Externality Mitigation with Dynamic Fuel Selection*

Emmanuel Murray Leclair†

Department of Economics, Western University

PRELIMINARY AND INCOMPLETE

February 28, 2023

Abstract

Different technologies allow firms to choose over different mix of inputs in production and inputs to vary in their efficiency. One important instance of this choice is over fuels used to provide energy. This poses challenges to the estimation of firms' production function and hampers our ability to evaluate the impact of policies that can induce firms to change their input mix. I propose a dynamic production model with multidimensional energy input (fuel) choices and heterogeneity in fuel productivity. I exploit recent development in the production function estimation literature to identify fuel productivity in the presence of inter-temporal switching between fuel sets, and I estimate the model with a panel of manufacturing establishment from the Indian Survey of Industries between 2009 and 2016. I use my estimates to assess firms' responses to carbon taxation implemented through fuel-specific tax rates, which distort their input choices. I then explore alternative tax designs on fossil fuels that balance public and private trade-offs from dynamic fuel selection. I show that they outperform standard implementation of carbon taxes at mitigating environmental externalities.

*I would like to thank my advisors David Rivers, Sergio Ocampo Diaz and Salvador Navarro for their help throughout the elaboration of this project. I also want to thank Rory McGee, Nourri Najjar, Nirav Mehta, Martin O'Connell, and Pascual Restrepo for their useful comments.

†email: emurrayl@uwo.ca

1 Introduction

Firms' production functions, which dictate how inputs such as labor, capital, and intermediate materials are combined with technology to produce output, play a crucial role in determining how firms respond to input price variation and taxes on inputs. While there is a significant body of literature on the estimation of production functions (Olley and Pakes, 1996; Levinsohn and Petrin, 2003; Ackerberg, Caves and Frazer, 2015; Gandhi, Navarro and Rivers, 2020), these studies often rely on aggregated inputs and may not fully capture the impact of policies on firms' demand for specific, disaggregated inputs which are typically unobserved. One important example of such policy-sensitive inputs are fossil fuels, which provide energy for production but also contribute to negative externalities such as pollution and climate change. This paper aims to provide a more nuanced understanding of the role of fossil fuels in firms' input mix and how they respond to policies that affect their demand such as carbon taxes.

Studying the role of fossil fuels in firms' production processes highlights two important measurement issues. First, energy that firms use in production, which is referred to *realized energy*, is unobserved because it is the outcome of combining fuels with technology. This is in contrast with physical quantity of fuels measured in common heating potential units, which I refer to *potential energy*.[1] The wedge between potential and realized energy underlie differences in the productivity of energy, due in parts to the quality of capital such as furnaces and kilns, and firms' capacity to mitigate energy waste (Christensen, Francisco and Myers, 2022; Gerarden, Newell and Stavins, 2017; Allcott and Greenstone, 2012). Second, it has long been recognized that establishment-level consumption of fossil fuels varies not only at the intensive margin (Joskow and Mishkin, 1977), but also at the extensive margin, through variation in fuel sets, both over time and across firms within narrowly defined industries (Atkinson and Halvorsen, 1976). Extensive margin choices underlie dynamic selection that hampers the evaluation of decisions under counterfactual fuel sets if not properly taken into account.

To address these issues, I propose a dynamic production model with multidimensional energy input (fuel) choices and heterogeneity in fuel-augmenting productivity. The model features two nests of production: an outer nest with capital, labor, intermediate inputs and energy, and an inner nest where firms combine fuels to create energy. Fuel choices are then separated between an inter-temporal fuel set choice subject to fixed costs (e.g. pipeline access, furnaces, kilns, generator) and a within-period relative fuel quantity choice conditional on the fuel set. Consistent with the

[1] The use of potential energy in the context of this paper should not be confused with potential energy in physics.

literature on input complementarity (Broda and Weinstein, 2006), there are gains from variety for having multiple fuels in a set which decrease marginal costs. In the absence of fixed costs, firms would always use all fuels. Fixed cost thus creates a trade-off between a reduction in contemporaneous profits and a decrease in expected future marginal costs.

To identify the within-period component of the production model, I rely on recent development in production function estimation in the presence of unobserved input quantities (Grieco, Li and Zhang, 2016) and factor-augmenting productivity (Demirer, 2020; Zhang, 2019) that exploit optimality and timing of firms' decisions. Identifying the inter-temporal component of the production model is non-trivial because firms likely use knowledge of their own (expected) productivity to make fuel set decisions, which create dynamic selection on unobservables. For example, a firm may use coal this year and oil next year. The decision may be based on a prediction that the firm will be quite successful if it switches to oil for a variety of reasons, some of which are unobserved to the researcher. To deal with this issue, I assume that unobserved productivity for counterfactual fuels has a permanent component that differentiate firms. Identification then proceeds following Arcidiacono and Jones (2003, 2011). I am able to recover all production function parameters, the distribution of fuel-augmenting productivity, and switching costs between fuel sets. This allows me to conduct policy counterfactual that affect firms' fuel choices at both margins and make novel predictions in terms of externality mitigation.

I then estimate the model with the Indian Survey of Industries (ASI), a panel of manufacturing establishments from India between 2009 and 2016. The panel features quantities and prices of disaggregated inputs that plants purchase, standard measures of revenue and costs, as well as plants' location into 775 districts. Every year, all manufacturing plants with more than 100 workers are part of the dataset, as well as a random sample of plants between 10 and 100 workers. Quantities of fuels such as coal, natural gas, oil and electricity are converted into British thermal units (mmBtu), a standard measure of potential energy in the literature (EPA). I narrow the focus on heavy manufacturing industries such as steel, cement and glass production, which are known to be both very energy and coal intensive in India.

I choose to investigate manufacturing plants because most studies of energy productivity have either studied households consumption or electricity generation. Meanwhile, manufacturing activity is an important contributor of pollution, accounting for 37% of global greenhouse-gas emissions (Worrell et al., 2009). I also choose to investigate Indian plants because many plants primarily use coal in energy-intensive industries such as Steel manufacturing and Cement manufacturing.

In this context, switching costs refer to capital that firms must purchase to use specific fuels. For example, recent developments in metal-casting allow steel-producing firms to use an electric arc furnace rather than the traditional coal-powered oxygen furnaces, which combines electricity, natural gas and recycled materials to cast iron within the steel-making process. Switching costs also refer to transportation infrastructure. For example, firms that want to use natural gas either need to pay a fix cost to connect high-pressure transmission pipelines directly to their establishment, or they need to invest in gasification terminals to convert liquified natural gas (LNG) into natural gas. These switching costs are identified from aggregate factors that affected fuel prices such as state-wide reforms to electricity supply and the global oil shock of 2014, as well as the expansion of the natural gas pipeline network between 2009 and 2016.

Preliminary results suggest large and persistent heterogeneity in fuel productivity, consistent with stylized facts about productivity found by Bartelsman and Doms (2000) and Syverson (2011). Moreover, I find that firms with more fuels in their set face a lower marginal cost of energy. I am able to decompose the different factors that contribute to this difference in marginal costs, and I find that it is due to a combination of higher fuel productivity, lower fuel prices, and the gains from variety. Moreover, the gains from variety explain the majority of this difference and are the main motivation behind firms' choice to expand their fuel set over time.

On the policy side, I explore various tax schemes that affect fuel prices to mitigate externality damages from carbon dioxide (CO_2) and related emissions that contribute to climate change. These tax schemes include a Pigouvian carbon tax levied on fossil fuels, a carbon tax on every fuel but natural gas, a tax on coal only, and various non-Pigouvian taxes. These policy counterfactual affect both margin of choices on the mix of fuels (extensive margin) and the relative intensity of each fuel within a mix (intensive margin). These choice margins highlight important trade-offs that a public decision maker ought to consider. On one hand, a policy that incentivizes firms to add a fuel to their set provides a private benefit to firms. Indeed, a larger fuel set allow firms to better allocate fuels in the context of decreasing marginal returns, creating a gain from variety. Moreover, it allow firms to hedge against various shocks. These shocks include persistent fuel price fluctuations which are ubiquitous in this market due to external geopolitical events, as well as supply disruption such as electricity shortages which are prevalent in India (Allcott, Collard-Wexler and O'Connell, 2016; Mahadevan, 2022; Ryan, 2021). On the other hand, this private benefit can be costly to society because it increases firms' energy demand, potentially increasing CO_2 emissions.

In a benchmark model where firms cannot switch between fuel sets, I find that a Pigouvian carbon tax allow firms to internalize marginal externality damages, achieving the first-best allocation, as in (Golosov, Hassler, Krusell and Tsyvinsky, 2014). However, such Pigouvian taxes can be improved when firms are allowed to inter-temporally switch between fuel sets, which creates discontinuity in the planner's objective function. Preliminary results suggest that the optimal tax rate on natural gas relative to coal should be larger than relative marginal externality damages because natural gas is subject to more frequent and persistent price shocks, whereas the price of coal is stable. However, the tax rate on both coal and natural gas relative to oil and electricity should be larger than relative marginal externality damages to disincentives firms to add a new fuel to their set because the social loss overweight the private gain from variety. Quantitatively, I find that the optimal tax rate increases the net present value (NPV) of welfare by 0.08% and decreases the NPV of pollution damages by 10%, in the welfare formulation of Fowlie, Reguant and Ryan (2016). While the change in welfare isn't large relative to a carbon tax, most of the welfare gains come from decreases in externality damages rather than increases in government revenue. Since the tax is assumed to be revenue neutral, any policy that achieves welfare gains through reduction in externality damages is preferable to a policy that achieves the same welfare gains through government revenue (Kotchen, 2022).

Literature and Contribution

I contribute to the longstanding empirical literature on energy input/fuel substitution in manufacturing industries by combining the two canonical approaches of Joskow and Mishkin (1977), who consider fuel switching as a discrete choice between sets of fuels, and Atkinson and Halvorsen (1976), who use a continuous fuel demand approach. I associate these choices with multiple new implications. Indeed, matching the intensive margin of observed fossil fuel consumption has implication for fuel-specific productivity, while matching the extensive margin of observed inter-temporal fuel set choices has implication for the option value that different sets provide, and creates dynamic selection à la Roy (1951). Together, these margins of choice underlie novel predictions for the optimal design of externality taxes on fossil fuels to improve social welfare. Along the way, I contribute to multiple strain of literature.

First, I contribute to the literature on production function estimation (Olley and Pakes, 1996; Levinsohn and Petrin, 2003; Ackerberg et al., 2015; Grieco et al., 2016; Zhang, 2019; Gandhi et al., 2020; Demirer, 2020). I make a methodological contribution by showing identification of a production function with input-augmenting productivity, where some of the inputs are not always used by firms and can change over time, creating dynamic selection. I solve this selection problem by combining the aformentioned literature with methods from the dynamic discrete choice literature in the presence of unobserved heterogeneity (Arcidiacono and Jones, 2003, 2011).

Second I contribute to the literature on energy productivity/efficiency which has put much attention to the consumer/residential sector (Fowlie and Meeks, 2021; Chan and Gillingham, 2015) and the power generation sector (Cicala, 2022; Davis and Wolfram, 2012; Fabrizio, Rose and Wolfram, 2007).[2] Yet, manufacturing activities contribute to 37% of global greenhouse-gas emissions (Worrell, Bernstein, Roy, Price and Harnisch, 2009), and energy productivity improvements through more efficient furnaces and better heat waste management in this sector can help dealing with climate change. While there is a literature on energy efficiency that extends to the industrial sector (Gerarden, Newell and Stavins, 2017; Allcott and Greenstone, 2012), this literature studies the adoption (or lack thereof) of specific physical technologies. Considering the heterogeneous nature of industrial activity, my paper interprets energy productivity from a more general perspective, where technology can be both physical and intangible (such as worker's knowledge) and where energy productivity can be decomposed into relative efficiency of different fuels.

Third, I contribute to the literature investigating gains from variety in the composition of intermediate inputs, which underlie complementarity between inputs (Ramanarayanan, 2020; Goldberg, Khandelwal, Pavcnik and Topalova, 2010; Kasahara and Rodrigue, 2008; Broda and Weinstein, 2006; Romer, 1990; Ethier, 1982). My application of this theory to the consumption of fossil fuels is consistent with the large literature on electricity shortages in India and other developing economies (Allcott, Collard-Wexler and O'Connell, 2016; Mahadevan, 2022; Ryan, 2021). Indeed, Fisher-Vanden, Mansur and Wang (2015) find that substitution towards other inputs is an important margin of adjustment when Chinese firms face electricity shortages.

Fourth, my paper also contributes to the literature investigating the effects of environmental policies on firm-level pollution, and the optimal design of policies aimed at mitigating climate change and other externalities. I relax the canonical assumptions of a pollution function that underlie a uni-dimensional choice of pollution abatement that has been staple in this literature (Copeland and Taylor, 2004; Shapiro and Walker, 2018). I also contribute the very large literature on the social cost of carbon which aims to find the optimal level of a carbon tax (Golosov, Hassler, Krusell and Tsyvinsky, 2014; Hambel, Kraft and Schwartz, 2021; Miftakhova and Renoir, 2021; Dietz, van der Ploeg, Rezai and Venmans, 2021). I complement this literature by studying the optimal relative tax rate across polluting units such as fossil fuels.

[2]It is important to mention that the terms "productivity" and "efficiency" can be used interchangeably in this context.

On the policy side, closest to me is the work of Fowlie, Reguant and Ryan (2016) who quantify the optimal carbon price in cement manufacturing when there is an additional market inefficiency caused by imperfect competition. While the role of imperfect competition in the optimal design of externality taxes has been recognized since Buchanan (1969), I abstract from concerns of imperfect competition to instead focus on the role of dynamic fuel set selection.

Closest to my paper is the work of Scott (2021), who studies how pipeline expansion in the US encourages power plants to switch towards natural gas.

Paper outline

In section 2, I discuss relevant features of the data. In Section 3, I provide evidence on emissions and fuel usage by ASI establishments. In section 4, I elaborate on the model. In section 5, I show identification and estimation of the outer and inner production model. In section 6, I show identification of fixed costs. In section 7, I present optimal taxes on fossil fuels.

2 Data

I use longitudinal data on inputs, location, and emission of manufacturing establishments in India. This allows for a comprehensive understanding of the role of fuels in the production processes of these establishments and the evolution of fuel use over time. I then link this data to India's vast Natural Gas pipeline infrastructure network, providing a unique level of detail and enabling estimation of a rich model of establishment dynamics. The findings therein offer novel insights into the impact of policies, such as carbon taxes, in reducing the negative effects of climate change.

Manufacturing Establishments I use a panel of establishments from the Indian Survey of Industries (ASI) covering 2009-2016 for 300,000 establishment-year observations. The panel ASI is a restricted-use dataset that covers all manufacturing establishments with at least 100 workers, and a representative sample of establishments with less than 100 workers. The sample is stratified at various levels, including number of workers and location. Some of the sampling rules changed over time, and more detailed can be found in Appendix A.1. The ASI features measures of costs and revenues such as inputs, outputs, and prices. In particular, it contains information on prices and quantities of Coal, Oil, Electricity, and Natural Gas, which I convert to million British thermal unit (mmBtu) using standard scientific calculations from the U.S. Environmental Protection Agency (EPA, 2022).

Location I obtain detailed location information from the publicly available version of the ASI. I locate establishments in one of 775 districts across 28 Indian states. This information allows me to relate cross-sectional variation in fuel prices from the panel ASI to spatial variation in transportation costs. Additionally, I map the entire natural gas pipeline network to these districts from public records by the Petroleum and Natural Gas Regulatory Board (PNGRB, 2023), which oversees all pipelines in India. This allows me to capture variation in the fixed costs of using natural gas, and helps explaining fuel set choices, as I show in Section 3.

Emissions To get establishment-level measures of greenhouse gas emissions, I convert units of potential energy (mmBtu) of each fuel into metric tons of carbon dioxide equivalent. Since I focus on heavy manufacturing industries that use fossil fuels for combustion, each mmBtu of fuel releases some quantity of carbon dioxide CO_2, methane CH_4, and nitrous oxide N_2O in the air, which varies by industry based on standard practices in the Indian context Gupta, Biswas, Janakiraman and Ganesan (2019). I then convert emissions of these three chemicals into carbon dioxide equivalent (CO_{2e}) using the Global Warming Potential method (GWP, see Appendix A.4.1). The conversion rate of carbon dioxide, methane and nitrous oxide to carbon dioxide equivalent is the same across establishments and industries.

Deflation Lastly, I follow standard procedures to deflate all production variables in the Indian manufacturing context (Harrison, Hyman, Martin and Nataraj, 2016). Particularly, I deflate output values by industry-specific wholesale price indices (WPI). I deflate material inputs by the aggregate wholesale price index following Martin, Nataraj and Harrison (2017). Capital stock is deflated by the both the WPI for machinery and an India-specific capital deflator from the Penn World Table Feenstra, Inklaar and Timmer (2015). Both measures yield largely equivalent results. Labor spending is deflated using the consumer price index (CPI) to get a measure of wages that reflects the value it provides to consumers.

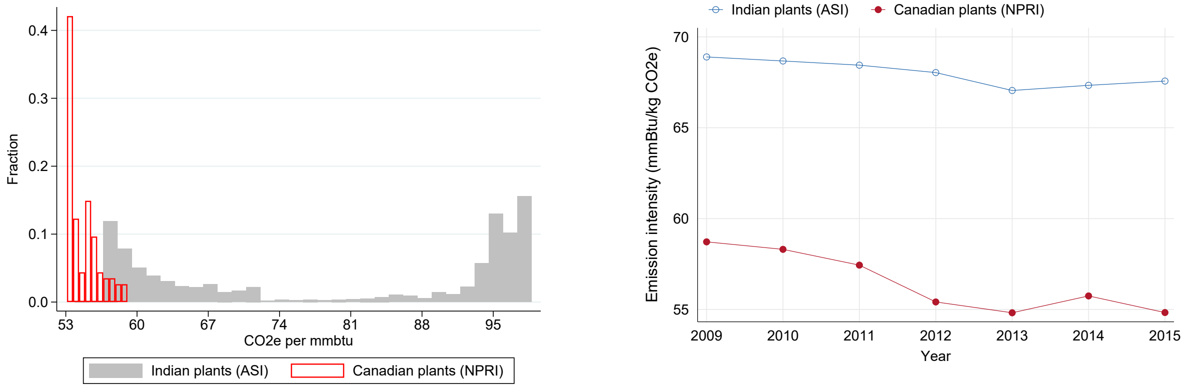
3 Facts about Emissions and Fuels in India

Using these datasets, I highlight a set of facts about fuel usage and carbon emissions, that motivate my choice of India's manufacturing sector to conduct this analysis, and influences modeling choices to capture plants' fuel decisions.

Fact 1: High Pollution Levels from Indian Manufacturing Establishments

Indian manufacturing establishments exhibit a higher level of pollution intensity compared to their counterparts in developed economies. As demonstrated in Figure 1a, half of Indian cement manufacturers emit twice the amount of carbon dioxide per unit of energy compared to the average of Canadian cement manufacturer. This trend is not limited to the cement industry, but prevails across all heavy manufacturing industries that use fuels as primary means of combustion (Figure 1b).

The main reason underlying this gap in emission intensity is the use of different fuels. The cluster of Canadian establishments that emit on average 55 kg of CO_{2e} per $mmBtu$ in Figure 1a reflect establishments that primarily use natural gas. Indeed, switching from coal to gas has been a large contributor to the manufacturing clean-up in developed economies (Rehfeldt, Fleiter, Herbst and Eidelloth, 2020). In contrast, a large portion of Indian plants primarily use coal, which pollutes twice as much as gas. In Figure 2, I show that coal consistently contributes to 40% of all fuels used by Indian Establishments. This prevalence of coal usage among Indian manufacturers explains the cluster of plants that emit on average 95 kg of CO_{2e} per $mmBtu$ in figure 1a.

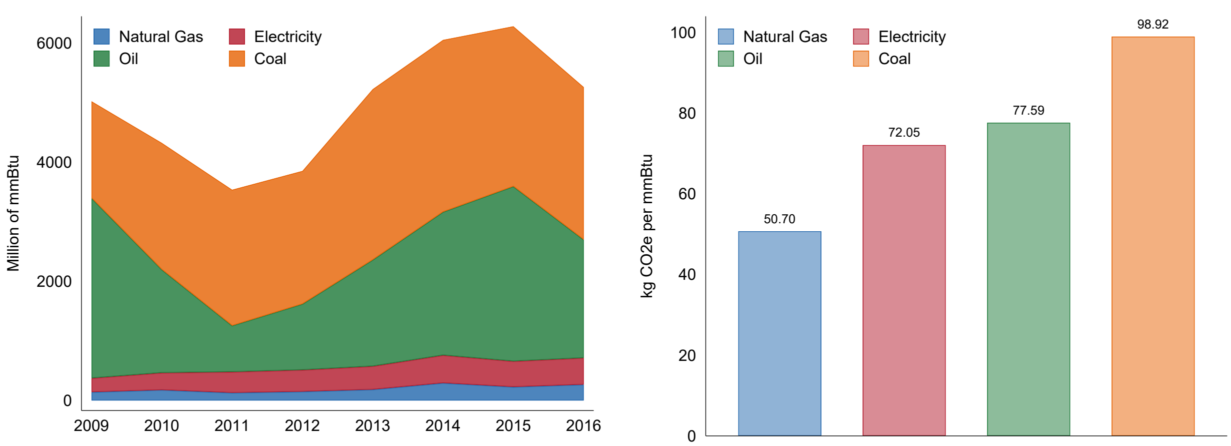


(a) Pollution Intensity - Cement Manufaturing

(b) Evolution of Pollution Intensity - Heavy Manufacturing Industries

Figure 1: Pollution Intensity of Energy in India and Canada ($kg\ CO_{2e}/mmBtu$).

Note: Information from Canadian plants come from the National Pollutant Release Inventory (NPRI) (Government of Canada, 2022). This is a publicly available dataset that records emission of specific pollutants by Canadian manufacturing plants, which I convert into CO_{2e} emissions using the Global Warming Potential (GWP) method. In Figure 1b, I compare the within industry average pollution intensity for 5 heavy manufacturing industries: Pulp & paper, cement, steel, aluminium, and glass.



(a) Total Emissions by Fuel

(b) Average Emission Intensity by Fuel

Figure 2: Comparison of Fuels used by ASI establishments

Note: Figure 2a aggregates across all manufacturing establishments in the ASI by year, and suggests a much lower usage of natural gas compared to coal. Figure 2b shows the average emission intensity of each fuel, where the average is taken across industries according to scientific calculations made by Gupta et al. (2019).

Fact 2: Indian Manufacturing Establishments often Switch Between Fuel Sources

I find that plants operating in narrowly defined industries use different fuel mixes at any given time. Moreover, I find that 40% of all plants add at least one fuel to their mix at some point in the sample, and 40% of all plants drop an existing fuel at some point in the sample. Additionally, I find that plants who switch tend to switch on average two times, which is likely to be an under-estimate of switching because I only observe plants for a maximum of 8 years. See Appendix A.5. Importantly, this isn't a feature of Indian plants, but rather a prevalent feature of fuel consumption in manufacturing across the world. In Appendix A.5, I look at fuel switching in U.S. based plants and find similar results. Below I document two facts that help explaining fuel switching.

	Steel	Casting of Metal	Cement	Glass
Oil, Electricity	51.3	51.5	42.1	53.6
Oil, Electricity, Coal	19.3	23.7	42.00	3
Oil, Electricity, Gas	10.8	12.2	1	31
Oil, Electricity, Coal, Gas	7.4	3.5	1.3	1.2
Other	11	9.2	13.7	11.2

Table 1: Percentage (%) of ASI Establishment That use Different Fuel Sources - Selected Industries

	Adds New Fuel (%)	Drops Existing Fuel (%)
No	58.5	60.4
Yes	41.5	39.6

Table 2: Percentage of unique plants that **add** and **drop** a fuel.

Notes: To construct Table 2, I balanced the panel, keeping establishments that operate in all 8 years between 2009-2016. I did this to get a sense of the prevalence of aggregate fuel switching.

Fact 3: Establishments Increase the Number of Fuels as they age

As firms become older, the average number of fuels in their set rises. See Figure 3. Consistent with a standard firm life-cycle model in which older firms are moe productive and have accumulated capital, making it possible for them to finance fixed costs such as new furnaces to expand their activity. I interpret as the suggestive of the importance of fixed costs for firms' choice of fuel sets. In this context, firms would like to use more fuels, but may not always find it profitable to pay the fixed costs. My model will include many factors explaining why firms want to use more fuels. These include gains from variety, an option value against fuel-specific price shocks and an option value against fuel-specific quantity shortages, all of which are relevant in the Indian context.

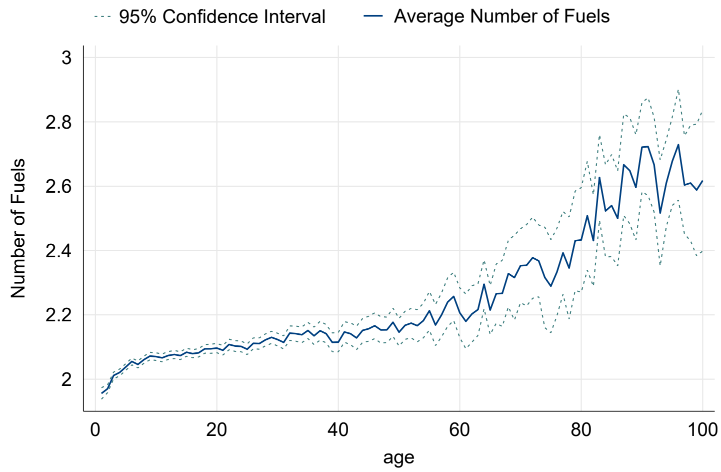


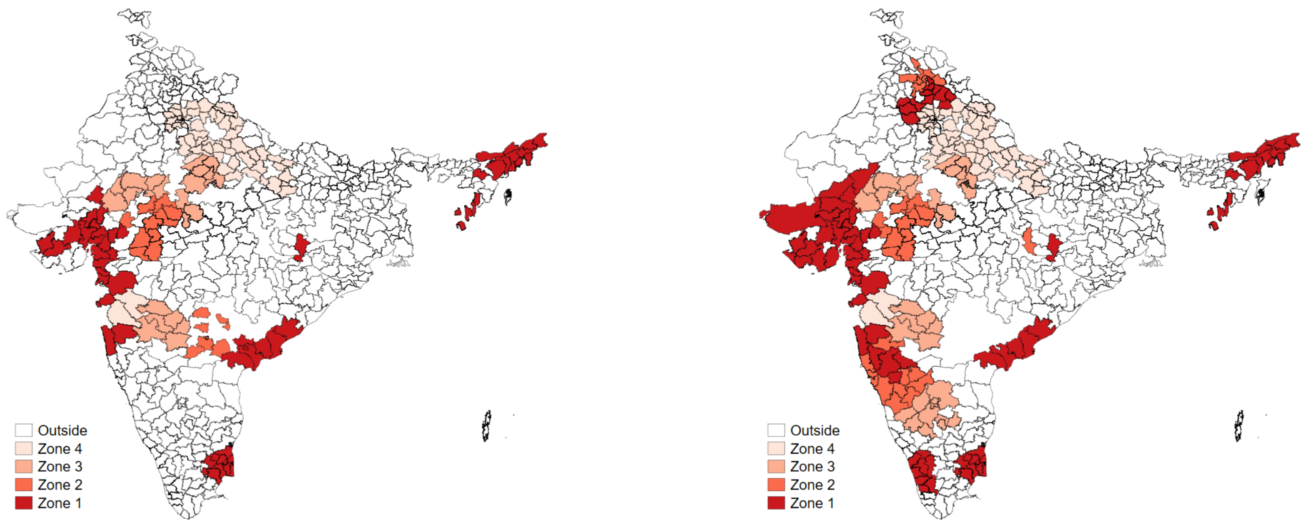
Figure 3: Number of fuels by plant age, average of all ASI plants

Fact 4: Proximity of Establishments to Pipelines Correlates with Gas Usage

Plants located near pipelines have access to the main distribution network through direct connection to transmission pipelines. This access reduces costs compared to those located far from the pipeline network, who need to construct expensive gasification terminals to convert liquified natural gas to its usable form. This is because natural gas can only be transported as a gas through high pressure pipelines. In the Indian context, I investigated the impact of the natural gas pipeline network expansion between 2009 and 2016 on the likelihood of adding natural gas as a fuel source.

I used a simple logit regression where the dependent variable is an indicator for whether a plant in district j added natural gas between year t and $t+1$. The dependent variable of interest is whether the pipepline network expanded in that district between t and $t+1$. The results indicate that an expansion in the pipeline network within a plant's district leads to a 2.2 percentage point increase in the probability of adding natural gas. These results are consistent with Scott (2021) who provides evidence that proximity of power plants to gas pipelines in the U.S. is a critical factor in determining the fixed costs of adding natural gas as a fuel source, which affects the probability that a power plant adds natural gas.

Figure 4: Indian districts by zone of access to natural gas pipelines.



(a) 2009

(b) 2016

Notes: Zone 1 is closest to source and Zone 4 is furthest from source. Zones are defined according to regulations under the Petroleum and Natural Gas Regulatory Board (PNGRB)

Added Natural Gas	(1)	(2)	(3)
Pipeline Expanded	0.013**	0.013**	0.02***
	(0.004)	(0.004)	(0.005)
Industry Fixed effects		Y	Y
District Fixed effects			Y
Observations	128,496	128,496	128,496

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4: Probability of adding natural gas, logit average marginal effects from pipeline expansion between two years.

4 Model

Consistent with the evidence provided so far, I develop and estimate a rich dynamic production model which allows me to quantify firms' choice of fuel sources, both at the intensive and the extensive margin. Decisions in the model are reliant on the channels discussed above, and includes idiosyncratic reasons why establishments use different fuel sources. Particularly relevant are differences in the productivity of using different fuels such as heat management practices and the types of furnaces/kilns that plants operate, and heterogeneity in fuel prices. The model improves upon the literature on externality mitigation at predicting the effect of various policies such as carbon taxes on the emissions of manufacturing establishments.

Each period, firms have access to a set of fuels from a combination of oil, natural gas, coal and electricity. They combine fuels to produce realized energy that goes into the outer nest of production. Each fuel in the firm's set has its own productivity term, and the production model for energy is the same across fuel sets. Then, firms can choose which set of fuels to use for the next period. Since fuels can only be used in conjunction with technological capital such as furnaces and kilns, firms have to pay a fixed cost to use a new fuel, but get a salvage value from dropping an existing fuel. The presence of gains from variety implies that all firms would use all fuels in the absence of fixed costs. Thus, fixed costs create a trade-off between reduction in contemporaneous profits and decreases in expected marginal cost. I first present the structure of production for a given plant in a static setting, and then consider inter-temporal decisions. Throughout the exposition, subscript i refers to a plant and t refers to a year. The analysis is conducted industry-by-industry, so I omit the industry subscript going forward.

4.1 Static Production Model

There are two levels of production which correspond to two nests. The outer nest is a standard CES production function and features Hicks-neutral productivity z_{it}, labor $(L_{it}/\overline{L})$, capital $(K_{it}/\overline{K})$, intermediate inputs $(M_{it}/\overline{M})$ and realized energy $(E_{it}/\overline{E})$ inputs.[3] Following Grieco et al. (2016), the production function is explicitly normalized around the geometric mean of each variable $\overline{X} = \left(\prod_{i=1}^{n}\prod_{t=1}^{T} X_{it}\right)^{\frac{1}{nT}}$.[4]

[3] The particular functional form of the CES is not necessary. Identification works for a large class of production functions.

[4] It has been known for a long time that all CES functions are either implicitly or explicitly normalized around a point (León-Ledesma, McAdam and Willman, 2010). I choose the geometric mean as a normalization point to be consistent with the literature, but the choice of any particular normalization does not carry any meaning beyond mathematical convenience, or lack thereof. Details on the explicit derivation of the CES normalization can be found in the appendix

$$\frac{Y_{it}}{\overline{Y}} = z_{it}\left(\alpha_k\left(\frac{K_{it}}{\overline{K}}\right)^{\frac{\sigma-1}{\sigma}} + \alpha_L\left(\frac{L_{it}}{\overline{L}}\right)^{\frac{\sigma-1}{\sigma}} + \alpha_M\left(\frac{M_{it}}{\overline{M}}\right)^{\frac{\sigma-1}{\sigma}} + \alpha_E\left(\frac{E_{it}}{\overline{E}}\right)^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\eta\sigma}{\sigma-1}} \tag{1}$$
$$s.t. \quad \alpha_L + \alpha_K + \alpha_M + \alpha_E = 1$$

Where $\sigma \geq 0$ is the elasticity of substitution between inputs, and η is the returns to scale. In the outer nest, firms choose the quantity of all inputs given input prices, including realized energy, $\frac{E_{it}}{\overline{E}}$. Then, given the current fuel set $\mathcal{F}_{it} \subseteq \mathbb{F}$, firms combine all fuels available in the set to produce a quantity of realized energy $\frac{E_{it}}{\overline{E}}$ in the inner nest of production.

$$\frac{E_{it}}{\overline{E}} = \left(\sum_{f\in\mathcal{F}_{it}}\left(\psi_{fit}\frac{e_{fit}}{\overline{e_f}}\right)^{\frac{\lambda-1}{\lambda}}\right)^{\frac{\lambda}{\lambda-1}} \tag{2}$$

e_{fit} refers to quantity of fuel f for plant i in year t, and p_{fit} is the corresponding price. λ is the elasticity of substitution between fuels. This parameter plays a crucial role in this model because it determines the magnitude of gains from variety that a firms would get by expanding its fuel set $\mathcal{F}_{it}$. As long as $\lambda > 1$, there are gains from variety. However, the more complements fuels are conditional on being gross substitutes, the larger are the gains from variety. This is because a lower λ implies that marginal products from a given fuel decrease faster, so there are larger marginal gains from switching. In section 3.4, I explore these comparative statics in more details. The production model of realized energy in equation 2 features a fuel-specific productivity term for each firm in each year, which allows for flexible variation in input usage at the intensive margin.

4.2 Static Decisions

Assumption 1-. *Capital is rented flexibly every period at user cost of capital r_{kit}*[5]

Assumption 2-. *Firm takes output price in year t as given and engage in perfect competition (I discuss this assumption in section 3.3)*

For a given set of fuels $\mathcal{F}_{it}$, the firm's static problem is then to maximize profits. To avoid notation clutter, I will define $\tilde{X_{it}} \equiv \frac{X_{it}}{\overline{X}}$ for normalized quantities and $\tilde{p}_{xit} \equiv p_{xit}\overline{X}$ for normalized prices from now on.

[5]This is not needed for identification of the production function, but it simplifies the computation of firms' dynamic choice of fuel set.

$$\max_{K_{it},M_{it},L_{it},\{e_{fit}\}_{f\in\mathcal{F}_{it}}} \Big\{P_tY_{it} - w_{it}L_{it} - r_{kit}K_{it} - p_{mit}M_{it} - \sum_{f\in\mathcal{F}_{it}} p_{fit}e_{fit}\Big\}$$

$$s.t. \quad \tilde{Y}_{it} = z_{it}\Bigg[\alpha_K\tilde{K}_{it}^{\frac{\sigma-1}{\sigma}} + \alpha_L\tilde{L}_{it}^{\frac{\sigma-1}{\sigma}} + \alpha_M\tilde{M}_{it}^{\frac{\sigma-1}{\sigma}} + \alpha_E\Big(\sum_{f\in\mathcal{F}_{it}}(\psi_{fit}\tilde{e}_{fit})^{\frac{\lambda-1}{\lambda}}\Big)^{\frac{\lambda}{\lambda-1}\frac{\sigma-1}{\sigma}}\Bigg]^{\frac{\eta\sigma}{\sigma-1}}$$

The nested structure of production is such that it can be express in two stages:

1. Fuel choices to minimize cost given quantity of realized energy (Inner nest):

Given fuel prices, firms find the combination of fuels that minimizes the cost of producing a given unit of realized energy. Note that fuel prices in mmBtu are observed allowed to vary across plants here. Appendix A.2 discusses the main reason underlying cross-sectional price variation.

$$\min_{\{e_{fit}\}_{f\in\mathcal{F}_{it}}} \Big\{\sum_{f\in\mathcal{F}_{it}} p_{fit}e_{fit}\Big\} \quad s.t. \quad \tilde{E}_{it} = \Big(\sum_{f\in\mathcal{F}_{it}}(\psi_{fit}\tilde{e}_{fit})^{\frac{\lambda-1}{\lambda}}\Big)^{\frac{\lambda}{\lambda-1}} \tag{3}$$

The solution to this problem is an energy cost function $\mathcal{C}(\tilde{E}_{it})$ that satisfies:

$$\begin{aligned}\mathcal{C}(\tilde{E}_{it}) &= \Big(\sum_{f\in\mathcal{F}_{it}}\big(\frac{\tilde{p}_{fit}}{\psi_{fit}}\big)^{1-\lambda}\Big)^{\frac{1}{1-\lambda}}\tilde{E}_{it} \\ &= \tilde{P}_{Eit}\tilde{E}_{it} = \sum_{f\in\mathcal{F}_{it}} p_{fit}e_{fit}\end{aligned}$$

Where the unobserved price of realized energy $\tilde{p}_{E_{it}}$ correspond to a CES price index in fuel prices over productivity. Constant returns in the energy production function implies that the marginal cost of realized energy is the price of realized energy and is constant $MC(\tilde{E}_{it}) = \tilde{p}_{Eit}$.

2. Input choices to maximize profit (outer nest):

Given a cost-minimizing allocation of fuels that produce a quantity of realized energy, firms pay a price p_{Eit} for each unit of realized energy. They take this price as given when choosing quantity realized energy because p_{Eit} is only a function of the optimal *relative* allocation of fuels, not the scale of energy. This is due to the constant returns assumption in equation 2. Firms also take other inputs and ouput prices as given, and maximize profits:

$$
\max_{K_{it},M_{it},L_{it},E_{it}} \left\{ P_t Y_{it} - w_{it} L_{it} - r_{kit} K_{it} - p_{mit} M_{it} - p_{Eit} E_{it} \right\}
$$
$$
s.t. \tilde{Y}_{it} = z_{it} \left[\alpha_K \tilde{K}_{it}^{\frac{\sigma-1}{\sigma}} + \alpha_L \tilde{L}_{it}^{\frac{\sigma-1}{\sigma}} + \alpha_M \tilde{M}_{it}^{\frac{\sigma-1}{\sigma}} + \alpha_E \tilde{E}_{it}^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\eta\sigma}{\sigma-1}} \tag{4}
$$

The solution to this problem is a profit function:

$$
\pi_{it} = \frac{(P_t \overline{Y} z_{it})^{\frac{1}{1-\eta}}}{\tilde{p}_{it}^{\frac{\eta}{1-\eta}}} \left[\eta^{\frac{\eta}{1-\eta}} - \eta^{\frac{1}{1-\eta}} \right]
$$

Where $\tilde{p}_{it}$ is the CES input price index from the outer nest:

$$
\tilde{p}_{it} = \left(\alpha_K^\sigma \tilde{r}_{kit}^{1-\sigma} + \alpha_M^\sigma \tilde{p}_{Mit}^{1-\sigma} + \alpha_L^\sigma \tilde{w}_{it}^{1-\sigma} + \alpha_E^\sigma \tilde{p}_{Eit}^{1-\sigma} \right)^{\frac{1}{1-\sigma}}
$$

4.3 Full model and inter-temporal fuel set choices

Aggregation

All analysis, including welfare evaluations is done at the industry level. In each industry, there is a representative consumer with quasi-linear utility over the total output produced in a given period Y_t and an outside good Y_{0t}. The representative consumer owns the firms and derives income I_t from total profits Π_t. This quasi-linear utility specification is standard in the literature on externality taxation (Fowlie et al., 2016) and allows researchers to use the social cost of carbon (SCC) in dollars to construct externality damages. As such, externality damages and firms profit will act as shift on the representative consumer's aggregate income.

$$
U(Y_t, Y_{0t}) = \max_{Y_t, Y_{0t}} u(Y_t) + Y_{0t} \tag{5}
$$

s.t.

$$
\begin{aligned}
Y_{0t} + P_t Y_t = I_t = \Pi_t \\
Y_t = \int_i Y_{it} di \\
\Pi_t = \int_i \pi_{it} di
\end{aligned}
$$

In each period, firms take output prices as given and engage in perfect competition. In Appendix C.1, I show that identification of this production economy is equivalent to an economy in which firms produce differentiated goods and engage in monopolistic competition, but have constant returns to scale (CRS) technology.[6] However, there is one key distinction. When setting optimal taxes in an imperfectly competitive economy, the government takes into account both the externality and the inefficient allocation caused by imperfect competition (Buchanan, 1969). This assumption allows me to isolate the government's role in correcting the externality rather than the inefficiency caused by imperfect competition. The interaction between externality mitigation and imperfect competition has been studied quantitatively by Fowlie et al. (2016) in the Portland Cement industry.

Inter-temporal fuel set choice

At the beginning of each period, plants start with a set of fuels $\mathcal{F}_{it} \subseteq \mathbb{F}$, observe their hicks-neutral productivity z_{it}, productivity for each fuels $\{\psi_{fit}\}_{f\in\mathcal{F}_{it}}$, and all input prices $\{w_{it}, r_{kit}, p_{mit}, \{p_{fit}\}_{f\in\mathcal{F}_{it}}\}$. Together, these form a set of state variables s_{it}. Firms then take expectation over the evolution of state variables, and choose a fuel set for next period $\mathcal{F}'$ to maximize expected lifetime profits:

$$
V(\mathbf{s}_{it}, \mathcal{F}_{it} \in \mathbb{F}) = \max_{\mathcal{F}'} \Big\{ \pi(\mathbf{s}_{it}, \mathcal{F}_{it}) - \Phi(\mathcal{F}' \mid \mathcal{F}_{it}) + \epsilon_{\mathcal{F}'it} + \beta\, \mathbb{E}[V(\mathbf{s}_{it+1}, \mathcal{F}') \mid s_{it}] \Big\}
$$

Where $\Phi(\mathcal{F}' \mid \mathcal{F}_{it})$ is the net cost of switching from fuel set $\mathcal{F}$ to $\mathcal{F}'$ and $\epsilon_{\mathcal{F}'_{it}}$ allow for unobserved variation in fixed costs for all fuel sets. Fuel switching costs are composed of two terms. First, there are fixed costs of adding a fuel κ_f which underlie the technologies that plants need to buy to use that fuel. Second, there are salvage values of dropping a fuel γ_f that plants obtain by selling technologies. In later sections, I show that the fixed cost to add natural gas κ_g can vary based on plants' distance to the nearest distribution pipeline to account for expansion of the natural gas pipeline network.

[6] In Appendix C.1, I show how aggregation with quasi-linear utility can be adapted to monopolistic competition.

$$\Phi(\mathcal{F}' \mid \mathcal{F}_{it}) = \Big(\underbrace{\sum_{f \in \mathcal{F}_{it}} \mathbb{I}(f \notin \mathcal{F}')\gamma_f}_{\text{value of dropping fuels}} - \underbrace{\sum_{f \notin \mathcal{F}_{it}} \mathbb{I}(f \in \mathcal{F}')\kappa_f}_{\text{cost of adding fuels}} \Big)$$

Since 90% of plants in the dataset always use Electricity and Oil, I assume that the choice set of firms is as follow, where e = electricity, o = oil, g = gas, c = coal:

$$\mathbb{F} = \Big\{(oe); (oec); (oeg); (oecg)\Big\}$$
$$\mathcal{K} = \{\kappa_g, \kappa_c, \gamma_g, \gamma_c\}$$

4.4 Comparative statics: gains from variety and option value

In this section I show why a firm would want to pay a fixed cost to add a new fuel to its set. The price that firms pay for realized energy is a CES price index:

$$p_{\tilde{E}_{it}} = \Big(\sum_{f \in \mathcal{F}_{it}} (\frac{\tilde{p}_{fit}}{\psi_{fit}})^{1-\lambda} \Big)^{\frac{1}{1-\lambda}}$$

Broda and Weinstein (2006) and others show that this CES price index is decreasing in the number of inputs it contains, here $\mid \mathcal{F}_{it} \mid$, as long as inputs are gross substitutes ($\lambda > 1$). This means that absent of fixed costs, all firms would always include all fuels in their set. The intuition underlying the gains from variety and can be understood through decreasing marginal products. Indeed, the energy production function is concave in each inputs, so fuel-specific marginal products are decreasing in fuel quantities. Adding an additional fuel allows to substitute away from the least productive units of existing fuels to the more productive units of the new fuel due to gross substitution ($\lambda > 1$), which in terms increases the marginal product of each existing fuels. The net effect is an overall decrease in the total quantity of fuels required to produce a unit of realized energy $\tilde{E}_{it}$, which decreases marginal costs p_{Eit}. In section 4, I show evidence consistent with this conceptualization of these gains from variety. In appendix C.3, I show how similar comparative statics can be derived from a task-based model for realized energy similar to Acemoglu and Restrepo (2021).

Proposition 1. *Gains from variety: ceteris-paribus, if a fuel set $\mathcal{F}$ is a strict subset of $\mathcal{F}'$ and fuels are gross subsitute ($\lambda > 1$), then the marginal cost to produce energy is higher under $\mathcal{F}$.*
$\mathcal{F} \subset \mathcal{F}' \rightarrow p_{\tilde{E}_{it}(\mathcal{F})} > p_{\tilde{E}_{it}(\mathcal{F}')}$

Proof. Assume not:

$$\Big(\sum_{f\in\mathcal{F}} \big(\frac{\tilde{p}_{fit}}{\psi_{fit}}\big)^{1-\lambda}\Big)^{\frac{1}{1-\lambda}} < \Big(\sum_{f\in\mathcal{F}'} \big(\frac{\tilde{p}_{fit}}{\psi_{fit}}\big)^{1-\lambda}\Big)^{\frac{1}{1-\lambda}}$$

$$0 > \sum_{f\in\mathcal{F}'\setminus\mathcal{F}} \big(\frac{\tilde{p}_{fit}}{\psi_{fit}}\big)^{1-\lambda} > 0 \Rightarrow\Leftarrow$$

$\square$

In addition, by expanding its fuel set, a firm also gains the option value of being able to hedge against negative price shocks and quantity shortages. The following two propositions demonstrate the differential effects of a fuel price increase and a binding quantity shortage on a firm, depending on the size of its fuel set.

Proposition 2. *Option value against positive fuel price shock: ceteris-paribus, if a fuel set $\mathcal{F}$ is a strict subset of $\mathcal{F}'$, an increase in the price of a fuel in both sets will increase marginal costs under $\mathcal{F}$ by a larger amount.* $\mathcal{F} \subset \mathcal{F}' \rightarrow \frac{\partial p_{\tilde{E}_{it}(\mathcal{F})}}{\partial p_{fit}} > \frac{\partial p_{\tilde{E}_{it}(\mathcal{F}')}}{\partial p_{fit}}$

$$\frac{\partial p_{E_{it}(\mathcal{F})}}{\partial \tilde{p}_{fit}} - \frac{\partial p_{E_{it}(\mathcal{F}')}}{\partial \tilde{p}_{fit}} = \Big(\frac{p_{fit}}{\psi_{fit}}\Big)^{-\lambda} \frac{1}{\psi_{fit}} \Big[p^{\lambda}_{\tilde{E}_{it}(\mathcal{F})} - p^{\lambda}_{\tilde{E}_{it}(\mathcal{F}')}\Big] > 0$$

$$\text{if} \quad \mathcal{F} \subset \mathcal{F}' \text{ since } \quad \lambda > 1 \quad \text{and} \quad p_{\tilde{E}_{it}(\mathcal{F})} > p_{\tilde{E}_{it}(\mathcal{F}')} \text{ by Proposition 1}$$

The idea behind Proposition 2 is that a larger set of fuels can act as a form of insurance against negative price shocks, which is particularly relevant in a world where many fossil fuels are susceptible to geopolitical turmoil that can have long-lasting effects on fuel prices. Consider, for instance, the Ukraine-Russia war of 2021, which sent natural gas prices soaring worldwide, or the oil shock of 2014, which resulted in a significant drop in the price of both oil and natural gas (figure 5). In this volatile geopolitical landscape, firms' decisions carry significant weight in terms of how they plan to weather potential future shocks.

The final proposition demonstrates that a larger fuel set allows firms to weather binding quantity shortages of a particular fuel more effectively. This proposition is particularly applicable in the In-

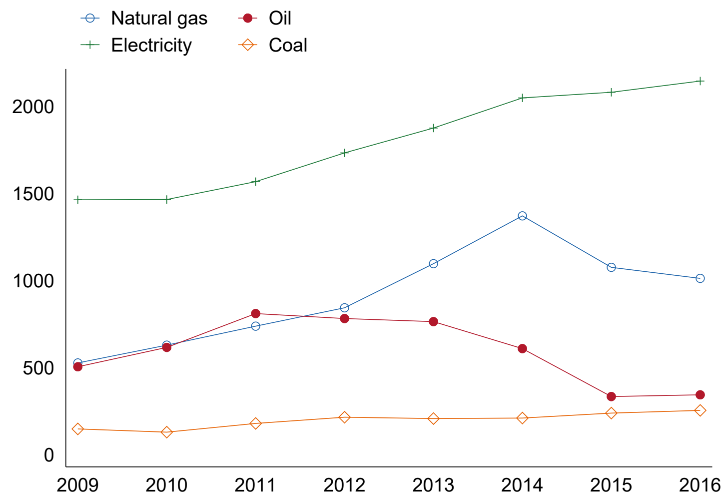


Figure 5: Yearly Median Fuel Prices (INR/mmBtu)

dian context, where the economy frequently experiences disruptions in its electricity supply (Allcott et al., 2016; Mahadevan, 2022; Ryan, 2021) due to financial struggles faced by state-owned utility suppliers, which can lead to the discontinuation of electricity supply (Mahadevan, 2022). In such circumstances, fuel substitution can serve as a means of adjustment for firms to insure themselves against these shortages.

Proposition 3. *Option value against binding fuel shortage: ceteris-paribus, a binding shortage on the quantity of a specific fuel $\overline{e}_f$ will increase the perceived marginal cost to produce energy. Moreover, if a fuel set $\mathcal{F}$ is a strict subset of $\mathcal{F}'$, the increase in marginal costs will be larger under $\mathcal{F}$.* $\quad \mathcal{F} \subset \mathcal{F}' \rightarrow p_{\tilde{E}_{it}(\mathcal{F},\overline{e}_f)} > p_{\tilde{E}_{it}(\mathcal{F}',\overline{e}_f)}$

$$\min_{\{e_{fit}\}_{f\in\mathcal{F}_{it}}} \left\{ \sum_{f\in\mathcal{F}_{it}} p_{fit}e_{fit} \right\} \quad s.t. \quad \tilde{E}_{it} = \left(\sum_{f\in\mathcal{F}_{it}} (\psi_{fit}\tilde{e}_{fit})^{\frac{\lambda-1}{\lambda}} \right)^{\frac{\lambda}{\lambda-1}}$$

$$e_{fit} \leq \overline{e}_f \quad \text{for some } f$$

The Lagrangian can be written as:

$$\mathcal{L} = \sum_{f\in\mathcal{F}_{it}} p_{fit}e_{fit} + \mu_1 \left[\tilde{E}_{it} - \left(\sum_{f\in\mathcal{F}_{it}} (\psi_{fit}\tilde{e}_{fit})^{\frac{\lambda-1}{\lambda}} \right)^{\frac{\lambda}{\lambda-1}} \right] + \mu_2(e_{fit} - \overline{e}_f)$$

In the first order condition for fuel f, the Lagrange multiplier for the supply constraint μ_2 acts as an increase in the shadow price of fuel f. Since I assume that the constraint is binding, the value of this shadow price will be such that the quantity purchased of fuel f would be $\overline{e}_f$ if the firm was facing $p_{fit} + \mu_2$ as the true price.

$$p_{fit} + \mu_2 = \mu_1 \underbrace{\left(\sum_{f \in \mathcal{F}} (\psi_{fit} \tilde{e}_{fit})^{\frac{\lambda-1}{\lambda}} \right)^{\frac{1}{\lambda-1}} \psi_{fit}^{\frac{\lambda}{\lambda-1}} e_{fit}^{\frac{-1}{\lambda}}}_{\text{Marginal Product of } e_{fit}}$$

Then, the perceived marginal cost of energy, $p_{E_{it}(\mathcal{F},\overline{e}_f)}$ will include the shadow cost of fuel f. By proposition 2, the increase in marginal costs will be larger under $\mathcal{F}$ than $\mathcal{F}'$. Thus, the firm is better of under the larger fuel set, $\mathcal{F}'$ when facing a shortage.

5 Estimating the production model

Identification of the model is done is three steps, each of which rely on different methods. First, I jointly identify and estimate the unobserved quantity and prices of realized energy in the outer nest of the production function by exploiting variation between observed expenditures on energy, expenditure on labor, and the quantity of labor. Identification relies on the mapping between observed substitution patterns between inputs and implied substitution patterns induced by optimality of the firm's choice, and works for a large class of parametric production functions (Grieco et al., 2016). Second, I jointly identify the inner nest of production, including all fuel-specific productivity. To do so, I rely on recent development in production function estimation in the presence of input-augmenting productivity that exploit firms' optimality conditions and Markovian assumptions on productivity (Demirer, 2020; Zhang, 2019).

Third, While the previous method allows me to recover fuel productivity for fuels that firms are using in a given period, it doesn't allow me to point identify counterfactual productivity for fuels that firms have never used, which may underlie selection in the distribution of observed fuel productivity. To uncover the unselected distribution of fuel productivity and perform counterfactual policy experiments, I follow Arcidiacono and Jones (2003, 2011) by assuming that the distribution of unobserved fuel productivity comes from a finite mixture with known support. I use the full information likelihood coupled with the EM algorithm to iteratively estimate the probability that a firm is in each point of the support conditional on observables and choices made . This method is then embedded into an otherwise standard dynamic discrete choice model in fuel sets to recover fixed switching costs.

Similarly for unobserved prices in counterfactual fuel sets, I use firms' location information into Indian districts and geopolitical shocks to deduce counterfactual fuel prices. Indeed, after mapping the entire natural gas pipeline network of India, I find significant spatial variation in fuel prices based on the distance between firms' location and the source of the nearest pipeline which captures transportation fees. For example, firms located nearby the Midwest Coast of India in the states of Gujarat and Maharashtra face significantly cheaper gas prices because most gas is imported from the Middle East to the West Coast, where most pipelines begin. I explain this in detail in Appendix A.2.

5.1 Identification of outer production function

In the outer nest, the main unobserved quantity that departs from standard models is realized energy $\tilde{E}_{it}$. In contrast to heating potential of fuels, realized energy is the output of combining different fuels in production, and is unobserved by construction. Fortunately, using a method developed by Grieco et al. (2016), there is a way to uniquely recover $\tilde{E}_{it}$ by exploiting structural variation in the quantity of another flexible input (labor) induced by unobserved variation in the price of realized energy that underly firm's choice of the scale of energy. To see this, one can look at the ratio of first-order conditions for labor and energy from profit maximization in equation 4:

$$\frac{w_{it}}{p_{E_{it}}} = \frac{\alpha_L}{\alpha_E}\Big(\frac{L_{it}/\overline{L}}{E_{it}/\overline{E}}\Big)^{-1/\sigma}\frac{\overline{E}}{\overline{L}}$$

Multiply both sides by L/E:

$$\frac{w_{it}L_{it}}{p_{E_{it}}E_{it}} = \frac{\alpha_L}{\alpha_E}\Big(\frac{L_{it}/\overline{L}}{E_{it}/\overline{E}}\Big)^{(\sigma-1)/\sigma} \tag{6}$$

Given production function parameters, $\frac{E_{it}}{\overline{E}}$ can be recovered from (6) because I observe expenditures for both inputs (recalling that energy expenditure is the sum of fuel expenditures from the energy production function: $p_{E_{it}}E_{it} = \sum_{f\in\mathcal{F}_{it}} p_{fit}e_{fit}$) and I observe quantity of labor. Identification of $\tilde{E}_{it}$ comes from variation in the relative input price of labor to energy induces variation in the expenditure ratio that isn't one-for-one with relative prices. For a given σ, observed variation in spending on energy $S_{E_{it}}$, spending on labor $S_{L_{it}}$ and the quantity in labor L_{it} implies a unique quantity of realized energy by the optimality condition between both inputs. Only when $\sigma = 1$ (Cobb-Douglas), the percentage change in relative prices is always offset by an equivalent percentage change in expenditure shares, such that expenditure shares are constant.

$$\frac{E_{it}}{\overline{E}} = \Big(\frac{p_{eit}E_{it}}{w_{it}L_{it}}\Big)^{\frac{\sigma}{\sigma-1}}\Big(\frac{\alpha_L}{\alpha_E}\Big)^{\frac{\sigma}{\sigma-1}}\frac{L_{it}}{\overline{L}} \tag{7}$$

In this setting, Grieco et al. (2016) show that one can identify production parameters by replacing E_{it} for (7) in the production function and exploiting first-order conditions to control for the transmission bias from unobserved hicks-neutral productivity z_{it} to observed inputs, a method that is also used by Doraszelski and Jaumandreu (2013, 2018). Detailed derivations can be found in **appendix X**. Following Grieco et al. (2016), I also use the same method to control for unobserved price dispersion in the bundle of material inputs, which has been known to vary significantly:

$$\frac{M_{it}}{\overline{M}} = \Big(\frac{p_{mit}M_{it}}{w_{it}L_{it}}\Big)^{\frac{\sigma}{\sigma-1}}\Big(\frac{\alpha_L}{\alpha_M}\Big)^{\frac{\sigma}{\sigma-1}}\frac{L_{it}}{\overline{L}}$$

The main dependent variable is revenues, where $e^{u_{it}}$ is an unobserved iid shock which is meant to capture measurement error and unanticipated demand & productivity shocks to the plant (Klette and Griliches, 1996).

$$\begin{aligned} R_{it} &= e^{u_{it}}p(Y_{it})Y_{it} \\ &= \frac{1}{\eta}\left[w_{it}L_{it}\Big(1+\frac{\alpha_k}{\alpha_L}\Big(\frac{K_{it}/\overline{K}}{L_{it}/\overline{L}}\Big)^{\frac{\sigma-1}{\sigma}}\Big) + p_{mit}M_{it} + p_{eit}E_{it}\right]e^{u_{it}} \end{aligned}$$

Taking logs of revenues yields the main estimating equation:

$$\ln R_{it} = \ln\frac{1}{\eta} + \ln\left[w_{it}L_{it}\Big(1+\frac{\alpha_k}{\alpha_L}\Big(\frac{K_{it}/\overline{K}}{L_{it}/\overline{L}}\Big)^{\frac{\sigma-1}{\sigma}}\Big) + p_{mit}M_{it} + p_{eit}E_{it}\right] + u_{it} \tag{8}$$

Note that due to the substitution for E_{it} and M_{it} in the production function, the main estimating equation (8) does not recover α_E and α_M. To recover α_E and α_M, I can rearrange the ratio of first-order conditions to show that expenditure shares of labor to energy must satisfy:

$$\frac{w_{it}L_{it}}{p_{eit}E_{it}} = \frac{\alpha_L}{\alpha_E}\Big(\frac{L_{it}/\bar{L}}{E_{it}/\bar{E}}\Big)^{\frac{\sigma-1}{\sigma}}$$

An analogous expression exists for the ratio of expenditures on materials and labor. Taking the geometric mean of relative expenditures yields the first parameter restriction[7], whereas the second is implied by constant returns to scale:

$$\begin{aligned} \overline{wL}/\overline{p_e E} &= \frac{\alpha_L}{\alpha_E} \\ \overline{wL}/\overline{p_m M} &= \frac{\alpha_M}{\alpha_E} \\ \alpha_K + \alpha_L + \alpha_M + \alpha_E &= 1 \end{aligned} \tag{9}$$

Then, I estimate (8) subject to (9) with non-linear least squares.

5.1.1 Outer Production function estimation results

Preliminary estimates of the production function parameters can be found in Table 5. I also report the average output elasticity with respect to each input to be consistent with the literature (Gandhi et al., 2020). The output elasticity with respect to intermediate materials (and the coefficient on materials) is much larger than other inputs, which is consistent with previous estimates in the literature. Moreover, the share of energy in the production function is larger for plants in industries that are expected to be more energy intensive such as the manufacturing of cement (Ryan, 2012).

5.1.2 Reduced-form evidence of gains from variety

One important quantitative goal of this paper is to study how much of the variation in the marginal cost of energy across different fuel sets can be explained by the gains from variety/option value vs. variation in prices and fuel/energy productivity due to the selection of more productive firms into large fuel sets. At this point, I have an estimate of energy $E_{it}/\overline{E}$ which gives me an estimate of the

[7] This is the convenience given by the geometric mean normalization of the CES. However, any other normalization would work, but would require some more albegra to recover the distribution parameters

Table 5: Production Function Estimation (selected industries)

	Glass	Cement	Basic Steel	Casting of Steel & Iron
Returns to scale: $\hat{\eta}$	0.87	0.76	0.92	0.88
	[0.857,0.880]	[0.749,0.769]	[0.916,0.923]	[0.876,0.887]
Elasticity of substitution: $\hat{\sigma}$	1.51	3.36	1.81	1.50
	[1.410,1.646]	[2.646,4.978]	[1.707,1.948]	[1.413,1.589]
Capital coef: $\hat{\alpha}_K$	0.12	0.08	0.06	0.08
	[0.102,0.128]	[0.058,0.084]	[0.055,0.062]	[0.080,0.091]
Labor coef: $\hat{\alpha}_L$	0.12	0.10	0.03	0.08
	[0.107,0.114]	[0.080,0.085]	[0.025,0.026]	[0.076,0.079]
Materials coef: $\hat{\alpha}_M$	0.64	0.61	0.82	0.74
	[0.644,0.669]	[0.594,0.619]	[0.822,0.830]	[0.736,0.747]
Energy coef: $\hat{\alpha}_E$	0.12	0.21	0.09	0.09
	[0.112,0.124]	[0.231,0.249]	[0.088,0.092]	[0.092,0.097]
N	1,553	2,145	7,177	4,654

Bootstrap 95% confidence interval in bracket (499 reps)

Table 6: Average Output Elasticities (selected industries)

	Glass	Cement	Basic Steel	Casting of Steel & Iron
$\bar{\epsilon}_{y,l}$	0.11	0.09	0.03	0.09
	[0.11,0.12]	[0.08,0.09]	[0.03,0.04]	[0.09,0.09]
$\bar{\epsilon}_{y,k}$	0.11	0.07	0.07	0.09
	[0.10,0.13]	[0.06,0.09]	[0.07,0.08]	[0.09,0.10]
$\bar{\epsilon}_{y,m}$	0.60	0.55	0.76	0.70
	[0.59,0.62]	[0.54,0.57]	[0.76,0.77]	[0.69,0.70]
$\bar{\epsilon}_{y,e}$	0.17	0.29	0.13	0.12
	[0.16,0.18]	[0.28,0.30]	[0.13,0.13]	[0.12,0.13]

price of energy $\tilde{p}_{E_{it}}\overline{E} = \tilde{p}_{E_{it}} = \frac{s_{E_{it}}\overline{E}}{E_{it}}$. Before imposing any further assumptions on the production model, I look at the relationship between the price of energy and the number of fuels available to firms. I find a large and monotone negative relationship between energy prices and the number of fuels. Since this evidence comes before imposing any assumption on the inner nest of production, it provides evidence in favor of the energy production model presented earlier.

Table 7: Relationship between $\ln \tilde{p}_{E_{it}}$ and the number of fuels available to plants.

	(1)		(2)		(3)		(4)	
	$\ln p_{Eit}$		$\ln p_{Eit}$		$\ln p_{Eit}$		$\ln p_{Eit}$	
Two fuels	-0.57***	(0.014)	-0.51***	(0.014)	0	(.)	0	(.)
Three fuels	-1.48***	(0.017)	-1.42***	(0.017)	-0.91***	(0.012)	-0.93***	(0.01)
Four fuels	-1.92***	(0.036)	-1.87***	(0.035)	-1.37***	(0.032)	-1.42***	(0.028)
Industry Dummies	Yes		Yes		Yes		Yes	
Year Dummies			Yes		Yes		Yes	
Controlling for fuel prices					Yes		Yes	
Controlling for TFP							Yes	
N	222,104		222,104		197,512		197,512	

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: the third and fourth columns control for the prices of electricity and oil, and are based on plants that always use these two fuels. This means that the benchmark number of fuels in these columns is two, rather than one in the first and second columns.

5.2 Identification of inner production function for energy

The energy production function in equation 2 can be rewritten by taking out the productivity of a fuel that plants always use, such as oil, and redefining the productivity of all other fuels relative to oil, $\tilde{\psi}_{fit} = \frac{\psi_{fit}}{\psi_{oit}}$:

$$\tilde{E}_{it} = \psi_{oit} \left(\sum_{f \in \mathcal{F}_{it}} \left(\tilde{\psi}_{fit} \tilde{e}_{fit} \right)^{\frac{\lambda-1}{\lambda}} \right)^{\frac{\lambda}{\lambda-1}} \tag{10}$$

At this point, I observe $\tilde{E}_{it}, \{\tilde{e}_{fit}\}_{f \in \mathcal{F}_{it}}$, and all prices: $\tilde{p}_{Eit} = \frac{S_{Eit}}{\tilde{E}_{it}}, \{\tilde{p}_{fit} = \frac{s_{fit}}{\tilde{e}_{fit}}\}_{f \in \mathcal{F}_{it}}$. I show how to recover the elasticity of substitution λ, and the distribution of all productivity terms ψ_{fit}. To do so, I rely on optimality of the firms' cost-minimization coupled with a Markovian assumption on the productivity of one of the input (e.g. oil). This is similar to the method proposed by Zhang (2019) and Demirer (2020), but relies on insights from the dynamic panel literature rather than the proxy variable/control variable approach to deal with endogeneity of productivity (transmission bias). As a reminder, the cost-minimization problem of the firm is as follows:

$$\min_{\{e_{fit}\}_{f\in\mathcal{F}_{it}}} \sum_{f\in\mathcal{F}_{it}} \tilde{p}_{fit}\tilde{e}_{fit} \quad s.t. \quad \tilde{E}_{it} = \psi_{oit}\left(\sum_{f\in\mathcal{F}_{it}} \left(\tilde{\psi}_{fit}\tilde{e}_{fit}\right)^{\frac{\lambda-1}{\lambda}}\right)^{\frac{\lambda}{\lambda-1}}$$

The relative first order conditions identifies relative productivity of fuel f as a function of observables up to parameter values:

$$\tilde{\psi}_{fit} = \left(\frac{\tilde{p}_{fit}}{\tilde{p}_{oit}}\right)^{\frac{\lambda}{\lambda-1}}\left(\frac{\tilde{e}_{fit}}{\tilde{e}_{oit}}\right)^{\frac{1}{\lambda-1}} \tag{11}$$

I then impose this optimality condition by plugging back the implied relative fuel productivity terms (11) into the energy production function (10) and rearrange (where $s_{fit} \equiv p_{fit}e_{fit}$ is spending on fuel f):

$$\tilde{E}_{it} = \psi_{oit}\tilde{e}_{oit}\left(\sum_{f\in\mathcal{F}_{it}} \frac{s_{fit}}{s_{oit}}\right)^{\frac{\lambda}{\lambda-1}} \tag{12}$$

At this point, the only unobservable in the energy production function is the productivity of oil, which is correlated with current period quantities and spending on inputs since it is assumed to be known to the firm at the time of choosing fuel quantities. This is the standard transmission bias. To deal with this issue, I assume that the productivity of oil follows an AR(1) Markov process with both years t and location s fixed effects. [8]

$$\ln\psi_{oit} = (1-\rho_{\psi_o})(\mu_0^{\psi_o} + \mu_s^{\psi_o}) + \mu_t^{\psi_o} - \rho_{\psi_o}\mu_{t-1}^{\psi_o} + \rho_{\psi_o}\ln\psi_{oit-1} + \epsilon_{it}^{\psi_o} \tag{13}$$

I then take log of equation 12 and use the Markov process above to get an estimating equation:

[8] The choice of these modified AR(1) processes is to ensure that the average of each state variables observed in the data corresponds to the unconditional average of this process. This means that many even though the model is estimated from a short panel (between 2 and 8 years), forward simulations multiple years ahead will match the support of the data.

$$\ln \tilde{E}_{it} - \ln \tilde{e}_{oit} = \mu_0^{\psi_o}(1-\rho_{\psi_o}) + \mu_t^{\psi_o} - \rho_{\psi_o}\mu_{t-1}^{\psi_o} + \rho_{\psi_o}(\ln \tilde{E}_{it-1} - \ln \tilde{e}_{oit-1}) + \frac{\lambda}{\lambda-1}\Big(\ln \sum_{f\in\mathcal{F}_{it}} \frac{s_{fit}}{s_{oit}} - \rho_{\psi_o} \ln \sum_{f\in\mathcal{F}_{it-1}} \frac{s_{fit-1}}{s_{oit-1}}\Big) + \epsilon_{it}^{\psi_o} \quad (14)$$

Since $\epsilon_{it}^{\psi_o}$ is a shock to productivity of oil at time t, it is uncorrelated with choices made at time $t-1$:

$$\mathbb{E}(\epsilon_{it}^{\psi_o} \mid \mathcal{I}_{it-1}) = 0$$

Moreover, estimating equation (14) is a linear structural regression, and I can use linear IV with $t-1$ choices as instruments to control for endogeneity between relative spending on all fuels $\ln \sum_{f\in\mathcal{F}_{it}} \frac{s_{fit}}{s_{oit}}$ and the shock to productivity of oil u_{oit}. In the paper, I use the price of oil and electricity, and the quantity of oil and electricity purchased at $t-1$ to instrument for $\ln \sum_{f\in\mathcal{F}_{it}} \frac{s_{fit}}{s_{oit}}$. The underlying assumption is that fuel price variation is persistent, which is consistent with the many geopolitical shocks which are pervasive in this market. In this dataset, this includes, for example, the oil crash of 2014 which persisted until 2016. For the price of electricity, Mahadevan (2022) documents many state-specific reforms to electricity markets which had persistent increases on the price of electricity. In the dynamic section, I specify a Markov process for the price of oil and electricity which is consistent with these assumptions.

Remark

So far, I haven't used the first-order condition for oil in the cost-minimization problem. This is not an issue because firms choose the level of realized energy in the first stage of production, given some price of energy. Once I recover the price of energy and the quantity of energy that firms want to buy, cost minimization implies that one of the input choice is "free". That is, I only need to recover the optimal quantity of all fuels relative to oil, whereas the quantity of oil will be pinned down by the firm's choice of realized energy. The first order condition for oil in the cost-minimization problem is as follows, where I sub in equation 11 for all relative fuel-augmenting productivity:

$$
\begin{aligned}
\tilde{p}_{oit} &= \mu_{it}\psi_{oit}\left(\sum_{f\in\mathcal{F}_{it}}\left(\tilde{\psi}_{fit}\tilde{e}_{fit}\right)^{\frac{\lambda-1}{\lambda}}\right)^{\frac{1}{\lambda-1}}\tilde{e}_{oit}^{-1/\lambda} \\
&= \mu_{it}\psi_{oit}\Big(\sum_{f\in\mathcal{F}_{it}}\frac{s_{fit}}{s_{oit}}\Big)^{\frac{1}{\lambda-1}}
\end{aligned} \tag{15}
$$

Once I take into account all first-order conditions, firms' optimality condition implies that the shadow cost of oil (Lagrange multiplier μ_{it}) is the marginal cost of realized energy. Plugging the equilibrium condition for the shadow cost of oil into equation (15) implies that the first order condition for oil is always satisfied.

$$
\mu_{it} = \tilde{p}_{Eit} = \frac{1}{\psi_{oit}}\Big(\sum_{f\in\mathcal{F}_{it}}\big(\frac{\tilde{p}_{fit}}{\tilde{\psi}_{fit}}\big)^{1-\lambda}\Big)^{\frac{1}{1-\lambda}} = \Big(\sum_{f\in\mathcal{F}_{it}}\frac{s_{fit}}{s_{oit}}\Big)^{\frac{1}{1-\lambda}}\frac{\tilde{p}_{oit}}{\psi_{oit}}
$$

5.2.1 Estimation of the inner production function

Energy production function

Table 8: Estimates of Energy Production Function (Selected parameters and industries)

	(1) Casting of Steel & Iron	(2) Glass	(3) Basic Steel
Elasticity of sub. λ	1.754***	2.191*	1.901***
	(0.140)	(0.902)	(0.105)
Persistence of oil prod ρ_{ψ_o}	0.859***	0.873***	0.911***
	(0.0162)	(0.0288)	(0.0121)
Observations	1,737	652	2,746

Standard errors in parentheses
$^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

To get the distribution of all fuel productivity, I multiply the estimates of oil productivity with the relative productivity of all other fuels $\hat{\psi}_{fit} = \hat{\tilde{\psi}}_{fit}\hat{\psi}_{oit}$. I then construct a measure of energy productivity by taking the firm-by-year sample average of all fuel productivity terms:

$$\ln \overline{\psi}_{it} = \frac{1}{\mid \mathcal{F}_{it} \mid} \sum_{f \in \mathcal{F}_{it}} \ln \psi_{fit}$$

Of course this is a selected sample of productivity, because I only observe gas and coal productivity for firms that use gas and coal. Yet, it is still indicative of productivity differences across different fuel sets. In the dynamic section, I show how to recover the unselected distribution of fuel productivity.

Distribution of energy and fuel productivity, by fuel sets

The average and median energy productivity tend to increase as firms use more fuels, which is consistent with a productivity-efficiency argument in which more productive firms select into large fuel sets (Table 8). However, firms who use only gas with oil and electricity (rather than gas, coal, oil and electricity) tend to have a comparative advantage in using gas relative to their overall productivity of energy (Figures 6):

Table 9: Distribution of log energy productivity $\ln \overline{\psi}_{it}$ by fuel sets (Steel manufacturing)

	Mean	Median	s.d.	N
oil, electricity	0.31	0.22	1.93	1,501
oil, coal, electricity	0.29	0.50	2.05	667
oil, gas, electricity	0.38	0.24	1.53	353
oil, gas, coal, electricity	0.55	0.57	1.43	225

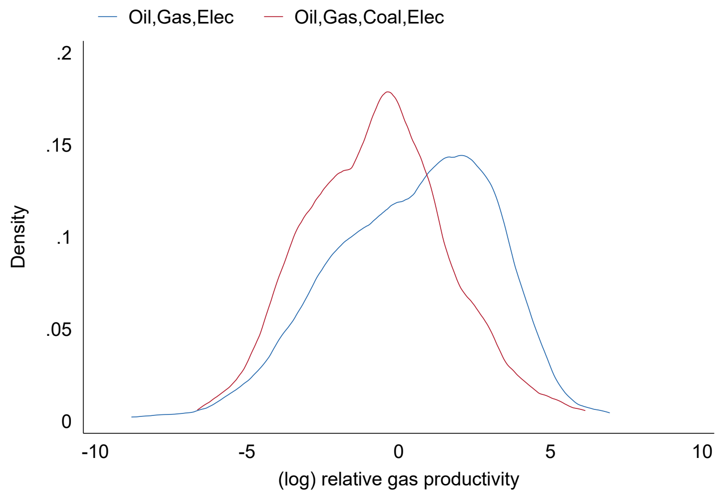


Figure 6: Kernel density of log relative productivity of gas ($\ln \tilde{\psi}_{git}$), Steel manufacturing

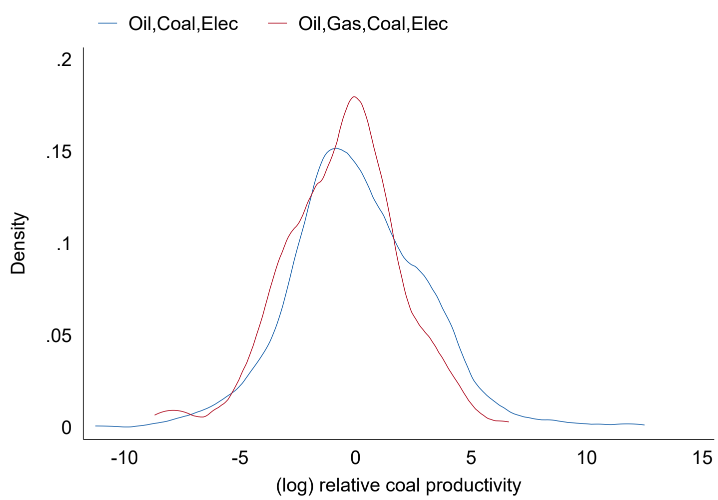


Figure 7: Kernel density of log relative productivity of coal ($\ln \tilde{\psi}_{cit}$), Steel manufacturing

6 Estimating fuel switching costs

The firm has access to a set of fuels $\mathcal{F}_{it}$ and is considering all alternative fuel sets in the next period $\mathcal{F}' \equiv \mathcal{F}_{it+1} \subseteq \mathbb{F} \equiv \{oe, oge, oce, ogce\}$. Since all state variables s_{it} will be assumed to follow a Markovian process, I can start from the recursive formulation of the problem, where the firm chooses $\mathcal{F}'$ to maximize a bellman equation, the net present value of lifetime profits:

$$V(s_{it}, \epsilon_{it}, \mathcal{F}_{it}) = \max_{\mathcal{F}' \subseteq \mathbb{F}} \Big\{ \underbrace{\pi(s_{it}, \mathcal{F}_{it})}_{\text{static profit}} + \underbrace{\Phi(\mathcal{F}' \mid \mathcal{F}_{it}) + \epsilon_{\mathcal{F}'it}}_{\text{Fixed switching costs}} + \beta \int_{\Omega_s} \int_{\epsilon} V(s_{it+1}, \epsilon_{it+1}, \mathcal{F}') dF(s_{it+1} \mid s_{it}) dG(\epsilon_{it+1}) \Big\} \tag{16}$$

Where the constant component of switching costs are defined as:

$$\Phi(\mathcal{F}' \mid \mathcal{F}_{it}) = \Big(\underbrace{\sum_{f \in \mathcal{F}_{it}} \mathbb{I}(f \notin \mathcal{F}') \gamma_f}_{\text{salvage value for dropping fuels}} - \underbrace{\sum_{f \notin \mathcal{F}_{it}} \mathbb{I}(f \in \mathcal{F}') \kappa_f}_{\text{cost of adding new fuels}} \Big)$$

From now on, I define the parameters governing the switching cost function $\theta_1 = \{\kappa_g, \kappa_c, \gamma_g, \gamma_C\}$ for coal c and gas g, and θ_2 the parameters underlying the evolution of state variables. I make the assumption that cost shocks are iid and come from a standardized Type 1 Extreme value $\epsilon_{\mathcal{F}'it} \sim Gumbel(0, 1)$. This allows me to analytically integrate over these shocks and work with the expected value function:

$$\begin{aligned} W(s_{it}, \mathcal{F}_{it}) &= \int \max_{\mathcal{F}' \in \mathbb{F}} \Big\{ \pi(\mathbf{s}_{it}, \mathcal{F}_{it}) + \Phi(\mathcal{F}' \mid \mathcal{F}_{it}) + \epsilon_{\mathcal{F}'_{it}} + \beta\, \mathbb{E}[V(\mathbf{s}_{it+1}, \mathcal{F}') \mid s_{it}] \Big\} g(\epsilon) d\epsilon \\ &= \gamma + \log \left(\sum_{\mathcal{F}' \in \mathbb{F}} \exp \Big(\pi(s_{it}, \mathcal{F}_{it}) + \Phi(\mathcal{F}' \mid \mathcal{F}_{it}) + \beta \int W(s_{it+1}, \mathcal{F}') f(s_{it+1} \mid s_{it}) ds_{it+1} \Big) \right) \\ &= \gamma + \log \left(\sum_{\mathcal{F}' \in \mathbb{F}} \exp \Big(\upsilon_{\mathcal{F}'}(s_{it}, \mathcal{F}_{it}) \Big) \right) \end{aligned}$$

Where $\gamma \approx 0.5772$ is the Euler–Mascheroni constant. Then, the probability of choosing fuel $\mathcal{F}'$ has a logit formulation, which simplifies the likelihood. Note that This probability is implicitely a function of both θ_1 and θ_2

$$Pr(\mathcal{F}' \mid \mathcal{F}_{it}, s_{it}; \theta_1, \theta_2) = \frac{exp\Big(\upsilon_{\mathcal{F}'}(s_{it}, \mathcal{F}_{it}; \theta_1, \theta_2)\Big)}{\sum_{\mathcal{F}'\in\mathbb{F}} \exp\Big(\upsilon_{\mathcal{F}'}(s_{it}, \mathcal{F}_{it}; \theta_1, \theta_2)\Big)}$$

Where $\upsilon()$ are the choice-specific expected value functions. Below is the main assumption underlying firms' expectation over state variables:

Assumption 1-. *I assume that firms are agnostic about the evolution of labor, capital and material prices (w_{it}, r_{kit}, p_{mit}). However, firms take expectation over all productivity terms and fuel prices ($\{\psi_{fit}, p_{fit}\}_{f\in\mathbb{F}}, z_{it}$)*

6.1 Evolution of non-selected state variables

I assume a specific process for the evolution of state variables that is consistent with previous sections of the paper. The productivity and price of both electricity and oil follow a persistent AR(1) process with time (t) and location (s) fixed effects. Since firms always use these two fuels, they are referred as non-selected state variables. $\forall f = \{e, o\}$:

$$\begin{aligned}
\ln \psi_{fit} &= (1-\rho_{\psi_f})(\mu_0^{\psi_f} + \mu_s^{\psi_f}) + \mu_t^{\psi_f} - \rho_{\psi_f}\mu_{t-1}^{\psi_f} + \rho_{\psi_f}\ln \psi_{fit-1} + \epsilon_{it}^{\psi_f} \\
\ln p_{fit} &= (1-\rho_{p_f})(\mu_0^{p_f} + \mu_s^{p_f}) + \mu_t^{p_f} - \rho_{p_f}\mu_{t-1}^{p_f} + \rho_{p_f}\ln p_{fit-1} + \epsilon_{it}^{p_f}
\end{aligned}$$

I also assume a similar persistent AR(1) process for hicks-neutral productivity z_{it}:

$$\ln z_{it} = (1-\rho_z)\mu_z + \mu_{zt} - \rho_z\mu_{zt-1} + \rho_z \ln z_{it-1} + \epsilon_{zit}$$

I allow all shocks to productivity and prices to be arbitrarily correlated in a multivariate normal distribution. Baseline Electricity prices are set by state-owned electricity utilities, but vary non-linearly based on demand across the grid. Moreover, due to widespread electricity shortages, many states have made reforms in different years to modernize the electricity sector, at the expense of higher prices. Mahadevan (2022) shows that these reforms also increased firm productivity, thereby increasing demand for electricity, which motivates a joint process for shocks to prices and productivity.

$$\left(\epsilon_{it}^{\psi_o}, \epsilon_{it}^{p_o}, \epsilon_{it}^{\psi_e}, \epsilon_{it}^{p_e}, \epsilon_{zit}\right) \equiv \epsilon_{\mathbf{it}} \sim \mathcal{N}(\mathbf{0}, \mathbf{\Sigma})$$

6.2 Evolution of selected state variables - Permanent heterogeneity

I assume prices and productivity for coal and gas follow a joint process with permanent plant-specific heterogeneity in productivity. $\forall f = \{c, g\}$,

$$\begin{aligned}
\ln \psi_{fit} &= \mu_0^{\psi_f} + \mu_t^{\psi_f} + \mu_s^{\psi_f} + \mu_{fi} + \epsilon_{it}^{\psi_f} \\
\ln p_{fit} &= \mu_0^{p_f} + \mu_t^{p_f} + \mu_s^{p_f} + \epsilon_{it}^{p_f}
\end{aligned}$$

Both price and productivity may vary over time and across locations. For example, Firms typically face lower prices of natural gas when closer to the source of a natural gas pipeline (Table 11, Appendix A.2). Moreover, I allow shocks to productivity and prices to be correlated:

$$\begin{pmatrix} \epsilon_{it}^{\psi_f} \\ \epsilon_{if}^{p_f} \end{pmatrix} \sim N \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_{\psi_f}^2 & \sigma_{\psi_f p_f} \\ \sigma_{\psi_f p_f} & \sigma_{p_f}^2 \end{pmatrix} \right] \qquad \forall f = \{g, c\}$$

The main difference between selected and unselected state variables is the presence of firm-specific comparative advantage at using fuel f, μ_{fi}. Whereas unselected variables feature persistent heterogeneity in the form of an AR(1), I assume that the productivity of selected fuels feature permanent heterogeneity in the form of random effects. The main difference is methodological and stems from limitation in the literature to handle persistent unobserved heterogeneity.

6.3 Identification

To learn about the extent to which the distribution of comparative advantage for natural gas and coal is selected, I follow Arcidiacono and Jones (2003, 2011). I assume that the distribution of comparative advantages comes from a finite mixtures, where the initial guess of the mean and variance of the finite mixture are the same as the mean and variance of some initial guesses $(\tilde{\mu}_f, \tilde{\sigma}_{\mu_f}^2)$ of an underlying continuous process:

$$\sum_{k}^{K} \pi^0_{fk} \mu_{fk} = \tilde{\mu}_f$$
$$\sum_{k}^{K} (\mu_{fk} - \tilde{\mu}_f)^2 \pi^0_{fk} = \tilde{\sigma}^2_{\mu_f}$$

Where $\pi^0_{fk} = Pr(\mu_{fk})$ is the unconditional probability of being type k, and $\sum_k \pi^0_{fk} = 1$. I make an informed initial guess of both distributions from the selected sample of firms who use these fuels.

$$\mu^0_{fi} \sim N(\tilde{\mu}_f, \tilde{\sigma}^2_{\tilde{\mu}_f})$$

In this context, external estimation of parameters governing the distribution of random effects from a selected sample of firms who use these fuels leads to biased estimates of $\tilde{\mu}_g, \tilde{\mu}_c, \tilde{\sigma}^2_{\mu_g}, \tilde{\sigma}^2_{\mu_c}$. This is because different firms face different distributions of expected fuel-specific productivity on the basis of their comparative advantage, which affects their decisions of which fuel sets to choose for next period. Indeed, firms with larger comparative advantage to use coal are more likely to use coal, and likewise for gas. Thus, I expect to get upward biases in both the mean of coal and gas.

Discretization methods such as Rouwenhorst (1995) allow me to initialize the finite mixture from the continuous distribution of observed productivity from selected firms. For now, the true probability weights π_{fk} over the support of the finite mixture are unknown due to selection, but Arcidiacono and Jones (2003, 2011) provide a method to recover the unselected distribution by sequentially iterating over the fixed costs to maximize the likelihood and updating the probability weights $\pi^0_{kf}, \pi^1_{fk}, \pi^2_{fk}, ...$ using an EM algorithm. For now I assume the support of the finite mixture is know. In later versions, I will also allow the support points to vary.

Using the law of total probability, I can write the full information (log) likelihood by integrating the log of choice probabilities over this distribution (assuming there is only one finite mixture over both coal and gas for notational convenience):

$$\ln \mathcal{L}(\mathcal{F}, s \mid \theta_1, \theta_2) = \sum_{i=1}^{n} \ln \left[\sum_{k} \pi_{fk} \left[\prod_{t=1}^{T} Pr(\mathcal{F}_{it+1} \mid \mathcal{F}_{it}, s_{it}, \mu_{fi} = \mu_{fk}; \theta_1, \theta_2) \right] \right] + \sum_{i=1}^{n} \sum_{t=1}^{T} \ln f(s_{it} \mid s_{it-1}; \theta_2)$$

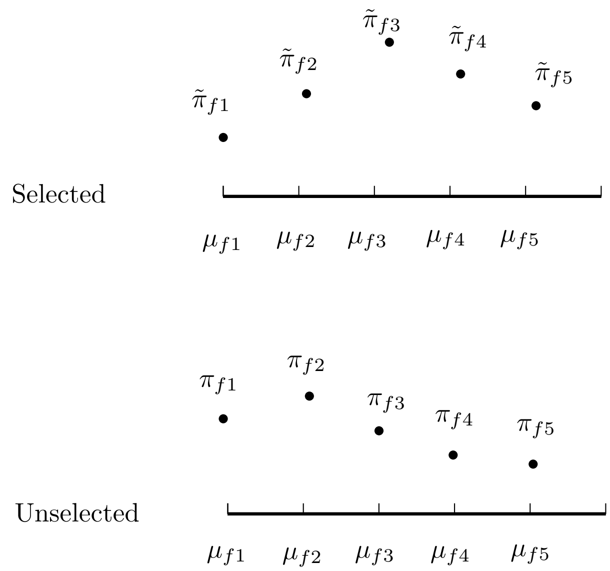


Figure 8: Example of how updating the probability weights can shift the distribution

I assume that the state transitions are independent of the random effects. This is possible if the parameter estimates $\hat{\theta}_2$ are unbiased from selected data, and if I slightly redefine the state space. In Appendix C.2, I show Monte-Carlo simulation results that are consistent with this assumption. Then, following Baye's law, one can show that the solution to this maximum likelihood problem is the same as the solution to a sequential EM algorithm that uses the posterior conditional probabilities that firm i is of type k given all observable, including choices made:

$$
\begin{aligned}
\hat{\theta}_1 &= \underset{\theta_1,\theta_2,\pi}{\arg\max} \sum_{i=1}^{n} \ln \left[\sum_{k} \pi_{fk} \left[\prod_{t=1}^{T} Pr(\mathcal{F}_{it+1} \mid \mathcal{F}_{it} s_{it}, \mu_{fi} = \mu_{fk}; \theta_1, \theta_2) \right] \right] \\
&\equiv \underset{\theta_1}{\arg\max} \sum_{i=1}^{N} \sum_{t=1}^{T} \sum_{k} \rho(\mu_{fk} \mid \mathcal{F}_i, s_i; \hat{\theta}_1, \hat{\theta}_2, \hat{\pi}) \ln Pr(\mathcal{F}_{it+1} \mid \mathcal{F}_{it}, s_{it}, \mu_{fi} = \mu_{fk}; \theta_1, \hat{\theta}_2)
\end{aligned}
$$

Where $\mathcal{F}_i$ is all the choices that we observe firm i making. Using Baye's rule, the conditional probability of that firm i is of type k is given by the current guess of the unconditional probability π_{fk} weighted by the probability that the firm makes the observed sequence of fuel set choices conditional being type k:

$$\rho(\mu_{fk} \mid \mathcal{F}_i, s_i; \theta_1, \theta_2, \pi) = \frac{\pi_{fk}\left[\prod_{t=1}^{T}\left[\prod_{\mathcal{F}\subseteq\mathbb{F}}\left[Pr(\mathcal{F}_{it} \mid s_{it}, \mu_{fi} = \mu_{fk}; \theta_1, \theta_2)\right]^{\mathbb{I}(\mathcal{F}_{it}=\mathcal{F})}\right]\right]}{\sum_k \pi_{fk}\left[\prod_{t=1}^{T}\left[\prod_{\mathcal{F}\subseteq\mathbb{F}}\left[Pr(\mathcal{F}_{it} \mid s_{it}, \mu_{fi} = \mu_{fk}; \theta_1, \theta_2)\right]^{\mathbb{I}(\mathcal{F}_{it}=\mathcal{F})}\right]\right]} \tag{17}$$

The Procedure to estimate the fixed costs parameters θ_1 and the unselected, unconditional distribution of fuel-specific random effects goes as follows. I experimented with both the Arcidiacono and Jones (2003) version that relies on a nested fixed point algorithm to update the value function and the Arcidiacono and Jones (2011) that uses the conditional choice probabilities (CCP) and forward simulations to update the value function. Going forward, I will be using the nested fixed point version with a large grid for the state space, unless specified otherwise:

1. Initialize fixed cost parameters θ_1^1 and guess some initial probabilities $\{\pi_{f1}^1, \pi_{f2}^1, ..., \pi_{fK}^1\}$. I will use the distribution of selected random effects to initialize.

2. Do VFI/Foward Simulation to update the expected value function W conditional on these guesses, where different realizations of the random effects μ_{fk} are just another state variable that is fixed over time.

3. Get posterior conditional probabilities that firm i is of type k according to (equation 17): $\rho^1(\mu_{fk} \mid \mathcal{F}_i, s_i; \theta_1^1, \hat{\theta}_2, \pi^1)$

4. Update the unconditional probabilities as follows:

$$\pi_{fk}^2 = \frac{\sum_{i=1}^{n} \rho^1(\mu_{fk} \mid \mathcal{F}_i, s_i; \theta_1^1, \hat{\theta}_2, \pi^1)}{n}$$

5. Find fixed cost parameters θ_1^2 that maximize the (log)-likelihood conditional on current guess of unconditional and conditional probabilities $\pi_{fk}^1, \rho^1(\mu_{fk} \mid .)$

6. Repeat 2-5 until convergence

7 Preliminary Estimation Results - Steel manufacturing (in progress)

First, I find that fixed costs are consistent with ballpark estimates of capital typically required to use various fuels in steel manufacturing. For example, an electric arc furnace that uses natural gas typically costs from several hundred thousands to several million US dollars depending on the scale of a plant's operations. I find that the average fixed cost of adding natural gas is equal to \$ 615,595 USD. I also find that the fixed cost to add coal is 30% cheaper than natural gas, consistent with coal-based methods being older and more archaic. Lastly, I find that salvage values are much lower than fixed costs, but 93% larger for natural gas relative to coal. Indeed, many plants who drop coal salvage older and cheaper equipment than firms who drop gas. These results are qualitatively consistent with depreciating capital over time.

	Fixed Costs (Thousand \$USD)	**Salvage Values (Thousand \$USD)**
Natural Gas	615.9	208.2
Coal	474.6	107.6
N	2,385	2,385

Table 10: Estimates of fixed costs and salvage values

Regarding the distribution of unselected comparative advantage for natural gas and coal, I find a significant upward bias in the mean of both distributions. Moreover, I find that coal is much more selected than gas.

	Natural Gas			Coal		
	Mean	Variance	N	Mean	Variance	N
All sample	0.064	2.94	2,385	-0.38	1.81	2,385
Selected sample	0.1	2.91	508	0.16	2.73	763

Table 11: Distribution of unobserved heterogeneity (comparative advantages) for natural gas and coal, selected vs. unselected sample

Overall, the estimates of switching costs and the distribution of comparative advantage allow the model to predict quite well the unconditional empirical distribution of fuel set choices and the observed transition patterns between fuel sets. In all figures below, the blue bars (Data) are constructed as follows:

$$N_{\mathcal{F}'}(data) = \sum_i \sum_t \mathbb{I}(\mathcal{F}_{it+1} = \mathcal{F}')$$

$$N_{\mathcal{F}'|\mathcal{F}}(data) = \sum_i \sum_t \mathbb{I}(\mathcal{F}_{it+1} = \mathcal{F}' \mid \mathcal{F}_{it} = \mathcal{F})$$

The orange bars (model) are constructed by adding the predicted probability that each firm uses each fuel sets, integrated over the conditional distribution of comparative advantages:

$$N_{\mathcal{F}'}(model) = \sum_i \sum_t \sum_k \rho(\mu_{fk} \mid \mathcal{F}_i, s_i; \hat{\theta}_1, \hat{\theta}_2, \hat{\pi}) Pr(\mathcal{F}' \mid \mathcal{F}_{it}, s_{it}, \mu_{fi} = \mu_{fk}; \hat{\theta}_1, \hat{\theta}_2)$$

$$N_{\mathcal{F}'|\mathcal{F}}(model) = \sum_i \sum_t \sum_k \underbrace{\rho(\mu_{fk} \mid \mathcal{F}_i, s_i; \hat{\theta}_1, \hat{\theta}_2, \hat{\pi})}_{\text{Conditional probability of comparative advantage}} \underbrace{Pr(\mathcal{F}' \mid \mathcal{F}, s_{it}, \mu_{fi} = \mu_{fk}; \hat{\theta}_1, \hat{\theta}_2)}_{\text{Conditional choice probability}}$$

Graphs using conditional probability of comparative advantage

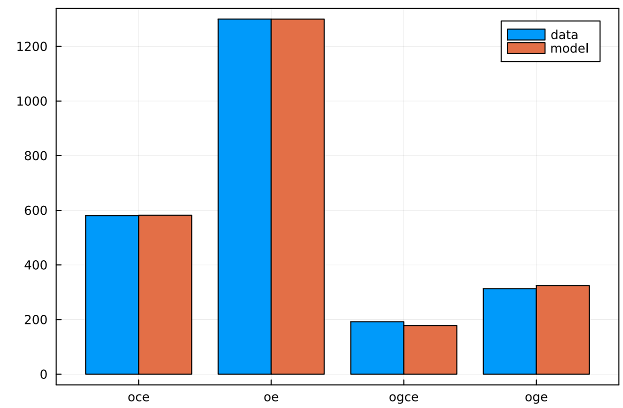


Figure 9: Unconditional distribution of fuel sets, model vs. data

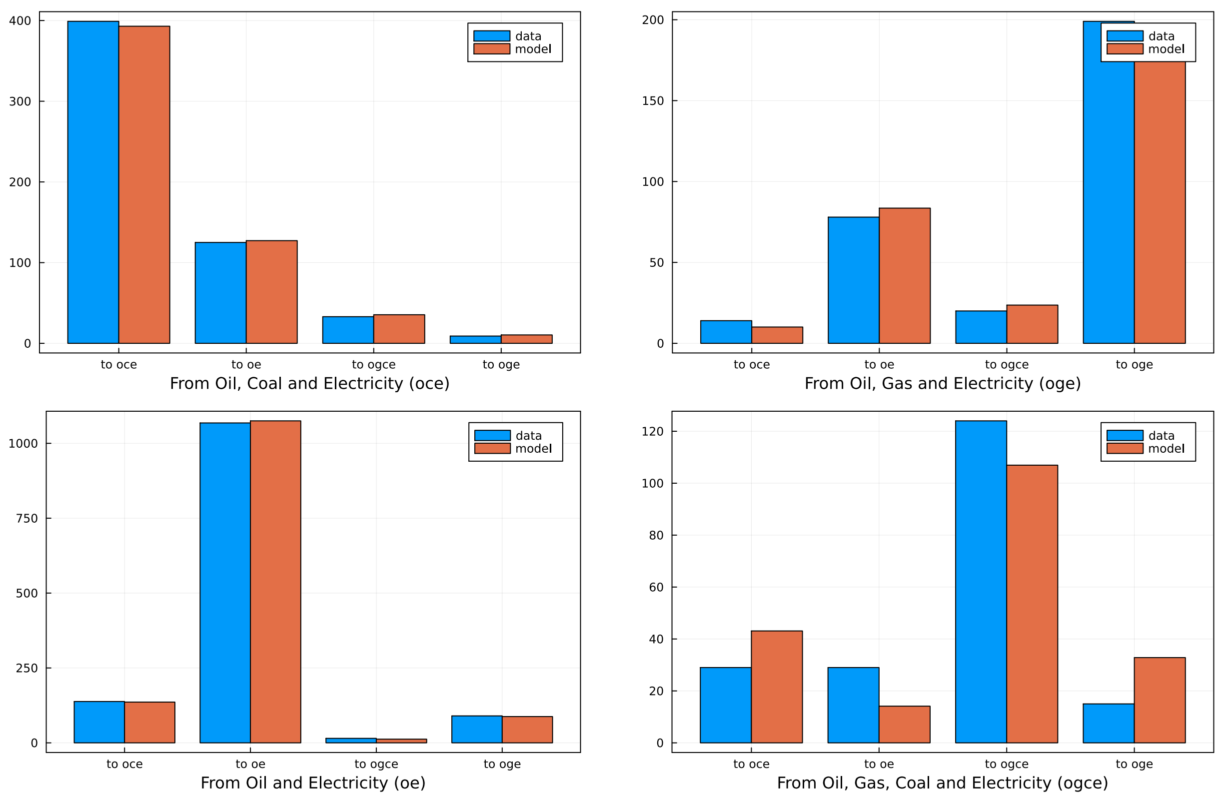


Figure 10: Conditional distribution of fuel sets (transition), model vs. data

8 Optimal Fossil Fuels Taxes (in progress)

Externality Damages

Externality comes from the release of pollutants in the air by the combustion of fuels. All pollutants are converted into carbon dioxide equivalent (CO_{2e}) using standard scientific calculations from the US EPA. Then, each unit of potential energy of fuel f contributes to contemporaneous greenhouse gas emissions as follows: 1 $mmBtu$ of e_f releases γ_f short tons of CO_{2e}. γ_f are fuel-specific emission intensity, and are calculated using the global warming potentital (GWP) method detailed in Appendix A.3. For example, 1 $mmBtu$ of coal releases roughly twice as much carbon dioxide equivalent in the air as 1 $mmBtu$ of natural gas $\frac{\gamma_c}{\gamma_g} \approx 2$. Fuel-specific emission intensity γ_f are then multiplied by the social cost of carbon (SCC) to get a monetized value of externality damages: $\tilde{\gamma}_f = SCC * \gamma_f$. Following Fowlie et al. (2016), I use the most recent estimates of the SCC from the IPCC which is \$53 USD per short ton of CO_{2e}, assuming a social discount factor of 3%.[9] Marginal externality damages are economically significant, For example, the marginal externality damage for coal is almost twice as high as the average price of coal. Below I compare the average price of all fuels if a planner were to impose a pigouvian carbon tax where the tax rate on each fuel is equal to marginal externality damages $\tau_f = \tilde{\gamma}_f \;\; \forall f \in \{o, g, c, e\}$.

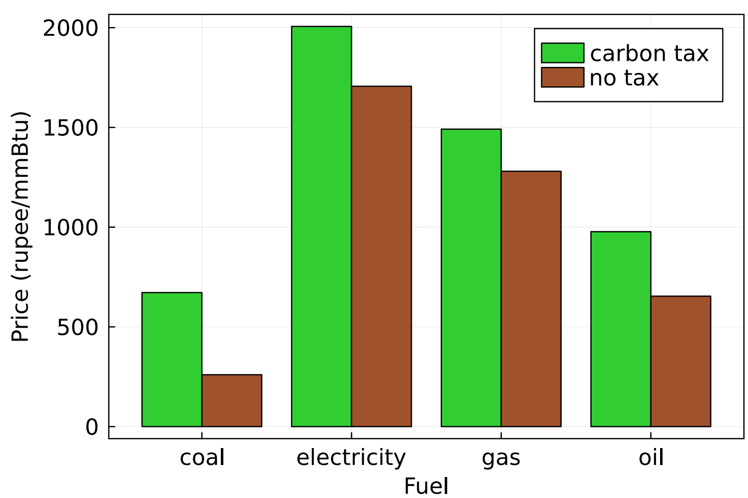


Figure 11: Average Fuel prices with and without carbon tax - Steel Manufacturing

Since I assume there is a representative consumer with quasi-linear utility, when expressed in dollars using the social cost of carbon, externality damages effectively act as reduction in aggregate income. This is one of the two standard approach in the literature to evaluate externality taxes.[10]

[9]While 3% is standard in the literature, I also experiment with different values of the social discount factor which yield different SCC. The more society discounts the future, the lower the SCC will be.

[10]The other approach is to use integrated assessment models (IAM), in which the social cost of carbon is endogenously determined by the interaction of economic activity and the atmospheric concentration of CO_2. These are

Per-period welfare is composed of four terms (Fowlie et al., 2016):

$$w_t(\tau) = \underbrace{\nu_t(\tau)}_{\text{consumer surplus}} + \underbrace{\sum_i \pi_{it}(\tau)}_{\text{Producers surplus}} + \underbrace{\sum_f \sum_i \tau_f e_{fit}}_{\text{Gov. revenues}} - \underbrace{\sum_f \sum_i \tilde{\gamma}_f e_{fit}}_{\text{Externality damages}} \tag{18}$$

Where τ_f are fuel-specific tax rates, $\tilde{\gamma}_f$ are marginal externality damages, $\nu_t(\tau)$ is the indirect utility of the representative consumer, and $\sum_i \pi_{it}(\tau)$ is the sum of firm profits including switching costs. As in Fowlie et al. (2016), I assume that the social cost of carbon (SCC) is constant and the government imposes a permanent tax to maximize the net present value (NPV) of expected period welfare functions over an infinite horizon:

$$\max_{\{\tau_f\}_{f \in \mathbb{F}}} \sum_{t=0}^{\infty} \beta^t \, \mathbb{E}_0(w_t(\tau))$$

To evaluate the social welfare, I follow Fowlie et al. (2016) by simulating firms' path of fuel set choices and production decisions over an horizon of 30 years, which I average over S simulations. Standard Pigouvian taxation dictates that a carbon tax can be implemented with taxes on different fuels, where relative tax rate reflects the relative emission intensities of different fuels $\tau_f = \tilde{\gamma}_f$. This is also how carbon taxes are implemented in practice.

8.1 Benchmark with literature

To evaluate how the introduction of fuel-augmenting productivity and costly switching contrasts with previous findings in the literature, I use a benchmark version of my model in which all firms have access to different fuel sets to produce energy, but cannot switch between sets and do not differ in fuel productivity. In that context, a firm with fuel set $\mathcal{F}$ produces energy according to the following CES production function:

$$E(\mathcal{F}) = \left(\sum_{f \in \mathcal{F}} \beta_f e_f^{\frac{\lambda - 1}{\lambda}} \right)^{\frac{\lambda}{\lambda - 1}}$$

dynamic general equilibrium models of the aggregate economy and are not used to study the behavior of individual agents.

In **Appendix B.2** I show that these fuel-set specific energy production functions can be aggregate into a single aggregate CES production function for energy, where ψ_f are endogenous loading of each fuels. They are a function of both the average productivity of each fuel β_f and the share of firms who use each fuels.

$$\tilde{E} = \left(\sum_{f \in \{o,g,c,e\}} \psi_f^{\frac{1}{\lambda}} e_f^{\frac{\lambda-1}{\lambda}} \right)^{\frac{\lambda}{\lambda-1}}$$

This production function allow me to compare results with **Golosov et al. 2014** who study optimal fossil fuel taxes in a general equilibrium IAM model with an aggregate CES production function for energy. Below I compare the full model with the restricted models to study the effect of a carbon tax on the different components of welfare. First, allowing firms to switch leads to a decrease in fuel sets that include coal when faced with a carbon tax because the tax on coal almost doubles its price. In particular, the probability of choosing a fuel set that contains coal goes from 0.37 to 0.31, a decrease of 6 percentage points. There is also some movements to and off natural gas, but the net effect is a decrease of in the probability of using a fuel set that contains natural gas by 2.7 percentage points.

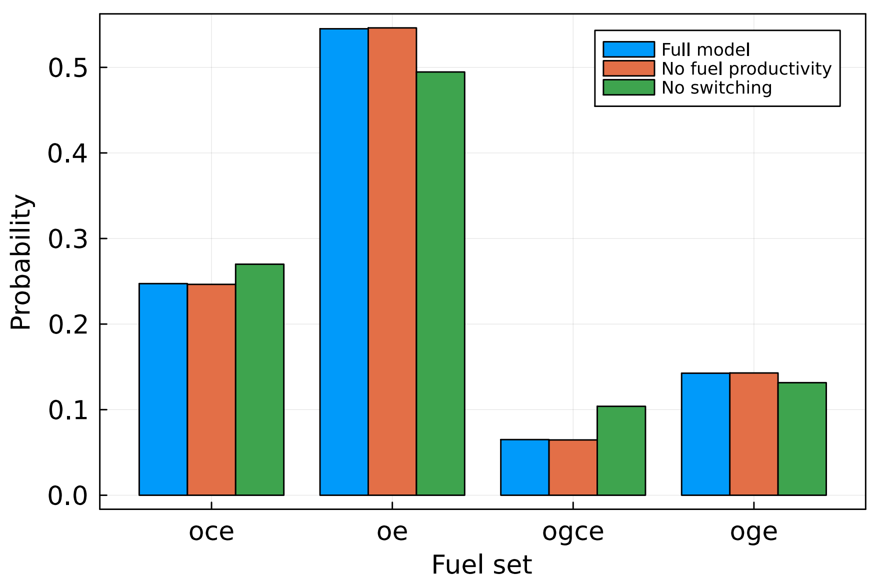


Figure 12: Predicted probability of using different fuel sets when facing a carbon tax - Steel manufacturing

Notes: To compute these predicted probabilities, I averaged across all fuel choices made by firms in the forward simulation. To compute the predicted probability in the no switching model, I used the predicted probabilities under no carbon tax which closely corresponds to the data, and did the simulation under the carbon tax by imposing a no switching rule.

In particular, I find that disallowing firms to switch significantly underestimate the reduction in externality damages from a carbon tax by 122% and overestimates the reduction in operating profits by 8.8 %. This is because fuel switching adds an additional margin of substitution that firms can exploit to hedge against the tax. Since firms internalize externality damages when facing a carbon tax, this additional margin of substitution improves both the private benefits (operating profits) and the social costs (externality damages), which improves overall welfare. When I turn off fuel productivity in the model, I significantly underestimate efficiency gains from fuel combustion, which reduces fuel demand across the board. This leads to a reduction in both operating profits due to larger marginal costs of realized energy and an a decrease in externality damages.

		Effect of a Carbon tax on Welfare Components				
		Full Model (1)	No switching (2)	No fuel prod (3)	% difference (1) vs. (2)	% difference (1) vs. (3)
Operating profits (excl. switching costs) (billion rupees)		21.4	19.6	10.6	8.8	67.5
Externality Damages (billion rupees)	Coal	51.15	217.9	0.15	123	199
	Gas	0.1	0.9	0.01	160	160
	Oil	1.14	1.94	0.3	52	116
	Electricity	1.05	2.65	0.31	86	108
	Total	53.4	223.4	0.5	122.8	196

Table 12: Effect of a carbon tax - Full Model vs. No switching vs. No fuel productivity - Steel Manufacturing

Notes: Given a $53 USD social cost of carbon, the carbon tax is 211 rupee/mmBtu for gas, 411 rupee/mmBtu for coal, 323 rupee/mmBtu for oil and 300 rupee/mmBtu for electricity.

8.2 Policy Evaluation

The vast majority of externality damages come from the combustion of coal, even under a carbon tax. See Table 12. Firms who use coal use very large quantities and pollute twice as much as natural gas. While a pigouvian carbon tax maximizes social welfare, this type of first-best argument has been recently questioned by Kotchen (2022). Indeed, a key element to the pigouvian argument rests on the tax being revenue-neutral. Since tax revenues are returned to the representative consumer as aggregate income, it exactly cancel the loss of aggregate income from externality damages when $\tau_f = \tilde{\gamma}_f$. However, in a developing economy like India, government intervention may not be frictionless, and the revenue-neutral argument of the carbon tax may fall. For this reason, Kotchen (2022) proposes to evaluate the effectiveness of a policy by looking at the ration of welfare to tax revenues. The idea is to find a policy that achieves most of its welfare gains by reducing externality damages. In this context, a policy whose primary aim is to reduce coal consumption may fair better than a carbon tax.

I choose to investigate the effectiveness of two policies in addition to a carbon tax. See Table 14. The first one is a large tax on coal where the tax rate is equal to the sum of marginal externality damages for all fuels $\tau_c = \sum_{f\in\{o,g,c,e\}} \tilde{\gamma}_f$,

	Coal	**Gas**	**Oil**	**Electricity**
Carbon Tax	211	411	323	300
Coal Tax	0	1245	0	0
Coal Tax and Gas Subsidy	-211	1245	0	0

Table 14: Different tax proposals ($rupees/mmBtu$)

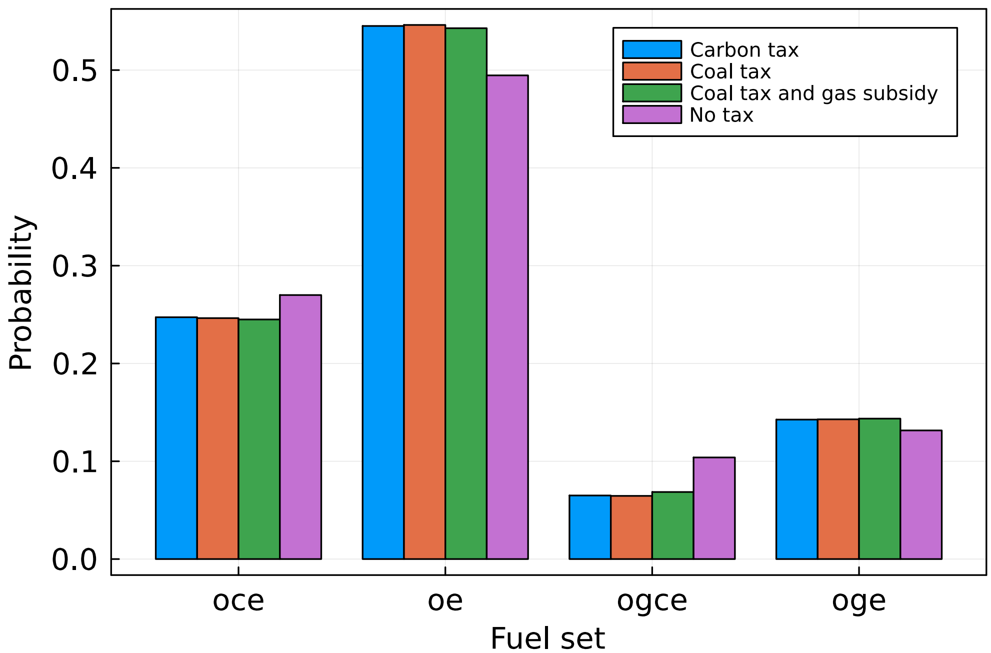


Figure 14: Probability of each fuel set choice

		No tax	Carbon Tax	Coal Tax
Externality damage (ED)	Gas	394	0.1	0.05
	Coal	437,049	51	0.42
	Oil	483	1.14	16
	Elec	394	1.05	0.88
	Total	438,320	53.4	17.35
Producer surplus (PS)	Operating profit	29,015	21.4	11.3
	Net switching costs	-6,296	-7,237.6	-7,236.7
	Total	35,311	7,259	7,248
Tax revenue (TR)		0	53.4	17.67
Consumer surplus (CS)		42.4	32.8	32.7
Welfare (CS+PS+TR-ED)		-403,352	7,291	7,264.8

Table 15: Welfare Analysis with Different Tax Regimes - Full Model

Notes: All units in the table (except consumer surplus) are in billion rupees. The negative net switching cost implies that firms get paid for switching. This may be because firms drop fuels more often than they add new fuels, but requisite a more careful investigation.

		No tax	Carbon Tax	Coal Tax
Externality damage (ED)	Gas	0.025	0.01	0.024
	Coal	4.34	0.15	0.026
	Oil	0.09	0.032	0.08
	Elec	0.44	0.31	0.43
	Total	4.89	0.5	0.57
Producer surplus (PS)	Operating profit	7.8	7.68	7.64
	Net switching costs	-7,171.7	-7,171.8	-7,171.8
	Total	7,179	7,179	7,179
Tax revenue (TR)		0	0.5	0.08
Consumer surplus (CS)		30.9	30.3	30.5
Welfare (CS+PS+TR-ED)		7,205.6	7209.9	7,209.5

Table 17: Welfare Analysis with Different Tax Regimes - No Fuel Productivity ($\psi_f = 1\ \forall f$)

Notes: All units in the table (except consumer surplus) are in billion rupees. The negative net switching cost implies that firms get paid for switching. This may be because firms drop fuels more often than they add new fuels, but requisite a more careful investigation.

References

Acemoglu, Daron and Pascual Restrepo, "Tasks, automation, and the rise in US wage inequality," *NBER Working Paper*, 2021, p. No. 28920.

Ackerberg, Daniel A., Kevin Caves, and Garth Frazer, "Identification Properties of Recent Production Function Estimators," *Econometrica*, 2015, *83* (6), 2411–2451.

Allcott, Hunt, Allan Collard-Wexler, and Stephen D. O'Connell, "How do electricity shortages affect industry? Evidence from India," *American Economic Review*, 2016, *106* (3), 587–624.

_ **and Michael Greenstone**, "Is There an Energy Efficiency Gap?," *Journal of Economic Perspectives*, 2012, *26* (1), 3–28.

Arcidiacono, Peter and John Bailey Jones, "Finite mixture distributions, sequential likelihood and the EM algorithm," *Econometrica*, 2003, *71* (3), 933–946.

_ **and Robert Jones**, "Conditional Choice Probability Estimation of Dynamic Discrete Choice Models With Unobserved Heterogeneity," *Econometrica*, 2011, *79* (6), 1823–1867.

Atkinson, Scott E and Robert Halvorsen, "Interfuel Substitution in Steam Electric Power Generation," *Journal of Political Economy*, 1976, *84* (5), 959–978.

Bagwell, Kyle and Seung Hoon Lee, "Trade Policy under Monopolistic Competition with Heterogeneous Firms and Quasi-linear CES Preferences," 2018.

Bartelsman, Eric J. and Mark Doms, "Understanding productivity: Lessons from longitudinal microdata," *Journal of Economic Literature*, 2000, *38* (3), 569–594.

Boehm, Johannes and Ezra Oberfield, "Misallocation in the Market for Inputs: Enforcement and the Organization of Production," *The Quarterly Journal of Economics*, 2020, *135* (4), 2007–2058.

Broda, Christian and David E Weinstein, "Globalization and the Gains from Variety," *The Quarterly Journal of Economics*, 2006, *121* (2), 541–585.

Buchanan, J. M., "External Diseconomies , Corrective Taxes, and Market Structure," *The American Economic Review*, 1969, *59* (1), 174–177.

Chan, Nathan W. and Kenneth Gillingham, "The microeconomic theory of the rebound effect and its welfare implications," *Journal of the Association of Environmental and Resource Economists*, 2015, *2* (1), 133–159.

Christensen, Peter, Paul Francisco, and Erica Myers, "Can Incentive-Based Pay Increase the Marginal Value of Public Spending on Energy Efficiency ? Experimental Evidence from the Weatherization Assistance Program," 2022.

Cicala, Steve, "Imperfect Markets versus Imperfect Regulation in US Electricity Generation," *American Economic Review*, 2022, *112* (2), 409–411.

Copeland, Brian R and M Scott Taylor, "Trade , Growth , and the Environment," *Journal of Economic Literature*, 2004, *42* (1), 7–71.

Davis, By Lucas W and Catherine Wolfram, "Deregulation, Consolidation , and Efficiency : Evidence from US Nuclear Power," *American Economic Journal : Applied Economics*, 2012, *4* (4), 194–225.

Demirer, Mert, "Production Function Estimation with Factor-Augmenting Technology: An Application to Markups," *Working Paper*, 2020.

Dietz, Simon, Frederick van der Ploeg, Armon Rezai, and Frank Venmans, "Are Economists Getting Climate Dynamics Right and Does it Matter?," *SSRN Electronic Journal*, 2021, *8* (5).

Doraszelski, Ulrich and Jordi Jaumandreu, "R and D and productivity: Estimating endogenous productivity," *Review of Economic Studies*, 2013, *80* (4), 1338–1383.

_ **and** _ , "Measuring the Bias of Technological Change," *Journal of Political Economy*, 2018, *126.*

Ethier, By Wilfred J, "National and International Returns to Scale in the Modern Theory of International Trade," *The American Economic Review*, 1982, *72* (3), 389–405.

Fabrizio, Kira R., Nancy L. Rose, and Catherine D. Wolfram, "Do markets reduce costs? Assessing the impact of regulatory restructuring on US electric generation efficiency," *American Economic Review*, 2007, *97* (4), 1250–1277.

Feenstra, Robert C., Robert Inklaar, and Marcel P. Timmer, "The Next Generation of the Penn World Table," *American Economic Review*, 2015, *105* (10), 3150–3182. available for download at www.ggdc.net/pwt.

Environmental Protection Agency - Center for Corporate Climate Leadership

Environmental Protection Agency - Center for Corporate Climate Leadership, "Emission Factors for Greenhouse Gas Inventories," 2022. data retrieved from the GHG Emission Factors Hub https://www.epa.gov/system/files/documents/2022-04/ghg_emission_factors_hub.pdf.

Fisher-Vanden, Karen, Erin T. Mansur, and Qiong Juliana Wang, "Electricity shortages and firm productivity: Evidence from China's industrial firms," *Journal of Development Economics*, 2015, *114*, 172–188.

Fowlie, Meredith and Robyn Meeks, "The economics of energy efficiency in developing countries," *Review of Environmental Economics and Policy*, 2021, *15* (2), 238–260.

_ , Mar Reguant, and Stephen P. Ryan, "Market-based emissions regulation and industry dynamics," *Journal of Political Economy*, 2016, *124* (1), 249–302.

Gandhi, Amit, Salvador Navarro, and David A. Rivers, "On the identification of gross output production functions," *Journal of Political Economy*, 2020, *128* (8), 2973–3016.

Gerarden, Todd D, Richard G Newell, and Robert N Stavins, "Assessing the Energy-Efficiency Gap," *Journal of Economic Literature*, 2017, *55*, 1486–1525.

Goldberg, Pinelopi Koujianou, Amit Kumar Khandelwal, Nina Pavcnik, and Petia Topalova, "Imported intermediate inputs and domestic product growth: Evidence from India," *Quarterly Journal of Economics*, 2010, *125* (4), 1727–1767.

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

Government of Canada, "National Pollutant Release Inventory," 2022.

Grieco, Paul L.E., Shengyu Li, and Hongsong Zhang, "Production function estimation with unobserved input price dispersion," *International Economic Review*, 2016, *57* (2), 665–690.

Gupta, Vaibhav, Tirtha Biswas, Deepa Janakiraman, and Karthik Ganesan, "Methodology Note - Greenhouse Gas Emissions of India - Subnational Estimates Industry Sector 2005 to 2015 series," Technical Report September 2019.

Hambel, Christoph, Holger Kraft, and Eduardo Schwartz, "The social cost of carbon in a non-cooperative world," *Journal of International Economics*, 2021, *131.*

Harrison, Ann, Ben Hyman, Leslie Martin, and Shanthi Nataraj, "When do Firms Go Green? Comparing Price Incentives with Command and Control Regulations in India," 2016.

Helpman, Elhanan and Oleg Itskhoki, "Labour Market Rigidities, Trade and Unemployment," *Review of Economic Studies*, 2010, *77* (3), 1100–1137.

Joskow, Paul L. and Frederic S. Mishkin, "Electric Utility Fuel Choice Behavior in the United States," *International Economic Review*, 1977, *18* (3), 719–736.

Kasahara, Hiroyuki and Joel Rodrigue, "Does the use of imported intermediates increase productivity? Plant-level evidence," *Journal of Development Economics*, 2008, *87* (1), 106–118.

Klette, Tor Jakob and Zvi Griliches, "The inconsistency of common scale estimators when output prices are unobserved and endogenous," *Journal of Applied Econometrics*, 1996, *11* (4), 343–361.

Kortum, Samuel, "Research, Patenting, and Technological Change," *Econometrica*, 1997, *65* (6), 1389–1419.

Kotchen, Matthew, "Taxing Externalities: Revenue vs. Welfare Gains with an Application to U.S. Carbon Taxes," 2022.

León-Ledesma, Miguel A., Peter McAdam, and Alpo Willman, "Identifying the elasticity of substitution with biased technical change," *American Economic Review*, 2010, *100* (4), 1330–1357.

Levinsohn, James and Amil Petrin, "Estimating production functions using inputs to control for unobservables," *Review of Economic Studies*, 2003, *70* (2), 317–341.

Luh, Sandro, Sara Budinis, Sara Giarola, Thomas J. Schmidt, and Adam Hawkes, "Long-term development of the industrial sector – Case study about electrification, fuel switching, and CCS in the USA," *Computers and Chemical Engineering*, 2020, *133*, 106602.

Mahadevan, Meera, "You Get What You Pay For: Electricity Quality and Firm Response," *SSRN Electronic Journal*, 2022, pp. 1–31.

Martin, Leslie A., Shanthi Nataraj, and Ann E. Harrison, "In with the big, out with the small: Removing small- scale reservations in India," *American Economic Review*, 2017, *107* (2), 354–386.

Miftakhova, Alena and Clément Renoir, "Economic Growth and Equity in Anticipation of Climate Policy," *SSRN Electronic Journal*, 2021, (June).

Oberfield, Ezra, "A Theory of Input-Output Architecture," *Econometrica*, 2018, *86* (2), 559–589.

Olley, G. Steven and Ariel Pakes, "The Dynamics of Productivity in the Telecommunications Equipment Industry," *Econometrica*, 1996, *64* (6), 1263–1297.

Petroleum and Natural Gas Regulatory Board

Petroleum and Natural Gas Regulatory Board, "Data Bank," 2023. data retrieved from the PNGRB's data bank, https://www.pngrb.gov.in/eng-web/data-bank.html#.

Ramanarayanan, Ananth, "Imported inputs and the gains from trade," *Journal of International Economics*, 2020, *122*, 103260.

Rehfeldt, Matthias, Tobias Fleiter, Andrea Herbst, and Stefan Eidelloth, "Fuel switching as an option for medium-term emission reduction - A model-based analysis of reactions to price signals and regulatory action in German industry," *Energy Policy*, 2020, *147* (September), 111889.

Romer, Paul M, "Endogenous Technological Change," *Journal of Political Economy*, 1990, *98* (5).

Roy, A. D., "Some thoughts on the distribution of earnings," *Oxford Economic Papers*, 1951, *3* (2), 135–146.

Ryan, Nicholas, "The Competitive Effects of Transmission Infrastructure in the Indian Electricity Market," *American Economic Journal: Microeconomics*, 2021, *13* (2), 202–242.

Ryan, Stephen P, "The Costs of Environmental Regulation in a Concentrated Industry," *Econometrica*, 2012, *80* (3), 1019–1061.

Scott, Jonathan B., "Positive Spillovers from Infrastructure Investment: How Pipeline Expansions Encourage Fuel Switching," *The Review of Economics and Statistics*, 2021, pp. 1–43.

Shapiro, Joseph S. and Reed Walker, "Why is pollution from us manufacturing declining? the roles of environmental regulation, productivity, and trade," *American Economic Review*, 2018, *108* (12), 3814–3854.

Syverson, Chad, "What determines productivity," *Journal of Economic Literature*, 2011, *49* (2), 326–365.

Worrell, Ernst, Lenny Bernstein, Joyashree Roy, Lynn Price, and Jochen Harnisch, "Industrial energy efficiency and climate change mitigation," *Energy Efficiency*, 2009, *2* (2), 109–123.

Zhang, Hongsong, "Non-neutral technology, firm heterogeneity, and labor demand," *Journal of Development Economics*, 2019, *140* (May), 145–168.

A Data

A.1 Details on sampling rules

A.2 Fuel Productivity and Distinction Between Potential and Realized Energy

Energy inputs are measured in different units. For example, coal is typically measured in weight whereas natural gas is typically measured in volume. As a result, scientific calculations converts baseline fuel quantities into equivalent heating potential (million British thermal units, mmBtu). In this paper, I call this *potential energy.* This is because it captures what energy may be extracted from combustion of a particular fuel.

However, what firms get in terms of energy service from the combustion process, which I call *realized energy*, depends on a variety of factors, such as the technology used for combustion and firms' knowledge on wasting energy. In essence, realized energy is what firms get after combining fuels with some technology. As such, there is a conceptual gap between potential and realized energy, which underlies productivity differences. These differences come in many forms, and I highlight three examples:

1. **Across fuel types:** In the transformation of liquid iron into liquid steel, electric-arc furnaces which use a combination of electricity, natural gas and recycled materials, are more productive than coal furnaces at using heating potentials of the underlying fuels (**Worrell, Blinde, Neelis, Blomen and Masanet 2010**).

2. **Within fuel types:** In coal manufacturing, Coal used in rotary kilns is more productive than in vertical shaft kilns for the production of clinker as part of cement manufacturing **(Christina Galitsky and Lynn Price,2007)**.

3. **Wasted resources:** Energy retrofit programs underlie large heterogeneity differences on how efficiently agents in the economy use the heating potential of fuels **(Christensen, Francisco and Myers 2022)**. Examples include keeping lights opened unnecessarily or forgetting to turn off machinery.

A.3 Fuel Prices and Transportation Costs

Identification of firms' responses to changes in fuel prices rests on two important sources of price variation. First, it relies on persistent shocks that are largely driven by worldwide variation in supply and demand related to macroeconomic conditions and geopolitical events such as wars, trade agreements, and sanctions. **Figure Y** shows the evolution in the median fuel prices paid

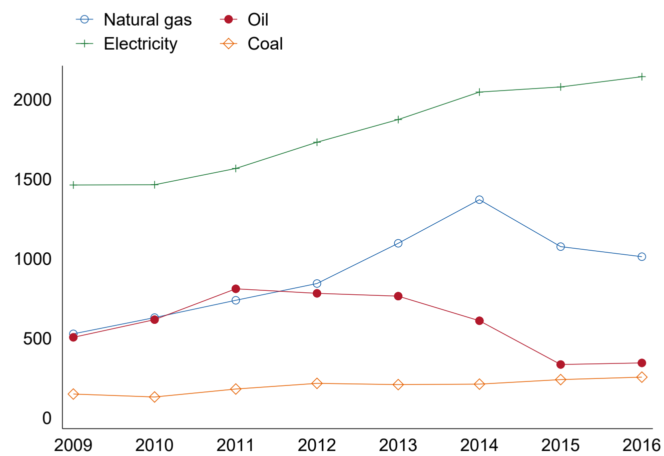


Figure 15: Yearly Median Fuel Prices (INR/mmBtu)

by ASI plants. Notably, the oil shock of 2014 led to a 50% decrease in the price of oil and a 30% decrease in the price of natural gas. At the same time, the price of coal is much more stable. This is largely because there are less geopolitical events surrounding coal due to the general phase-out of coal in developed countries (**source**). This will play an important role in the government's provision of insurance against price shocks through taxation.

Second, identification relies on spatial variation in fuel prices, which I argue is related to transportation costs. As an example, natural gas is expensive to transport because it needs to be carried in high pressure pipelines. The Petroleum and Natural Gas Regulatory Board of India (PNGRB) sets transportation prices according to a 4 zone schedule in a vicinity of 10 km on both sides of the pipeline: 1 being the closest to the source and 4 being farthest from the source of the pipeline[11]. By 2016, there was 13 gas pipeline networks, each with their own 1-4 zone tariffs (depending on the length of the pipeline). However, different pipelines have different baseline transportation costs, such that it is possible for a plant in the zone 4 of a pipeline to pay less than a plant in the zone 1 of another pipeline. For example, transportation costs the zone 4 of the integrated Hazira-Vijaipur-Jagdishpur pipeline costs 49 INR/mmBtu, whereas transportation costs in the zone 1 of the East West Gas Pipeline (PNGRB) is 65.5 INR/mmBtu. If the plant is not in a vicinity of a pipeline, it can carry liquefied natural gas (LNG), but it needs to re-gasify it which is costly. Below is a schema describing how the natural gas pipeline tariffs work:

Overall, the transportation cost structure of natural gas should lead to large dispersion in the price of natural gas that plants pay. On the contrary, coal is much simpler to transport because it

[11] The Indian government is considering changing its pricing structure, and it would be an interesting counterfactual to consider

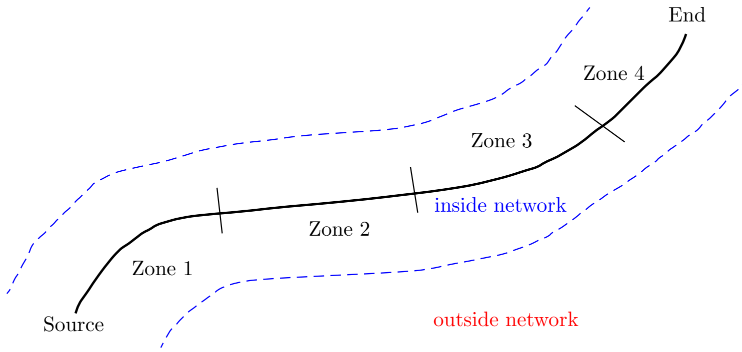


Figure 16: Hypothetical Structure of Transportation Costs for Natural gas Pipeline

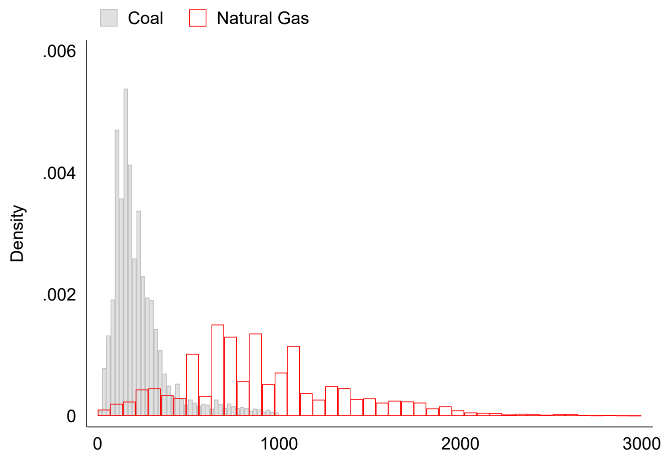


Figure 17: Histogram comparing price of natural gas and coal (INR/mmBtu)

is a solid and because it is mostly extracted domestically[12]. As such, 17% of all coal is transported directly from the mine to plants through conveyor belts, 33% is transported by road, and 50% is transported by train. These cheaper and simpler transportation methods should lead to lower dispersion in the price of coal. If fuel prices in the ASI reflect differences in transportation costs, then the price distribution should reflect this difference in dispersion. This is indeed what I find, as **figure Y** suggests a much larger dispersion in the price of gas relative to that of coal.

Moreover, I find that accounting for pipeline fixed effects, there is a positive and significant jump in the price of natural gas from being in zone 2-4 relative to zone 1. However, the effect for zone 4 does not seem robust. Zones and pipeline data were constructed by mapping the entire natural gas pipeline network to the districts in which they pass, directly or indirectly. Thus these results are subject to measurement error.

Lastly, I argue that spatial fuel price variation captures some exogenous variation from the

[12]**Khanna 2021** shows that Coal India Limited (CIL) is the largest coal mining company in the world

Table 19: Relationship between (log) natural gas prices and proximity to pipelines

	(1) (log) P_{natgas}		(2) (log) P_{natgas}		(3) (log) P_{natgas}	
Zone 2	0.278***	(0.044)	0.235***	(0.045)	0.219***	(0.045)
Zone 3	0.214***	(0.046)	0.176***	(0.047)	0.163***	(0.047)
Zone 4	0.119***	(0.033)	0.052	(0.036)	0.038	(0.035)
year dummies	Yes		Yes		Yes	
Pipeline dummies	Yes		Yes		Yes	
Industry dummies			Yes		Yes	
Additional controls					Yes	
Observations	11,780		11,780		11,780	

Standard errors in parentheses

Baseline zone is 1 (closest to source of pipeline).

Additional controls: number of workers and quantity of gas purchased.

$^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

firms' perspective because plants location decisions are somewhat constrained by the language locals speak. Indeed, there are 22 official regional languages in India, which are broadly related to one of 28 States. For examples, Bengali is the main language in West Bengal, Gujarati is the main language in Gujarat, Punjabi is the main language in Punjab, and so on. For this reason, I will use States as the main driver of spatial price variation in the model.

A.4 Emissions Data

A.4.1 Calculation of Emissions

Excluding Electricity

To get establishment-level measures of greenhouse gas emissions, I convert units of potential energy (mmBtu) of each fuel into metric tons of carbon dioxide equivalent, as a result of combustion. Each mmBtu of fuel releases some quantity of carbon dioxide CO_2, methane CH_4, and nitrous oxide N_2O in the air, which may vary by industry based on standard practices and technology. Emissions of chemical k for a plant in industry j can be calculated as follows:

$$emissions_{jk} = \sum_f \sum_k \zeta_{fkj} * e_f$$

$$\forall\ k = \{CO_2, CH_4, N_2O\} \quad \forall\ f = \{\text{Natural Gas, Coal, Oil}\}$$

The fuel-by-industry emission factors of each chemical ζ_{fkj} are found in the database provided by GHG Platform India, and come from two main sources: India's Second Biennial Update Report (BUR) to United Nations Framework Convention on Climate Change (UNFCCC) and IPCC Guidelines. Quantities in mmBtu of each fuel e_f are observed for each establishment in each year. Then, quantities of each chemical is converted into carbon dioxide equivalent CO_2e using the Global Warming Potential (GWP) method as follows:

$$CO_2e = \underbrace{GWP_{co2}}_{=1} * CO_2 + GWP_{ch4} * CH_4 + GWP_{n2o} * N_2O$$

From the calculations above, I can define fuel-specific emission factors which will be used to directly convert fuels to CO_2e (or GHG):

$$\gamma_{fj} = GWP_{co2} * \zeta_{f,co2,j} + GWP_{ch4} * \zeta_{f,ch4,j} + GWP_{n2o} * \zeta_{f,n2o,j}$$

Total greenhouse gas emissions in units of CO_2e for plant i in industry j and year t is defined as follows:

$$GHG_{ijt} = \sum_{f \in \{natgas, coal, oil\}} \gamma_{fj} * e_{fijt}$$

Including Electricity

Calculations of emissions from electricity is done slightly differently than from fossil fuels because emissions comes from production rather than end usage of electricity. Figure 1 shows that coal is used to consistently generate above 60% of total electricity in India, which increased in 2010 and started to decrease after 2012.

To construct measures of emissions from electricity, I will take the distribution of emissions from different fuels used to produce electricity, averaged across years for the entire grid. Let $\omega_{ef} \in [0, 1]$ $\forall f \in \{Coal, Gas\}$ be the share of fuel f used to generate electricity, then

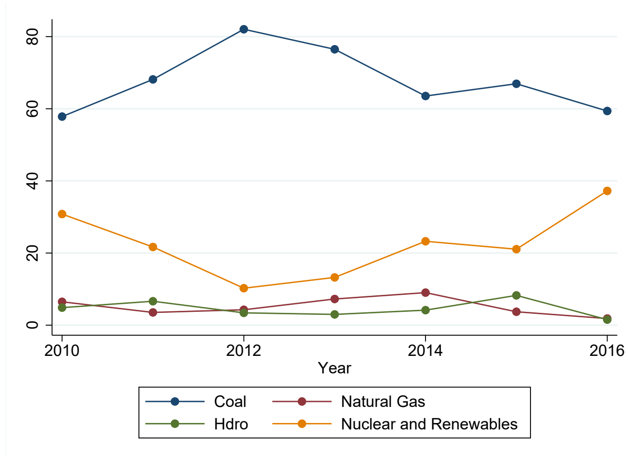


Figure 18: Annual Indian Electricity Generation by Source (% of Total)
Source: International Energy Agency (IEA)

$$\gamma_e = \sum_{f \in \{coal,gas\}} \omega_{ef} * \gamma_{ef}$$

Where γ_{ef} is calculated exactly as in 2.2.1 for the electricity generation industry. Including electricity, total GHG emissions for plant i in industry j and year t is defined as:

$$GHG_{ijt} = \gamma_e * e_{eijt} + \sum_{f \in \{natgas,coal,oil\}} \gamma_{fj} * e_{fijt}$$

Below are the tables detailing emissions factors. Note that for oil, I take hte average over all pretoleum fuels. The dispersion between oil types is much lower than the dispersion between the average of oil and coal/gas.

A.5 Evidence on Switching and Mixing

Indian Plants

U.S. plants

Here I show some of the evidence presented in the main text from manufacturing plants located in the U.S. The data is from the Greenhouse Gas Reporting Program (GHGRP), which reports fuel consumption (oil, gas, coal) from large manufacturing plants in selected industries. Below I show evidence from the Pulp & Paper industry between 2010 and 2018.

		Emission factors (kg CO_2e/mmBtu)			
Fuel	*Industry*	CO_2	CH_4	N_2O	Total (γ_{fj})
	Cement	100.90	0.03	0.42	101.34
	Non-ferrous metals	101.67	0.03	0.42	102.11
Coal	Pulp and paper	101.59	0.03	0.42	102.04
	Electricity generation	102.09	0.03	0.42	102.54
	Other	98.84	0.03	0.42	99.29
Oil	All	77.34	0.09	0.17	77.59
Natural Gas	All	50.64	0.03	0.03	50.70

Table 20: Emission factors from fuels to carbon dioxide equivalent $\zeta_{fkj} * GWP_k$ (kg CO_2e/mmBtu). Source: GHG Platform India

Share of Electricity Generated by Source				
Natural Gas	Coal	Hydro	Other	Emission factor (kg CO_2e/mmBtu)
0.052	0.68	0.046	0.23	72.05

Table 21: Emission factors from Electricity

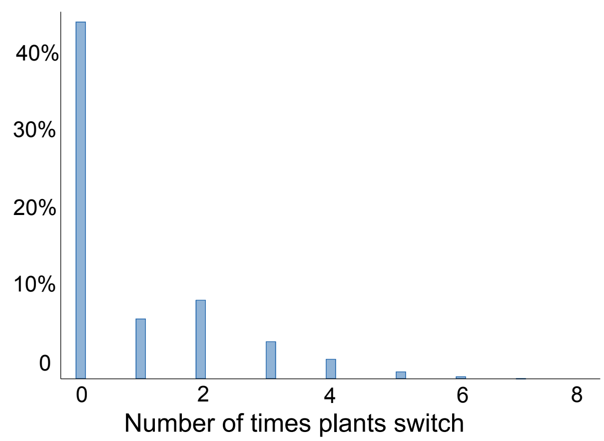


Figure 19: Number of Times Unique Plants add or drop a Fuel (ASI)

Table 22: Different Fuel Sets

	Frequency	%
Natural Gas	602	50.76
Oil	36	3.04
Natural Gas, Coal	72	6.07
Natural Gas, Oil	332	27.99
Coal, Oil	9	0.76
Natural Gas, Coal, Oil	135	11.38
Total	1186	100.00

Table 23: Percentage of unique plants that **add** and **drop** a fuel

	Adds New Fuel (%)	Drops Existing Fuel (%)
No	77.13	76.68
Yes	22.87	23.32
Total	100.0	100.0

Pulp and Paper Manufacturing (U.S. GHGRP)

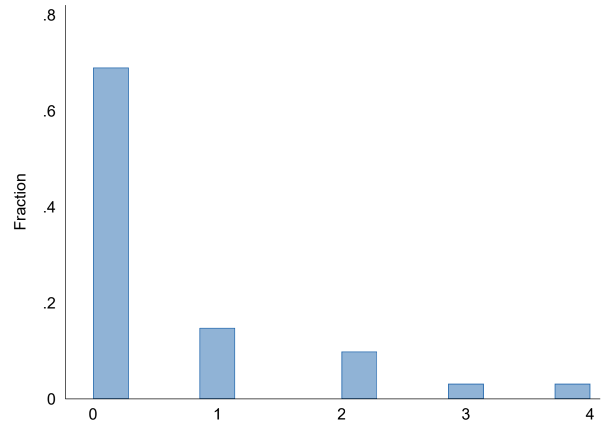


Figure 20: Number of Times Unique Plants add or drop a Fuel (U.S. GHGRP - Pulp & Paper Manufacturing)

A.5.1 Fuel Expenditure Shares - ASI

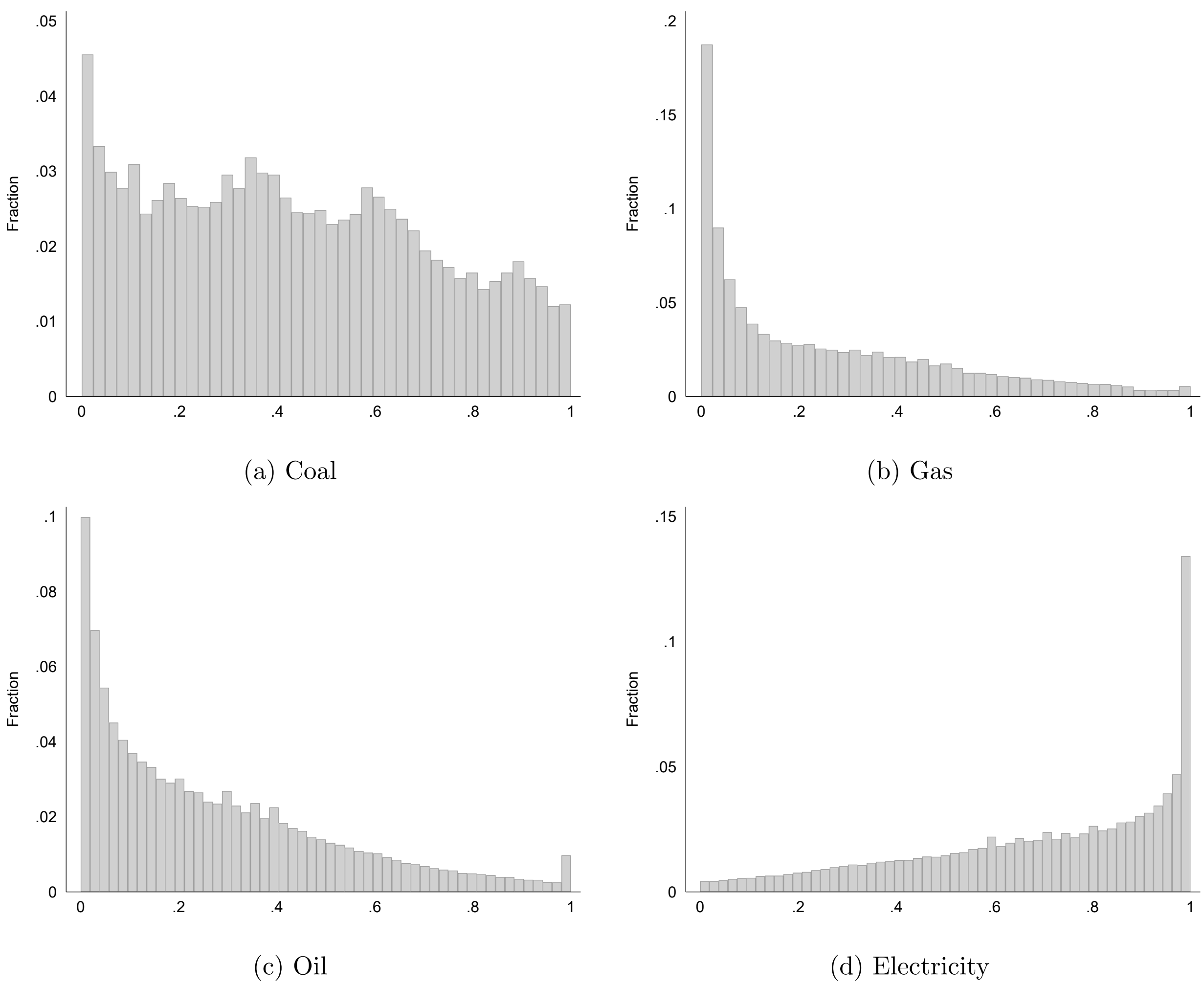


Figure 21: Distribution of Expenditure Shares

B Model

B.1 Aggregation with Monopolistic Competition

Each period, plants in a given industry produce differentiated goods and engage in monopolistic competition. In each industry, there is a representative consumer with quasi-linear CES utility function that aggregates all differentiated goods:

$$U(Y_{it}, Y_{0t}) = \max_{Y_{0t},\{Y_{it}\}} Y_{0t} + \frac{1}{\zeta}\Big(\int_i Y_{it}^{\frac{\rho-1}{\rho}} di\Big)^{\frac{\zeta\rho}{\rho-1}} \tag{19}$$

s.t.

$$Y_{0t} + \int_i P_{it} Y_{it} d_i = I_t = \Pi_t \tag{20}$$

Where ρ is the demand elasticity, ζ is the elasticity between firms' output and the outside good, and I is the consumer's income. This quasi-linear specification has been used by Bagwell and Lee (2018) and Helpman and Itskhoki (2010), and will be useful to study optimal tax policies later. From **(11)** and **(12)**, plant i faces the following inverse demand:

$$P_{it}(Y_{it};\rho,\zeta) = P_t^{\frac{\rho(1-\zeta)-1}{1-\zeta}} Y_{it}^{-\frac{1}{\rho}} \tag{21}$$

$P_t = \Big(\int_i P_{it}^{1-\rho} di\Big)^{\frac{1}{1-\rho}}$ is the industry-level price index and p_{it} is firm's i output price at time t.

B.2 Aggregation of production function without switching

In this section, I show that when firms do not have fuel-augmenting productivity and cannot switch between fuel sets, the economy can be aggregated into a single CES production function similar to the one of **Golosov et al. 2014** who study optimal externality taxes on fossil fuels in an aggregate economy. This allows me to benchmark results from my model with the existing literature. For reference, **Golosov et al. 2014** postulate the existence of an aggregate production Cobb-Douglas production function which nests an aggregate CES production function for energy. The aggregate

CES production function for energy takes the following form, where f indexes fuels

$$
\begin{aligned}
E &= \left(\sum_{f \in \{o,g,c,e\}} \beta_f e_f^{\frac{\lambda-1}{\lambda}} \right)^{\frac{\lambda}{\lambda-1}} \\
\sum_f \beta_f &= 1
\end{aligned} \tag{22}
$$

In my paper, there are multiple firms with a different fuel sets $\mathcal{F}$ available to them. To be consistent with **Golosov et al. 2014** and get aggregation results, I assume that all firms are identical but differ in the fuel set available to them $\mathcal{F} \subset \mathbb{F} = \Big(\{o,e\}, \{o,g,e\}, \{o,c,e\}, \{o,g,c,e\} \Big)$. Then, firms in each fuel set have the following production function:

$$
E_{\mathcal{F}} = \left(\sum_{f \in \mathcal{F}} \beta_f e_f^{\frac{\lambda-1}{\lambda}} \right)^{\frac{\lambda}{\lambda-1}} \tag{23}
$$

From equation (23) and cost-minimization, I can solve for the quantity of each fuel demanded $e_f(\mathcal{F})$ given fuel prices and fuel sets as

$$
e_f(\mathcal{F}) = E \left(\frac{\beta_f}{p_f} \right)^{\lambda} P_E(\mathcal{F})^{\lambda}
$$

For pre-determined quantity of energy E, where $P_E(\mathcal{F})$ is the energy price index of firms using fuel set $\mathcal{F}$.

$$
P_E(\mathcal{F}) = \Big(\sum_{f \in \mathcal{F}} \beta_f^{\lambda} p_f^{1-\lambda} \Big)^{\frac{1}{1-\lambda}}
$$

Let $s_{oe}, s_{oge}, s_{oce}, s_{ogce}$ be the share of firms that use each fuel sets such that $s_{oe} + s_{oge} + s_{oce} + s_{ogce} = 1$. I can then use these share of firms in each fuel set to define the total quantity of each fuel demanded by summing over all fuel set that use fuel f.

$$
\begin{aligned}
e_f &= \sum_{\mathcal{F}} \mathbb{I}(f \in \mathcal{F}) s_{\mathcal{F}} e_f(\mathcal{F}) \\
&= E\left(\frac{\beta_f}{p_f}\right)^{\lambda} \left(\sum_{\mathcal{F}} \mathbb{I}(f \in \mathcal{F}) s_{\mathcal{F}} P_E(\mathcal{F})\right)
\end{aligned} \tag{24}
$$

I postulate that there exist an aggregate CES energy production function $\tilde{E}$ such that the total quantity demanded of each fuel is equal to (24).

Proposition 4. *There exist an aggregate energy production function in all fuels $\tilde{E}$ with aggregate productivity Ψ such that cost-minimizing input quantities $\tilde{e}_f$ are the same as cost-minimizing input quantities in equation (24).*

Proof. I show this proposition by constructing the following production function:

$$
\tilde{E} = \left(\sum_{f \in \{o,g,c,e\}} \psi_f^{\frac{1}{\lambda}} e_f^{\frac{\lambda-1}{\lambda}} \right)^{\frac{\lambda}{\lambda-1}} \tag{25}
$$

Where ψ_f is the endogenous loading of each fuel into the production function. As I show below, it takes into account both the share of each fuel in the original production function β_f as well as the share of firms who are using each fuels. The cost-minimizing quantity of each fuel from the production function in (25) for a given quantity of energy E is

$$
\tilde{e}_f = s_f E \left(\frac{P_{\tilde{E}}}{p_f} \right)^{\lambda} \tag{26}
$$

Where $P_{\tilde{E}}$ is the price index of energy:

$$
P_{\tilde{E}} = \left(\sum_{f \in \{o,g,c,e\}} \psi_f \beta_f^{\lambda} p_f^{1-\lambda} \right)^{\frac{1}{1-\lambda}}
$$

Then, the loadings on each fuels are implicitely defined by

$$\psi_f = \frac{\beta_f^\lambda \sum_{\mathcal{F}} I(f \in \mathcal{F}) s_{\mathcal{F}} P_E(\mathcal{F})}{P_{\tilde{E}}^\lambda}$$

Then, $\tilde{e}_f$ from (26) is equal to e_f from (24).

$\square$

C Identification

C.1 Equivalence (non-Identification) of Monopolistic Competition and Variable Returns to Scale

Constant Returns to Scale and Monopolistic Competition

Production function:

$$\frac{Y_{it}}{\overline{Y}} = e^{\omega_{it}} \left(\alpha_k \left(\frac{K_{it}}{\overline{K}} \right)^{\frac{\sigma-1}{\sigma}} + \alpha_l \left(\frac{L_{it}}{\overline{L}} \right)^{\frac{\sigma-1}{\sigma}} + \alpha_m \left(\frac{M_{it}}{\overline{M}} \right)^{\frac{\sigma-1}{\sigma}} + \alpha_e \left(\frac{E_{it}}{\overline{E}} \right)^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}} \tag{27}$$

$$= e^{\omega_{it}} \left(\alpha_k \tilde{K}_{it}^{\frac{\sigma-1}{\sigma}} + \alpha_l \tilde{L}_{it}^{\frac{\sigma-1}{\sigma}} + \alpha_m \tilde{M}_{it}^{\frac{\sigma-1}{\sigma}} + \alpha_e \tilde{E}_{it}^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}} \tag{28}$$

Where I define $\frac{X_{it}}{\overline{X}} = \tilde{X}_{it}$

Demand:

$$P(Y_{it}) = P_t \left(\frac{Y_t}{Y_{it}} \right)^{\frac{1}{\rho}} \tag{29}$$

Revenues:

$$\begin{aligned} R_{it} &= P(Y_{it}) Y_{it} e^{u_{it}} \\ &= P_t Y_t^{\frac{1}{\rho}} Y_{it}^{\frac{\rho-1}{\rho}} e^{u_{it}} \end{aligned}$$

Assumption 1-. *L_{it}, M_{it}, E_{it} are flexible inputs*

Assumption 2-.. *I observe the quantity for L_{it} and K_{it} but only spending for materials and energy: $S_{L_{it}}, S_{E_{it}}$*

Profit-maximization:

$$\max_{L_{it},M_{it},E_{it}} \Big\{ P_t Y_t^{\frac{1}{\rho}} Y_{it}^{\frac{\rho-1}{\rho}} - p_{mit}M_{it} - p_{eit}E_{it} - p_{lit}L_{it} \Big\}$$
$$s.t. Y_{it} = \overline{Y} e^{\omega_{it}} \Big(\alpha_k \tilde{K}_{it}^{\frac{\sigma-1}{\sigma}} + \alpha_l \tilde{L}_{it}^{\frac{\sigma-1}{\sigma}} + \alpha_m \tilde{M}_{it}^{\frac{\sigma-1}{\sigma}} + \alpha_e \tilde{E}_{it}^{\frac{\sigma-1}{\sigma}} \Big)^{\frac{\sigma}{\sigma-1}}$$

First-order conditions:

M_{it}/L_{it}:

$$\frac{M_{it}}{\overline{M}} = \Big(\frac{\alpha_l}{\alpha_m} \frac{S_{Mit}}{S_{Lit}} \Big)^{\frac{\sigma}{\sigma-1}} \frac{L_{it}}{\overline{L}} \tag{30}$$

E_{it}/L_{it}:

$$\frac{E_{it}}{\overline{E}} = \Big(\frac{\alpha_l}{\alpha_e} \frac{S_{Eit}}{S_{Lit}} \Big)^{\frac{\sigma}{\sigma-1}} \frac{L_{it}}{\overline{L}} \tag{31}$$

L_{it}:

$$\frac{\rho-1}{\rho} P_t Y_t^{\frac{1}{\rho}} Y_{it}^{-\frac{1}{\rho}} e^{\omega_{it}} \Big(\alpha_k \tilde{K}_{it}^{\frac{\sigma-1}{\sigma}} + \alpha_l \tilde{L}_{it}^{\frac{\sigma-1}{\sigma}} + \alpha_m \tilde{M}_{it}^{\frac{\sigma-1}{\sigma}} + \alpha_e \tilde{E}_{it}^{\frac{\sigma-1}{\sigma}} \Big)^{\frac{1}{\sigma-1}} \alpha_l \tilde{L}_{it}^{\frac{\sigma-1}{\sigma}} = S_{Lit}$$

Solve for TFP:

$$e^{\omega_{it} \frac{\rho-1}{\rho}} = \frac{1}{\overline{Y}^{\frac{\rho-1}{\rho}}} \frac{\rho}{\rho-1} \frac{S_{Lit}}{\alpha_L \tilde{L}_{it}^{\frac{\sigma-1}{\sigma}}} \frac{1}{P_t Y_t^{\frac{1}{\rho}}} \Big(\alpha_k \tilde{K}_{it}^{\frac{\sigma-1}{\sigma}} + \alpha_l \tilde{L}_{it}^{\frac{\sigma-1}{\sigma}} + \alpha_m \tilde{M}_{it}^{\frac{\sigma-1}{\sigma}} + \alpha_e \tilde{E}_{it}^{\frac{\sigma-1}{\sigma}} \Big)^{\frac{\sigma-\rho}{\rho(\sigma-1)}} \tag{32}$$

Plug (6) into revenue equation:

$$R_{it} = P_t Y_t^{\frac{1}{\rho}} e^{\omega_{it}\frac{\rho-1}{\rho}} \left(\alpha_k \tilde{K}_{it}^{\frac{\sigma-1}{\sigma}} + \alpha_l \tilde{L}_{it}^{\frac{\sigma-1}{\sigma}} + \alpha_m \tilde{M}_{it}^{\frac{\sigma-1}{\sigma}} + \alpha_e \tilde{E}_{it}^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\rho-1}{\rho}\frac{\sigma}{\sigma-1}} e^{u_{it}}$$
$$= \frac{\rho}{\rho-1} \frac{S_{Lit}}{\alpha_L \tilde{L}_{it}^{\frac{\sigma-1}{\sigma}}} \left(\alpha_k \tilde{K}_{it}^{\frac{\sigma-1}{\sigma}} + \alpha_l \tilde{L}_{it}^{\frac{\sigma-1}{\sigma}} + \alpha_m \tilde{M}_{it}^{\frac{\sigma-1}{\sigma}} + \alpha_e \tilde{E}_{it}^{\frac{\sigma-1}{\sigma}}\right) e^{u_{it}}$$

Plug ratio of FOCs (4) and (5) into the previous equation:

$$R_{it} = \frac{\rho}{\rho-1} S_{Lit} \left(\frac{\alpha_k}{\alpha_L} \left(\frac{\tilde{K}_{it}}{L_{it}}\right)^{\frac{\sigma-1}{\sigma}} + 1 + \frac{S_{Mit}}{S_{Lit}} + \frac{S_{Eit}}{S_{Lit}} \right) e^{u_{it}}$$
$$= \frac{\rho}{\rho-1} \left(S_{Lit} \left(1 + \frac{\alpha_k}{\alpha_L} \left(\frac{\tilde{K}_{it}}{\tilde{L}_{it}}\right)^{\frac{\sigma-1}{\sigma}} \right) + S_{Mit} + S_{Eit} \right) e^{u_{it}}$$

Estimating Equation:

$$\ln R_{it} = \ln \frac{\rho}{\rho-1} + \ln \left(S_{Lit} \left(1 + \frac{\alpha_k}{\alpha_L} \left(\frac{\tilde{K}_{it}}{\tilde{L}_{it}}\right)^{\frac{\sigma-1}{\sigma}} \right) + S_{Mit} + S_{Eit} \right) + u_{it} \tag{33}$$

Comparing this equation with the main estimating equation in the text (**X**) shows that defining the returns to scale as $\eta = \frac{\rho-1}{\rho}$ yields the exact same estimating equation in both economy.

Adding Variable Returns to Scale

Production Function:

$$Y_{it} = \overline{Y} e^{\omega_{it}} \left(\alpha_k \tilde{K}_{it}^{\frac{\sigma-1}{\sigma}} + \alpha_l \tilde{L}_{it}^{\frac{\sigma-1}{\sigma}} + \alpha_m \tilde{M}_{it}^{\frac{\sigma-1}{\sigma}} + \alpha_e \tilde{E}_{it}^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\eta\sigma}{\sigma-1}} \tag{34}$$

Equation (6) for productivity revised :

$$e^{\omega_{it}\frac{\rho-1}{\rho}} = \frac{1}{\eta} \frac{\rho}{\rho-1} \frac{S_{Lit}}{\alpha_L \tilde{L}_{it}^{\frac{\sigma-1}{\sigma}}} \frac{1}{P_t Y_t^{\frac{1}{\rho}}} \left(\alpha_k \tilde{K}_{it}^{\frac{\sigma-1}{\sigma}} + \alpha_l \tilde{L}_{it}^{\frac{\sigma-1}{\sigma}} + \alpha_m \tilde{M}_{it}^{\frac{\sigma-1}{\sigma}} + \alpha_e \tilde{E}_{it}^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\rho(\sigma-1)-\eta\sigma(\rho-1)}{\rho(\sigma-1)}} \tag{35}$$

Estimating Equation (7) revised:

$$\ln R_{it} = \ln \frac{\rho}{\rho - 1} + \ln \frac{1}{\eta} + \ln \left(S_{Lit} \Big(1 + \frac{\alpha_k}{\alpha_L} \Big(\frac{\tilde{K}_{it}}{\tilde{L}_{it}} \Big)^{\frac{\sigma-1}{\sigma}} \Big) + S_{Mit} + S_{Eit} \right) + u_{it} \tag{36}$$

Hence, the markup ρ is not separately identified from the returns to scale ν.

C.2 Monte-carlo simulations to recover the distribution of comparative advantages over selected fuels

I create a sample of firms with all state variables present in the main model. External estimation of parameters governing the distribution of random effect from the sample of firms who use gas and/coal leads to upward biased estimates. Indeed, firms with larger comparative advantage to use coal are more likely to use coal, and likewise for gas. Monte-Carlo simulations confirms this intuition:

	Natural Gas					
	μ_{p_g}	$\sigma^2_{\psi_g}$	$\sigma^2_{p_g}$	$\sigma_{\psi_g p_g}$	μ_g	$\sigma^2_{\mu_g}$
Unselected Sample ($N = 3,000$)	-0.0004 (0.003)	0.03	0.03	-0.0122	0.28	0.48
Selected Sample ($N = 694$)	0.005 (0.006)	0.03	0.028	-0.0128	0.41	0.48
True value	0	0.03	0.03	-0.0120	0.3	0.5

Table 24: Selected and unselected state transition parameters for price and productivity of natural gas (Monte-carlo data, standard errors in parenthesis)

	Coal					
	μ_{p_c}	$\sigma^2_{\psi_c}$	$\sigma^2_{p_c}$	$\sigma_{\psi_c p_c}$	μ_c	$\sigma^2_{\mu_c}$
Unselected Sample ($N = 3,000$)	0.003 (0.002)	0.21	0.01	-0.018	0.19	0.39
Selected Sample ($N = 936$)	0.002 (0.003)	0.21	0.01	-0.019	0.25	0.39
True value	0	0.2	0.01	-0.0179	0.2	0.4

Table 25: Selected and unselected state transition parameters for price and productivity of natural gas (Monte-carlo data, standard errors in parenthesis)

In this Monte-Carlo simulation, there isn't much selection going on for coal because gas has a higher average productivity so firms are mostly selecting on the basis of gas, and almost all firms either use oil and electricity or oil,electricity, gas and coal.

C.3 Alternative Energy Production Models

In this section, I show how the model and identification can be adapted to different energy production function models, including a task-based model and a non-parametric model. I worked out identification and estimation of the energy task model and I show that the gains from variety argument also holds in the energy task model, but some additional assumptions are required to identify the non-parametric model

C.3.1 Energy Task Model

for a given quantity of realized energy, firms allocate fuels to energy tasks that compose a given unit of E in an inner nests. This task model is very similar to Acemoglu and Restrepo (2021) which I adapt to study energy substitution. It features the assignment of a mix of energy inputs to energy tasks which allows for flexible variation in input usage. Specifically, the inner nest features a continuum of energy tasks in the $\omega \in (0, 1)$ interval that are perfect complements in producing a unit of E.

$$E = \inf \left\{ \tau(\omega) : \omega \in [0, 1] \right\} \tag{37}$$

I assume that tasks are perfect complements, where firms have to complete steps that are necessary for production. However, the task production function can be relaxed to more general functional forms like CES. Each energy task $\tau(\omega)$ can be performed with physical quantities of fuels e_f that are available in the firm's fuel set $\mathcal{F}$, where fuels are in principle perfectly substitutable at performing each task:

$$\tau(\omega) = \sum_{f \in \mathcal{F}} \psi_f(\omega) e_f(\omega) \tag{38}$$

Where $\psi_f(\omega)$ are fuel-by-task specific productivity terms. The latter is important distinction from standard task-based production models and allows for very flexible input usage. To motivate this framework one should be thinking of tasks such as the steps required to produce crude steel: preparation of raw material, conversion of iron ore into iron, and conversion of iron into crude steel (Luh, Budinis, Giarola, Schmidt and Hawkes, 2020). The preparation of raw materials typically requires coal whereas the two subsequent steps can be done with different fuels and the $\psi_f(\omega)$ terms can reflect the different fuel-specific technologies that the plant can use for each step. Moreover, these steps are complementary and require high amounts of energy.

Going back to the model, the inner nest problem can be solved in two steps:

1. Find the cheapest fuel to perform each task
2. Aggregate across tasks into fuel categories to get fuel demand.

C.4 Task choices:

Minimize the cost of producing one unit of energy E given task prices $p(\omega)$

$$\min_{\tau(\omega)} \left\{ \int_0^1 p(\omega)\tau(\omega)d\omega \right\}$$
$$s.t.\ E = \inf \left\{ \tau(\omega) : \omega \in [0, 1] \right\}$$

Which implies that demand for each task is the same and equals total energy demand. Then, the cost of producing one unit of realized energy p_e is given by aggregating across all tasks and is equivalent to the price index of tasks:

$$\begin{aligned} \int_0^1 p(\omega)\tau(\omega)d\omega &= E \int_0^1 p(\omega)d\omega \\ &= Ep_e \end{aligned} \tag{39}$$

Assignment of fuels to energy tasks:

Given the set of fuels available to the plant, $\mathcal{F}$, and fuel prices, a plant finds the fuel that minimize the cost of performing task ω:

$$\mathcal{C}(\omega) = \min_{e_1(\omega),\ldots,e_F(\omega)} \sum_f p_f e_f(\omega)$$
$$s.t.\ \sum_f \psi_f(\omega)e_f(\omega) = \tau(\omega)$$

The linearity of the constraint implied by perfect substitution across fuels is such that the plant chooses the fuel that has the lowest unit cost to produce the task. Hence, the task price and fuel choices follow a discrete choice:

$$p(\omega) = \min_{f \in \mathcal{F}} \left\{ \frac{p_f}{\psi_f(\omega)} \right\} \tag{40}$$

Aggregation from tasks to fuels

From the problem before, I can define the set of tasks that are performed by each fuel:

$$\mathcal{T}_f = \left\{ \omega : \frac{p_f}{\psi_f(\omega)} \leq \frac{p_j}{\psi_j(\omega)} \ \forall j \neq f \in \mathcal{F} \right\} \tag{41}$$

From the optimal assignment of fuels to energy tasks, I also get fuel demand for each task:

$$e_f(\omega) = \begin{cases} E\psi_f(\omega)^{-1} & \text{if } \omega \in \mathcal{T}_f \\ 0 & \text{otherwise} \end{cases} \tag{42}$$

I can then aggregate fuel demand across all tasks to get fuel demand at the plant-level (conditional on some level of realized energy E):

$$e_f = E \underbrace{\int_{\mathcal{T}_f} \psi_f(\omega)^{-1} d\omega}_{\Gamma_f^{-1}} \tag{43}$$

Rearranging terms, I can defined realized energy E as the product of physical fuel quantities e_f times a terms that converts fuel quantities into realized energy Γ_f.

$$E = e_f \Gamma_f \qquad \forall f \in \mathcal{F}$$

The Γ_f term is an important novelty of this model, and contains information about both the share of tasks performed by fuel f, and the average productivity of fuel f. One one hand, Γ_f could be large if there are many tasks are allocated fuel f which would happen if fuel f is relatively cheap (*price/task channel*). On the other hand, Γ_f could be large if the productivity for each task is high (*productivity channel*). An important empirical challenge will be to separate these two channels to separately identify the share of tasks performed by a fuel from the average productivity of that fuel.[13] I can now rewrite the price index of realized energy as a weighted sum of fuel prices:

$$\begin{aligned} p_e &= \frac{1}{E}\int_0^1 p(\omega)\tau(\omega)d\omega \\ &= \frac{1}{E}\sum_f \int_{\mathcal{T}_f} \frac{p_f}{\psi_f(\omega)}E \\ &= \sum_f p_f \Gamma_f \end{aligned} \tag{44}$$

Both the price of realized energy and the quantity of realized energy are a function of fuel prices and unobserved fuel efficiency terns, hence are by definition unobserved. However, I observe energy spending which equals fuel spending:

$$\begin{aligned} p_e E &= \sum_f \int_{\mathcal{T}_f} \frac{p_f}{\psi_f(\omega)} \frac{E}{\mathcal{M}} \\ &= \sum_f p_f e_f \end{aligned} \tag{45}$$

This identity is very important and will play an important role in identifying the production function, from which I will identify the weighted share of tasks performed by fuel f, Γ_f and later on to identify the underlying distribution of fuel efficiency. This production model in energy inputs is fairly flexible because it allows for very large variation in relative fuel quantity shares, an important feature of plant-level fuel consumption.

Proposition 5. *Ceteris-Paribus, increasing the number of fuels available $\mathcal{F}$ weakly decreases the*

[13] note that both channels interact with each others. Ceterus-paribus, higher task productivity also implies a higher share of tasks performed by a fuel.

price of energy $p_e(\mathcal{F})$.

$$| \mathcal{F}' | > | \mathcal{F} | \rightarrow p_e(\mathcal{F}') \leq p_e(\mathcal{F})$$

Proposition 1 highlights the option value that an additional fuel provide. Indeed, when a fuel is added, firms have more productivity draws to choose from for each tasks. Since fuels are perfect substitutes within tasks and tasks are perfect complements, this means that the overall productivity of energy sources will increase, leading to a lower marginal cost of realized energy

Identification

My approach to identifying fuel productivity is novel and exploit the task-based nature of production. I show how to simultaneously recover the production function and the normalized quantity of realized energy $\frac{E_{it}}{\overline{E}}$. I can use this result to recover the weighted share of tasks performed by each fuel (also the cost-minimizing quantity of potential energy from fuel f required to produce one unit of realized energy), up to the normalization of Grieco et al. (2016):

$$\widehat{\overline{E}\Gamma_{fit}} = e_{fit}\Big(\frac{\overline{E}}{E_{it}}\Big) \tag{46}$$

I need to separate the unweighted share of tasks performed by each fuel ($\mathcal{T}_{fit}$) from the productivity of each fuels at performing each tasks Ψ_{it}. To do so, I rely on two assumptions which allow me to aggregate fuels across tasks and exploit observed fuel price variation in order to separate the share of tasks performed by each fuel from the efficiency of each fuel. The first assumption is standard in the task-based production function literature (Acemoglu and Restrepo, 2021). The second assumption is standard in the literature on technological choice (Boehm and Oberfield, 2020; Oberfield, 2018; Kortum, 1997).

Assumption 1-Symmetric tasks. *Energy tasks are all equivalent and for a given fuel, plants draw from the same productivity distribution across tasks.*

Under this assumption coupled with a continuum of energy tasks in the $\omega \in [0, 1]$ interval, the (unweighted) share of tasks performed by each fuel $\mathcal{T}_{fit}$ can also be interpreted as the probability that fuel f is is preferable over all other fuels, where the probability is taken over the distribution of fuel efficiency:

$$\mathcal{T}_{it} = Pr\Big(\frac{p_{fit}}{\psi_{fit}} \leq \frac{p_{jit}}{\psi_{jit}} \forall j \neq f \in \mathcal{F}_{it}\Big) \tag{47}$$

Then, Γ_f is the joint distribution of the inverse productivity of a fuel when that fuel is chosen. By Bayes's rule, it is also the distribution of inverse fuel productivity conditional on fuel f being chosen times the probability that fuel f is chosen. Since the fuel f needs to outperform all other fuels to be chosen, the distribution of observed fuel efficiency is a truncated version of the underlying true fuel productivity. In this sense, realized fuel efficiency is an endogenous outcome of the choices that firms make.

$$\begin{aligned}
\Gamma_{fit} &= \int_{\mathcal{T}_{it}} \psi_{fit}(\omega)^{-1} d\omega \\
&= \mathbb{E}_\omega \left(\psi_{fit}^{-1}(\omega), \omega \in \mathcal{T}_{fit}\right) \\
&= \quad \underbrace{\mathbb{E}_\omega(\psi_{fit}^{-1}(\omega) \mid \omega \in \mathcal{T}_{fit})}_{\text{inverse fuel efficiency when f is chosen}} \quad \times \quad \underbrace{\mathcal{T}_{fit}}_{\text{Probability f is chosen}}
\end{aligned}$$

For a given plant, Γ_{fit} integrates out the inverse of productivity for fuel f over the probability that each draw makes fuel f chosen over all other alternative fuels:

$$\Gamma_f = \int \psi_f^{-1} \underbrace{\Big[\prod_{j \neq f} \int \mathbb{I}\Big(\psi_f \geq p_f \max\{\psi_j / p_j\}\Big) f(\psi_j) d\psi_j\Big]}_{\text{Pr(efficiency draw for f is chosen over all other fuels)}} f(\psi_f) d\psi_f \tag{48}$$

However, I am interested in recovering the underlying exogenous distribution of fuel efficiency which doesn't vary with fuel prices. Otherwise, I cannot separate fuel price variation (needed for counterfactual tax experiments) from fuel efficiency. To do so, I make the following assumption:

Assumption 2-Pareto Distribution. *I assume that the distribution of fuel productivity/efficiency across tasks follows a Pareto distribution with plant and year-specific scale and common shape.*

$$\psi_{fit} \sim Pareto(\overline{\psi}_{fit}, \theta)$$

From now, the scale of the fuel efficiency distribution will be referred as fuel efficiency. Under assumption 2, Γ_{fit} has a closed-form solution. If the plan has access to two fuels, e.g. gas (g) and coal (c), then

$$\Gamma_{g,it} = \underbrace{\frac{\theta}{\theta+1}\Omega_{git}^{-\theta-1}\overline{\psi}_{git}^{\theta}}_{\text{Direct task displacement}} - \underbrace{\left(\frac{p_{cit}}{p_{git}}\right)^{-\theta}\frac{\theta}{2\theta+1}\Omega_{git}^{-2\theta-1}(\overline{\psi}_{git}\overline{\psi}_{cit})^{\theta}}_{\text{Indirect task displacement through fuel c}} \tag{49}$$

Where $\Omega_{git} = \max\left\{\frac{p_{git}}{p_{cit}}\overline{\psi}_{cit}, \overline{\psi}_{gict}\right\}$. Note that there is an analogous expression for $\Gamma_{c,it}$. For a given shape parameter (θ), this is a system of two equations ($\Gamma_{g,it}, \Gamma_{c,it}$) and two unknowns ($\overline{\psi}_{git}, \overline{\psi}_{cit}$), which can be solved easily to recover the scale of the exogenous fuel efficiency distribution that each plant has for each fuel it is using. In the appendix, I show extension of equation (23) to the case with more than 2 fuels. I also show that in the 2 fuels case, there is a unique solution ($\overline{\psi}_{git}, \overline{\psi}_{cit}$) that solves the system of equation in (23). This means that there for a given set of prices, there is a unique optimal allocation of fuels to tasks. In the case of more than 2 fuels, Monte-Carlo simulations also suggest uniqueness.

When there are more than 2 fuels, the number of interaction terms increases exponentially with the number of fuels. For example, if one adds oil (o) to gas and coal, then there will be two second-order task displacement terms (the interaction of oil with gas, oil with coal and coal with gas), and one third-order task displacement term (e.g. the task displacement of tasks performed by gas induced by the price of oil caused by changes in the price of oil). This proves to be a fairly general micro-foundation for the production of realized energy under different fuel sets. I only observe Γ_{fit} up to an industry-specific normalization (the geometric mean of realized energy, which is unobserved), $\overline{E}\Gamma_{fit}$. While I can use (22) to recover the scale of fuel efficiency for each plant, I cannot compare fuel efficiency across plants in different industries.

Lastly, I normalize the shape of the Pareto distribution θ to 1. This is because for different shape parameters, I can always recover different scale parameters that will exactly solve the system of equations in (23). Since the weighted share of tasks captures all information about the substitution of fuels to task, any moment related to fuel consumption/expenditure shares will not recover the common shape θ separately from the individual-specific scale $\overline{\psi}_{fit}$. Intuitively, this is because the same fuel substitution patterns can be achieved with a high Pareto tail (low θ) and low scale parameters, or with a low Pareto tail (high θ) and large scale parameters.

C.4.1 Non-parametric

Given a fuel set $\mathcal{F}$, firms produce realized energy according to the following unspecified production function:

$$\begin{aligned} E_{it} &= g(\psi_{1it}e_{1it}, \psi_{2it}e_{2it}, ..., \psi_{fit}e_{fit}) \\ &= g(\mathbf{\Psi}_{it}\mathbf{e}_{it}) \end{aligned}$$

$\forall f \in \mathcal{F}$, where ψ_{fit} is the productivity of fuel f. Taking input prices $\{p_{fit}\}_{f\in\mathcal{F}}$ and the set of fuels $\mathcal{F}$ as given, fuel quantity choices are static. Given a unit of realized energy E_{it}, the cost-minimization problem of the firm is as follows:

$$\min_{\{e_{fit}\}_{f\in\mathcal{F}_{it}},\lambda} \sum_{f\in\mathcal{F}} p_{fit}e_{fit} + \lambda\Big(E_{it} - g(\mathbf{\Psi}_{it}\mathbf{e}_{it})\Big) \tag{50}$$

In this approach, I do not seek to recover the production function g directly. Rather, I seek to recover a structural equation for the endogenous price of realized energy $p_{eit}(\mathbf{p_{it}}, \mathbf{\Psi}_{it})$, which is what the task-based model allows me to do in the main text. Indeed, knowing the endogenous price of realized energy is sufficient to perform all counterfactual. Given standard assumptions on g, the solution to (34) gives a cost function, which can be mapped to total spending on energy. Then, I know that the endogenous price of realized energy is the unit cost function:

$$\begin{aligned} p_{eit}E_{it} &= \mathcal{C}(E_{it}, \mathbf{p}_{it}, \mathbf{\Psi}_{it}) \\ p_{eit} &= \mathcal{C}(1, \mathbf{p}_{it}, \mathbf{\Psi}_{it}) \end{aligned}$$

To identify the unit cost of energy, I exploit firms' optimality conditions. Using Sheppard's Lemma, I can characterize optimal fuel choices as the derivative of the unit cost with respect to fuel prices:

$$e_{fit} = \frac{\partial \mathcal{C}(E_{it}, \mathbf{p}_{it}, \mathbf{\Psi}_{it})}{\partial p_{fit}} \; \forall f \in \mathcal{F}_{it}$$
$$= \mathcal{H}_f(E_{it}, \mathbf{p}_{it}, \mathbf{\Psi}_{it})$$

Which gives me a system of structural equations:

$$\begin{aligned}
e_{1it} &= \mathcal{H}_1(E_{it}, p_{1it}, p_{2it}, ..., \psi_{1it}, \psi_{2it}, ..., \psi_{fit}) \\
e_{2it} &= \mathcal{H}_2(E_{it}, p_{1it}, p_{2it}, ..., \psi_{1it}, \psi_{2it}, ..., \psi_{fit}) \\
&... \\
e_{fit} &= \mathcal{H}_f(E_{it}, p_{1it}, p_{2it}, ..., \psi_{1it}, \psi_{2it}, ..., \psi_{fit})
\end{aligned}$$

Matzkin (???) shows that identification of both the structural equations $\{\mathcal{H}_f\}_{f\in\mathcal{F}}$ and unobserved terms $\{\psi_{fit}\}_{f\in\mathcal{F}}$ is possible under certain conditions.[14] First, fuels must be ordered such that there is monotonicity in ψ_{fit} and ψ_{-fit}. Second, unobserved productivity terms must be separable from the observed prices. Unfortunately, this is not usually the case, even in log. For examples, only the ratio of (log) fuel quantities admits separability with a CES production function. With the task-based model used in the main text, there is no separability at all. In this context, if may be difficult to rationalize what kind of production functions this approach admits. Relaxing this assumption is an interesting question, but it is realistically beyond the scope of this paper. Lastly, these structural equations must be integrated to recover the unit cost.

[14] I omit some technical assumptions which are standard in non-parametric identification of systems of structural equations.