Balancing Economic Activity and Carbon Emissions: A Study of Fuel Substitution in India*

Emmanuel Murray Leclair†

Department of Economics, Western University

This Version: *August 6, 2023*

Abstract

Policies such as Pigouvian carbon taxes that increase fuel prices in proportion to their emission intensity are often considered a solution to internalize the social cost of pollution. In this paper, I develop a dynamic production model incorporating multidimensional energy input choices and fuel productivity heterogeneity. Leveraging advancements in production function estimation, I identify fuel productivity while accounting for the costs associated with inter-temporal switching between fuel sets. I estimate the model using a panel of steel establishments from the Indian Survey of Industries (2009-2016) to examine establishments' responses to carbon taxation via fuel-specific tax rates, which affect input choices. I demonstrate that accounting for heterogeneity in fuel productivity and inter-temporal switching between fuel sets significantly reduce the economic cost of decreasing emissions through policy. Additionally, I show that using proceeds from carbon taxation to subsidize the fixed cost of cleaner fuel adoption only minimally improves welfare.

*I would like to thank my advisors David Rivers, Salvador Navarro and Sergio Ocampo Diaz for their incredible support throughout the elaboration of this project. I also want to thank Rory McGee, Nouri Najjar, Nirav Mehta, Charles Saunders, Martin O'Connell, Matt Gordon, Yanyou Chen, Richard Sweeney, Pascual Restrepo, Elie Tamer, Yongseok Shin, and Andrew McCallum for their useful comments. I also want to thank participants that gave me valuable feedback at the 56th and 57th Annual Conference of the Canadian Economics Association, the 2023 UM-MSU-UWO Labour Day Conference, the 21st Annual International Industrial Organization Conference (IIOC 2023), and the 2022 Young Economist Symposium (YES). All errors are my own.

†email: emurrayl@uwo.ca

1 Introduction

Fossil fuels are widely used in manufacturing industries, where large quantities of fuels are burnt as part of industrial processes such as the melting of iron ore to produce steel and the calcination of limestone to produce cement. While providing energy for production, fossil fuels also contribute to negative externalities such as air pollution and climate change. Overall, manufacturing activity accounts for 37% of global greenhouse gas emissions (Worrell, Bernstein, Roy, Price and Harnisch, 2009), which led to a number of important studies on fuel substitution and carbon-based policies across diverse manufacturing industries (Ganapati, Shapiro and Walker, 2020; Hyland and Haller, 2018; Fowlie, Reguant and Ryan, 2016; Ryan, 2012; Stern, 2012). Following recent evidence calling for a more in-depth investigation into fuel substitution, this paper aims to refine our understanding of this phenomena.

One key finding emerging from Lyubich, Shapiro and Walker (2018) is that firms vary substantially in how productive they are at using energy inputs. These disparities in productivity may stem from divergent heat efficiency, inherent to different fuel-burning technologies (Allcott and Greenstone, 2012). Furthermore, the proficiency of a plant's workforce can play an important role in how optimally resources are allocated towards the operation of specific fuel-burning technology, and certain industrial facilities implement energy retrofit programs aimed at curbing energy waste (Christensen, Francisco and Myers, 2022) while others stick with old and inefficient energy consumption practices (Indian Ministry of Steel, 2023).

Another significant finding pertains to the enduring impact of large fixed costs and time commitments associated with the adoption of new fuels (Scott, 2021). This encompasses a spectrum of changes, including technological adaptations, new storage facilities, and the establishment of requisite transportation infrastructure. For instance, consider the scenario where a steel plant seeks to transition from coal to natural gas. Such transition necessitate going from coal-based blast furnaces to electric arc furnaces, as well as the installation of dedicated distribution pipelines, directly linking the plant to existing transmission networks. Consequently, the inertia stemming from entrenched technologies, commonly referred to as technological lock-in (Hawkins-Pierot and Wagner, 2022), impedes the seamless migration from dirty to cleaner fuel alternatives. Together, these findings challenge the prevailing assumptions of fully flexible and static fuel substitution (Ganapati, Shapiro and Walker, 2020; Hyland and Haller, 2018; Wang and Lin, 2017; Ma, Oxley, Gibson and Kim, 2008; Cho, Nam and Pagan, 2004; Pindyck, 1979), which may have implications for the economic cost of reducing emissions.

In this context, this paper aims to provide a more nuanced understanding of the role of fossil fuels in plants' input mix and how fuel demand responds to policies such as carbon taxation. Second, it aims to quantify the welfare implications of such policies, particularly emphasising the trade-off between social benefit of emission reduction and economic cost of increased fuel prices. As a preview of results, I find that the economic cost of reducing emissions through a carbon tax is significantly lower than previously thought – largely due to heterogeneity in plants' exposure to the tax. To alleviate fixed costs, I then show how proceeds from the tax can be used to subsidize the cost of cleaner fuel adoption, and find positive but minimal welfare effects.

To get at these results, I develop and estimate a dynamic production model with multidimensional energy input (fuel) choices and heterogeneity in fuel-augmenting productivity. The model features monopolistic competition and two nests of production: an outer nest with capital, labor, intermediate inputs and energy, and an inner nest where plants combine fuels to produce energy. Fuel choices are then separated between an inter-temporal fuel set choice subject to fixed switching costs and a within-period relative fuel quantity choice conditional on the fuel set. Consistent with the literature on input complementarity (Broda and Weinstein, 2006), there is an option value from adding more fuels to a set, which decrease marginal costs. In the absence of fixed costs, plants 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.

Quantifying the role of fossil fuels in this production model highlights two important measurement issues. First, energy that plants 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 is referred to as *potential energy*.[1] The wedge between potential and realized energy underlie differences in the productivity of fuels that compose energy. Second, studying switching between fuel sets underlie dynamic selection that hampers the evaluation of counterfactual costs and production under different fuel sets. For example, consider a plant deliberating whether to use coal and/or gas in the upcoming period. If the plant chooses to use coal exclusively, its choice might hinge on an anticipation of high coal productivity and low gas productivity. However, the researcher lacks insight into the plant's actual gas productivity as it abstains from using gas. This holds significance due to the potential for policy to prompt plants to adopt counterfactual fuel combinations.

To address these issues, I rely on important development on the estimation of production function

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

and dynamic discrete choice models in the presence of unobserved heterogeneity. My identification method follows three steps. First, I identify the quantity and price of *realized energy*, jointly with demand and the outer nest of production following the work of Grieco, Li and Zhang (2016) and Ganapati, Shapiro and Walker (2020). This method relies on optimality conditions from profit maximization to map observed relative input spending to unobserved relative input quantities. Second, I identify the distribution of fuel productivity across plants and the inner nest of production following the work of Zhang (2019) and Blundell and Bond (1998, 2000). This allows me to exploit first-order conditions to recover relative fuel productivity that equate relative fuel prices to relative marginal products. The energy production function is then identified using lagged inputs and prices as instruments. Third, I follow Arcidiacono and Jones (2003); Arcidiacono and Miller (2011) to recover fixed costs and the distribution of fuel productivity for counterfactual fuel sets. Using this three-step approach, I am able to recover all production function parameters, the distribution of fuel productivity, and switching costs between fuel sets. This allows me to conduct policy counterfactual that affect plants' fuel choices at the intensive and extensive margin.

I then apply this model to the Indian steel industry between 2009 and 2016 using data from the Indian Survey of Industries (ASI), a panel of manufacturing establishments. The panel features quantities and prices of disaggregated inputs that plants purchase and outputs that plants manufacture, as well as plants' location into 775 districts, which I map to the entire network of natural gas pipelines. 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 to steel manufacturing because it is one of the most environmentally damaging industries in India, with coal accounting for nearly 70% of its energy sources. In this context, there are various shocks affecting plants that substantively help estimating the model. These include aggregate factors that affected fuel prices such as the global oil shock of 2014, as well as the expansion of the natural gas pipeline network between 2009 and 2016.

Preliminary results suggest a higher elasticity of substitution among fuels than between fuels and non-energy inputs. I also find large and persistent heterogeneity in fuel productivity, consistent with stylized facts about productivity found by Bartelsman and Doms (2000) and Syverson (2011). Moreover, plants with more fuels in their set face a significantly lower marginal cost of energy even after controlling for fuel prices. This phenomenon can be explained by factors such as higher fuel-specific productivity, and the option value that an additional fuel provides. Overall, plants that do not utilize natural gas would be 30% less productive at using gas compared to their gas-utilizing counterparts. This disparity is accentuated by significant fixed costs of natural gas adoption, averaging between 28 and 40 million U.S. dollars, and natural gas prices quintupling coal prices. These interrelated factors collectively point towards a technological lock-in, shedding light on the prevalent utilization of coal within the Indian steel sector.

On the policy side, I first explore a carbon tax levied on fossil fuels, where the relative tax rate on each fuel is equal to its marginal externality damage. I explore various levels of the tax corresponding to different social cost of carbon (SCC) to characterize the trade-off between emission reduction and output along various percentages of emission reduction. While the trade-off is nonlinear because the marginal cost of emission reduction is increasing in output, I find that reducing aggregate emissions by 50% implies a reduction in aggregate output of only 7% relative to a *laissez faire* economy. In contrast, If heterogeneity in fuel productivity was omitted from the model, output would decrease by 12% for the same reduction in emissions. Plants who are productive at using high emission fuels such as coal and oil specialize in those fuels, and are thus more exposed to the carbon tax. Consequently, these plants become less competitive, and some output reallocates from high emission to low emission plants. This composition effect induced by heterogeneity in fuel productivity reduces the economic cost of emission reduction.

I then show how proceeds from the carbon tax can be used to subsidize the adoption of natural gas to help alleviate technology lock-in induced in parts by large fixed costs. I show that a 10% subsidy can be fully financed by a carbon tax with a social cost of carbon of $51 per ton of carbon dioxide. The subsidy leads to a 26% increase in natural gas uptake, from 19 % to 24 % of establishments. This increases the net present value of variable profits in the economy by USD 19 millions (0.09%), consumer surplus by USD 14 millions (0.05%), and emission damages by USD 10 millions (0.38 %). The net effect is a positive but small welfare increase of USD 1.18 millions (0.003%). However, this welfare effect is very small when compared to the total subsidized amount of USD 2.79 billions, as the net societal benefits only account for 0.04% of the subsidy's cost.

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 show that empirically matching these choices has multiple new implications, requiring a more flexible model. Indeed, matching the intensive margin of observed fossil fuel consumption, particularly the heterogeneity in relative fuel shares, has implication for fuel-specific productivity. Additionally, matching observed inter-temporal fuel set switching has implication for fixed costs, the option value that different sets provide, and dynamic selection à la Roy (1951) when combined with fuel productivity. Together, these choice margins underlie novel welfare implications of policies aimed at mitigating externality damages from emissions, such as carbon taxes. Along the way, I contribute to multiple strain of literature.

First, I contribute to the literature on production function estimation (Olley and Pakes, 1996; Blundell and Bond, 2000; 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 how to identify and estimate a dynamic production function with input-augmenting productivity, where some of the inputs are not always used by plants and can change over time, which creates dynamic selection on unobservables. I solve this selection problem by combining the aforementioned literature with methods from the dynamic discrete choice literature in the presence of unobserved heterogeneity (Arcidiacono and Jones, 2003; Arcidiacono and Miller, 2011).

Second, I contribute 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 pollution 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 macroeconomic literature on climate change using integrated assessment models (IAM) (Golosov, Hassler, Krusell and Tsyvinsky, 2014; Hambel, Kraft and Schwartz, 2021; Miftakhova and Renoir, 2021; Dietz, van der Ploeg, Rezai and Venmans, 2021). I show that commonly made assumptions on the aggregate production function for energy that combines different fuels may understate the extent of fuel substitution in the economy.

Third, 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. There is also one exception by Hawkins-Pierot and

[2]The terms "productivity" and "efficiency" can be used interchangeably in this context, and throughout the paper.

Wagner (2022), who estimate the energy productivity of manufacturing plants and its implication for technology lock-in. 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 the relative productivity of different fuels. I show that the distinction between fuel productivity and energy productivity is crucial to understand the impact of carbon policy on the economy.

Fourth, 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 paper provides some evidence in support of this theory. Indeed, the prevalence of switching between fuel sets combined with large fixed costs of fuel adoption can be explained by such gains from variety, in which marginal plants are willing to pay a fixed cost to reduce variable production costs through input complementarity and ability to substitute when facing fuel-specific shocks. This is particularly relevant in the Indian context which often face electricity shortages(Allcott, Collard-Wexler and O'Connell, 2016; Mahadevan, 2022; Ryan, 2021). This argument is strengthen by the observation that plants with more fuels tend to face lower marginal cost of production.

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 how the outer and inner (energy) production model can be identified and estimated. In section 6, I show how fixed costs and fuel productivity for counterfactual fuel sets can be identified and estimated. In section 7, I discuss estimation results for for Indian steel industry. In section 8, I discuss policy counterfactuals.

2 Data

I use longitudinal data on inputs, location, and emission of manufacturing establishments in India. This allows for an in-depth analysis of the role of fuels in the production processes of these establishments and the evolution of fuel usage 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 with 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. More details on sampling rules, including changes over time, 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). Since the ASI is known to contain large number of extreme outliers (Bollard, Klenow and Sharma, 2013), I follow standard practices by top-coding and bottom-coding the 1% tails of plant-level inputs and output by industry (winsoring).[3] The analysis is applicable to industries that uses fossil fuels for combustion such as steel, non-ferrous metals, cement, glass, pulp & paper, etc.

Location I extract detailed location data from the publicly available version of the ASI, enabling me to assign establishments to 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 (CO_{2e}). This method is applicable to manufacturing industries that use fossil fuels for combustion, where each

[3] Such outliers are typically due to reporting errors, and are inconsistent with a wide range of official statistics.

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, Annexure 3). 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 using the WPI for machinery as well as an India-specific capital deflator from the Penn World Table Feenstra, Inklaar and Timmer (2015). Both deflation methods produce similar outcomes. 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 this data, 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 choices.

Fact 1: High Pollution Levels from Indian Manufacturing Establishments

Indian establishments in heavy manufacturing industries exhibit high levels of pollution. Total annual greenhouse gas emissions averages 25 million tons of CO_{2e} in Steel manufacturing, and 55 million tons of CO_{2e} in Cement manufacturing, together accounting for nearly half of annual emissions in Indian manufacturing. See table 1. This is primarily due to two reasons. First, the aggregate share of coal as part of the energy mix is significantly larger than other fuels, and is much larger than in developed countries, who typically rely on 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 appendix A.5.1, I provide evidence comparing Indian with Canadian establishments in similar industries, and I find stark differences in pollution intensity. Second, the aggregate share of energy as part of the input mix averages 23% in these industries, significantly larger than the average across all other Indian manufacturing industries. At a high level, this evidence reinforces the scope of this paper in studying fuel substitution among highly polluting and energy-intensive establishments.

Industry	Annual Average Number of Plants	Average Annual Revenue (Million USD)	Annual Total Emissions (million tons co_{2e})	Aggregate Energy Input Share	Aggregate Coal Fuel Share
Paper	275	6.35	6.52	0.16	0.89
Glass	199	4.16	0.92	0.22	0.01
Cement	317	21.87	54.79	0.41	0.93
Steel	1,080	17.61	24.76	0.12	0.69
Other	32,224	5.67	157.95	0.11	0.17

Table 1: Descriptive Statistics for Selected Industry (2009-2016)

Note: The energy input share is calculated as the aggregate spending on energy by industry, as a fraction of total spending on labor, materials and energy. It is then averaged across years. Similarly, the coal fuel share is calculated as the aggregate share of coal (in mmBtu) relative to other fuels in each industry, averaged across years.

Fact 2: Indian Manufacturing Establishments use Different Fuel Sets

I find that plants operating in narrowly defined industries use different fuel mixes at any given time. Moreover, the vast majority of fuel sets include both oil and electricity. Most of the variation in fuel sets thus comes from whether plants also use coal, natural gas, neither or both. There are many reasons why plants in the same industry use different fuel sets. For example, plants in steel manufacturing can use different types of furnace to melt iron ore. They can use blast furnaces which relies on coke (coal) as a primary fuel, or electric arc furnaces which can burn natural gas, oil, or coal to generate high voltage electricity, which is then discharged through an electric arc to melt iron ore.

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

Notes: The *Other* category comprises of any other mix of these 4 fuels. The total number of plant-year observations is $10,360$ in Steel, $5,889$ for Casting of Metal, $3,162$ in Cement and $2,009$ for Glass.

This heterogeneity in fuel sets has many implications. First, plants that use coal are likely to release more CO_{2e} than plants who use gas. Second, plants who have more fuels have access to additional margins of substitution, which are 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 due to the civil war in Libya and economic sanctions against Iran (see figure 1). In this volatile geopolitical landscape, plants' decisions carry significant weight in terms of how they plan to weather potential future shocks and coal may be particularly desirable due to the stability of its price relative to other fuels.

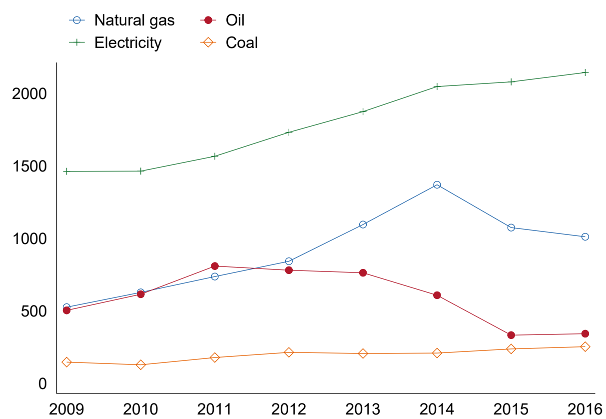


Figure 1: Median Fuel Prices

Additionally, the Indian 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. (Mahadevan, 2022). Fuel substitution can serve as a means of adjustment for plants to insure themselves against these shortages, for example, by using other fuels to perform tasks that were previously performed by electricity, or even by using other fuels to generate electricity in-house.

Fact 3: Indian Manufacturing Establishments Often Switch Between Fuel Sets

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 underestimate of switching because I only observe plants for a maximum of 8 years. See Appendix A.5.2. 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.5.2, I also look at fuel switching in U.S. based plants and find similar results.

There may be many reasons why plants switch between fuel sets.[4] The development of new technologies may increase the productivity of some fuels (e.g. electric arc furnaces vs. blast furnaces, rotary vs vertical shaft kilns. See Appendix A.2 for more details.). Large and persistent fuel

[4]In addition to the explanations provided in the text, some establishments may change their position in the supply chain, which requires them to either perform new energy-intensive tasks, or outsource existing tasks to other plants. While this paper does not tackle this explanation directly, Boehm and Oberfield (2020) provide a way to measure the level of vertical integration of an establishment from the ASI dataset. In Appendix A.5.6, I construct this measure and assess whether switching between fuel sets is correlated with changes in vertical integration.

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

Table 3: Percentage of unique plants that **add** and **drop** a fuel.

Notes: To construct the table, I partially balanced the panel, keeping establishments that operate in at least 5 years between 2009-2016 ($N = 120,916$). When keeping plants that operate in all 8 years, the prevalence of adding and dropping fuels increases to 50% for both. When keeping plants that operate for at least 2 years, the prevalence of adding fuels reduces to 16% and the prevalence of dropping fuels reduces to 18%. Nevertheless, this is an underestimate of switching because many plants exist before and after the sample period.

price shocks may provide incentives for plants to readjust their input mix, and the expansion of transportation infrastructures, particularly pipeline networks, may ease the access to new fuels. see Fact 5. The next two facts provide further clarity to understand the heterogeneity in fuel sets and the prevalence of fuel set switching.

Fact 4: Establishments Increase the Number of Fuels as they Become More Productive

As plants produce more output per worker, which can be interpreted as a proxy for productivity, they tend to increase the number of fuels in their set. Moreover, I show in Appendix A.5.3 that this pattern also persist across plant age. I interpret these findings as suggestive of the importance of fixed costs for fuel adoption. Indeed, establishments would like to use more fuels to leverage their complementarity, but may not always find it profitable to pay the necessary fixed costs. More productive plants have marginally more to gain from combining multiple variety of fuels since any gain in marginal products will lead to larger increase in total product[5]. This is the same type of mechanism that explains why more productive plants are more likely to pay fixed costs to enter international markets in trade models (Melitz, 2003).

Moreover, there are plenty of substantive reasons why fixed costs may be relevant. Expensive Furnaces/kilns are required to extract the heating potential of fuels, and plants who want to use natural gas may need to pay for distribution pipelines that connect their plant directly to main transmission lines (Scott, 2021). In this context, a positive gradient between the number of fuels and productivity coupled with large fixed costs of fuel adoption may create situations of technological lock-in where plants are not productive enough to overcome fixed costs. Similar technological lock-in have been previously documented in manufacturing by Hawkins-Pierot and Wagner (2022), and

[5]This is one particular mechanism that can explain why more productive/older plants use more fuels, and will be the primary machanism in this paper. Other complementary explanations include borrowing constraints due to a lack of collateral that prevents small plants from investing in new infrastructure to diversify their fuel choices. This explanation could rationalize Figure 2 if plant size is correlated with productivity. This is something that I can include in the paper, albeit at a large computational cost.

has important implications for policy to help plants overcome this technological lock-in.

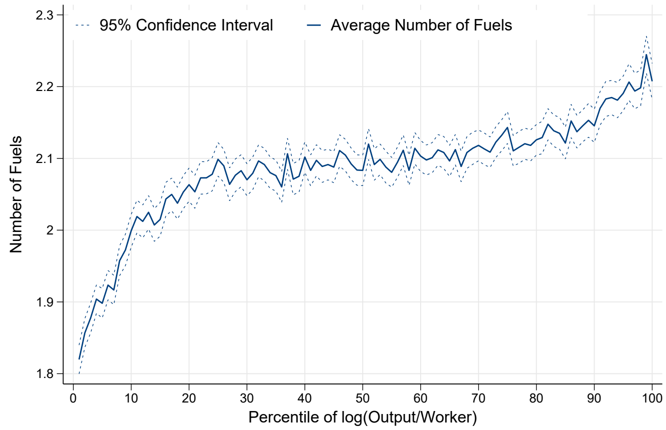


Figure 2: Number of Fuels by Output per Worker, Average of all ASI Plants

Fact 5: Proximity of Plants to Pipelines Increases the Probability of Gas Usage

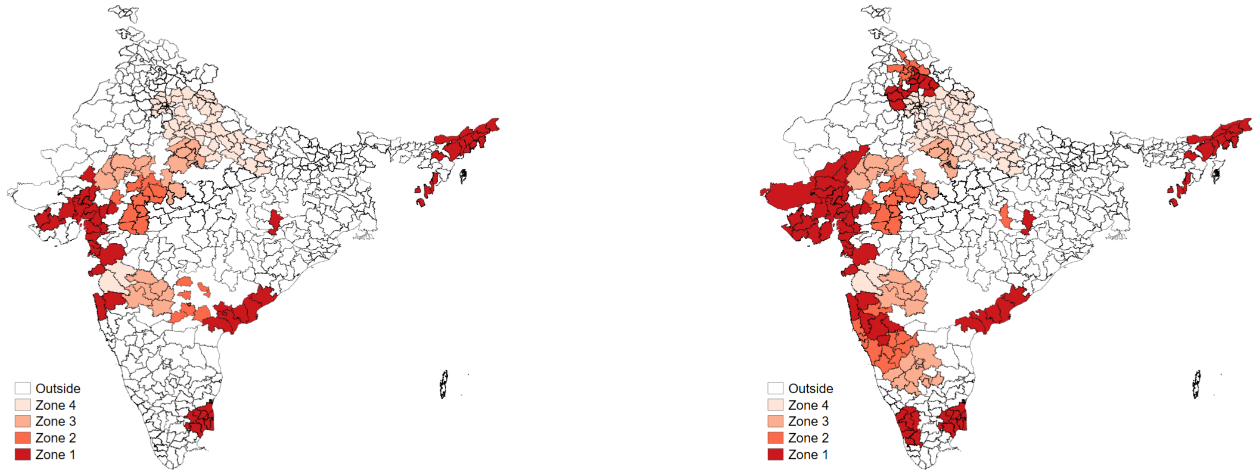
Plants located near transmission pipelines have access to the main distribution network. This access reduces costs compared to those located far from the pipeline network, who need to either have access to or construct expensive gasification terminals to convert liquified natural gas (LNG) to its usable 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 $t+1$.[6] 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.

4 Model

Consistent with the evidence provided so far, I develop and estimate a rich dynamic production model which allows me to quantify establishments' fuel choices, both at the intensive and the

[6] While pipeline expansions are not exogenous, I show in Appendix A.5.4 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.



(a) 2009 (b) 2016

Figure 3: 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.

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.

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 $Pr(\Delta Dgas_{it} = 1 \mid \Delta Dpipeine_{it} = 1, controls_{it}) - Pr(\Delta Dgas_{it} = 1 \mid \Delta Dpipeine_{it} = 0, controls_{it})$.

extensive margin. Decisions in the model are reliant on the channels discussed above, and include many reasons why establishments use different fuels. Particularly relevant are differences in fuel productivity due to variation in heat management practices and technologies (furnaces/kilns) that plants operate, as well as heterogeneity in fuel prices and fixed costs. This model enhances the existing literature on externality mitigation by more accurately predicting the impact of policies like carbon taxes on manufacturing establishments' emissions, achieved through alterations in plants' input mix.

Each period, plants have access to a set of fuels from a combination of oil, natural gas, coal and electricity. Fuels are combined to produce energy that goes into the outer nest of production. The production model for energy is the same across fuel sets, but each fuel in the plant's set has its own productivity term. Plants can choose to change fuel sets across periods, either by adding new fuels and/or dropping existing fuels. Ideally, a plant would use all fuels to leverage their option value, and gain an additional margin of substitution to hedge against various shocks. However, adopting a new fuel requires paying a fixed cost, because the utilization of fuels involves complex heating processes that require fuel-burning technologies such as furnaces/kilns, storage facilities, and transportation infrastructures. For similar reasons, a plant can get a salvage value from dropping a fuel. As a result, adding a new fuel involves a trade-off between between reduction in contemporaneous profits and decreases in expected future marginal costs, and vice-versa for dropping an existing fuel.

To understand these dynamics better, 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 at the industry level, so I omit the industry subscript going forward.

4.1 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 $(E_{it}/\overline{E})$.[7] Following Grieco et al. (2016), the production function is explicitly normalized around the geometric mean of each variable $\overline{X} = \left(\prod_{i=1}^{n}\prod_{t=1}^{T} X_{it}\right)^{\frac{1}{nT}}$.[8]

[7] The particular functional form of the CES is not necessary. Identification works for a large class of production functions. This is true for both nests of the production function.

[8] 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

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

$$s.t. \quad \alpha_L + \alpha_K + \alpha_M + \alpha_E = 1$$

Where $\sigma \geq 0$ is the elasticity of substitution between inputs, and $\eta > 0$ is the returns to scale. In the outer nest, plants choose input quantities given input prices, which includes realized energy, $\frac{E_{it}}{\overline{E}}$. Then, given the current fuel set $\mathcal{F}_{it} \subseteq \mathbb{F} = \{\text{oil, gas, coal, elec}\}$, plants combine all fuels available in the set to produce a quantity of 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. p_{fit} and ψ_{fit} are the corresponding fuel price and productivity, respectively.[9] The fuel-specific productivity terms are novel as they allow for flexible variation input usage at the intensive margin and heterogeneity in fuel substitution. This is a significant departure from the literature, where most previous papers that estimate a production function with fuels do not allow for fuel-specific productivity (Hyland and Haller, 2018; Ma et al., 2008; Pindyck, 1979; Joskow and Mishkin, 1977; Atkinson and Halvorsen, 1976). More recently, Hawkins-Pierot and Wagner (2022) allowed for the productivity of the total energy bundle to vary across plants. While this allows for heterogeneity in the substitution between energy and other inputs, it does not capture salient features of fuel consumption and differential responses to fuel price changes.

$\lambda > 0$ is the elasticity of substitution between fuels. This parameter plays a crucial role in this model because it determines the magnitude of the option value that a plant would get by expanding its fuel set $\mathcal{F}_{it}$. As long as $\lambda > 1$, there is an option value to have more fuels. However, the more complements fuels are conditional on being gross substitutes, the larger is the option value. This is because a lower λ implies that marginal products from a given fuel decrease faster with quantity, so there are larger marginal gains from adding a new fuel. In section 4.4, I explore these comparative statics in more details.

in the appendix

[9] Note that fuel productivity terms ψ_{fit} are in units of normalized fuel quantities $e_{fit}/\overline{e}_f$, but can always be rewritten in units of physical fuel quantities (mmBtu): $\tilde{\psi}_{fit} = \psi_{fit}/\overline{e}_f$

On the policy side, both elasticity of substitution (σ, λ) and the return to scale η will play an important role in the own and cross price elasticity of fuel demand, which will be paramount to understand how plants readjust their input mix when facing different policies. Additionally, heterogeneity in fuel productivity may induce distributional effects from taxation which will affect the aggregate trade-off between emission reduction and output. Next, I show how plants compete and set prices.

4.2 Static Decisions

Assumption 1. *Plants produce different output varieties and engage in monopolistic competition.*

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}. Total output is produced by aggregating all the varieties with standard Dixit-Stiglitz preferences across varieties. Given a mass of N_t operating plants, income I_t and an aggregate demand shock e^{Γ_t}, the representative consumer solves:

$$\begin{aligned}
&\max_{\{Y_{it}\}_{i=1}^{N_t}, Y_{0t}} \mathbb{U} = Y_{0t} + \frac{e^{\Gamma_t}}{\theta}\left(\frac{1}{N_t}\int_{\Omega_i}(N_t Y_{it})^{\frac{\rho-1}{\rho}}di\right)^{\frac{\theta\rho}{\rho-1}} \\
&\quad s.t. \quad Y_{0t} + \int_{\Omega_i} P_{it}Y_{it}di \leq I_t
\end{aligned} \tag{3}$$

Where $\rho > 1$ is the elasticity of substitution between varieties, and $\theta \in (0,1)$ indexes the substitution between consumption of the differentiated varieties and the outside good. Following Helpman and Itskhoki (2010), I restrict $\theta < \frac{\rho-1}{\rho}$, which ensures that output varieties are more substitutable between each others than with the outside good. These quasi-linear CES preferences were first proposed by Helpman and Itskhoki (2010), and provides analytical convenience for welfare evaluation. Indeed, quasi-linear preferences are standard in the literature on externality taxation (Fowlie et al., 2016) and allow researchers to use the social cost of carbon (SCC), which expresses the net present value of expected future damages from carbon emissions in dollars.[10] As such, externality damages affect consumption of the outside good by varying aggregate income, and thus directly affect consumer surplus. Solving the representative consumer's problem in (3) yields the following downward slopping demand for each varieties Y_{it}, which is augmented with an ex-post

[10]This is the approach typically taken in applied microeconomics. However, there is an alternative approach in macroeconomics which relies on integrated assessment models (IAM) to explicitly study the dynamic relationship between aggregate emissions and concentration of CO_2 in the atmosphere, which affects future aggregate output in various ways. See Nordhaus (2008); Golosov et al. (2014); Hassler et al. (2019, 2020); Golosov et al. (2014).

idiosyncratic demand shock $e^{\epsilon_{it}}$:

$$Y_{it} = \frac{e^{\tilde{\Gamma}_t}}{N_t} P_{it}^{-\rho} P_t^{\frac{\rho(1-\theta)-1}{1-\theta}} e^{\epsilon_{it}} \tag{4}$$

Where $e^{\tilde{\Gamma}_t} = e^{\Gamma_t \frac{1}{1-\theta}}$ and $P_t = \left(\frac{1}{N_t}\int P_{it}^{1-\rho} di\right)^{\frac{1}{1-\rho}}$ is the CES aggregate price index across all varieties. Detailed derivations can be found in Appendix B.1.

Plants maximize profits

Given set of fuels $\mathcal{F}_{it} \subseteq \mathbb{F}$, technological constraints, inverse demand, and all input prices, the plant's problem is a static profit maximization.[11] 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.

$$\begin{aligned}
&\max_{K_{it}, M_{it}, L_{it}, \{e_{fit}\}_{f\in\mathcal{F}_{it}}} \left\{ P_{it}(Y_{it})Y_{it} - w_t L_{it} - r_{kt}K_{it} - p_{mit}M_{it} - \sum_{f\in\mathcal{F}_{it}} p_{fit}e_{fit} \right\} \\
s.t. \quad & \tilde{Y}_{it} = z_{it}\left[\alpha_K \tilde{K}_{it}^{\frac{\sigma-1}{\sigma}} + \alpha_L \tilde{L}_{it}^{\frac{\sigma-1}{\sigma}} + \alpha_M \tilde{M}_{it}^{\frac{\sigma-1}{\sigma}} + \alpha_E \Big(\sum_{f\in\mathcal{F}_{it}} (\psi_{fit}\tilde{e}_{fit})^{\frac{\lambda-1}{\lambda}} \Big)^{\frac{\lambda}{\lambda-1}\frac{\sigma-1}{\sigma}} \right]^{\frac{\eta\sigma}{\sigma-1}} \\
& P_{it}(Y_{it}) = \left(\frac{e^{\tilde{\Gamma}_t}}{N_t Y_{it}}\right)^{\frac{1}{\rho}} P_t^{\frac{1+\rho(\theta-1)}{(\theta-1)\rho}}
\end{aligned}$$

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, plants find the combination of fuels that minimizes the cost of producing a given unit of energy. Note that fuel prices in mmBtu are observed and allowed to vary across plants. Appendix A.3 discusses the main reasons underlying cross-sectional price variation.

$$\min_{\{e_{fit}\}_{f\in\mathcal{F}_{it}}} \left\{ \sum_{f\in\mathcal{F}_{it}} p_{fit}e_{fit} \right\} \quad s.t. \quad \tilde{E}_{it} = \Big(\sum_{f\in\mathcal{F}_{it}} (\psi_{fit}\tilde{e}_{fit})^{\frac{\lambda-1}{\lambda}} \Big)^{\frac{\lambda}{\lambda-1}} \tag{5}$$

[11] I expose the decision of plants under the assumption that plants flexibly rent capital with unit cost of capital r_{kt}. While I use this assumption to reduce the computational burden in the dynamic discrete choice model of fuel sets, I do not need nor use this assumption to estimate the production function.

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

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

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 $MC(\tilde{E}_{it}) = p_{\tilde{E}_{it}}$.

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

Given a cost-minimizing allocation of fuels that produce a quantity of energy, plants pay a price p_{Eit} for each unit of energy. They take this price as given when choosing quantity of 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. Then, 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 with location identifiers and year of production, these form a set of state variables s_{it}. Given these state variables, plants maximize profits which yield a period profit function $\pi(\mathbf{s}_{it}, \mathcal{F}_{it})$.

$$
\begin{aligned}
\pi(\mathbf{s}_{it}, \mathcal{F}_{it}) = &\max_{K_{it}, M_{it}, L_{it}, E_{it}} \Big\{ \Big(\frac{e^{\Gamma_t}}{N_t}\Big)^{\frac{1}{\rho}} P_t^{\frac{1+\rho(\theta-1)}{(\theta-1)\rho}} Y_{it}^{\frac{\rho-1}{\rho}} - w_t L_{it} - r_{kt} K_{it} - p_{mit} M_{it} - p_{Eit} E_{it} \Big\} \\
s.t. \quad \tilde{Y}_{it} = &\, z_{it} \Bigg[\alpha_K \tilde{K}_{it}^{\frac{\sigma-1}{\sigma}} + \alpha_L \tilde{L}_{it}^{\frac{\sigma-1}{\sigma}} + \alpha_M \tilde{M}_{it}^{\frac{\sigma-1}{\sigma}} + \alpha_E \tilde{E}^{\frac{\sigma-1}{\sigma}} \Bigg]^{\frac{\eta\sigma}{\sigma-1}}
\end{aligned} \tag{6}
$$

4.3 Inter-temporal Fuel Set Choices

Inter-temporal fuel set choice

Every period, plants take expectation over the evolution of state variables, and choose a fuel set for next period $\mathcal{F}'$ to maximize expected discounted lifetime profits:

$$V(\mathbf{s}_{it}, \mathcal{F}_{it} \in \mathbb{F}) = \max_{\mathcal{F}'} \Big\{ \underbrace{\pi(\mathbf{s}_{it}, \mathcal{F}_{it})}_{\text{static profits}} - \underbrace{\mathcal{K}(\mathcal{F}' \mid \mathcal{F}_{it}, s_{it}) + \sigma_\epsilon \epsilon_{\mathcal{F}'it}}_{\text{fixed switching costs}} + \underbrace{\beta\, \mathbb{E}[V(\mathbf{s}_{it+1}, \mathcal{F}') \mid s_{it}]}_{\text{continuation value}} \Big\}$$

Where $\mathcal{K}(\mathcal{F}' \mid \mathcal{F}_{it}, s_{it})$ is the net cost of switching from fuel set $\mathcal{F}$ to $\mathcal{F}'$ and $\epsilon_{\mathcal{F}'_{it}}$ capture idiosyncratic shocks to these switching costs. Fuel set switching costs are allowed to vary with some state variables. In particular, I allow these costs to vary by plant size (proxied by hicks-neutral productivity z_{it}) and whether a plant is in a district d that has access to natural gas pipelines:

$$\mathcal{K}(\mathcal{F}' \mid \mathcal{F}_{it}, s_{it}) = k(\mathcal{F}' \mid \mathcal{F}_{it}, d_{it}) + \gamma \ln z_{it}$$

The switching cost function $k(\mathcal{F}' \mid \mathcal{F}_{it}, d_{it})$ is composed of two types of arguments. First, there are fixed costs of adding a fuel κ_f. Second, there are salvage values of dropping a fuel γ_f that plants obtain by selling technologies. Since 90% of plants in the dataset always use electricity and oil, I assume that the choice set of plants is as follow, where e = electricity, o = oil, g = gas, c = coal. Given this fixed costs structure, I show in the next section what motivates plants to switch between fuel sets.

$$\mathbb{F} = \Big\{ (oe); (oge); (oce); (ogce) \Big\}$$

$\mathbb{F}$	**oe**	**oge**	**oce**	**ogce**
oe	0	κ_g	κ_c	$\kappa_g + \kappa_c$
oge	$-\gamma_g$	0	$-\gamma_g + \kappa_c$	κ_c
oce	$-\gamma_c$	$-\gamma_c + \kappa_g$	0	κ_g
ogce	$-\gamma_g - \gamma_c$	$-\gamma_c$	$-\gamma_g$	0

Table 5: $k(\mathcal{F}' \mid \mathcal{F})$

Notes: rows correspond to fuel sets today $\mathcal{F}$, whereas columns correspond to fuel sets next period $\mathcal{F}'$. I assume that fixed costs and salvage values for coal are the same across districts. However, fixed costs and salvage values for natural gas vary by plants' proximity to the natural gas pipeline network in a binary fashion, where I define $d = 0$ if plants have no access to pipelines and $d = 1$ if plants have access to pipelines. Then $\kappa_g = \kappa_{g1}$ if $d = 1$ and $\kappa_g = \kappa_{g0}$ if $d = 0$, and likewise for γ_g. I define plants having access to pipelines if they are located in a district in which a pipeline directly passes, or in a district immediately adjacent to a district in which a pipeline passes.

4.4 Comparative statics: Option Value of Larger Fuel Set

In this section I show why a plant would want to pay a fixed cost to add a new fuel to its set. The price that plants pay for its energy bundle is a CES price index in all fuels available $\mathcal{F}_{it}$:

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

Broda and Weinstein (2006) 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 $p_{\tilde{E}_{it}}$. In Section 3, I showed evidence consistent with this conceptualization of this option value. In Appendix E.1, I show how similar comparative statics can be derived from a task-based model for realized energy similar to Acemoglu and Restrepo (2021). 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,

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

$\square$

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.* $\quad \mathcal{F} \subset \mathcal{F}' \rightarrow \frac{\partial p_{\tilde{E}_{it}(\mathcal{F})}}{\partial p_{fit}} > \frac{\partial p_{\tilde{E}_{it}(\mathcal{F}')}}{\partial p_{fit}}$

$$\frac{\partial p_{E_{it}(\mathcal{F})}}{\partial \tilde{p}_{fit}} - \frac{\partial p_{E_{it}(\mathcal{F}')}}{\partial \tilde{p}_{fit}} = \left(\frac{p_{fit}}{\psi_{fit}}\right)^{-\lambda} \frac{1}{\psi_{fit}} \left[p^{\lambda}_{\tilde{E}_{it}(\mathcal{F})} - p^{\lambda}_{\tilde{E}_{it}(\mathcal{F}')}\right] > 0$$
$$\text{if} \quad \mathcal{F} \subset \mathcal{F}' \text{ since} \quad \lambda > 1 \quad \text{and} \quad p_{\tilde{E}_{it}(\mathcal{F})} > p_{\tilde{E}_{it}(\mathcal{F}')} \text{ by Proposition 1}$$

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 (Allcott et al., 2016; Mahadevan, 2022; Ryan, 2021).

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 $\mathcal{F}$.* $\quad \mathcal{F} \subset \mathcal{F}' \rightarrow p_{\tilde{E}_{it}(\mathcal{F},\overline{e}_f)} > p_{\tilde{E}_{it}(\mathcal{F}',\overline{e}_f)}$

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

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

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

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.

5 Identification of the Production Function

Estimation of the model is done is three steps, each of which rely on different methods, and require solving non-standard challenges. First, I estimate the outer production function in the presence of an unobserved input (energy). Following Grieco et al. (2016), identification relies on the mapping between observed expenditure share of inputs and optimal quantity share of inputs under profit maximization. This method allows me to identify the outer production function and both the price and quantity of energy. Since I observe plant-level output quantity and prices, I also separate the curvature of the production between demand and returns to scale by estimating demand using the Bartik style shift-share instruments of Ganapati et al. (2020), which exploit aggregate fuel price variation as cost shifters.

Second, I jointly identify the inner nest of production, including all fuel-specific productivity terms for fuels that plants are using. To do so, I rely on recent development in production function estimation in the presence of input-augmenting productivity (Demirer, 2020; Zhang, 2019) coupled with dynamic panel methods (Blundell and Bond, 1998, 2000, 2023). More specifically, I exploit plants' optimality conditions, allowing me to infer fuel productivity that would make fuel marginal products equal to observed fuel prices. I then assume a Markovian structure for the evolution of fuel productivity to consistently estimate parameters of the production function using lagged inputs quantities and prices as instruments

Third, while the previous method allows me to recover fuel productivity for fuels that plants are using in a given period, selection bias in this distribution may arise if plants are using their fuel productivity as a basis for fuel set choices. As a result, the distribution of observed fuel productivity may not correspond to the distribution of counterfactual fuel productivity, which is needed to identify switching costs between fuel sets and to perform counterfactual policies. To tackle this selection bias, I allow for systematic difference in fuel productivity across plants and I infer the distribution of counterfactual fuel productivity that would rationalize observed fuel set choices in a full information likelihood, borrowing from the literature on unobserved heterogeneity in dynamic discrete choice models (Arcidiacono and Jones, 2003; Arcidiacono and Miller, 2011).

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, Grieco et al. (2016) show there is a way to uniquely recover the price and quantity of such unobserved input if other flexible inputs are observed and if plants are price-takers in the input market.[12] The key method relies on using relative first-order conditions to map observed expenditure shares to unobserved input quantity shares. To see this, one can look at the ratio of first-order conditions for labor and energy from profit maximization in equation 6, and rearranging:

$$\underbrace{\frac{w_{it}L_{it}}{p_{E_{it}}E_{it}}}_{\text{Expenditure ratio}} = \frac{\alpha_L}{\alpha_E}\Bigg(\underbrace{\frac{L_{it}/\overline{L}}{E_{it}/\overline{E}}}_{\text{Quantity ratio}}\Bigg)^{(\sigma-1)/\sigma} \tag{7}$$

Given production function parameters, $\frac{E_{it}}{\overline{E}}$ can be recovered from (7) 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. Identifica-

[12] The assumption of price-taking in the input market allows for unobserved variation in input prices (the main motivation underlying the Grieco et al. (2016) paper), which could be related to plant size, productivity, location, and any other state variables. However, this assumption rules out quantity discounts.

tion of $\tilde{E}_{it}$ comes from variation in the relative price of labor to energy, which 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.

$$\frac{E_{it}}{\overline{E}} = \Big(\frac{p_{eit}E_{it}}{w_{it}L_{it}}\Big)^{\frac{\sigma}{\sigma-1}}\Big(\frac{\alpha_L}{\alpha_E}\Big)^{\frac{\sigma}{\sigma-1}}\frac{L_{it}}{\overline{L}} \tag{8}$$

In this setting, one can identify production parameters by replacing $\tilde{E}_{it}$ for (8) 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). Following Grieco et al. (2016), I also use the same method to control for unobserved price dispersion in the bundle of material inputs:

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

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). Detailed derivations of the estimating equation can be found in Appendix C.1. Taking logs of revenues yields the main estimating equation:

$$\ln R_{it} = \ln\frac{\rho}{\rho-1} + \ln\frac{1}{\eta} + \ln\left[w_{it}L_{it}\Big(1+\frac{\alpha_K}{\alpha_L}\Big(\frac{K_{it}/\overline{K}}{L_{it}/\overline{L}}\Big)^{\frac{\sigma-1}{\sigma}}\Big) + p_{mit}M_{it} + p_{eit}E_{it}\right] + u_{it} \tag{9}$$

The main parameter of interests are the elasticity of substitution (σ), the elasticity of demand (ρ) and the returns to scale (η) in (9). While the elasticity of substitution is identified from observed variation in the capital to labor ratio, the elasticity of demand/markup is not separately identified from the returns to scale. This is a standard problem with revenue production function, whereby the curvature in the revenue function is driven by both technology (returns to scale) and market power (markup). Fortunately, I observe output prices and quantities, and I have access to exogenous cost shifters, which I use to to recover the elasticity of demand ρ in section 5.1.1. Lastly, since $\tilde{E}_{it}$ and $\tilde{M}_{it}$ were substituted out of the production function, the main estimating equation (9) does not recover α_E and α_M. To recover α_E and α_M, I take the geometric mean of relative first-order conditions in equation (7) for energy and labor, and likewise for materials and labor.[13]

$$\begin{aligned} \overline{wL}/\overline{p_E E} = \frac{\alpha_L}{\alpha_E}; \qquad \overline{wL}/\overline{p_m M} = \frac{\alpha_M}{\alpha_E} \\ \alpha_K + \alpha_L + \alpha_M + \alpha_E = 1 \end{aligned} \tag{10}$$

Then, I estimate (9) subject to (10) with non-linear least squares.[14]

5.1.1 Estimating Elasticity of Demand

To separate the demand elasticity ρ from the returns to scale η in estimating equation (9), I estimate demand from observed output prices and quantities using the demand equation (4).

$$\ln Y_{it} = \Lambda_t - \rho \ln P_{it} + \epsilon_{it} \tag{11}$$

Where $\Lambda_t = \tilde{\Gamma}_t + \ln\left(\frac{1}{N_t}\right) + \frac{\rho(1-\theta)-1}{1-\theta} \ln P_t$, and contains both the unobserved aggregate output price index and aggregate demand shocks. Due to standard simultaneity bias, the elasticity of demand ρ is not identified from price and quantity data alone. To solve this issue, I instrument output prices with a Barktik style shift-share cost shifter proposed by Ganapati et al. (2020) and used by Hawkins-Pierot and Wagner (2022). The instruments have two components: an exogenous shock to aggregate fuel prices (the shift) and pre-shock variation in exposure to aggregate fuel prices by Indian States (the share).

$$z_{s,t} = \left[\overline{p}_{-s,t,f} * \sigma_{s,2008,f}\right], \quad f \in \{\text{coal, gas, oil}\}$$

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

[14]Consistency of the parameters is shown by Grieco et al. (2016) using the first-order conditions of the NLLS objective function as moment conditions.

$\overline{p}_{-s,t,f}$ is the average price (leaving out state s) of fuel f in year t, and acts as an exogenous shock to production cost. This is because much of aggregate fuel price variation stems from worldwide variation in demand and supply induced by geopolitical turmoil, aggregate technological evolution & growth, all of which are unrelated to Indian manufacturing plants. $\sigma_{s,2008,f}$ is the pre-sample aggregate share of fuel f used to generate electricity in state s. Since electricity prices in India are set by state-owned utilities, variation in the price of a fuel is going to induce more variation in electricity prices in states that use more of that fuel to generate electricity. This is going to create exogenous exposure to aggregate fuel price shocks since all plants use electricity as an input. Moreover, the shares are taken in 2008 (before the sample starts), and are thus unaffected by shocks to fuel prices.

For the remaining parts of the demand equation, the aggregate output price index P_t is part of the year fixed effect in equation (11), and is endogenously determined by the elasticity of demand ρ. I first estimate demand using year dummies $\tilde{\Lambda}_t$, and then solve for the price index ex-post, which is a fixed point given the estimate of $\hat{\rho}$ and observe output prices P_{it}. Once I have the output price index, I can recover the elasticity of the outside good θ and the aggregate demand shifter $\tilde{\Gamma}_t$ by exploiting the restrictions on θ.

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 electricity, and redefining the productivity of all other fuels relative to electricity, $\tilde{\psi}_{fit} = \frac{\psi_{fit}}{\psi_{eit}}$:

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

At this point, I have an estimate of the quantity and price of energy, $(\hat{E}_{it}, p_{\hat{E}_{it}})$ from the previous step, fuel quantities, $\{e_{fit}\}_{f\in\mathcal{F}_{it}}$, and fuel prices: $p_{\hat{E}_{it}} = \frac{S_{Eit}}{\hat{E}_{it}}, \{p_{fit} = \frac{s_{fit}}{e_{fit}}\}_{f\in\mathcal{F}_{it}}$. I show how to recover the elasticity of substitution λ, and all productivity terms ψ_{fit}. To do so, I rely on optimality conditions from the energy cost-minimization problem coupled with a Markovian assumption on the productivity of electricity. This is similar to the method proposed by Zhang (2019) and Demirer (2020), but relies on insights from the dynamic panel literature (Blundell and Bond, 2000, 1998) rather than the proxy variable/control variable approach to deal with endogeneity of productivity (transmission bias). As a reminder, the energy cost-minimization problem of the plant is as follows:

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

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

$$\tilde{\psi}_{fit} = \left(\frac{p_{fit}}{p_{eit}}\right)^{\frac{\lambda}{\lambda-1}}\left(\frac{e_{fit}}{e_{eit}}\right)^{\frac{1}{\lambda-1}}\frac{\overline{e}_f}{\overline{e}_e} \tag{13}$$

The intuition underlying equation (13) is that relative fuel productivity equate relative fuel price to relative marginal products $\frac{p_{fit}}{p_{eit}} = \tilde{\psi}_{fit}^{\frac{\lambda-1}{\lambda}}\left(\frac{e_{eit}}{e_{fit}}\right)^{\frac{1}{\lambda}}$. I then exploit these optimality conditions by plugging back the implied relative fuel productivity terms (13) into the energy production function (12) and rearrange:

$$\frac{\tilde{E}_{it}}{\tilde{e}_{eit}} = \psi_{eit}\left(\sum_{f\in\mathcal{F}_{it}}\frac{s_{fit}}{s_{eit}}\right)^{\frac{\lambda}{\lambda-1}} \tag{14}$$

Where $s_{fit} \equiv p_{fit}e_{fit}$ is spending on fuel f. Recalling that $\tilde{e}_{eit} \equiv \frac{e_{eit}}{\overline{e}_e}$ The intuition underlying equation 14 is fairly straightforward. The LHS is the value added of an additional unit of electricity in terms of realized energy. Naturally, plants that are more productive at electricity ($\tilde{e}_{eit}$) will generate more value added. Moreover, the more complement fuels are, the more important are spending on fuels other than electricity to explain the value added of electricity ($\lambda \to 0$ implies that $(\frac{\lambda}{\lambda-1}) \to \infty$). Note that I haven't used the first-order condition (in level) for electricity in the cost-minimization problem. This is because given some quantity of energy E_{it}, one of the input choice is free. Details in Appendix C.2.

At this point, the only unobservable left in the energy production function is the productivity of electricity, which is correlated with current period quantities and spending on fuels since it is assumed to be known to the plant at the time of choosing fuel quantities. This is a standard transmission bias. To deal with this issue, I assume that the productivity of electricity follows an

AR(1) Markov process with year fixed effects.[15] Moreover, I allow for plant-specific fixed effects in the productivity of electricity $\mu_i^{\psi_e}$

$$\ln \psi_{eit} = (1-\rho_{\psi_e})(\mu_0^{\psi_e} + \mu_i^{\psi_e}) + \mu_t^{\psi_e} - \rho_{\psi_e}\mu_{t-1}^{\psi_e} + \rho_{\psi_e} \ln \psi_{eit-1} + \epsilon_{it}^{\psi_e} \tag{15}$$

I then take log of equation (14) and use the Markov process above to get an estimating equation:

$$\ln \hat{\tilde{E}}_{it} - \ln \tilde{e}_{eit} = \Gamma_t + \rho_{\psi_e}(\ln \tilde{E}_{it-1} - \ln \tilde{e}_{eit-1}) + \frac{\lambda}{\lambda-1}\Big(\ln \sum_{f\in\mathcal{F}_{it}} \frac{s_{fit}}{s_{eit}} - \rho_{\psi_e} \ln \sum_{f\in\mathcal{F}_{it-1}} \frac{s_{fit-1}}{s_{eit-1}}\Big) + \mu_i^* + \epsilon_{it}^{\psi_e} \tag{16}$$

Where $\Gamma_t = \mu_0^{\psi_e}(1-\rho_{\psi_e}) + \mu_t^{\psi_e} - \rho_{\psi_e}\mu_{t-1}^{\psi_e}$ is a year fixed-effect and $\mu_i^* = (1-\rho_{\psi_e})\mu_i^{\psi_e}$ is the normalized plant fixed effect. Since $\epsilon_{it}^{\psi_e}$ is a shock to productivity of electricity at time t, it is uncorrelated with choices made at time $t-1$:

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

The estimating equation (16) is very similar to the canonical model in Blundell and Bond (2000) who estimate a Cobb-Douglas production function, and can be written in its canonical form as a linear dynamic panel regression with a common factor restriction ($\beta_2' = -\rho\beta_1'$):

$$y_{it} = \alpha_0 + \alpha_t + \rho y_{it-1} + \beta_1' x_{it} + \beta_2' x_{it-1} + u_i + \epsilon_{it}$$

There are two main endogeneity concerns in this model. First, the lagged value added of energy and the lagged relative spending on other fuels are correlated with the plant fixed effect μ_i^*, which

[15] The choice of these modified AR(1) processes where the mean is normalized by the persistence are standard in the dynamic panel literature with short panels (Blundell and Bond, 2023). It ensures 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. It is equivalent to the assumption that the residuals of the productivity distribution follows an AR(1) process, rather than electricity productivity itself.

biases the persistence of electricity productivity ρ_{ψ_e}. This is the standard concern in the dynamic panel literature. See Bun and Kleibergen (2022) for a review. Second, contemporaneous relative spending on other fuels are correlated with both the fixed effect μ_i^* and the innovation term $\epsilon_{it}^{\psi_e}$ to electricity productivity, which biases the estimate of the elasticity of substitution λ. Blundell and Bond (2000, 1998) and many others show that these concerns can be addressed with properly specified moment conditions conditions. Following Blundell and Bond (1998), I use the system GMM approach which combines both level and difference moment conditions as follows:

$$\mathbb{E}(\Delta X_{i,t-1}(\mu_i^* + \epsilon_{it}^{\psi_e})) = 0$$
$$\mathbb{E}(X_{i,t-1}\Delta\epsilon_{it}^{\psi_e}) = 0$$

For $X_{i,t-1} \in \{\ln \tilde{E}_{i,t-1} - \ln \tilde{e}_{e,i,t-1}, \ln \sum_{f\in\mathcal{F}_{i,t-1}} \frac{s_{fit-1}}{s_{eit-1}}\}$ and likewise for $\Delta X_{i,t-1}$. Moreover, these moment conditions yield a consistent estimate of the elasticity of substitution λ under the assumption that shocks affecting relative fuel spending are persistent. This assumption is consistent with many geopolitical shocks affecting fuel prices that are prevalent in the fuel 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. Lastly, I get standard errors on the elasticity of substitution using the delta method.

6 Identification and Estimation of Fixed Fuel Switching Costs

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

$$V(s_{it}, \epsilon_{it}, \mathcal{F}_{it}) = \max_{\mathcal{F}'\subseteq\mathbb{F}} \Big\{ \pi(s_{it}, \mathcal{F}_{it})/\sigma_\epsilon - \mathcal{K}(\mathcal{F}' \mid \mathcal{F}_{it}, s_{it})/\sigma_\epsilon + \epsilon_{\mathcal{F}'it} + \beta\,\mathbb{E}(V(s_{it+1}, \epsilon_{it+1}, \mathcal{F}') \mid s_{it}) \Big\} \tag{17}$$

Where the fuel set switching cost function, $\mathcal{K}(\mathcal{F}' \mid \mathcal{F}_{it}, s_{it})$, was defined in Table 5. It is a function of productivity z_{it} and access to natural gas pipelines. σ_ϵ is a parameter that maps units of profits (dollars) to units of the shocks to fixed costs.[16] From now on, I define the parameters governing the switching cost function $\theta_1 = \{\kappa_{g1}, \kappa_{g0}, \kappa_c, \gamma_{g1}, \gamma_{g0}, \gamma_c\}$ for coal c and gas g, and θ_2 the parameters underlying the evolution of state variables. I use κ_{g1} to denote the fixed cost of adding natural gas for plants that are located in a district near the pipeline network, and κ_{g0} for plants that are located in a district that isn't immediately adjacent to the pipeline network, and likewise for salvage values. 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, $W(s_{it}, \mathcal{F}_{it}) = \mathbb{E}(V(s_{it}, \epsilon_{it}, \mathcal{F}_{it}))$:

$$
\begin{aligned}
W(s_{it}, \mathcal{F}_{it}) &= \gamma + \ln\left(\sum_{\mathcal{F}' \in \mathbb{F}} \exp\left(\pi(s_{it}, \mathcal{F}_{it})/\sigma_\epsilon - \mathcal{K}(\mathcal{F}' \mid \mathcal{F}_{it}, s_{it})/\sigma_\epsilon + \beta \int W(s_{it+1}, \mathcal{F}') f(s_{it+1} \mid s_{it}) ds_{it+1}\right)\right) \\
&= \gamma + \ln\left(\sum_{\mathcal{F}' \in \mathbb{F}} \exp\left(\upsilon_{\mathcal{F}'}(s_{it}, \mathcal{F}_{it})\right)\right)
\end{aligned}
$$

Where $\gamma \approx 0.5772$ is the Euler–Mascheroni constant and $\upsilon_{\mathcal{F}'}(s_{it}, \mathcal{F}_{it})$ is the choice-specific value function. 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. Below is the main assumption underlying plants' expectation over state variables:

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

Assumption 2. *Plants take expectation over the evolution of all productivity terms, fuel prices and material prices ($\{\psi_{fit}, p_{fit}\}_{f \in \mathbb{F}}, z_{it}, p_{mit}$). However, they are agnostic about the evolution of wages and rental rate of capital (w_t, r_{kt}).*

This assumption for wages and the rental rate of capital is only to reduce computational burden, because the state-space is already extremely large with 12 state variables. See Appendix C.3 for all computational details on the expected value function. I then separate state variables

[16] An equivalent approach would be to map units of the fixed cost shocks to units of profits (dollars) because once σ_ϵ is known, I can always switchback to dollars by multiplying everything by σ_ϵ.

into two categories: non-selected state variables, which I observe for every plant in every year $(\psi_{oit}, \psi_{eit}, p_{ot}, p_{eit}, z_{it}, p_{mit})$ and selected state variables, which I only observe when plants are using the relevant fuel $(\psi_{cit}, \psi_{git}, p_{cit}, p_{git})$. Next, I discuss the evolution of each state variable.

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) fixed effects. $\forall f = \{e, o\}$:[17]

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

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

$$\ln z_{it} = (1-\rho_z)\mu_0^z + \mu_t^z - \rho_z \mu_{zt-1} + \rho_z \ln z_{it-1} + \epsilon_{zit}$$

I allow all shocks to productivity and prices to be arbitrarily correlated in a multivariate normal distribution. For example, 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 plant productivity, thereby increasing demand for electricity, which motivates a joint process for shocks to prices and productivity.

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

[17] I allow for the shock to productivity $\epsilon_{it}^{\psi_f}$ to include a plant fixed effect $\mu_i^{\psi_f}$ when estimating parameters of the productivity transition process to get a consistent estimate of the auto-correlation parameters (Blundell and Bond, 1998). However, I do not currently separate this plant fixed-effect from the shocks $\epsilon_{it}^{\psi_f}$ when simulating the choice probability due computational reasons. Indeed, it would effectively double the state space or would require to solve the model plant-by-plant. Future versions of this paper will allow for these plants fixed-effects in the estimation of fixed costs.

6.2 Evolution of selected state variables – Systematic heterogeneity

Selected state variables refer to price and productivity pertaining to specific fuels that are only observed when a plant uses that fuel. This selection creates a non-trivial challenge to both the estimation of fixed costs and counterfactual policy experiments. Allowing for systematic differences in the productivity of coal and gas across plants, which I call *fuel comparative advantage*, I show that both fuel switching costs and the conditional and unconditional distribution of these comparative advantages can be recovered. The transition for these state variables follows a Markovian process.[18] $\forall f \in \{c, g\}$,

$$\ln \psi_{fit} = \mu_0^{\psi_f} + \mu_t^{\psi_f} + \mu_i^{\psi_f} + \epsilon_{it}^{\psi_f}$$
$$\ln p_{fit} = \mu_0^{p_f} + \mu_t^{p_f} + \epsilon_{it}^{p_f}$$

$\mu_i^{\psi_f}$ is plant's i comparative advantage for coal and gas, which are assumed to come from some distribution. These comparative advantages, which are assumed to be known to the plants when making decisions but unknown to the researcher, underlie a central econometric issue of selection on unobservables. Indeed, recovering the distribution of $\mu_i^{\psi_f}$ only from plants that use coal and/or gas in a given year may not reflect the true distribution of comparative advantage due to selection bias. Not accounting for this selection bias would bias counterfactual fuel choice predictions under alternative policy regimes as well as fixed cost estimates.[19] Lastly, I also allow for shocks to productivity and prices to be correlated. In the next section, I show how to recover the unselected distribution of gas and coal comparative advantages jointly with fixed costs.

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

[18] There is one key distinction between the assumptions underlying the distribution of selected and non-selected state variable, which is due to methodological/computational constraints. For non-selected states, I allow for shocks to be persistent at a decaying rate with an AR(1). Allowing for such decay in the persistence of shocks that affect selected states is currently difficult because it would require to keep track of the hidden Markov process underlying the evolution of these shocks when the state variables are not observed. An alternative specification would be to allow for shocks to states only when plants are using the particular fuels, and to draw initial conditions when they are not. I am currently considering on this approach as well.

[19] I assume that only the distribution of fuel productivity is biased, rather than both prices and productivity. This is for two reasons. First, fuel prices and productivity always enter together as p_{fit}/ψ_{fit} in the static profit function of plants. Second, much of the heterogeneity in coal and gas prices can be explained by observable. For example, much of the historical variation in fuel prices is due to aggregate shocks that affect all establishments such as the oil crash of 2014.

6.3 Identification

To learn about the extent to which the distribution of comparative advantage for natural gas and coal is selected, I follow the algorithm proposed by Arcidiacono and Jones (2003); Arcidiacono and Miller (2011). I assume that the distribution of comparative advantages comes from a finite mixtures with $K = 3$ points of support for each fuel. I parameterize the initial guess of the mean and variance of the finite mixture to the mean and variance of the empirical (selected) distribution $(\tilde{\mu}_f, \tilde{\sigma}^2_{\mu_f})$:

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

Where $\pi^0_{fk} = Pr(\mu_{fk})$ is the unconditional probability of being type k, where types refer to support point of the fuel comparative advantage distribution, and $\sum_k \pi^0_{fk} = 1$. In this context, external estimation of parameters governing the distribution of random effects from a selected sample of plants who use these fuels leads to biased estimates of $\tilde{\mu}_g, \tilde{\mu}_c, \tilde{\sigma}^2_{\mu_g}, \tilde{\sigma}^2_{\mu_c}$. Indeed, plants with larger comparative advantage at 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. Using the law of total probability, I can integrate over the unconditional distribution of comparative advantages using the full information (log) likelihood. Assuming there is only one finite mixture over both coal and gas for notation convenience, and where the distribution of comparative advantages are independent across fuels such that $\pi_k \in \Pi = vec(\Pi_g \otimes \Pi_c)$, where $\pi_{kg} \in \Pi_g$ and $\pi_{kc} \in \Pi_c$:

$$\ln \mathcal{L}(\mathcal{F}, s \mid \theta_1, \theta_2) = \sum_{i=1}^{n} \ln \left[\sum_k \pi_k \left[\prod_{t=1}^{T} Pr(\mathcal{F}_{it+1} \mid \mathcal{F}_{it}, s_{it}, \mu_i = \mu_k; \theta_1, \theta_2) \right] \right] + \sum_{i=1}^{n} \sum_{t=1}^{T} \ln f(s_{it} \mid s_{it-1}; \theta_2) \tag{18}$$

In particular, the likelihood in (18) assumes that the state transitions are independent of the distribution of comparative advantages for coal and gas.[20] This is possible if the parameter estimates $\hat{\theta}_2$ are unbiased from selected data. In Appendix C.4, I show Monte-Carlo simulation results that are consistent with this assumption. Initially, the true probability weights π_k over the support of the finite mixture are unknown due to selection, but Arcidiacono and Jones (2003); Arcidiacono and Miller (2011) provide a method to recover the unselected distribution by sequentially iterating

[20]This assumption isn't necessary but it simplify computation of the model in the presence of these comparative advantages.

over the fixed costs to maximize the likelihood and updating the probability weights $\pi_k^0, \pi_k^1, \pi_k^2, ...$ using an EM algorithm.[21] 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 plant i is of type k given all observables, including choices made:

$$\begin{aligned}\hat{\theta}_1 &= \arg\max_{\theta_1,\theta_2,\pi} \sum_{i=1}^{n} \ln \left[\sum_k \pi_k \left[\prod_{t=1}^{T} Pr(\mathcal{F}_{it+1} \mid \mathcal{F}_{it} s_{it}, \mu_i = \mu_k; \theta_1, \theta_2) \right]\right] \\ &\equiv \arg\max_{\theta_1} \sum_{i=1}^{N} \sum_{t=1}^{T} \sum_k \rho(\mu_k \mid \mathcal{F}_i, s_i; \hat{\theta}_1, \hat{\theta}_2, \hat{\pi}) \ln Pr(\mathcal{F}_{it+1} \mid \mathcal{F}_{it}, s_{it}, \mu_i = \mu_k; \theta_1, \hat{\theta}_2)\end{aligned}$$

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

$$\rho(\mu_k \mid \mathcal{F}_i, s_i; \theta_1, \theta_2, \hat{\pi}) = \frac{\hat{\pi}_k \left[\prod_{t=1}^{T} \left[\prod_{\mathcal{F} \subseteq \mathbb{F}} \left[Pr(\mathcal{F}_{it} \mid s_{it}, \mu_i = \mu_k; \theta_1, \theta_2) \right]^{\mathbb{I}(\mathcal{F}_{it} = \mathcal{F})} \right]\right]}{\sum_k \hat{\pi}_k \left[\prod_{t=1}^{T} \left[\prod_{\mathcal{F} \subseteq \mathbb{F}} \left[Pr(\mathcal{F}_{it} \mid s_{it}, \mu_i = \mu_k; \theta_1, \theta_2) \right]^{\mathbb{I}(\mathcal{F}_{it} = \mathcal{F})} \right]\right]} \tag{19}$$

The idea underlying the EM algorithm is to iteratively estimate fixed costs parameters θ_1 given some guess of the distribution of comparative advantages $\{\pi_k\}_k$ – M step, draw new comparative advantages using Baye's law from (19), which are used to then update the unconditional distribution of comparative – E step, and repeat this procedure until the likelihood in (18) is minimized. Details of the algorithm can be found in Appendix C.5.

7 Application to Steel manufacturing

I then apply this production model to Steel manufacturing. I do so for several compelling reasons. First, steel production is one of the most environmentally damaging industries in India, with coal accounting for nearly 70% of its energy sources. See figure 21 in Appendix A.5.1. Second, compared to other heavy manufacturing industries like Cement, Steel manufacturing is highly competitive.

[21] For now I assume the support of the finite mixture is known. In future versions, I will also allow the support points to vary.

Steel plants significantly outnumber similar industries in India, and face intense international competition, particularly from the Chinese Steel industry. Moreover, I show in Appendix A.5.7 that steel plants produce a number of differentiated varieties of steel, such as alloy and non-alloy steel ingots, billets, blooms, bars and rods. Overall, there are 404 unique varieties produced by Steel plants between 2009 and 2016, where varieties are defined from the Indian National Product Classification for Manufacturing Sector (NPCMS, 2011). High level of competition with a large number of varieties makes the assumption of monopolistic competition compelling in this context.

From a policy perspective, studying steel manufacturing is particularly interesting because a majority of Steel plants are situated in Eastern and Central India, particularly in regions that form the "steel belt", where most iron ore and coal mines are located. See Figure 17. However, there is limited access to natural gas pipelines in these regions. See Figure 3. Considering these significant geographical features, exploring policies that subsidize fixed costs to complement carbon taxation may be relevant in order to generate more substitution from coal to natural gas. I first show estimation results for Steel manufacturing, then discuss policy counterfactuals.

7.0.1 Outer Production function estimation results

Preliminary estimates of the production function parameters can be found in Table 6. I also report the average output and revenue elasticities with respect to each input to be consistent with the literature (Gandhi et al., 2020).

Table 6: Production Function Estimation (Steel Manufacturing)

Production and Demand Parameters		Average Output Elasticities		Average Revenue Elasticities
Elasticity of substitution $\hat{\sigma}$	1.80 [1.374,3.054]	Labor	0.040 [0.037,0.048]	0.030 [0.029,0.030]
Returns to scale $\hat{\eta}$	1.23 [1.137,1.444]	Capital	0.023 [0.014,0.035]	0.017 [0.010,0.025]
Elasticity of demand $\hat{\rho}$	3.84 [2.695,4.914]	Materials	1.008 [0.935,1.185]	0.745 [0.741,0.749]
Elasticity of outside good $\hat{\theta}$	0.63 [0.552,0.683]	Energy	0.155 [0.144,0.182]	0.115 [0.113,0.117]
Observations	8,547			

Bootstrap 95% confidence interval in bracket (499 reps)

Notes: the average output (revenue) elasticities are defined as the average of the individual output (revenue) elasticity, where the output elasticity is $\frac{\partial y_{it}}{\partial x_{jit}} \frac{x_{jit}}{y_{it}}$ for $y_{it} \in \{Y_{it}, R_{it}\}$ and $x_{jit} \in \{L_{it}, K_{it}, M_{it}, E_{it}\}$

65

The average output and revenue elasticities with respect to intermediate materials is much larger than other inputs, which is consistent with the vast majority of manufacturing production function estimation in the literature (Gandhi et al., 2020; Grieco et al., 2016; Doraszelski and Jaumandreu, 2013, 2018).[22] In Steel manufacturing, this is because iron ore is the primary raw material that goes into steel production. Moreover, average output and revenue elasticities are considerably larger for energy than labor and capital, which is expected in such an energy-intensive industry. The estimated demand elasticity is also consistent with estimates by Zhang (2019) who finds a demand elasticity of around 4 in the Chinese Steel industry. Using these estimates I can construct estimates of the price $p_{\hat{E}_{it}}$ and quantity of the energy bundle for each plant $\hat{E}_{it}$ from the relation first-order conditions in equation 8, which I use to provide some evidence in favor of the energy production function.

Relationship between Estimated Price of Energy and Number of Fuels

To motivate the energy production function in the Indian steel context, 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 7: Relationship between $\ln p_{\hat{E}_{it}}$ and the number of fuels available to plants.

	(1)	(2)	(3)	(4)
	$\ln p_{\hat{E}_{it}}$	$\ln p_{\hat{E}_{it}}$	$\ln p_{\hat{E}_{it}}$	$\ln p_{\hat{E}_{it}}$
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.

[22]Note that the average output elasticity is greater than 1. This is because it only captures a technological constraints. However, increasing say materials by 1% will not lead to an increase in output by 1% in practice because of downward sloping demand. In this context, the revenue elasticity is more appropriate because it captures both technological and demand constraints.

Moreover, the evidence provided in Table 7 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 show in the next section that selection channels, particularly fuel productivity, dominate the option value.

7.0.2 Energy production function estimation results

Table 8: Estimates of Energy Production Function

	Steel
Elasticity of substitution $\hat{\lambda}$	2.224***
	(0.231)
Persistence of electricity productivity $\hat{\rho}_{\psi_e}$	0.649***
	(0.118)
Observations	3,459

Standard errors in parentheses

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

Notes: Since the estimating equation (16) is a linear reduced-form IV, I get an estimate of $\hat{\gamma} = \frac{\hat{\lambda}}{\hat{\lambda}-1}$ with corresponding standard errors $\hat{\sigma}_\gamma$. I use the delta method to recover the standard error of $\hat{\lambda}$ where $\hat{\sigma}_\lambda = (\hat{\lambda}-1)\hat{\sigma}_\gamma$. Moreover, The number of observations in the energy production function (3,459) is lower than in the outer production function (8,547). This is because the method to estimate the energy production function constructs moments that require at least 3 years of observation per plant to yield consistent estimates (Blundell and Bond, 2000, 1998).

Turning to the energy production function, results suggest that the elasticity of substitution across fuels $\hat{\lambda}$ is at least as large as the elasticity of substitution across labor, capital, materials and energy $\hat{\sigma}$ from Table 6.[23] This is important because the larger is the elasticity of substitution between fuels, the larger are the aggregate gains from carbon taxation (Acemoglu, Aghion, Bursztyn and Hemous, 2012). Indeed, more substitution possibilities means that more emission reduction can be achieved by substituting away from polluting fuels rather than by reducing output, which is a key trade-off when evaluating carbon policy. Next, I can construct estimates of the fuel-specific productivity for each plant in each year $\hat{\psi}_{fit}$ and I find large and persistent heterogeneity in the distribution of fuel productivity.

Heterogeneity in fuel productivity across fuel sets

[23] While the point estimate for $\hat{\lambda}$ is 25% larger than $\hat{\sigma}$, the 95% confidence interval is much larger for $\hat{\sigma}$, and includes both the lower bound and upper bound of the 95% confidence interval for $\hat{\lambda}$. Hence, I make the conservative claim that fuels are at least as substitutable than other inputs

In Table 7, I provided evidence of systematic differences in the marginal cost of energy across different fuel sets, which can be driven by a combination of the option value of having more fuels and fuel prices/productivity heterogeneity. Using estimates of fuel-specific productivity, I do In Table 9 a full Shapley decomposition of 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.

		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	Contribution of Total Difference (%)	62.6	97.75	94.84
Fuel Prices		1.25	-3.18	-1.14

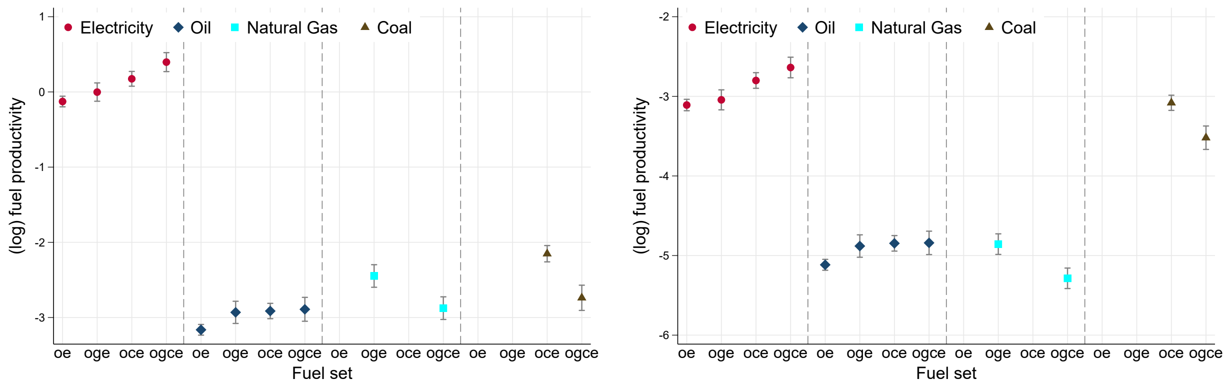
Table 9: Shapley Decomposition of the Difference in Average Marginal Cost of Energy Between Fuel Sets

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).

I find that fuel productivity explains the majority of observed differences in the price of energy (between 60 and 97%). This is consistent with a productivity-efficiency argument in which more productive plants have more to gain from paying the fixed costs, thus select into larger fuel sets, similar to the argument made in markets with fixed costs of entry (Melitz, 2003; Hopenhayn, 1992). General details of the Shapley decomposition can be found in Appendix D.1 and specific details to this particular decomposition can be found in Appendix D.1.2. To understand this result better, I provide some evidence of systematic differences in fuel productivity across plants and across fuel sets.

Results are summarized in Table 4. There are few takeaways. First, electricity is by far the most productive fuel, both in terms of physical quantity and dollar invested. On average, one mmBtu of electricity is three times more productive than one mmBtu of oil, while one dollar of electricity is two times more productive than one dollar oil. Additionally, while 1 mmBtu of coal is on average 30 % more productive than 1 mmBtu of gas, this gap widely expands when looking at productivity per dollar. Indeed, one dollar invested in coal is on average 1.85 times more productive than one dollar invested in natural gas. This is because coal is significantly cheaper than other fuels – averaging one fifth of the price of natural gas.

Second, when looking at electricity and oil, there is a positive gradient between the number of fuels and productivity, consistent with the decomposition of Table 7. However, I find that more productive plants at using coal tend to specialize in coal, and likewise for gas, whereas plants who use both coal and gas tend to be less productive at either fuel. This specialization mechanism can explain why, for example, plants who have four fuels tend to have slightly higher energy prices after controlling for all controlling for year fixed effects, fuel prices and hicks-neutral productivity in Table 7.



(a) Fuel productivity per mmBtu – $\ln(\psi_{fit}/\overline{e}_f)$ (b) Fuel productivity per dollar – $\ln(\psi_{fit}/(\overline{e}_f * p_{fit}))$

Figure 4: Mean and 95% CI for (log) fuel productivity, by fuel set

Notes: The figure is created by taking the sample average of the estimated log productivity for all four fuels, by fuel sets. Fuel set labels are created as follows: oe = oil and electricity, oge = oil,gas, and electricity. oce = oil, coal, and electricity. ogce = oil, gas, coal, and electricity. The reason I divide by the geometric mean of fuel quantities $\overline{e}_f$ is because fuel productivity are originally in normalized units due to the normalization in estimation. There is no observation for gas and coal productivity for fuel sets that exclude these fuels.

Third, I investigate how important is the heterogeneity in fuel productivity across plants. To do so in compelling way, I compute the marginal return to one mmBtu of each fuel in dollars of revenue, keeping every other input constant, I find significant heterogeneity. While plants at the bottom of the distribution get between \$0.5 to \$2 in revenue per mmBtu, these returns can go up to \$1,000 per mmBtu, which is considerably larger than the unit price of the most expensive fuel.

Percentile	Natural Gas	Coal	Oil	Electricity
10^{th}	2.2	.5	.7	2.2
25^{th}	7.1	1.5	3.4	8.5
50^{th}	24.7	5.0	16.8	43.72
75^{th}	86.5	16.8	72.8	209.0
90^{th}	283.4	53.5	246.7	927.17

Table 10: Distribution of *Ceteris-Paribus* marginal returns (\$) to one mmBtu of each fuel

Notes: The returns is computed by taking the derivative of the revenue function with respect to physical quantities of fuels: $\frac{\partial R_{it}}{\partial e_{fit}}$, which is evaluated at the observed quantities of all inputs.

In summary, the previous evidence suggests that fuel productivity is highly heterogeneous, and can explain the majority of differences in marginal costs across fuel sets. In a context where plants with different fuel sets systematically differ in how efficiently they use fuels, technology lock-in may be important. This argument is reinforced by the Indian Ministry of Steel's own statements on energy management which reports that inefficient plants face difficulties in transitioning out of old technologies:

> "*The higher rate of energy consumption is mainly due to obsolete technologies including problems in retrofitting modern technologies in old plants, old shop floor & operating practices*" Indian Ministry of Steel (2023)

Moreover, technology lock-in will be exacerbated if fixed costs of fuel adoption are very large, which I document in the next section.

7.0.3 Estimation of Fixed Costs and Selection Bias in Fuel Productivity

Fixed costs are report in Table 11. The estimates of fixed costs encompass both the tangible expenses related to new fuel-burning technologies, and intangible costs associated with fuel adoption. This includes logistical challenges, new contractual agreements for transportation and storage, as well as potential opportunity costs from diverting labor away from production. On average, these costs are substantial, ranging from 28 to 40 million dollars, and align well with upper echelon of existing accounting estimates.[24]

		Fixed Costs (Million USD)	**Salvage Values (Million USD)**
Natural Gas	*Pipeline Access*	28.83	9.02
	No Pipeline Access	40.46	17.21
	Coal	28.82	8.33
Total Factor Productivity (100 % Increase)		0.82	0.25
Observations		2,393	

Table 11: Estimates of Fuel Set Fixed Costs and Salvage Values

Notes: this table shows the fixed cost and salvage value estimates for each fuel in million U.S dollars. For natural gas, these costs vary based on whether plants are in a district with access to a natural gas pipeline. The parameter in front of "Total Factor Productivity" is the effect of doubling productivity on the fixed costs and salvage values of any fuel, and is meant to capture how these costs vary with plant size. The sample size is lower than the energy production function because I remove the last year of observation since I don't observe subsequent fuel set choices.

In the context of this paper, raising productivity by 1% leads to an $\$8,200$ increase in the fixed cost of adopting any fuel and a $\$2,500$ increase in the salvage value of dropping an existing fuel. Additionally, coal adoption tends to be 30% cheaper than gas, consistent with coal-based methods

[24]While recent comprehensive accounting estimates of switching costs are hard to find, a single electric arc furnace may cost between a few hundred thousand dollars and a few million dollars (Source: alibaba's listings `https://www.alibaba.com/product-detail/WONDERY-Custom-Made-Siemens-PLC-Industrial_1600732474634.html`), whereas switching from pig iron, typically produced with coal-powered blast furnace, to direct reduced iron, typically produced with gas-powered oxygen furnaces would historically cost upwards of USD 70 millions Miller (1976)

being more outdated. Moreover, plants without access to high pressure natural gas pipelines incur 40% higher costs of adoption due to the need for alternative transportation methods, such as liquefied natural gas (LNG) which can be costly to gasify. This effect of pipeline accessibility is consistent with findings from Scott (2021) in their study of U.S. Power Plants. The observed salvage values for both coal and natural gas are significantly lower, ranging from 57% to 71% below the fixed costs. These salvage values are in line with depreciating capital over time. Importantly, the combination of substantial fixed costs and relatively low salvage values likely contributes to situations of technology lock-in.

To further understand factors that prevent plants from transitioning out of old technologies, I revisit the distribution of fuel productivity by taking into account selection bias. I find significant evidence of selection bias for both coal and gas. Indeed, plants who do not use gas would be 30% less productive at using gas than plants who do, whereas this effect goes up to 80% for coal. This evidence of selection bias is also likely to exacerbate technology lock-in and undermine how much switching between fuel sets we might except from policy.

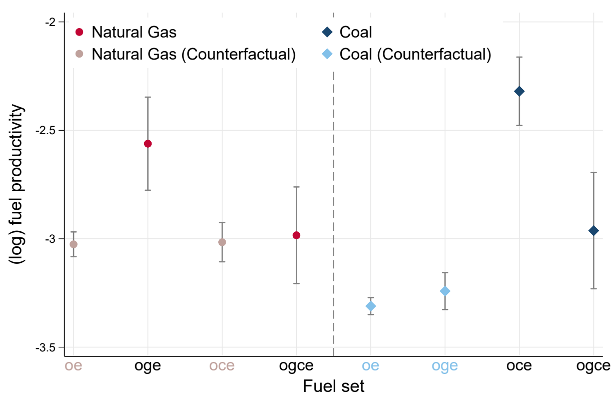


Figure 5: Distribution of fuel productivity – Including counterfactual fuel sets

Notes: This figure shows the distribution of fuel productivity per mmBtu ($\ln \psi_{fit}/\overline{e}_f$) with 95% confidence intervals for coal and natural gas, and includes counterfactual productivity for plants with fuel sets that exclude gas and/or coal. The distribution of fuel productivity for counterfactual fuel sets was computed by simulating draws from the estimated distribution of unobserved heterogeneity (*comparative advantages*) in the dynamic discrete choice model, using the conditional probability distribution $\rho(\mu_k \mid \mathcal{F}_i, s_i; \hat{\theta}_1, \hat{\theta}_2, \hat{\pi})$.

Model fit

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. The model does worse at predicting the transitions for plants that start with all four fuels because it only represents 8% of the sample. In all figures below, the organe bars (Data) are constructed as follows:

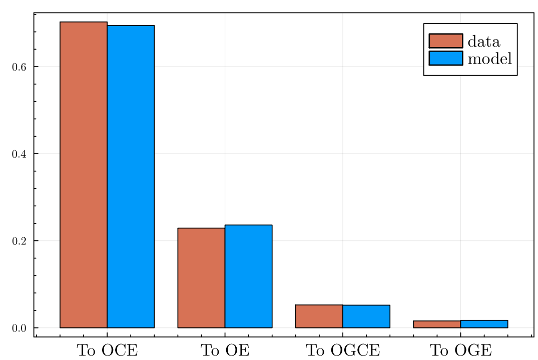
$$\mathbb{P}_{\mathcal{F}'}(data) = \frac{1}{NT}\sum_i \sum_t \mathbb{I}(\mathcal{F}_{it+1} = \mathcal{F}'), \qquad \mathbb{P}_{\mathcal{F}'|\mathcal{F}}(data) = \frac{1}{N_{\mathcal{F}}T}\sum_i \sum_t \mathbb{I}(\mathcal{F}_{it+1} = \mathcal{F}' \mid \mathcal{F}_{it} = \mathcal{F})$$

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

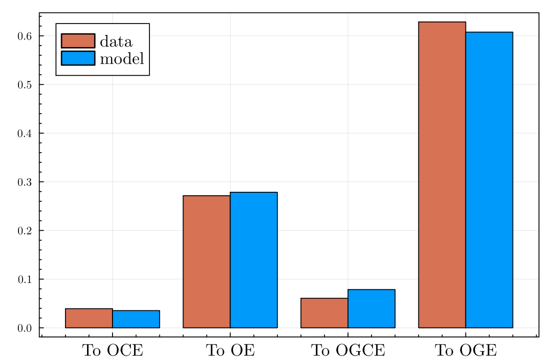
$$\mathbb{P}_{\mathcal{F}'}(model) = \frac{1}{NT}\sum_i \sum_t \sum_k \rho(\mu_{fk} \mid \mathcal{F}_i, s_i; \hat{\theta}_1, \hat{\theta}_2, \hat{\pi}) Pr(\mathcal{F}' \mid \mathcal{F}_{it}, s_{it}, \mu_{fi} = \mu_{fk}; \hat{\theta}_1, \hat{\theta}_2)$$

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

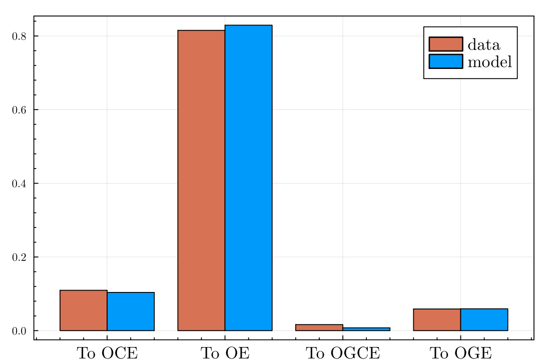
Graphs using conditional probability of comparative advantage



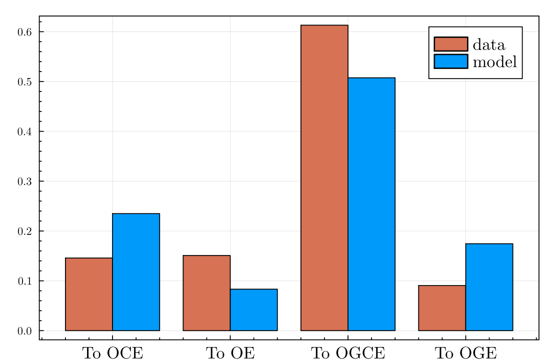
From Oil,Coal,Electricity (OCE) − $N = 572$



From Oil,Gas,Electricity (OGE) − $N = 280$



From Oil,Electricity (OE) − $N = 1,342$



From Oil,Gas,Coal,Electricity (OGCE) − $N = 199$

Figure 6: Conditional distribution of fuel sets (transition), model vs. data

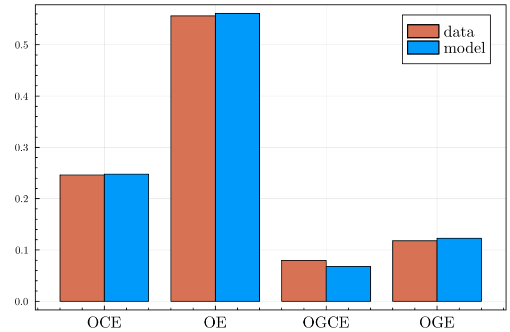


Figure 7: Unconditional distribution of fuel sets, model vs. data ($N = 2,393$)

8 Externality Mitigation - Steel Manufacturing

In this section, I study the effectiveness of various policies at mitigating externality damages from fuel combustion to improve social welfare. I detail how externality damages are constructed, and I perform two counterfactual policy experiments. First, I quantify the trade-off between emission reduction and output for various levels of a carbon tax levied on fossil fuels. I then compare this trade-off with an economy where plants cannot switch between fuel sets and an economy without heterogeneity in fuel productivity. I find that under the no-switching restriction, the aggregate loss of output required for any reduction in emission slightly larger than in the unrestricted economy. In contrast, I find that under the restriction that imposes no heterogeneity in fuel productivity, the aggregate loss of output required for any reduction in emission is much larger. Overall, accounting for both those channels jointly improves how the economy responds to a carbon tax, by reducing aggregate costs of emission reduction.

Second, I allow the government to use carbon tax revenues to finance a subsidy that reduces the fixed cost of adopting natural gas, in order to alleviate technology lock-in. I find as the subsidy rate increases (in percentage of the fixed cost), externality damages slightly increases, because more and more plants add natural gas, but don't drop coal. This additional fuel provides them with an option value, which reduces marginal costs and improves both consumer and producer surplus, at the expense of increased emission damages and total investment costs. I then investigate the welfare implications of this trade-off with a permanent 10% subsidy. Despite the subsidy being nominally very large, the welfare effects are positive but quite small relative to the economy with a carbon tax that rebates its proceeds as lump-sum transfers to consumers.

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.4. 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: $\tilde{\gamma}_f = SCC * \gamma_f$. As an example, a conservative social cost of carbon for India would be \$5.74 USD per short ton of CO_{2e} (Tol, 2019), and would imply the following carbon tax, where the tax rate on each fuel is equal to marginal externality damages $\tau_f = \tilde{\gamma}_f \ \forall f \in \{o, g, c, e\}$.

Fuel Prices (rupee/mmBtu)			
	No Tax	**Carbon Tax**	**% Change**
Coal	262	308	17.5
Oil	665	701	5.4
Elec	1,681	1,715	2
Gas	1,307	1,331	1.8

Table 12: Example of Average Fuel Prices With and Without Carbon Tax

Notes: These prices are averaged across all sample periods. Coal is by far the most polluting fuel, so the average price change of coal is an increase of 17.5%. Interestingly, since 50% of Indian electricity is generated with coal, natural gas is slightly less polluting than electricity, making it the cleanest of the four fuels.

8.1 Carbon Tax and the Trade-off between Output and Emission Reduction

In figure 8, I trace the trade-off between output and emission reduction for various carbon tax rates. Each point on the curve corresponds to a different level of the carbon tax, and together they form a production frontier in output and emission reduction. In particular I simulate the economy with and without the carbon tax for 40 years and look at the net present value (NPV) of outcomes along the entire path. If the level of the tax approaches zero, then the model converge to the no-tax or *laissez-faire* economy with 100% of output and 0% of emission reduction. As the level of the tax increases, the aggregate output decreases but so does emission reduction.

The production frontier is concave because at the aggregate level, there is an increasing marginal cost of reducing emissions, consistent with previous findings by Fowlie et al. (2016). This is because fuel substitution (and more generally input substitution) is a low hanging fruit, where much of emission reduction can first be achieved by substituting away from coal to cleaner fuels such as natural gas and electricity. However, as the carbon tax rate increases, more emission reduction comes at the cost of plants scaling down their operation which decreases aggregate output because marginal plants already substituted towards cleaner fuels.

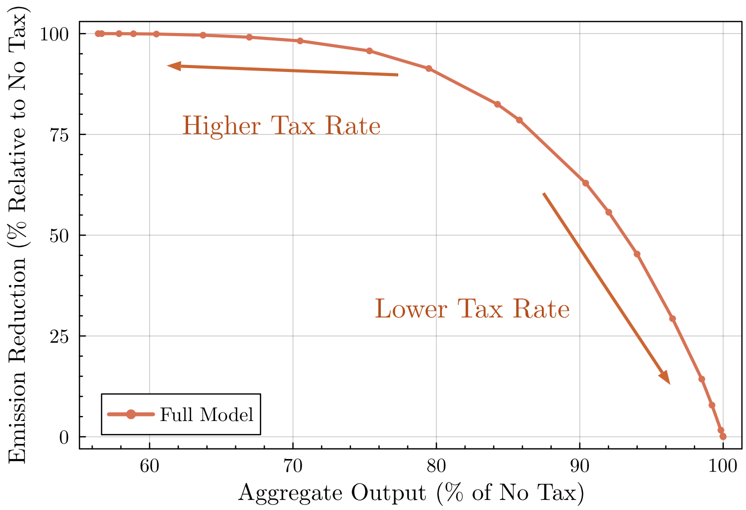


Figure 8: Production Frontier in Output and Emission Reduction for Various Carbon Tax Rates

Notes: This production frontier was constructed by simulating the economy under 21 different levels of the carbon tax, ranging from 0 (no tax) to approximately infinity. Linear interpolation is assumed for the trade-off in-between each tax level. As the level of the tax approaches infinity, aggregate output does not reach 0. This is a feature of the CES production function. Indeed, as fuel prices are extremely high, fuel consumption approaches zero but plants always use some positive amount of fuels.

The role of inter-temporal switching between fuel sets

I then compare in figure 9 what happens when I don't allow for inter-temporal switching between fuel sets and don't allow for heterogeneity in fuel productivity in the economy to highlight their role in this trade-off. First, removing the ability of plants to pay fixed costs and switch between fuel sets removes a substitution channel. As a result, one may expect that more emission reduction comes at the cost of plants scaling down, such that any desired reduction in emissions will result in lower aggregate output. However, the opposite is happening. This is because the carbon tax leads to a net decrease in the fraction of plants that use coal and gas. Natural gas is significantly more expensive than coal, and has an average fixed costs of adoption twice as large. As a result, plants are more likely to salvage their coal and/or gas technologies rather than replace coal with gas. The net effect is an increase in the fixed portion of their profit through salvage values, but a decreases in variable profits. Indeed, as plants salvage expensive fuels, they lose an option value which increases their marginal cost through a higher price of energy $p_{E_{it}}$. In Appendix D.2.1, I confirm this intuition by showing that as the level of the tax raise, the average fraction of plants that use coal and gas decreases, and the average price of energy increases relative to the economy in which plants are not able to switch between fuel sets. Overall, this leads to an increase in the marginal costs which increases output prices and decreases aggregate output for any level of emission reduction.

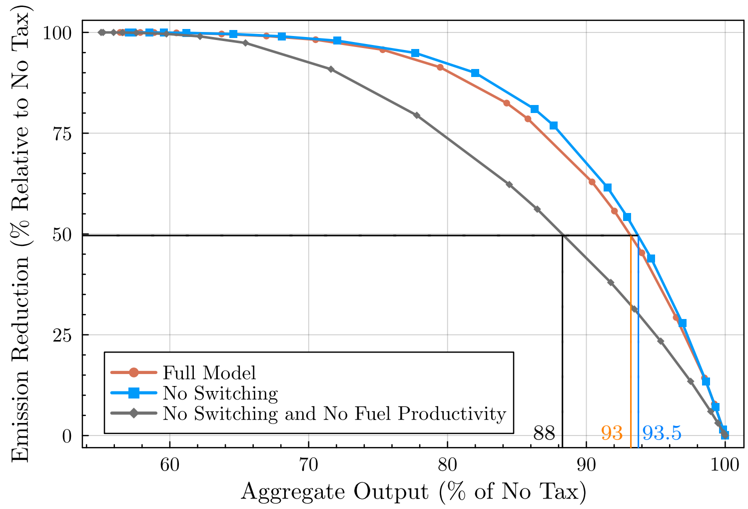


Figure 9: Comparison of Production Frontier Across Model Specification

However, while qualitatively valid, this aggregate phenomena is not quantitatively large. For example, to reduce emissions by 50%, the economy originally produces 93% of the *laissez-faire* output in the full model, with an implied elasticity between emission reduction and output of 7.14.[25] This is in contrast to the economy without fuel set switching which can produce 93.5% of its *laissez-faire* output for a 50% decrease in emissions, with an implied elasticity of 7.7. This small difference can be partially attributed to the inability of carbon taxes at incentivizing much inter-temporal switching between fuel sets. As as shown in Figure 27, it would take a carbon that raises the price of coal by 400% to incentivize a 10% decrease in plants who use coal a 2% decrease in plants who use natural gas.

The role of heterogeneity in fuel productivity

Second, I do the same exercise while removing heterogeneity in fuel productivity on top of removing fuel set switching. To do so, I re-estimate an energy production function in which plants have heterogeneous energy productivity ψ_{Eit}, but have the same average fuel productivity. captured by β_f. Details on estimation of this production function are in Appendix D.2.2, and follow the dynamic panel approach, similarly to the energy production function with fuel-augmenting productivity. Crucially, estimating this production function matches average levels of fuel quantities and aggregate levels of emissions, but misses the heterogeneity in fuel shares across plants.

[25]This elasticity between emission reduction and output goes down to around 2 as the level of the tax approaches infinity.

$$E_{it} = \psi_{Eit} \left(\sum_{f \in \mathcal{F}_{it}} \beta_f e_{fit}^{\frac{\lambda-1}{\lambda}} \right)^{\frac{\lambda}{\lambda-1}}$$

The green production frontier in figure 9 corresponds to this economy, and shrinks inwards even more than in the model without fuel set switching. For example, to reduce emissions by 50%, the economy operates at 80% of *laissez-faire* output, with an implied elasticity of 2.5. Allowing for heterogeneity in fuel productivity diminishes how much output must decrease to achieve any reduction in emissions because it increases the aggregate elasticity of substitution between fuels. Intuitively, even though the individual elasticity of substitution is the same across plants, the aggregate elasticity of substitution is larger because output reallocates from high emission to low emission plants. There are two channels that explain this reallocation.

First, conditional on fuel prices and fuel set, the elasticity of the cost energy with respect to relative fuel prices isn't constant across plants. For example, as the price of coal increases relative to the price of gas, plants who are more productive at using coal relative to gas face a larger percentage increase in their cost of energy. This is because larger coal productivity induces specialization in coal, making them more exposed to the relative price increase, as long as fuels are gross substitutes $\lambda > 1$. Since the carbon tax is effectively an increase in the relative price of polluting fuels, plants who are more productive at using polluting fuels and initially use more of those fuels face a larger increase in their marginal costs. To see this, let $\tilde{p}_{fit} = p_{fit}/p_{git}$ be the price of fuel f relative to gas and likewise for relative fuel productivity $\tilde{\psi}_{fit} = \psi_{fit}/\psi_{git}$. Then,

$$\frac{\partial \ln p_{E_{it}}}{\partial \ln \tilde{p}_{cit}} = \frac{\left(\tilde{p}_{cit}/\tilde{\psi}_{cit}\right)^{1-\lambda}}{\sum_{f \in \mathcal{F}_{it}} \left(\tilde{p}_{fit}/\tilde{\psi}_{fit}\right)^{1-\lambda}} = \frac{p_{cit} e_{cit}}{\sum_{f \in \mathcal{F}_{it}} p_{fit} e_{fit}} \tag{20}$$

The elasticity of the marginal of energy with respect to relative price of any fuel (here coal relative to gas) is just the plant-specific spending share of that fuel relative to all fuels. Details in Appendix D.2.3. Moreover, this elasticity is increasing in relative fuel productivity when $\lambda > 1$, which is the key result here. This means, conditional on fuel prices and fuel set, plants who are more productive at using coal spend more on coal, and are more sensitive to relative changes in the price of coal:

$$\frac{\partial^2 \ln p_{E_{it}}}{\partial \ln \tilde{p}_{cit} \partial \tilde{\psi}_{cit}} = \frac{(\lambda-1)\psi_{cit}^{\lambda-2}\tilde{p}_{cit}^{1-\lambda}\Big[\sum_{f\in\mathcal{F}_{it}\backslash c}\big(\tilde{p}_{fit}/\tilde{\psi}_{fit}\big)\Big]}{\big(\sum_{f\in\mathcal{F}_{it}}(\tilde{p}_{fit}/\tilde{\psi}_{fit})^{1-\lambda}\big)^2} > 0 \qquad \text{if} \quad \lambda > 1$$

In contrast, in the model without fuel-specific productivity, the elasticity of the price of energy with respect to relative fuel prices is constant up to fuel prices and fuel set. Second, this heterogeneous increase in marginal costs of energy makes polluting plants less competitive relative to cleaner plants, consistent with the aggregation result by Oberfield and Raval (2021). The tax thus induces a reallocation of output from more polluting to less polluting plants, which increases aggregate fuel substitution and diminishes how much aggregate output must be reduced to achieve any emission reduction target[26]. Moreover, this reallocation channel is function of the elasticity of demand and the returns to scale. Indeed, as the elasticity of demand increases and different output varieties become more substitutable, any variation in relative marginal costs across plants will lead to larger reallocation of output. In Appendix D.2.5, I confirm this intuition by showing that the difference between both production frontier expands as the elasticity of demand increases. In summary, not allowing for this rich heterogeneity between fuel productivity shuts down the reallocation of output from high emission to low emission plants, and decreases the effectiveness of a carbon tax.

To benchmark this result with the literature, I do two exercises. First, I compare the aggregate trade-off between output and emissions with Fowlie et al. (2016) who conduct similar policy exercises for U.S. cement plants. Crucially, their margin of interest is plant entry/exit and dynamic investments in output capacity. However, they do not allow for input substitution. I show in Table 14 that a version of my model without input substitution yields an average elasticity between emission reduction and output more than half as large as in the full model, and closer to Fowlie et al. (2016).[27] Comparing this to the more flexible economy in which plants can substitute at both margin in Figure 13 sheds light on the important role that input substitution plays in mitigating the loss of output for any emission reduction target.

Second, I show that heterogeneity in fuel productivity and inter-temporal switching between fuel

[26]Note that the correlation between fuel productivity and total factor productivity (TFP) also matters for this result. Indeed, if plants who are more affected by the carbon tax were also initially more productive overall, this reallocation effect may reduce aggregate TFP and aggregate output. However, I show in Appendix D.2.4 that the opposite is true. Plants with higher fuel productivity tend to be less productive overall.

[27]Note that a gap still remains and the production frontier is still concave without input substitution at the plant level. This is for two reasons. First, even without input substitution, plants are differently affected by the tax based on their fuel sets, which affects aggregate input substitution due to the reallocation of output across plants. Second, the difference between the two elasticities is also attributed to the entry/exit margin in Fowlie et al. (2016), which decreases both output and emissions through plants exiting in the aftermath of carbon policy.

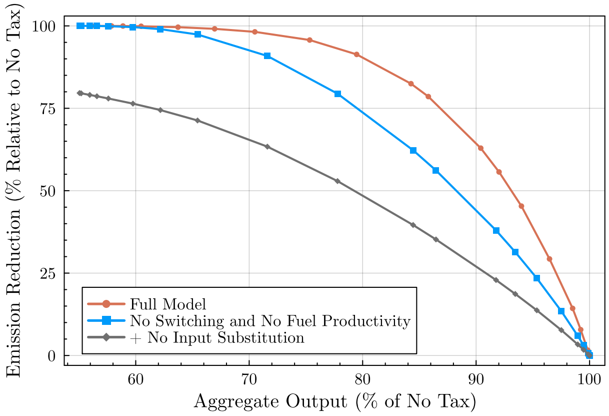


Table 13: Comparison of Trade-off Including No Input Substitution

	Average Elasticity $\frac{\%\Delta CO_{2e}}{\%\Delta Y}$
Full Model	5.46
No Switching	5.77
No Switching and No Fuel Producitivty	3.89
No Input Substitution	2.48
Fowlie et al. (2016) – U.S Cement	1.04

Table 14: Comparison of Average Elasticity

Notes: The average elasticity U.S. Cement plants is constructed by approximating Figure 2 A (aggregate output capacity) and C (aggregate emissions) in Fowlie et al. (2016). They do various carbon policy exercises across different carbon price. I specifically approximate their *Auctioning* policy which is isomorphic to a carbon tax.

sets also serve as a cautionary tale for larger-scale climate models in which an aggregate production function in different fossil fuels is typically assumed as part of a larger integrated assessment model (IAM). Such models are used to study the relationship between climate change and economic growth. For example, Golosov et al. (2014) rely on an aggregate CES production for a composite energy which combines oil, gas and coal. In Appendix D.2.6, I show that such an aggregate CES production function can be micro-founded from an economy in which many plants have an energy production function without heterogeneity in fuel productivity, and where plants have access to different fuel sets but cannot switch between them. Given the results of this section, the use of such a production function may understate the extent of fuel substitution as a response of policy because it does not capture the underlying heterogeneity in fuel productivity, and the ability of plants to inter-temporally switch between fuel sets.

Alleviating Technology Lock-in - Combining Carbon tax with Fixed Cost Subsidy

To complement the carbon tax, I now investigate how proceeds from the tax can be used to finance subsidies to the fixed cost of natural gas. The underlying motivation is that fixed costs are found to be economically large, which can lead to situations of technology lock-in in which some plants cannot substitute away from polluting fuels such as coal. Particularly, the vast majority natural gas pipeline infrastructure in India is on the West Coast. See figure 15. Meanwhile, a large fraction of Steel plants that form the "Steel Belt" are located in Central and Eastern India (Odisha, West Bengal, Jharkhand, and Chhattishgarh) nearby India's largest reserves of coal and iron ore. See figure 17. Plants in these regions are far away from the natural gas pipeline network, and face a fixed cost of gas adoption 50% larger (see Table 11), with gas prices averaging 5 times the price of coal. These plants could potentially benefit from fixed cost subsidies to incentivize natural gas adoption, which can take many forms.[28]

Below I show results for various permanent subsidy, ranging from 0% to 100% of the average fixed cost of natural gas. I do these experiments jointly with a carbon tax. To choose a representative social cost of carbon (SCC), I first set the social discount rate to 3% ($\beta = 0.7$) to match the average real interest rate in India during the sampled period. Then, following the most recent estimates from the Inter-agency Working Group on the Social Cost of Carbon (IWG, 2021), I set the SCC to the 2020 estimates for a social discount rate of 3% at USD $51/tCO_{2e}$. This SCC corresponds to a mid-range estimate in the literature.[29] in Figure 10, I show what happens to carbon tax revenues/externality damages and total subsidy paid out as the subsidy rate increases

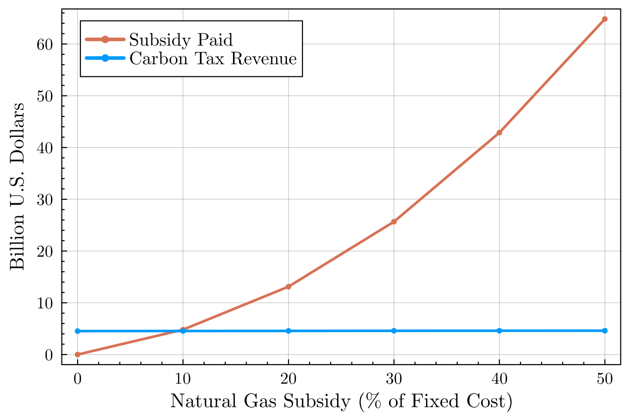


Figure 10: Carbon tax revenue and subsidy paid out along subsidy rate

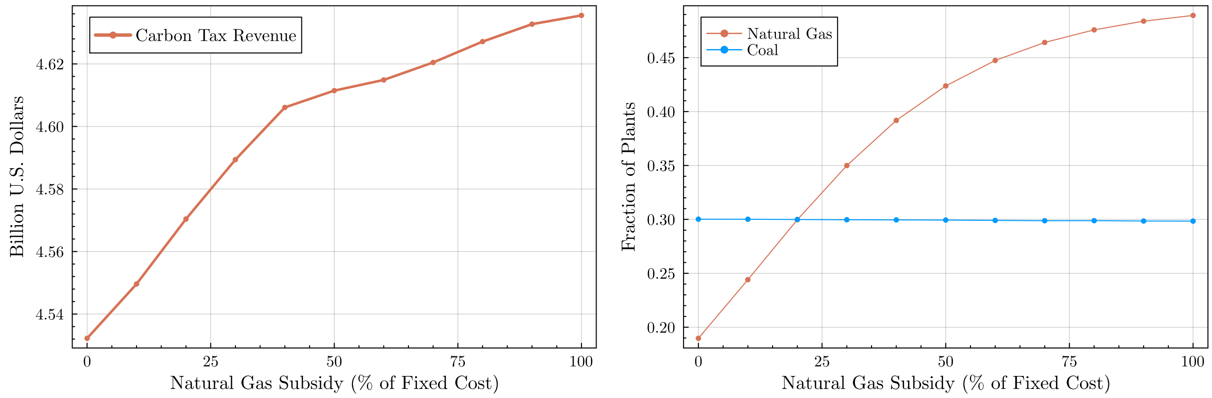
Notes: this figure was calculated by simulating the expected total tax revenues and subsidy paid out to plants for an horizon of 40 years. Note that carbon tax revenues corresponds to externality damages since the carbon tax rates equates marginal externality damages. As such, evolution of the carbon tax revenues is indicative of emissions.

First, up to 10% the subsidy can be fully financed by a carbon tax in expectation. More importantly, carbon tax revenues monotonically increase as the subsidy rate increases. I show this more clearly in Figure 11, where I compare the evolution of the tax revenues with the fraction of plants who use natural gas and coal.

These results are important. Indeed, as the subsidy rate increases, more plants add natural gas, but there is almost no change in the fraction of plants who use coal. This isn't surprising because coal is still significantly cheaper than gas, and the salvage values of coal are much lower than gas. In this context, it makes more sense for plants to keep coal for the option value it provides. In

[28]For example, it can be done through infrastructure subsidies to expand the natural gas pipeline network or the development of liquefied natural gas (LNG) in Eastern India. However, such projects often involve large investments that go beyond the scope of this paper. For analytical tractability, I will assume there exists some technology that allows the government to directly subsidize plants' fixed costs by utilizing carbon tax revenues.

[29]I also experiment with other social cost of carbon, ranging from USD $5.74/tCO_{2e}$ (Tol, 2019) to higher end estimates of USD $196/tCO_{2e}$ by the IPCC (IWG, 2021). All results are with different carbon prices are qualitatively similar but have different quantitative implications.



(a) Carbon Tax Revenues/Externality Damages (b) Fraction of Plants who use Natural Gas/Coal

Figure 11: Comparison of selected outcomes along subsidy rate

Notes: carbon tax revenues corresponds to externality damages since the carbon tax rates equates marginal externality damages. As such, evolution of the carbon tax revenues is indicative of aggregate emissions.

the aggregate, these results create two countervailing effects on carbon tax revenues/externality damages. On one hand, plants who add natural gas substitute away from more polluting fuels such as oil, electricity and coal. This substitution effect reduces tax revenues. On the other hand, as plants add natural gas, they have more fuels which increases their option value and decreases the price of energy. The net effect is a decrease in marginal costs of production and an increases in output, which increases all input demand. Thus, even with substitution towards natural gas, the input demand for other fuels such as coal, oil and electricity goes up, which increases tax revenues. In Figure 13, I do a full Shapley decomposition of the change in tax revenues as the subsidy rate increases between these two channels. Details of the Shapley decomposition can be found in Appendix **??**.

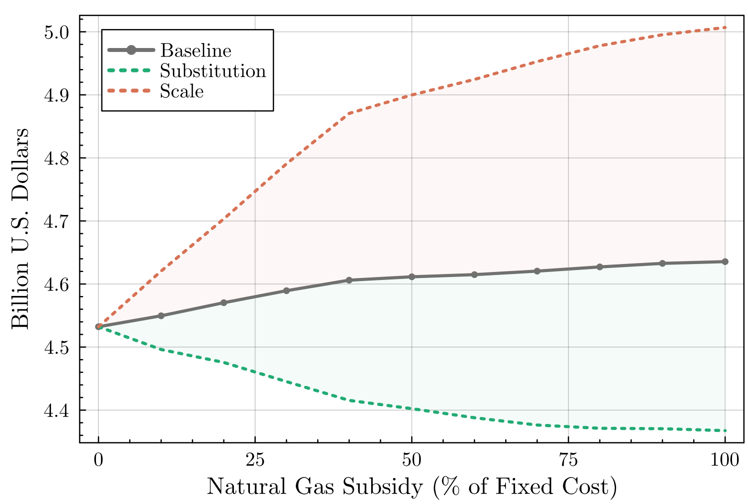


Figure 13: Shapley Decomposition of Changes in Tax Revenues

Unsurprisingly, I find that the scale effect dominates. From a welfare perspective, it is unclear whether the subsidy is preferable to an economy with only a carbon tax because while both profits and consumer surplus increase, this comes at the cost of more externality damages and considerable investment subsidies which could be allocated towards more profitable ventures. For this reason, I do a formal welfare analysis of this policy in the next section.

Welfare Analysis of a 10% Subsidy

I choose to narrow the focus on a 10% subsidy because it can be financed by a carbon tax and internally satisfies the government's budget constraint. With such a policy, per-period welfare is standard and features four components: consumer surplus, producer surplus, net government revenues and externality damages (Fowlie et al., 2016):

$$w_t(\tau, s) = \underbrace{\nu_t(\tau, s)}_{\text{consumer surplus}} + \underbrace{\Pi(\tau, s)}_{\text{producer surplus}} + \underbrace{G(\tau, s)}_{\text{net gov. revenue}} - \underbrace{\sum_f \sum_i \gamma_f e_{fit}(\tau, s)}_{\text{externality damages}}$$

Where consumer surplus is the indirect utility function, which is decreasing in the aggregate output price index P_t. This is due to quasi-linear aggregate utility: $\nu_t(\tau, s) = \frac{\theta}{1-\theta} P_t(\tau, s)^{-\frac{\theta}{1-\theta}}$. As such, we can think of the remaining three parts of this welfare function as shifting aggregate income of the consumers if it owns all plants and gets aggregate profits net of fixed costs, government revenues as lump-sum transfers and suffer externality damages from pollution in dollars from the social cost of carbon. To include a fixed-cost subsidy towards natural gas adoption, I make some simplifying assumptions for tractability. I assume that the subsidy is financed by government revenue from the carbon tax, and that every plant faces the same permanent subsidy amount s. In this context, producer surplus is defined as the sum of total profits net of subsidized fixed costs, and net government revenue is total tax revenues minus subsidy paid out.

$$\Pi(\tau, s) = \sum_{i=1}^{N} \Big(\underbrace{\pi_{it}(\tau, s)}_{\text{variable profits}} - \sum_{\mathcal{F}' \subseteq \mathbb{F}} \big[\underbrace{\mathcal{K}(\mathcal{F}' \mid \mathcal{F}_{it}) + s\mathbf{I}(gas \in \mathcal{F}' \setminus \mathcal{F}_{it}) \big] \mathbf{I}(\mathcal{F}_{it+1} = \mathcal{F}' \mid \tau, s)}_{\text{subsidized fixed costs}} \Big)$$

$$G(\tau, s) = \sum_{i=1}^{N} \Big(\underbrace{\sum_f \tau_f e_{fit}(\tau, s)}_{\text{tax revenue}} - \underbrace{s\mathbf{I}(gas \in \mathcal{F}_{it+1} \setminus \mathcal{F}_{it})}_{\text{subsidy}} \Big)$$

Note that externality damages cancel out with tax revenue, and the subsidy cancels out because it is a transfer from $G(\tau, s)$ to $\Pi(\tau, s)$. As a result, period welfare is effectively equal to consumer surplus plus variable profits minus total fixed costs:

$$w_t(\tau, s) = \underbrace{\nu_t(\tau, s)}_{\text{consumer surplus}} + \underbrace{\sum_{i=1}^{N} \pi_{it}(\tau, s)}_{\text{variable profits}} - \underbrace{\sum_{i}^{N} \Big(\sum_{\mathcal{F}' \subseteq \mathbb{F}} \mathcal{K}(\mathcal{F}' \mid \mathcal{F}_{it}) \mathbf{I}(\mathcal{F}_{it+1} = \mathcal{F}' \mid \tau, s) \Big)}_{\text{total fixed costs}} \tag{21}$$

Total welfare is then defined as the net present value of expected period welfare. I approximate total welfare by averaging multiple Monte-Carlo simulations of the economy (indexed by k) over an horizon of 40 years. Lastly, the subsidy rate s was chosen such that expected net government revenues is weakly positive.

$$\begin{aligned} \mathcal{W}(\tau, s) &= \mathbb{E}_0 \Big(\sum_{t=0}^{\infty} \beta^t w_t(\tau, s) \Big) \qquad\qquad \mathbb{E}_0(G(\tau, s)) \geq 0 \\ &\approx \frac{1}{K} \sum_{k} \sum_{t=0}^{40} \beta^t \omega_{tk}(\tau, s) \end{aligned}$$

Below are the welfare results. In net, there is a small but positive welfare effect from the subsidy, which means that using carbon tax revenues to subsidize the adoption of natural gas is slightly better than rebating it as a lump sum transfer to consumers. This welfare effect is explained by two countervailing effects. On one hand, variable profits and consumer surplus increase by 19 and 13 million dollars, respectively. This is because more plants add natural gas, but the fraction of plants using coal remains constant. This leads to a decrease in the average price of energy, a decrease in average marginal costs and a decrease in output prices. Thus, more steel is produced at a lower cost which benefits producers, and sold at lower price which benefits consumers. On the other hand, there is an increase in externality damages and an increase in total fixed costs paid in the economy by 10 and 31 million dollars, respectively. While externality damages cancel out with tax revenue, total fixed costs do not.

To understand how small the welfare effects really are, I compare in Table 16 how much of the total fixed costs are financed by the subsidy. While 2.7 billion dollars go towards the adoption of natural gas, the fraction of plants who use natural gas only goes up by 20% from 0.18 to 0.24, while

	Carbon Tax *Billion U.S Dollars*	Carbon Tax + 10% Subsidy *Billion U.S Dollars*	Difference *Million U.S. Dollars*
Total Welfare	63.415	63.417	1.18
Variable Profit	22.58	22.60	19.18
Consumer Surplus	22.21	22.22	13.62
Total Fixed Costs (*Paid by plants + subsidy*)	-18.633	-18.601	31.61
Externality Damages/Tax Revenue	2.64	2.65	9.78

Table 15: Decomposition of Welfare Effects – Carbon Tax with and without Subsidy

Notes: All welfare components are reported by their net present value (NPV) over an horizon of 40 years from the last year of observation in the data (2016) with a social discount rate of 3%. Also, externality damages and tax revenue exactly cancel out in the welfare function. the subsidy also cancels out because it is simply a transfer from net government revenue towards producer surplus. As a result, variable profits, consumer surplus and total fixed costs are the remaining components in the welfare function such that $Welfare = ConsumerSurplus + VariableProfit - TotalFixedCost$

variable profits and consumer surplus jointly increase by 31 million dollars. Hence, private gains from the subsidy are only 1.1% of the policy's cost.

	Carbon Tax	Carbon Tax + 10% Subsidy	Difference
Fraction of Gas Users	0.19	0.24	0.05
Total Subsidy paid *(Billion U.S. Dollars)*	0.0	2.793	2.793

Table 16: Total Subsidy paid

Notes: This table reports the long-run fraction of plants who use natural gas after the policy, and the net present value of expected total subsidy paid to plants.

There are a few reasons explaining this small effect. First, by virtue of being a universal subsidy, the government effectively finances the adoption of natural gas for plants who would have still adopted natural gas in the absence of the subsidy. This can be seen from the increase in total fixed costs by 31.61 million dollars in Table 15 after the policy, which is considerably lower than the subsidy's cost of 2.7 billion. Second, plants at the margin who are actually incentivized to adopt natural gas in the aftermath of the policy are on average 30% less productive at using natural gas than plants who already use natural gas. See Figure 5. At the same time, natural gas is on average less productive per dollar invested into it than any other fuel. See Figure 4b This raises the question of whether the government could find more profitable avenues to invests proceeds from the carbon tax revenue. For example, it could invest in energy efficiency training programs to increase energy and fuel productivity, or carbon capture technologies that reduce emissions ex-post. While outside the scope of this paper, I believe this is an interesting avenue for future research.

9 Conclusion

In this paper, I develop a rich dynamic production model to capture salient features of fuel consumption in manufacturing. This includes the prevalence of switching between fuel sets and heterogeneity in fuel productivity. By combining various methods from the production function estimation and the dynamic discrete choice literature, I show how this model can be estimated with a panel of plant-level data that features both output and input prices/quantities. I then apply this model to the Indian Steel industry, which is high in energy and emission intensity due to the prevalence of coal usage, among other factors. From a normative standpoint, I perform various counterfactual policy experiments aimed at reducing emissions in this industry, which include a carbon tax and a carbon tax with a subsidy towards the adoption of cleaner fuels.

I show that novel features of this model have important quantitative implications for the scope of these policies. Indeed, carbon taxation is much more targeted towards high emission plants than previously thought due to the presence of multiple layers of heterogeneity. As a result, high emission plants become relatively less competitive, which reallocates output towards low emission plants. This considerably reduces the overall economic cost of reducing emissions. However, a carbon tax alone does not lead to more adoption of cleaner fuels such as natural gas. For this reasons, I show how proceeds from the carbon tax can be used to subsidize the fixed cost of natural gas adoption. Doing so, there is a small but positive effect welfare effect, unexpectedly through larger private surplus (producer and consumer) at the expense of higher emissions. This is due to the option value that an additional provides, which lowers production costs. Overall, these results highlight the importance of producer heterogeneity and inter-temporal decisions when quantifying the impact of carbon policy.

References

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

_ **, Philippe Aghion, Leonardo Bursztyn, and David Hemous**, "The Environment and Directed Technological Change," *American Economic Review*, 2012, *102* (1), 131–166.

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 Miller**, "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.

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

Blundell, Richard and Stephen Bond, "Initial conditions and moment restrictions in dynamic panel data models," *Journal of Econometrics*, 1998, *87*, 115–143.

_ **and** _ , "GMM Estimation with persistent panel data: an application to production functions," *Econometric Reviews*, 2000, *19* (3), 321–340.

_ **and** _ , "Initial conditions and Blundell-Bond estimators," *Journal of Econometrics*, 2023, *234*, 101–110.

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.

Bollard, Albert, Peter J. Klenow, and Gunjan Sharma, "Indias mysterious manufacturing miracle," *Review of Economic Dynamics*, 2013, *16*, 59–85.

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

Bun, Maurice J.G. and Frank Kleibergen, "Identification Robust Inference for Moment-based Analysis of Linear Dynamic Panel Data Models," *Econometric Theory*, 2022, *38*, 689–751.

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.

Cho, Won G., Kiseok Nam, and José A. Pagan, "Economic growth and interfactor/interfuel substitution in Korea," *Energy Economics*, 2004, *26*, 31–50.

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.

Indian Ministry of Steel

Indian Ministry of Steel, "Energy and Environmet Management in Iron & Steel Sector," 2023. Retrieved from https://steel.gov.in/en/technicalwing/energy-and-environment-management-iron-steel-sector.

Interagency Working Group on Social Cost of Greenhouse Gases

Interagency Working Group on Social Cost of Greenhouse Gases, United States Government, "Technical Support Document: Social Cost of Carbon, Methane, and Nitrous Oxide; Interim Estimates under Executive Order 13990," Technical Report February 2021.

Government of India Ministry of Statistics and Programme Implementation

Government of India Ministry of Statistics and Programme Implementation, "National Product Classification for Manufacturing Sector (Revised)," 2011. Retrieved from https://www.mospi.gov.in/sites/default/files/main_menu/national_product_classification/NPC-MS_21sep11.pdf.

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.

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.

Ganapati, Sharat, Joseph S. Shapiro, and Reed Walker, "Energy Cost Pass-Through in US Manufacturing: Estimates and Implications for Carbon Taxes," *American Economic Journal: Applied Economics*, 2020, *12* (2), 303–342.

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.

Hassler, John, Conny Olovsson, and Michael Reiter, "Integrated Assessment in a Multi-region World with Multiple Energy Sources and Endogenous Technical Change," 2019.

— , **Per Krusell, Conny Olovsson, and Michael Reiter**, "On the effectiveness of climate policies," 2020, pp. 1–43.

Hawkins-Pierot, Jonathan T. and Katherine R. H. Wagner, "Technology Lock-In and Optimal Carbon Pricing," 2022.

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

Hopenhayn, Hugo A., "Entry, Exit, and firm Dynamics in Long Run Equilibrium," *Econometrica*, 1992, *60* (5), 1127–1150.

Hyland, Marie and Stefanie Haller, "Firm-level Estimates of Fuel Substitution," *The Energy Journal*, 2018, *39* (6), 71–98.

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.

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

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.

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.

Lyubich, Eva, Joseph Shapiro, and Reed Walker, "Regulating Mismeasured Pollution: Implications of Firm Heterogeneity for Environmental Policy," *AEA Papers and Proceedings*, 2018, *108*, 136–142.

Ma, Hengyun, Les Oxley, John Gibson, and Bonggeun Kim, "China's energy economy: Technical change, factor demand and interfactor/interfuel substitution," *Energy Economics*, 2008, *30*, 2167–2183.

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.

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

Melitz, Marc J., "The impact of trade on intra-industry reallocations and aggregate industry productivity," *Econometrica*, 2003, *71* (6), 1695–1725.

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

Miller, Jack Robert, "The Direct Reduction of Iron Ore," *Scientific American*, 1976, *235* (1), 68–81.

Nordhaus, William, *A Question of Balance: Weighing the Options of Global Warming Policies*, Vol. 22, New Haven London: Yale University Press, 2008.

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

_ **and Devesh Raval**, "Micro Data and Macro Technology," *Econometrica*, 2021, *89* (2), 703–732.

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

Pindyck, Robert S., "Interfuel Substitution and the Industrial Demand for Energy," *The Review of Economics and Statistics*, 1979, *61* (2), 169–179.

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.

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.

Stern, David I, "INTERFUEL SUBSTITUTION: A META-ANALYSIS," *Journal of Economic Surveys*, apr 2012, *26* (2), 307–331.

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

Tol, Richard S.J., "A social cost of carbon for (almost) every country," *Energy Economics*, 2019, *83*, 555–566.

Wang, Xiaolei and Boqiang Lin, "Factor and fuel substitution in China's iron & steel industry: Evidence and policy implications," *Journal and Cleaner Production*, 2017, *141*, 751–759.

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

In the ASI, Manufacturing plants are surveyed either as part of a census or as part of a sample. All plants who qualify for the census are required to fill the survey by the Government of India's Central Statistics Office. The remaining plants are surveyed based on stratified sampling rules. The definition of census vs. sample and the sampling rules went through some changes over the years. In 2008, all plants with more than 100 workers and multi-plant firms, as well as plants in the lesser industrialized states (Manipur, Meghalaya, Nagaland, Tripura, Sikkim and Andaman Nicobar Islands) were part of the census. For the remaining plants, strata were constructed by state/industry pairs and 20% of plants were sampled within each stratum.

By 2016, the rules for a plant to be considered in the census expanded. Plants in the following states with more than 75 workers were part of the census: Jammu Kashmir, Himachal Pradesh, Rajasthan, Bihar, Chhattisgarh and Kerala. Plants in the following states with more than 50 workers were part of the census: Chandigarh, Delhi and Puducherry. Plants in the seven less industrialized states where part of the census: Arunachal Pradesh, Manipur, Meghalaya, Nagaland, Sikkim, Tripura and Andaman Nicobar Islands. Lastly, plants with more than 100 workers in all other states were part of the census.

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 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 (Worrell, Bernstein, Roy, Price and Harnisch, 2009).

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 (Galitsky and 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 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 14 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.

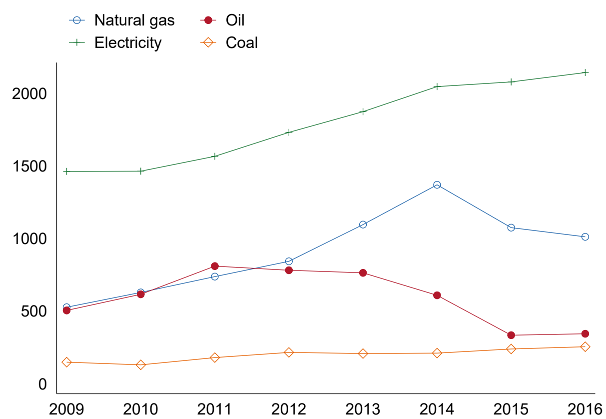


Figure 14: Yearly Median Fuel Prices (INR/mmBtu)

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[30]. 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:

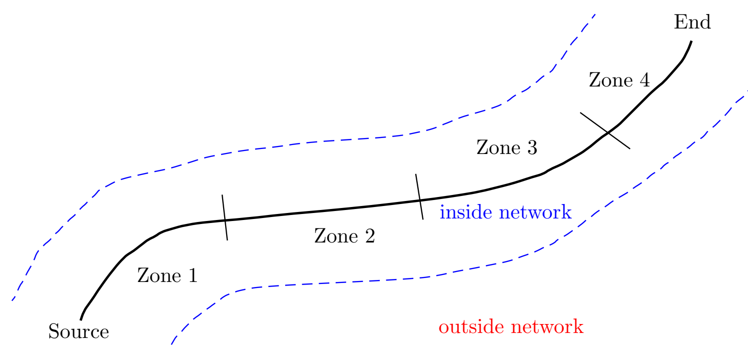


Figure 15: Hypothetical Structure of Transportation Costs for Natural gas Pipeline

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[31]. 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 16 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

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

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

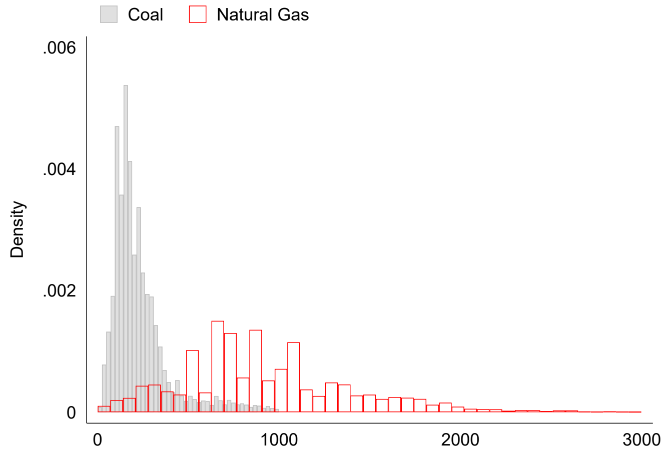


Figure 16: Histogram comparing price of natural gas and coal (INR/mmBtu)

gas pipeline network to the districts in which they pass, directly or indirectly. Thus these results are subject to measurement error.

Table 17: 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 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"

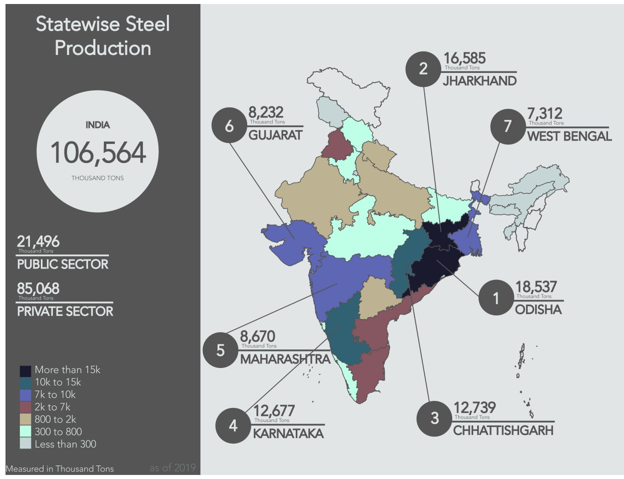


Figure 17: Concentration of steel plants by regions

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

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.

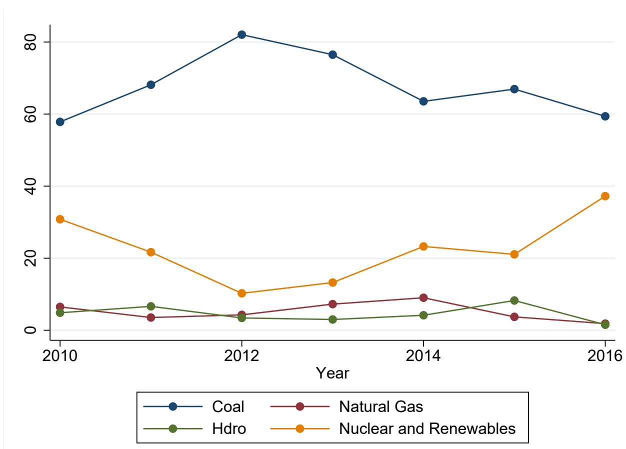


Figure 18: Annual Indian Electricity Generation by Source (% of Total)
Source: International Energy Agency (IEA)

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

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

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.

		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 18: Emission factors from fuels to carbon dioxide equivalent $\zeta_{fkj} * GWP_k$ (kg CO_2e/mmBtu). Source: (Gupta et al., 2019, Annexure 3)

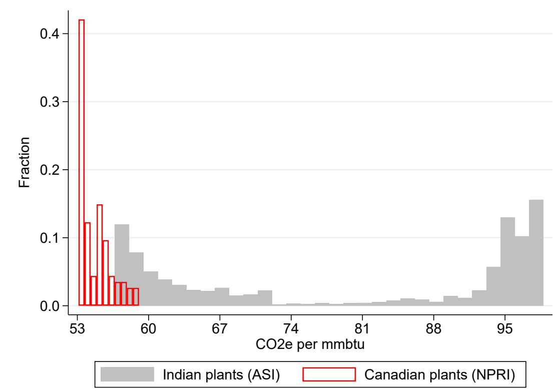
Share of Electricity Generated by Source				
Natural Gas	Coal	Hydro	Other	Emission factor (kg CO_2e/mmBtu)
0.052	0.68	0.046	0.23	72.05

Table 19: Emission factors from Electricity

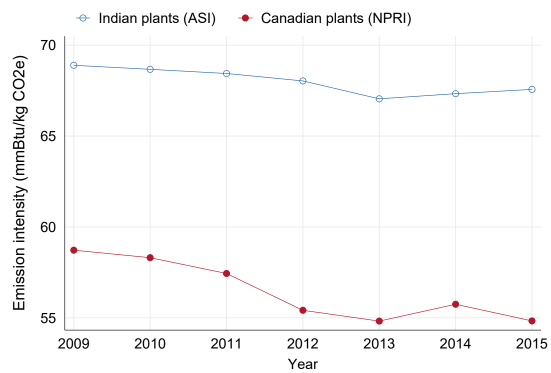
A.5 Further Evidence

A.5.1 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 19a, 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 19b).



(a) Pollution Intensity - Cement Manufaturing

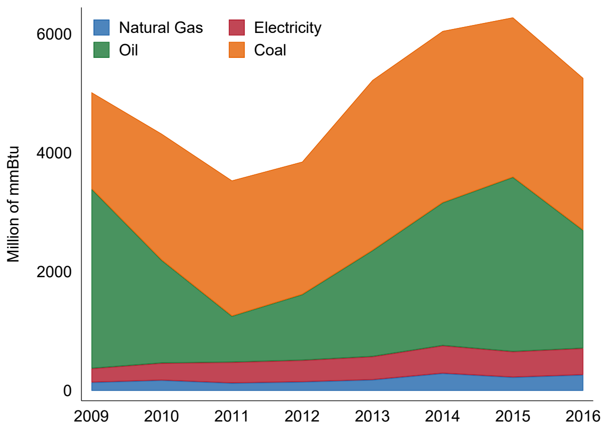


(b) Evolution of Pollution Intensity - Heavy Manufacturing Industries

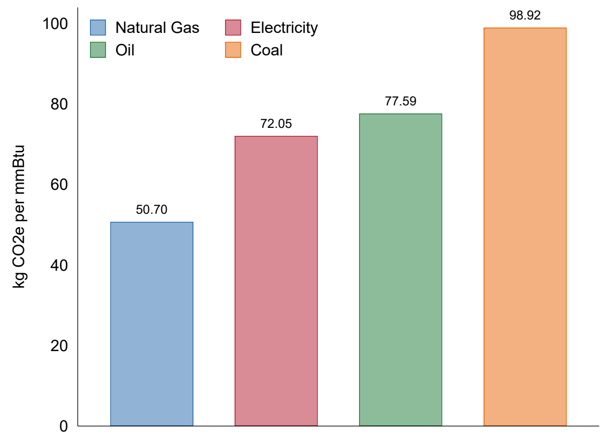
Figure 19: Pollution Intensity of Energy in India and Canada ($kg\ CO_{2e}/mmBtu$).

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 19b, I compare the within industry average pollution intensity for 5 heavy manufacturing industries: Pulp & paper, cement, steel, aluminium, and glass.

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 19a 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 20, 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 19a.



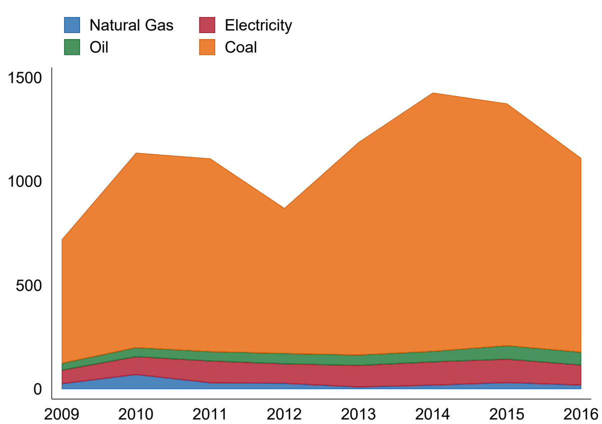
(a) Total Emissions by Fuel



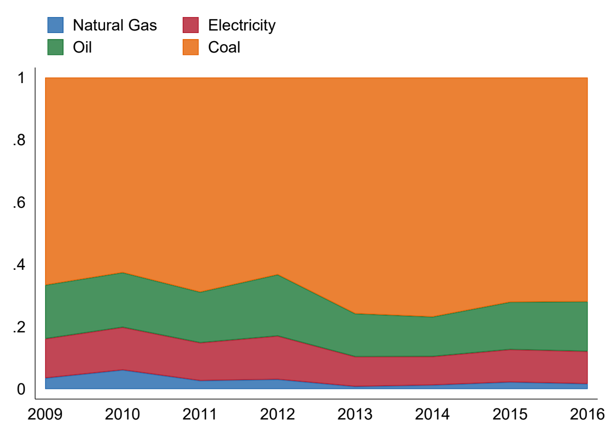
(b) Average Emission Intensity by Fuel

Figure 20: Comparison of Fuels used by ASI establishments

Note: Figure 20a aggregates across all manufacturing establishments in the ASI by year, and suggests a much lower usage of natural gas compared to coal. Figure 20b 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).



(a) Total Quantity by Fuel (mmBtu)



(b) Percent of Total Quantity by Fuel

Figure 21: Comparison of Fuels used by Steel Establishment

A.5.2 Evidence on Switching and Mixing

Indian Plants

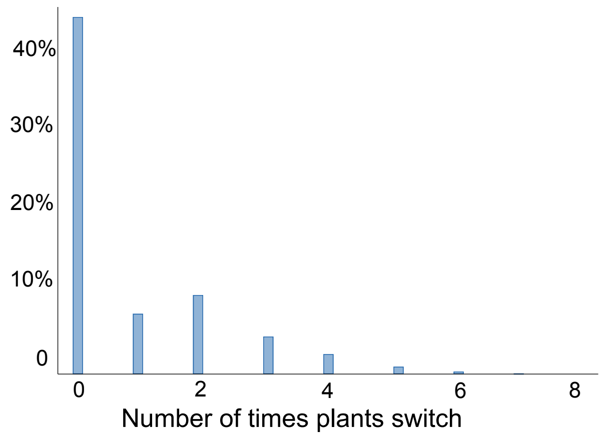


Figure 22: Number of Times Unique Plants add or drop a Fuel (ASI)

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 20: 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 21: 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)

A.5.3 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.

A.5.4 Natural Gas Demand by Sector

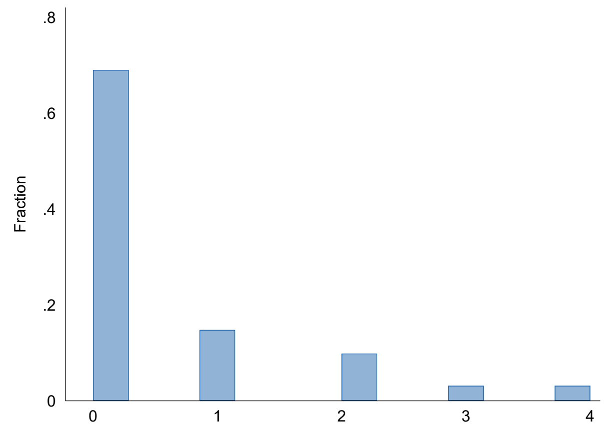


Figure 23: Number of Times Unique Plants add or drop a Fuel (U.S. GHGRP - Pulp & Paper Manufacturing)

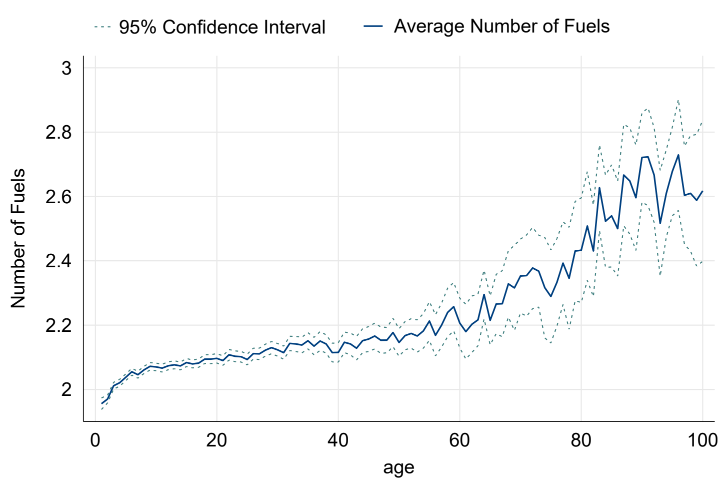


Figure 24: Number of fuels by plant age, average of all ASI plants

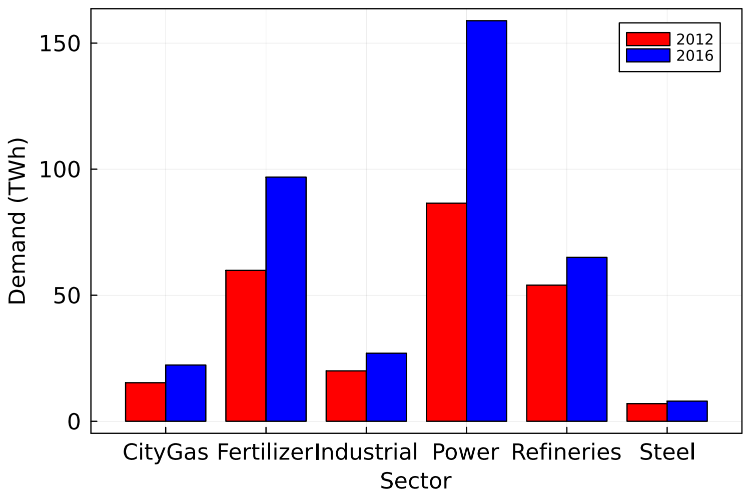
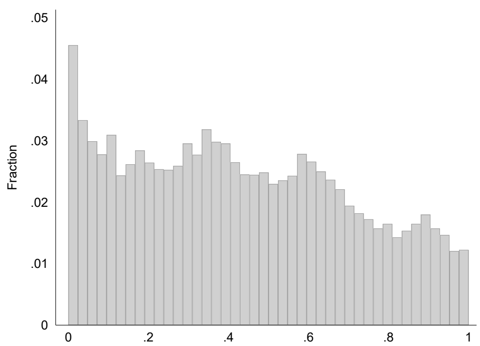


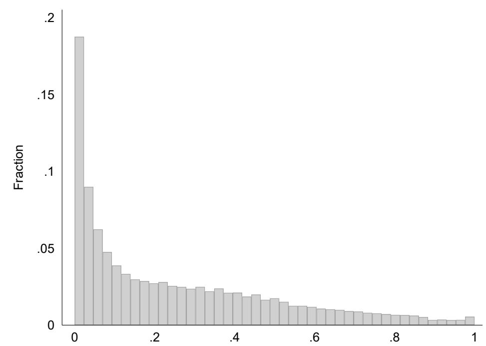
Figure 25: Projected Natural Gas Demand by Sector - all of India (2012 and 2016)

Notes: Data retrieved from the Petroleum and Natural Gas Regulatory Board's Data Bank (Petroleum and Natural Gas Regulatory Board, 2023)

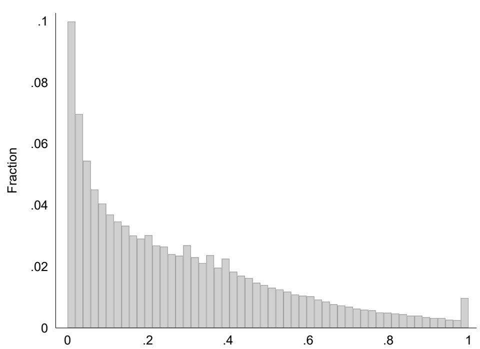
A.5.5 Fuel Expenditure Shares - ASI



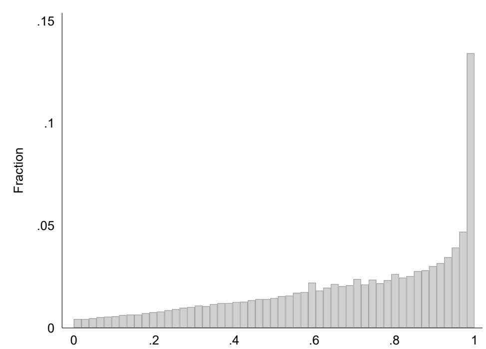
(a) Coal



(b) Gas



(c) Oil



(d) Electricity

Figure 26: Distribution of Expenditure Shares

A.5.6 Relationship between changes in vertical integration and fuel switching

TBD

A.5.7 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 22: 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 Closing the Model: Aggregation details

Given a mass of N_t operating plants, income I_t and aggregate demand shock e^{Γ_t}, the representative consumer solves:

$$\begin{aligned} \max_{\{Y_{it}\}_{i=1}^{N_t}, Y_{0t}} \mathbb{U} &= Y_{0t} + \frac{e^{\Gamma_t}}{\theta} \left(\frac{1}{N_t} \int_{\Omega_i} (N_t Y_{it})^{\frac{\rho-1}{\rho}} di \right)^{\frac{\theta\rho}{\rho-1}} \\ s.t. \quad Y_{0t} &+ \int_{\Omega_i} P_{it} Y_{it} di \leq I_t \end{aligned} \tag{22}$$

Following Helpman and Itskhoki (2010), this can be separated in two problems. First, the consumers chooses consumption of the aggregate final good Y_t, given some aggregate price index P_t and aggregate demand shock e^{Γ_t}:

$$\begin{aligned} &\max_{Y_{0t}, Y_t} Y_{0t} + \frac{e^{\Gamma_t}}{\theta} Y_t^{\theta} \\ &s.t. \quad Y_{0t} + P_t Y_t \leq I_t \end{aligned}$$

Optimal consumption of the aggregate final good is given by $Y_t(P_t) = \left(\frac{P_t}{e^{\Gamma_t}}\right)^{\frac{-1}{1-\theta}}$, and consumption of the outside good is given by $Y_{0t}(P_t) = I_t = P_tY_t(P_t) = I_t - e^{\Gamma_t \frac{1}{1-\theta}} P_t^{\frac{-\theta}{1-\theta}}$. Putting the two together yields the indirect utility $\mathbb{V}$, which corresponds to the consumer surplus due to quasi-linear preferences:

$$\mathbb{V} = I_t + \left(\frac{1}{1-\theta}\right)\Gamma_t^{\frac{1}{1-\theta}} P_t^{\frac{-\theta}{1-\theta}}$$

This is the same indirect utility function as in Helpman and Itskhoki (2010), augmented with an aggregate demand shock. Keeping income constant, consumer surplus is decreasing in the aggregate price index. Then, the representative consumer chooses which varieties to allocate for a given quantities of good Y_t by minimizing the cost of different varieties:

$$\min_{\{Y_{it}\}_{i=1}^{N_t}} \int_{\Omega_i} P_{it}Y_{it} \quad s.t. \quad Y_t = \left(\frac{1}{N_t}\int_{\Omega_i} (N_tY_{it})^{\frac{\rho-1}{\rho}} di\right)^{\frac{\rho}{\rho-1}}$$

Solving this cost-minimization problem yields the following conditional demand of each varieties:

$$Y_{it}(Y_t) = \frac{Y_t}{N_t}\left(\frac{P_{it}}{P_t}\right)^{-\rho} \tag{23}$$

Combining both steps together yields the demand for each varieties, corresponding to equation 4 in the main text:

$$Y_{it} = \frac{e^{\Gamma_t \frac{1}{1-\theta}}}{N_t} P_t^{\frac{\rho(1-\theta)-1}{1-\theta}} P_{it}^{-\rho}$$

Where the aggregate price index is such that $\int_{\Omega_t} P_{it}Y_{it}di = P_tY_t$ and is given by $P_t = \left(\frac{1}{N_t}\int_{\Omega_i} P_{it}^{1-\rho}\right)^{\frac{1}{1-\rho}}$.

C Identification

C.1 Derivation of Estimating Equation for Outer Production Function

Production function:

$$\frac{Y_{it}}{\overline{Y}} = e^{\omega_{it}}\left(\alpha_k\Big(\frac{K_{it}}{\overline{K}}\Big)^{\frac{\sigma-1}{\sigma}} + \alpha_l\Big(\frac{L_{it}}{\overline{L}}\Big)^{\frac{\sigma-1}{\sigma}} + \alpha_m\Big(\frac{M_{it}}{\overline{M}}\Big)^{\frac{\sigma-1}{\sigma}} + \alpha_e\Big(\frac{E_{it}}{\overline{E}}\Big)^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\eta\sigma}{\sigma-1}} \tag{24}$$

$$= e^{\omega_{it}}\Big(\alpha_k\tilde{K}_{it}^{\frac{\sigma-1}{\sigma}} + \alpha_l\tilde{L}_{it}^{\frac{\sigma-1}{\sigma}} + \alpha_m\tilde{M}_{it}^{\frac{\sigma-1}{\sigma}} + \alpha_e\tilde{E}_{it}^{\frac{\sigma-1}{\sigma}}\Big)^{\frac{\eta\sigma}{\sigma-1}} \tag{25}$$

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

Assumption 3. *L_{it}, M_{it}, E_{it} are flexible inputs*

Assumption 4. *I observe the quantity for L_{it} and K_{it} but only spending for materials and energy: $S_{M_{it}}, S_{E_{it}}$*

Profit-maximization subject to technology and demand constraint:

$$\begin{aligned}
&\max_{L_{it},M_{it},E_{it}} \Big\{P_{it}(Y_{it})Y_{it} - p_{Mit}M_{it} - p_{Eit}E_{it} - w_tL_{it}\Big\}\\
s.t.\quad & Y_{it} = \overline{Y}e^{\omega_{it}}\Big(\alpha_K\tilde{K}_{it}^{\frac{\sigma-1}{\sigma}} + \alpha_L\tilde{L}_{it}^{\frac{\sigma-1}{\sigma}} + \alpha_M\tilde{M}_{it}^{\frac{\sigma-1}{\sigma}} + \alpha_E\tilde{E}_{it}^{\frac{\sigma-1}{\sigma}}\Big)^{\frac{\eta\sigma}{\sigma-1}}\\
& P_{it}(Y_{it}) = \left(\frac{e^{\Gamma_t}}{N_tY_{it}}\right)^{\frac{1}{\rho}} P_t^{\frac{1+\rho(\theta-1)}{(\theta-1)\rho}}
\end{aligned}$$

First-order conditions:

M_{it}/L_{it}:

$$\frac{M_{it}}{\overline{M}} = \Big(\frac{\alpha_L}{\alpha_M}\frac{S_{Mit}}{S_{Lit}}\Big)^{\frac{\sigma}{\sigma-1}}\frac{L_{it}}{\overline{L}} \tag{26}$$

E_{it}/L_{it}:

$$\frac{E_{it}}{\overline{E}} = \Big(\frac{\alpha_L}{\alpha_E}\frac{S_{Eit}}{S_{Lit}}\Big)^{\frac{\sigma}{\sigma-1}}\frac{L_{it}}{\overline{L}} \tag{27}$$

L_{it}:

$$\Big(\frac{e^{\Gamma_t}}{N_t}\Big)^{\frac{1}{\rho}} P_t^{\frac{\rho(1-\theta)-1}{(1-\theta)\rho}} \frac{\rho-1}{\rho}\eta(e^{\omega_{it}}\overline{Y})^{\frac{\rho-1}{\rho}}\alpha_L L_{it}^{\frac{\sigma-1}{\sigma}} ces_{it}^{\frac{\rho[\sigma(\eta-1)+1]-\eta\sigma}{(\sigma-1)\rho}} = S_{L_{it}}$$

Where $ces_{it} = \Big(\alpha_k \tilde{K}_{it}^{\frac{\sigma-1}{\sigma}} + \alpha_l \tilde{L}_{it}^{\frac{\sigma-1}{\sigma}} + \alpha_m \tilde{M}_{it}^{\frac{\sigma-1}{\sigma}} + \alpha_e \tilde{E}_{it}^{\frac{\sigma-1}{\sigma}}\Big)$

using the FOC for labor, I can solve for total factor productivity $e^{\omega_{it}}$:

$$e^{\omega_{it}\frac{\rho-1}{\rho}} = \overline{Y}^{\frac{\rho-1}{\rho}} \frac{\rho}{\rho-1}\frac{1}{\eta}\Big(\frac{N_t}{e^{\Gamma_t}}\Big)^{\frac{1}{\rho}} P_t^{\frac{1-\rho(1-\theta)}{(1-\theta)\rho}} \frac{S_{L_{it}}}{\alpha_L L_{it}^{\frac{\sigma-1}{\sigma}}} ces_{it}^{\frac{\eta\sigma-\rho[\sigma(\eta-1)+1]}{(\sigma-1)\rho}} \tag{28}$$

Plug (28) into revenue equation:

$$\begin{aligned}
R_{it} &= P_{it}(Y_{it})Y_{it}e^{u_{it}} \\
&= \Big(\frac{e^{\Gamma_t}}{N_t}\Big)^{\frac{1}{\rho}} P_t^{\frac{1+\rho(\theta-1)}{(\theta-1)\rho}} Y_{it}^{\frac{\rho-1}{\rho}} e^{u_{it}} \\
&= \Big(\frac{e^{\Gamma_t}}{N_t}\Big)^{\frac{1}{\rho}} P_t^{\frac{1+\rho(\theta-1)}{(\theta-1)\rho}} \Bigg(e^{\omega_{it}}\Big(\alpha_K \tilde{K}_{it}^{\frac{\sigma-1}{\sigma}} + \alpha_L \tilde{L}_{it}^{\frac{\sigma-1}{\sigma}} + \alpha_M \tilde{M}_{it}^{\frac{\sigma-1}{\sigma}} + \alpha_E \tilde{E}_{it}^{\frac{\sigma-1}{\sigma}}\Big)^{\frac{\eta\sigma}{\sigma-1}}\Bigg)^{\frac{\rho-1}{\rho}} e^{u_{it}} \\
&= \frac{\rho}{\rho-1}\frac{1}{\eta}\Big(\alpha_K \tilde{K}_{it}^{\frac{\sigma-1}{\sigma}} + \alpha_L \tilde{L}_{it}^{\frac{\sigma-1}{\sigma}} + \alpha_M \tilde{M}_{it}^{\frac{\sigma-1}{\sigma}} + \alpha_E \tilde{E}_{it}^{\frac{\sigma-1}{\sigma}}\Big)e^{u_{it}}
\end{aligned}$$

Plug ratio of FOCs (26) and (27) into the previous equation:

$$\begin{aligned}
R_{it} &= \frac{\rho}{\rho-1}\frac{1}{\eta}S_{Lit}\Bigg(\frac{\alpha_k}{\alpha_L}\Big(\frac{\tilde{K}_{it}}{L_{it}}\Big)^{\frac{\sigma-1}{\sigma}} + 1 + \frac{S_{Mit}}{S_{Lit}} + \frac{S_{Eit}}{S_{Lit}}\Bigg)e^{u_{it}} \\
&= \frac{\rho}{\rho-1}\frac{1}{\eta}\Bigg(S_{Lit}\Big(1+\frac{\alpha_k}{\alpha_L}\Big(\frac{\tilde{K}_{it}}{\tilde{L}_{it}}\Big)^{\frac{\sigma-1}{\sigma}}\Big) + S_{Mit} + S_{Eit}\Bigg)e^{u_{it}}
\end{aligned}$$

Estimating Equation:

$$\ln R_{it} = \ln \frac{\rho}{\rho - 1} + \ln \frac{1}{\eta} + \ln \Bigg(S_{Lit}\Big(1 + \frac{\alpha_k}{\alpha_L}\Big(\frac{\tilde{K}_{it}}{\tilde{L}_{it}}\Big)^{\frac{\sigma-1}{\sigma}}\Big) + S_{Mit} + S_{Eit}\Bigg) + u_{it} \tag{29}$$

C.2 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 (13) for all relative fuel-augmenting productivity:

$$\begin{aligned} \tilde{p}_{eit} &= \mu_{it}\psi_{eit}\Bigg(\sum_{f\in\mathcal{F}_{it}}\Big(\tilde{\psi}_{fit}\tilde{e}_{fit}\Big)^{\frac{\lambda-1}{\lambda}}\Bigg)^{\frac{1}{\lambda-1}}\tilde{e}_{eit}^{-1/\lambda} \\ &= \mu_{it}\psi_{eit}\Big(\sum_{f\in\mathcal{F}_{it}}\frac{s_{fit}}{s_{eit}}\Big)^{\frac{1}{\lambda-1}} \end{aligned} \tag{30}$$

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 (30) implies that the first order condition for electricity is always satisfied:

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

C.3 Computational Details on Solving the Dynamic Discrete Choice Model

I show how to iterate over the expected value function $\vec{W}$ until $|| \vec{W}^{n+1} - \vec{W}^{n} ||$ is small enough with a very large state space, where for any set of states today $s, \mathcal{F}$.

$$W^n(s, \mathcal{F}) = \gamma + \log\left(\sum_{\mathcal{F}' \in \mathbb{F}} \exp\left(\pi(s, \mathcal{F}) + \Phi(\mathcal{F}' \mid \mathcal{F}) + \beta \int W^n(s', \mathcal{F}')dF(s' \mid s)\right)\right)$$

To evaluate the expected value function, note that there are originally 12 state variables: prices and productivity of all 4 fuels, hicks neutral productivity, the price of material inputs, year of observation, and whether a plan is located near a pipeline. I can reduce the dimension of the state space to 8 state variables, 2 of which are deterministic and 6 of which are follow a Markov process. The 6 Markovian state variables are hicks-neutral productivity z, price of materials p_m, price/productivity of electricity p_e/ψ_e, price/productivity of oil p_o/ψ_o, price/productivity of gas p_g/ψ_g, and price/productivity of coal p_c/ψ_c, which are allowed to be correlated. Then,

$$\int W^{n+1}(s', \mathcal{F}')dF(s' \mid s) = \int_z \int_{p_m} \int_{\frac{p_e}{\psi_e}} \int_{\frac{p_o}{\psi_o}} \int_{\frac{p_g}{\psi_g}} \int_{\frac{p_c}{\psi_c}} W^n\left(z', p'_m \frac{p'_e}{\psi'_e}, \frac{p'_o}{\psi'_o}, \frac{p'_g}{\psi'_g}, \frac{p'_c}{\psi'_c}, \mathcal{F}', t, d\right) \times$$
$$f_{z', p'_m, \frac{p'_e}{\psi'_e}, \frac{p'_o}{\psi'_o}, \frac{p'_g}{\psi'_g}, \frac{p'_c}{\psi'_c}}\left(z', p'_m, \frac{p'_e}{\psi'_e}, \frac{p'_o}{\psi'_o}, \frac{p'_g}{\psi'_g}, \frac{p'_c}{\psi'_c} \middle| z, p_m, \frac{p_e}{\psi_e}, \frac{p_o}{\psi_o}, \frac{p_g}{\psi_g}, \frac{p_c}{\psi_c}\right) d_z d_{p_m} d_{\frac{p_e}{\psi_e}} d_{\frac{p_o}{\psi_o}} d_{\frac{p_g}{\psi_g}} d_{\frac{p_c}{\psi_c}}$$

Where t corresponds to year of observation and d is an indicator for access to natural gas pipeline. I approximate this high dimensional expected value function by discretizing the state space and the underlying Markov process. Since most state variables are highly persistent AR(1) processes, I use Rouwenhorst (1995) to discretize the process. Let M be the number of points on each grid. I am currently using $M = 4$ which gives me $4^6 = 4{,}096$ grid points for the Markovian state variables. When adding the 6 years of observations between 2010 and 2015 as well as the access to pipeline indicator, I get $(4^6) * 6 * 2 = 49{,}152$. However, the curse of dimensionality really starts to kick in when I add the distribution of comparative advantages for gas and coal (see later sections). With three possible values for gas and coal, this gives me 9 possible combination of comparative advantages. Ultimately, I am left with $(4^6) * 6 * 2 * 9 = 442{,}368$ grid points. Using this discretization process, I can then represent the value function as a block matrix $\vec{W}^n$ that contains all combinations of states. Let S be the set of all state variable combinations, $\Gamma(s' \mid s)$ be the vector of all state transition probabilities when starting at state s (in vectorized form), $\mathbf{\Pi}$ be the vector of all possible profit combinations, $\vec{\mathcal{K}}$ be the vector of all possible fuel set switching costs. Then

$$\vec{W} \approx \gamma + \log\left(\sum_{\mathcal{F}'\in\mathbb{F}} \exp\left(\mathbf{\Pi} + \vec{\mathcal{K}}(\mathcal{F}') + \beta\Big[\bigotimes_{s\in S}\Gamma(s' \mid s)\Big]^T \vec{W}\right)\right) \tag{31}$$

Lastly, to reduce computational burden, I iterate over equation (31) by paralellizing across all possible combination of starting states using graphics processing units (GPU) Arrays with CUDA. Computational gains using GPU Arrays are significant order of magnitude over standard CPU parallelization. Detailed Julia code is available on my Github.

C.4 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 23: 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 24: 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.

C.5 Details of EM Algorithm to recover distribution of fixed costs and comparative advantages

The procedure to estimate the fixed costs parameters θ_1 and the unselected, unconditional distribution of fuel-specific random effects is explained below. 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 Miller (2011) that uses the conditional choice probabilities (CCP) and forward simulations to update the value function. In the main version of the paper, I am using the nested fixed point version with a large grid for the state space as discussed in Appendix C.3.

$$\begin{aligned}
\ln \mathcal{L}(\mathcal{F}, s \mid \theta_1, \theta_2) &= \sum_{i=1}^{n} \ln \left[\sum_k \pi_k \left[\prod_{t=1}^{T} Pr(\mathcal{F}_{it+1} \mid \mathcal{F}_{it}, s_{it}, \mu_{fi} = \mu_k; \theta_1, \theta_2) \right] \right] + \sum_{i=1}^{n} \sum_{t=1}^{T} \ln f(s_{it} \mid s_{it-1}; \theta_2) \\
&= \sum_{i=1}^{n} \ln \left[\sum_k \pi_k \left(\prod_{t=1}^{T} \frac{e^{\upsilon_{\mathcal{F}_{it+1}}(\mathcal{F}_{it}, s_{it}, \mu_i = \mu_k; \theta_1, \theta_2)}}{\sum_{\mathcal{F} \subseteq \mathbb{F}} e^{\upsilon_{\mathcal{F}}(\mathcal{F}_{it}, s_{it}, \mu_i = \mu_k; \theta_1, \theta_2)}} \right) \right] + \sum_{i=1}^{n} \sum_{t=1}^{T} \ln f(s_{it} \mid s_{it-1}; \theta_2)
\end{aligned}$$

In principle, one can directly estimate both the fixed costs θ_1 and the distribution of comparative advantages from the full information likelihood above. However, this is computationally very expensive and rarely used in practice. For this reason, Arcidiacono and Jones (2003) use Baye's law to show that the first-order conditions of the full information likelihood with respect to all parameters are the same as the first-order conditions of the posterior likelihood with respect to fixed costs θ_1 given some prior guess of the distribution of unobserved heterogeneity. This is the key result that allows me to use the EM algorithm

$$\begin{aligned}
\hat{\theta}_1 &= \arg\max_{\theta_1, \theta_2, \pi} \sum_{i=1}^{n} \ln \left[\sum_k \pi_k \left[\prod_{t=1}^{T} Pr(\mathcal{F}_{it+1} \mid \mathcal{F}_{it} s_{it}, \mu_i = \mu_k; \theta_1, \theta_2) \right] \right] \\
&\equiv \arg\max_{\theta_1} \sum_{i=1}^{N} \sum_{t=1}^{T} \sum_k \rho(\mu_k \mid \mathcal{F}_i, s_i; \hat{\theta}_1, \hat{\theta}_2, \hat{\pi}) \ln Pr(\mathcal{F}_{it+1} \mid \mathcal{F}_{it}, s_{it}, \mu_i = \mu_k; \theta_1, \hat{\theta}_2)
\end{aligned}$$

Estimation then proceeds iteratively as follows:

1. Estimate the distribution of state variables externally $\hat{\theta}_2$. These stay fix throughout the procedure.

2. Initialize fixed cost parameters θ_1^0 and guess some initial probabilities $\{\pi_{f1}^0, \pi_2^0, ..., \pi_K^0\}$. I use the distribution of selected random effects to initialize this distribution.

3. Do value function iteration (VFI) to update the expected value function W for all combination of state variables conditional on these guesses, where different realizations of the random effects μ_k are just another state variable that is fixed over time.

$$W(s, \mathcal{F}, \mu_k; \theta_1^0, \hat{\theta}_2) = \gamma + \log \left(\sum_{\mathcal{F}' \in \mathbb{F}} \exp \left(\pi(s, \mathcal{F}) + \mathcal{K}(\mathcal{F}' \mid \mathcal{F}, s; \theta_1^0) + \beta \int W(s', \mathcal{F}', \mu_k; \theta_1^0, \hat{\theta}_2) dF(s' \mid s; \hat{\theta}_2) \right) \right)$$

4. Get posterior conditional probabilities that plant i is of type k, $\rho^1(\mu_k \mid \mathcal{F}_i, s_i; \theta_1^0, \hat{\theta}_2, \pi^0)$, according to Baye's law:

$$\rho^1(\mu_k \mid \mathcal{F}_i, s_i; \theta_1^0, \hat{\theta}_2, \pi^0) = \frac{\pi_{fk}^0 \left[\prod_{t=1}^T \left[\prod_{\mathcal{F} \subseteq \mathbb{F}} \left[Pr(\mathcal{F}_{it} \mid s_{it}, \mu_i = \mu_k; \theta_1^0, \hat{\theta}_2) \right]^{\mathbb{I}(\mathcal{F}_{it} = \mathcal{F})} \right] \right]}{\sum_k \pi_k^0 \left[\prod_{t=1}^T \left[\prod_{\mathcal{F} \subseteq \mathbb{F}} \left[Pr(\mathcal{F}_{it} \mid s_{it}, \mu_i = \mu_k; \theta_1^0, \hat{\theta}_2) \right]^{\mathbb{I}(\mathcal{F}_{it} = \mathcal{F})} \right] \right]}$$

5. **E-step:** Update the unconditional comparative advantage probabilities as follows:

$$\pi_k^1 = \frac{\sum_{i=1}^n \rho^1(\mu_k \mid \mathcal{F}_i, s_i; \theta_1^0, \hat{\theta}_2, \pi^0)}{n} \qquad \forall k$$

6. **M-step:** Find fixed cost parameters θ_1^1 that maximize the (log)-likelihood conditional on current guess of unconditional and conditional probabilities $\pi_k^1, \rho^1(\mu_k \mid .)$

7. Repeat 3-6 until the full information likelihood is minimized.

D Counterfactuals

D.1 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 $f(\cdot)$,

$$\sum_{j=1}^{n} C_j = f(X_1, X_2, ..., X_n), \tag{32}$$

so that $C_j f(\cdot)$ can be interpreted as the proportion of $f(.)$ that can be attributed to X_j.[32] 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

$$C_j = \sum_{k=0}^{n-1} \frac{(n-k-1)!k!}{n!} \left(\sum_{s \subseteq S_k \setminus \{X_j\} : |s| = k} [f(s \cup X_j) - f(s)] \right), \tag{33}$$

where n is the total number of arguments in the original function f, $S_k \setminus \{X_j\}$ is the set of all *"sub-models"* that contain k arguments and exclude argument X_j.[33] 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 (33) 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: $\{\underbrace{(0,0)}_{k=0}, \underbrace{(1,0),(0,1)}_{k=1}, \underbrace{(1,1)}_{k=2}\}$.

[32]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 $C_j < 0$.

[33]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.

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

D.1.1 Application to Decomposition of the Carbon Tax Revenues Along Subsidy Rate

Given a subsidy rate $s \in [0, 1]$ towards the fixed costs of natural gas, the function of interest is the change in tax revenues/externality damages between the economy with subsidy s and the economy with no subsidy:

$$
\begin{aligned}
\Delta\mathcal{T}(s) &= \mathcal{T}(s) - \mathcal{T}(0) \\
&= \mathbb{E}_0 \Big(\sum_{t=0} \beta^t \sum_i \tau_f e_{fit}(s) \Big) - \mathbb{E}_0 \Big(\sum_{t=0} \beta^t \sum_i \tau_f e_{fit}(0) \Big)
\end{aligned}
$$

Note that this function is implicitely a function of all state variables and the carbon tax rate. As such, it can be decomposed into multiple arguments. The two arguments of interests here are the total quantity of energy E, which corresponds to the scale effect (higher quantity of energy equals higher quantities of all fuels), and the distribution of fuel sets in the economy $\mathcal{F}$, which corresponds to the substitution effect. Then, expected tax revenues can be rewritten as:

$$
\begin{aligned}
\Delta\mathcal{T}(\underbrace{E(s)}_{\text{scale}}, \underbrace{\mathcal{F}(s)}_{\text{substitution}}) &= \mathcal{T}(E(s), \mathcal{F}(s)) - \mathcal{T}(E(0), \mathcal{F}(0)) \\
&= \mathbb{E}_0 \Big(\sum_{t=0} \beta^t \sum_i \tau_f e_{fit}\big(E_{it}(s), \mathcal{F}_{it}(s)\big) \Big) - \mathbb{E}_0 \Big(\sum_{t=0} \beta^t \sum_i \tau_f e_{fit}\big(E_{it}(0), \mathcal{F}_{it}(0)\big) \Big)
\end{aligned}
$$

Here I define the null case for both scale $\emptyset_E$ and substitution $\emptyset_{\mathcal{F}}$ arguments as the case with no substitute, such that the function is well defined, satisfying the criteria laid out in Shorrocks (2013).

$$
\Delta\mathcal{T}(\emptyset_E, \emptyset_{\mathcal{F}}) = \mathcal{T}(E(0), \mathcal{F}(0)) - \mathcal{T}(E(0), \mathcal{F}(0)) = 0
$$

Since there are two arguments, there will always be only two submodels that can exclude each argument: when the other argument is present and when it is not, with associated probability of $\frac{1}{2}$ for each submodel. It is then quite easy to show that the total partial effect of adding the scale and substitution effect, respectively is as follows:

$$\begin{aligned}
\mathcal{C}_{\text{scale}} &= \frac{1}{2}\underbrace{\Big(\Delta\mathcal{T}(E(s),\mathcal{F}(s)) - \Delta\mathcal{T}(\emptyset_E,\mathcal{F}(s))\Big)}_{\text{adding scale with substitution}} + \frac{1}{2}\underbrace{\Big(\Delta\mathcal{T}(E(s),\emptyset_{\mathcal{F}}) - \Delta\mathcal{T}(\emptyset_E,\emptyset_{\mathcal{F}})\Big)}_{\text{adding scale without substitution}} \\
&= \frac{1}{2}\Big(\Delta\mathcal{T}(s) - \Delta\mathcal{T}(\emptyset_E,\mathcal{F}(s))\Big) + \frac{1}{2}\Big(\Delta\mathcal{T}(E(s),\emptyset_{\mathcal{F}})\Big)
\end{aligned}$$

$$\begin{aligned}
\mathcal{C}_{\text{substitution}} &= \frac{1}{2}\underbrace{\Big(\Delta\mathcal{T}(E(s),\mathcal{F}(s)) - \Delta\mathcal{T}(E(s),\emptyset_{\mathcal{F}})\Big)}_{\text{adding substitution with scale}} + \frac{1}{2}\underbrace{\Big(\Delta\mathcal{T}(\emptyset_E,\mathcal{F}(s)) - \Delta\mathcal{T}(\emptyset_E,\emptyset_{\mathcal{F}})\Big)}_{\text{adding substitution without scale}} \\
&= \frac{1}{2}\Big(\Delta\mathcal{T}(s) - \Delta\mathcal{T}(E(s),\emptyset_{\mathcal{F}}\Big) + \frac{1}{2}\Big(\Delta\mathcal{T}(\emptyset_E,\mathcal{F}(s))\Big)
\end{aligned}$$

Lastly, it can be seen that this decomposition satisfies the additive criteria laid out by Shorrocks (2013):

$$\begin{aligned}
\mathcal{C}_{\text{scale}} + \mathcal{C}_{\text{substitution}} &= \Delta\mathcal{T}(s) - \frac{1}{2}\Delta\mathcal{T}(\emptyset_E,\mathcal{F}(s)) + \frac{1}{2}\Delta\mathcal{T}(\emptyset_E,\mathcal{F}(s)) + \frac{1}{2}\Delta\mathcal{T}(E(s),\emptyset_{\mathcal{F}}) - \frac{1}{2}\Delta\mathcal{T}(E(s),\emptyset_{\mathcal{F}}) \\
&= \Delta\mathcal{T}(s)
\end{aligned}$$

D.1.2 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 $\mathcal{F}$:

$$f(\mathbf{p},\mathbf{\Psi},\mathcal{F}) = \frac{1}{N_{\mathcal{F}}}\sum_{i:\mathcal{F}}\Big(\sum_{f\in\mathcal{F}}\left(\tilde{p}_{fit}/\psi_{fit}\right)^{1-\lambda}\Big)^{\frac{1}{1-\lambda}} - \frac{1}{N_{oe}}\sum_{i:oe}\Big(\sum_{f\in oe}\left(\tilde{p}_{fit}/\psi_{fit}\right)^{1-\lambda}\Big)^{\frac{1}{1-\lambda}} \tag{34}$$

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 (34) is equal to zero when all its arguments are null, and thus satisfies the criteria laid out in Shorrocks (2013):

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

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

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

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.

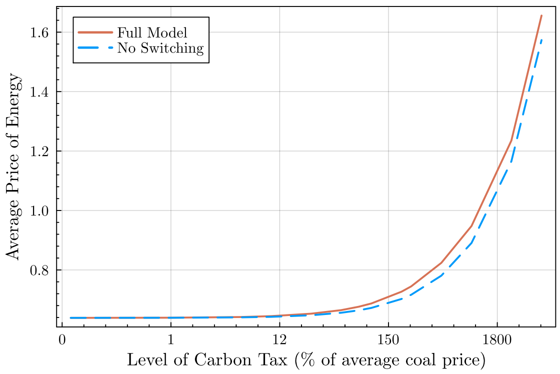
		OCE	OGE	OGCE
Total Difference	Percent (%) Difference with OE	-65.65	-71.54	-86.97
Option Value		36.14	5.42	6.3
Fuel Productivity	Percent (%) of Total	62.6	97.75	94.84
Fuel Prices		1.25	-3.18	-1.14

Table 25: Shapley Decomposition of the Difference in Average Marginal Cost of Energy Between Fuel Sets

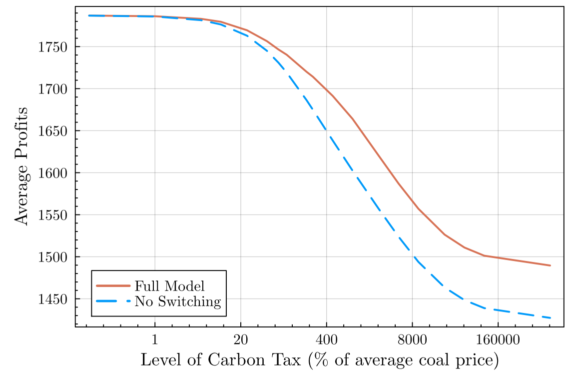
Notes: I compare the observed differences in the average (across plants) marginal cost of realized energy between plants who use coal and/or gas on top of oil and electricity (OCE,OGE,OGCE) relative to plants who only use oil and electricity (OE).

D.2 Details to Compare Trade-off Between Emissions Reduction and Output

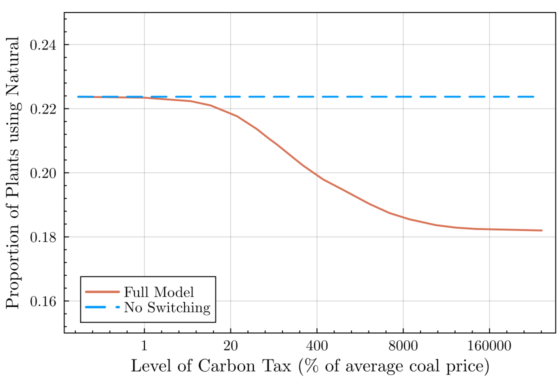
D.2.1 More Details on the Comparison between Switching and No Switching



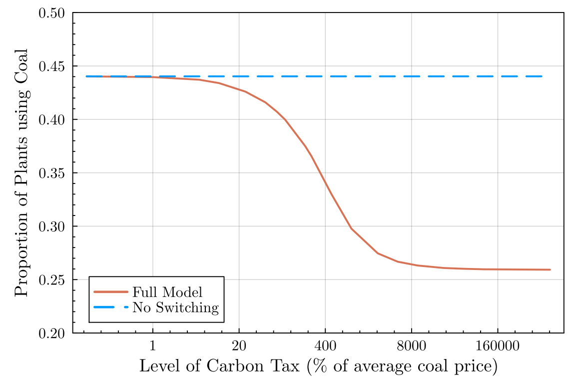
(a) Average Price of Energy $p_{E_{it}}$



(b) Average Profits (Including Fixed Costs)



(c) Aggregate Share of Gas Users



(d) Aggregate Share of Coal Users

Figure 27: Comparison of Model with Switching and Without Switching for Selected Variables Across Carbon Tax Levels

D.2.2 Estimation of Energy Production Function with Energy Productivity

The energy production function is as follows:

$$E_{it} = \psi_{Eit}\Big(\sum_{f\in\mathcal{F}_{it}} \beta_f e_{fit}^{\frac{\lambda-1}{\lambda}}\Big)^{\frac{\lambda}{\lambda-1}} \qquad \sum_{f\in\{o,g,c,e\}} \beta_f = 1$$

Assuming that the log of energy productivity follows and AR(1) process with year dummies $\ln\psi_{Eit} = \mu_0^{\psi_E} + \mu_t^{\psi_E} + \rho_{\psi_E}\ln\psi_{Eit-1} + \epsilon_{it}^{\psi_E}$, the production function can be written in log as

$$\ln E_{it} = \mu_0^{\psi_E} + \mu_t^{\psi_E} + \frac{\lambda}{\lambda-1}\Big(\sum_{f\in\mathcal{F}_{it}} \beta_f e_{fit}^{\frac{\lambda-1}{\lambda}}\Big) + \rho_{\psi_E}\ln E_{it-1} - \rho_{\psi_E}\frac{\lambda}{\lambda-1}\ln\Big(\sum_{f\in\mathcal{F}_{it}} \beta_f e_{fit-1}^{\frac{\lambda-1}{\lambda}}\Big) + \epsilon_{it}^{\psi_E}$$

This is very similar to the estimating equation for the fully flexible energy production function in the main text, where $\epsilon_{it}^{\psi_E}$ is the innovation to energy productivity between $t-1$ and t. As such, it is independent of period $t-1$ decisions:

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

However, this innovation is correlated with fuel choices at t. I instrument fuel choices at t with aggregate variation in fuel prices due to exogenous reasons such as geopolitical events, which I interact with the share of each fuel to generate electricity by Indian States. These shift-share instruments are the same instruments proposed by Ganapati et al. (2020), which I also use to estimate demand in the main model. Together, these instruments and fuel choices at $t-1$ form a set of moment conditions that satisfy exogeneity and identify the relevant parameters of the production function: $\lambda, \beta_o, \beta_g, \beta_c, \beta_e$. Below are the estimates of the production function:

Table 26: Estimates of Energy Production Function with Energy Productivity

	Steel	
Elasticity of substitution $\hat{\lambda}$	2.173***	(0.240)
Relative productivity of oil $\hat{\beta}_o$	0.099***	(0.011)
Relative productivity of gas $\hat{\beta}_g$	0.049***	(0.012)
Relative productivity of coal $\hat{\beta}_c$	0.426***	(0.033)
Observations	3459	

Standard errors in parentheses

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

D.2.3 Elasticity of the Price of Energy with Respect to Relative Fuel Prices

The price of energy in the fully flexible model is as follows:

$$p_{E_{it}} = \Big(\sum_{f\in\mathcal{F}_{it}} \big(p_{fit}/\psi_{fit}\big)^{1-\lambda}\Big)^{\frac{1}{1-\lambda}}$$

and can be written in terms of price ratios for a given fuel (say gas), where $\tilde{p}_{fit} = p_{fit}/p_{git}$ and likewise for $\tilde{\psi}_{fit}$

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

The the elasticity of this price of energy with respect to relative fuel prices (say coal relative to gas) is as follows:

$$\begin{aligned}
\frac{\partial \ln p_{E_{it}}}{\partial \ln(p_{cit}/p_{git})} &= \frac{1}{p_{E_{it}}}\left[\frac{p_{git}}{1-\lambda}\Big(\sum_{f\in\mathcal{F}_{it}} (\tilde{p}_{fit}/\tilde{\psi}_{fit})\Big)^{\frac{\lambda}{\lambda-1}}\frac{\partial \exp((1-\lambda)(\ln\tilde{p}_{cit} - \ln\tilde{\psi}_{cit}))}{\partial \ln p_{fit}}\right] \\
&= \frac{1}{p_{E_{it}}}\left[p_{git}\Big(\sum_{f\in\mathcal{F}_{it}} (\tilde{p}_{fit}/\tilde{\psi}_{fit})\Big)^{\frac{\lambda}{\lambda-1}}(\tilde{p}_{cit}/\tilde{\psi}_{cit})^{1-\lambda}\right] \\
&= \frac{(\tilde{p}_{cit}/\tilde{\psi}_{cit})^{1-\lambda}}{\sum_{f\in\mathcal{F}_{it}}(\tilde{p}_{fit}/\tilde{\psi}_{fit})^{1-\lambda}}
\end{aligned}$$

which is found in the main text. Additionally, the derivative of this elasticity with respect to coal productivity can be easily found as (**SHOW THIS**):

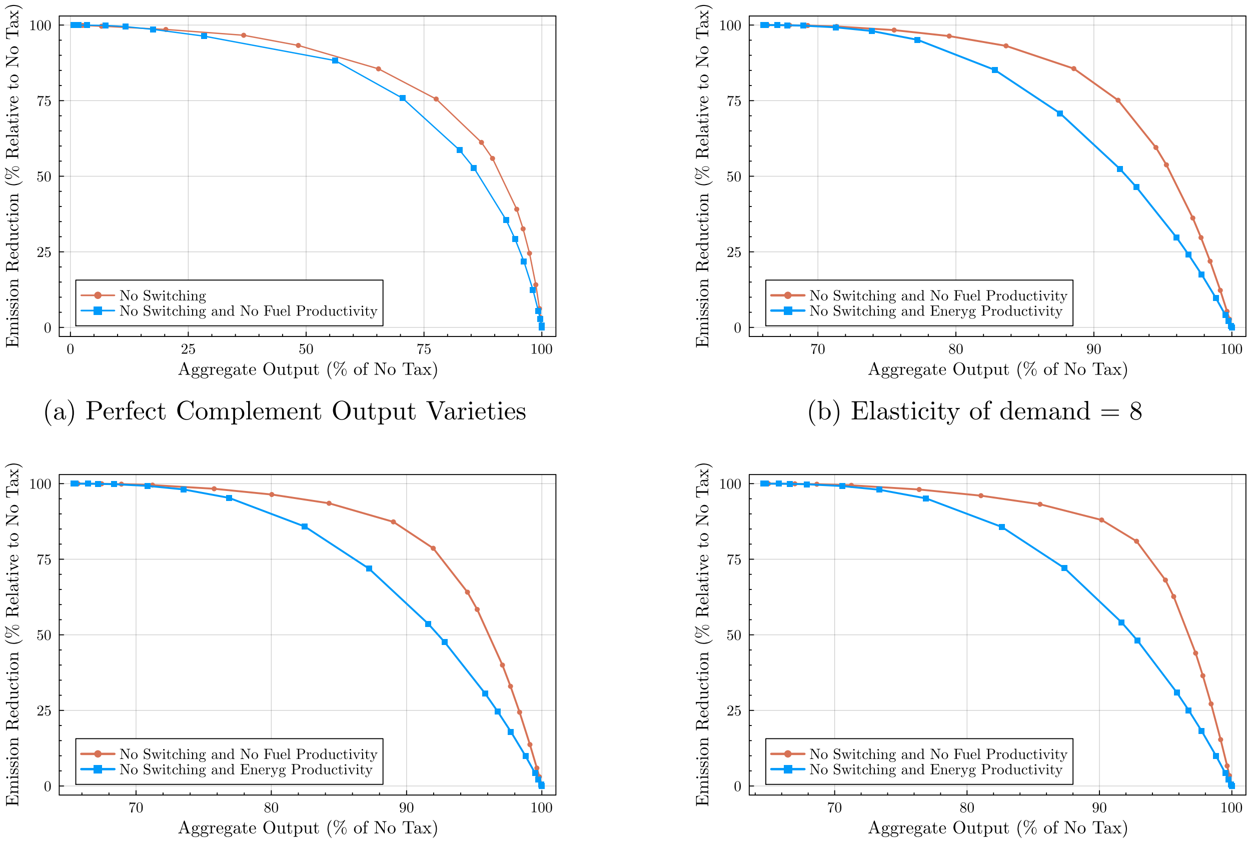
D.2.4 Correlation between Fuel Productivity and TFP

	Total Factors	Gas	Coal	Oil	Elec
Total Factors	1				
Gas	-0.19	1			
Coal	-0.11	0.14	1		
Oil	-0.24	0.02	0.02	1	
elec	-0.56	0.13	0.07	0.06	1

Table 27: Correlation Matrix of Fuel Productivity and Total Factor Productivity

D.2.5 Trade-off Across Values of Demand Elasticity

Below I show the trade-off between aggregate output and emissions for different levels of the carbon tax, comparing how much better the economy fairs when allowing for heterogeneity in fuel productivity rather than just energy productivity. I do this exercise by varying the elasticity of demand, which affects the extent of output reallocation across plants. Consistent with **Oberfield and Raval (2021)**, I find that more elastic demand increases output reallocation across plants when allowing for fuel productivity, which increases the gap between the two production frontiers. This confirms the importance of the output reallocation channel in explaining the aggregate trade-off in the main text.



(a) Perfect Complement Output Varieties

(b) Elasticity of demand = 8

(c) Elasticity of demand = 10

(d) Elasticity of demand = 12

Figure 28: Graphs comparing and contrasting the difference in the aggregate trade-off between emission reduction and output, across different elasticities of demand.

Notes: Figure (a) corresponds to perfect complement output varieties, and refers to a representative consumer who has Leontief preferences across different varieties. While both production frontier get closer to each others with such preferences, there is still a gap is due to the fact due to initial differences in fuel concentration. The economy with heterogeneity in fuel productivity allows for more concentration of fuels across plants, particularly coal, and thus yields more substitution away from coal in *level*, even though the elasticity is the same.

D.2.6 Aggregation of production function without switching

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 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

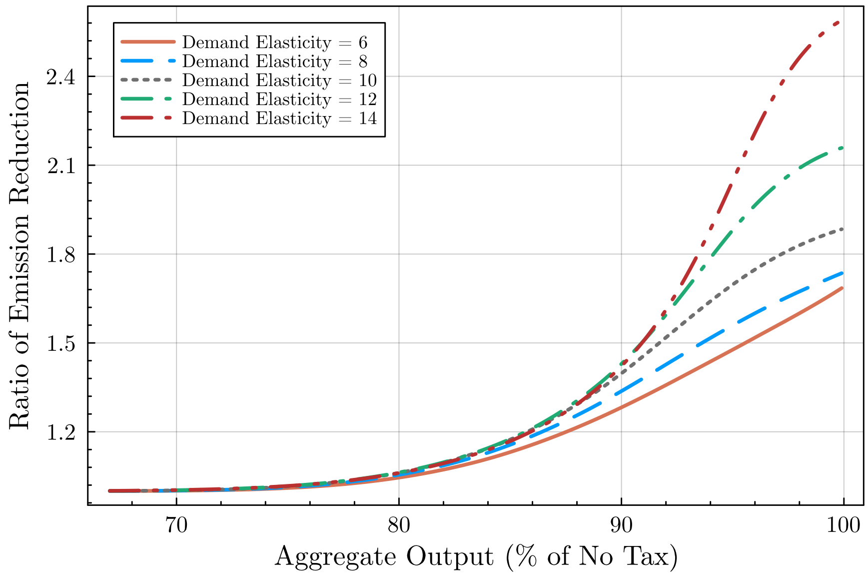


Figure 29: Comparison of the gap between the economy with fuel productivity and the economy with energy productivity.

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

$$
\begin{aligned}
E &= \left(\sum_{f \in \{o,g,c,e\}} \beta_f e_f^{\frac{\lambda-1}{\lambda}} \right)^{\frac{\lambda}{\lambda-1}} \\
\sum_f \beta_f &= 1
\end{aligned}
\tag{35}
$$

In my paper, there are multiple plants 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 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:

$$
E_{\mathcal{F}} = \left(\sum_{f \in \mathcal{F}} \beta_f e_f^{\frac{\lambda-1}{\lambda}} \right)^{\frac{\lambda}{\lambda-1}} \tag{36}
$$

From equation (36) 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.

$$\begin{aligned} e_f &= \sum_{\mathcal{F}} \mathbb{I}(f\in\mathcal{F}) s_{\mathcal{F}} e_f(\mathcal{F}) \\ &= E\left(\frac{\beta_f}{p_f}\right)^{\lambda}\left(\sum_{\mathcal{F}} \mathbb{I}(f\in\mathcal{F}) s_{\mathcal{F}} P_E(\mathcal{F})\right) \end{aligned} \tag{37}$$

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 (37).

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 (37).*

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

$$\tilde{E} = \left(\sum_{f\in\{o,g,c,e\}} \psi_f^{\frac{1}{\lambda}} e_f^{\frac{\lambda-1}{\lambda}}\right)^{\frac{\lambda}{\lambda-1}} \tag{38}$$

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 (38) for a given quantity of energy E is

$$\tilde{e}_f = s_f E \left(\frac{P_{\tilde{E}}}{p_f} \right)^{\lambda} \tag{39}$$

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 (39) is equal to e_f from (37).

□

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 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.

$$E = \inf\left\{\tau(\omega) : \omega \in [0,1]\right\} \tag{40}$$

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:

$$\tau(\omega) = \sum_{f \in \mathcal{F}} \psi_f(\omega) e_f(\omega) \tag{41}$$

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.

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:

$$\begin{aligned} \int_0^1 p(\omega)\tau(\omega)d\omega &= E \int_0^1 p(\omega)d\omega \\ &= Ep_e \end{aligned} \tag{42}$$

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:

$$p(\omega) = \min_{f\in\mathcal{F}} \left\{ \frac{p_f}{\psi_f(\omega)} \right\} \tag{43}$$

Aggregation from tasks to fuels

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

$$\mathcal{T}_f = \left\{\omega : \frac{p_f}{\psi_f(\omega)} \leq \frac{p_j}{\psi_j(\omega)} \ \forall j \neq f \in \mathcal{F}\right\} \tag{44}$$

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

$$e_f(\omega) = \begin{cases} E\psi_f(\omega)^{-1} & \text{if } \omega \in \mathcal{T}_f \\ 0 & \text{otherwise} \end{cases} \tag{45}$$

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):

$$e_f = E \underbrace{\int_{\mathcal{T}_f} \psi_f(\omega)^{-1} d\omega}_{\Gamma_f^{-1}} \tag{46}$$

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

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

$$
\begin{aligned}
p_e &= \frac{1}{E}\int_0^1 p(\omega)\tau(\omega)d\omega \\
&= \frac{1}{E}\sum_f \int_{\mathcal{T}_f} \frac{p_f}{\psi_f(\omega)} E \\
&= \sum_f p_f \Gamma_f
\end{aligned} \tag{47}
$$

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:

$$
\begin{aligned}
p_e E &= \sum_f \int_{\mathcal{T}_f} \frac{p_f}{\psi_f(\omega)} \frac{E}{\mathcal{M}} \\
&= \sum_f p_f e_f
\end{aligned} \tag{48}
$$

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 Grieco et al. (2016):

$$\widehat{\overline{E}\Gamma_{fit}} = e_{fit}\Big(\frac{\overline{E}}{E_{it}}\Big) \tag{49}$$

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 5. *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:

$$\mathcal{T}_{it} = Pr\Big(\frac{p_{fit}}{\psi_{fit}} \leq \frac{p_{jit}}{\psi_{jit}} \forall j \neq f \in \mathcal{F}_{it}\Big) \tag{50}$$

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:

$$
\Gamma_f = \int \psi_f^{-1} \underbrace{\left[\prod_{j \neq f} \int \mathbb{I}\Big(\psi_f \geq p_f \max\{\psi_j/p_j\}\Big) f(\psi_j) d\psi_j\right]}_{\text{Pr(efficiency draw for f is chosen over all other fuels)}} f(\psi_f) d\psi_f \tag{51}
$$

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 6. *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

$$
\Gamma_{g,it} = \underbrace{\frac{\theta}{\theta+1}\Omega_{git}^{-\theta-1}\overline{\psi}_{git}^{\theta}}_{\text{Direct task displacement}} - \underbrace{\left(\frac{p_{cit}}{p_{git}}\right)^{-\theta} \frac{\theta}{2\theta+1}\Omega_{git}^{-2\theta-1}(\overline{\psi}_{git}\overline{\psi}_{cit})^{\theta}}_{\text{Indirect task displacement through fuel c}} \tag{52}
$$

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:

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

$\forall f \in \mathcal{F}$, 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:

$$\min_{\{e_{fit}\}_{f\in\mathcal{F}_{it}},\lambda} \sum_{f\in\mathcal{F}} p_{fit}e_{fit} + \lambda\Big(E_{it} - g(\mathbf{\Psi}_{it}\mathbf{e}_{it})\Big) \tag{53}$$

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.[35] 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:

[35] 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 $k \geq 1$ (Grieco et al., 2016).

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

Matzkin (2008) 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.[36] 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.

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