University of Glasgow

Adam Smith Business School

WORKING PAPER SERIES

Extracting Wedges: Misallocation and Taxation in the Oil Industry

Radek Stefanski, Lassi Ahlvik, Jorgen Juel Andersen, Torfinn Harding and Alex Trew

Paper No. 2025-10
May 2025

Extracting Wedges: Misallocation and Taxation in the Oil Industry*

Radek Stefanski, Lassi Ahlvik,
Jørgen Juel Andersen, Torfinn Harding and Alex Trew

May 30, 2025

Abstract

How large are the productivity differences arising from micro-level distortions, and how much of that is due to tax policy? Using over a century of field-level data (1900-2023), this paper examines the role of field-level revenue taxes in explaining misallocation in the oil and gas industry, a single large sector that produces a homogeneous, globally-traded good. A key advantage is our ability to link model-implied distortions directly to these observed tax rates. We show that misallocation is significant in the oil industry, and that over half of this misallocation can be accounted for by the dispersion in revenue tax rates across fields, exceeding the 2-25% explanatory power typical in studies of misallocation sources. We show that nearly all of the impact of this tax dispersion operates through the intensive margin (the inputs allocated at a field) rather than the extensive margin (the choice to enter a field). These findings have direct implications for tax policy.

Keywords: Misallocation; productivity; distortions; tax policy.
JEL codes: O47, O11, D24, Q32.

*Stefanski (author for correspondence): University of St Andrews, rls7@st-andrews.ac.uk. Ahlvik: University of Helsinki and Helsinki GSE. Andersen: BI Norwegian Business School. Harding: University of Stavanger Business School. Trew: University of Glasgow, IZA, CESifo. Stefanski thanks Gerhard Toews for providing the data and for the initial collaboration without which this paper would not have been possible. A previous version of this paper circulated with the title "What's in a Wedge?" We would also like to thank Christiano Eichenbaum, Victoire Girard, Doug Gollin, Lutz Kilian, Alexander Monge-Naranjo, Alexander Naumov, Rick van der Ploeg, Steven Poelhekke, Chiara Ravetti, Akos Valentinyi, and Tony Venables as well as participants at seminars at Yale, Glasgow, Oxford, St. Andrews, Amsterdam, Aberdeen, Heriot-Watt, the OECD and the Environmental Conference at the University of Orleans, the CEBRA workshop at the Dallas FED, the SED conference in Mexico City and the 7th NBP Summer workshop in Warsaw, the 2023 Scottish Economic Society Annual Conference as well as the 2024 Structural Change and Productivity Conference at the Cambridge Janeway Institute.

1 Introduction

Large differences in income across countries have been a longstanding focus of research in economics. A key factor driving these disparities is heterogeneity in total factor productivity (TFP) across countries. In turn, a substantial portion of these TFP differences can be attributed to the misallocation of resources across firms and sectors within countries. Hsieh and Klenow (2009) estimate that if capital and labor inputs in China and India were allocated as efficiently as in the United States, TFP would be 30-60% higher in these countries. While the magnitude of misallocation is now well-established, identifying its specific causes and quantifying their relative importance remains a significant challenge. This is in part because the unobserved distortions (or 'wedges') that models infer often serve as imperfect proxies for the actual policies or frictions researchers aim to measure. Unobserved heterogeneity across sectors often further compounds these challenges of mis-specification.

To make progress on identifying the specific sources of misallocation, we need settings where the model distortions can be directly observed and measured, and where sector-specific heterogeneity can be controlled for. One potentially important but understudied factor is dispersion in the tax rates faced by different producers at different locations. While it is well-known that corporate tax systems differ across countries and in their design, the specific impact of tax rate dispersion at the field level on misallocation has, to our knowledge, not been quantified. In this paper, we use field-level global data and focus on the role of the dispersion in revenue taxes—taxes on the gross income earned by firms—in driving misallocation in the upstream oil and gas industry. A typical concern with inferring and explaining distortions is unobserved heterogeneity at the sector level. Our approach addresses this problem by focusing on a single, large and globally-traded sector.

The oil and gas industry is a natural setting to study this question for several reasons. First, the industry is globally integrated, with many large firms operating production fields across multiple countries, producing a homogeneous good. This homogeneity in output provides a cleaner setting for productivity comparisons than typical multi-sector studies. Second, common international prices provide clear benchmarks for measuring distortions, reducing concerns about measurement error. Third, there is substantial variation in distortions within the oil and gas sector, with intra- and inter-national heterogeneity in the tax rates imposed by different governments on oil producers, especially revenue taxes such as royalties. As we find, there can even be important within-firm dispersion in distortions across fields within the same country.

To study the impact of revenue taxes on misallocation in the oil industry, we develop a model of firm investment and production. Multinational firms choose whether to develop new fields (the extensive margin) or to expand existing fields (the intensive margin). Fields can differ in their underlying productivity, and firms face fixed drilling costs to begin production on new fields. Once operating, firms combine capital and labor to extract output, taking output prices and input costs as given.Crucially, firms face dispersion in revenue tax rates across the fields they operate, varying by country and sometimes even within the same country. Firms also face country-specific non-tax distortions such as variations in trade or transportation costs, borrowing constraints, political connectedness, institutional frameworks or even differences in geography and climate. We derive expressions for the marginal revenue products of capital and labor as a function of not only technological parameters but also tax and non-tax wedges. Unlike standard approaches that typically infer distortions as residuals and use them as a proxy for a separate (unmodelled) observed distortion, our model incorporates directly observed tax rates, which we can decompose into an observed tax and unobserved non-tax component.

We then use rich, field-level, worldwide data on oil fields from Rystad Energy, spanning the period 1900-2023, to estimate the key parameters of the model. Using the United States as a benchmark, we infer the underlying productivity of fields and the magnitude of distortions. We find that misallocation is substantial in the global oil industry—if inputs across all fields were efficiently allocated global output would increase by 61%. Strikingly, the dispersion in revenue tax rates across countries and firms accounts for over half of this misallocation. Even if we control for the extensive margin by holding fixed the number of fields, this tax rate dispersion still explains the large bulk of misallocation. This suggests that revenue taxes impact productivity primarily by distorting the intensive margin of capital and labor inputs.

Literature This paper contributes to several strands of the literature. Methodologically, our approach builds on the framework of Hsieh and Klenow (2009) for indirectly measuring misallocation using model-based wedges. Compared to Hsieh and Klenow, we also consider the extensive margin at which distortions may operate (i.e., the firm's decision to enter a field), in addition to the intensive margin. We then follow Restuccia and Rogerson (2017) in trying to directly understand the causes of that misallocation. The role of taxes in misallocation has been studied in Fajgelbaum et al. (2019), though that paper considers only state-level variation in taxes and so the explanatory power is relatively low. A number of other frictions have also been quantified, including financial (Buera et al., 2011; Midrigan and Xu, 2014), firm size constraints (Garicano et al., 2016), trade restrictions (Edmond et al., 2015; Brandt et al., 2017) and property rights (Adamopoulos and Restuccia, 2014; Chen et al., 2023). Typically these direct approaches account for around only 2-25% of the total distortions. We focus specifically on the role of tax rate dispersion, which allows us to quantify the productivity costs of heterogeneous fiscal policies. Our finding that the dispersion in revenue taxes at the intensive and extensive margin can explain over 50% of the total inferred wedges represents a significant advance on our previous understanding of distortions.

We also add to a growing literature that focuses on productivity and misallocation in the petroleum sector. Asker et al. (2019) quantifies misallocation in the oil sector, specifically focusing on market power. Our work complements theirs by identifying tax policy as a specific source of the misallocation they document. Coulomb et al. (2021) study the climate implications of misallocation, when oil is extracted from emission intensive, high-cost deposits. Our results complement these studies by pointing at a specific source of misallocation—dispersion in production taxes—that directly connects to policy levers. The results imply that changes to tax policy may be a powerful lever for boosting productivity in this important industry, freeing up resources that may be used more productively elsewhere in the economy. More broadly, our paper provides a general framework to study the implication of misallocation, in a way that is consistent with the earlier findings that petroleum investments respond strongly to the business climate (Bohn and Deacon, 2000; Cust and Harding, 2020; Arezki et al., 2019; Hamang, 2024) and taxes (Black et al., 2018; Anderson et al., 2018; Brown et al., 2020; Ahlvik et al., 2022; Ahlvik and Harding, 2024).

Outline The remainder of the paper is organized as follows. Section 2 presents the model and derives the key theoretical results. Section 3 describes the data and measurement. Section 4 details the calibration approach and parameter estimates. Section 5 first employs an accounting approach to analyze the dispersion of distortions and the role of revenue taxes. Section 6 then presents results from our structural model, including counterfactuals that quantify the aggregate output gains from eliminating these distortions, and explores the robustness of our findings. Section 7 concludes.

2 Model

We develop a model of the global oil industry that permits us to isolate the role of wedges on resource allocation. The model is structured as a sequence of static equilibria, solved independently for each year. Within any given year, a set of firms is indexed by $f = 1, \dots, F$, and countries by $i = 1, \dots, I$.[1]

A firm f that operates in country i can develop a continuum of potential oil assets. Each potential asset, if developed, becomes an operational "field". These potential assets are distinguished by their intrinsic productivity, A. We model this as the firm f drawing an asset's productivity from a firm-country specific cumulative distribution, $G_{fi}(A)$ (with corresponding PDF $g_{fi}(A)$). A draw A from this distribution represents the inherent quality of a specific potential oil discovery for that firm in that country, capturing factors like geological richness, ease of extraction, and firm-specific expertise in exploiting such an asset.[2] Figure 1 illustrates how a firm may face distinct productivity distributions across the different countries where it might operate.

Firm decisions involve two margins. At the *extensive margin*, the firm decides whether to incur a fixed development cost, ω_{fi}, to turn a potential asset with productivity A into an operational field. At the *intensive margin*, for each operational field, the firm determines the optimal input levels of capital, $K_{fi}(A)$, and labor, $L_{fi}(A)$.

Since the model is solved year by year, time subscripts (t) are generally omitted from variables and parameters for notational clarity. All variables and distributions

[1] Modeling the extensive margin of country entry/exit is beyond this paper's scope, partly due to the significant data challenges in systematically observing the firm-country specific factors and negotiation processes that govern these complex, often confidential, decisions. Abstracting from this margin means our estimated impacts of tax dispersion are specific to the field-level operations studied; the net effect on the total magnitude of misallocation from omitting country choice is ambiguous.

[2] The firm-specific nature of $G_{fi}(A)$ allows for firms to have heterogeneous capabilities or technologies that make them differentially effective even with similar geological endowments. Firms bring a firm-specific productivity to the table – be it in terms of technology, human capital, skills, financing or other factors that enable some firms to be better at running certain oil fields than others. This is akin to firm-specific efficiency in broader productivity models.

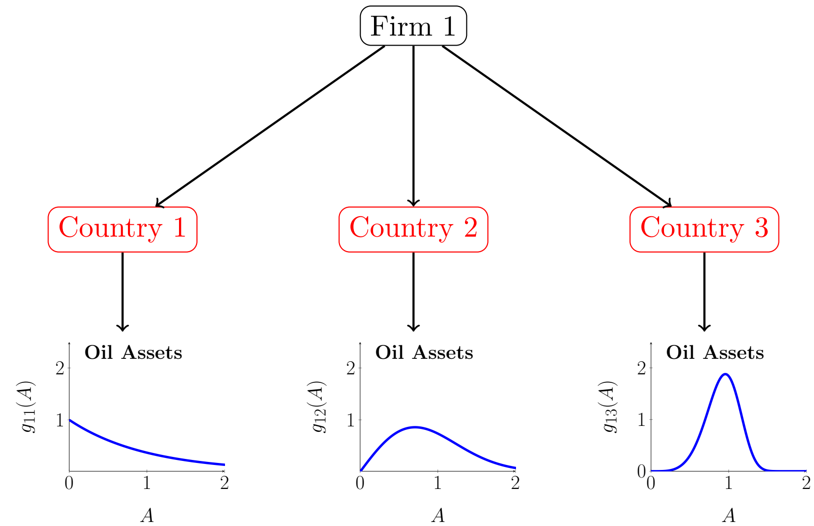


Figure 1: Firm structure. A firm $f = 1$ operates in a set of countries $i = 1, 2, 3$, facing different distributions of asset productivity draws A (illustrated by PDFs $g_{fi}(A)$) in each country. Notice that the countries the firm operates in are fixed. The firm then decides which fields to operate in each country and how many inputs to assign to each operational field. In the above $g_{fi}(\cdot)$ represents the corresponding probability density function to the CDF $G_{fi}(\cdot)$.

should be understood as year specific unless otherwise indicated. Notice that this includes productivity distributions which may change exogenously over time capturing any external shifts in productive efficiency.

Firms are price-takers in output and factor markets.[3] The development of any asset into an operational field by firm f in country i requires incurring a fixed cost ω_{fi}. This cost, paid in units of capital, is assumed to be the same regardless of the specific productivity A of the asset being considered.[4]

A firm will choose to develop a potential asset with productivity A if the expected

[3]This assumption is a standard simplification. While the global oil market has oligopolistic features, individual firms often make field-level development and production decisions conditional on a perceived world price. This approach allows us to isolate the impact of heterogeneous firm-country distortions $(\tau^Y_{fi}, \tau^K_{fi})$ on resource allocation, which is the central focus of this paper. Incorporating strategic pricing behavior would introduce considerable complexity beyond our current scope.

[4]While in reality development costs might vary with geological complexity (which could be correlated with A), assuming a uniform ω_{fi} simplifies the extensive margin decision considerably without losing the core trade-off. Heterogeneity in A remains the primary driver of differential profitability across assets.

operating profit from the resulting field is non-negative:

$$\pi_{fi}(A,\omega_{fi}) = \max_{L_{fi}(A),K_{fi}(A)} \Bigg(p(1-\tau^{Y}_{fi})AK_{fi}(A)^{\gamma}L_{fi}(A)^{\alpha} \\ - wL_{fi}(A) - r(1+\tau^{K}_{fi})K_{fi}(A) \Bigg) - \omega_{fi} \geq 0. \tag{1}$$

Here, p is the oil price, w the wage, and r the capital rental rate. The production technology for a field is $Y_{fi}(A) = AK_{fi}(A)^{\gamma}L_{fi}(A)^{\alpha}$. We assume decreasing returns to scale in the variable inputs, i.e., $0 < \alpha + \gamma < 1$, motivated by the presence of unmodeled field-specific fixed factors (like the fixed size of oil reserves) or managerial span-of-control limitations at the field level. This assumption ensures a determinate optimal operating scale for each developed field. The terms τ^{Y}_{fi} and τ^{K}_{fi} represent firm-country specific wedges that distort effective revenues and capital costs, respectively. These wedges are crucial: they capture the combined effect of taxes, subsidies, regulations, geography, climate, trade or transportation costs and other institutional factors that differ across firms and countries.

Simplification In reality, distortions can vary even among fields of the same firm within one country. A general model would allow for the intrinsic productivity A and field-specific distortions (i.e., $\tau^{Y}(A), \tau^{K}(A)$) to be drawn from a joint distribution. The key challenge with such an approach is empirical: reliably calibrating this high-dimensional distribution, including all its correlations and higher moments, would be exceptionally demanding, requiring vast data and raising difficult questions about whether observed correlations are truly structural.

Given these empirical hurdles, we proceed with a more tractable framework and focus on firm-country average distortions. A key simplifying assumption of our model is thus that the wedges, τ^{Y}_{fi} and τ^{K}_{fi}, are the same across all the fields developed by firm f in country i. This focuses our analysis on the impact of *average* distortions at the firm-country level and greatly simplifies the calibration of the model.

Importantly, this simplification does not mean that we neglect the rich heterogeneity that exists across fields in our data. On the contrary, our calibration of the firm-country specific productivity distribution, $G_{fi}(A)$, is designed to incorporate the influence of these field-level distortion variations, albeit in a reduced-form way. As we explain in Section 4, the *mean* of our modeled $G_{fi}(A)$ will be disciplined by the observed aggregate productivity of each firm-country unit, which inherently reflects any misallocation internal to its portfolio of operational fields. Furthermore, the *variance* of $G_{fi}(A)$ will be informed by constructing an "effective" productivity measure for each operational field, a measure that accounts for that field's specific distortions relative to the firm-country average. By matching the moments of productivity in our model to this empirically-derived effective productivity, our calibrated $G_{fi}(A)$ will embed the consequences of the unmodeled field-level distortion heterogeneity.

Solution The profit maximization in (1) yields optimal input choices for an active field with productivity A:

$$L_{fi}(A) = \left[p \left(\frac{\gamma}{r}\right)^{\gamma} \left(\frac{\alpha}{w}\right)^{1-\gamma} A \frac{(1-\tau^Y_{fi})}{(1+\tau^K_{fi})^{\gamma}} \right]^{\frac{1}{1-\alpha-\gamma}}, \tag{2}$$

$$K_{fi}(A) = \left[p \left(\frac{\gamma}{r}\right)^{1-\alpha} \left(\frac{\alpha}{w}\right)^{\alpha} A \frac{(1-\tau^Y_{fi})}{(1+\tau^K_{fi})^{1-\alpha}} \right]^{\frac{1}{1-\alpha-\gamma}}. \tag{3}$$

An asset is developed if, after substituting the optimal input choices from (2) and (3) into (1), the resulting maximized profit $\pi_{fi}(A, \omega_{fi})$ is non-negative. The break-even point, where this maximized profit (net of fixed costs) equals zero, defines the cutoff productivity $\bar{A}_{fi}$:

$$\bar{A}_{fi} \equiv \underbrace{\left(\frac{1}{1-\alpha-\gamma}\right)^{1-\alpha-\gamma} \left(\frac{r}{\gamma}\right)^{\gamma} \left(\frac{w}{\alpha}\right)^{\alpha} \frac{1}{p}}_{\equiv \delta} \underbrace{\omega_{fi}^{1-\alpha-\gamma}}_{\equiv \bar{\omega}_{fi}} \underbrace{\frac{(1+\tau^K_{fi})^{\gamma}}{(1-\tau^Y_{fi})}}_{\equiv T_{fi}}. \tag{4}$$

The threshold $\bar{A}_{fi}$ depends on oil prices and input costs (δ), effective fixed costs ($\bar{\omega}_{fi}$), and a composite distortion term (T_{fi}). The economic viability of developing a field with a given intrinsic productivity, A, depends on these factors. Favorable conditions—such as high output prices, low input costs, and minimal distortions—lower the effective barrier for development, making it profitable to operate even fields with relatively low intrinsic productivity. Conversely, under less favorable circumstances, firms will be more selective, focusing only on the most promising, high-productivity fields.

Figure 2 illustrates this extensive margin cutoff by plotting field productivities against country-wide distortions. The upward-sloping, blue line represents the cutoff

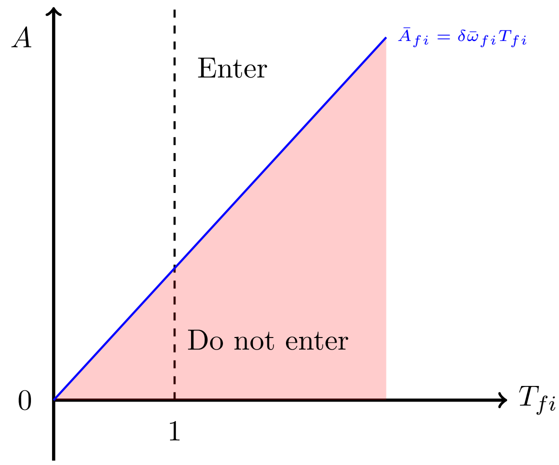


Figure 2: Extensive margin cutoff. For given δ and $\bar{\omega}_{fi}$, the upward-sloping blue line shows the minimum field productivity $\bar{A}_{fi}$ required for development, as a function of the composite distortion T_{fi}. The vertical, dashed, black line shows where $T_{fi} = 1$, the undistorted benchmark and hence what field-productivities would be operated without distortions.

condition (4) — any field above this line is profitable to operate, while any field below is not. Changes in global conditions (δ) or fixed costs ($\bar{\omega}_{fi}$) cause the cutoff line to swivel up or down, while changes in distortions (T_{fi}) move us along the x-axis. The vertical, dashed, black line at $T_{fi} = 1$ represents the undistorted benchmark. A firm-country unit faces this benchmark distortion level, will develop its potential assets into operational fields if and only if their intrinsic productivity A is higher than the blue cutoff line, $\bar{A}_{fi}$.

We can solve for the total labor capital and output of the firm-country unit by aggregating over developed fields (i.e., where $A \geq \bar{A}_{fi}$) for firm f in country i:

$$L_{fi} = \int_{\bar{A}_{fi}}^{\infty} L_{fi}(A)\, dG_{fi}(A), \tag{5}$$

$$K_{fi} = \int_{\bar{A}_{fi}}^{\infty} K_{fi}(A)\, dG_{fi}(A), \tag{6}$$

$$Y_{fi} = \int_{\bar{A}_{fi}}^{\infty} Y_{fi}(A)\, dG_{fi}(A). \tag{7}$$

The realized aggregate productivity for firm f in country i is then:

$$\mathcal{A}_{fi} \equiv \frac{Y_{fi}}{K_{fi}^{\gamma} L_{fi}^{\alpha}} = \left(\mathrm{E}\left[A^{\frac{1}{1-\alpha-\gamma}} \,\middle|\, A > \bar{A}_{fi} \right] \right)^{1-\alpha-\gamma}. \tag{8}$$

This measure, $\mathcal{A}_{fi}$, is a key empirical object. Notice also that the second equality in (8) shows that, due to aggregation over heterogeneous fields under decreasing returns to scale, $\mathcal{A}_{fi}$ is a generalized mean of the intrinsic productivities of active fields. This relationship will be used later in the calibration section. Finally, total fixed costs are $C_{fi} = \omega_{fi} \int_{\bar{A}_{fi}}^{\infty} dG_{fi}(A)$. Capital for exploration and development is $K_{fi}^{E} \equiv C_{fi}/r$.

The model is closed by an exogenous rental rate of capital, r, and fixed total quantities of capital ($\bar{K}$) and labor ($\bar{L}$) employed by the sector. The world oil price, p, and the wage rate, w, adjust endogenously to ensure that aggregate firm demands for these inputs match their fixed sectoral totals:

$$\sum_{f=1}^{F} \sum_{i=1}^{I} \left(K_{fi} + K_{fi}^{E} \right) = \bar{K}, \tag{9}$$

$$\sum_{f=1}^{F} \sum_{i=1}^{I} L_{fi} = \bar{L}. \tag{10}$$

This closure, similar in spirit to the approach in Hsieh and Klenow (2009), frames our misallocation analysis around quantifying gains from reallocating these observed sectoral factor totals more efficiently, holding r constant. Counterfactuals thus measure changes in aggregate output or TFP from the optimal deployment of these given inputs.

3 Data and Measurement

We use a comprehensive, proprietary micro-level dataset from Rystad Energy, a leading energy consultancy and data provider. This dataset spans the years 1900 to 2023 and covers the vast majority of global upstream activity: it includes information on over 6,000 firms and accounts for 99.5% of global oil reserves and 98.4% of oil production. The data are recorded at the level of individual assets—an active license, a specific field, or a discovery. For our purposes, we focus on operational fields. Figure 3 provides a geographical overview of these assets.

While the ultimate decision-making unit in our model is the firm operating within a country, the fundamental heterogeneity and many of the distortions we aim to understand are at the level of individual operational fields. Therefore, our primary data construction occurs at this granular level. As detailed in Section 4, this field-

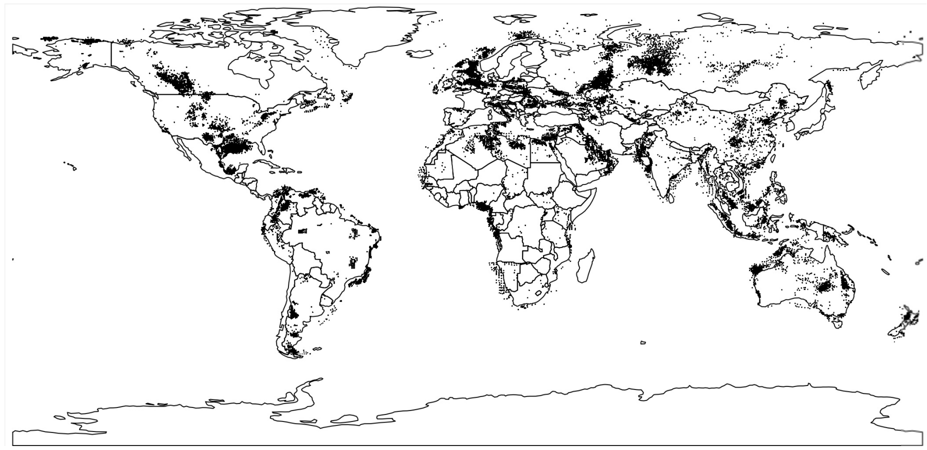


Figure 3: Map of Assets in the Rystad Energy Dataset

specific information is crucial for calibrating the underlying productivity distributions within our model.

For each operational field j, operated by firm f in country i, during year t, we construct the following key economic variables from the Rystad cash flow statements. All monetary values are converted to constant 2010 U.S. dollars to ensure comparability across time.[5]

Variables that we observe in the data are denoted x^{obs} to distinguish them from model objects. First, total revenue ($p_t Y^{\text{obs}}_{fijt}$) represents the value firm f generates from selling the oil and gas produced by field j. Second, production capital expenditures (I^{obs}_{fijt}) refer to the firm's spending on durable goods (assets with a lifetime > 1 year) for production activities at field j; these expenditures are instrumental in constructing the field-level production capital stock, K^{obs}_{fijt}, via the perpetual inventory method, a process detailed further in Section 4. Third, operational costs ($w_t L^{\text{obs}}_{fijt}$) are calculated as the amount spent on labor and non-durable inputs for field j. While primarily composed of salaries and wages, this category also includes other variable costs like materials and maintenance, and throughout the paper, we often refer to this variable as "labor costs," although this is, of course, a simplification. Fourth, total government take (TGT^{obs}_{fijt}) is the sum of all tax payments received by the government from firm f

[5] Nominal values are deflated using the U.S. GDP deflator series from (Johnston and Williamson, 2025). We apply this single deflator across all countries because firms in this globally integrated industry generally face international prices for both their inputs and outputs. As will become clear later, the specific choice of deflator does not affect our core results on estimated wedges or the outcomes of our counterfactual exercises, as it cancels out in the relevant calculations.

associated to field j, typically in exchange for the rights to explore and extract oil and gas. Finally, exploration capital expenditures ($I^{E,\text{obs}}_{fijt}$) denote the firm's investment in durable goods (assets with a lifetime > 1 year) for exploring and appraising field j or its broader asset area; these expenditures are analogously used to construct the field-level exploration capital stock, $K^{E,\text{obs}}_{fijt}$.

These granular, field-level measurements of revenues, expenditures, derived capital stocks, and taxes form the empirical basis for our analysis. For parts of our analysis that connect to firm-country level decisions in our model (such as average distortions or aggregate outcomes), these field-level variables (e.g., Y^{obs}_{fijt}, K^{obs}_{fijt}, L^{obs}_{fijt} etc.) are summed across all operational fields j of firm f in country i to obtain their firm-country aggregate counterparts (e.g., Y^{obs}_{fit}, K^{obs}_{fit}, L^{obs}_{fit} etc.).

With these constructed measures we are now equipped to bring our theoretical model to the data. The next section details this crucial step: the calibration of the model's parameters.

4 Calibration

To calibrate the model we proceed in two steps. First, we calibrate the production function elasticities and infer the average firm-country level distortions directly from the data and the model's first-order conditions. Second, we calibrate the parameters governing the distribution of intrinsic field-level productivities and the fixed costs of field development for each firm-country pair.

Elasticities and Firm-Level Wedges For each firm f operating in country i, we use data on total revenues (pY^{obs}_{fi}), operational expenditures (wL^{obs}_{fi}), and the capital stock (K^{obs}_{fi}) to infer the output and capital wedges, τ^Y_{fi} and τ^K_{fi}. We set the real rental price of capital (excluding distortions) to $r = 0.1099$. This includes a 3.7% real interest rate and a 7.36% depreciation rate which we obtain as the average Implied Depreciation Rate for Private Nonresidential Fixed Assets of the Oil and Gas sector between 1920 and 2023 from the BEA (2020).[6] We calculate the firm-country specific

[6]We construct the capital stock series using the perpetual inventory method at the field level, assuming that capital evolves according to $K_{fij,t+1} = (1-\delta)K_{fij,t} + I^{\text{obs}}_{fij,t}$, where $\delta = 0.0729$ is the depreciation rate and $I^{\text{obs}}_{fij,t}$ is real capital expenditure. We initialize field-level capital stock series assuming zero capital at the start of a field's observed operations. We then aggregate these measures up to obtain capital at the country-firm level, $K_{fi,t}$.

revenue tax rate (τ^V_{fi}) from the data as the ratio of total government take (TGT^{obs}_{fi}) to total revenues (pY^{obs}_{fi}). As we will see, the overall output wedge in the model is made up of this revenue-tax rate and an implicit non-revenue-tax.

To determine the production elasticities, α (for labor) and γ (for capital), we follow Hsieh and Klenow (2009) and assume the United States serves as a relatively undistorted benchmark. Specifically, we assume that wedges for the median *field* operating in the U.S. over our entire (pooled) sample, are zero. This normalization is standard in the misallocation literature and allows us to express wedges in other countries relative to this U.S. benchmark. Importantly, since all our counterfactuals will be based on relative comparisons, this has no quantitative or qualitative effect on our results. Given this normalization, from the first order conditions of the firm we obtain:

$$\alpha = \overline{\left(\frac{wL^{\text{obs}}_{fijt}}{(1-\tau^V_{fijt})pY^{\text{obs}}_{fijt}}\right)}_{i=US} = 0.406 \quad \text{and} \quad \frac{\alpha}{\gamma} = \overline{\left(\frac{wL^{\text{obs}}_{fijt}}{rK^{\text{obs}}_{fijt}}\right)}_{i=US} = 1.025, \qquad (11)$$

where $\bar{\cdot}$ represents the median over all firms, fields and years. These calculations imply $\alpha = 0.406$ and $\gamma = 0.396$ so that the returns to scale parameter is $1-\alpha-\gamma = 0.198 > 0$.

With these parameters in hand, we can use the firm's first-order conditions to calculate the *firm-country average* wedges τ^Y_{fi} and τ^K_{fi} as:

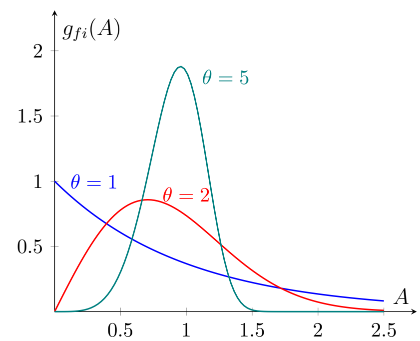
$$1-\tau^Y_{fi} = \frac{1}{\alpha}\frac{wL^{\text{obs}}_{fi}}{pY^{\text{obs}}_{fi}} \qquad (12)$$

$$1+\tau^K_{fi} = \frac{\gamma}{\alpha}\frac{wL^{\text{obs}}_{fi}}{rK^{\text{obs}}_{fi}} \qquad (13)$$

The overall output wedge $1-\tau^Y_{fi}$ can further be decomposed into the observed firm-country revenue tax $1-\tau^V_{fi}$ and an implicit non-revenue-tax wedge $1-\tau^W_{fi}$:

$$1-\tau^Y_{fi} \equiv (1-\tau^V_{fi})(1-\tau^W_{fi}). \qquad (14)$$

These wedges, τ^Y_{fi} and τ^K_{fi}, represent the average distortions faced by firm f in country i and will serve as the baseline distortions in our model. Notice that if we allowed wedges to vary at the field level, then corresponding equations to 12 and 13 could easily be used to measure field-level wedges τ^Y_{fij} and τ^K_{fij}. This will prove useful in the later parts of our calibration.



Productivity A follows a Weibull distribution with shape θ_{fi} and scale μ_{fi}

Figure 4: Illustrative Weibull Productivity Distributions (PDF $g_{fi}(A)$)

Productivity Distributions and Development Costs The second stage of our calibration determines the parameters governing the intrinsic productivity distribution of oil-assets and their fixed development costs. For each firm f in country i, we assume these asset productivities, A, are drawn from a firm-country specific Weibull distribution, $G_{fi}(A) \sim \text{Weibull}(\theta_{fi}, \mu_{fi})$, characterized by shape parameter θ_{fi} and scale parameter μ_{fi}. Figure 4 illustrates this distribution for a fixed scale parameter and a changing shape parameter.

As noted in Section 2, explicitly estimating field-level heterogeneity would require specifying and calibrating a high-dimensional joint distribution for individual asset productivities and their unique distortions, a step that poses considerable empirical challenges. To bridge this gap, our strategy is to ensure that our model's simple productivity distribution, $G_{fi}(A)$, reflects the combined economic impact of both distortions and intrinsic productivity. Our model assumes *constant* firm-country average distortions $(\tau^Y_{fi}, \tau^K_{fi})$, even though actual operational fields exhibit significant heterogeneity in both their underlying productivities and their observed, field-level distortions $(\tau^Y_{fij}, \tau^K_{fij})$.

To reconcile the model with the data, we first use our detailed field-level data to construct a "effective productivity" (A^{eff}_{fij}) for each observed operational field j in the data. This A^{eff}_{fij} adjusts a field's raw measured productivity (A^{obs}_{fij}, derived from its output and inputs) to account for how its specific distortions $(\tau^Y_{fij}, \tau^K_{fij})$ deviate from the firm-country average.[7] In essence, A^{eff}_{fij} captures the true productivity signal from

[7]The calculation of A^{obs}_{fij} from field-level data and the precise formula for the adjustment factor to obtain A^{eff}_{fij} (our Equation 28) are detailed in Appendix A. The intuition is that A^{eff}_{fij} is constructed such that its generalized mean, when aggregated across a firm's fields, corresponds to the firm-

each field relevant for our model, which only has firm-country average distortions.

We then calibrate three key parameters for each firm-country unit: the fixed development cost ω_{fi}, as well as the Weibull distribution parameters θ_{fi} and μ_{fi}. First, ω_{fi} is set so the model's entry threshold, $\bar{A}_{fi}$, matches the lowest observed A^{eff}_{fij} among active fields. Since $\bar{A}_{fi}$ is a function of ω_{fi} this also pins down ω_{fi}. Given this $\bar{A}_{fi}$, the Weibull parameters θ_{fi} and μ_{fi} are jointly chosen to ensure the model's truncated productivity distribution (for $A > \bar{A}_{fi}$) matches both the mean (via aggregate TFP) and variance of the observed A^{eff}_{fij} values for that country-firm.[8]

This approach ensures that our model's "intrinsic" productivity distribution $G_{fi}(A)$ effectively absorbs not only true differences in asset productivity but also the average impact of unmodeled field-level distortion heterogeneity. The resulting parameters $(\theta_{fi}, \mu_{fi}, \omega_{fi})$ are thus best interpreted as reduced-form estimates that allow our tractable model to replicate key aggregate features of the data. While this means the model may not perfectly predict allocations for any single, specific field (if driven by unique distortions), it provides a robust foundation for analyzing firm-country aggregate behavior and the impact of changes in average firm-country distortions.

Our calibrated model provides a lens through which to analyze the distortions within the global oil and gas industry. We begin this analysis in the next section with an accounting exercise focused on the dispersion of these distortions, particularly the role played by heterogeneous revenue taxes. The broader implications for aggregate output and productivity are then explored in the subsequent section using model-based counterfactuals.

5 An Accounting Approach to Distortions

Following Hsieh and Klenow (2009), we first measure misallocation by examining wedges in the marginal revenue products (MRPs) of capital and labor. For a firm f in country i, operating with Cobb-Douglas production, these MRPs are proportional

country aggregate TFP, $\mathcal{A}_{fi}$, that would arise in a world where field-level distortions are present and heterogeneous, thus ensuring consistent aggregation from field-level effective productivities to the firm-country TFP.

[8]Formally, we match the model's implied aggregate TFP, $\mathcal{A}_{\text{model},fi}$, to the aggregate TFP constructed from the A^{eff}_{fij} data, and we match the variance of A (conditional on $A > \bar{A}_{fi}$) in the model to the variance of A^{eff}_{fij} in the data. This system of three moment conditions identifies ω_{fi}, θ_{fi}, and μ_{fi}.

to average revenue products:

$$MRP_{fi}^{K} \equiv \gamma \frac{p_t Y_{fi}^{\text{obs}}}{K_{fi}^{\text{obs}}} = \frac{r_t(1+\tau_{fi}^{K})}{(1-\tau_{fi}^{V})(1-\tau_{fi}^{W})} \tag{15}$$

$$MRP_{fi}^{L} \equiv \alpha \frac{p_t Y_{fi}^{\text{obs}}}{L_{fi}^{\text{obs}}} = \frac{w_t}{(1-\tau_{fi}^{V})(1-\tau_{fi}^{W})}. \tag{16}$$

In an efficient allocation, these MRPs would be equalized to input costs across all country-firm units. Any measured differences observed MRPs thus represent wedges driving the economy from efficiency.

To isolate the impact of non-revenue tax distortions, we define "After-Tax" MRPs (MRP_{AT}), which notionally remove the revenue tax τ_{fi}^{V}:

$$MRP_{AT,fi}^{K} \equiv MRP_{fi}^{K}(1-\tau_{fi}^{V}) = \frac{r_t(1+\tau_{fi}^{K})}{1-\tau_{fi}^{W}} \tag{17}$$

$$MRP_{AT,fi}^{L} \equiv MRP_{fi}^{L}(1-\tau_{fi}^{V}) = \frac{w_t}{1-\tau_{fi}^{W}} \tag{18}$$

We summarize the overall distortion for each firm-country, fi, using a geometric average of its factor MRPs, which we term $MRP_{\text{Total},fi}$. Similarly, we define $MRP_{AT,fi}$ using the after-tax factor MRPs:

$$MRP_{\text{Total},fi} \equiv (MRPK_{fi})^{\frac{\gamma}{\alpha+\gamma}} (MRP_{fi}^{L})^{\frac{\alpha}{\alpha+\gamma}} \tag{19}$$

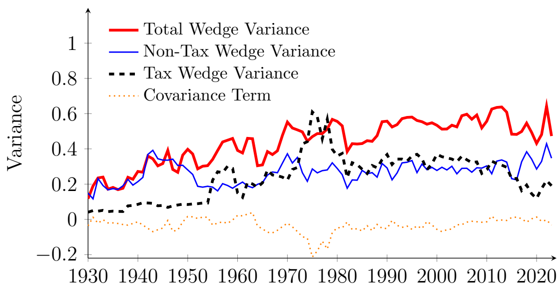
$$MRP_{AT,fi} \equiv (MRP_{AT,fi}^{K})^{\frac{\gamma}{\alpha+\gamma}} (MRP_{AT,fi}^{L})^{\frac{\alpha}{\alpha+\gamma}}. \tag{20}$$

From these definitions, it follows directly that:

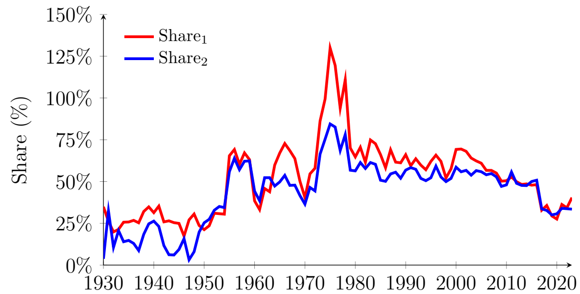
$$\boxed{MRP_{\text{Total},fi} = MRP_{AT,fi} \times \frac{1}{1-\tau_{fi}^{V}}} \tag{21}$$

This expression cleanly decomposes the total measured wedge into a component reflecting non-revenue tax distortions ($MRP_{AT,fi}$) and a component due to the revenue tax ($1/(1-\tau_{fi}^{V})$).

Variance Decomposition To quantify the contribution of revenue tax dispersion to overall wedge dispersion, we take logarithms of Equation 21 and decompose its



(a) Variance Decomposition of Log MRPs



(b) Contribution of Revenue Tax Dispersion

Figure 5: Variance Decomposition of Log MRPs and the Role of Revenue Taxes Over Time (Revenue Weighted)

Note (a): MRPs are calculated at the firm-country-year level. Variances are computed across these units each year, weighted by their share of total global revenue in that year. *Note (b):* "Share 1" is the direct share of tax variance in total wedge variance. "Share 2" includes the share of tax variance plus the covariance term. See Equations 23 and 24.

variance across firm-country units:

$$\begin{aligned}\mathrm{Var}(\log(MRP_{\mathrm{Total},fi})) = &\mathrm{Var}(\log(MRP_{AT,fi})) + \mathrm{Var}\left(\log\left(\frac{1}{1-\tau^V_{fi}}\right)\right) \\ &+ 2\,\mathrm{Cov}\left(\log(MRP_{AT,fi}), \log\left(\frac{1}{1-\tau^V_{fi}}\right)\right)\end{aligned} \tag{22}$$

We use two measures to assess the contribution of revenue taxes to the total dispersion. First we consider the fraction of total MRP variance directly explained by the variance of the revenue tax term:

$$\mathrm{Share}_1 = \frac{\mathrm{Var}(\log(\frac{1}{1-\tau^V_{fi}}))}{\mathrm{Var}(\log(\mathrm{MRP}_{\mathrm{Total},fi}))}. \tag{23}$$

Second, we consider the fraction explained by the revenue tax term plus its covariance with the non-revenue-tax distortions:

$$\mathrm{Share}_2 = \frac{\mathrm{Var}(\log(\frac{1}{1-\tau^V_{fi}})) + \mathrm{Cov}\left(\log(\mathrm{MRP}_{\mathrm{AfterTax},fi}), \log(\frac{1}{1-\tau^V_{fi}})\right)}{\mathrm{Var}(\log(\mathrm{MRP}_{\mathrm{Total},fi}))}. \tag{24}$$

Share_1 gives a conservative estimate, while Share_2 attributes part of the interaction term to taxes, which is common in such decompositions.

Figure 5 illustrates this variance accounting for the period 1930-2023. Panel (a) of the Figure decomposes the variance of the log of the MRPs, our proxy for overall misallocation. The total variance (red line) exhibits a clear upward trend over the sample period, indicating a worsening of misallocation. Critically, the variance of the revenue tax term (black dashed line) not only constitutes a substantial portion of this total variance but also mirrors its upward trajectory, particularly before 2015. In contrast, the variance of non-tax wedges (blue line) remains largely flat. Importantly, the covariance between tax and non-tax terms (orange dotted line) is quantitatively small which implies that the tax and non-tax components contribute largely independently to the total variance.

Panel (b) of Figure 5 presents measures of the contribution of revenue tax dispersion to the total variance. The "Success_1" measure (red line), represents the direct share of tax variance, and reveals that revenue taxes account for a dominant portion of total MRP variance, averaging 62 percent after 1970 and itself trending upwards for a significant part of the sample. The "Success_2" measure (blue line), which includes the covariance term, corroborates this with an average contribution of 53% after 1970.

This accounting suggests that overall misallocation is substantial and has been increasing and that revenue tax dispersion is a key statistical correlate. However, to quantify the aggregate output and TFP consequences of these distortions and evaluate any potential policy reforms, we turn to counterfactual exercises using our calibrated structural model in the next section.

6 Results and Counterfactuals

To quantify the aggregate impact of these distortions, and to quantify the role played by taxation, we perform three counterfactuals for each year using our calibrated structural model. These experiments aim to measure the output gains from removing wedge *dispersion* by equalizing wedges to a common constant across firms and countries. Note that the resulting percentage output gains are independent of the *level* of this constant, as they stem from improved *relative* resource allocation. For simplicity, and without loss of generality, we will set this common benchmark to zero.

Alternatively, our counterfactuals could hold global output constant and calculate the required reduction in aggregate inputs (capital and labor, scaled proportionally).

This perspective is particularly relevant given concerns about climate change and the desirability of increasing aggregate oil production. However, as this approach yields nearly identical productivity gains in our model, we proceed with the current setup focused on output gains. The three counterfactuals are:

1. **Full Efficiency (No Wedge Dispersion):** We eliminate all dispersion in output and capital wedges by setting $\tau_{fi}^{Y} = 0$ and $\tau_{fi}^{K} = 0$ for all firm-country units fi. This scenario quantifies the total potential output gain from a first-best allocation of resources within the global oil industry, given observed aggregate inputs.

2. **No Revenue Tax Dispersion:** We remove dispersion in revenue taxes by setting the revenue tax wedge $\tau_{fi}^{V} = 0$ for all firm-country units, while keeping other non-revenue-tax output distortions (τ_{fi}^{W}) and capital wedges (τ_{fi}^{K}) at their baseline calibrated levels. This isolates the misallocation attributable specifically to the dispersion in revenue taxes.

3. **No Revenue Tax Dispersion (Intensive Margin Only):** We again eliminate revenue tax dispersion by setting $\tau_{fi}^{V} = 0$ for all firm-country units. However, to isolate the impact on the intensive margin of production, we simultaneously adjust the fixed development costs ω_{fi} by introducing a counterbalancing wedge τ_{fi}^{ω}. This new wedge τ_{fi}^{ω} is chosen such that the extensive margin decision (the number of operational fields for each firm-country unit) remains unchanged from the baseline. This experiment highlights the extent to which revenue tax dispersion operates through the choice of input quantities for existing fields, as opposed to the decision to develop new fields.

For each counterfactual, we re-solve the model for the new general equilibrium, including adjustments in the world oil price p and wage rate w that ensure factor markets clear for the given global supplies of capital and labor. The results are typically presented as the percentage gain in global oil output (or, for given fixed global inputs, equivalently productivity) relative to the baseline observed in the data.

Figure 6 plots the potential output gains from these counterfactuals over the period 1930-2023. The red line (Full Efficiency) shows that the costs of overall misallocation are substantial and, for a large part of the sample, increasing. Eliminating all distortions could, in many recent years, more than double global oil output. This

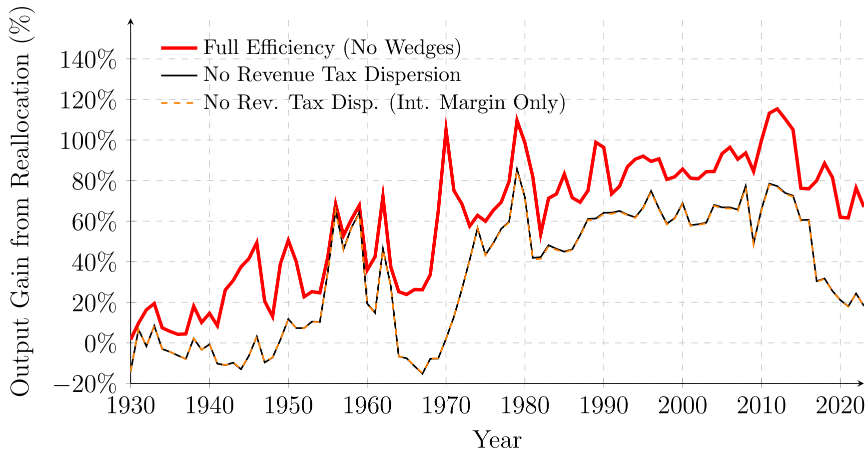


Figure 6: Potential Output Gains from Counterfactual Reforms, 1930-2023. The figure shows the percentage increase in global oil output relative to the baseline under three scenarios: elimination of all wedges, elimination of revenue tax dispersion, and elimination of revenue tax dispersion affecting only the intensive margin.

aggregate gain is similar to that found under full efficiency for India and China in Hsieh and Klenow (2009). A key novel finding here is shown in the black line (No Revenue Tax Dispersion). This reveals that a significant fraction of these potential gains can be achieved simply by eliminating the dispersion in revenue taxes. That is, revenue tax heterogeneity is the primary driver of the observed misallocation and increasing tax dispersion has contributed to rising misallocation. Furthermore, the orange dashed line (No Revenue Tax Dispersion, Intensive Margin Only) is nearly indistinguishable from the black line. The vast majority of the misallocation induced by revenue tax dispersion operates through the intensive margin by distorting firms' choices of capital and labor for already operational fields rather than distorting which fields are developed.

Table 1 summarizes our key findings by averaging the potential output gains over two distinct periods. Several interesting results emerge. First, misallocation in the global oil industry is indeed large and has been growing. Over the entire 1930-2023 period, efficient reallocation of resources could have increased global output by an average of 60.6%. This figure rises to a striking 85.1% for the more recent 1970-2023 period, underscoring the increasing severity of the problem. Second, heterogeneity in revenue taxes has become the dominant factor behind this misallocation. Eliminating revenue tax dispersion alone would have captured, on average, 55.7% of the total

Table 1: Misallocation Costs and the Role of Revenue Taxes: Summary by Period

Period	Counterfactual Output Gains (%)			Contribution (%)	
	(1) No Distortions	(2) No Rev. Tax Disp.	(3) No Rev. Tax Disp. (Int. Margin Only)	of Taxes ((2)/(1))	of Int. Margin ((3)/(2))
1930-2023	60.62	33.77	33.53	**55.70**	**99.30**
1930-1969	30.52	7.00	6.96	**22.94**	**99.38**
1970-2023	85.13	56.21	55.83	**66.03**	**99.33**

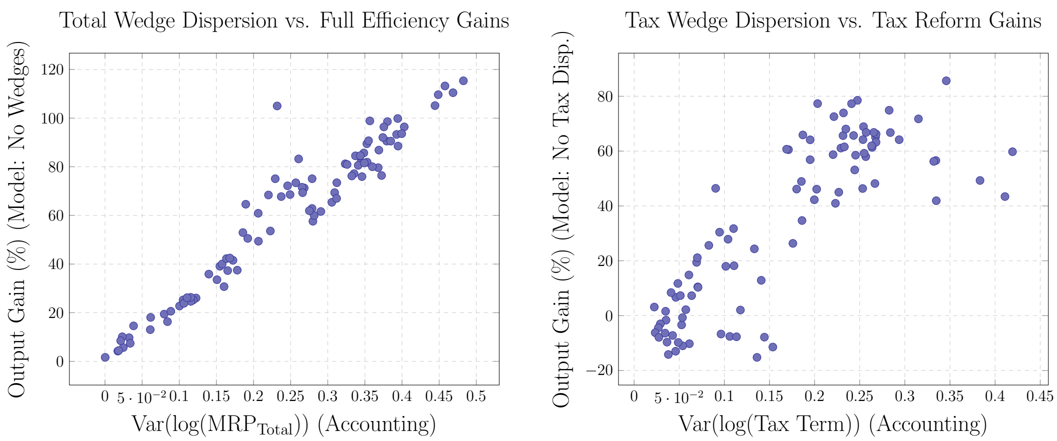
Note: Column (1) shows average output gains if all distortions (τ^Y_{fi}, τ^K_{fi}) are set to zero. Column (2) shows gains if only revenue tax dispersion is removed ($\tau^V_{fi} = 0$). Column (3) shows gains if revenue tax dispersion is removed but the extensive margin is held constant via an offsetting exploration cost wedge τ^ω_{fi}. Contributions are calculated as the ratio of relevant gains. Data from authors' calculations based on model simulations.

potential gains over 1930-2023, and this share increases to 66.0% in the 1970-2023 period. This is a remarkably high explanatory power for a single, observable policy distortion. Third, the impact of these revenue tax distortions is overwhelmingly channeled through the intensive margin. Across all periods, more than 99% of the output gains from eliminating revenue tax dispersion are realized even when the extensive margin of field entry is held constant. This implies that policies aimed at regulating new field development may miss the primary channel through which tax-induced inefficiencies arise in this sector.

These counterfactuals paint a clear picture: the global oil industry is characterized by very significant resource misallocation, a large share of which can be traced to the large and rising dispersion of revenue taxes across countries and firms. The primary impact of these taxes appears to be the distortion of input choices at existing fields, rather than barriers to entry for new ones.

Linking Accounting Dispersion to Modeled Output Gains While our structural model provides quantitative estimates of misallocation costs, it is instructive to connect these back to the simpler accounting measures of wedge dispersion discussed in Section 5. If the variance of MRPs (our accounting proxy for misallocation) truly reflects economically significant distortions, we should expect it to correlate with the output gains predicted by our model when those distortions are removed.

Figure 7 explores this connection. Panel (a) plots the year-by-year variance of the log total MRP (weighted by revenue, from our accounting exercise) against the potential output gain from achieving full efficiency (eliminating all wedges, from our model counterfactual). Panel (b) similarly plots the variance of the log revenue tax



(a) Total Wedge Dispersion

(b) Revenue Tax Wedge Dispersion

Figure 7: Linking Accounting Wedge Dispersion to Modeled Output Gains. Each point represents a year. Panel (a) plots total MRP dispersion against modeled gains from full efficiency. Panel (b) plots revenue tax term dispersion against modeled gains from eliminating tax dispersion.

Note: Data for accounting dispersion measures (variances) are revenue-weighted annual calculations. Modeled output gains are from the counterfactual experiments described in this section. The data file 'Images/MisallocationandCorrectVars.csv' combines these annual series.

term against the potential output gain from eliminating revenue tax dispersion.

Both panels reveal a strong positive correlation. Years with higher accounting-based dispersion in total wedges (Panel a) or in the revenue tax term specifically (Panel b) are systematically associated with larger potential output gains from efficient reallocation in our structural model. This correlation suggests that the distortions captured by our simpler accounting measures are indeed economically significant and manifest as substantial aggregate output losses. Regressing the modeled output gains on the accounting dispersion measures confirms this visual impression. For instance, a 0.01 unit increase in the variance of the log total MRP is associated with an approximate 2.35 percentage point increase in the potential output gain from achieving full efficiency. Similarly, a 0.01 unit increase in the variance of the log revenue tax term is associated with an approximate 2.45 percentage point increase in the output gain from eliminating revenue tax dispersion. These relationships underscore the real economic costs implied by the observed heterogeneity in tax policies.

In sum, our model-based counterfactuals solidify and quantify the insights from the accounting exercises. They demonstrate that misallocation, driven to a large extent by dispersed revenue taxes acting primarily on the intensive margin, imposes a significant burden on the productivity of the global oil industry.

Alternative Approach: Net Present Value of Projects Our baseline model assumes a sequence of static, year-by-year optimization problems. While aggregating to the firm-country level helps smooth potential year-to-year volatility in individual field operations (e.g., alternating phases of exploration and production), one might be concerned that this static framing does not fully capture the long-term investment nature of oil field development, where initial costs are weighed against a stream of future revenues and distortions might be front or back-loaded over a project's life.

To address this, we consider an alternative approach where each potential oil asset, characterized by its intrinsic lifetime productivity $\bar{A}$, is modeled as a single, multi-year "project" if developed. The firm's decision to develop an asset with lifetime productivity $\bar{A}$ is then based on whether the expected net present value (NPV) of profits from that entire project is non-negative. The production function and profit maximization are re-cast in terms of the total discounted lifetime values of outputs and inputs for a project of type $\bar{A}$.

Specifically, the problem for firm f in country i considering developing an asset with lifetime productivity $\bar{A}$ (which becomes a project if developed) at the start of the project (period 0) becomes:

$$\begin{aligned}\bar{\pi}_{fi}(\bar{A}) = \max_{\bar{K}_{fi}(\bar{A}),\bar{L}_{fi}(\bar{A})} \Bigg(& \bar{P}_0 \bar{Y}_{fi}(\bar{A})(1-\bar{\tau}^Y_{fi}) \\ & - \bar{W}_0 \bar{L}_{fi}(\bar{A}) - \bar{R}_0 \bar{K}_{fi}(\bar{A})(1+\bar{\tau}^K_{fi}) \Bigg) - \bar{\Omega}_{fi} \geq 0,\end{aligned} \tag{25}$$

where $\bar{Y}_{fi}(\bar{A}) = \bar{A} \cdot (\bar{K}_{fi}(\bar{A}))^{\gamma}(\bar{L}_{fi}(\bar{A}))^{\alpha}$. Here, $\bar{A}$ is the intrinsic lifetime productivity of the project derived from the asset. The terms $\bar{K}_{fi}(\bar{A})$ and $\bar{L}_{fi}(\bar{A})$ represent the total lifetime (or effective NPV) units of capital and labor, respectively, employed for a project of type $\bar{A}$, resulting in total lifetime output $\bar{Y}_{fi}(\bar{A})$. The effective NPV prices for output, labor, and capital at the project's inception are denoted by $\bar{P}_0, \bar{W}_0$, and $\bar{R}_0$. The terms $\bar{\tau}^Y_{fi}$ and $\bar{\tau}^K_{fi}$ are the effective lifetime output and capital wedges for firm f in country i, assumed to be uniform for all projects undertaken by that firm-country unit, reflecting the discounted impact of all taxes and distortions over

Table 2: Misallocation and the Role of Revenue Taxes: NPV Approach Sensitivity

Period	Counterfactual NPV Output Gains (%)			Contribution (%)	
	(1) No Distortions	(2) No Rev. Tax Disp.	(3) No Rev. Tax Disp. (Int. Margin Only)	of Taxes ((2)/(1))	of Int. Margin ((3)/(2))
1930-2023	29.36	15.38	15.34	**52.38**	**99.77**
1970-2023	38.91	23.80	23.73	**61.18**	**99.69**

Note: This table presents results from the Net Present Value (NPV) modeling approach. "NPV Output Gains" refers to the percentage increase in the aggregate NPV of output. Column (1) shows gains if all distortions are removed. Column (2) shows gains if only revenue tax dispersion is removed. Column (3) shows gains if revenue tax dispersion is removed but the extensive margin (number of projects) is held constant. "Full Period" uses NPVs of all cash flows from 1930-2023, discounted/compounded to a 1930 base year. "Post-1970 Period" uses NPVs of cash flows from 1970-2023, discounted to a 1970 base year. Contributions are calculated as the ratio of relevant gains. Data from authors' calculations.

a typical project's duration. Finally, $\bar{\Omega}_{fi}$ is the NPV of all fixed development costs, also assumed to be the same for any project undertaken by firm f in country i.

This setup allows for periods of negative cash flow within the project, as long as the overall NPV (represented by $\bar{\pi}_{fi}(\bar{A})$) is positive. The calibration of this model involves using the observed NPVs of revenues, operational expenditures, capital expenditures, and government take, aggregated appropriately to the firm-country level or representative project level, to inform these lifetime parameters and wedges. All cash flows are discounted using the real annual rate of 3.7%, which is consistent with the real interest rate component used in our main calibration.

To close this version of the model, we assume that firms draw these lifetime productivities $\bar{A}$ from a distribution $\bar{G}_{fi}(\bar{A})$. The aggregate NPV of global "effective" capital and labor (i.e., $\sum_{f,i}\int_{\bar{A}_{fi}^{\text{cutoff}}}^{\infty}\bar{K}_{fi}(\bar{A})d\bar{G}_{fi}(\bar{A})$ and similarly for labor) would sum to their observed global NPV counterparts, allowing us to infer the implicit effective prices $\bar{P}_0, \bar{W}_0, \bar{R}_0$. We then repeat our counterfactual exercises.

To assess the sensitivity of our findings to this NPV framework and the long historical period, Table 2 presents the counterfactual results under two scenarios for the NPV calculation. The first row shows results when project NPVs are calculated over the full 1930-2023 period, with all cash flows brought to a 1930 base year. The second row restricts the analysis to cash flows from 1970-2023, with NPVs calculated relative to a 1970 base year. In both scenarios, the counterfactuals measure the potential percentage increase in the aggregate NPV of output if resources (total lifetime inputs, held constant) are reallocated efficiently.

The results in Table 2 indicate that our main conclusions are robust to these variations in the NPV approach. For the full period (1930-2023), eliminating all distortions could increase the aggregate NPV of output by 29.4%. Eliminating revenue tax dispersion alone accounts for 52.4% of this potential gain, and 99.8% of that tax-related gain is attributable to the intensive margin. When focusing on the more modern post-1970 period, the potential gain from full efficiency is 38.9%. Revenue tax dispersion explains an even larger share (61.2%) of this gain, with the intensive margin again being overwhelmingly dominant (99.7%).

Thus, even when adopting a lifetime project perspective and considering different historical windows for the NPV calculation, the core findings persist: revenue tax dispersion is a major driver of misallocation in the oil industry, and its impact is primarily felt through distortions to the intensive margin of existing projects.

7 Conclusion

This paper quantified the degree of misallocation in the global oil industry and assessed the role of fiscal policy, specifically the dispersion in revenue taxes, in driving this misallocation. Using a newly constructed dataset on oilfield-level operations and a model-based approach that directly incorporates these observed tax rates, we find substantial effects of these dispersed tax policies on productivity.

Our analysis yields three main conclusions. First, misallocation in the oil sector is severe: efficient reallocation of capital and labor could increase global output by approximately 61% on average over our 1930-2023 study period. Second, dispersion in revenue tax rates across fields accounts for over half of this misallocation—a significantly larger share than attributed to other specific distortions in prior literature—establishing fiscal policy as a key driver of productivity differences. Finally, the impact of this tax dispersion operates primarily through the intensive margin—distorting investment and extraction at existing fields—rather than the extensive margin of field development.

These findings represent a first step in understanding the aggregate consequences of varied fiscal regimes in the extractive sector. A promising direction for future work is to explicitly model the strategic interactions between firms and governments in a dynamic setting, potentially exploring how tax competition contributes to the observed dispersion. As global energy markets evolve, further research will be crucial to assess how fiscal policies can be designed to promote efficiency while ensuring resource-rich countries can equitably share in the gains from production.

Appendix

A Calibration Details for Productivity Distributions and Development Costs

This appendix provides the detailed methodology for the second stage of our calibration procedure, where we determine the parameters of the intrinsic asset productivity distribution, $G_{fi}(A)$, and the fixed development costs, ω_{fi}, for each firm f in country i (for each year t, though the time subscript is omitted here for brevity). As stated in the main text, we assume asset productivities A are drawn from a firm-country specific Weibull distribution, $G_{fi}(A) \sim \text{Weibull}(\theta_{fi}, \mu_{fi})$.

A.1 The Challenge: Model Simplification vs. Data Richness

Our theoretical model (Section 2) assumes that the firm-country average distortions, τ^Y_{fi} and τ^K_{fi} (derived in Section 4), apply uniformly to all fields operated by firm f in country i. This simplification is crucial for tractability. However, our granular Rystad dataset provides rich information at the individual operational field level (indexed by j), revealing that both underlying (intrinsic) productivities and specific distortions can vary significantly even *within* a given firm-country unit. Our calibration strategy must bridge this gap, ensuring that the parameters of our simpler model effectively capture the consequences of this real-world heterogeneity.

A.2 Constructing Empirical Field-Level Productivities

Observed Intrinsic Productivity (A^{obs}_{fij}) The first step is to derive an empirical measure of "observed intrinsic productivity" for each operational field j of firm f in country i (for year t, though t is omitted for brevity), denoted A^{obs}_{fij}. This measure aims to be an empirical counterpart to the intrinsic productivity A drawn from the distribution $G_{fi}(A)$ in our continuous theoretical model, where field output is $Y_{fij} = AK^{\gamma}_{fij}L^{\alpha}_{fij}$.

To construct A^{obs}_{fij}, we start by calculating a field-level productivity measure from observed data. Using observed real revenue $p_tY^{\text{obs}}_{fij}$ (where p_t is the aggregate oil price and Y^{obs}_{fij} is physical output), the constructed real production capital stock K^{obs}_{fij}, and real operational (labor) expenditures $w_t L^{\text{obs}}_{fij}$ (where w_t is the aggregate wage rate and L^{obs}_{fij} is physical labor), a "naive" productivity estimate for field j is:

$$A_{\text{naive},fij} = \frac{p_t Y^{\text{obs}}_{fij}}{(K^{\text{obs}}_{fij})^{\gamma}(w_t L^{\text{obs}}_{fij})^{\alpha}}. \tag{26}$$

This $A_{\text{naive},fij}$ is proportional to the field's physical TFP, scaled by p_t/w_t^{α}. Since p_t and w_t are aggregate prices common to all fields in year t, this common scaling does not affect relative productivity comparisons or wedge calculations. $A_{\text{naive},fij}$ reflects not only the intrinsic productivity of field j (scaled by these prices) but also effects related to the firm operating a discrete number of fields, N_{fi}. Our theoretical model, with its continuum of fields, implicitly embeds how a firm's resources or managerial capacity might be "spread thin" across operations, particularly through the decreasing returns to scale at the field level ($1 - \alpha - \gamma > 0$). If a firm operating N_{fi} discrete fields faces such span-of-control or complexity costs, effectively scaling down the measured productivity of each field by a factor related to N_{fi} (e.g., by $(1/N_{fi})^{1-\alpha-\gamma}$ if a scarce managerial input is spread across fields), then $A_{\text{naive},fij}$ would be $A^{\text{scaled}}_{\text{true},fij} \times (1/N_{fi})^{1-\alpha-\gamma}$. Here, $A^{\text{scaled}}_{\text{true},fij}$ is the fundamental price-scaled productivity analogous to the A in our continuous model.

To recover an estimate of this fundamental (price-scaled) productivity, $A^{\text{scaled}}_{\text{true},fij}$, from our discrete data, we adjust $A_{\text{naive},fij}$ to "undo" this scaling effect. We define the **observed intrinsic productivity** A^{obs}_{fij} as:

$$A^{\text{obs}}_{fij} = A_{\text{naive},fij} \times (N_{fi})^{1-\alpha-\gamma} = \left(\frac{p_t Y^{\text{obs}}_{fij}}{(K^{\text{obs}}_{fij})^{\gamma}(w_t L^{\text{obs}}_{fij})^{\alpha}} \right) \times (N_{fi})^{1-\alpha-\gamma}. \tag{27}$$

Here, N_{fi} is the number of active fields for firm f in country i (taken from Rystad data for each firm-country-year unit). This A^{obs}_{fij} is our data-based estimate of the fundamental (price-scaled) productivity of field j, intended to be directly comparable to a draw A from the model's $G_{fi}(A)$ distribution. This adjustment ensures that the productivity distributions we calibrate are distributions of fundamental field efficiencies (implicitly scaled by p_t/w_t^{α}), consistent with the theoretical underpinnings where the aggregation over the continuum implicitly handles the "scale of operations" effects. This A^{obs}_{fij} then serves as the basis for calculating the empirical effective productivity A^{eff}_{fij} (Equation 28), which further accounts for relative field-level distortions.

The aggregate prices p_t and w_t are determined in general equilibrium (Equations 9-10). While their specific levels influence the absolute level of A_{fij}^{obs}, they do not affect the relative dispersion of A_{fij}^{obs} across fields within a year, nor the calculation of wedges, which are based on ratios.

Data Cleaning for A_{fij}^{obs} calculation: Several data cleaning steps are performed. We exclude observations with non-positive revenues, oil production, or operational expenditures. To mitigate the influence of extreme outliers on subsequent calculations (especially wedge derivations that feed into effective productivity), the top and bottom 1% of key financial ratios (e.g., ratios of opex or capex to revenue) are trimmed based on their oil-production weighted distribution across all firm-field-year observations. Specific entity-year combinations identified as anomalous (e.g., Mazarine Energy for ROU in 2011-2012 due to data irregularities) are also excluded from this part of the analysis.

Empirical Effective Productivity (A_{fij}^{eff}) With A_{fij}^{obs} for each field, we next construct the *empirical effective productivity*, A_{fij}^{eff}. This is the crucial measure that our model's intrinsic productivity distribution $G_{fi}(A)$ will be calibrated against. A_{fij}^{eff} adjusts A_{fij}^{obs} to account for how field j's specific, observed distortions deviate from the firm-country average distortions (τ_{fi}^Y, τ_{fi}^K) that are assumed in our model.

The rationale is as follows: our model (Equation 8) implies that aggregate TFP, $\mathcal{A}_{fi}$, is a generalized mean of the *intrinsic* productivities A because the distortions τ_{fi}^Y, τ_{fi}^K are assumed constant across all fields of firm f in country i. If, hypothetically, field-level distortions $\tau_{fij}^Y, \tau_{fij}^K$ did vary, then aggregate TFP would instead be a generalized mean of an "effective" productivity $A_{fij}^{\text{eff}} = A_{fij}^{\text{model}} \times \text{AdjustmentFactor}_{fij}$, where A_{fij}^{model} is the true underlying productivity and the adjustment factor accounts for $\tau_{fij}^Y, \tau_{fij}^K$ relative to the firm-country averages τ_{fi}^Y, τ_{fi}^K.

To ensure our model's $G_{fi}(A)$ (which represents A^{model}) implicitly accounts for these real-world, field-level distortion variations, we target its moments to an empirical construct that mirrors this theoretical A_{fij}^{eff}. Thus, for each observed operational field j, we calculate:

$$A_{fij}^{\text{eff}} = A_{fij}^{\text{obs}} \times \underbrace{\left(\frac{1-\tau_{fij}^Y}{1-\tau_{fi}^Y}\right)^{\alpha+\gamma} \left(\frac{1+\tau_{fi}^K}{1+\tau_{fij}^K}\right)^{\gamma}}_{\text{Relative Distortion Adjustment Factor}}. \quad (28)$$

Here, τ^Y_{fij} and τ^K_{fij} are the observed distortions specific to field j (calculated using field-level data pY^{obs}_{fij}, wL^{obs}_{fij}, K^{obs}_{fij} in a manner analogous to Equations 12 and 13), while τ^Y_{fi} and τ^K_{fi} are the firm-country average distortions.

The intuition is that A^{eff}_{fij} measures what field j's productivity contribution would be if its idiosyncratic distortions were neutralized and it operated under the firm-country average distortion assumed by our model. It reflects the combined heterogeneity from both underlying intrinsic productivity (A^{obs}_{fij}) and these relative distortions. This A^{eff}_{fij} is the appropriate empirical target for calibrating our model's $G_{fi}(A)$.

A.3 Moment Matching Calibration Strategy

We calibrate the three parameters for each firm f in country i—the fixed development cost ω_{fi} (which, via Equation 4, determines the model's entry threshold $\bar{A}_{fi}$) and the Weibull parameters θ_{fi} (shape) and μ_{fi} (scale)—by matching three moments derived from our model to their empirical counterparts constructed from the distribution of A^{eff}_{fij} values for all active fields of firm f in country i.

The three conditions are:

1. **Entry Threshold (identifies ω_{fi} via $\bar{A}_{fi}$):** The model's minimum productivity threshold for entry, $\bar{A}_{fi}$, is set to align with the lowest observed empirical effective productivity among active fields, less a small positive constant ε (e.g., $\varepsilon = 0.001 \times \text{median}(A^{\text{eff}}_{fij})$ to avoid issues with zero or negative minimums if data is noisy):

$$\bar{A}_{fi}(\omega_{fi}, T_{fi}, \delta) = \min_{j \in \text{active fields for (f,i)}} (A^{\text{eff}}_{fij}) - \varepsilon. \tag{29}$$

   Given T_{fi} (from τ^Y_{fi}, τ^K_{fi}) and δ (from global prices/costs), this equation can be solved for ω_{fi}. This $\bar{A}_{fi}$ serves as the left-truncation point for the model's Weibull distribution $G_{fi}(A; \theta_{fi}, \mu_{fi})$.

2. **Variance of Productivity (helps identify θ_{fi}, μ_{fi}):** The variance of intrinsic productivities A among active fields in the model (i.e., $A > \bar{A}_{fi}$) is matched to the sample variance of the observed A^{eff}_{fij} values for firm f in country i:

$$\text{Var}_{\text{model}}[A|A > \bar{A}_{fi}](\theta_{fi}, \mu_{fi}) = \text{Var}_{\text{data}}[A^{\text{eff}}_{fij}|j \in \text{active fields for (f,i)}]. \tag{30}$$

3. **Level of Aggregate TFP (helps identify θ_{fi}, μ_{fi}):** The model's implied aggregate TFP for firm f in country i, $\mathcal{A}_{\text{model},fi}$, is calculated as the generalized mean of A from the truncated Weibull distribution (see Equation **??**):

$$\mathcal{A}_{\text{model},fi}(\theta_{fi}, \mu_{fi}, \bar{A}_{fi}) = \left(\mathrm{E}\left[A^{\frac{1}{1-\alpha-\gamma}} \,\middle|\, A > \bar{A}_{fi}; \theta_{fi}, \mu_{fi} \right] \right)^{1-\alpha-\gamma}. \tag{31}$$

This $\mathcal{A}_{\text{model},fi}$ is matched to a target aggregate TFP, $\mathcal{A}_{\text{target},fi}$, which is constructed as the corresponding generalized mean of the observed A^{eff}_{fij} values for that firm-country unit:

$$\mathcal{A}_{\text{model},fi}(\theta_{fi}, \mu_{fi}, \bar{A}_{fi}) = \mathcal{A}_{\text{target},fi} \equiv \left(\frac{1}{N_{fi}} \sum_{j=1}^{N_{fi}} (A^{\text{eff}}_{fij})^{\frac{1}{1-\alpha-\gamma}} \right)^{1-\alpha-\gamma}. \tag{32}$$

(Here N_{fi} is the number of active fields for firm f in country i.)

Solving this system of three conditions (Equations 29, 30, and 32) simultaneously for ω_{fi} (or equivalently $\bar{A}_{fi}$), θ_{fi}, and μ_{fi} yields the calibrated parameters for each firm-country-year unit. The expectation in Equation 31 involves an integral over the truncated Weibull PDF, which is handled numerically.

A.4 Interpretation of Calibrated Parameters

As emphasized in the main text, the calibrated intrinsic productivity parameters θ_{fi} and μ_{fi} are "reduced-form." By construction, they ensure that our model, which assumes uniform firm-country distortions, replicates key aggregate features (average productivity level and dispersion) of an economy where true field-level distortions vary. These parameters therefore effectively absorb the combined impact of true underlying asset quality variation and the unmodeled heterogeneity in field-level distortions that is present in the data's A^{eff}_{fij}. This approach allows for a tractable analysis of aggregate misallocation while acknowledging the richness of the microdata.

References

Adamopoulos, Tasso and Diego Restuccia, "The Size Distribution of Farms and International Productivity Differences," *American Economic Review*, June 2014, *104* (6), 1667–1697.

Ahlvik, Lassi and Torfinn Harding, "Resource investments and the timing of tax deductions," 2024.

_, Jørgen Juel Andersen, Jonas Hveding Hamang, and Torfinn Harding, "Quantifying supply-side climate policies," 2022.

Anderson, Soren T, Ryan Kellogg, and Stephen W Salant, "Hotelling under pressure," *Journal of Political Economy*, 2018, *126* (3), 984–1026.

Arezki, Rabah, Frederick van der Ploeg, and Frederik Toscani, "The shifting natural wealth of nations: The role of market orientation," *Journal of Development Economics*, 2019, *138*, 228–245.

Asker, John, Allan Collard-Wexler, and Jan De Loecker, "(Mis)Allocation, Market Power, and Global Oil Extraction," *American Economic Review*, 2019, *109* (4).

BEA, *Implied rates of depreciation of private nonresidential fixed assets* 2020.

Black, Katie Jo, Shawn J McCoy, and Jeremy G Weber, "When externalities are taxed: The effects and incidence of Pennsylvania's impact fee on shale gas wells," *Journal of the Association of Environmental and Resource Economists*, 2018, *5* (1), 107–153.

Bohn, Henning and Robert T Deacon, "Ownership risk, investment, and the use of natural resources," *American Economic Review*, 2000, *90* (3), 526–549.

Brandt, Loren, Johannes Van Biesebroeck, Luhang Wang, and Yifan Zhang, "WTO Accession and Performance of Chinese Manufacturing Firms," *American Economic Review*, September 2017, *107* (9), 2784–2820.

Brown, Jason P, Peter Maniloff, and Dale T Manning, "Spatially variable taxation and resource extraction: The impact of state oil taxes on drilling in the US," *Journal of Environmental Economics and Management*, 2020, *103*, 102354.

Buera, Francisco J, Joseph P Kaboski, and Yongseok Shin, "Finance and Development: A Tale of Two Sectors," *American Economic Review*, August 2011, *101* (5), 1964–2002.

Chen, Chaoran, Diego Restuccia, and Raül Santaeulàlia-Llopis, "Land Misallocation and Productivity," *American Economic Journal: Macroeconomics*, April 2023, *15* (2), 441–465.

Coulomb, Renaud, Fanny Henriet, and Léo Reitzmann, "'Bad'Oil,'Worse'Oil and Carbon Misallocation," 2021.

Cust, James and Torfinn Harding, "Institutions and the location of oil exploration," *Journal of the European Economic Association*, 2020, *18* (3), 1321–1350.

Edmond, Chris, Virgiliu Midrigan, and Daniel Yi Xu, "Competition, Markups, and the Gains from International Trade," *American Economic Review*, October 2015, *105* (10), 3183–3221.

Fajgelbaum, Pablo D., Eduardo Morales, Juan Carlos Suarez Serrato, and Owen M. Zidar, "State Taxes and Spatial Misallocation," *Review of Economic Studies*, 2019, *86* (1).

Garicano, Luis, Claire Lelarge, and John Van Reenen, "Firm Size Distortions and the Productivity Distribution: Evidence from France," *American Economic Review*, November 2016, *106* (11), 3439–3479.

Hamang, Jonas Hveding, "Economic development and known natural resource endowment: Discovery rate differentials of oil," *Journal of Development Economics*, 2024, p. 103290.

Hsieh, Chang-Tai and Peter J. Klenow, "Misallocation and Manufacturing TFP in China and India," *Quarterly Journal of Economics*, 2009, *124* (4), 1403–1448.

Johnston, Louis and Samuel H. Williamson, "What Was the U.S. GDP Then?," MeasuringWorth 2025.

Midrigan, Virgiliu and Daniel Yi Xu, "Finance and Misallocation: Evidence from Plant-Level Data," *American Economic Review*, February 2014, *104* (2), 422–458.

Restuccia, Diego and Richard Rogerson, "The causes and costs of misallocation," *Journal of Economic Perspectives*, 2017, *31* (3), 151–74.