}_t^{ik} = \begin{cases} \text{Flare}_t^{ik} - \text{NRF}_t^{ik} & \text{if } \text{Flare}_t^{ik} - \text{NRF}_t^{ik} > 0 \\ 0 & \text{otherwise ,} \end{cases}$$

for all the oilfields for which we have information about flaring. Given the two previous definitions, we can compare non-routine with routine flaring. According to our calculation, on average a Light & Medium deposit non-routinely flares 2.36 BOE/Day and routinely flares 24.70 BOE/Day. In other words, routine flaring is more than ten times bigger than non-routine one. The same is true if the medians are compare. Both of which are

[14]While this assumption allows us to construct the reference case, its impact on the economic and environmental consequences of the policy are negligible. We study what happens with increasingly bigger thresholds and the results are virtually unchanged. All the results are available upon request.

significantly smaller due to the presence of few outliers, which flare up to 827.6 BOE/Day, see Figure 6.6. All these proportion are similar for Heavy & Extra Heavy and Other Oil, with Other Oil having a particularly fat right tail, see Table 6.4.

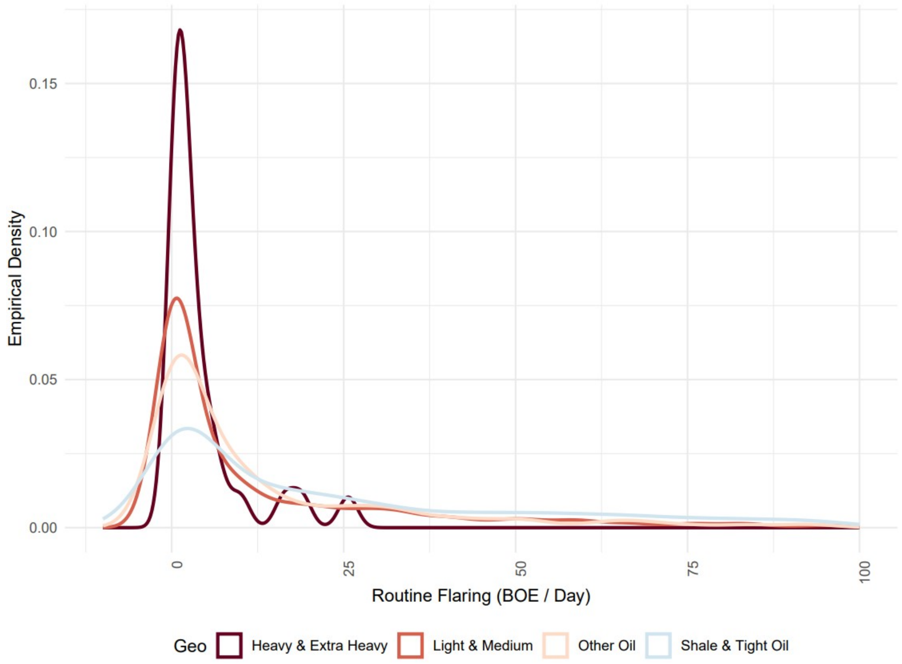

Figure 6.6: The empirical density function of the routine flaring for oilfields. The colors define the four types of formation (Heavy & Extra Heavy, Light & Medium, Other Oil, and Shale & Tight). The shape of the density is the one of a left censored variable.

Table 6.4: Summary Statistics Non-Routine vs Routine Flaring

	Non-Routine Flaring (BOE/Day)					Routine Flaring (BOE/Day)				
	Mean	Median	SD	Min	Max	Mean	Median	SD	Min	Max
Light & Medium	2.36	0.53	8.99	0.00	129.50	24.70	3.80	64.87	0.00	698.10
Heavy & Extra Heavy	0.20	0.09	0.24	0.00	0.66	4.30	1.75	6.24	0.00	25.65
Shale & Tight	2.38	0.81	4.22	0.00	40.95	237.00	44.00	558.67	0.00	9656.00
Other Oil	3.23	0.37	7.97	0.00	50.81	73.60	8.30	214.50	0.00	2433.90

Having defined routine and non-routine flaring, we can run model (6.238) using the Henry Hub spot price P_t^{Gas}, measured in (United States) Dollars/BOE. The model requires the estimation of a field-specific coefficient α_0^{ik}, which measures the unobserved intrinsic characteristics that lead a field to flare, and a population coefficient α_1, which measures the sensitivity of flaring to a change in the price of natural gas. To estimate these two parameters, along with the residuals, we cannot use the within transformation since the fixed-effects panel estimation is affected by the incidental parameters problem[15] Neyman and Scott [1948], Lancaster [2000]. Therefore, we assume that the individual effects are independent from the Henry Hub spot price and estimate the parameters consistently using a random effect model. In particular, we assume $\alpha_0^{ik} \overset{iid}{\sim} \mathcal{N}(0,1)$, and maximize the likelihood,

$$\mathcal{L}^{ik} = \int_{-\infty}^{\infty} \left\{ \prod_{t=1}^{T^{ik}} \left[\Phi\left(\frac{-\alpha_0^{ik} - \alpha_1 P_t^{\text{Gas}}}{\sigma_\eta} \right) \right]^{D_t^{ik}} \left[\frac{1}{\sigma_\eta} \phi\left(\frac{\text{Flare}_t^{ik} - \alpha_0^{ik} - \alpha_1 P_t^{\text{Gas}}}{\sigma_\eta} \right) \right]^{1-D_t^{ik}} \right\} \phi\left(\frac{\alpha_0^{ik}}{\sigma_{\alpha_0^{ik}}} \right) d\alpha_0^{ik} ,$$

to estimates the five parameters of interest $(\alpha_0^{ik}, \alpha_1, \sigma_{\alpha_0^{ik}}, \sigma_\eta)$, where D_t^{ik} is a dummy variable, which takes value equal to one if the field is doing routine flaring and zero otherwise, and σ_η is the standard deviation of the error term

[15]Note that, even if the magnitude of the coefficients could be estimated consistently with T small (in our case 16) using special maximization routines as the ones described in Greene [2001] and Webel [2011]. Their variance would still be inconsistent Henningsen [2010].

η. We set as initial values of the optimization $(\alpha_0^{ik} = 0, \alpha_1 = 0)$ and $\sigma_{\alpha_0^{ik}}$ as the standard deviation of the (column) mean of Flare_t^{ik} (i.e. the standard deviation of the mean of flare across oilfields), and $\sigma_\eta = \text{sd}(\text{Flare}_t^{ik})$. Using the `nlminb` package, we obtain an unconstrained optimization using a quasi-Newton method optimizer running the FORTRAN PORT library. The likelihood converges after 18 iterations (value of 35,141). The resulting $\hat{\eta}$ has the empirical density function shown in Figure 6.7.

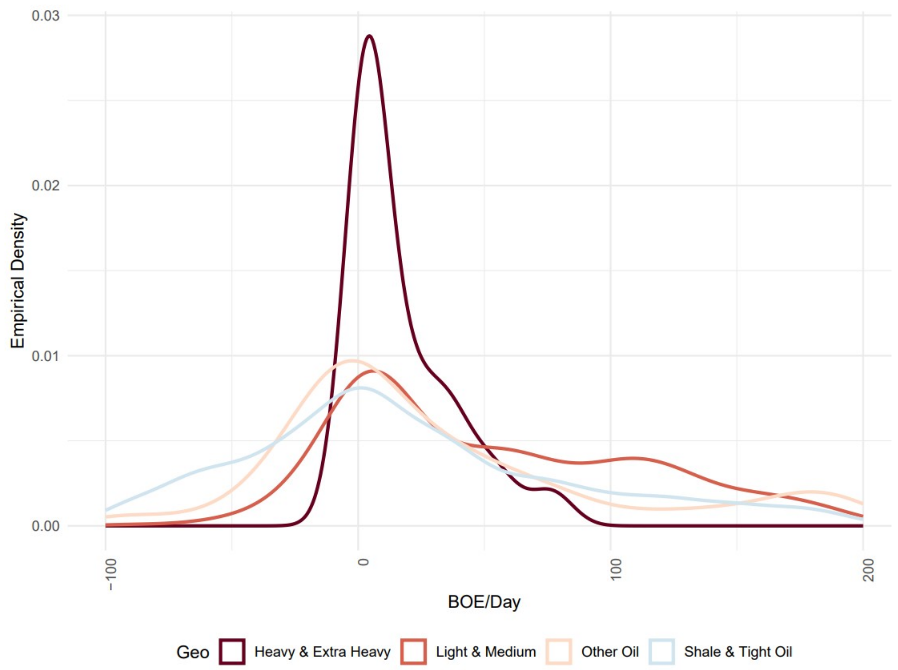

Figure 6.7: The empirical density function of the estimated unobserved part of natural gas marginal production costs for oilfields with positive routine flaring. For all four types of formation the expected value of $\hat{\eta}_t^{ik}$ is positive. The maxima are on average four times bigger than the minima highlighting a positive skewness of the distribution as suggested by the theoretical model.

The empirical density of η_t^{ik} is in line with what the theoretical model predicts. η_t^{ik} represents the non-visible part of the marginal costs of producing natural gas. If routine flaring is bigger than zero, it means that the oilfield faces 'high' marginal costs of gas production. Therefore, the distribution of $\hat{\eta}_t^{ik}$ must be centered around positive numbers. Table 6.5 breaks the results for the different types of formations.

Table 6.5: Summary Statistics $\hat{\eta}_t^{ik}$

	Routine Flaring > 0 (BOE/Day)					Routine Flaring = 0 (BOE/Day)				
	Mean	Median	SD	Min	Max	Mean	Median	SD	Min	Max
Light & Medium	63.70	45.30	82.87	-262.20	526.40	-254.00	-280.00	90.04	-568.00	-80.00
Heavy & Extra Heavy	17.10	6.65	20.84	-0.32	76.64	NA	NA	NA	NA	NA
Shale & Tight	57.00	7.00	431.66	-1785.00	7746.00	-326.00	-322.00	146.77	-1894.00	-107.00
Other Oil	58.10	14.50	107.06	-94.10	451.00	-264.00	-265.00	98.13	-460.00	-119.00

According to our estimates, when routine flaring is positive, $\hat{\eta}_t^{ik}$ is negative for the first three deciles of its distribution, but then shifts to positive with a median expected value of 15 BOE/Day and an average of 58 BOE/Day. When we break down these results across different types of oil formations, they remain consistent with the theoretical model. All four formation types show a positive mean, with Light & Medium, Shale & Tight, and Other Oil formations displaying similar values around 60 BOE/Day, while Heavy & Extra Heavy formations show a significantly lower mean of 17.10 BOE/Day. This is not surprising since the natural gas content of heavy deposits is minimal. Similarly, all medians are positive, and the spread between the minimum and maximum values suggests a right-skewed distribution. In contrast, the magnitude of $\hat{\eta}_t^{ik}$ when $\text{RoutineFlare}_t^{ik} = 0$ is negative by construction, amounting to $-391.50\frac{\phi(-\widehat{\text{Flare}t}^{ik}/391.50)}{\Phi(-\widehat{\text{Flare}t}^{ik}/391.50)}$. Lastly, the fixed coefficient $\hat{\alpha}_1$ is estimated at -1.74, indicating that for every 1 \$ increase in the price of natural gas (in Dollar/BOE), flaring at the oilfield declines by 1.74 BOE/Day.

6.4.2 Second Step Estimation

The second step of the estimation process involves running a panel linear model,

$$\begin{aligned}\text{OtherGas}_t^{ik} =& \delta_0^{ik} + \delta_1 D_t^{ik} + \delta_2 P_t^{\text{Gas}} D_t^{ik} + \delta_3 \hat{\eta}_t^{ik} D_t^{ik} + \delta_4 \text{Oil}_t^{ik} + \delta_5 P_t^{\text{Gas}} + \delta_6 \hat{\eta}_t^{ik} + \\ &+ \delta_7 \text{GasInj}_t^{ik} + \delta_8 P_t^{\text{Gas}} \text{GasInj}_t^{ik} + \delta_9 \hat{\eta}_t^{ik} \text{GasInj}_t^{ik} + \\ &+ \delta_{10} \text{GasInS}_t^{i} + \delta_{11} P_t^{\text{Gas}} \text{GasInS}_t^{ik} + \delta_{12} \hat{\eta}_t^{ik} \text{GasInS}_t^{ik} + \epsilon_t^{ik} \ , \end{aligned} \tag{6.239}$$

where the dependent variable represents the quantity of extracted gas that is not sold or flared, as defined in equation (6.157), measured in BOE/Day. In other words, the dependent variable equals the quantity of natural gas vented intentionally and unintentionally, plus the quantity (re-)injection, plus the quantity used onsite to generate heat or electricity, as described in equation (6.158). D_t^{ik} is a dummy variable, which takes value equal to one if the field is doing routine flaring and zero otherwise, P_t^{Gas} is the price of natural gas, as defined in section 6.4.1, while $\hat{\eta}_t^{ik}$ are the residuals obtained in the first step regression. Finally, all the other terms control for the volumes of natural gas injected or used in situ. δ_0^{ik} is an unobserved field specific effect, which might correlate with the other parameters as well as with the other explanatory variables. All the other are fixed coefficients. Finally, ϵ_t^{ik} is an error term normally distributed with mean zero and finite variance.

Table 6.6: Other Gas among Oil Producers (BOE/Day)

	Mean	Median	SD	Min	Max
Light & Medium	238.00	30.00	851.25	-60.00	8555.00
Heavy & Extra Heavy	1.29	0.40	1.53	0.00	4.30
Shale & Tight	204.00	38.00	473.17	-1.00	5925.00
Other Oil	171.20	33.40	276.55	0.00	1269.50

The dependent variable has an expected value of 210 BOE/Day and a median of 34 BOE/Day. The mean is significantly larger than the median due to a few outliers, particularly in the Light & Medium and Shale & Tight formations, which produce up to 8,555 and 5,925 BOE/Day, respectively, that are neither sold nor flared. Additionally, the standard deviation is notably higher for Light & Medium formations compared to other types of oil. According to a covariate-augmented Dickey-Fuller test with one and two lags (p-value = 0.01), the dependent variable is stationary. This allows us to run equation (6.239) in levels without encountering issues related to non-stationarity. We run three standard linear panel regressions: pooled ordinary least squares (OLS), a random effects model assuming that δ_0^{ik} is normally distributed with homoskedastic variance $\delta_0^{ik} \overset{iid}{\sim} \mathcal{N}(0, \sigma^2_{\delta_0})$ and independent of the regressors, and a fixed effects model. To check for autocorrelation among the estimated residuals $\hat{\epsilon}_t^{ik}$, we perform three Breusch-Godfrey tests for panel models, all of which indicate the presence of autocorrelation. Additionally, an F-test for cross-sectional and time effects suggeste the presence of fixed effects in the model. Finally, we run a Hausman test to differentiate between the random and fixed effects models, and the results rejected the random effects model. These findings collectively suggest that: 1) the error terms may be autocorrelated, 2) the variance of the residuals is heteroskedastic across the cross-sectional dimension, 3) the panel structure of the dataset matters, and 4) the unobserved individual effects are likely correlated with the explanatory variables.

Therefore, we run equation (6.239) using a feasible generalized least squared model, which included field-level fixed effects. This method uses a two-step estimation process. In the first step an ordinary least square estimation is done on equation (6.239) using the fixed effect option. Then, the resulting residuals $\hat{\epsilon}_t^{ik}$ are used to estimate an error covariance matrix to be used in a feasible generalized least square analysis. In this way, the error covariance structure inside each oilfield is fully unrestricted and is therefore robust against any type of intra-group heteroskedasticity and serial correlation[16]. Table 6.7 reports the results of the within estimation of the feasible generalized least squared model[17].

Variables	Estimate	Std. Error	z-Value	Pr($> \|z\|$)
Dummy	**103.63*** **	7.41	14.00	0.00
Spot Price · Dummy	**-2.37*** **	0.10	-22.60	0.00
First Stage Residual · Dummy	**0.19*** **	0.02	11.10	0.00
Oil	**0.01*** **	0.00	49.90	0.00
Lag Future Price	**-0.68*** **	0.05	-12.50	0.00
First Stage Residual	**-0.19*** **	0.02	-11.00	0.00
Residual standard error:	$\sqrt{MSE} = 0.32$ on 4770 df			
Adjusted R-Squared:	0.68			

Table 6.7: GLS Regression Results to quantify Venting in Flaring Oilfields

The first three coefficients, denoted as $(\delta_1, \delta_2, \delta_3)$, measure the extent of intentional venting in flaring oilfields. The coefficient $\hat{\delta}_1$ is positive, with an estimated value of 103.66 BOE/Day, indicating the maximum amount of gas a flaring oilfield is willing to intentionally vent when the expected natural gas price is zero. This value represents an upper limit for intentional venting, assuming the oilfield is flaring and faces marginal costs for selling natural gas that align with its unconditional expectations. From this upper bound, intentional venting decreases by 2.37 BOE/Day for each one Dollar/BOE increase in the natural gas price, as reflected by the coefficient $\hat{\delta}_2 = -2.37$ BOE^2/Dollar, conditional on flaring being greater than zero. Additionally, $\hat{\delta}_3$, estimated at 0.19, suggests that for every one BOE increase in flaring at constant prices, intentional venting rises by 0.19 BOE/Day, again conditional on flaring being greater than zero.

The second set of three coefficients, denoted as $(\delta_4, \delta_5, \delta_6)$, characterizes the magnitude of unintentional venting in both flaring and non-flaring oilfields. The coefficient $\hat{\delta}_4$ is positive, with an estimated value of 0.01, suggesting that, all else being equal, higher oil production leads to an increase in unintentional venting. The coefficient $\hat{\delta}_5$ is negative, with an estimated value of 0.68 BOE^2/Dollar, indicating that the level of maintenance of natural gas equipment increases as the price of natural gas rises. It is interesting to notice that the magnitude of this effect is smaller of the ones obtained by an increase in natural gas prices for flaring (-1.75 BOE/Day), and for intentional venting (-2.37 BOE/Day), suggesting that maintenance plays a smaller indirect role in shaping the responsiveness of GHG emissions to a change in natural gas prices. Finally, $\hat{\delta}_6$ is negative, signifying that lower expected future gas production costs incentivize increased maintenance today in preparation for more efficient future operations. The statistical significance of this last coefficient indirectly demonstrates that a joint taxation of flaring and venting would be ineffective, as oilfields under this framework would reduce maintenance activities.

Table 6.8: Summary Statistics of Estimated Venting

	$IVent_t^{ik}$ (BOE/Day)					$UVent_t^{ik}$ (BOE/Day)				
	Mean	Median	SD	Min	Max	Mean	Median	SD	Min	Max
Light & Medium	43.00	52.00	35.48	0.00	150.00	170.00	0.00	646.75	0.00	4500.00
Heavy & Extra Heavy	64.00	67.00	8.40	46.00	78.00	0.00	0.00	0.00	0.00	0.00
Shale & Tight	69.00	62.00	76.18	0.00	1500.00	36.00	0.00	274.55	0.00	1900.00
Other Oil	44.00	51.00	37.36	0.00	150.00	36.00	0.00	90.72	0.00	370.00

The resulting intentional venting and unintentional venting estimates offer a clear contrast in the scale and distribution of gas venting across different types of oil formations. For intentional venting, Shale & Tight for-

[16]Note that this method requires the estimation of $T(T+1)/2$ variance parameters. Therefore, for an individual fixed effect efficiency requires $N >> T$. This requirement is respected since the dataset has a cross-sectional dimension of 556 oilfields and a time dimension of 16 years.

[17]Only the coefficients relevant for intentional and unintentional venting are reported. The remaining ones, which are second-order Taylor approximations of the (re-)injection and in situ functions do not have a direct interpretation. Their sign and magnitude is only relevant to net out these two options and do not over-estimate the impact of the policy reform.

mations exhibit the highest mean venting rate at 69 BOE/Day, with significant variability, as indicated by a high standard deviation of 76.18 BOE/Day. This suggests that intentional venting practices in these formations are more inconsistent, with some fields experiencing substantially higher venting rates. In contrast, Heavy & Extra Heavy oil formations show more uniform behavior, with a mean of 64 BOE/Day and a relatively low standard deviation of 8.40 BOE/Day. Light & Medium and Other Oil formations have similar mean venting rates, around 43-44 BOE/Day, but Light & Medium fields demonstrate a slightly higher spread in values, suggesting more variation in venting practices. For unintentional venting, the results are more varied. Light & Medium formations exhibit the highest mean UVent rate at 170 BOE/Day, driven by a few extreme outliers, as reflected in the large standard deviation of 646.75 BOE/Day. Shale & Tight formations also show notable unintentional venting, with a mean of 36 BOE/Day and a significant standard deviation, again suggesting variability across fields. Notably, Heavy & Extra Heavy formations show no unintentional venting, possibly due to the inherent characteristics of the formation (i.e. extremely low gas-oil ratio). Similarly, Other Oil formations have low UVent values, with a mean of 36 BOE/Day and less variation compared to other types. Overall, our estimates reveal that intentional venting tends to be higher and more consistent in Shale & Tight formations, while unintentional venting is more pronounced in Light & Medium formations, likely driven by equipment degradation and operational practices.

Economic Outcome By averaging across the time dimension to create a cross-sectional dataset, we can evaluate the average economic performance of the 556 observed oilfields over the study period. This approach allows us to compare the energy wasted vs the energy recovered by the policy. Summing the energy lost through non-routine flaring, routine flaring, intentional venting, and unintentional venting, and then dividing this total by the sum of oil and natural gas extracted, we find that, on average, 2.78% of all energy is wasted (4.44% of oil energy and 7.09% of natural gas energy). This equates to an aggregate average waste of 0.19 million BOE/Day, with an average waste of 350 BOE/Day and a median waste of 140 BOE/Day. As in previous analyses, the mean is significantly skewed by a small number of outliers. If 100% of routine flaring and intentional venting would have been saved by implementing the reform in it most ambitious version, on average 1.09% of all the extracted energy would have been wasted (1.74% of oil energy and 2.92% of natural gas energy). This equates to an aggregate average waste of 0.08 million BOE/Day, with an average waste of 139 BOE/Day and a median one of 18 BOE/Day. In other words, for an average natural gas price of 21.07 Dollars/BOE the average waste would decline from 7,374 Dollars/Day to 2,929 Dollars/Day (the median waste would shift from 2,950 Dollars/Day to 379.3 Dollars/Day). The savings are not equally divided from routine flaring and intentional venting. The former account for a total of 0.0860 million BOE/Day and 0.0334 million BOE/Day of intentional venting. The part of unrecovered waste is in minimal part due to non-routine flaring, which amounts to a total of 0.0012 million BOE/Day and 0.0759 million of BOE/Day of unintentional venting.

Table 6.9: Delta in Oil and Natural Gas Tax for different Methane Savings Scenarios

	Oil Tax ($/BBL)					Natural Gas Tax ($/BOE)
	Mean	Median	SD	Min	Max	
25%	6.75	2.50	12.56	0.00	84.29	-5.23
50%	13.45	4.98	25.03	0.00	168.01	-10.43
75%	20.92	7.75	38.95	0.00	261.40	-16.23
100%	101.30	37.50	188.70	0.00	1266.10	-78.63

The summary statistics presented in Table 6.9 illustrate the required changes in oil and natural gas taxes under different methane-saving scenarios. Table 6.10 shows that, as the percentage of methane emissions saved increases, so does the magnitude of energy savings. For flaring, the mean savings range from 36.70 BOE/Day at a 25% savings scenario to 80.30 BOE/Day at 100%, with median values displaying a similar upward trend. The substantial standard deviations, especially in the 100% savings scenario (91.53 BOE/Day), indicate significant variability across oilfields. This is likely due to differences in the existing levels of waste. The same pattern is observed for venting savings, where mean values rise from 15.10 BOE/Day at 25% savings to 60.00 BOE/Day at 100%. However, the greater divergence between the mean and median values in the 100% savings scenario underscores the uneven distribution of venting reduction across oilfields. The high maximum venting savings of 239.90 BOE/Day in the 100% scenario further emphasize the potential for significant reductions in emissions if the most ambitious methane-saving policies are implemented.

Table 6.10: Summary Statistics of Energy Savings for Different Methane Savings Scenarios

	Flaring Savings (BOE/Day)					Venting Savings (BOE/Day)					Total Savings (BOE/Day)
	Mean	Median	SD	Min	Max	Mean	Median	SD	Min	Max	Total Savings
25%	36.70	23.70	48.26	0.00	366.90	15.10	16.00	3.55	0.00	16.00	28806
50%	41.40	28.70	50.67	0.00	375.90	29.90	31.80	7.33	0.00	31.80	39637
75%	46.30	34.80	53.66	0.00	386.00	45.10	49.50	11.89	0.00	49.50	50812
100%	80.30	41.00	91.53	0.00	494.50	60.00	62.70	25.63	0.00	239.90	78013

From an economic standpoint, the potential financial impact of these energy savings is substantial. With every saved BOE valued at 21.07 Dollars, achieving a 100% methane savings rate could result in daily savings of approximately 1,691 Dollars from reduced flaring and 1,264 Dollars from reduced venting, on average. These savings represent a meaningful offset to the costs associated with higher taxes on oil production under methane-reduction policies. At the 100% savings level, oil taxes rise sharply, with an average increase to 101.30 Dollars/BBL. However, the natural gas tax simultaneously decreases to -78.63 Dollars/BBL, signaling a balanced approach to promoting environmental sustainability while maintaining economic viability for oilfields. The significant reduction in natural gas taxes mitigates the cost burden of achieving higher methane savings, encouraging firms to adopt technologies and practices that reduce flaring and venting. At the same time, the rising oil tax ensures that the policy retains a level of economic rigor, making it costly for firms to ignore potential savings opportunities. Furthermore, the variability in the savings across oilfields suggests that some fields, particularly those with higher venting and flaring rates, stand to benefit more substantially from the reforms, potentially driving industry-wide adoption of more sustainable practices.

Environmental Outcome This energy waste translates to aggregate emissions of 0.46 million tons of CO_2e/Day, of which 0.16 million tons are currently in play due to routine flaring and intentional venting. Eliminating routine flaring entirely would reduce emissions by 0.03 million tons of CO_2e/Day, while completely stopping intentional venting would save an additional 0.16 million tons of CO_2e/Day. Consequently, the emissions from intentional venting are approximately five times those from routine flaring. Of the emissions not in play, 99.87% are attributable to unintentional venting, amounting to 0.30 million tons of CO_2e/Day. The average recoverable emissions account for 35.31% of the total, with routine flaring representing 98.63% of all observable flaring-related CO_2e emissions and intentional venting accounting for 30.05% of all methane emissions.

Table 6.11: Summary Statistics of Emissions Savings for Different Methane Savings Scenarios

	Flaring Savings (TCO_2e/Day)					Venting Savings (TCO_2e/Day)					Total Savings (TCO_2e/Day)
	Mean	Median	SD	Min	Max	Mean	Median	SD	Min	Max	Total Savings
25%	13.76	8.87	18.09	0.00	137.55	59.80	63.20	14.07	0.00	63.20	40883
50%	15.50	10.77	19.00	0.00	140.94	118.00	126.00	29.01	0.00	126.00	74401
75%	17.35	13.05	20.12	0.00	144.72	179.00	196.00	47.07	0.00	196.00	108956
100%	30.10	15.39	34.42	0.00	185.39	238.00	248.00	101.40	0.00	950.00	148857

The Table 6.11 presents the emissions savings in terms of tonnes for the same methane savings, offering an environmental perspective on the energy savings highlighted in the previous table. For flaring the mean savings range from 13.76 TCO_2e/Day at 25% savings to 30.10 TCO_2e/Day at 100%. The median values follow a similar pattern, increasing from 8.87 TCO_2e/Day to 15.39 TCO_2e/Day across the scenarios. While the standard deviations are large, particularly in the 100% scenario, they highlight the variation in potential emissions savings across oilfields. This variability suggests that some oilfields are contributing disproportionately to flaring emissions and would benefit more from aggressive methane reduction policies. The maximum flaring savings at 100%, reaching 185.39 TCO_2e/Day, indicate that substantial reductions are possible under ideal conditions.

The mean venting savings range from 59.80 TCO_2e/Day at 25% savings to 238.00 TCO_2e/Day at 100%. This significant increase across scenarios, especially at the upper end, highlights the environmental importance of reducing methane emissions. The maximum venting savings at 100% (950 TCO_2e/Day) are notably high. When these emissions savings are compared to the energy savings discussed earlier, it is clear that focusing on methane reduction (through venting savings) can lead to far greater environmental gains relative to flaring reductions. While the previous table indicated that venting savings could yield significant energy conservation, the current table shows that these reductions have an even more pronounced impact in terms of mitigating climate change.

6.5 Discussion & Further Policy Proposals

6.5.1 Efficiency of the Proposed Solution

It is easy to verify that the allocation generated by our proposed tax scheme is generally not Pareto-efficient. The reader may wonder why this scheme is preferable to Piguouvian taxation or other traditional approaches based on the Polluter-Pays principle, such as emission markets, which are well-known for inducing the First-best allocation in some circumstances.

The answer to this question lies in two key assumptions of our model, which closely mimic two core features of oil & gas markets and cause the First-best allocation to be unattainable in this economy. The first assumption is that intentional venting of natural gas is not perfectly observable and/or not contractible by the regulator. This assumption is not only justified by the fact that methane emissions are not easy to quantify and monitor. Perhaps more importantly, it is extremely difficult for the regulator to prove whether a certain amount of methane emission is “deliberate" or not in a legally binding way, because oil firms often claim that venting is justified by safety concerns (e.g., fire or explosion risk) or independent of their control to avoid punishment. Given this issue, one may wonder why the regulator does not tax all methane emissions, regardless of their (intentional or unintentional) nature. To see why, note that the existence of safety concerns implies that a regulator committing to punish venting even if they cannot prove it to be “avoidable" may induce firms to adopt a risky behavior with respect to fire and explosion hazard. Moreover, a tax on unintentional methane emissions my encourage emission misreporting, as we argue later in this section. The second assumption is that the level of maintenance of an oilfield is also not observable and/or contractible. This assumption follows the fact that the regulator may perhaps observe the firm's monetary investment in maintenance, but cannot easily assess whether such investment targets leakages reduction and/or detection in an effective way. As a result, if leakage-reducing maintenance becomes unprofitable, firms can either waste their maintenance investment in ineffective activities of divert some of it towards targets other than leakage reduction, in a way that is hard for the regulator to detect. These two assumptions together have dramatic consequences for the effectiveness of traditional pricing schemes.

First, any scheme that increases the cost of flaring relative to that of venting causes unwanted substitution between these two practices. This means that a flaring tax typically results in an increase in intentional venting, as illustrated in Proposition 1. For instance, our empirical results suggests that during the period 2005-2020 the introduction of a flaring tax of 1$ per BOE of flared natural gas would have caused an average increase in intentional venting by the firms included in our sample equal to at least $\hat{\delta}_5/\hat{\delta}_6\mathbf{1}[\text{Flare}_t^{ik} > 0] = 3.58$ BOE/Day.

Second, unless the production for commercial purposes of the co-extracted gas (or the alternative uses such as re-injection and in-situ use) is so profitable for a given firm that routine flaring and intentional venting are not a concern, any scheme that increases the cost of natural gas disposal (either flaring or intentional venting, or both) reduces the incentives for the firm to carry out effective maintenance aiming at reducing and detecting leakages. This implies that the introduction of a flaring tax or an increase in either the fines for intentional venting or the effectiveness of its detection by the regulator (or both) typically result in lower maintenance and increased natural gas leaking (a.k.a. “unintentional venting"), as illustrated in Proposition 2. For instance, our empirical results suggests that during the period 2005-2020 the introduction of a flaring tax of 1$ per BOE of flared natural gas would have caused an average increase in unintentional venting by the firms included in our sample equal to at least $\hat{\delta}_5 = 0.68$ BOE/Day.

These two pieces of evidence illustrate how in our setup - characterized by asymmetric information and limited enforcement - the Polluter-Pays principle does not work. Thus, a price scheme can achieve the elimination of both routine flaring and intentional venting without causing either substitution between the two practices or an increase in non-voluntary natural gas leakages only if it does not increase the overall marginal cost of gas disposal. Our proposed scheme possesses this feature. As a consequence, it eliminates both routine flaring and intentional venting while simultaneously reducing leaking (unintentional venting).

6.5.2 Political Economy & Implementation

One of the major shortcomings of most traditional pricing policies based on the Polluter-Pays principle, such as carbon taxes and emission markets, is that they typically result in lower output and higher equilibrium consumer prices for the goods affected directly or indirectly by the pricing scheme. This has important consequences that often undermine both their effectiveness and the political support they enjoy. First, lower output and higher consumer prices typically result in lower corporate profits and reduced consumption of certain goods. These undesirable outcomes incentivize firms to lobby against the implementation of such policies and consumers to support political parties that oppose them. Moreover, if these policies are introduced in a single country (or in a limited group of countries), they tend to reduce the competitiveness of domestically produced products on global markets, as they become more expensive relative to similar goods produced abroad. In turn, this may result in lower output and unemployment. Moreover, it may cause *carbon leakage*: the production of emission-intense goods may move from countries that apply a pricing scheme to those that do not, causing free-riding and resulting in limited or no effect of those policies on global GHG emissions. Our proposed reform is immune to these side effects because it has zero impact on all equilibrium prices and (approximately) no effect on firms' profitability, consumers' purchasing power, and government revenue. Figure 6.8 show the Strengths, Weaknesses, Opportunities, and Threats of the proposed reform versus two standard alternatives: the introduction of a flaring tax and the combination of a flaring tax with venting regulation and/or taxation.

The intuition underpinning these desirable outcomes is simple. Our reform requires oil producing firms to pay an extra tax which is proportional to the total natural gas extracted from the oilfield net of the quantity of natural gas that is re-injected or used in-situ for electricity and heating generation or that cannot be recovered using currently available technology and the industry's best practices (i.e., non-routine flaring and the amount of unintentional venting that cannot be avoided even with adequate maintenance). The resulting extra oil tax amount is proportional to the maximum avoidable methane footprint that the firm's oil extraction activity generates; i.e., the pricing scheme makes the firm *internalize* the potential social cost of their methane emissions, in the same way as a traditional Pigouvian tax would. However, the price scheme also provides a tax rebate which is proportional to the amount of natural gas that is sold by the firm. This second component of the reform serves two purposes. Firstly, it makes natural gas production relatively more profitable than both flaring and venting, inducing oil & gas firms to capture and sell on the market all the co-extracted natural gas that they would have flared and/or vented otherwise. Secondly, it exactly offsets the additional marginal cost of oil production caused by the increase in the oil tax. As a result, the overall marginal cost of oil production - including the cost of managing the co-extracted natural gas - in unchanged after the reform is implemented. This implies in turn that the oil production choices of all firms are unaffected at given market prices. Moreover, as long as the tax rebate on natural gas production is sufficiently large to eliminate both routine flaring and intentional venting, the rebate also compensates each firm such that its profits are approximately unaffected by the reform.

The reform also prescribes a small increase in the marginal tax rate on natural gas produced by gas-only firms

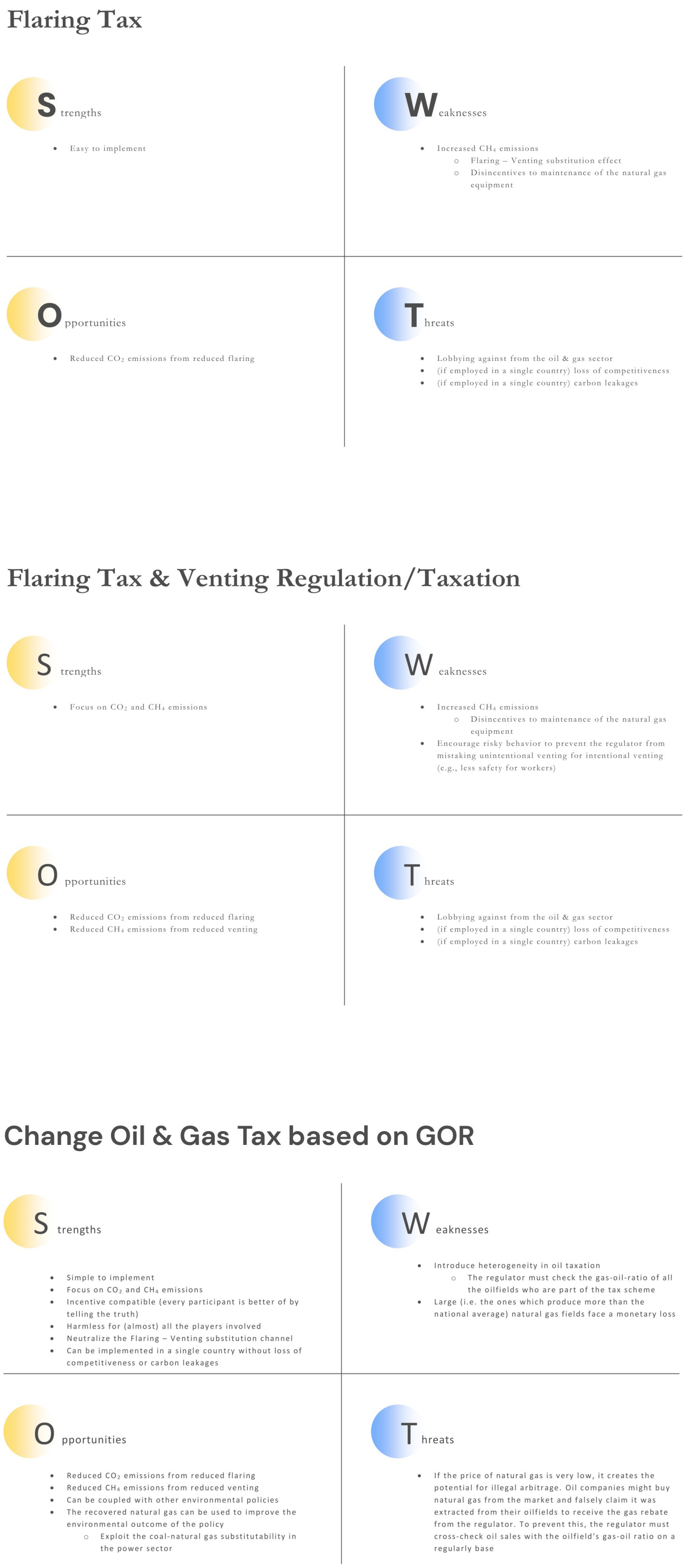

Figure 6.8: Differences in Strengths, Weaknesses, Opportunities, and Threats of the proposed policy versus two standard alternatives.

and a small increase in deductions for gas-only fields. The former ensures that the additional gas supply generated by the elimination of routine flaring and intentional venting is exactly compensated by a fall in the supply of natural gas from gas-only fields. The latter compensates gas-only firms for the small profit loss they face because of the extra tax and should help in preventing lobbying by this type of firms against the implementation of the reform. As a result of these corrective taxes, the aggregate supply of both oil and natural gas and, in turn, their equilibrium market prices are unaffected by the reform. Moreover, all oil & gas firms' profits are approximately unchanged, meaning that the oil & gas industry has little or no incentive to lobby against the implementation of the reform. For the same reason, our proposed reform is immune to loss of competitiveness and carbon leakage. Its introduction in a single country does not changes the firms' incentive to produce domestically and/or relocate production abroad. Similarly, the fact that the reform has no effect on all equilibrium prices of consumption goods and a weakly positive effect on government revenue implies that consumers and taxpayers have no incentive to oppose it though voting and/or collective action.

From a mechanism design perspective, note that - as long as the gas-oil ratio, the sales of oil and natural gas, and the quantity of natural gas that are either flared, re-injected or used in-situ for electricity production are fully observable by the regulator - each firm's total methane emission can be easily and accurately calculated using a simple formula, meaning that any attempt of cheating would be immediately detected. One may object that the quantity of natural gas re-injected and used in-situ may not be easily observable by the regulator, who may in fact have to rely on self-reported measures. While this is a valid concern, there are strong arguments suggesting that it is not a major one. First of all, it is relatively simple and cheap for the regulator to monitor ex-post the quantity of natural gas re-injected or used in-situ by a firm to detect substantial misreporting. For instance, natural gas injections typically affect the field's gas-oil ratio, whereas in-situ use to produce electricity is driven by the firm's electricity needs net of its purchases from the power grid. Any inconsistency between these measures and the reported quantities of co-extracted natural gas re-injected and used in-situ by the firm would constitute a strong signal of a likely attempt of cheating. There is an even more compelling theoretical argument that should reassure the reader regarding this potential issue. If the tax rates are set equal to their recommended level stated in section 1.4.1, then by Proposition 3 the quantity of co-extracted natural gas, which is intentionally vented by each oil & gas firm tends to zero. Recall that under our proposed tax regime any firm, which does not perform illegal venting, face no fines or fees for its unintentional methane leakages. Thus, because no methane is released intentionally, the firm's management has no strict incentive to manipulate the self-reported values of natural gas emissions, because at the firm's optimal choice the expected cost due to venting regulation is equal to zero and cannot be reduced any further. In fact, the introduction of a small fine for detected misreporting is sufficient to make it *strictly* unprofitable in expectation. In other words, truth-telling is *incentive compatible*. Even in the prudent empirical scenarios illustrated in section 5.2, the majority of the oil & gas firms is shown to react to the reform by either eliminating or drastically reducing intentional venting, meaning that the extent of misreporting is likely to be either null or extremely limited. Together with the safety concerns mentioned in the main body of the paper, this result represents a further theoretical reason to recommend no taxation on unintentional venting (leaking).

From a theoretical perspective, the main potential weakness of our proposed reform is that, if the tax rebate on natural gas sales from oilfields required to eliminate routine flaring and (intentional) venting is larger than the market price of natural gas, then the scheme may promote illegal arbitrage on the natural gas market. That is, oil & gas firms may have an incentive to purchase natural gas from non-monitored sources (e.g., black market) and pretend it has been extracted from an oilfield to obtain the rebate and earn a positive profit. Because of that, our results explore four different venting reduction targets (25%, 50%, 75% and 100%), each corresponding to different values for the tax rebate. Our empirical estimates show that during the 2005-2020 time period the complete elimination of intentional venting would have required a tax rebate rate which is larger than the average market price of natural gas over the period of interest and therefore potentially prone to promote illegal arbitrage. This finding is a direct consequence of the presence in our sample of a very small number of oilfields that feature an extremely large gas-oil ratio. However, all the other venting reduction targets (25%, 50%, 75%) could have been achieved with a tax rebate rate which is significantly lower than the average market price of natural gas during the time span of interest. Thus, for the sake of prudence in the main body of the article we present the results for the 75% target, which delivers the largest emission reductions without generating incentives to perform illegal arbitrage.

6.5.3 Link to the Current Tax Structure

For the sake of simplicity and ease of interpretation the baseline model presented in section 1 of this appendix assumes that oil and gas taxation is levied through *specific taxes* on oil and natural gas sales. However, the proposed setup does not represent an accurate description of the US tax system, which is based mostly on *ad valorem* taxes, with a few exceptions. In this section, we show how the results presented in section 6.2 and 6.3 hold true even if a more realistic tax system is assumed. In detail, we borrow the setup in Kunce et al. [2003], which provides a stylized but sufficiently realistic model of the US tax system with respect to oil and gas firms. First, we assume a tax system featuring two linear corporate tax rates on firm's profits: one at Federal level and one at State level, denoted by $T_{t,US}^{ks}$ and $T_{t,S}^{ks}$, respectively. After noticing that, relative to the baseline model, $\left(1-T_{t,US}^{ks}\right)\left(1-T_{t,S}^{ks}\right)$ simply replaces $1-T_t^{ks}$ in each firm k's objective function and that $T_{t,S}^{ks}T_{t,US}^{ks}$ replaces T_t^{ks} in the formula for total tax revenues, it is easy to show that this first change in the tax system is fully innocuous for our predictions. Secondly, we introduce taxes on oil and gas production that mimic those that are imposed by most US States and the Federal Government. Specifically, let $\tau_R^{ks,\text{Oil}}$ and $\tau_R^{ks,\text{Gas}}$ the royalty rates on production of oil and gas from public (state and federal) land. Moreover, we denote with $\tau_P^{ks,\text{Oil}}$ and $\tau_P^{ks,\text{Gas}}$ the production (severance) tax rate on production of oil and gas, respectively. Lastly, δ_{US} denotes the federal percentage depletion allowance weighted by the percentage of production attributable to eligible producers (non-integrated independents). Given these assumptions, the formula for firm k's revenue from oil and gas production in field i (before corporate taxes) writes:

$$
\begin{aligned}
&P_t^{iks,\text{Oil}}\left[(1-\tau_R^{ks,\text{Oil}})(1-\tau_P^{ks,\text{Oil}})+\frac{T_{t,US}^{ks}(1-\tau_R^{ks,\text{Oil}})}{(1-T_{t,US}^{ks})(1-T_{t,S}^{ks})}\delta_{US}\right]\text{Oil}_t^{iks}+\\
&+P_t^{s,\text{Gas}}\left[(1-\tau_R^{ks,\text{Gas}})(1-\tau_P^{ks,\text{Gas}})+\frac{T_{t,US}^{ks}(1-\tau_R^{ks,\text{Gas}})}{(1-T_{t,US}^{ks})(1-T_{t,S}^{ks})}\delta_{US}\right]\text{Gas}_t^{iks}+\\
&-\tau_t^{iks,\text{Flare}}\text{Flare}_t^{iks}-PP_t^{iks,\text{Gas}}\text{PInS}_t^{iks}-\text{IM}_t^{iks}+\text{Z}_t^{iks}\ .
\end{aligned}
$$

Replacing $1-T_t^{ks}$ with $(1-T_{t,US}^{ks})(1-T_{t,S}^{ks})$ in (6.18), we find that the F.O.C.s w.r.t. Oil_t^{iks} and Gas_t^{iks} in (6.19) become:

$$
\begin{aligned}
\frac{\partial\mathcal{L}_u^{ks}}{\partial\text{Oil}_t^{iks}} = \;& P_t^{\text{Oil}}\left(1-\varsigma_t^{\text{Oil}}MS_t^{ks,\text{Oil}}+\frac{\sigma^{iks}}{P_t^{\text{Oil}}}\right)\\
&\left[\left(1-\tau_R^{ks,\text{Oil}}\right)\left(1-\tau_P^{ks,\text{Oil}}\right)\left(1-T_{t,US}^{ks}\right)\left(1-T_{t,S}^{ks}\right)+T_{t,US}^{ks}\left(1-\tau_R^{ks,\text{Oil}}\right)\delta_{US}\right]\\
&-\phi_{1t}^{iks}\frac{\partial TF_{1t}^{iks}(.)}{\partial\text{Oil}_t^{iks}}-\phi_{2t}^{iks}GOR^{iks}-\phi_{3t}^{iks}-\phi_{4t}^{iks}\zeta+\phi_{5t}^{iks}=0\\
\frac{\partial\mathcal{L}_u^{ks}}{\partial\text{Gas}_t^{iks}} = \;& P_t^{s,\text{Gas}}\left(1-\varsigma_t^{s,\text{Gas}}MS_t^{ks,\text{Gas}}\mathbf{1}\left[\text{Oil}_t^{isk}=0\right]\right)\\
&\left[\left(1-\tau_R^{ks,\text{Gas}}\right)\left(1-\tau_{t,P}^{ks,\text{Gas}}\right)\left(1-T_{t,US}^{ks}\right)\left(1-T_{t,S}^{ks}\right)+T_{t,US}^{ks}\left(1-\tau_R^{ks,\text{Gas}}\right)\delta_{US}\right]\\
&-\phi_{1t}^{iks}\frac{\partial TF_{1t}^{iks}(.)}{\partial\text{Gas}_t^{iks}}+\left(\phi_{2t}^{iks}-\phi_{3t}^{iks}-\phi_{4t}^{iks}\zeta\right)\left(1-\vartheta^{iks}\right)^{-1}=0
\end{aligned}
\tag{6.240}
$$

Moreover, we find that all the other F.O.C.s of firm k in (6.19) are unchanged, except for featuring the term $\left(1-T_{t,US}^{ks}\right)\left(1-T_{t,S}^{ks}\right)$ instead of $\left(1-T_t^{ks}\right)$. Following the same steps as those presented in section 6.2, we find that the incentives generated by $\varDelta\tau_t^{jks,\text{Oil}}$, $\varDelta\tau_t^{s,\text{Gas}}$ for oil fields in the baseline model are replicated in the alternative model if the following equations hold true:

$$\begin{aligned} \Delta\tau_{t,P}^{s,\text{Gas}} &= \min_{i\in\{\{1,2,\dots,I_k\}\}_{k=1}^{K}} \left\{1-\tau_{t,P}^{ks,\text{Gas}}+T_{t,US}^{ks}\delta_{US}-\frac{MC_t^{iks,\text{Gas}}}{P_t^{s,\text{Gas}}\left(1-\tau_R^{ks,\text{Gas}}\right)}\right\} \\ \Delta\tau_{t,P}^{iks,\text{Oil}} &= -\Delta\tau_{t,P}^{s,\text{Gas}}GOR^{iks}\left(1-\vartheta^{iks}\right)\frac{P_t^{s,\text{Gas}}}{P_t^{\text{Oil}}\left(1-\varsigma_t^{\text{Oil}}MS_t^{ks,\text{Oil}}\right)+\sigma^{iks}} \end{aligned} \tag{6.241}$$

That is, it is sufficient to adjust the severance tax on oil and gas production at Federal level to obtain the same optimal choices obtained in the baseline setup under tax adjustments (6.43) and (6.44). In a similar way, it is possible to derive the formulas for the adjustment in the severance tax rates of gas-only fields, as well as in the rate of deduction of non-commercial gas use and unavoidable gas losses and in the lump-sum deduction amount for oil fields, in a way that generates the same incentives produced by the changes in the specific tax rates listed in (6.45) and (6.46). This ensures that and all the equilibrium outcomes are the same as those generated by the baseline model under the tax scheme outlined in (6.43), (6.44), (6.45) and (6.46). The intuition underpinning the formulas in (6.241) is unchanged with respect to those in (6.43) and (6.44). Namely, the reduction in the tax rate on natural gas production must be exactly compensated in terms of marginal profits for firm k by an increase in the tax rate on oil production, which is proportional to the gas/oil ratio. However, because severance taxes are *ad valorem* rather than specific taxes, their marginal effect on firm's profits is a function of oil and gas prices. Thus, in order to exactly compensate the firm at the margin, the optimal tax rate on oil production must also be multiplied by a term that is a function of the prices of natural gas and oil and that also adjusts for the market power of firm k on the crude market.

6.5.4 Alternative Solutions

The core of our tax reform proposal consists in the adjustment of two tax rates: the tax rate on crude sales and on gas sales by oil fields. The other tax provisions, such as the change in the tax rate faced by gas-only fields, are not crucial. They are meant to offset the excess supply of natural gas caused by the reduction of gas waste (i.e., the elimination of flaring and intentional venting and the reduction of unintentional venting) and avoid in turn any possible effect of the policy reform on equilibrium prices. However, the tax on gas production imposed on gas-only fields is not the only possible way to offset such excess natural gas supply. One could obtain a similar result through increasing the tax rate on the purchase of goods that are gross substitutes to natural gas in some midstream industry. An example is given by the rise of a specific tax on coal use in electricity production. If the cross-price elasticity of the demand for gas by power plants with respect to the price of coal is positive (i.e., coal and natural gas are gross substitutes in the production of electricity) and sufficiently large in magnitude, then there exists a specific tax rate on coal purchases by power plants which exactly offsets the excess natural gas supply mentioned above, ensuring that the natural gas price is unchanged by the introduction of the tax reform. However, note that there is no guarantee that such a policy would deliver the same level of power plants' profits that prevail under the original tax scheme. Thus, this solution only preserves some of the results stated in Proposition 3. An alternative approach is that of eliminating the excess natural gas supply via direct government purchases or via subsidies to alternative uses, such as the production of blue hydrogen. This approach would avoid losses for all firms and ensures weakly larger consumption of consumption goods $\mathbf{c}_t^s$, but would drain government revenue, implying that the policy may not be revenue-neutral and cause a fall in other consumption C_t^s. Thus, this solution also preserves some but not all the results stated in Proposition 3.

6.5.5 Tax on Coal and Gas Purchases

Let us consider an alternative tax scheme that allows for the excess gas supply due to the reduction of flaring and venting performed by oil firms to be offset by the demand from the power sector, with no effect on electricity output and price. In particular, the alternative scheme is identical to the baseline reform with respect to the taxation of oil firms, but does not prescribe any change in the taxation of gas firms. Conversely, the alternative scheme introduces linear taxes on the purchase of natural gas and thermal coal by firms operating in the power sector. Let the l-th net output of midstream firms y_{lt}^{js} be electricity, and the k-th net output y_{kt}^{js} be thermal coal. The excess gas supply due to the effect of the reform on the oil extraction sector at constant oil and gas prices is

$$\text{ExcessGas}_t^s = -\sum_{k=1}^{K^s}\sum_{i=1}^{I^{ks}}\left(\Delta\text{Flare}_t^{iks} + \Delta\text{Vent}_t^{iks}\right) , \tag{6.242}$$

where $\Delta\text{Flare}_t^{iks}$ and $\Delta\text{Vent}_t^{iks}$ represent the reduction in flaring and venting by oil&gas firm i due to the introduction of the reform at constant prices. Let $E^s \subseteq \{1, 2, ..., J^s\}$ be the set of midstream firms operating in the power sector of country s and $\eta_{xy}^{s,E}$ denote the cross-price elasticity of the net supply of commodity x with respect to the price of commodity y within the power sector of country s. For instance,

$$\eta_{Gk}^{s,E} = \frac{\partial \sum_{j\in E^s} \mathrm{G}_t^{js}}{\partial p_{kt}^s} \frac{p_{kt}^s}{\sum_{j\in E^s} \mathrm{G}_t^{js}} \tag{6.243}$$

is the elasticity of the net supply of natural gas from the power sector with respect to the price of thermal coal. Firstly, we assume that thermal coal is supplied to firms in country s at a global market price $p_{kt}^s = p_{kt}^s\left(\text{Coal}_t\right)$, where Coal_t represents the global supply of thermal coal. Secondly, we assume that thermal coal is used solely by power firms in this economy. As a consequence, the cross-price elasticities of the net supplies of midstream firms, other than coal-fueled power plants with respect to the price of thermal coal, equal zero. This assumption implies that, for instance, the equilibrium net output choices of oil refineries is unaffected by changes in the price of coal as long as the prices of crude and refined oil products are unchanged. Moreover, we assume that the cross-price elasticities of the net supply of other inputs used by the power sector with respect to p_{kt}^s and $P_t^{s,\text{Gas}}$ are equal to zero. This is equivalent to assume that power plants other than fossil fuel operated ones (e.g., nuclear power plants) cannot use either gas nor coal as inputs. Under these assumptions, the net output supply of all firms other than natural gas- and coal-powered power plants is unaffected by the tax scheme. However, the scheme may, in principle, affect the market price of coal because it implies a fall in the demand for coal by power plants. In turn, because the own-price elasticity of the supply of electricity from coal-fueled power plant is typically different from zero, this implies that a change in the specific tax on coal consumption from coal-fueled power plant may affect the equilibrium price of coal. Specifically, the total effect of a change in the specific tax on coal purchases on the price of coal equals

$$\frac{\eta_{kk}^{s,S}}{\eta_{kk}^{s,S} - \eta_{kk}^{s,E}} , \tag{6.244}$$

where η_{kk}^{sS} denotes the own-price supply elasticity of coal in country s. The alternative scheme aims to achieve two targets: (1) eliminating the excess natural gas supply due to the taxation on oil firms and (2) delivering zero effect on the prices of consumer goods, such that the consumption of energy-related goods is unaffected by the policy change. Firstly, in order to offset the excess supply of gas, the scheme must solve:

$$\text{ExcessGas}_t^s = \sum_{j\in E^s} \frac{\partial \mathrm{G}_t^{js}}{\partial P_t^{s,\text{Gas}}} \Delta b_t^{js} + \frac{\partial \mathrm{G}_t^{js}}{\partial p_{kt}^s} \frac{\eta_{kk}^{s,S}}{\eta_{kk}^{s,S} - \eta_{kk}^{s,D}} \Delta a_{kt}^{js} \tag{6.245}$$

where a_{kt}^{js} and b_t^{js} are linear specific taxes on coal and gas transactions made by power sector firms, respectively. Given that natural gas and coal are net outputs that typically have negative values for power sector firms (i.e., they are net inputs), the values of a_{kt}^{js} and b_t^{js} should be interpreted as (possibly negative-valued) subsidies. Thus, an increase in a_{kt}^{js} and b_t^{js} corresponds to a reduction in the tax rate on coal (natural gas) purchased by power firms. Secondly, we want the scheme to ensure unchanged electricity price for consumers. For small price changes, a zero effect of the tax scheme on electricity prices is obtained if the following equation is satisfied:

$$\sum_{j\in E^s} \frac{\partial \mathrm{y}_{lt}^{js}}{\partial P_t^{s,\mathrm{Gas}}} \Delta b_t^{js} + \frac{\partial \mathrm{y}_{lt}^{js}}{\partial p_{kt}^{s}} \frac{\eta_{kk}^{s,S}}{\eta_{kk}^{s,S} - \eta_{kk}^{s,D}} \Delta a_{kt}^{js} = 0 \tag{6.246}$$

Using the definitions of the price elasticities, the equations (6.245) and (6.246) can be rewritten as:

$$\mathrm{ExcessGas}_t^s = \eta_{GG}^E \frac{\sum_{s\in E} \mathrm{G}_t^{js}}{P_t^{s,\mathrm{Gas}}} \Delta b_t^s + \eta_{Gk}^E \frac{\sum_{s\in E} \mathrm{G}_t^{js}}{p_{kt}^s} \frac{\eta_{kk}^{s,S}}{\eta_{kk}^{s,S} - \eta_{kk}^{s,D}} \Delta a_{kt}^s \tag{6.247}$$

and

$$\eta_{lG}^E \frac{\sum_{s\in E} \mathrm{y}_{lt}^{js}}{P_t^{\mathrm{Gas}}} \Delta b_t^s + \eta_{lk}^E \frac{\sum_{s\in E} \mathrm{y}_{lt}^{js}}{p_{kt}^{js}} \frac{\eta_{kk}^{s,S}}{\eta_{kk}^{s,S} - \eta_{kk}^{s,D}} \Delta a_t^s = 0\ . \tag{6.248}$$

The scheme solving the system of equations (6.247) and (6.248) ensures that the electricity output y_{lt}^{js} is unchanged at constant natural gas market price $P_t^{s,\mathrm{Gas}}$, implying in turn that the equilibrium price of electricity is also unaffected by the policy. We assume that the technology of a fossil fuel-operated power plant is captured by the transformation function

$$MTF_t^{js}\left(\mathbf{y}_t^{js}, \mathrm{O}_t^{js}, \mathrm{G}_t^{js}, \mathrm{F}_t^{js}, \mathrm{MInS}_t^{js}, \mathrm{MZ}_t^{js}\right) \tag{6.249}$$

satisfies $\frac{\partial MTF_t^{js}(\cdot)}{\partial \mathrm{x}_{ht}^{js}} \to 0$ for $\mathrm{x}_{ht}^{js} \leq 0$ and $\frac{\partial MTF_t^{js}(\cdot)}{\partial \mathrm{x}_{ht}^{js}} \to +\infty$ $\mathrm{x}_{ht}^{js} > 0$ for any argument x_{ht}^{js} other than y_{lt}^{js}, y_{kt}^{js}, and G_t^{js}. This assumption ensures that the optimal choice of any net output x_{ht}^{js} other than y_{lt}^{js}, y_{kt}^{js}, and G_t^{js} equals zero. Setting $MTF_t^{js}(\cdot)$ equal to zero and differentiating it with respect to the price of a commodity, e.g., P_t^{Gas} and rearranging the resulting equation, we obtain:

$$\frac{\partial \mathrm{y}_{lt}^{js}}{\partial P_t^{\mathrm{Gas}}} + \left(\frac{\partial MTF_t^{js}(\cdot)}{\partial \mathrm{y}_{kt}^{js}} \Bigg/ \frac{\partial MTF_t^{js}(\cdot)}{\partial \mathrm{y}_{lt}^{js}}\right) \frac{\partial \mathrm{y}_{kt}^{js}}{\partial P_t^{\mathrm{Gas}}} + \left(\frac{\partial MTF_t^{js}(\cdot)}{\partial \mathrm{G}_t^{js}} \Bigg/ \frac{\partial MTF_t^{js}(\cdot)}{\partial \mathrm{y}_{lt}^{js}}\right) \frac{\partial \mathrm{G}_t^{js}}{\partial P_t^{\mathrm{Gas}}} = 0\ . \tag{6.250}$$

The equilibrium change in the output of electricity is equal to the equilibrium change in the amount of natural gas and coal used in electricity production multiplied by the marginal rate of transformation between natural gas and electricity and coal and electricity, respectively. Note that the F.O.C.s of the firm's maximization problem imply that at the optimal choice the marginal rate of transformation between natural gas (thermal coal) and electricity is the same for all the firms that consume a positive amount of natural gas (thermal coal) as long as all such firms face the same marginal tax rate on natural gas and thermal coal consumption, and electricity production. Let $\zeta^{s,\mathrm{Gas}}$ ($\zeta^{s,k}$) denote the equilibrium industry-level time-invariant marginal rate of transformation between natural gas (thermal coal) and electricity. Under these assumptions, at the optimal choice for all firms the formula for the marginal effect of a change in natural gas price on the aggregate electricity production writes:

$$\frac{\partial \sum_{j\in E^s} \mathrm{y}_{lt}^{js}}{\partial P_t^{\mathrm{Gas}}} = -\left(\zeta^{s,\mathrm{Gas}} \frac{\partial \sum_{j\in E^s} \mathrm{G}_t^{js}}{\partial P_t^{\mathrm{Gas}}} + \zeta^{s,k} \frac{\partial \sum_{j\in E^s} \mathrm{y}_{kt}^{js}}{\partial P_t^{\mathrm{Gas}}}\right) \tag{6.251}$$

Similarly, with respect to the price of coal we get:

$$\frac{\partial \sum_{j\in E^s} \mathrm{y}_{lt}^{js}}{\partial p_{kt}^{js}} = -\left(\zeta^{s,\mathrm{Gas}} \frac{\partial \sum_{j\in E^s} \mathrm{G}_t^{js}}{\partial p_{kt}^{js}} + \zeta^{s,k} \frac{\partial \sum_{j\in E^s} \mathrm{y}_{kt}^{js}}{\partial p_{kt}^{js}}\right) \tag{6.252}$$

Using the two formulas above, we can obtain the formulas for η_{lG}^E,

$$\eta_{lG}^E = -\left(\zeta^{s,\mathrm{Gas}} \eta_{GG}^E \frac{\sum_{j\in E^s} \mathrm{G}_t^{js}}{P_t^{s,\mathrm{Gas}}} + \zeta^{s,k} \eta_{kG}^E \frac{\sum_{j\in E^s} \mathrm{y}_{kt}^{js}}{P_t^{s,\mathrm{Gas}}}\right) \frac{\partial P_t^{\mathrm{Gas}}}{\sum_{j\in E^s} \mathrm{y}_{lt}^{js}}\ , \tag{6.253}$$

and η^E_{lk},

$$\eta^E_{lk} = -\left(\zeta^{s,\text{Gas}}\eta^E_{Gk}\frac{\sum_{j\in E^s}\text{G}^{js}_t}{p^{js}_{kt}} + \zeta^{s,k}\eta^E_{kk}\frac{\sum_{j\in E^s}\text{y}^{js}_{kt}}{p^{js}_{kt}}\right)\frac{p^{js}_{kt}}{\sum_{s\in E}\text{y}^{js}_{lt}} , \tag{6.254}$$

We substitute formulas (6.253) and (6.254) into equation (6.248) to get:

$$\begin{aligned} &-\left[\zeta^{s,\text{Gas}}\eta^E_{GG}\left(\textstyle\sum_{j\in E}\text{G}^{js}_t\right) + \zeta^{s,k}\eta^E_{kG}\left(\textstyle\sum_{j\in E}\text{y}^{js}_{kt}\right)\right]\frac{\Delta b^s_t}{P^{s,\text{Gas}}_t} \\ &-\left[\zeta^{s,\text{Gas}}\eta^E_{Gk}\left(\textstyle\sum_{j\in E}\text{G}^{js}_t\right) + \zeta^{s,k}\eta^E_{kk}\left(\textstyle\sum_{j\in E}\text{y}^{js}_{kt}\right)\right]\frac{\eta^{s,S}_{kk}}{\eta^{s,S}_{kk}-\eta^{s,D}_{kk}}\frac{\Delta a^s_t}{p^{js}_{kt}} = 0 \end{aligned} \tag{6.255}$$

Solve the system of equations (6.247) and (6.255) for $\frac{\Delta a^s_t}{p^s_{kt}}$ and $\frac{\Delta b^s_t}{P^{s,\text{Gas}}_t}$ to by how much taxation on coal and natural gas purchases should change

$$\begin{cases} \frac{\Delta a^s_t}{p^s_{kt}} &= -\frac{\eta^{s,S}_{kk}-\eta^{s,E}_{kk}}{\eta^{s,S}_{kk}\left(\eta^E_{kk}\eta^E_{GG}-\eta^E_{kG}\eta^E_{Gk}\right)}\left(\eta^E_{kG} + \frac{\zeta^{s,\text{Gas}}\sum_{j\in E^s}\text{G}^{js}_t}{\zeta^{s,k}\sum_{j\in E^s}\text{y}^{js}_{kt}}\eta^E_{GG}\right)\frac{\text{ExcessGas}^s_t}{\sum_{j\in E}\text{G}^{js}_t} \\ \frac{\Delta b^s_t}{P^{s,\text{Gas}}_t} &= \frac{1}{\eta^E_{kk}\eta^E_{GG}-\eta^E_{kG}\eta^E_{Gk}}\left(\eta^E_{kk} + \eta^E_{Gk}\frac{\zeta^{s,\text{Gas}}\sum_{j\in E^s}\text{G}^{js}_t}{\zeta^{s,k}\sum_{j\in E^s}\text{y}^{js}_{kt}}\right)\frac{\text{ExcessGas}^s_t}{\sum_{s\in E}\text{G}^{js}_t} \end{cases} \tag{6.256}$$

Estimation of Net Supply Elasticities

In order to estimate values for the net supply elasticities of interest, $\eta^E_{kk}, \eta^E_{GG}, \eta^E_{kG}, \eta^E_{Gk}$ we start from the elasticities of conditional factor demands. Let w^s_j denote the cost share of input j in non-renewable electricity production,

$$w^s_{jt} = \frac{p_{jt}q_{jt}}{\sum_{h=1}^3 p_{ht}q_{ht}} \tag{6.257}$$

where q_{jt} denotes the aggregate quantity of input $j \in \{1,2,3\}$ used by the power sector in country s. Specifically, subscript 1 corresponds to natural gas, 2 to thermal coal, and 3 to nuclear fuel (uranium). For instance, q_{1t} denotes the amount of natural gas demanded by the power sector of country s in period t, and must satisfy $q_{1t} = -\sum_{j\in E^s}\text{G}^{js}_t$ at any given price vector (i.e., at given prices, the values of the factor demands must be equal to the negative of the values of the net supply functions). Following Considine [1989] and EIA [2012], we assume that w^s_{jt} has the functional form

$$w^s_{jt} = \frac{\exp\left\{\chi^s_j + \sum_{z=1}^3 \psi^s_{jz}\ln p^s_{zt} + \phi^s_j \ln e^s_t\right\}}{\sum_{h=1}^3 \exp\left\{\chi^s_h + \sum_{z=1}^3 \psi^s_{hz}\ln p^s_{zt} + \phi^s_h \ln e^s_t\right\}} \tag{6.258}$$

for each $j \in \{1,2,3\}$, where $e^s_t = \sum_{j\in E^s}^{J^s} y_{lt}$ is the aggregate supply of electricity from non-renewable sources of country s. We differentiate (6.258) w.r.t. p_{kt} to obtain the elasticity of cost share w^s_{jt} with respect to input price p_{kt}, denoted by ϑ^s_{jk},

$$\vartheta^s_{jk} = \frac{\partial w^s_{jt}}{\partial p_{kt}}\frac{p_{kt}}{w^s_{jt}} = \psi^s_{jk} - \sum_{h=1}^3 \psi^s_{hk}w^s_{ht} \tag{6.259}$$

In a similar way, we can differentiate (6.257) w.r.t. p_{kt} to obtain another formula for ϑ^s_{jk},

$$\vartheta^s_{jk} = \epsilon^{s,E}_{jk} - w^s_{kt} - \sum_{h=1}^3 w^s_{ht}\epsilon^{s,E}_{hz} + \mathbf{1}\left[k=j\right] . \tag{6.260}$$

The homogeneity of the conditional demand functions implies:

$$\sum_{j=1}^3 \epsilon^{s,E}_{kj} = \sum_{j=1}^3 \frac{\partial q^s_{kt}}{\partial p_{jt}}\frac{p_{jt}}{q^s_{kt}} = 0\ \forall k \tag{6.261}$$

Secondly, the symmetry of the substitution matrix implies

$$\frac{\partial q_{kt}^s}{\partial p_{jt}} = \frac{\partial q_{jt}^s}{\partial p_{kt}} \ \forall j, k \tag{6.262}$$

using condition (6.262) into (6.261) we obtain:

$$\sum_{j=1}^{3} \epsilon_{jk}^{s,E} w_{jt}^s = 0 \ \forall k \tag{6.263}$$

Using this result in (6.260) and imposing the normalization $\sum_{h=1}^{3} \psi_{hk}^s w_{ht}^s = 0$, we can equate the RHS of (6.259) with the RHS of (6.260) and solve for $\epsilon_{jk}^{s,E}$ to get:

$$\epsilon_{jk}^{s,E} = \left(\Psi_{jk}^s + 1\right) w_{kt}^s - \mathbf{1}\left[k = j\right] \tag{6.264}$$

where $\Psi_{jk}^s = \psi_{jk}^s / w_{kt}^s$. Note that the formula for the cost share (6.258) implies:

$$\ln w_{jt}^s = \chi_1^s + \psi_{j1t}^s \ln p_{1t}^s + \psi_{j2t}^s \ln p_{2t}^s + \psi_{j3t}^s \ln p_{3t}^s + \phi_j^s \ln e_t^s - \ln C\left(e_t^s, \mathbf{p}_t^s\right) \tag{6.265}$$

for $j = \{1, 2, 3\}$ where $C\left(e_t^s, \mathbf{p}_t^s\right) = \sum_{h=1}^{3} p_{ht} q_{ht}^*$ is the cost function. Moreover, the homogeneity condition (6.261) is satisfied if and only if:

$$\psi_{jj}^s = -\sum_{k \neq j} \psi_{jk}^s \ \forall j \tag{6.266}$$

Furthermore, the symmetry condition (6.262) implies:

$$\frac{\psi_{kj}^s}{w_{jt}^s} = \frac{\psi_{jk}^s}{w_{kt}^s} \ \forall j, k \tag{6.267}$$

Lastly, the fact that the cost share must add up to one for any possible value of e_t^s implies:

$$\phi_3^s = -\phi_1^s \frac{w_{1t}^s}{w_{3t}^s} - \phi_2^s \frac{w_{2t}^s}{w_{3t}^s} \tag{6.268}$$

Define $\Phi_j^s = \frac{\phi_j^s}{w_{3t}^s}$ for $j = \{1, 2, 3\}$. Using conditions (6.266) and (6.267) into (6.265) for $j = 1, 2, 3$ and combining the three resulting equations we obtain the following system of equations

$$\begin{cases} \ln\left(\frac{w_{1t}^s}{w_{3t}^s}\right) = & \left(\chi_{1t}^s - \chi_{3t}^s\right) - \left[\Psi_{12}^s w_{2t}^s + \Psi_{13}^s \left(w_{1t}^s + w_{3t}^s\right)\right] \ln \frac{p_{1t}^s}{p_{3t}^s} + \left(\Psi_{12}^s w_{2t}^s - \Psi_{23}^s w_{3t}^s\right) \ln \frac{p_{2t}^s}{p_{3t}^s} + \\ & + \left[\Phi_1^s \left(w_{1t}^s + w_{3t}^s\right) + \Phi_2^s w_{2t}^s\right] \ln e_t^s \\ \ln\left(\frac{w_{2t}^s}{w_{3t}^s}\right) = & \left(\chi_{2t}^s - \chi_{3t}^s\right) + \left(\Psi_{12}^s w_{1t}^s - \Psi_{13t}^s w_{1t}^s\right) \ln \frac{p_{1t}^s}{p_{2t}^s} - \left[\Psi_{23}^s \left(w_{2t}^s + w_{3t}^s\right) + \Psi_{13t}^s w_{1t}^s\right] \ln \frac{p_{2t}^s}{p_{3t}^s} + \\ & + \left[\Phi_2^s \left(w_{2t}^s + w_{3t}^s\right) + \Phi_1^s w_{1t}^s\right] \ln e_t^s \end{cases} \tag{6.269}$$

where the third equation for $\ln\left(\frac{w_{1t}^s}{w_{2t}^s}\right)$ is omitted because costs share must add up to one, implying that the inclusion of the third equation would result in over-identification. Lastly, we define ω_{1t}^s and ω_{2t}^s as

$$\begin{aligned} \omega_{1t}^s = & \ \chi_{1t}^s - \chi_{3t}^s - \chi_1^s + \chi_3^s + \alpha_{14}^s \ln \frac{w_{1t-1}^s}{w_{3t-1}^s} \\ \omega_{2t}^s = & \ \chi_{2t}^s - \chi_{3t}^s - \chi_2^s + \chi_3^s + \alpha_{24}^s \ln \frac{w_{2t-1}^s}{w_{3t-1}^s} \end{aligned} \tag{6.270}$$

and we assume they are i.i.d. shocks. Under these assumptions, we can write the empirical equations:

$$\begin{cases} \ln\left(\frac{w^s_{1t}}{w^s_{3t}}\right) = & \alpha^s_{10} + \alpha^s_{11}\ln\frac{p^s_{1t}}{p^s_{3t}} + \alpha^s_{12}\ln\frac{p^s_{2t}}{p^s_{3t}} + \alpha^s_{13}e^s_t + \alpha^s_{14}\ln\frac{w^s_{1t-1}}{w^s_{3t-1}} + \omega^s_{1t} \\ \ln\left(\frac{w^s_{2t}}{w^s_{3t}}\right) = & \alpha^s_{20} + \alpha^s_{21}\ln\frac{p^s_{1t}}{p^s_{3t}} + \alpha^s_{22}\ln\frac{p^s_{2t}}{p^s_{3t}} + \alpha^s_{23}e^s_t + \alpha^s_{24}\ln\frac{w^s_{2t-1}}{w^s_{3t-1}} + \omega^s_{2t} \end{cases} \tag{6.271}$$

where the parameters $\theta^s = \left\{\alpha^s_{1j}, \alpha^s_{2j}\right\}^4_{j=0}$ in (6.271) map into the structural parameters in (6.269), delivering a linear system of ten equations and ten unknowns that can be solved to obtain formulas for all the structural parameters of interest as functions of $\left\{\alpha^s_{1j}, \alpha^s_{2j}\right\}^4_{j=0}$ and, in turn, for the elasticities of interest $\{\{\epsilon^{s,E}_{jk}\}^3_{j=1}\}^3_{k=1}$.

In order to fit the system of equations (6.271), we collect power plant level data on fuel consumption, electricity generation, fuel costs, fuel quantities received, and indicators of the quality of the fuel received using the forms-923 and -860 of the EIA database for the years 2005-2020 [EIA, a,b][18].

In order to compute cost shares, we need information about fuel consumption and fuel costs. Fuel consumption data are available in volumetric units and MMBTU in the form-923. We take into account only the volumes of fuel consumed, measured in MMBTU, to generate electricity. As some Combined Heat and Power (CHP) plants generate not only electricity but also district heating, we break the raw data into total power consumption and fuel consumption for electricity generation. Once the power plant level quantities are collected, we aggregate them across time (from month to year), energy source (from energy source to fuel type[19]), power plants, and regions (from NERC regions to the entire US).

Prices for delivered fuels, measured in \$/MMBTU, are derived from the monthly receipts of received fuel quantities by the power plants and plant-specific monthly fuel costs including transportation again using form-923 [EIA, a]. This approach takes into account long transportation routes to the power plant, which increase the variable costs of the plants and thus the cost of electricity generation [Hughes and Lange, 2018]. Furthermore, high fuel consumption gets higher weight, which results from the use of lower-quality fuels with low heat content. In the absence of raw data on specific monthly fuel costs, these were imputed with the costs of the nearest power plants using the same energy source measured by the Haversine distance using the following algorithm:

The monthly fuel costs per specific energy source is obtained using a volume-weighted aggregate, which aggregates across time, energy source, power plants, and regions,

$$p_{j,es,t,m,n} = \frac{\sum\limits_{b\in B}\left(p_{j,es,t,m,n,b} \times o_{j,es,t,m,n,b} \times \rho_{j,es,t,m,n,b}\right)}{\sum\limits_{b\in B}\left(o_{j,es,t,m,n,b} \times \rho_{j,es,t,m,n,b}\right)} \tag{6.272}$$

$$p_{jt} = \frac{1}{N} \times \sum_{n=1}^{N}\left(\frac{p_{j,es,t,m,n}}{100}\right) \tag{6.273}$$

where o is the volumes of fuel delivered to power plant n, located in block b, in month m, year t, of fuel type j belonging to energy source es, and ρ is the heat content of the fuel, measured in MMBTU/Unit of Fuel.

Fuel prices for uranium used in nuclear power plants in \$/MMBTU are derived from the public wholesale prices provided by the EIA [EIA, d]. The volume-weighted mean total purchase price in \$/pound of U_3O_8 equivalent per year is used as the basis for calculation, as it includes all potential countries of origin for the uranium, and it integrates short-term, medium-term, and long-term purchase contracts. Electricity generation costs in \$/kWh were subsequently derived from the regression and additional cost information provided by the WNA [EIA, c] and converted into \$/MMBTU.

Putting together all the previous information, it is possible to fit equation (6.271) as a seemingly unrelated

[18]We filter out power plants located in Alaska and Hawaii, as these two states tend to be independent in terms of their electricity generation and have little connection to the rest of the US.

[19]We decide to aggregate between the specific energy source instead of the higher-level fuel type because of the qualitative differences between coal types such as lignite or anthracite, which can result in high deviations in the fuel costs.

Algorithm 1 Imputation of Missing Fuel Costs for Power Plants

```
function IMPUTEFUELCOSTSTIMEBASED(df)
    df ← Group df by plantId, blockNumber, and energySource
    for each entry of grouped df do
        if fuelCost_t == -9999 and (fuelCost_{t-1} and fuelCost_{t+1}) > 0 then
            fuelCost ← 1/2 × (fuelCost_{t-1} + fuelCost_{t+1})
        end if
    end for
    return df
end function

function IMPUTEFUELCOSTSGEOBASED(df)
    missingIndices ← Which entries in df have fuelCost == −9999
    validIndices ← Which entries in df have fuelCost > 0
    if missingIndices and validIndices! = ∅ then
        distances ← Haversine distances on longitude and latitude with R-function
  geosphere::distm(missing entries, valid entries)
        distances[distances > 200 km] ← NA
        for each entry in missingIndices do
            closest ← Which index with minimum distance[distance != NA]
            imputedFuelCost ← fuelCost [closest]
        end for
    end if
    return df
end function

df ← IMPUTEFUELCOSTSTIMEBASED(df)
df ← IMPUTEFUELCOSTSGEOBASED(df)
```

Year	Coal	Natural Gas	Nuclear
2005	1.872	8.754	0.439
2006	1.755	7.566	0.459
2007	1.852	7.480	0.525
2008	2.306	9.339	0.587
2009	2.615	5.067	0.587
2010	2.413	5.336	0.603
2011	2.524	5.012	0.633
2012	2.609	3.669	0.630
2013	2.630	4.569	0.615
2014	2.522	5.532	0.588
2015	2.387	3.740	0.579
2016	2.318	3.250	0.571
2017	2.256	3.749	0.554
2018	2.310	3.947	0.554
2019	2.209	3.113	0.539
2020	2.378	2.757	0.528

Table 6.12: Weighted average fuel prices [US-Dollar/MMBTU]

regression (SUR) in a system of equations that is estimated simultaneously comparable to [EIA, 2012]. Correlation of the error terms between equations is explicitly taken into account for the estimation of coefficients, what extends the SUR model from the assumptions of conventional OLS. In this way, the complex mutual and partially time-shifted dependencies between prices and demand for different fuels for electricity generation can be considered [Considine and Mount, 1984, Jones, 1995]. The system of equations is implemented in using the `R` package *systemfit*. The results of the regression are shown in 6.13.

Equation 1	Estimate	Std. Error	t-Value	Pr(> $\|t\|$)
(Intercept)	15.53	18.66	0.83	0.42
Log price ratio gas/nuclear	**0.70*****	(0.08)	8.86	0.00
Log price ratio coal/nuclear	-0.52	(0.32)	-1.58	0.14
Log total electricity generation	-0.64	(0.84)	-0.76	0.46
Log lagged consumption gas/nuclear	0.34	(0.17)	2.07	0.06
Residual standard error:	0.06 on 10 df (15 observations for 5 parameters)			
MSE:	0.003			
Adjusted R-Squared:	0.93			
Equation 2	Estimate	Std. Error	t-Value	Pr(> $\|t\|$)
(Intercept)	-23.34	24.58	-0.95	0.36
Log price ratio gas/nuclear	0.14	(0.08)	1.74	0.11
Log price ratio coal/nuclear	**1.31****	(0.38)	3.48	0.01
Log total electricity generation	1.03	(1.11)	0.93	0.38
Log lagged consumption coal/nuclear	**0.96*****	(0.17)	5.59	0.00
Residual standard error:	0.063 on 10 df (15 observations for 5 parameters)			
MSE:	0.004			
Adjusted R-Squared:	0.95			

Table 6.13: SUR model results

From Conditional Factor Demands to Net Supplies

In order to obtain values for the elasticities of the aggregate net supply functions of coal and natural gas from the power sector, we use the assumption that fossil fuel-operated power plants only use natural gas and/or thermal coal as inputs and only produce electricity as outputs. Given these assumptions, the following equations must hold true:

$$\frac{\partial}{\partial p_{ht}}\left[\sum_{j\in E^s}\mathrm{G}_t^{js}\right] = -\frac{\partial q_{1t}^s}{\partial p_{ht}} - \frac{\partial q_{1t}^s}{\partial e_t^s}\frac{\partial}{\partial p_{ht}}\left[\sum_{j\in E^s} y_{lt}^{js}\right] \tag{6.274}$$

and

$$\frac{\partial}{\partial p_{ht}}\left[\sum_{j\in E^s} y_{kt}^{js}\right] = -\frac{\partial q_{2t}^s}{\partial p_{ht}} - \frac{\partial q_{2t}^s}{\partial e_t^s}\frac{\partial}{\partial p_{ht}}\left[\sum_{j\in E^s} y_{lt}^{js}\right] \tag{6.275}$$

for each $h = 1, 2, 3$. Note that our assumptions imply $\eta_{zh}^{s,E} = 0$ for all h other than $h = z$ and $h = l$ for nuclear fuels. Using the formulas in (6.251) and (6.252) and the definition of net supply elasticity, the equations above rewrite as follows.

$$\eta_{Gh}^{s,E} = -\epsilon_{1h}^s + \frac{\epsilon_{1e}^s}{e_t^s}\left[\zeta^{s,\mathrm{Gas}}\eta_{Gh}^{s,E}q_{1t}^s + \zeta^{s,\mathrm{k}}\eta_{kh}^{s,E}q_{2t}^s\right] \tag{6.276}$$

$$\eta_{kh}^{s,E} = -\epsilon_{2h}^s + \frac{\epsilon_{2e}^s}{e_t^s}\left[\zeta^{s,\mathrm{Gas}}\eta_{Gh}^{s,E}q_{1t}^s + \zeta^{s,\mathrm{k}}\eta_{kh}^{s,E}q_{2t}^s\right] \tag{6.277}$$

Solving the system of equations (6.276) and (6.277), we obtain the formulas for the net supply elasticities of interest, which write:

$$\begin{aligned}
\eta_{GG}^{s,E} &= -\epsilon_{11}^s - \epsilon_{1e}^s\frac{\zeta^{s,\mathrm{k}}\epsilon_{21}^s q_{2t}^s + \zeta^{s,\mathrm{Gas}}\epsilon_{11}^s q_{1t}^s}{e_t^s - \zeta^{s,\mathrm{k}}\epsilon_{2e}^s q_{2t}^s - \zeta^{s,\mathrm{Gas}}\epsilon_{1e}^s q_{1t}^s}\\
\eta_{Gk}^{s,E} &= -\epsilon_{12}^s - \epsilon_{1e}^s\frac{\zeta^{s,\mathrm{k}}\epsilon_{22}^s q_{2t}^s + \zeta^{s,\mathrm{Gas}}\epsilon_{12}^s q_{1t}^s}{e_t^s - \zeta^{s,\mathrm{k}}\epsilon_{2e}^s q_{2t}^s - \zeta^{s,\mathrm{Gas}}\epsilon_{1e}^s q_{1t}^s}\\
\eta_{kG}^{s,E} &= -\epsilon_{21}^s - \epsilon_{2e}^s\frac{\zeta^{s,\mathrm{k}}\epsilon_{21}^s q_{2t}^s + \zeta^{s,\mathrm{Gas}}\epsilon_{11}^s q_{1t}^s}{e_t^s - \zeta^{s,\mathrm{k}}\epsilon_{2e}^s q_{2t}^s - \zeta^{s,\mathrm{Gas}}\epsilon_{1e}^s q_{1t}^s}\\
\eta_{kk}^{s,E} &= -\epsilon_{22}^s - \epsilon_{2e}^s\frac{\zeta^{s,\mathrm{k}}\epsilon_{22}^s q_{2t}^s + \zeta^{s,\mathrm{Gas}}\epsilon_{12}^s q_{1t}^s}{e_t^s - \zeta^{s,\mathrm{k}}\epsilon_{2e}^s q_{2t}^s - \zeta^{s,\mathrm{Gas}}\epsilon_{1e}^s q_{1t}^s}
\end{aligned} \tag{6.278}$$

Lastly, substituting the formulas for the conditional demand elasticities from (6.264) into the formulas in (6.278) we obtain the formulas for the net supply elasticities of interest as functions of estimated parameters and known quantities, namely:

$$\begin{aligned}
\eta_{GG}^{s,E} &= \Psi_{12}^s w_{2t}^s + \Psi_{13}^s w_{3t}^s - w_{1t}^s + 1 - \Phi_1^s w_{3t}^s \frac{\zeta^{s,\mathrm{k}}(\Psi_{12}^s w_{1t}^s + w_{1t}^s) q_{2t}^s - \zeta^{s,\mathrm{Gas}}(\Psi_{12}^s w_{2t}^s + \Psi_{13}^s w_{3t}^s - w_{1t}^s + 1) q_{1t}^s}{e_t^s - \zeta^{s,\mathrm{k}} \Phi_2^s w_{3t}^s q_{2t}^s - \zeta^{s,\mathrm{Gas}} \Phi_1^s w_{3t}^s q_{1t}^s} \\
\eta_{Gk}^{s,E} &= -(\Psi_{12}^s w_{2t}^s + w_{2t}^s) + \Phi_1^s w_{3t}^s \frac{\zeta^{s,\mathrm{k}}(\Psi_{12}^s w_{1t}^s + \Psi_{23}^s w_{3t}^s - w_{2t}^s + 1) q_{2t}^s - \zeta^{s,\mathrm{Gas}}(\Psi_{12}^s w_{2t}^s + w_{2t}^s) q_{1t}^s}{e_t^s - \zeta^{s,\mathrm{k}} \Phi_2^s w_{3t}^s q_{2t}^s - \zeta^{s,\mathrm{Gas}} \Phi_1^s w_{3t}^s q_{1t}^s} \\
\eta_{kG}^{s,E} &= -(\Psi_{12}^s w_{1t}^s + w_{1t}^s) - \Phi_2^s w_{3t}^s \frac{\zeta^{s,\mathrm{k}}(\Psi_{12}^s w_{1t}^s + w_{1t}^s) q_{2t}^s - \zeta^{s,\mathrm{Gas}}(\Psi_{12}^s w_{2t}^s + \Psi_{13}^s w_{3t}^s - w_{1t}^s + 1) q_{1t}^s}{e_t^s - \zeta^{s,\mathrm{k}} \Phi_2^s w_{3t}^s q_{2t}^s - \zeta^{s,\mathrm{Gas}} \Phi_1^s w_{3t}^s q_{1t}^s} \\
\eta_{kk}^{s,E} &= \Psi_{12}^s w_{1t}^s + \Psi_{23}^s w_{3t}^s - w_{2t}^s + 1 + \Phi_2^s w_{3t}^s \frac{\zeta^{s,\mathrm{k}}(\Psi_{12}^s w_{1t}^s + \Psi_{23}^s w_{3t}^s - w_{2t}^s + 1) q_{2t}^s - \zeta^{s,\mathrm{Gas}}(\Psi_{12}^s w_{2t}^s + w_{2t}^s) q_{1t}^s}{e_t^s - \zeta^{s,\mathrm{k}} \Phi_2^s w_{3t}^s q_{2t}^s - \zeta^{s,\mathrm{Gas}} \Phi_1^s w_{3t}^s q_{1t}^s}
\end{aligned} \quad (6.279)$$

which we can use in (6.256) to calculate the adjustment in the tax rates that offsets any effect of the tax reform on natural gas and electricity prices.

Using the regression coefficients, we can calculate the net-supply elasticities for natural gas and coal[20]. That means we are measuring to what extent the supply of fuels (taking into account imports and exports) reacts relatively to a change in the fuel price of one percentage point. Since the data is aggregated to an annual resolution and collected over a period of 16 years, the elasticities can be considered as long-term stable. Regulatory changes and the reduction or expansion of production capacities (through the exploitation of additional gas fields or the opening of mines) are explicitly reflected in the elasticity. The results for the own-price elasticities (η_{GG} and η_{CC}) as well as the cross-price elasticities (η_{GC} and η_{CG}) are shown in Table 6.14. The positive net-supply own-price elasticity confirm the expectation that power plants will react to a rise in fuel prices by reducing production volumes in the long term. That is, the (negative) net supply of fuel by power plants becomes smaller in magnitude, resulting in turn in a fall in the firm's electricity output. Regarding the strength of the elasticity, we find that for natural gas, overall, the suppliers' reaction is inelastic. That is in line with results from the literature [Ponce and Neumann, 2014, Mason and Roberts, 2018]. The own-price elasticity of coal suggests that coal mine operators react roughly proportionally to relative price changes. Again this result is in line with previous findings [Dahl, 2009, Coglianese et al., 2020]. The cross-price elasticities illustrate that e.g., coal mine operators supply less coal to power plants in the long term and presumably also reduce supply capacity when the price of natural gas rises. The magnitude of both elasticities is roughly the same and lower than one, which means that the price-induced reactions of supply to the electricity sector are more or less similar and inelastic.

η_{GG}	η_{GC}	η_{CG}	η_{CC}
0.910	-0.826	-0.895	1.033

Table 6.14: Net-supply elasticities for coal and natural gas

Substituting the long-term net-supply price elasticities into the system of equations (6.256), we can calculate the change in the coal and natural gas tax, which would make sure that all the excess natural gas supply due to the elimination of routine flaring and venting is absorbed and that the prices of consumer goods is unchanged. The results per year and long-term average can be seen in Table 6.15.

In the case of the tax change for natural gas, we see that the values from 2005 to 2019 are consistently negative but decreasing in magnitude. This ensures that the additional natural gas is purchased by the electricity sector, thanks to the reduced net average gas price. Only in 2020 the positive sign for the natural gas tax indicates that suppliers would need to pay higher taxes for the delivery of the excess gas. The annual change in the tax rate for coal purchases is positive up to and including 2018, which means that coal-operated power plants pay an extra amount in \$/MMBTU of coal. However, from 2019 to 2020 the sign is reversed, so that theoretically coal-operated power plants would have faced a lower tax rate for the delivery of coal.

[20]The only elasticity not calculated is $\eta_{kk}^{s,S}$, which we assume equal to 0.89 following Dahl [2009].

	Δ tax coal	Δ tax natural gas
2005	0.000273	-0.0115
2006	0.000483	-0.0166
2007	0.000523	-0.0170
2008	0.00124	-0.0408
2009	0.00196	-0.0310
2010	0.00262	-0.0480
2011	0.00663	-0.114
2012	0.00734	-0.106
2013	0.0138	-0.230
2014	0.0178	-0.362
2015	0.00859	-0.165
2016	0.00368	-0.0856
2017	0.00525	-0.129
2018	0.000790	-0.129
2019	-0.0122	-0.0108
2020	-0.00764	0.0271
∅	**0.00243**	**-0.0724**

Table 6.15: Yearly tax rate changes on coal and natural gas purchases by power firms

Environmental Effects

We can calculate the change in emissions relative to the baseline policy reform using the formula

$$\text{ExcessGas}_t^s CI^{\text{Gas}} - \left[\frac{\partial \sum_{j\in E^s} \text{y}_{kt}^{js}}{\partial p_{kt}^s} \frac{\eta_{kk}^{s,S}}{\eta_{kk}^{s,S} - \eta_{kk}^{s,E}} \Delta a_t^s + \frac{\partial \sum_{j\in E^s} \text{y}_{kt}^{js}}{\partial P_t^{s,\text{Gas}}} \Delta b_t^s \right] CI^{\text{k}} \,. \tag{6.280}$$

Substituting (6.256) allows us to rewrite (6.280) as

$$\text{ExcessGas}_t^s \left(CI^{\text{Gas}} - \frac{\zeta^{s,\text{Gas}}}{\zeta^{s,k}} CI^{\text{k}} \right) \,, \tag{6.281}$$

where $\zeta^{s,\text{Gas}}/\zeta^{s,k}$ represents the marginal rate of technical substitution between natural gas and coal of the power sector in country s[21]. In order to compute equation 6.281, we need to calculate the carbon intensity of the conversion parameters ζ. To obtain these numbers, we collect data for fuel consumption, electricity generation, and emission rates for each region were the extra gas could have been produced combining information from the EIA with the ones recorded by the eGRID Program (i.e., the program responsible for collecting the emissions of every NERC region for every fuel type used in the power sector) [EPA, 2020]. We compute the CI^{Gas} multiplying the output emission rates by the efficiency of the different gas power plant. The output emission rates are calculated as total annual adjusted emissions divided by annual net generation

$$\text{Output Emission Rate} = \frac{\text{Total annual adjusted emissions [TonCO}_2]}{\text{Annual net electricity generation [MWh]}} \tag{6.282}$$

as in EPA [2020]. Since, the emission data are not fully available from for the entire analysed period, we use the 2020 emission rates to obtain a conservative estimate [EPA, 2020]. The efficiency of the gas power plants, expressed as the amount of electrical energy produced (megawatt-hours, MWh) per unit of thermal energy input

[21] As mention in section 6.5.4, the alternative scheme presented in section 6.5.5 may not be revenue- and profit-enhancing. This scheme only ensures unchanged consumption of energy-related goods in country s, but cannot ensure that the policy is welfare-improving. Moreover, if this alternative reform is implemented solely in country s, the change in the thermal coal price, even if it is likely to be modest in magnitude, may cause carbon leakages towards countries other than s. Thus, equation 6.281 measures the extra emissions savings in country s. However, it does not account for possible carbon leakages.

(MMBTU)

$$\text{Gas Power Plant Efficiency} = \frac{\text{Annual net electricity generation [MWh]}}{\text{Units of thermal energy input [MMBTU]}} \quad (6.283)$$

measure how efficient different power plants in converting thermal energy into electrical energy. Both output emission rates and gas power plant efficiency are calculated for individual market areas using specific emission rates[22]. Then, we average them for the period under consideration to obtain the time-invariant carbon intensity

$$CI^{\text{Gas}} = \text{Gas Output Emission Rate} \cdot \text{Gas Power Plant Efficiency}. \quad (6.284)$$

In a similar way, we can compute the carbon intensity multiplied by the ratio of the substitution coefficients calculating the emissions avoided by generating electricity from natural gas instead of coal. The joint product $\zeta^{s,\text{Gas}}/\zeta^{s,\text{Coal}} CI^{\text{Coal}}$ is determined by the amount of electricity generated from natural gas and the emission factor for coal-based power generation under the assumption that the amount of generated electricity remains constant,

$$\frac{\zeta^{\text{Gas}}}{\zeta^{\text{Coal}}} CI^{\text{Coal}} = \text{Coal Output Emission Rate} \cdot \text{Gas Power Plant Efficiency}. \quad (6.285)$$

The additional emission savings resulting from the conversion of the natural gas into electricity and substituting coal in the electricity sector are -135 million metric Ton of Carbon Dioxide Equivalent (MM TCO_2e) for the entire period under investigation from 2005 to 2020, -8,4 MM TCO_2e on average per year and 0.023 MM TCO_2e on average per day. The largest emission savings can be achieved in the TRE NERC region, as the amount of surplus natural gas and thus the substitution of coal is highest there.

NERC Region	**Power Sector Emissions Savings 2005-2020 (t)**	**Power Sector Emissions Savings per year (t/year)**	**Power Sector Emissions Savings per Day (t/year/day)**
MRO	-32,775,224	-2,048,452	-5,612
TRE	-92,185,563	-5,761,598	-15,785
WECC	-9,546,946	-596,684	-1,635
$\sum$	-134,507,733	-8,406,734	-23,032

Table 6.16: Delta Emissions Results

Model's Limitations

There are some limitations in our approach, which should be mentioned.

First, considering nuclear power as qualitatively similar to a thermal power plant type in the analysis of net supply substitution elasticities in the power sector is not common in the literature. Nuclear power plants are unlikely to be switched in the dispatch decision of the power plant portfolio by the utilities based on fuel cost changes, preferring coal or natural gas combustion over nuclear power. Moreover, differently from some papers in the literature [Jones, 1995, EIA, 2012], our analysis does not allow for electricity generation using oil or oil derivatives such as petroleum coke. This simplifying assumption may potentially distort our estimates for the net supply elasticities in the power sector. However, given the extremely small share of power generation that the excluded inputs accounts for in the US during the last two decades, we believe the size of the aforementioned distortion should be negligible.

Second, we assume that the price of natural gas only depends on (national or regional) aggregate demand and

[22] Efficiencies lower than 0 or higher than 0.7 have been filtered out due to physical and technical limitations of the power plants [Fu et al., 2015].

supply, and does not vary with the quantity purchased by a given plant. This assumption could be interpreted as a stylized representation of an economic environment in which there is sufficient transport capacity and infrastructure for any additional quantity of natural gas delivered from the oilfields to the power plants, as otherwise additional and increasing cost surcharges for the transport of natural gas would have to be added to the fuel prices as natural gas sales by oil & gas firms increase.

Third, the Seemingly Unrelated Regression (SUR) model we employ assumes linearity, which may not accurately capture the potentially non-linear relationships in the energy market. Furthermore, we use yearly averages to compute the net supply elasticities, which do not account for seasonal monthly variations in energy consumption and production. This simplification may have important consequences because dispatch decisions in electricity markets are based on hourly (but mostly daily) commodity prices. Thus, even if at a yearly granularity, gas fired generation is cheaper (or more expensive) than coal fired generation, during significant parts of the year the reality could be vice versa.

Fourth, the calculation of emissions does not take into account upstream emissions, which are generated for the exploration and development of oil fields or gas wells and any upstream technology used. However, this limitation only affects the estimation of the emission levels, but not the quantification of the economic and environmental effects of our proposed policy, because the reform does not affect oil & gas firm's incentives to invest in exploration and development. As a result, those types of emissions should not be affected by the policy change and be largely irrelevant for the purpose of our policy evaluation exercise.

These limitations indicate areas for future research to refine the methodology and ideas to mitigate flaring and venting in the oil industry.

Part III

Conclusions

Chapter 7

Thesis Outcomes, Limitations, and Future Work

7.1 Outcomes

The present thesis explores the interplay between the economic decisions of oil firms, including petroleum refineries, and their resulting environmental impacts, with a particular focus on greenhouse gas emissions. By integrating extensive economic data from sources such as Wood Mackenzie and Rystad Energy with environmental data from both private and public datasets, this work establishes a foundation for techno-economic analysis across various phases of oil production and refining. This interdisciplinary approach facilitates a detailed examination of sector-specific emissions and their driving factors.

The economic analysis employs an Applied Industrial Organization (AIO) framework, combining economic theory with empirical modelling to investigate how oil firms respond to market forces and regulatory constraints. The analysis provides key insights into the strategic decisions oil companies make regarding exploration, extraction, production, and emissions management under diverse economic and policy conditions. Central to the thesis is the evaluation of the pricing mechanisms for discovered oil reserves and the opportunity costs associated with reducing natural gas flaring and venting. To quantify these two unobserved quantities, it is necessary to estimate two key parameters by fitting multiple equations:

- <u>The Shadow Price of Discovered Oil</u>: The thesis merges four quantities to compute this unobserved quantity. First, the demand elasticity is drawn from existing literature to understand market sensitivity to price changes. Second, the field-level selling price is estimated using a novel pricing equation that links the microfoundations of global oil demand to volume-weighted variations in the API gravity and sulfur content of traded crude oil. Specifically, the field-level price deviates from the global average based on the degree to which the API gravity and sulfur content differ from global averages at a given time. Fields with characteristics closer to the global average align more closely with global prices, while significant deviations — either higher or lower — result in price premiums or discounts. Third, the marginal extraction costs are derived using both traditional and new econometric specifications. Traditional models assume that costs are convexly increasing with the quantities produced while decreasing with recoverable reserves. This reflects geological dynamics, where reservoir pressure declines as extraction continues, leading to higher costs for incremental production. Depleting reserves often necessitate costly secondary and tertiary recovery methods, such as fluid or gas injection, to sustain production levels. In contrast, larger reserves provide greater initial pressure and more accessible hydrocarbons, lowering marginal costs. Traditional econometric models capture these dynamics by incorporating reserve levels and production rates into cost functions, enabling analyses of how depletion and recovery investments influence the economic behaviour of oil firms under varying conditions.

However, the new model presented in the thesis goes further by incorporating additional factors, such as development and depletion rates, to better capture the complexities of oil extraction costs. Unlike traditional approaches, the new specification explicitly addresses the imperfect substitutability between newly discovered oil and depleted reserves, particularly in unconventional formations like heavy, extra-heavy, shale, and tight oil. For example, heavy and extra-heavy oil deposits require significant input costs, such as steam injection or other enhanced recovery methods, to maintain production levels as reservoir pressure declines. Within this framework, the discovery of one additional barrel reduces marginal costs by less than the corresponding increase in marginal costs resulting from the extraction of one barrel. Consequently, it becomes crucial to disentangle the dual effects of discovery and depletion from their combined impact on recoverable reserves. By independently analysing the contributions of discoveries, which expand field capacity, and depletion, which reduces reservoir pressure, this new model provides a flexible and comprehensive framework for understanding the economic behaviour of oil firms. It accommodates diverse geological, technological, market, and regulatory conditions, offering a more nuanced analysis of cost structures and decision-making processes. Furthermore, this approach not only enhances our comprehension of resource economics but also facilitates the identification of strategic opportunities to mitigate environmental impacts arising from oil extraction and production activities.

- Intentional Venting Supply Function: The thesis makes this quantity function of six variables: 1) whether the oilfield engages in flaring, 2) the price of natural gas, 3) the quantity of oil extracted, 4) the amount of natural gas injected into the oilfield, 5) the volume of natural gas used on-site to generate electricity and/or heat for field operations, and 6) the degree of substitutability between flaring and venting. The first factor is represented as a dummy variable, equal to one if routine flaring occurs (i.e., flaring volumes are greater than zero) and zero otherwise. The price of natural gas is taken as the Henry Hub benchmark price. Oil extraction volumes are sourced from Rystad Energy, while the quantities of natural gas injected and used on-site are derived using field-specific data fed into the Oil Production Greenhouse Gas Emissions Estimator (OPGEE). Finally, the degree of substitutability between flaring and venting is computed as the residual of a separate production function, which models routine flaring volumes as a function of natural gas prices. This residual reflects how changes in gas prices influence the balance between flaring and venting practices, capturing the economic and operational dynamics of substitutability. Together, these variables provide a comprehensive framework for understanding the drivers of intentional venting in oilfield operations.

The environmental analysis employs a bottom-up life cycle assessment (LCA) designed to estimate greenhouse gas (GHG) emissions across the oil production and refining processes. The Oil Production Greenhouse Gas Emissions Estimator (OPGEE) is utilized to quantify upstream emissions, from exploration to the refinery gate. This model integrates geological, technological, and logistical variables, capturing field-specific differences in emissions intensity. By considering factors such as reservoir properties, extraction techniques, and transportation logistics, OPGEE provides a granular understanding of emissions variability across oilfields. This enables the identification of key drivers of emissions, such as energy use in extraction and flaring or venting of associated gas, which are critical for formulating targeted reduction strategies.

The Petroleum Refinery Life Cycle Inventory Model (PRELIM) is applied to assess midstream emissions, focusing on the refining of various crude oil types under diverse configurations. PRELIM accounts for the energy-intensive nature of refining processes and the chemical transformations required to produce usable products such as gasoline, diesel, and jet fuel. The model captures the emissions implications of refining heavier, sour crude oils versus lighter, sweet ones, providing insights into how feedstock characteristics and refinery designs influence emissions. For instance, refining heavy crudes often demands extensive upgrading processes, such as hydrocracking and coking, which contribute to higher emissions compared to simpler processing of light crudes.

Together, the AIO and the LCA approach offer a comprehensive assessment of emissions along the oil value chain, enabling a systematic identification of emissions hotspots and opportunities for targeted mitigation. By integrating these LCA models, the thesis bridges the gap between economic and environmental analyses, linking emissions to specific operational and decision-making contexts. This integrated approach not only highlights the environmental costs of upstream and midstream oil operations but also provides actionable insights for optimizing processes and reducing emissions in line with global climate goals.

The integration of economic and environmental analyses yields three novel findings:

- Emissions Disparities at the Extensive Margin: The thesis uncovers significant disparities in emissions among oilfields at the extensive margin, defined as those marginal oilfields that are economically viable only under favourable market conditions. These fields often operate with older, less efficient technologies and lack the infrastructure for effective emissions management. As a result, their greenhouse gas emissions per barrel are substantially higher than the industry average. This finding has profound implications for global climate policy: by reducing overall oil demand, these high-emission marginal sources are likely to be the first to become uneconomical, leading to a disproportionately large reduction in emissions relative to the decline in oil consumption. The thesis provides quantitative analysis demonstrating that targeted demand-side interventions, such as carbon pricing or demand-reducing policies, can effectively phase out these high-emission sources while minimizing disruption to the broader energy supply. This approach aligns economic incentives with environmental goals, emphasizing the importance of prioritizing emissions reductions at the extensive margin to maximize climate benefits.

- High Emissions among non-Marginal Oilfields: The thesis identifies a subset of oilfields that, despite low extraction energy requirements and potentially low marginal extraction costs, exhibit exceptionally high greenhouse gas emissions due to routine flaring and venting of co-extracted natural gas. These oilfields, often located in regions lacking adequate gas infrastructure or market access, waste valuable energy resources while releasing significant quantities of methane—a greenhouse gas with a global warming potential far exceeding that of carbon dioxide. The research highlights the dual environmental and economic inefficiencies of these practices, noting that flaring and venting not only contribute to climate change but also represent a failure to monetize associated gas effectively. Addressing this issue requires well-designed regulatory interventions that incentivize gas capture and utilization while discouraging environmentally harmful practices. The thesis examines the technical and economic barriers to reducing flaring and venting, such as the high upfront costs of gas infrastructure and market constraints for associated gas, and proposes policy measures tailored to overcome these challenges. These include subsidies for gas capture technologies, stricter enforcement of emissions standards, and fiscal incentives to encourage the development of gas markets. By focusing on this subset of high-emission oilfields, policymakers can achieve substantial emissions reductions while promoting more sustainable energy practices.

- A Revenue-Neutral Tax Reform to Reduce Routine Flaring and Venting: The thesis proposes a comprehensive tax reform framework designed to address the dual environmental and economic challenges associated with routine flaring and venting of natural gas. This policy links fiscal measures to the reservoir gas-oil ratio and the costs of marketing otherwise unprofitable gas, thereby incentivizing firms to reduce emissions without compromising profitability or market stability. By adjusting the tax on oil production in proportion to the reservoir's gas-oil ratio and reducing taxes on natural gas sales by the costs incurred in preventing waste, the reform achieves a balance between environmental accountability and economic incentives. This dynamic structure encourages the adoption of emissions-reducing technologies, such as gas capture and reinjection, while mitigating undesirable substitution effects between flaring and venting. Importantly, the proposed policy avoids adverse economic outcomes like carbon leakage or increased product prices, maintaining the equilibrium in global energy markets. Unlike traditional regulatory approaches or Pigouvian taxes, the framework is designed to be revenue-neutral, aligning with existing tax structures in countries like the United States while maintaining government revenue and consumer purchasing power. The reform is politically feasible, as it preserves corporate profitability and reduces incentives for lobbying or regulatory evasion. Furthermore, it addresses the competitive and legal limitations of current methane management policies by leveraging readily measurable metrics, such as the gas-oil ratio, to ensure compliance and transparency. This scalable, stakeholder-friendly solution not only aligns with international efforts like the World Bank's Zero Routine Flaring Initiative but also offers a robust pathway to achieving substantial methane emission reductions in the oil and gas industry. Through careful modeling and empirical validation, the thesis demonstrates that the proposed reform is capable of cutting emissions by more than 20%, underscoring its potential to contribute significantly to global climate goals.

Overall, this thesis underscores the significant emissions disparities across oilfields and advocates for targeted policies that address high-emission practices without necessarily curbing overall production levels. It leverages established economic frameworks to balance environmental goals with practical feasibility, providing a foundation for future research to expand upon these findings.

7.2 Limitations and Future Work

While the thesis successfully integrates supply-side economic modelling and life-cycle environmental assessments, the limitations inherent in the complexity of modelling the global oil market remain. Future research could address these limitations at least across three dimensions.

Micro-Modelling of the Market Power In all the papers, the market power of oil companies is taken as exogenous. It is quantified as the firm's market share divided by the demand elasticity. A more sophisticated solution would be to model the origin of the market power, in particular the one originating from the coordination exercised by the members of the Organization of Petroleum Exporting Countries (OPEC).

OPEC's influence on the global oil market is profound, as it controls a significant portion of the world's oil reserves and can manipulate production levels to influence prices. Future research could model the global oil market as a repeated game with asymmetric information. In this framework, each OPEC country aims to maximize its discounted expected profits while possessing incomplete information about the aggregate marginal production costs of other member countries.

This research could solve each country's problem in two stages. First, each country would determine its aggregate production level. Second, it would allocate this production across its oilfields, solving a cost minimization problem in the style of a multi-plant monopolist, which ignores environmental considerations. The model would feature a self-enforcing dynamic contract between a principal (e.g., Saudi Arabia) or a super-principal in the case of OPEC+ (Saudi Arabia and Russia) and various agents (other OPEC members), with a large number of small price-taker producers, the so called fringe, made of international oil companies. The contract would include annual quotas for each OPEC country, with violations leading to a punishment phase characterized by temporary breakdowns of the cartel and price reductions.

The presence of asymmetric information implies that quotas could be violated with positive probability, influencing field-level production. The optimal contract would be characterized by a perfect Bayes-Nash equilibrium, where countries decide to violate quotas based on their aggregate marginal production costs. This modelling strategy could yield structural equations for compliance and non-compliance choices, providing micro-founded estimations of field-level supply price elasticity. The latter could be used to perform relative and absolute misallocation measures, such as:

- Relative Environmental Missallocation

  1. CO_2e relative to no-OPEC scenario: difference between total CO_2e emitted and the amount emitted in an alternative scenario in which OPEC does not exist in that period and all major oil companies play a

game à la Cournot.

(a) within country: difference between total CO_2e emitted by country i and the amount emitted in an alternative scenario in which OPEC does not exist in that period (Cournot) and the expected price is adjusted, such that total country production is unchanged.

(b) within OPEC: difference between total CO_2e emitted by OPEC and the amount produced in an alternative scenario in which OPEC does not exist in that period (Cournot) and the expected price is adjusted, such that total OPEC production is unchanged, minus (a).

(c) globally: difference between total CO_2e emitted and the amount emitted in an alternative scenario in which OPEC does not exist in that period (Cournot), such that global production is unchanged, minus (a)+(b).

2. CO_2e relative to perfect competition scenario: difference between total CO_2e and the amount emitted in an alternative scenario in which every field behaves as a price-taking firm.

(a) within country: difference between total CO_2e emitted by country i and the amount emitted in an alternative scenario in which each field behaves as a price-taker firm and the expected price is adjusted, such that total country production is unchanged.

(b) within OPEC: difference between total CO_2e emitted by OPEC and the amount emitted in an alternative scenario in which OPEC does not exist in that period (Cournot) and the expected price is adjusted, such that total OPEC production is unchanged, minus (a).

(c) globally: difference between total CO_2e emitted and the amount produced in an alternative scenario in which OPEC does not exist in that period (Cournot), minus (a)+(b).

- Absolute Environmental Missallocation

(a) within country: difference between total CO_2e emitted by country i and the amount produced in an alternative scenario in which the country allocates its total production in a socially optimal way.

(b) within OPEC: difference between total CO_2e emitted by OPEC and the amount produced in an alternative scenario in which OPEC allocates its total production in a socially optimal way, minus (a).

(c) globally: difference between total CO_2e emitted and the amount produced in an alternative scenario in which OPEC allocates its total production in a socially optimal way, minus (a) + (b).

Micro-Modelling of the Demand for Crude Oil In this thesis, the demand for crude oil is represented by a "representative refinery" model, drawing on emissions data from over 300 refineries to simulate midstream behaviour. This approach was necessitated by limited access to granular economic data for individual refineries, but it introduces two challenges. First, the reliance on a representative refinery model restricts the economic analysis of the midstream sector, limiting insight into how specific refinery configurations, operational efficiencies, and regional differences impact costs and emissions. Second, the use of a representative firm model carries inherent limitations, as it assumes that all refineries behave uniformly, thereby overlooking potential variations in demand elasticity, production choices, and technological efficiencies among diverse refinery types. Future research could address these issues by building a more differentiated model of demand that captures refinery-specific characteristics, including adaptive responses to fluctuating crude prices, regulatory changes, and shifts in consumer demand for refined products. This could involve gathering detailed economic data from individual refineries or developing simulation models to reflect the heterogeneous nature of refineries within the broader crude oil demand framework.

Micro-Modelling of the Demand for Oil Refined Products In all the papers discussed, the demand elasticity for oil-refined products is treated as an exogenous parameter, estimated from existing literature. However, relying solely on historical estimates may overlook the nuances and dynamics of consumer behaviour. A more sophisticated approach would involve an in-depth examination of consumers' capacity to substitute petroleum-derived products with viable alternatives, which can vary significantly across different regions, economic contexts, and consumer demographics. To accurately capture the demand elasticity, it is crucial to consider both short-term and long-term consumer reactions to price changes. In the short term, consumers might respond to price increases by reducing their consumption volumes, seeking cheaper alternatives, or postponing non-essential travel. In the long term, they may invest in energy-efficient goods, such as hybrid or electric vehicles, or increase the use of public transportation or carpooling. Additionally, changes in consumer preferences and technological advancements, such as the rise of renewable energy sources and the increased availability of electric vehicles, can also significantly impact demand elasticity.

Future research could leverage micro-level data to explore these interactions in greater detail. For instance, analysing on-line marketplaces like AutoTrader could provide insights into how the prices of cars and trucks respond to expected oil price changes, indicating shifts in consumer preferences towards more fuel-efficient or alternative-fuel vehicles. Similarly, using mobility data from platforms like Apple and Google could help examine how variations in oil prices influence the number of kilometres driven, offering a detailed view of consumer behaviour in response to fuel cost fluctuations. Combining such micro-behavioural data with economic models could yield a more nuanced understanding of the transportation sector's environmental footprint. This approach would allow researchers to identify specific patterns and trends in consumer behaviour, enabling more accurate predictions of future demand for oil-refined products.

Integrating these three limitations with the modelling presented in the current thesis would provide a comprehensive framework to analyse the oil industry's supply and demand dynamics with unprecedented detail. By combining "micro-modelling of the market power," "micro-modelling of the demand for crude oil," and "micro-modelling of the demand for oil-refined products," future research could capture the interdependencies between supply-side decisions, midstream economic behaviours, and downstream consumer responses. For instance, modelling the origin of market power among oil producers, particularly OPEC's strategic decisions, could be linked to refinery-specific demand for crude oil, accounting for regional preferences and technological differences. Similarly, understanding how consumers adapt to price changes in oil-refined products could provide feedback loops that influence upstream production strategies. The resulting framework would integrate structural market modelling with behavioural analyses, allowing policymakers and industry stakeholders to simulate the effects of various economic and regulatory interventions across the entire oil value chain. This integrated approach would enhance the predictive accuracy of policy impacts, identify potential unintended consequences, and facilitate the design of balanced solutions that optimize environmental outcomes while maintaining economic stability.

Bibliography

Jessica P. Abella and Joule A. Bergerson. Model to investigate energy and greenhouse gas emissions implications of refining petroleum: Impacts of crude quality and refinery configuration. *Environmental Science Technology*, 46(24):13037–13047, 2012.

Jessica P. Abella, Kavan Motazedi, John Guo, and Joule A. Bergerson. Petroleum refinery life cycle inventory model (prelim) v1.0. *International Journal of Life Cycle Assessment*, 20(5):604–620, 2015.

Morris A. Adelman. *The International Oil Market: Perspectives and Prospects*. MIT Press, 1984.

Morris A. Adelman. *The Economics of Petroleum Supply*. The MIT Press, 1993.

Victor Aguirregabiria and Andres Luengo. A microeconometric dynamic structural model of copper mining decisions. *Unpublished manuscript, Department of Economics, University of Toronto*, 2016.

Rashed Al Dhaif, Ahmed F. Ibrahim, Saad Elkatatny, and Dham Al Shehri. Prediction of oil rates using machine learning for high gas oil ratio and water cut reservoirs. *Flow Measurement and Instrumentation*, 82, 2021.

Alaska Department of Administration. Alaska oil and gas conservation commission data. `https://www.commerce.alaska.gov/web/aogcc/Data.aspx`, 2017. Accessed: 2021-07-29.

Alberta Energy Regulator. Statistical reports. `https://portal.aer.ca/providing-information/data-and-reports/statistical-reports.html`, 2015. Accessed: 2021-07-29.

Robert Allen. Fugitive methane emissions in oil production. *Journal of Environmental Research*, 28:123–137, 2020.

Soren T Anderson, Ryan Kellogg, and Stephen W Salant. Hotelling under pressure. *Journal of Political Economy*, 126(3):984–1026, 2018.

Kenneth J. Arrow and Anthony C. Fisher. Natural resource extraction with irreversible environmental effects. *Journal of Economic Theory*, 7:64–73, 1974.

John Asker, Allan Collard-Wexler, and Jan De Loecker. (Mis) allocation, market power, and global oil extraction. *American Economic Review*, 109(4):1568–1615, 2019.

Robert Bacon and Silvana Tordo. Crude oil price differentials and differences in oil qualities: A statistical analysis. WB technical paper no. 81, 2005.

Jushan Bai. Panel data models with interactive fixed effects. *Econometrica*, 77(4):1229–1279, 2009.

Anindya Banerjee, Juan J. Dolado, John W. Galbraith, and David F. Hendry. *Co-integration, Error Correction, and the Econometric Analysis of Non-Stationary Data*. Oxford University Press, Oxford, 1993.

World Bank. Zero routine flaring by 2030 initiative. `https://www.worldbank.org/en/programs/zero-routine-flaring-by-2030/endorsers`, 2023.

Fernando Barbosa, Giacomo Bresciani, Peter Graham, Sven Nyquist, and Kate Yanosek. Oil and gas after covid-19: The day of reckoning or a new age of opportunity, 2020.

Douglas Bates. Fitting linear mixed models in R. *R News*, 5(1):27–30, 2005.

Douglas Bates, Martin Mächler, Ben Bolker, and Steve Walker. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1):1–48, 2015.

Nico Bauer, Irina Mouratiadou, Gunnar Luderer, Lavinia Baumstark, Robert J. Brecha, Ottmar Edenhofer, and Elmar Kriegler. Global fossil energy markets and climate change mitigation - an analysis with remind. *Climatic Change*, 136(1):69–82, 2016.

William J. Baumol and Wallace E. Oates. Environmental pollution and the theory of the second best. *Public Economics*, 1:1–29, 1971.

Steven M. Beattie. OPEC and non-OPEC oil supply: The role of OPEC in the new millennium. *Resources Policy*, 26(2):101–110, 2000.

Giacomo Benini, Adam R. Brandt, Valerio Dotti, and Hassan M. El-Houjeiri. The economic and environmental consequences of the petroleum industry extensive margin. *Working Paper*, 2023.

Jeremy Black. *War and Technology*. Indiana University Press, 1998.

Bloomberg. Costs of canadian oil sands projects fell dramatically in recent years; but pipeline constraints and other factors will moderate, 2019. Accessed: 2024-12-16.

Bloomberg. Electric vehicle outlook, 2020. Accessed: 2024-12-16.

BP Energy. *BP Statistical Review of World Energy 2022*. 2022.

BP Energy. *BP Statistical Review of World Energy 2023*. 2023.

Adam R. Brandt. Energy intensity and greenhouse gas emissions from thermal enhanced oil recovery. *Energy & Fuels*, 22(3):1680–1688, 2008.

Adam R. Brandt. Oil depletion and the energy efficiency of oil production: The case of California. *Sustainability*, 3(10):1833–1854, 2011.

Adam R. Brandt. Oil production and environmental impact. *Energy and Environment*, 6:75–93, 2013.

Adam R. Brandt, Yue Sun, and Kyriacos Vafi. Uncertainty in regional-average petroleum ghg intensities: Countering information gaps with targeted data gathering. *Environmental Science & Technology*, 49:679–686, 2014.

Adam R. Brandt, Yue Sun, Siddharth Bharadwaj, Daniel Livingston, and Eric Tan. Energy return on investment (eroi) for forty global oilfields using a detailed engineering-based model of oil production. *PLoS ONE*, 10: e0144141, 2015.

Adam R. Brandt, Mohammad S. Masnadi, Jacob G. Englander, Jonathan Koomey, and Deborah Gordon. Climate-wise choices in a world of oil abundance. *Environmental Research Letters*, 13(4):044027, 2018.

Adam R. Brandt, Mohammad S. Masnadi, Jeff Rutherford, and Jacob Englander. Updates to the OPGEE v3.0a candidate model. *Unpublished Manuscript*, 2020.

Adam R. Brandt, Hassan M. El-Houjeiri, Al-Hassan A. Badahdah, Mohammad S. Masnadi, Jean-Christophe Monfort, Giacomo Benini, Joule Bergerson, Ben Caldecott, Ken Gillingham, Deborah Gordon, Michael Jakob, Nigel Jenvey, Nathan Meehan, Jeanne-Mey Sun, and Sangwon Suh. Stranding of global oil reserves in a carbon-constrained world. *Working Paper*, 2021.

Roger M. Butler. *Thermal Recovery of Oil and Bitumen*. Prentice Hall, 1991.

Roger M. Butler. Steam-assisted gravity drainage: concept, development, performance and future. *Journal of Canadian Petroleum Technology*, 33(2):45–54, 1994.

Christoph Böhringer, Carolyn Fischer, Knut Einar Rosendahl, and Thomas F. Rutherford. Potential impacts and challenges of border carbon adjustments. *Nature Climate Change*, 12:22–29, 2022.

Rafaël Calel and Paasha Mahdavi. Opinion: The unintended consequences of antiflaring policies and measures for mitigation. *Proceedings of the National Academy of Sciences*, 117(23):12503–12507, 2020.

California Air Resources Board. LCFS crude oil lifecycle assessment. `https://www.arb.ca.gov/fuels/lcfs/crude-oil/crude-oil.htm`, 2017. Accessed: 2024-12-16.

Colin J. Campbell. Ultra-deepwater oil production and its implications for global energy security. *Energy Policy*, 39(11):7265–7273, 2011.

Canada-Newfoundland and Labrador Offshore Petroleum Board (C-NLOPB). Newfoundland and labrador offshore area gas flaring reduction implementation plan. `https://www.cnlopb.ca/legislation/regulations/`, 2017.

Canada-Newfoundland and Labrador Offshore Petroleum Board (C-NLOPB). Statistical information. `https://www.cnlopb.ca/information/statistics/#rm`, 2021.

M-C Oliver Chang, Judith C Chow, John G Watson, Philip K Hopke, Seung-Muk Yi, and Glenn C England. Measurement of ultrafine particle size distributions from coal-, oil-, and gas-fired stationary combustion sources. *Journal of the Air & Waste Management Association*, 54(12):1494–1505, 2004.

Lauren Choquette-Levy, Heather L. MacLean, and Joule A. Bergerson. *Oil Sands: The World's Greatest Reserve?* Canadian Energy Research Institute, 2005.

Steve Cicala, David Hémous, and Mathias G. Olsen. Adverse selection as a policy instrument: Unraveling climate change. Technical Report 30283, NBER Working Paper, 2023.

Dennis Clark. Aviation fuel composition and properties. *Fuel Processing Technology*, 86:1585–1594, 2004.

John Coglianese, Todd D. Gerarden, and James H. Stock. The effects of fuel prices, environmental regulations, and other factors on U.S. coal production, 2008–2016. *The Energy Journal*, 41(1):55–82, 2020.

William Collins, Raymond Orbach, Michael Bailey, Sebastien Biraud, Ian Coddington, David DiCarlo, John Peischl, Anand Radhakrishnan, and David Schimel. Monitoring methane emissions from oil and gas operations. *Optics Express*, 14:24326–24351, 2022.

Jon M. Conrad. Optimal resource extraction in the presence of environmental externalities. *Journal of Environmental Economics and Management*, 6:155–168, 1979.

Timothy J Considine. Separability, functional form and regulatory policy in models of interfuel substitution. *Energy Economics*, 11(2):82–94, 1989.

Timothy J. Considine and Timothy D. Mount. The use of linear logit models for dynamic input demand systems. *The Review of Economics and Statistics*, 66(3):434, 1984.

Gregory Cooney, Matthew Jamieson, Joe Marriott, Joule Bergerson, Adam Brandt, and Timothy J. Skone. Updating the U.S. life cycle ghg petroleum baseline to 2014 with projections to 2040 using open-source engineering-based models. *Environmental Science Technology*, 51(2):977–987, 2017.

Nigerian National Petroleum Corporation. Annual statistical: 2015 statistical bulletin, 2015.

Michele Costantini and Claudio Lupi. A simple panel-CADF test for unit roots. *Oxford Bulletin of Economics and Statistics*, 75(2):276–296, 2013.

Ed Crooks. Cost reductions help US shale oil industry pass first real test. `https://www.ft.com/content/65ebdd54-6b79-11e6-ae5b-a7cc5dd5a28c`, 2016. Online; accessed 30 June 2020.

Felipe Cruz. *Spatial Aspects of the Oil Market*. Routledge, 2013.

Maria Cruz and David Peterson. *Oil and Gas Logistics: Ensuring Supply Chain Stability*. Springer, 2013.

Jacques Crémer and Djavad Salehi-Isfahani. *Models of the oil market*. Taylor & Francis, 2013.

Carol Dahl. Energy demand and supply elasticities. In *Encyclopedia of Life Support Systems: Energy Policy*, volume 1 of *Energy Policy*, page 23. 2009.

Danish Energy Agency. 2015 production. `https://ens.dk/en/our-responsibilities/oil-gas`, 2016. Accessed: 2024-12-16.

Partha S. Dasgupta and Geoffrey M. Heal. Optimal extraction of nonrenewable resources with r&d toward perfect substitutes. *Journal of Environmental Economics and Management*, 6:113–126, 1979.

Paul De Boeck, Merel Bakker, Rob Zwitser, Michiel Nivard, Anne Hofman, Francis Tuerlinckx, et al. The estimation of item response models with the lmer function from the lme4 package in R. *Journal of Statistical Software*, 39(12):1–28, 2011.

Shantayanan Devarajan and Anthony C. Fisher. Exploration and scarcity. *Journal of Political Economy*, 90: 1279–1290, 1982.

Rebecca Dorsey-Palmateer and Bin Niu. The effect of carbon taxation on cross-border competition and energy efficiency investments. *Energy Economics*, 85:104602, 2020.

DOT. Highway statistics series, state motor-vehicle registrations - 2021. *https://www.fhwa.dot.gov/policyinformation/statistics/2021/mv1.cfm*, 2022.

DW and IEA. OPEC+ members in 2022. `https://www.dw.com/`, 2022.

J. Mason Earles and Anthony Halog. Consequential life cycle assessment: a review. *The International Journal of Life Cycle Assessment*, 16(5):445–453, 2011.

Markus Eberhardt and Christian Helmers. Addressing transmission bias in micro production function models: A survey for practitioners. Technical report, Working Paper, 2010.

Dan Eberhart. The oil sector will survive the arrival of the electric car just fine. `https://www.forbes.com/sites/daneberhart/2018/03/22/the-oil-sector-will-survive-the-arrival-of-the-electric-car-just-fine/`, 2018. Accessed: 2024-12-16.

EDGAR. GHG emissions of all world countries. `https://edgar.jrc.ec.europa.eu/report_2023`, 2023. Accessed: 2024-12-16.

EIA. Form EIA-923. `https://www.eia.gov/electricity/data/eia923/`, a.

EIA. Form EIA-860. `https://www.eia.gov/electricity/data/eia860/`, b.

EIA. Nuclear power economics. https://world-nuclear.org/information-library/economic-aspects/economics-of-nuclear-power.aspx, c.

EIA. Uranium marketing annual report. https://www.eia.gov/uranium/marketing/summarytable1b.php, d.

EIA. Fuel competition in power generation and elasticities of substitution. https://www.eia.gov/analysis/studies/fuelelasticities/pdf/eia-fuelelasticities.pdf, 2012.

EIA. Total petroleum and other liquids production. https://go.nature.com/3jfink3, 2017.

EIA. Average api and sulphur content. https://www.eia.gov/dnav/pet/petpnpcrqdcunusa.htm, 2020a.

EIA. Crude oil and lease condensate production by api gravity. https://www.eia.gov/dnav/pet/pet_crd_api_adc_mbblpd_m.htm, 2020b.

EIA. Price us refiners' acquisition costs for imported oils. https://www.eia.gov/totalenergy/data/browser/csv.php?tbl=T09.01, 2020c.

EIA. Drilling productivity report. https://www.eia.gov/petroleum/drilling/, 2021a.

EIA. Crude oil input qualities. https://www.eia.gov/dnav/pet/pet_pnp_crq_dcu_nus_a.htm, 2021b.

EIA. Total energy data. https://www.eia.gov/totalenergy/data/browser/csv.php?tbl=T09.01, 2021c.

Hassan M. El-Houjeiri, Adam R. Brandt, and James E. Duffy. Open-source LCA tool for estimating greenhouse gas emissions from crude oil production using field characteristics. *Environmental Science*, 47:5998–6006, 2013.

Hassan M. El-Houjeiri, Mohammad S. Masnadi, Kyriacos Vafi, James Duffy, and Adam R. Brandt. Oil production greenhouse gas emissions estimator OPGEE v2.0a: User guide & technical documentation, 2017.

EPA. Greenhouse gas emissions from a typical passenger vehicle. https://www.epa.gov/greenvehicles/greenhouse-gas-emissions-typical-passenger-vehicle, 2018a.

EPA. Greenhouse gas reporting program: EPA GHG emissions inventory guidance. https://www.epa.gov/ghgreporting, 2018b.

EPA. eGRID2020. https://www.epa.gov/egrid/historical-egrid-data, 2020.

EPA. Guidance for combustion-related greenhouse gas emissions. https://www.epa.gov/air-emissions-inventories/guidance-combustion-related-greenhouse-gas-emissions, 2022a.

EPA. Greenhouse gas inventory guidance: Direct emissions from stationary combustion sources. https://www.epa.gov/climateleadership, 2022b.

EPA. Inventory of U.S. greenhouse gas emissions and sinks by state. https://www.epa.gov/ghgemissions/state-ghg-emissions-and-removals, 2023.

EPA. Waste emissions charge for petroleum and natural gas systems. https://www.epa.gov/ghgemissions/waste-emissions-charge-petroleum-natural-gas-systems, 2024.

Dennis Epple. The econometrics of exhaustible resource supply: a theory and an application. In *Energy, Foresight, and Strategy*, pages 157–214. Routledge, 2014.

Edward W. Erickson and Robert M. Spann. Supply response in a regulated industry: the case of natural gas. *The Bell Journal of Economics and Management Science*, pages 94–121, 1971.

Surinder Parkash Farina. *Refining Processes Handbook*. Elsevier, Oxford, UK, 2011.

Y. Hossein Farzin. *Competition in the Market for an Exhaustible Resource*. Springer, 1986.

Bassam Fattouh. The dynamics of crude oil price differentials. *Energy Economics*, 32:334–342, 2010.

Bassam Fattouh. OPEC in a shale oil world: Where to next? *Energy Economics*, 46:511–519, 2014.

Carlo A. Favero. Taxation and the optimization of oil exploration and production: The UK continental shelf. *Oxford Economic Papers*, 44(2):187–208, 1992.

Franklin M. Fisher. Supply and costs in the U.S. petroleum industry: Two econometric studies. *The Johns Hopkins University Press*, 1974.

Chao Fu, Rahul Anantharaman, Kristin Jordal, and Truls Gundersen. Thermal efficiency of coal-fired power plants: From theoretical to practical assessments. *Energy Conversion and Management*, 105:530–544, 2015.

James H. Gary, Glenn E. Handwerk, and Mark J. Kaiser. *Petroleum Refining: Technology and Economics*. CRC Press, 5th edition, 2007.

Norsk Olje & Gass. Feltspesifikke utslippsrapporter, 2015.

Rolf Golombek, Alfonso A. Irrazabal, and Lin Ma. OPEC's market power: An empirical dominant firm model for the oil market. *Energy Economics*, 70:98–115, 2018.

Lewis Cecil Gray. Rent under the assumption of exhaustibility. *The Quarterly Journal of Economics*, 28(3): 466–489, 1914.

Murray Gray. Processing challenges of heavy and unconventional crude oils. *Fuel Processing Technology*, 73(1): 1–7, 2002.

William H Greene. Fixed and random effects in nonlinear models. *NYU Working Paper No. EC-01-01*, 2001.

Ronald C. Griffin and Daniel W. Bromley. Agricultural runoff as a nonpoint externality: A theoretical development. *American Journal of Agricultural Economics*, 64:547–542, 1982.

Environmental Assessment & Optimization Group. OPGEE: The oil production greenhouse gas emissions estimator. `https://eao.stanford.edu/research-areas/opgee`, 2017.

Joachim H. Güntner. How do oil producers respond to oil demand shocks? *Energy Economics*, 44:1–13, 2014.

James D. Hamilton. *Oil Prices, Energy Security, and Economic Stability*. Brookings Institution Press, 2013.

James D. Hamilton. *Time Series Analysis*. Princeton University Press, 2020.

Michael Harris. Crude oil transportation by rail: Economic and safety implications. *Transportation Economics Journal*, 42(3):145–162, 2019.

Michael Harris and Emily Taylor. Transporting refined products: Economic and environmental considerations. *Energy Transport Economics*, 45(2):102–118, 2019.

Arne Henningsen. Estimating censored regression models in R using the censreg package. *R package vignettes*, 5:12, 2010.

Jason Ho and Rachel Smith. Catalytic cracking and reforming: Enhancing gasoline quality and octane rating. *Journal of Petroleum Technology*, 71:89–95, 2019.

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

Jonathan E. Hughes and Ian Lange. Who (else) benefits from electricity deregulation? Coal prices, natural gas and price discrimination. Working paper, 7374, CESifo, 2018.

John C. Hull. The economics of OPEC: A game theoretic approach. *Journal of Political Economy*, 94(2):240–252, 1986.

IEA. Global EV outlook. `https://www.iea.org/reports/global-ev-outlook-2019#executive-summary`, 2019.

IEA. Curtailing methane emissions from fossil fuel operations. `https://www.iea.org/reports/curtailing-methane-emissions-from-fossil-fuel-operations`, 2021a.

IEA. *World Energy Outlook 2021*. 2021b.

IEA. CO_2 emissions 2023. `https://www.iea.org/reports/co2-emissions-in-2023/executive-summary`, 2023a.

IEA. The imperative of cutting methane from fossil fuels. `https://iea.blob.core.windows.net/assets/9efb310e-94d7-4c46-817b-9493fe5abb0a/Theimperativeofcuttingmethanefromfossilfuels.pdf`, 2023b.

IEA. Global methane tracker 2024. `https://www.iea.org/reports/global-methane-tracker-2024`, 2024.

Nkem Ihejirika. The Canadian oilsands and strategic approaches to profitability. *Oxford Institute for Energy Studies*, 56, 2019.

IMF. Policy responses to covid-19. `https://www.imf.org/en/Topics/imf-and-covid19/Policy-Responses-to-COVID-19`, 2020.

IMO. IMO 2020 and the transition to low-sulfur shipping fuels. `https://www.imo.org/`, 2020.

IPCC. *Climate Change 2021: The Physical Science Basis*. Cambridge University Press, 2021.

ISO. Environmental management — life cycle assessment — goal and scope definition and inventory analysis. (ISO 14041:1998), 1998.

Liang Jing, Hassan M. El-Houjeiri, Jean-Christophe Monfort, Adam R. Brandt, Mohammad S. Masnadi, Deborah Gordon, and Joule A. Bergerson. Carbon intensity of global crude oil refining and mitigation potential. *Nature Climate Change*, 10:526–532, 2020.

Clifton T. Jones. A dynamic analysis of interfuel substitution in U.S. industrial energy demand. *Journal of Business & Economic Statistics*, page 459, 1995.

Mark Jones and Arjun Patel. *Refining Processes: Principles and Applications*. Elsevier, 2019.

Sarah M. Jordaan. Land and water impacts of oil sands production in Alberta. *Environmental Science & Technology*, 43(20):8027–8033, 2009.

Michael Kah. Electric vehicle penetration and its impact on global oil demand: A survey of 2019 forecast trends, 2019.

Haikun Liu Kaihan Mintz-Woo, Francis Dennig and Thomas Schinko. Carbon pricing and covid-19. *Climate Policy*, 2020.

Kohei Kawaguchi. Topics in empirical industrial organization. `https://kohei-kawaguchi.github.io/EmpiricalIO/`, 2021.

David Keith. Emissions in upstream oil production. *Journal of Environmental Science*, 14:102–118, 2020.

Daniel J. Khazzoom. The FPC staff's econometric model of natural gas supply in the united states. *The Bell Journal of Economics and Management Science*, pages 51–93, 1971.

Lutz Kilian. The economic effects of energy price shocks. *Journal of Economic Literature*, 46(4):871–909, 2008.

Lutz Kilian. Not all oil price shocks are alike: Disentangling demand and supply shocks in the crude oil market. *American economic review*, 99(3):1053–1069, 2009.

Lutz Kilian. Understanding the estimation of oil demand and oil supply elasticities. *Energy Economics*, 107: 105844, 2022.

Lutz Kilian and Daniel P. Murphy. Why agnostic sign restrictions are not enough: Understanding the dynamics of oil market var models. *Journal of European Economic Association*, 10:1166–1188, 2012.

Lutz Kilian and Xiaoqing Zhou. The impact of OPEC's production announcements on oil prices: Does OPEC still matter? *Journal of Applied Econometrics*, 33(3):602–616, 2018.

George E. King. Thirty years of gas shale fracturing: What have we learned? *Society of Petroleum Engineers*, pages 1–19, 2010.

Tjalling C. Koopmans. Optimal growth in a nonrenewable resource using economy. *The Review of Economic Studies*, 41:1–14, 1974.

Clifford Krauss. The emergence of shale oil and gas as a transformative force. *The New York Times*, July 2012.

Mitch Kunce, Shelby Gerking, William Morgan, and Ryan Maddux. State taxation, exploration, and production in the us oil industry. *Journal of Regional Science*, 43(4):749–770, 2003.

Tony Lancaster. The incidental parameter problem since 1948. *Journal of Econometrics*, 95(2):391–413, 2000.

Alessandro Lanza, Matteo Manera, and Mario Giovannini. Modeling and forecasting cointegrated relationships among heavy oil and product prices. *Energy Economics*, 27(6):831–848, 2005.

H. Lee and J. Romero. IPCC - summary for policymakers. climate change 2023: Synthesis report. contribution of working groups i, ii and iii to the sixth assessment report of the intergovernmental panel on climate change. pages 1–34, 2023.

Sarah Lee and George Martin. Tanker vessels in global oil transportation: Trends and challenges. *Maritime Policy & Management*, 47(5):599–612, 2020.

Sarah Lee and George Martin. Specialized tankers for refined products: Efficiency and challenges. *Maritime Policy & Management*, 48(6):709–724, 2021.

Francois Lefebvre and Richard Deschenes. Desulfurization processes in petroleum refining. *Fuel Processing Technology*, 50(3):117–126, 1997.

James Levinsohn and Amil Petrin. Estimating production functions using inputs to control for unobservables. *The review of economic studies*, 70(2):317–341, 2003.

Kung-Yee Liang and Scott L Zeger. Longitudinal data analysis using generalized linear models. *Biometrika*, 73 (1):13–22, 1986.

Yan Liang and Wei Chen. Evolution of refining technology for unconventional crudes. *Energy Technology*, 5(5): 715–728, 2017.

John R. Livernois and Robert S. Uhler. Extraction costs and the economics of nonrenewable resources. *Journal of Political Economy*, 95(1):195–203, 1987.

Claudio Lupi. Unit root CADF testing with R. *Journal of Statistical Software*, 32:1–19, 2010.

Paul Webster MacAvoy and Robert S. Pindyck. The economics of the natural gas shortage (1960-1980). Technical report, MIT Energy Lab, 1974.

Wood Mackenzie. Upstream data tool. `https://www.woodmac.com/research/products/upstream/upstream-data-tool/`, 2018.

Mohamed Mahdi. *Oil and Empire: Middle Eastern Oil in Global Perspective*. Cambridge University Press, Cambridge, 2012.

IHS Markit. Canadian oil sands dialogue: Key insights and perspectives. `https://www.ihsmarkit.com/products/canadian-oil-sands-dialogue.html`, 2020.

Jacob Marschak and William H Andrews. Random simultaneous equations and the theory of production. *Econometrica*, pages 143–205, 1944.

Mohammad S. Masnadi and Adam R. Brandt. Climate impacts of oil extraction increase significantly with oilfield age. *Nature Climate Change*, 7:551–556, 2017a.

Mohammad S. Masnadi and Adam R. Brandt. Energetic productivity dynamics of global super-giant oilfields. *Energy & Environmental Science*, 10:1493–1504, 2017b.

Mohammad S. Masnadi, Hassan M. El-Houjeiri, Dominik Schunack, Yunpo Li, Jacob G. Englander, Alhassan Badahdah, Jean-Christophe Monfort, James E. Anderson, Timothy J. Wallington, Joule A. Bergerson, Deborah Gordon, Jonathan Koomey, Steven Przesmitzki, Inês L. Azevedo, Xiaotao Bi, James Duffy, Garvin Heath, Gregory A. Keoleian, Christophe McGlade, D. Meehan, Sonia Yeh, Fengqi You, Michael Wang, and Adam R. Brandt. Global carbon intensity of crude oil production. *Science*, 361(6405):851–853, 2018.

Mohammad S. Masnadi, Giacomo Benini, Hassan M. El-Houjeiri, Alice Milivinti, James E. Anderson, Timothy J. Wallington, Robert De Kleine, Valerio Dotti, Patrick Jochem, and Adam R. Brandt. Carbon implications of marginal oils from market-derived demand shocks. *Nature*, 599:80–84, 2021.

Charles F. Mason and Gavin Roberts. Price elasticity of supply and productivity: An analysis of natural gas wells in Wyoming. *The Energy Journal*, 39:79–100, 2018.

Leonardo Maugeri. Oil: The next revolution. *Belfer Center for Science and International Affairs, Harvard Kennedy School*, 2012.

Michael J. Mayerhofer, Ely Lolon, Jeffrey E. Youngblood, and John R. Heinze. What is stimulated reservoir volume? *Society of Petroleum Engineers*, 25:89–98, 2010.

James McCarthy and Ravi Singh. Impact of crude oil characteristics on refining processes. *Journal of Petroleum Science and Engineering*, 145:315–328, 2016.

Marjorie B McElroy. Additive general error models for production, cost, and derived demand or share systems. *Journal of Political Economy*, 95(4):737–757, 1987.

Christophe McGlade and Paul Ekins. Un-burnable oil: An examination of oil resource utilization in a decarbonized energy system. *Energy Policy*, 64:102–112, 2014.

Christophe McGlade and Paul Ekins. The geographical distribution of fossil fuels unused when limiting global warming to $2°C. Nature, 517(7533) : 187 - -190, 2015$.

Joseph W. McMillan. OPEC's influence on global oil prices: A quantitative analysis. *Energy Economics*, 30(5): 2560–2575, 2008.

Donella H. Meadows, Dennis L. Meadows, Jorgen Randers, and William W. Behrens III. *The Limits to Growth*. Universe Books, 1972.

Jean-François Mercure, Hector Pollitt, Alex Edward Bowman, Pierre Valin, Unnada Chewpreecha, Pablo Salas, J-F Vinuales, and Neil R. Edwards. Macroeconomic impact of stranded fossil fuel assets. *Nature Climate Change*, 8 (7):588–593, 2018.

Robert A. Meyers. *Handbook of Petroleum Refining Processes*. McGraw-Hill Professional, 2004.

Nikolaos T. Milonas and John Henker. Price uniformity in the global oil market. *Journal of Business*, 74:563–591, 2001.

Carl T. Montgomery and Michael B. Smith. Hydraulic fracturing: History of an enduring technology. *Journal of Petroleum Technology*, 62(12):26–32, 2010.

Geoff Morgan. Despite the oilsands' relentless cost cutting, investors are still wary of jumping back in. Financial Post, 2020. URL `https://financialpost.com`.

Gunnar Myhre, Drew Shindell, François-Marie Bréon, William Collins, Jan Fuglestvedt, Jianping Huang, Dorothy Koch, Jean-François Lamarque, David Lee, Bruce Mendoza, Teruyuki Nakajima, Alan Robock, Graeme Stephens, Toshihiko Takemura, and Hua Zhang. *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, 2013.

Shinichi Nakagawa and Holger Schielzeth. A general and simple method for obtaining r^2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution*, 4(2):133–142, 2013.

Tony N. Nasr and Olukayode R. Ayodele. Improving steam-assisted gravity drainage (SAGD) using solvents. *Society of Petroleum Engineers*, pages 1–10, 2002.

Natural Resources Canada (NRCan). Basin database. `https://www.nrcan.gc.ca/`, 2017.

National Energy Board (NEB). Estimated production of canadian crude oil and equivalent. `https://www.neb-one.gc.ca`, 2017.

Philip A. Neher. Optimal depletion with resource augmenting technical progress. *Econometrica*, 44:1035–1047, 1976.

John Ashworth Nelder and Robert WM Wedderburn. Generalized linear models. *Journal of the Royal Statistical Society: Series A (General)*, 135(3):370–384, 1972.

Jerzy Neyman and Elizabeth L Scott. Consistent estimates based on partially consistent observations. *Econometrica*, pages 1–32, 1948.

William D Nordhaus. The economics of an integrated world oil market. In *International Energy Workshop*, pages 17–19, 2009.

NPD. Norway petroleum directorate report, 2017.

Petroleum Services Association of Canada (PSA). The crude oils and their key characteristics. `https://www.psac.ca`, 2020.

Oil Climate Index (OCI). Oil Climate Index Plus (OCI+) Tool. `https://dxgordon.github.io/OCIPlus/`, 2020.

Steven Olley and Ariel Pakes. The dynamics of productivity in the telecommunications equipment industry. Working Paper w3977, National Bureau of Economic Research, 1992.

OPEC. *OPEC: 50 Years of Global Influence*. OPEC Secretariat, 2005.

OPEC. OPEC share of world crude oil reserves. `https://www.{O}{P}{E}{C}.org/{O}{P}{E}{C}_web/en/data_graphs/330.htm`, 2021a. Accessed: 2024-12-16.

OPEC. OPEC+ membership and collaboration. `https://www.{O}{P}{E}{C}.org/{O}{P}{E}{C}_web/en/press_room/5891.htm`, 2021b. Accessed: 2024-12-16.

Surinder Parkash. Refining processes handbook. *Chemical Engineering*, 2003.

PenWell Corporation. Worldwide oil field production survey. Technical Report, 2015.

Hashem M. Pesaran. An econometric analysis of exploration and extraction of oil in the UK continental shelf. *The Economic Journal*, 100(401):367–390, 1990.

Hashem M. Pesaran. A simple panel unit root test in the presence of cross-section dependence. *Journal of Applied Econometrics*, 22(2):265–312, 2007.

Robert S. Pindyck. The optimal exploration and production of nonrenewable resources. *Journal of Political Economy*, 86:841–861, 1978.

Grant Plant, Eric A. Kort, Adam R. Brandt, Yuanrong Chen, Grayson Fordice, Adrian M. Gorchov Negron, Stefan Schwietzke, Matthew Smith, and Daniel Zavala-Araiza. Inefficient and unlit natural gas flares both emit large quantities of methane. *Science*, 377:1566–1571, 2022.

Richard J. Plevin, Mark A. Delucchi, and Felix Creutzig. Using attributional life cycle assessment to estimate climate-change mitigation benefits misleads policy makers. *Journal of Industrial Ecology*, 18:73–83, 2014.

Micaela Ponce and Anne Neumann. Elasticities of supply for the US natural gas market. *SSRN Electronic Journal*, 2014.

Harold Ramcharran. Oil production responses to price changes: An empirical application of the competitive model to OPEC and non-OPEC countries. *Energy Economics*, 24(2):97–106, 2002.

Gavin Roberts, Edward Barbier, and Klaas van't Veld. Global emissions from crude oil: The effect of oil-deposit heterogeneity. *Energy Policy*, 132:654–664, 2019.

Rystad Energy. Rystad shale well database. `https://www.rystadenergy.com/energy-themes/oil--gas/shale/shale-well-cube/`, 2021. Accessed: 2021-07-29.

Reza Sadeghbeigi. *Fluid Catalytic Cracking Handbook: An Expert Guide to the Practical Operation, Design, and Optimization of FCC Units*. Elsevier, 2012.

Lutz Sager. The global consumer incidence of carbon pricing: Evidence from trade. *Energy Economics*, 127:107101, 2023.

Said E. Said and David A. Dickey. Testing for unit roots in autoregressive-moving average models of unknown order. *Biometrika*, 71(3):599–607, 1984.

Anthony Samuel. *The Seven Sisters: The Great Oil Companies and the World They Shaped*. Hodder and Stoughton, 1988.

Amjad Shah, Robert Fishwick, Joseph Wood, Gary Leeke, Sean Rigby, and Malcolm Greaves. A review of novel techniques for heavy oil and bitumen extraction and upgrading. *Energy & Environmental Science*, 3(6):700–714, 2010.

Gaurav Sharma. Are electric vehicles really about to plateau oil demand? `https://www.forbes.com/sites/gauravsharma/2019/11/25/are-electric-vehicles-really-about-to-plateau-oil-demand/`, 2019.

Alex Smith and Rachel Brown. Oil logistics: Pipelines, rail, and marine transport. *Journal of Energy Transport*, 29 (4):202–218, 2018.

Alex Smith and Rebecca Johnson. The distribution network for refined oil products: Pipelines, rail, and marine transport. *Journal of Energy Logistics*, 37(3):215–229, 2020.

James L. Smith. OPEC and the world oil market: The dynamics of price determination. *Energy Economics*, 5(2): 121–129, 1983.

John D. Smith and Emily R. Taylor. Oil as the backbone of the global energy system: Historical and modern perspectives. *Energy Policy*, 123:215–228, 2018.

John D. Smith and Emily R. Taylor. The keystone XL pipeline and its role in north american oil transport. *Energy Policy Review*, 34(2):67–78, 2021.

Robert M. Solow. The economics of exhaustible resources. *The Review of Economic Studies*, 41:1–14, 1974.

Robert M Solow and Frederic Y Wan. Extraction costs in the theory of exhaustible resources. *The Bell Journal of Economics*, pages 359–370, 1976.

James G. Speight. *The Chemistry and Technology of Petroleum*. CRC Press, 4th edition, 2004.

James G. Speight. *The Refining of Petroleum: A Reference Guide*. CRC Press, 2011.

Stanford Doerr School of Sustainability. Global carbon emissions from fossil fuels reached record high in 2023. `https://sustainability.stanford.edu/news/global-carbon-emissions-fossil-fuels-reached-record-high-2023`, 2023.

State of California Department of Conservation, Division of Oil, Gas, & Geothermal Resources. 2015 report of California oil and gas production statistics. `ftp://ftp.consrv.ca.gov/pub/oil/annual_reports/2015/PR03_2015.pdf`, 2015.

Statista. Natural gas production worldwide from 1998 to 2023. `https://www.statista.com/statistics/265344/total-global-natural-gas-production-since-1998/`, 2024.

P. Arvind Swamy. Efficient inference in a random coefficient regression model. *Econometrica*, pages 311–323, 1970.

SM Taheri and Gholamreza Hesamian. A generalization of the wilcoxon signed-rank test and its applications. *Statistical Papers*, 54(2):457–470, 2013.

Garth S. Treese and Daniel P. Treese. *Modern Petroleum Technology*. Wiley, 2015.

Trident Consulting. Hydrocarbon loss: Identifying opportunities. `https://www.trindent.com/hydrocarbon-loss-identifying-opportunities`, 2023.

Vivek Tripathi and Adam R. Brandt. Estimating decades-long trends in petroleum field energy return on investment (eroi) with an engineering-based model. *PLoS ONE*, 12:e0171083, 2017.

UK Government. Oil and gas: Field data. `https://www.gov.uk/guidance/oil-and-gas-uk-field-data`, 2017.

Kyriacos Vafi and Adam R. Brandt. GHGfrack: An open-source model for estimating greenhouse gas emissions from combustion of fuel during drilling and hydraulic fracturing. *Environmental Science & Technology*, 50:7913–7920, 2016.

Robert W. Veatch. An overview of hydraulic fracturing. *Journal of Petroleum Technology*, 35(4):677–687, 1983.

Philip K. Verleger. OPEC behavior and world oil prices: A study of the 1990s. *The Energy Journal*, 12(3):1–22, 1991.

Timothy J. Wallington, James E. Anderson, Robert D. De Kleine, Hyung Chul Kim, Heiko Maas, Adam R. Brandt, and Gregory A. Keoleian. When comparing alternative fuel-vehicle systems, life cycle assessment studies should consider trends in oil production. *Journal of Industrial Ecology*, 21:244–248, 2017.

Qinbin Wang, Yi Chen, Ronald L. Thompson, and Prabir K. Patra. Natural gas fugitive emissions rates constrained by global atmospheric methane and ethane. *Proceedings of the National Academy of Sciences*, 111(17):6237–6242, 2014.

A. Watkinson and J. W. Wilson. *Petroleum Refining Processes*. McGraw-Hill Education, 2012.

WBCSD. *The GHG Protocol Initiative: Guidelines for Quantifying Greenhouse Gas Emissions*. World Business Council for Sustainable Development, Geneva, Switzerland, 2005.

Karsten Webel. Greene, W. H., econometric analysis. *Statistical Papers*, 52(4):983–984, 2011.

J. Weissman and M. T. Klein. Catalytic cracking: A key process in petroleum refining. *Chemical Engineering Progress*, 81:37–47, 1985.

Harold F. Williamson and Arnold R. Daum. The first oil well: A study of the titusville discovery. *Pennsylvania History: A Journal of Mid-Atlantic Studies*, 26:2–20, 1959.

Alan C. Wills. The role of OPEC in the global oil market: Evidence from the 1990s. *Energy Policy*, 23(9):779–785, 1995.

Jeffrey M Wooldridge. *Econometric Analysis of Cross Section and Panel Data*. MIT press, 2010.

WRI. *The Greenhouse Gas Protocol: A Corporate Accounting and Reporting Standard*. World Resources Institute and World Business Council for Sustainable Development, Geneva, Switzerland, 2004.

Yi Yang and Reinout Heijungs. On the use of different models for consequential life cycle assessment. *International Journal of Life Cycle Assessment*, 23:751–758, 2018.

Daniel Yergin. *The Prize: The Epic Quest for Oil, Money, and Power*. Simon & Schuster, 1991.

Robert E. Yergin. The dynamics of oil prices and OPEC's market power. *Energy Journal*, 25(1):1–24, 2004.

Arnold Zellner. An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias. *Journal of the American statistical Association*, 57(298):348–368, 1962.

Martin B Zimmerman. Modeling depletion in a mineral industry: The case of coal. *The Bell Journal of Economics*, pages 41–65, 1977.