Online Appendix for "Balancing Economic Activity and Carbon Emissions: A Study of Fuel Substitution in India"

Emmanuel Murray Leclair*

Department of Economics, Western University

September 21, 2023

Contents

*email: emurrayl@uwo.ca

A Data

A.1 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), which I call this *potential energy*. This is because it captures what energy may be extracted from combustion of a particular fuel.

However, what plants 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 plants' knowledge on wasting energy. In essence, realized energy is what plants 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 blast furnaces (coal) at using heating potentials of the underlying fuels (**?**).
2. **Within fuel types:** In cement manufacturing, coal used in rotary kilns is more productive than in vertical shaft kilns for the production of clinker as part of cement manufacturing (**?**).
3. **Wasted resources:** Energy retrofit programs underlie large heterogeneity differences on how efficiently agents in the economy use the heating potential of fuels (**?**). Examples include keeping lights opened unnecessarily or forgetting to turn off machinery.
4. **Fuel Quality:** For example, different grades of coal have different heating values, implying that the same physical quantity of coal observed in the data will be more or less productive depending on its grade. Anthracite coal is the most productive and also the most expensive grade of coal, whereas lignite coal has the lowers energy content and is the cheapest coal (https://www.eia.gov/energyexplained/coal/). I show in the figure below that coal prices are positively correlated with coal productivity, indicating that unobserved coal grades are being captured by estimated coal productivity.

A.2 Fuel Prices and Transportation Costs

Identification of plants' 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 1 shows the evolution in the median fuel prices paid 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 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 trans-

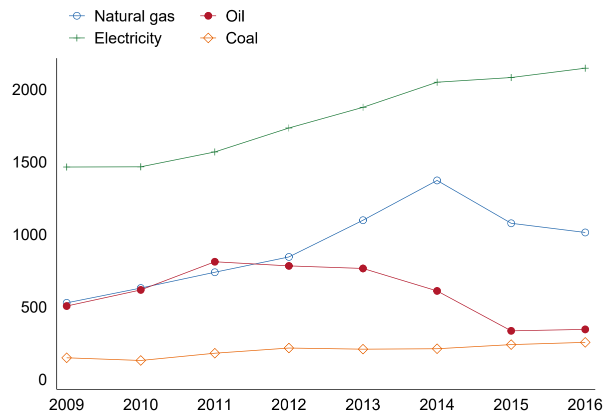


Figure 1: Yearly Median Fuel Prices (INR/mmBtu)

portation 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[1]. 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:

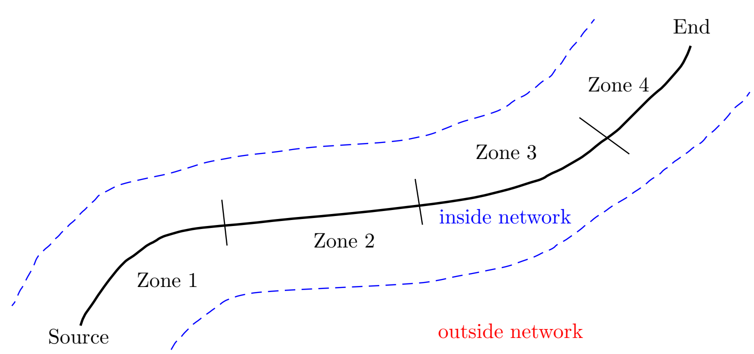


Figure 2: Hypothetical Structure of Transportation Costs for Natural gas Pipeline

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 is a solid and because it is mostly extracted domestically[2]. As such, 17% of all coal is transported directly from the mine to plants through conveyor belts, 33% is transported by road, and 50%

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

[2]**?** shows that Coal India Limited (CIL) is the largest coal mining company in the world

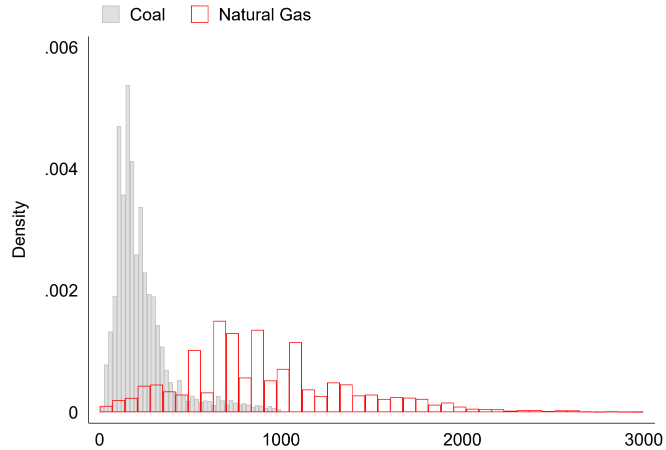


Figure 3: Histogram comparing price of natural gas and coal (INR/mmBtu)

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 3 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.

Table 1: 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.

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

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

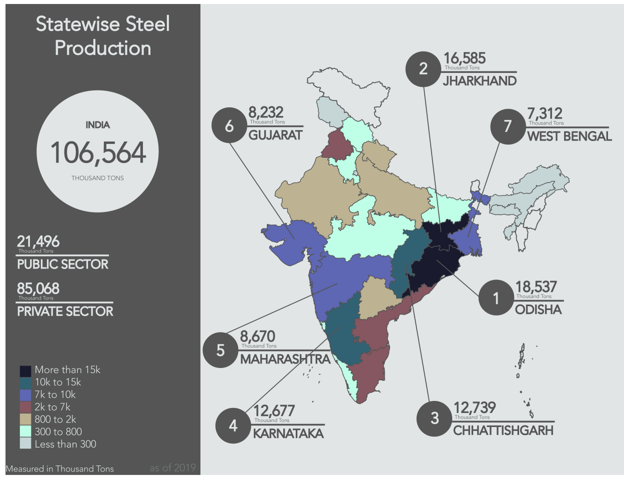


Figure 4: Concentration of steel plants by regions

Source: https://en.wikipedia.org/wiki/File:State_wise_Steel_Production_India,_2019.jpg

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.

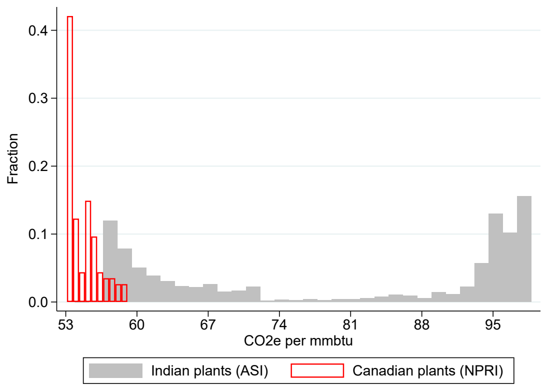
Location of Steel Plants and "Steel Belt"

A.3 Further Evidence

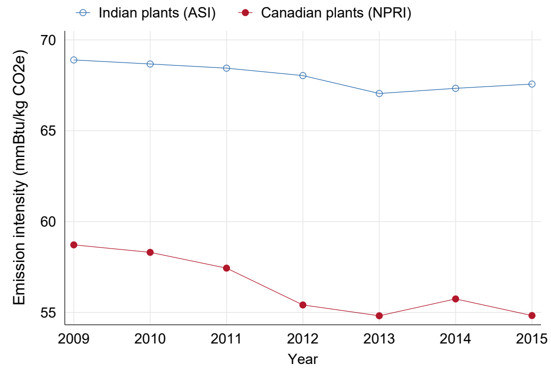
A.3.1 Evidence of Large Levels of Pollution in Indian Manufacturing

Indian manufacturing establishments exhibit a higher level of pollution intensity compared to their counterparts in developed economies. As demonstrated in Figure 5a, 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 5b).

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 5a 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 (**?**). In contrast, a large portion of Indian plants primarily use coal, which pollutes twice as much as gas. In Figure 6, 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 5a.



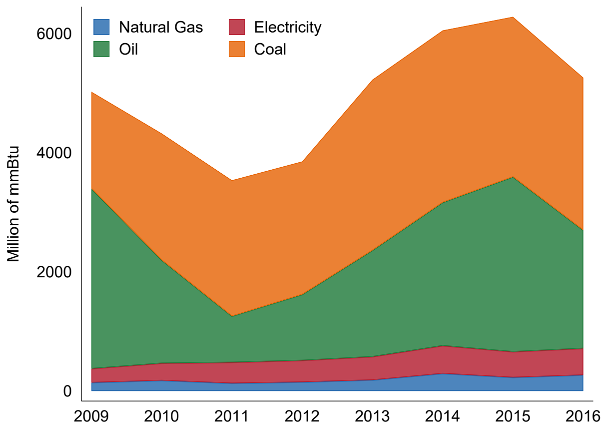
(a) Pollution Intensity - Cement Manufaturing



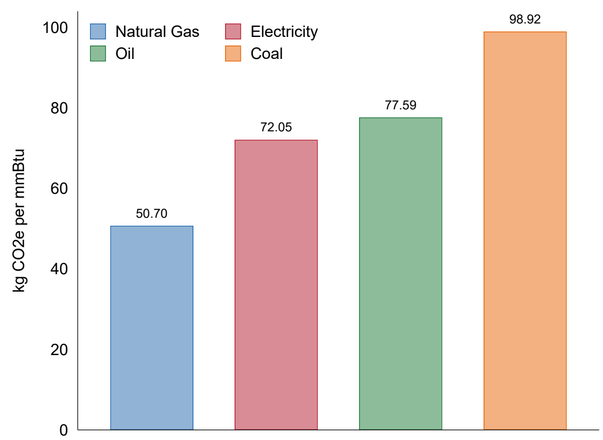
(b) Evolution of Pollution Intensity - Heavy Manufacturing Industries

Figure 5: 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) (?). 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 5b, I compare the within industry average pollution intensity for 5 heavy manufacturing industries: Pulp & paper, cement, steel, aluminium, and glass.



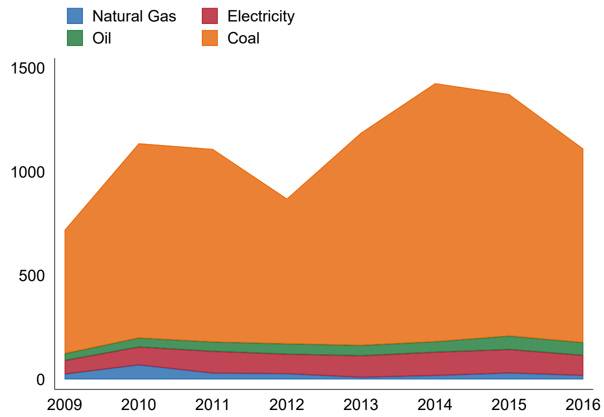
(a) Total Emissions by Fuel



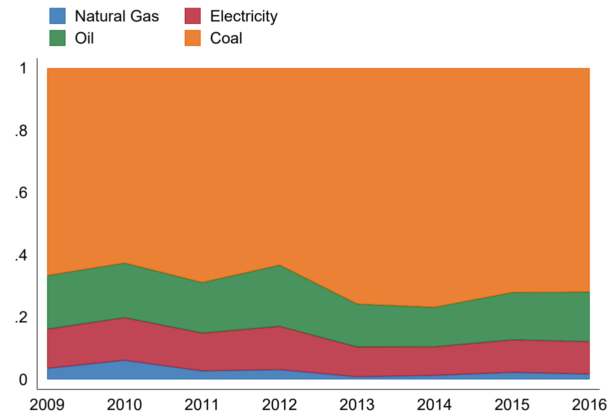
(b) Average Emission Intensity by Fuel

Figure 6: Comparison of Fuels used by ASI establishments

Note: Figure 6a aggregates across all manufacturing establishments in the ASI by year, and suggests a much lower usage of natural gas compared to coal. Figure 6b shows the average emission intensity of each fuel, where the average is taken across industries according to scientific calculations made by ?.



(a) Total Quantity by Fuel (mmBtu)



(b) Percent of Total Quantity by Fuel

Figure 7: Comparison of Fuels used by Steel Establishment

A.3.2 Evidence on Switching and Mixing

Indian Plants

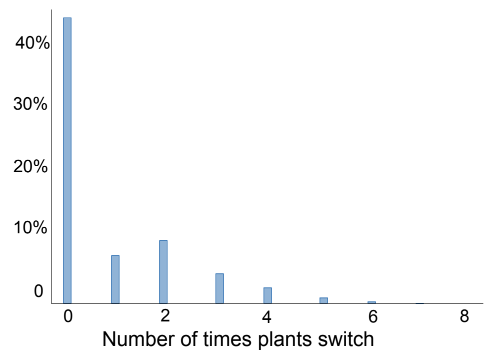


Figure 8: Number of Times Unique Plants add or drop a Fuel (ASI)

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.

Table 2: 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 3: 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)

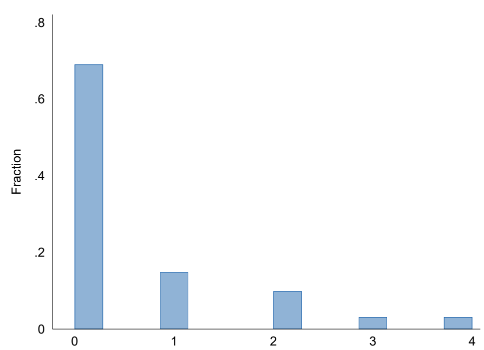
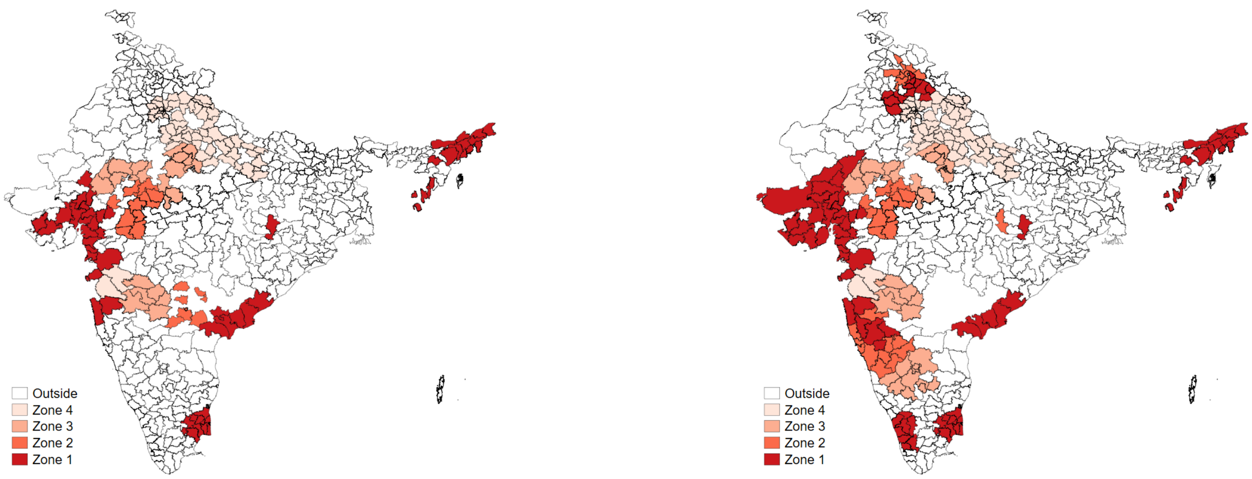


Figure 9: Number of Times Unique Plants add or drop a Fuel (U.S. GHGRP - Pulp & Paper Manufacturing)

A.3.3 Relationship Between Natural Gas Pipeline Expansion and Natural Gas Adoption

Plants located near transmission pipelines have access to the main distribution network. This access reduces costs compared to those located far from the pipeline network, who need to either have access to or construct expensive gasification terminals to convert liquified natural gas (LNG) to its usable form. In the Indian context, I investigate 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 use 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 pipeline network expanded in that district between t and $t+1$.[3] 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 **?** 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.



(a) 2009 (b) 2016

Figure 10: Indian districts by zone of access to natural gas transmission pipelines.

Notes: Zones are defined according to regulations under the Petroleum and Natural Gas Regulatory Board (PNGRB) and are defined by each pipeline segment of 250km from the source. Zone 1 is closest to source and Zone 4 is furthest from source. Moreover, unlike the U.S. and Canada, transportation tarrifs do not depend on long term contracts between the pipeline and suppliers. Instead, tarrifs are fully regulated and depend on the fixed and variable cost of each pipeline, and vary by zone.

[3]While pipeline expansions are not exogenous, I show in Appendix A.3.5 that the vast majority of aggregate demand for natural gas comes from fertilizers, power generation and oil refineries. In this context, the expansion of the pipeline network can be seen as a plausibly exogenous shock for the manufacturing sector.

	Added Natural Gas		
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.

Notes: The coefficient for "Pipeline Expanded" is the average marginal effect of increasing the pipeline network in a plant's district on the probability that the plant adds natural gas. Probabilities come from the following model, estimated with logit errors: $\Delta Dgas_{it} = \beta_0 + \beta_1 \Delta Dpipeline_{it} + controls_{it} + \epsilon_{it}$, where $\Delta Dgas_{it}$ is an indicator for whether plant i added natural gas in year t and $\Delta Dpipeline_{it}$ is an indicator for whether the natural gas pipeline network expanded to reach plant i's district in year t. Individual marginal effects are calculated as $Pr(\Delta Dgas_{it} = 1 \mid \Delta Dpipeine_{it} = 1, controls_{it}) - Pr(\Delta Dgas_{it} = 1 \mid \Delta Dpipeine_{it} = 0, controls_{it})$.

A.3.4 Relationship between Number of Fuels and Plant Age

As plants become older, he number of fuels in their set rises on average, which is similar to the pattern found with output per worker in the main text. Moreover, the magnitude of the relationship is larger with age.

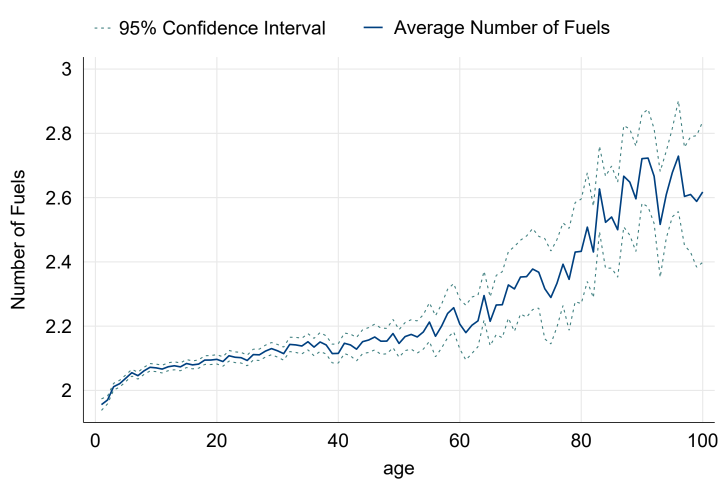


Figure 11: Number of fuels by plant age, average of all ASI plants

A.3.5 Natural Gas Demand by Sector

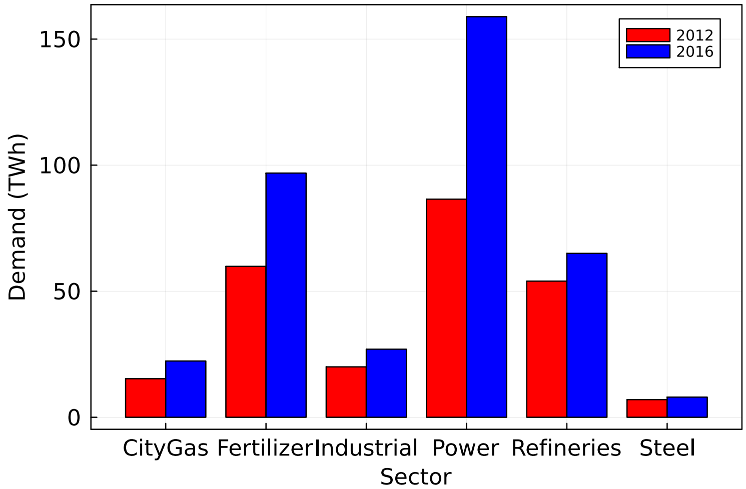
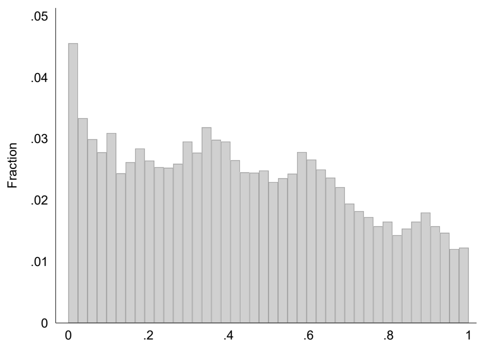


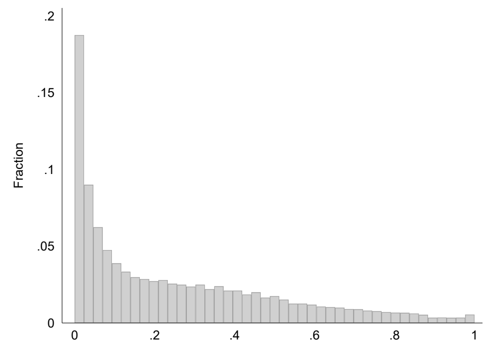
Figure 12: Projected Natural Gas Demand by Sector - all of India (2012 and 2016)

Notes: Data retrieved from the Petroleum and Natural Gas Regulatory Board's Data Bank (?)

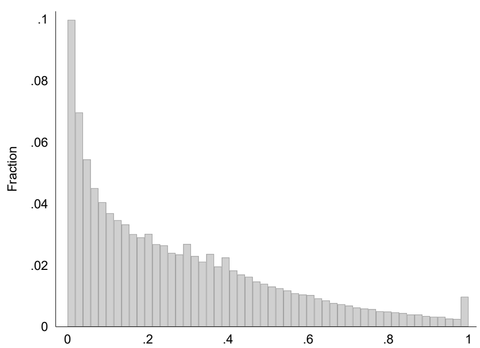
A.3.6 Fuel Expenditure Shares - ASI



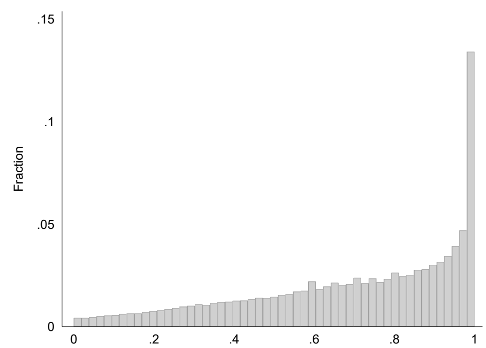
(a) Coal



(b) Gas



(c) Oil



(d) Electricity

Figure 13: Distribution of Expenditure Shares

A.3.7 Number of Steel Varieties

Product Category	Percentage of Occurence
Ferrous products from direct reduction of iron ore	5.5
Mild steel billets, blooms	4.5
Mild steel bright bar, rectangular cross section	4.3
Bars and rods, hot-rolled, in irregularly wound coils, of iron or non-alloy steel	3.8
Sponge Iron	3.6
Ingots alloy steel	3.5
Number of Unique Varieties	404

Table 5: Top 6 Output Varieties from Steel Plants in the ASI (NPCMS)

Notes: Unique varieties are taken from the primary product made by each plant.

B Model

B.1 Comparative Statics – Option Value of an Additional Fuel

Below are some propositions developing on the option value that an additional fuel provides. The price that plants pay for its energy bundle is a CES price index in all fuels available $\mathcal{F}_{it}$:

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

? and others show that this CES price index is decreasing in the number of input varieties it contains, here $|\mathcal{F}_{it}|$, as long as inputs are gross substitutes ($\lambda > 1$). This means that absent of fixed costs, all plants would always include all fuels in their set. The intuition underlying this option value 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, towards the more productive units of the new fuel due to gross substitution ($\lambda > 1$), which in terms increases the marginal product of all 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 $p_{\tilde{E}_{it}}$. The following three propositions formalize these ideas:

Proposition 1. *Gains from variety: ceteris-paribus, if a fuel set $\mathcal{F}$ is a strict subset of $\mathcal{F}'$ and fuels are gross substitute ($\lambda > 1$), 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, such that $\Big(\sum_{f\in\mathcal{F}} (\frac{\tilde{p}_{fit}}{\psi_{fit}})^{1-\lambda}\Big)^{\frac{1}{1-\lambda}} < \Big(\sum_{f\in\mathcal{F}'} (\frac{\tilde{p}_{fit}}{\psi_{fit}})^{1-\lambda}\Big)^{\frac{1}{1-\lambda}}$. By convexity of the function $f(x) = x^{\frac{1}{1-\lambda}}$ when $\lambda > 1$, we know that $\forall x, y \in dom(f), f(y) \geq f(x) + f'(x)(y-x)$ using a first-order Taylor expansion of f. Let $y = \sum_{f\in\mathcal{F}} (\frac{\tilde{p}_{fit}}{\psi_{fit}})^{1-\lambda}$ and $x = \sum_{f\in\mathcal{F}'} (\frac{\tilde{p}_{fit}}{\psi_{fit}})^{1-\lambda}$ Then,

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

$$\geq \Big(\sum_{f\in\mathcal{F}'}(\frac{\tilde{p}_{fit}}{\psi_{fit}})^{1-\lambda}\Big)^{\frac{1}{1-\lambda}} + \frac{1}{1-\lambda}\Big(\sum_{f\in\mathcal{F}'}(\frac{\tilde{p}_{fit}}{\psi_{fit}})^{1-\lambda}\Big)^{\frac{\lambda}{1-\lambda}}\Big(-\sum_{f\in\mathcal{F}'\setminus\mathcal{F}}(\frac{p_{fit}}{\psi_{fit}})\Big)$$

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

□

In addition, by expanding its fuel set, a plant 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 based on the size of a 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 relevant in a world where many fossil fuels, in particular oil and natural gas, are susceptible to geopolitical shocks that have persistent effects on fuel prices. The final proposition demonstrates that a larger fuel set allows plants to hedge against binding quantity shortages of a particular fuel more effectively. This proposition is particularly applicable in the Indian context, where the economy frequently experiences disruptions in its electricity supply (**???**).

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 perceived marginal costs will be larger under $\mathcal{F}$.* $\mathcal{F}\subset\mathcal{F}' \rightarrow p_{\tilde{E}_{it}(\mathcal{F},\overline{e}_f)} > p_{\tilde{E}_{it}(\mathcal{F}',\overline{e}_f)}$

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

The Lagrangian can be written as:

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

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 plant was facing $p_{fit} + \mu_2$ as the true price. Hence, the binding quantity shortage is analogous to a price increase.

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

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

C Identification

C.1 Remark on the Identification of the Energy Production Function

I haven't used the first-order condition (in level) for electricity in the energy cost-minimization problem. This is not an issue because plants choose the level of 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 plants 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 electricity, whereas the quantity of electricity will be pinned down by the plant's choice of energy. The first order condition for energy in the cost-minimization problem is as follows, where I sub in equation (**??**) for all relative fuel-augmenting productivity:

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

Once I take into account all first-order conditions, plants' optimality condition implies that the shadow cost of electricity (Lagrange multiplier μ_{it}) is the marginal cost of realized energy. Plugging the equilibrium condition for the shadow cost of electricity into equation (1) implies that the first order condition for electricity is always satisfied:

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

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

I create a sample of plants with all state variables present in the main model. External estimation of parameters governing the distribution of random effect from the sample of plants who use gas and/coal leads to upward biased estimates. Indeed, plants 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 6: 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 7: 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 plants are mostly selecting on the basis of gas, and almost all plants either use oil and electricity or oil,electricity, gas and coal.

D Results & Counterfactuals

D.1 Model Fit

Construction of fuel set choice probabilities and transition

$$\mathbb{P}_{\mathcal{F}'}(model) = \frac{1}{NT}\sum_i\sum_t\sum_k \rho(\mu_{fk} \mid \mathcal{F}_i, s_i; \hat{\theta}_1, \hat{\theta}_2, \hat{\pi}) Pr(\mathcal{F}' \mid \mathcal{F}_{it}, s_{it}, \mu_{fi} = \mu_{fk}; \hat{\theta}_1, \hat{\theta}_2)$$

$$\mathbb{P}_{\mathcal{F}'|\mathcal{F}}(model) = \frac{1}{N_{\mathcal{F}}T}\sum_i\sum_t\sum_k \underbrace{\rho(\mu_{fk} \mid \mathcal{F}_i, s_i; \hat{\theta}_1, \hat{\theta}_2, \hat{\pi})}_{\text{Conditional probability of comparative advantage}} \underbrace{Pr(\mathcal{F}' \mid \mathcal{F}, s_{it}, \mu_{fi} = \mu_{fk}; \hat{\theta}_1, \hat{\theta}_2)}_{\text{Conditional choice probability}}$$

Details explaning role of productivity in

D.2 Decomposition of Relationship between Estimated Price of Energy and Number of Fuels

I look at the relationship between the price of energy and the number of fuels available to plants. The findings reveal a significant and consistently negative relationship. It indicates that plants may be selecting larger fuel sets based on their productivity in utilizing those fuels, or that there might be a considerable option value associated with having a greater variety of fuels within a set. These factors, which are encompassed within the energy production model, could reasonably explain the observed negative relationship.

Table 8: Relationship between $\ln p_{\hat{E}_{it}}$ and the number of fuels available to plants.

	(1) $\ln p_{\hat{E}_{it}}$	(2) $\ln p_{\hat{E}_{it}}$	(3) $\ln p_{\hat{E}_{it}}$	(4) $\ln p_{\hat{E}_{it}}$
Three Fuels	-0.684***	-0.710***	-0.714***	-0.687***
	(0.0425)	(0.0420)	(0.0418)	(0.0319)
Four fuels	-0.794***	-0.803***	-0.828***	-0.571***
	(0.0698)	(0.0690)	(0.0686)	(0.0526)
Year Dummies		Yes	Yes	Yes
Controlling for fuel prices			Yes	Yes
Controlling for TFP				Yes
N	7,603	7,603	7,565	7,565

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. Since 90% of plants always use oil and electricity, and since the remainder of the analysis focus on these plants, I only kept plants who always both oil and electricity in this regression. This means that the benchmark number of fuels in these regressions is two rather than one.

Moreover, the evidence provided in Table 8 may also have policy implications. Indeed, a policy whose aim is to incentivize adoption of new fuels, such as a fixed cost subsidy, may be more effective if adding a fuel causes a decrease in energy marginal costs through the additional option value of the new fuel. On the other hand, if differences in energy marginal costs are explained by selection channels such as prices and productivity, then such a policy may not be as effective at incentivizing adoption of new fuels. I now show that selection channels, particularly fuel productivity, dominate the option value. I decompose the observed differences in the price of energy $p_{\tilde{E}_{it}}$ across fuel sets between three main factors: option value, fuel productivity and fuel prices.

Below I show results of the decomposition. I find that fuel productivity explains the majority of the average difference in energy marginal costs across fuel sets. This result acts as a cautionary tale against the effectiveness of fixed cost subsidies at incentivizing the adoption of new fuels, and particularly natural gas. Below are the details of the decomposition

		OCE	OGE	OGCE
Total Difference	Percent (%) Difference with OE	-65.65	-71.54	-86.97
Option Value		36.14	5.42	6.3
Fuel Productivity	Percent (%) of Total	62.6	97.75	94.84
Fuel Prices		1.25	-3.18	-1.14

Table 9: Shapley Decomposition of the Difference in Average Marginal Cost of Energy Between Fuel Sets

Notes: I compare the observed differences in the average (across plants) marginal cost of realized energy between plants who use coal and/or gas on top of oil and electricity (OCE,OGE,OGCE) relative to plants who only use oil and electricity (OE).

D.3 The Shapley-Owen-Shorrocks Decomposition

General Definition

Given an arbitrary function $Y = f(X_1, X_2, ..., X_n)$, the Shapley-Owen-Shorrocks decomposition is a method to decompose the value of $f(\cdot)$ into each of its arguments $X_1, X_2, ..., X_n$. Intuitively, the contribution of each argument if it were to be *"removed"* from the function. However, because the function can be nonlinear the order in which the arguments are removed matters in general for the decomposition. The function f can be the outcome of a regression, like the predicted values or sum of square residuals, or the output of a structural model, such as a counterfactual value for a variable given a list of model parameters or components, or a transformation of the sample, for example the Gini coefficient.

The Shapley-Owen-Shorrocks decomposition is the unique decomposition satisfying two important properties. First, the decomposition is exact decomposition under addition, letting C_j denote the contribution of argument X_j to the value of the function $f(\cdot)$,

$$\sum_{j=1}^{n} C_j = f(X_1, X_2, ..., X_n), \tag{2}$$

so that $C_j f(\cdot)$ can be interpreted as the proportion of $f(.)$ that can be attributed to X_j.[4] Second, the decomposition is symmetric with respect to the order of the arguments. That is, the order in which the variable X_j is removed from $f(\cdot)$ does not alter the value of C_j.

The decomposition that satisfies both those properties is

$$C_j = \sum_{k=0}^{n-1} \frac{(n-k-1)!k!}{n!} \left(\sum_{s \subseteq S_k \setminus \{X_j\} : |s|=k} [f(s \cup X_j) - f(s)] \right), \tag{3}$$

where n is the total number of arguments in the original function f, $S_k \setminus \{X_j\}$ is the set of all

[4]The interpretation holds as long as f is non-negative. If f can take negative values, then the interpretation of C_j under the exact additive rule can be misleading as some arguments can have $C_j < 0$.

"sub-models" that contain k arguments and exclude argument X_j.[5] For example,

$$
\begin{aligned}
S_{n-1} \setminus X_n &= f(X_1, X_2, ..., X_{n-1}) \\
S_1 \setminus X_n &= \{f(X_1), f(X_2), ..., f(X_{n-1})\}.
\end{aligned}
$$

The decomposition in (3) accounts for all possible permutations of the decomposition order. Thus, $\frac{(n-k-1)!k!}{n!}$ can be interpreted as the probability that one of the particular sub-model with k variables is randomly selected when all model sizes are all equally likely. For example, if $n = 3$, there are sub-models of size $\{0, 1, 2\}$. In particular, there are 2^2 permutation of models that exclude each variable: $\{\underbrace{(0,0)}_{k=0}, \underbrace{(1,0),(0,1)}_{k=1}, \underbrace{(1,1)}_{k=2}\}$.

$$
\begin{aligned}
k = 0 : \frac{(n-k-1)!k!}{n!} &= \frac{(3-0-1)!0!}{3!} = \frac{1}{3} \\
k = 1 : \frac{(n-k-1)!k!}{n!} &= \frac{(3-1-1)!1!}{3!} = \frac{1}{6} \\
k = 2 : \frac{(n-k-1)!k!}{n!} &= \frac{(3-2-1)!2!}{3!} = \frac{1}{3}
\end{aligned}
$$

D.3.1 Application to Decomposition of the Marginal Cost of Energy

The function of interest in this paper is the difference in sample average between the marginal cost of realized energy for a given fuel relative to oil and electricity only. The function takes three broad set of arguments: fuel prices $\mathbf{p}$, fuel productivity $\mathbf{\Psi}$ and a fuel set $\mathcal{F}$:

$$
f(\mathbf{p}, \mathbf{\Psi}, \mathcal{F}) = \frac{1}{N_{\mathcal{F}}} \sum_{i:\mathcal{F}} \Big(\sum_{f \in \mathcal{F}} \big(\tilde{p}_{fit}/\psi_{fit}\big)^{1-\lambda} \Big)^{\frac{1}{1-\lambda}} - \frac{1}{N_{oe}} \sum_{i:oe} \Big(\sum_{f \in oe} \big(\tilde{p}_{fit}/\psi_{fit}\big)^{1-\lambda} \Big)^{\frac{1}{1-\lambda}} \tag{4}
$$

The function excluding each of its argument requires a reference point for the excluded arguments, which plays an important role in this context. For the reference point of fuel productivity and prices, I will use the average productivity/price that plants using oil and electricity would get $\overline{\psi}_f(oe)$. For coal and gas, since there is no sample average productivity for plants that use oil and electricity only, I will simulate histories of coal and gas productivity for those plants, and take the sample average of the simulated history. I will do this with both the selected distribution of comparative advantages (from plants who use coal and/or gas in the data) and the estimated unselected distribution to see the extent to which selection on unobservables plays a role in explaining the difference in energy marginal cost. The function (4) is equal to zero when all its arguments are null, and thus satisfies the criteria laid out in **?**:

[5] We abuse notation here. A sub-model is an evaluation of function f with only some of its arguments. This language is motivated by the function corresponding in practice to the outcome of a regression or structural model. Formally when we write $f(X_1)$ we mean $f(X_1, \emptyset_2, ..., \emptyset_n)$, where we assume the j-th argument of the function can always take on a null value denoted $\emptyset_j$. The null value plays an important role in our context, because it serves as the reference point for the variable in question, which isn't 0.

$$
\begin{aligned}
f(\emptyset_p, \emptyset_\psi, \emptyset_{\mathcal{F}}) &= \frac{1}{N_{\mathcal{F}}}\sum_{i:\mathcal{F}}\Big(\sum_{f\in oe}\big(\overline{p}_f(oe)/\overline{\psi}_f(oe)\big)^{1-\lambda}\Big)^{\frac{1}{1-\lambda}} - \frac{1}{N_{oe}}\sum_{i:oe}\Big(\sum_{f\in oe}\big(\overline{p}_f(oe)/\overline{\psi}_f(oe)\big)^{1-\lambda}\Big)^{\frac{1}{1-\lambda}} \\
&= \Big(\frac{N_{\mathcal{F}}}{N_{\mathcal{F}}} - \frac{N_{oe}}{N_{oe}}\Big)\Big(\sum_{f\in oe}\big(\overline{p}_f(oe)/\overline{\psi}_f(oe)\big)^{1-\lambda}\Big)^{\frac{1}{1-\lambda}} = 0
\end{aligned}
$$

Since there are three arguments, there will always be 4 submodels which can exclude each of the arguments. For the first argument (prices), that would be: $\{0,0,0\},\{0,1,0\},\{0,0,1\},\{0,1,1\}$. Moreover, two of of the submodels exluding the variation of interest always contain one argument, and two of the submodels contain either 0 or two arguments. As such, the associated probabilities with each of the submodels will be as follows: $\{\frac{1}{3},\frac{1}{6},\frac{1}{6},\frac{1}{3}\}$. Overall, since the function takes 3 arguments, there are $2^3 = 8$ possible sub-models: $\{\underbrace{(0,0,0)}_{k=0},\underbrace{(1,0,0),(0,1,0),(0,0,1)}_{k=1},\underbrace{(1,1,0),(1,0,1),(0,1,1)}_{k=2},\underbrace{(1,1,1)}_{k=3}\}$. Then, as an example, the total partial effect of adding the gains from variety would be:

$$
\begin{aligned}
C_{\mathcal{F}} = \frac{1}{3}\Big(f(\emptyset_p,\emptyset_\psi,\mathcal{F}) - f(\emptyset_p,\emptyset_\psi,\emptyset_{\mathcal{F}})\Big) + \frac{1}{6}\Big(f(\mathbf{p},\emptyset_\psi,\mathcal{F}) - f(\mathbf{p},\emptyset_\psi,\emptyset_{\mathcal{F}})\Big) + \frac{1}{6}\Big(f(\emptyset_p,\mathbf{\Psi},\mathcal{F}) - f(\emptyset_p,\mathbf{\Psi},\emptyset_{\mathcal{F}})\Big) \\
+ \frac{1}{3}\Big(f(\mathbf{p},\mathbf{\Psi},\mathcal{F}) - f(\mathbf{p},\mathbf{\Psi},\emptyset_{\mathcal{F}})\Big)
\end{aligned}
$$

The total partial effect of adding fuel productivity is

$$
\begin{aligned}
C_{\mathbf{\Psi}} = \frac{1}{3}\Big(f(\emptyset_p,\mathbf{\Psi},\emptyset_{\mathcal{F}}) - f(\emptyset_p,\emptyset_\psi,\emptyset_{\mathcal{F}})\Big) + \frac{1}{6}\Big(f(\emptyset_p,\mathbf{\Psi},\mathcal{F}) - f(\emptyset_p,\emptyset_\psi,\mathcal{F})\Big) + \frac{1}{6}\Big(f(\mathbf{p},\mathbf{\Psi},\emptyset_{\mathcal{F}}) - f(\mathbf{p},\emptyset_\psi,\emptyset_{\mathcal{F}})\Big) \\
+ \frac{1}{3}\Big(f(\mathbf{p},\mathbf{\Psi},\mathcal{F}) - f(\mathbf{p},\emptyset_\psi,\mathcal{F})\Big)
\end{aligned}
$$

The total partial effect of adding fuel prices is

$$
\begin{aligned}
C_{\mathbf{p}} = \frac{1}{3}\Big(f(\mathbf{p},\emptyset_\psi,\emptyset_{\mathcal{F}}) - f(\emptyset_p,\emptyset_\psi,\emptyset_{\mathcal{F}})\Big) + \frac{1}{6}\Big(f(\mathbf{p},\emptyset_\psi,\mathcal{F}) - f(\emptyset_p,\emptyset_\psi,\mathcal{F})\Big) + \frac{1}{6}\Big(f(\mathbf{p},\mathbf{\Psi},\emptyset_{\mathcal{F}}) - f(\emptyset_p,\mathbf{\Psi},\emptyset_{\mathcal{F}})\Big) \\
+ \frac{1}{3}\Big(f(\mathbf{p},\mathbf{\Psi},\mathcal{F}) - f(\emptyset_p,\mathbf{\Psi},\mathcal{F})\Big)
\end{aligned}
$$

D.4 Aggregation of Production Function Without Switching and Without Heterogeneity in Fuel Productivity

In this section, I show that when plants 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 **?** 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, **?** 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

$$
\begin{aligned}
E &= \left(\sum_{f \in \{o,g,c,e\}} \beta_f e_f^{\frac{\lambda-1}{\lambda}} \right)^{\frac{\lambda}{\lambda-1}} \\
\sum_f \beta_f &= 1
\end{aligned}
\tag{5}
$$

In my paper, there are multiple plants with a different fuel sets $\mathcal{F}$ available to them. To be consistent with **?** and get aggregation results, I assume that all plants 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, plants in each fuel set have the following production function:

$$
E_{\mathcal{F}} = \left(\sum_{f \in \mathcal{F}} \beta_f e_f^{\frac{\lambda-1}{\lambda}} \right)^{\frac{\lambda}{\lambda-1}} \tag{6}
$$

From equation (6) 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 plants 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 plants that use each fuel sets such that $s_{oe} + s_{oge} + s_{oce} + s_{ogce} = 1$. I can then use these share of plants in each fuel set to define the total quantity of each fuel demanded by summing over all fuel set that use fuel f.

$$
\begin{aligned}
e_f &= \sum_{\mathcal{F}} \mathbb{I}(f \in \mathcal{F}) s_{\mathcal{F}} e_f(\mathcal{F}) \\
&= E \left(\frac{\beta_f}{p_f} \right)^{\lambda} \left(\sum_{\mathcal{F}} \mathbb{I}(f \in \mathcal{F}) s_{\mathcal{F}} P_E(\mathcal{F}) \right)
\end{aligned}
\tag{7}
$$

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 (7).

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 (7).*

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

$$\tilde{E} = \left(\sum_{f \in \{o,g,c,e\}} \psi_f^{\frac{1}{\lambda}} e_f^{\frac{\lambda-1}{\lambda}} \right)^{\frac{\lambda}{\lambda-1}} \tag{8}$$

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 plants who are using each fuels. The cost-minimizing quantity of each fuel from the production function in (8) for a given quantity of energy E is

$$\tilde{e}_f = s_f E \left(\frac{P_{\tilde{E}}}{p_f} \right)^{\lambda} \tag{9}$$

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 (9) is equal to e_f from (7).

□

E 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

E.1 Energy Task Model

for a given quantity of realized energy, plants allocate fuels to energy tasks that compose a given unit of E in an inner nests. This task model is very similar to **?** 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.

$$E = \inf \left\{ \tau(\omega) : \omega \in [0,1] \right\} \tag{10}$$

I assume that tasks are perfect complements, where plants 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 plant's fuel set $\mathcal{F}$, where fuels are in principle perfectly substitutable at performing each task:

$$\tau(\omega) = \sum_{f \in \mathcal{F}} \psi_f(\omega) e_f(\omega) \tag{11}$$

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 (**?**). 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.

E.2 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:

$$\begin{aligned} \int_0^1 p(\omega)\tau(\omega)d\omega &= E \int_0^1 p(\omega)d\omega \\ &= Ep_e \end{aligned} \tag{12}$$

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:

$$p(\omega) = \min_{f\in\mathcal{F}} \left\{ \frac{p_f}{\psi_f(\omega)} \right\} \tag{13}$$

Aggregation from tasks to fuels

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

$$\mathcal{T}_f = \left\{ \omega : \frac{p_f}{\psi_f(\omega)} \leq \frac{p_j}{\psi_j(\omega)}\ \forall j \neq f \in \mathcal{F} \right\} \tag{14}$$

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

$$e_f(\omega) = \begin{cases} E\psi_f(\omega)^{-1} & \text{if } \omega \in \mathcal{T}_f \\ 0 & \text{otherwise} \end{cases} \tag{15}$$

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):

$$e_f = E \underbrace{\int_{\mathcal{T}_f} \psi_f(\omega)^{-1} d\omega}_{\Gamma_f^{-1}} \tag{16}$$

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.[6] I can now rewrite the price index of realized energy as a weighted sum of fuel prices:

$$\begin{aligned} p_e &= \frac{1}{E} \int_0^1 p(\omega)\tau(\omega) d\omega \\ &= \frac{1}{E} \sum_f \int_{\mathcal{T}_f} \frac{p_f}{\psi_f(\omega)} E \\ &= \sum_f p_f \Gamma_f \end{aligned} \tag{17}$$

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:

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

$$
\begin{aligned}
p_e E &= \sum_f \int_{\mathcal{T}_f} \frac{p_f}{\psi_f(\omega)} \frac{E}{\mathcal{M}} \\
&= \sum_f p_f e_f
\end{aligned} \tag{18}
$$

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

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

Proposition 1 highlights the option value that an additional fuel provide. Indeed, when a fuel is added, plants 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 **?**:

$$
\widehat{\overline{E}\Gamma_{fit}} = e_{fit}\Big(\frac{\overline{E}}{E_{it}}\Big) \tag{19}
$$

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 (**?**). The second assumption is standard in the literature on technological choice (**???**).

Assumption 1. *1Symmetric 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:

$$\mathcal{T}_{it} = Pr\Big(\frac{p_{fit}}{\psi_{fit}} \leq \frac{p_{jit}}{\psi_{jit}} \forall j \neq f \in \mathcal{F}_{it}\Big) \tag{20}$$

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 plants make.

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

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:

$$\Gamma_f = \int \psi_f^{-1} \underbrace{\Big[\prod_{j \neq f} \int \mathbb{I}\Big(\psi_f \geq p_f \max\{\psi_j/p_j\}\Big) f(\psi_j) d\psi_j\Big]}_{\text{Pr(efficiency draw for f is chosen over all other fuels)}} f(\psi_f) d\psi_f \tag{21}$$

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. *2Pareto 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

$$\Gamma_{g,it} = \underbrace{\frac{\theta}{\theta+1}\Omega_{git}^{-\theta-1}\overline{\psi}_{git}^{\theta}}_{\text{Direct task displacement}} - \underbrace{\left(\frac{p_{cit}}{p_{git}}\right)^{-\theta}\frac{\theta}{2\theta+1}\Omega_{git}^{-2\theta-1}(\overline{\psi}_{git}\overline{\psi}_{cit})^{\theta}}_{\text{Indirect task displacement through fuel c}} \tag{22}$$

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.

E.3 Non-parametric Energy Production Function

Given a fuel set $\mathcal{F}$, plants 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 plant is as follows:

$$\min_{\{e_{fit}\}_{f\in\mathcal{F}_{it}},\lambda} \sum_{f\in\mathcal{F}} p_{fit}e_{fit} + \lambda\Big(E_{it} - g(\mathbf{\Psi}_{it}\mathbf{e}_{it})\Big) \tag{23}$$

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 outer production function allows me to do in the main text.[7] 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 plants' optimality conditions. Using Sheppard's Lemma, I can characterize optimal fuel choices as the derivative of the unit cost with respect to fuel prices:

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

Which gives me a system of structural equations:

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

? 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.[8] 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

[7]Note that in the main text, I assume an energy production function that is homogeneous of degree 1. However, it can be easily extended to homogeneity of degree $k \geq 1$ (**?**).

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

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.

References

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

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.

Boehm, Johannes and Ezra Oberfield, "Misallocation in the market for inputs: Enforcement and the organization of production," *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.

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.

Galitsky, Christina and Lynn Price, "Opportunities for Improving Energy Efficiency, Reducing Pollution and Increasing Economic Output in Chinese Cement Kilns," *Proceedings of the American Council for An Energy Efficient Economy's*, 2007.

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.

Khanna, Shefali, "Electricity Market Design and Market-Based Environmental Policy in India," 2021.

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

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.

Matzkin, Rosa, "Identification in Nonparametric Simultaneous Equations Models," *Econometrica*, 2008, *76*, 945–978.

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

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#.

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.

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

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

Shorrocks, Anthony F., "Decomposition procedures for distributional analysis: a unified framework based on the Shapley value," *The Journal of Economic Inequality*, 2013, *11*, 99—-126.

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.