Ca' Foscari University of Venice

Department of Economics

Working Paper

Giacomo Benini
Adam Brandt
Valerio Dotti
Hassan El-Houjeiri

The Economic and Environmental Consequences of the Petroleum Industry Extensive Margin

ISSN: 1827-3580
No. 14/WP/2023

Working Papers
Department of Economics
Ca' Foscari University of Venice
No. 14/WP/2023
ISSN 1827-3580

The Economic and Environmental Consequences of the Petroleum Industry Extensive Margin

Giacomo Benini
Department of Business and Management Science, NHH

Adam Brandt
Department of Energy Resource Engineering, Stanford University

Valerio Dotti
Ca' Foscari University of Venice

Hassan El-Houjeiri
Technology Outlook, Strategy & Planning Dept., Saudi Aramco

Abstract
The recent diffusion of novel oil technologies has increased the variability of petroleum resources. Today, it is possible to mine oil sands, to extract liquids from tight rocks, and to produce high-viscosity oils. Merging accounting and environmental data, we quantify the upstream emissions of the least profitable oilfields. According to our estimates thirteen fields, responsible for the production of 0.72 million barrels per day, represent the 1% extensive margin of the industry. These formations are Heavy & Extra Heavy and Sands deposits. Their average upstream carbon intensity is 114.61 KgCO2e per barrel versus a global average of 54.35. Similar results are obtained widening the extensive margin to 2.5% and 5%. This finding suggests that a fall in the global oil demand of 1% can reduce upstream emissions by 24.95 MMtCO2e per year, the annual footprint of 5.3% of all the cars registered in the United States.

Keywords
Oil Economics, Shadow Prices, Empirical Analysis of Firm Behavior, Panel Data, Co-integration, Endogeneity, Linear Mixed Models

JEL Codes
L23, D22, C23, C14

Address for correspondence:
Valerio Dotti
Department of Economics
Ca' Foscari University of Venice
Cannaregio 873, Fondamenta S.Giobbe
30121 Venezia - Italy
e-mail: valerio.dotti@unive.it

The Economic and Environmental Consequences of the Petroleum Industry Extensive Margin

Giacomo Benini[1], Adam Brandt[2], Valerio Dotti[3], and Hassan El-Houjeiri[4]

[1]Department of Business and Management Science, NHH
[2]Department of Energy Resource Engineering, Stanford University
[3]Department of Economics, Ca' Foscari University of Venice
[4]Technology Outlook, Strategy & Planning Dept., Saudi Aramco

May 22, 2023

Abstract

The recent diffusion of novel oil technologies has increased the variability of petroleum resources. Today, it is possible to mine oil sands, to extract liquids from tight rocks, and to produce high-viscosity oils. Merging accounting and environmental data, we quantify the upstream emissions of the least profitable oilfields. According to our estimates thirteen fields, responsible for the production of 0.72 million barrels per day, represent the 1% extensive margin of the industry. These formations are Heavy & Extra Heavy and Sands deposits. Their average upstream carbon intensity is 114.61 $KgCO_2e$ per barrel versus a global average of 54.35. Similar results are obtained widening the extensive margin to 2.5% and 5%. This finding suggests that a fall in the global oil demand of 1% can reduce upstream emissions by 24.95 $MMtCO_2e$ per year, the annual footprint of 5.3% of all the cars registered in the United States.

Keywords: Oil Economics, Shadow Prices, Empirical Analysis of Firm Behavior, Panel Data, Co-integration, Endogeneity, Linear Mixed Models.
JEL Classification: L23, D22, C23, C14.
Corresponding Author: valerio.dotti@unive.it

1 Introduction

The oil market connects tens of thousands of oilfields to billions of consumers via two million kilometers of pipelines and five hundred millions dead-weight tons of merchant shipping (Cruz & Taylor, 2013). This global web has progressively transformed many regional markets into a worldwide pool of crude where riskless price arbitrages are virtually unattainable (Adelman, 1984; Nordhaus, 2009).

The combined efforts of the logistic and of the financial sector have removed resale opportunities, standardizing the global oil demand (Milonas & Henker, 2001). Conversely, a set of technological advancements has diversified the global oil supply widening the production possibilities of the petroleum sector (Maugeri, 2012). Today, it is possible to extract liquids trapped tightly in impermeable shale rocks, to mine heavy-oil-bearing sands, to explore the far Arctic, and to access deep-water deposits located in the oceans depths (Gordon, Brandt, Bergerson, & Koomey, 2015). The increased variety of the industry supply side has two consequences. First, day-to-day expenses diverged making some oilfields marginally more profitable than others (Masnadi et al., 2021). Second, the upstream carbon footprint is now heterogeneous (Masnadi et al., 2018). The aim of the paper is to bundle these two quantities and study how they co-move to answer the following question: "How does a small shock on the global oil demand affect the carbon emissions of the global oil supply chain?"

We construct an extraction-exploration model where oilfields are risk-neutral firms facing heterogeneous revenues and costs (Pesaran, 1990). We quantify the firm's markup 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 microfounded field-level cost function, which disentangles the impact of natural pressure from the one of other factors of production. To achieve this goal, we build on the formulation proposed by Anderson, Kellogg, and Salant (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 (WoodMackenzie, 2020) to estimate the cost function.

Subtracting the marginal extraction costs from the selling price, 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. The least profitable segments of the merit base curve identify the extensive margin of the industry. According to our estimates, the least profitable 1% of the global production is made of Venezuelan Heavy & Extra Heavy deposits and of Canadian oil Sands. These formation are roughly responsible for the production of one million barrels per day.

Once identified the oil industry extensive margin, we analyze 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 to it 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 US). 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 behavior of a single well and ends with the impact of OPEC on the global average oil price.

2 The Oil Shadow Price

We study a general equilibrium economy featuring four types of players: oil extraction firms, refineries, consumers, and the government. In this section we focus solely on the key actors, the oil firms. A detailed description of the other parts of the model and of its competitive equilibrium solution, which are used solely to derive some of the rebustness results presented in section 4 and 5, is provided in the Appendix.

We assume that each oilfield i is managed by a risk-neutral firm k, which owns $n(k)$ fields. In this section, we assume that oil firms exert no market power, but we relax this assumption in section 5. The firm decides in period t its production and investment plan for all periods $t+s$ with $s = 0, 1, 2, ...$ Its intra-temporal profits, in each period $t+s$,

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

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 determined in a competitive equilibrium and is a function of the chemical characteristics of the crude produced by field i, as illustrated in section 3.1. The specific functional form for 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 3.2.1. 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,

$$L^i_{t+s} \leq L^i_{t+s-1} + D^i_{t+s}(W^i_{t+s}, L^i_{t+s-1}, \xi^i_{t+s}) \ , \tag{2}$$

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 $D^i_{t+s}(.)$[1], which are function

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

of the exploration expenditures W_{t+s}^i, of the quantity of discoveries made in the past L_{t+s-1}^i, and of an idiosyncratic shock ξ_{t+s}^i.

The second constraint ensures that the cumulative depletion exerted until $t+s-1$, denoted by M_{t+s-1}^i, plus the production at time $t+s$, equals or exceeds[2] the cumulative depletion at time $t+s$,

$$M_{t+s}^i \geq M_{t+s-1}^i + Q_{t+s}^i \; . \tag{3}$$

Each firm in period t decides the volumes of production, $Q_t^i, Q_{t+1}^i, ...$ and the rates of investment in exploratory $W_t^i, W_{t+1}^i, ...$ by maximizing the expected discounted future stream of profits. The decision is conditioned by the available information set Ω_{t+s-1}^k which includes previous prices, quantities, and shocks,

$$\Omega_{t+s-1}^k = \left\{ \left[P_s^i\right]_{s=0}^{t-1}, \; \left[Q_s^i, W_s^i, M_s^i, L_s^i\right]_{s=0}^{t-1}, \; \left[\epsilon_s^i, \xi_s^i\right]_{s=0}^{t-1} \right\}_{i=1}^{n(k)} .$$

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

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

where $0 \leq \kappa < 1$ is the inter-temporal discount factor. The shadow price of discovered oil in field i in period t,

$$\mathbb{E}_{t-1}[\lambda_t^i | \Omega_{t-1}^i] = \mathbb{E}_{t-1}[P_t^i | \Omega_{t-1}^i] - \mathbb{E}_{t-1} \left[\frac{\partial C_t^i(.)}{\partial Q_t^i} \middle| \Omega_{t-1}^i \right] , \tag{4}$$

is obtained by solving the firm's optimality condition and is equal to the difference between the price at which the field sells its output and its marginal production costs. 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.

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

3 Empirical Analysis

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, Manera, & Giovannini, 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,

$$\mathbb{E}_{t-1}[P_t^i|\Omega_{t-1}^i] = \bar{P}_t + \beta_{1t}(API^i - \overline{API}_t) + \beta_{2t}(S^i - \bar{S}_t) \ , \tag{5}$$

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 $(S^i - \bar{S}_t)$ Details on the derivation are provided in the Appendix. We take as reference price the average price at which United States refineries import different streams of crude (EIA, 2020b) 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[3] and the chemical

[3]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

characteristics of twenty-three oil classes over the time interval 1978-2018 (EIA, 2019; Management & BV, 2019). Table 1 provides the summary statistics of the future prices, the gravity, and the sulfur content of every class. Figure (1) shows how the future prices (FP) respond to the interaction between demand and supply. To the contrary, Figure (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 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 (2019); Management and BV (2019).

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 ($\Omega^i_{t-1} = \Omega^{pub}_{t-1}$), we derive a formula

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.

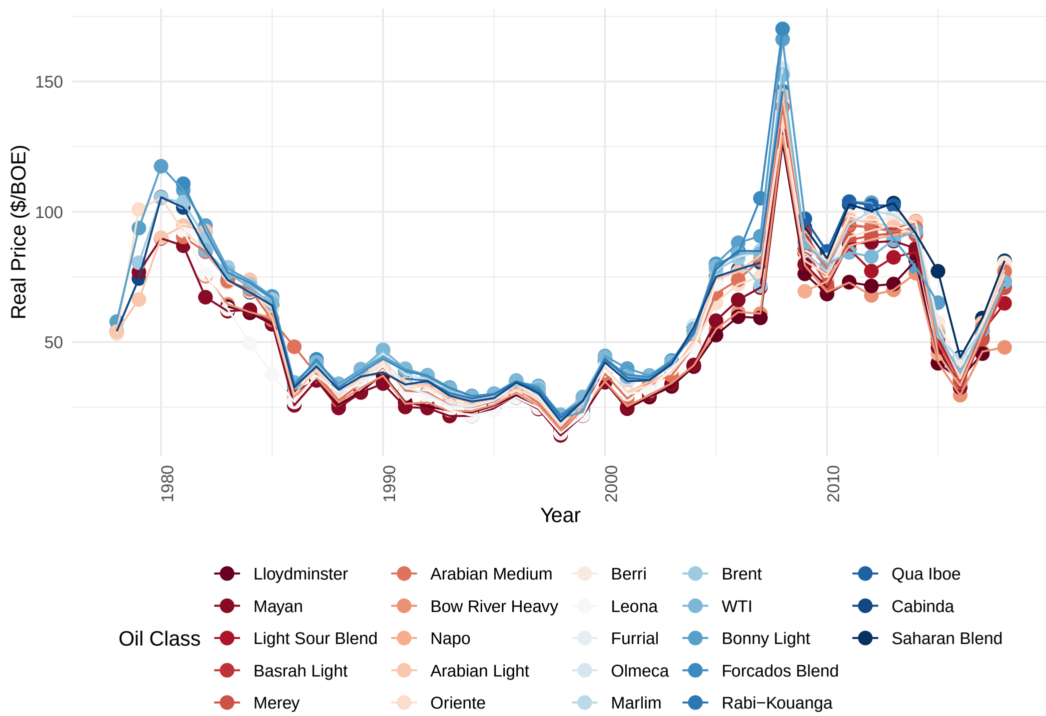

Figure 1: Future prices of twenty-three oil class over the time interval 1978-2018. Colors reflect the sulfur content, where dark red represents high content and dark blue low percentages.

for the future price of oil of class z in period t,

$$FP_t(z) = \mathbb{E}_{t-1}\left[\sum_{i\in z} w_t^i P_t^i | \Omega_{t-1}^{pub}\right] = \mathbb{E}_{t-1}\left[\sum_{i\in z} w_t^i P_t^i | \Omega_{t-1}^{i}\right] , \tag{6}$$

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) for the expected price of oil from field i into the formula (6). As a result, we can glue the twenty-three future prices time series into a panel structure and run the regression

$$FP_t(z) = \bar{P}_t + \beta_{1t}(API(z) - \overline{API}_t) + \beta_{2t}(S(z) - \bar{S}_t) + \varsigma_t \ , \tag{7}$$

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 (7) and use its structural coefficient to reverse-engineer $\mathbb{E}_{t-1}[P_t^i | \Omega_{t-1}^i]$, we check if the future and reference prices are stationary. Not all

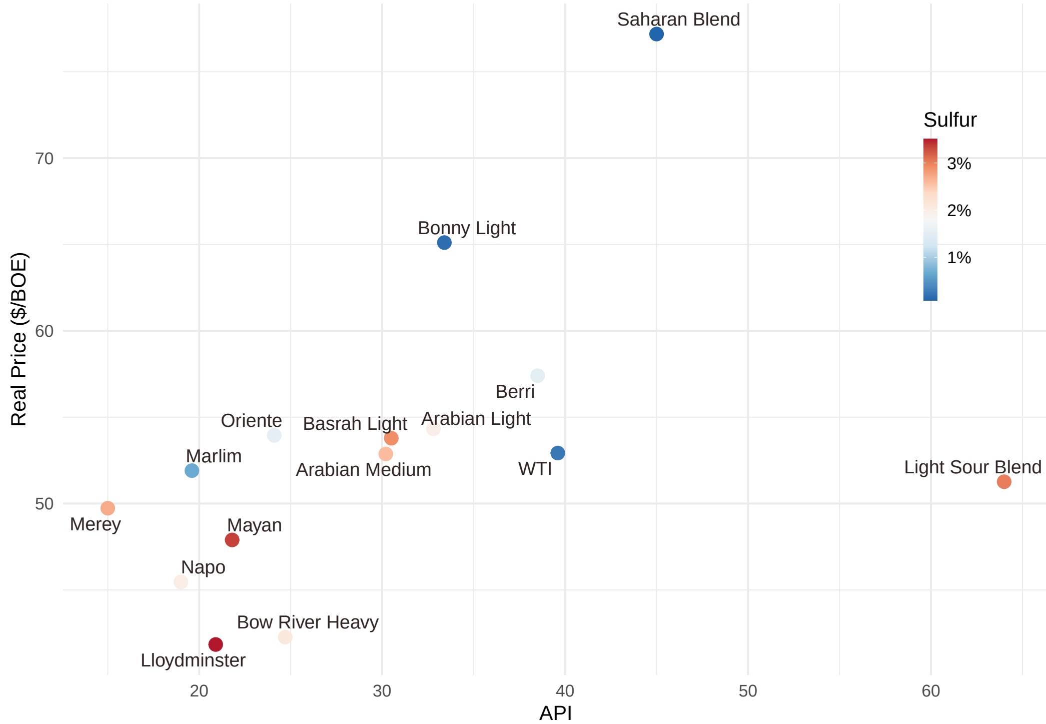

Figure 2: Spot prices of fourteen oil classes in 2015. The position on the horizontal axis reflects lightness. Colors reflect the sulfur content, where dark red represents high content and dark blue low percentages.

oil classes have time series long enough to perform unit root tests. However, a subgroup of ten does[4]. Following Pesaran (2007) and Costantini and Lupi (2013), we perform four types of Panel Covariate-Augmented Dickey-Fuller Test[5]: 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 $\sim$0.88 to $\sim$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 = $\sim$0.30) using an Augmented Dickey-Fuller test (Said & Dickey, 1984; Banerjee, Dolado, Galbraith, Hendry, et al., 1993). The lack of stationarity on both sides of equation (7) allows us to move

[4]The subgroup is Cabinda, Bow River Heavy, Lloydminster, Oriente, Mayan, Bonny Light, Forcados Blend, Arabian Light, Arabian Medium, and WTI

[5]The tests are performed using the R-package `CADFtest` running the command `pCADFtest` (Lupi, 2010).

$\bar{P}_t$ to the lefthandside of the future price equation,

$$FP_t(z) - \bar{P}_t = \beta_{1t}(API(z) - \overline{API}_t) + \beta_{2t}(S(z) - \bar{S}_t) + \varsigma_t \ , \tag{8}$$

and run a regression in which the dependent variable, being the difference between two variables cointegrated of order one, is stationary[6] (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 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 (8). 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 2 report the results obtained using a Random Coefficient Model (RCM) where β_{1t} and β_{2t} are normally distributed and vary across time[7] (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 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

[6]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.

[7]The RCM is estimated using the R-package `lme4` (Bates, Mächler, Bolker, & Walker, 2015) and the conditional modes are extracted with the command `ranef`.

R^2 to 0.40 (Nakagawa & 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.

Table 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

Using the estimates portrayed in Table 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

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

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

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

Notice that under the assumption stated in this section $\widehat{\mathbb{E}_{t-1}[P_t^i|\Omega_{t-1}^i]}$ 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}.

3.2 Marginal Costs

Several business intelligence companies collect data about oilfields revenues and costs. Among others, WoodMac classifies capital and operational expenditures

Table 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

of (parent) oilfields into twenty-three categories. Table 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",

$$C_t^i = \text{Abandonment Costs}_t^i + \text{Capital Receipts}_t^i + ... + \text{Terminal}_t^i \ , \tag{9}$$

and the last two to "find new oil",

$$W_t^i = \text{Development Drilling}_t^i + \text{Exploration and Appraisal}_t^i \ . \tag{10}$$

Table 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: WoodMackenzie (2020).

Figure 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

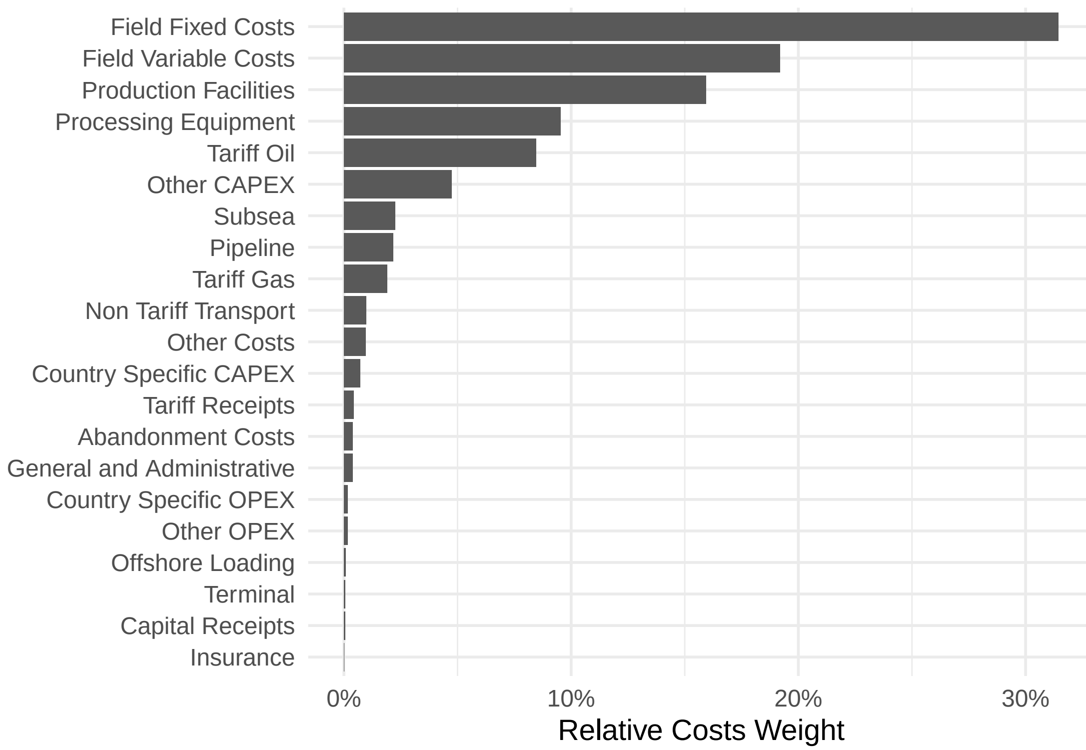

Figure 3: Relative weight of the different cost categories.

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 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 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 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 pressure in the reservoir by drilling a second type of well called injection wells. The later pump water, steam, or natural gas to keep the production process fluid, increase the reservoir

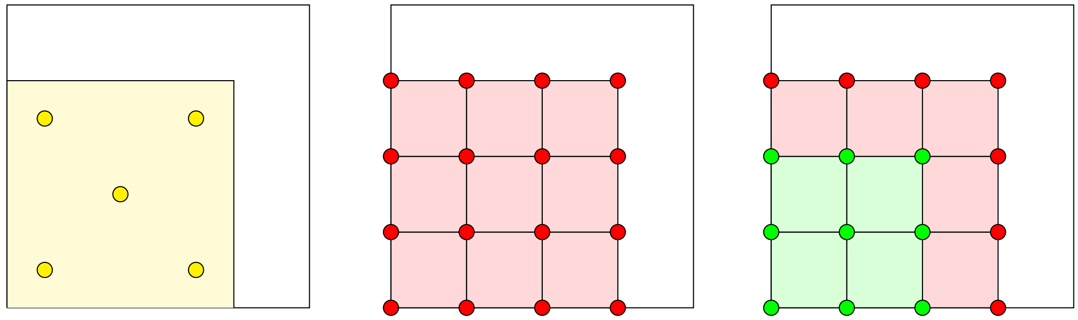

Figure 4: Exploration Wells ●, Explored Area ▬, Untapped Wells ●, Potentially Producing Area ▬, Tapped Wells ●, Producing Area ▬.

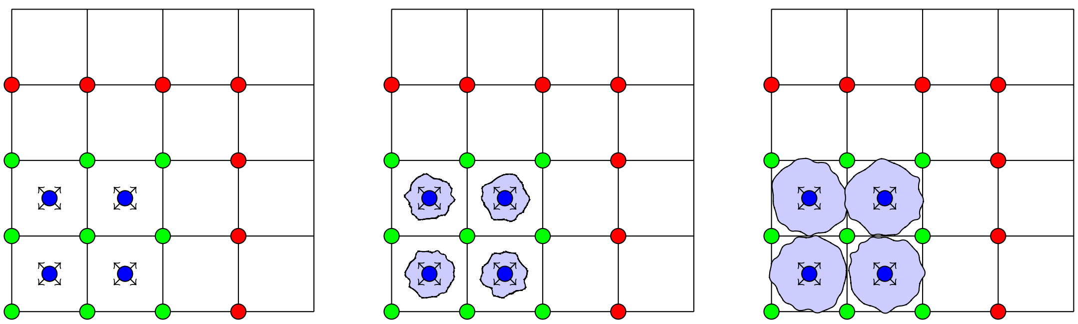

Figure 5: Untapped Wells ●, Tapped Wells ●, Injection Wells ●, Injected Water/Steam ▬

temperature, and its pressure, see Figure 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 & 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[8]. 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. 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[9]. 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: 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	✓	✗

[8]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.

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

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 process like steam, water, electricity, and labor. 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 (Livernois & Uhler, 1987; Pesaran, 1990; Favero, 1992; Masnadi et al., 2021). This last quantity should captures the entire extraction process becoming the proxy for the logical sequence: more reserves → more pressure → less inputs. In this framework discovering one barrel should compensate, in terms of marginal costs, the extraction of one. As a result, the oil sector becomes no different than any other exhaustible resources industry (Solow & Wan, 1976), like coal or copper mining (Zimmerman, 1977; Aguirregabiria & Luengo, 2016), where, once controlled for the ore grade, only the size of the mine impacts marginal extraction costs.

3.2.1 Extraction Costs

We assume that each firm i faces six types of costs, each corresponding to one or more cost classes listed in Table 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[10],

$$\begin{aligned} C_t^i = &\underbrace{ExtrCosts_t^i + TransCosts_t^i + Taxes_t^i}_{\text{Variable Costs}} + \\ &+ \underbrace{DisrCosts_t^i + MaintCosts_t^i + OtherCosts_t^i}_{\text{Fixed Costs}} \ . \end{aligned} \tag{11}$$

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 well's 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,

$$ExtrCosts_t^i = \sum_{j=1}^{n} \int_T p^j \ PInputs_t^j(\boldsymbol{\eta}) \, f^i\left(\boldsymbol{\eta} \mid h_t^i\right) d^3\boldsymbol{\eta} \ . \tag{12}$$

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

$$WellCapacity_t^i\left(\boldsymbol{\eta}, h_t^i\right) = F_t^i\left(\left\{PInputs_t^{i,j}(\boldsymbol{\eta})^*\right\}_{j=1,\boldsymbol{\eta}\in T}^{n}, \boldsymbol{\eta} \middle| \, h_t^i\right) \ , \tag{13}$$

[10]Note that the inter-temporal profit maximization problem described in section 2 and the within-period cost minimization problem outlined in section 3.2.1 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 (13) 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 (13) 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 (12). 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[11] 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, as illustrated by equation (4).

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,

$$FieldCapacity_t^i = \left\{ \int_T \left[WellCapacity_t^i\left(\boldsymbol{\eta}, h_t^i\right)\right]^\delta f^i\left(\boldsymbol{\eta} \mid h_t\right) d^3\boldsymbol{\eta} \right\}^{\frac{1}{\delta}}, \qquad (14)$$

[11](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.

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[12]. 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:

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

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 2 is well-behaved.[13] Using this setup, we show that at the optimal bundle of productive inputs the formula for extraction costs writes:

$$ExtrCosts_t^i = \left[\tilde{\theta}_2 + \theta_3^{Geo}\frac{L_{t-1}^i}{R^i} + \theta_3^{Geo}\frac{M_{t-1}^i}{R^i}\right] Q_t^{i\ 2} \ , \tag{15}$$

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

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 modeled as a linear-quadratic function of oil output, augmented by firm- and time-specific factors and a stochastic component, see the Appendix for a detailed description. The resulting cost function,

$$C_t^i = \theta_1 Q_t^i + \left(\theta_2 + \theta_3^{Geo}\frac{L_{t-1}^i}{R^i} + \theta_4^{Geo}\frac{M_{t-1}^i}{R^i}\right) Q_t^{i\ 2} + \theta_5^{Geo}\left(\frac{M_{t-1}^i}{R^i}\right)^2 + \epsilon_t^i \ , \tag{16}$$

disentangles the impact of production choices and average reservoir pressure on marginal production costs for different types of geological formations. More precisely, in equation (16) 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 (9), 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[14]. L_{t-1}^i/R^i and

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

[13]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.

[14]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.

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 (16) 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.

3.2.2 Estimation

The estimation of equation (16) faces three main econometric issues. First, the empirical probability density function of C_t^i is virtually always positive[15] 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 & Andrews, 1944; Wooldridge, 2010). Estimating (16) 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 & 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 & Wedderburn, 1972; Liang & Zeger, 1986). Furthermore, if we run on C_t^i the four Panel Covariate-Augmented Dickey-Fuller Tests presented in section 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)

[15] 43 out of 28,924 observations present negative costs (0,14% of the sample). This are mostly North American fields, which do to fiscal reasons had more rebates than expenditures during the first or second year of production, such that $Taxes_t^i < 0$.

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 = $\sim$0.06) and/or the development and depletion (p-value = $\sim$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 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, 2020), 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 6 summarizes the costs and the volumes of production for the different formations.

Table 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:*WoodMackenzie (2020).

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 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 7 assumes that discoveries and depletion play no role in 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 (16), 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

Table 7: Marginal Costs Regressions

	Geo Class	*Dependent variable: $C_t^i - C_{t-1}^i$ (MM \$)*			
		(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:
$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

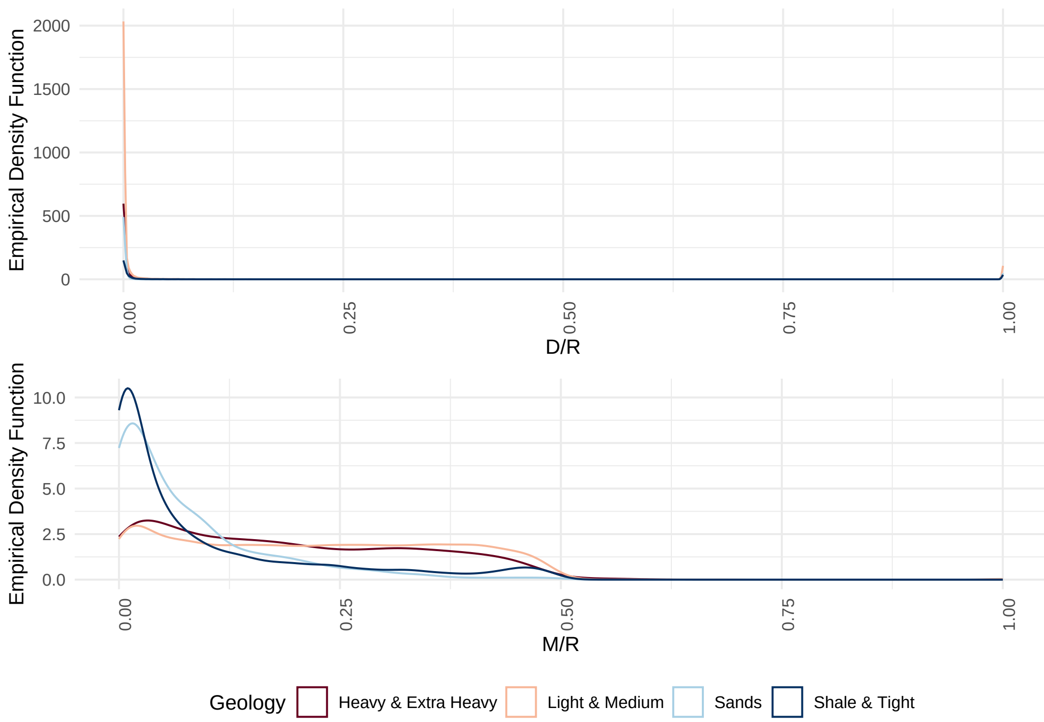

Figure 6: Empirical Density of the Discovery rate (D_t^i/R^i) and the Cumulative Depletion Rate (M_t^i/R^i).

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,

$$\mathbb{E}_{t-1}\left[\widehat{\frac{\partial C_t^i(.)}{\partial Q_t^i}}\middle|\Omega_{t-1}^i\right] = \underset{(0.15)}{4.98} + 2\left(\underset{(0.02)}{0.08} + \underset{(97.33)}{\hat{\theta}_3^{Geo}} \frac{L_{t-1}^i}{R^i} + \underset{(64.04)}{\hat{\theta}_4^{Geo}} \frac{M_{t-1}^i}{R^i}\right) Q_t^i \,, \tag{17}$$

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.

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[16]. 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[17]. 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 (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[18].

Upstream We analyze 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

[16]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).

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

[18]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 analyze 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.

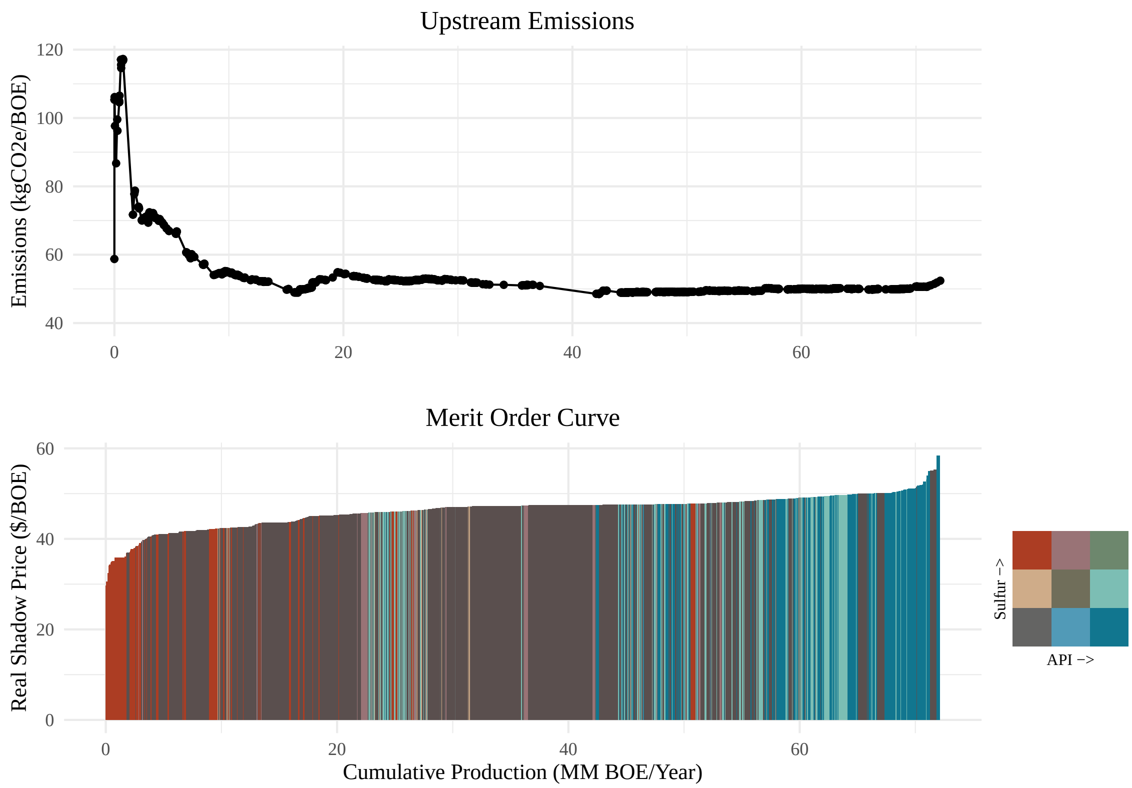

Figure 7: 2015 Global Merit Base Curve coupled with Upstream Carbon Intensity. Colors reflect the API gravity and the sulfur content, where dark red represents high sulfur content and low gravity and dark blue represents low sulfur content and high gravity.

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[19].

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

[19]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, 2018). The number of privately-owned automobiles for private and commercial use in the US in 2021 was equal to 101,601,344 (U.S. Department of Transportation, 2022).

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 (7). Table 8 summarizes all the results.

Table 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. Table 8 summarizes all the results. 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.

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 subunits 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,

$$\frac{\widehat{\mathrm{KgCO_2e}}^{r}}{\mathrm{BOE}} = \underset{(2.37)}{62.03} - \underset{(0.06)}{0.62} API^r - \underset{(0.64)}{0.95} S^r \ , \tag{18}$$

on the gravity and sulfur content of the processed crude using data from 343 refineries $r = 1, 2, ..., 343$, located in 83 countries[20] 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 & 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.

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 (18) 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%.

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

The pre-shock global pool has an 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 $KgCO_2e$/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 $KgCO_2e$/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 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 $MMtCO_2e$ / 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.

Table 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

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 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, 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, 2020) 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 (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 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 (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 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 Robustness to Non-Competitive Behavior

In the previous sections, we assume that each field is managed by a risk-neutral and price-taker firm. However, oilfields are owned by national and international oil companies, which may exert market power (Golombek, Irrazabal, & Ma, 2018; Asker, Collard-Wexler, & De Loecker, 2019). We introduce non-competitive behavior in the model assuming that the international oil companies play a game in quantity, while the national oil companies member of OPEC collude in the form of a cartel[21] In this framework, K oil firms (or groups of colluding firms) compete à la Cournot. Like in section 2, each firm $k = 1, 2, ..., K$ controls $n(k)$ oilfields and maximizes the present discounted value of the sum of present and future profits. Firm k decides in period t its production and investment plan for all periods $t + s$ with $s = 0, 1, 2, ...$ such that its intra-temporal profits equal

$$\Pi^k_{t+s} = \sum_{i=1}^{n(k)} \Pi^i_{t+s} = P^i_{t+s} Q^i_{t+s} - C^i_{t+s}(.) - W^i_{t+s} \ , \tag{19}$$

where all variables are defined as in equation (1). Each firm k anticipates the effect of its production choices on the equilibrium prices. In other word, we model the oil extraction market as a Cournot oligopoly where the equilibrium prices are identified by equations (5) and (7).

To obtain a tractable formula for the shadow price, we exploit the demand side of the economy (i.e., the behavior of oil refineries). Oil firms internalize the market-

[21]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.

clearing condition so that in each period the equilibrium price must be such that the demand of oil from field i equals its supply,

$$Q^i_{t+s} = AD^i_{t+s} \quad \forall i = 1, 2, 3, ...,$$

where AD^i_{t+s} 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 $t+s$,

$$\frac{\partial}{\partial Q^i_{t+s}}\left\{\mathbb{E}\left[\sum_{j=1}^{n(k)} P^j_{t+s}Q^j_{t+s}|\Omega^k_{t-1}\right]\right\} = \mathbb{E}\left[P^i_{t+s} + \sum_{j=1}^{n(k)} \frac{\partial P^j_{t+s}}{\partial AD^i_{t+s}}Q^j_{t+s}|\Omega^k_{t+s-1}\right], \qquad (20)$$

equals $(1 + MS^k_{t+s}/EL_P)\mathbb{E}_{t-1}[P^i_t|\Omega^k_{t-1}]$, where $MS^k_t = \sum_{j=1}^{n(k)} P^j_t Q^j_t / \sum_{i=1}^{n} P^i_t Q^i_t$ 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,

$$\mathbb{E}_{t-1}[\lambda^i_t|\Omega^k_{t-1}] = \left(1 + \frac{MS^k_t}{EL_P}\right)\mathbb{E}_{t-1}[P^i_t|\Omega^k_{t-1}] - \mathbb{E}_{t-1}\left[\frac{\partial C^i_t(.)}{\partial Q^i_t}\middle|\Omega^k_{t-1}\right], \qquad (21)$$

equals the perfect competition one only re-scaled by the market power correction term identified by 20. 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.

Upstream Using this alternative formula for the shadow price, we replicate the exercise proposed in section 4. Rank-ordering the new shadow prices from the lowest to the higher, we obtain the merit order curve resulting from a world in which every company non member of OPEC competes in quantities, while all the companies owned by governments member of OPEC behave as a perfect cartel, which maximizes the inter-temporal aggregate profits of the member countries. As a result, the entire OPEC production is now shifted to the right of the cumulative production, see the jump around the 20 million barrels shown in Figure (8). This finding becomes numerically more important as the price elasticity becomes smaller. However, within reasonable intervals, for example $EL_P \in (-0.25, -0.10)$ (Kilian, 2022), all the shadow prices remain positive suggesting that our estimates

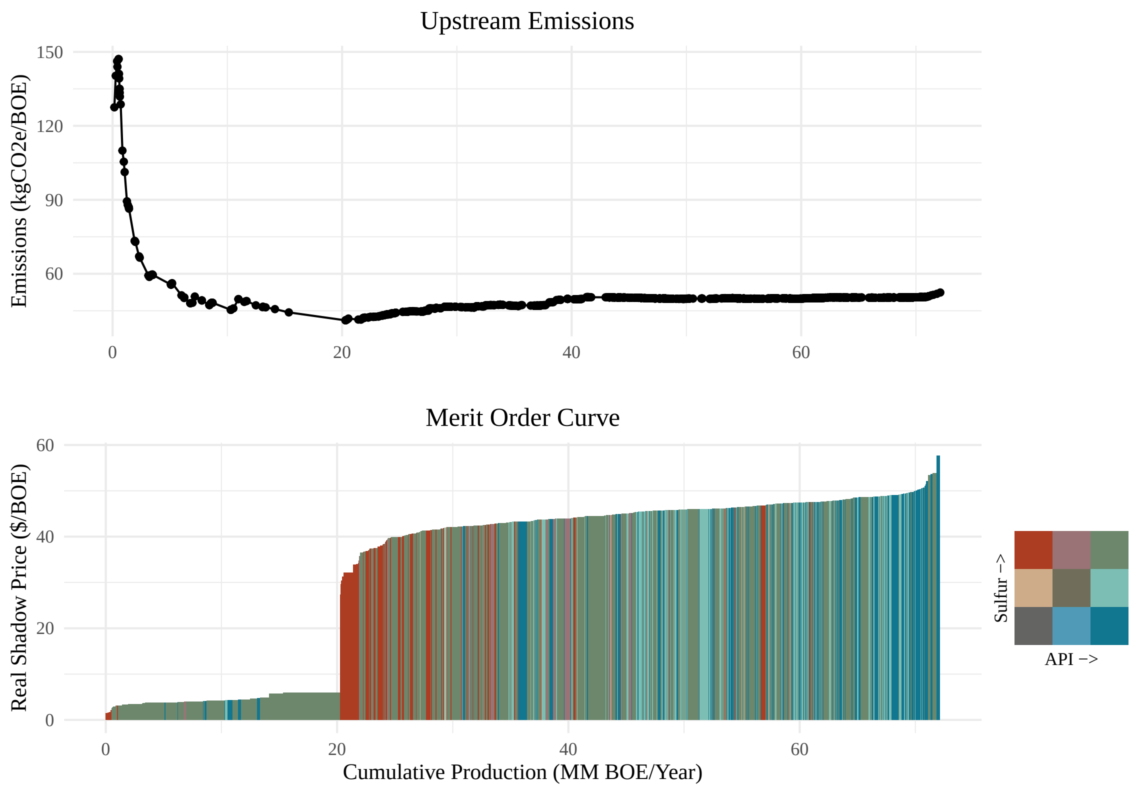

Figure 8: 2015 Global Merit Base Curve adjusted by the OPEC behavior coupled with Upstream Carbon Intensity. Colors reflect the API gravity and the sulfur content, where dark red represents high sulfur content and low gravity and dark blue represents low sulfur content and high gravity.

are in line with the one obtained using aggregate demand data[22]. In this new framework, the least profitable 1% of the global production is made out of six Heavy & Extra Heavy and seven Light & Medium formations. The volume-weighted average carbon intensity of this fraction of the global oil supply is 128.72 $KgCO_2e$. Like in the case of perfect competition the extensive margin is significantly more carbon intensive than the remaining production. However, its composition has changed. In perfect competition, Venezuela Heavy & Extra Heavy and Canadian Sands are the least competitive part of the global supply. In imperfect competition, Canadian Sands gain some competitiveness since they are operated by relatively small companies, which exert a small amount of market power. As a result, only the Heavy & Extra Heavy formations located in Venezuela determine the disproportionate carbon intensity of the extensive margin, since the rest is made of Light &

[22]If we are willing to assume that OPEC operates as an imperfect cartel, this finding remains true even for values of the demand elasticity closer to zero.

Medium formations located in Kuwait (one), Iran (three), and Iraq (three), whose carbon intensity is slightly above the global average. This description illustrates how the results in this section depend on the status of Venezuela as OPEC member. However, we are confident that they should be robust to endogenous changes in the structure of OPEC that may occur in response to shocks on the oil market conditions. This statement relies on the fact that Venezuela is a leading member of OPEC in terms of production volumes and available reserves and as such, its possible deviations from collusive behavior are expected to trigger a robust reaction by the other two key members, Saudi Arabia and Iran, on the ground of both theoretical considerations and historical records (Ghoddusi, Nili, & Rastad, 2017).

We obtain similar findings in the 2.5% and 5% scenarios, since none of the unconventional North American formations enters the extensive margin, which is exclusively made of OPEC members, see Table (12).

Table 12: 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%	128.72	0.72	33.17
2.5%	86.40	1.80	45.11
5%	59.53	3.60	76.18

While not shown in the paper, the mid- and downstream effects are virtually unchanged from the perfect competition case.

6 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.

References

Abella, J. P., & Bergerson, J. A. (2012). 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.

Adelman, M. A. (1984). International oil agreements. *The Energy Journal*, *5*(3), 1–9.

Aguirregabiria, V., & Luengo, A. (2016). A microeconometric dynamic structural model of copper mining decisions. *Unpublished manuscript, Department of Economics, University of Toronto*.

Anderson, S. T., Kellogg, R., & Salant, S. W. (2018). Hotelling under pressure. *Journal of Political Economy*, *126*(3), 984-1026. Retrieved from `https://doi.org/10.1086/697203` doi: 10.1086/697203

Asker, J., Collard-Wexler, A., & De Loecker, J. (2019). (mis)allocation, market power, and global oil extraction. *American Economic Review*, *109*(4), 1568–1615.

Bai, J. (2009). Panel data models with interactive fixed effects. *Econometrica*, *77*(4), 1229–1279.

Banerjee, A., Dolado, J. J., Galbraith, J. W., Hendry, D., et al. (1993). Co-integration, error correction, and the econometric analysis of non-stationary data. *OUP Catalogue*.

Bates, D. (2005). Fitting linear mixed models in r. *R news*, *5*(1), 27–30.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. doi: 10.18637/jss.v067.i01

BP. (2022). Statistical review of world energy.

Cooney, G., Jamieson, M., Marriott, J., Bergerson, J., Brandt, A., & Skone, T. J. (2017). 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.

Costantini, M., & Lupi, C. (2013). A simple panel-cadf test for unit roots. *Oxford Bulletin of Economics and Statistics*, *75*(2), 276–296.

Cruz, J. M., & Taylor, M. S. (2013). *A spatial approach to energy economics* (Tech. Rep.). National Bureau of Economic Research.

De Boeck, P., Bakker, M., Zwitser, R., Nivard, M., Hofman, A., Tuerlinckx, F., ... others (2011). The estimation of item response models with the lmer function from the lme4 package in r. *Journal of Statistical Software*, *39*(12), 1–28.

EIA. (2019). Eia. *EIA OPEN DATA*.

EIA. (2020a). Average api and sulphur content. $https://www.eia.gov/dnav/pet/pet_pnp_crq_dcu_nus_a.htm$.

EIA. (2020b). Price us refiners' acquisition costs for imported oils. *https://www.eia.gov/totalenergy/data/browser/csv.php?tbl=T09.01*.

EPA. (2018). Greenhouse gas emissions from a typical passenger vehicle. *Office of Transportation and Air Quality EPA-420-F-18-008*.

EPA. (2020). Greenhouse gas inventory guidance direct emissions from stationary combustion sources.

EPA. (2023). Inventory of u.s. greenhouse gas emissions and sinks by state. *https://www.epa.gov/ghgemissions/state-ghg-emissions-and-removals*.

Fattouh, B. (2010). The dynamics of crude oil price differentials. *Energy Economics*, *32*(2), 334–342.

Favero, C. A. (1992). Taxation and the optimization of oil exploration and production: the uk continental shelf. *Oxford Economic Papers*, *44*(2), 187–208.

Ghoddusi, H., Nili, M., & Rastad, M. (2017). On quota violations of opec members. *SSRN Research Papers Series, Available at SSRN: https://ssrn.com/abstract=3050369 or or http://dx.doi.org/10.2139/ssrn.3050369*.

Golombek, R., Irrazabal, A. A., & Ma, L. (2018). Opec's market power: An empirical dominant firm model for the oil market. *Energy Economics*, *70*, 98–115.

Gordon, D., Brandt, A. R., Bergerson, J., & Koomey, J. (2015). *Know your oil: Creating a global oil-climate index.* Carnegie Endowment for International Peace Washington, DC.

Hamilton, J. D. (2020). *Time series analysis.* Princeton university press.

Jing, L., El-Houjeiri, H. M., Monfort, J.-C., Brandt, A. R., Masnadi, M. S., Gordon, D., & Bergerson, J. A. (2020). Carbon intensity of global crude oil refining and mitigation potential. *Nature Climate Change*, 1–7.

Kawaguchi, K. (2020). *Topics in empirical industrial organization.* Retrieved from `https://kohei-kawaguchi.github.io/EmpiricalIO/`

Kilian, L. (2022). Understanding the estimation of oil demand and oil supply elasticities. *Energy Economics*, *107*, 105844.

Lanza, A., Manera, M., & Giovannini, M. (2005). Modeling and forecasting cointegrated relationships among heavy oil and product prices. *Energy Economics*, *27*(6), 831–848.

Liang, K.-Y., & Zeger, S. L. (1986). Longitudinal data analysis using generalized linear models. *Biometrika*, *73*(1), 13–22.

Livernois, J. R., & Uhler, R. S. (1987). Extraction costs and the economics of nonrenewable resources. *Journal of Political Economy*, *95*(1), 195–203.

Lupi, C. (2010). Unit root cadf testing with r. *Journal of Statistical Software*, *32*, 1–19.

Management, P., & BV, S. (2019). The crude oils and their key characteristics. *https://www.psa-bv.nl/files/CrudeOils.pdf*.

Marschak, J., & Andrews, W. H. (1944). Random simultaneous equations and the theory of production. *Econometrica, Journal of the Econometric Society*, 143–205.

Masnadi, M. S., Benini, G., El-Houjeiri, H. M., Milivinti, A., Anderson, J. E., Wallington, T. J., ... Brandt, A. R. (2021). Carbon implications of marginal oils from market-derived demand shocks. *Nature*, *599*, 80–84.

Masnadi, M. S., El-Houjeiri, H. M., Schunack, D., Li, Y., Englander, J. G., Badahdah, A., ... others (2018). Global carbon intensity of crude oil production. *Science*, *361*(6405), 851–853.

Maugeri, L. (2012). Oil: the next revolution. *Belfer Center for Science and International Affairs Discussion Paper*(2012-10).

McElroy, M. B. (1987). Additive general error models for production, cost, and derived demand or share systems. *Journal of Political Economy*, *95*(4), 737–757.

Milonas, N. T., & Henker, T. (2001). Price spread and convenience yield behaviour in the international oil market. *Applied Financial Economics*, *11*(1), 23–36.

Nakagawa, S., & Schielzeth, H. (2013). A general and simple method for obtaining r2 from generalized linear mixed-effects models. *Methods in ecology and evolution*, *4*(2), 133–142.

Nelder, J. A., & Wedderburn, R. W. (1972). Generalized linear models. *Journal of the Royal Statistical Society: Series A (General)*, *135*(3), 370–384.

Nordhaus, W. D. (2009). The economics of an integrated world oil market. In *International energy workshop* (pp. 17–19).

Pesaran, M. H. (1990). An econometric analysis of exploration and extraction of oil in the uk continental shelf. *The Economic Journal*, *100*(401), 367–390.

Pesaran, M. H. (2007). A simple panel unit root test in the presence of cross-section dependence. *Journal of applied econometrics*, *22*(2), 265–312.

Said, S. E., & Dickey, D. A. (1984). Testing for unit roots in autoregressive-moving average models of unknown order. *Biometrika*, *71*(3), 599–607.

Shah, A., Fishwick, R., Wood, J., Leeke, G., Rigby, S., & Greaves, M. (2010). A review of novel techniques for heavy oil and bitumen extraction and upgrading. *Energy & Environmental Science*, *3*(6), 700–714.

Solow, R. M., & Wan, F. Y. (1976). Extraction costs in the theory of exhaustible resources. *The Bell Journal of Economics*, 359–370.

Swamy, P. A. (1970). Efficient inference in a random coefficient regression model. *Econometrica: Journal of the Econometric Society*, 311–323.

Taheri, S., & Hesamian, G. (2013). A generalization of the wilcoxon signed-rank test and its applications. *Statistical Papers*, *54*(2), 457–470.

Trident Consulting. (2023). Hydrocarbon loss: identifying opportunities. *https://www.trindent.com/hydrocarbon-loss-identifying-opportunities*.

U.S. Department of Transportation, U. (2022). Highway statistics series, state motor-vehicle registrations - 2021. *https://www.fhwa.dot.gov/policyinformation/statistics/2021/mv1.cfm*.

WoodMackenzie. (2020). Wood mackenzie upstream data tool. *WoodMackenzie Upstream Oil & Gas Analysis.(https://www.woodmac.com/our-expertise/capabilities/upstream-oil-and-gas/)*.

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data.* MIT press.

Zimmerman, M. B. (1977). Modeling depletion in a mineral industry: The case of coal. *The Bell Journal of Economics*, 41–65.