Balancing Production and Carbon Emissions with Fuel Substitution*

Emmanuel Murray Leclair†

Development Innovation Lab, University of Chicago

This Version: *August 9, 2024*

Click here for the latest version and here for the online appendix.

Abstract

The economic cost of carbon pricing depends on the ability and incentives of firms to switch towards cleaner fuels. Yet, many fundamental economic forces that drive firms' decisions to use different fuels are unobserved, causing significant uncertainty over the effectiveness of carbon policies. In this paper, I propose a new dynamic production model with multidimensional unobserved heterogeneity that underly technology differences and captures how firms' fuel choices respond to price changes. These differences cause heterogeneity in abatement costs, which generates heterogeneous responses to carbon pricing. Leveraging minimal assumptions about optimal input choice and the technology frontier, I quantify the model from a detailed panel of Indian steel establishments. Based on these estimates, implementing a carbon tax equivalent to 2,000 INR/ton (25 USD/ton) of CO_{2e} leads to a 70% reduction in emissions. But only 18% of this reduction comes from fuel-switching within existing firms. I find that the larger reductions come from reallocation of output across firms (58%) and costly reduction in aggregate output (24%). Substantial heterogeneity in the fuel efficiency of existing furnaces coupled with the limited geographical reach of natural gas pipelines towards high-emission firms explains the prevalence of output reallocation relative to fuel switching.

*Most of this paper was written while I was a Ph.D. student in Economics at the University of Western Ontario. 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 Lance Lochner, Rory McGee, Michael Sullivan, Nouri Najjar, Nirav Mehta, Charles Saunders, Martin O'Connell, Matt Gordon, Geoffrey Barrows, Yanyou Chen, Richard Sweeney, Pascual Restrepo, Elie Tamer, Yongseok Shin, Andrew McCallum, Leslie Martin, David Byrne, Yao Luo, and Victor Aguirregabiria for their insightful comments. I also want to thank participants who gave me valuable feedback at the 2024 Society of Economic Dynamics (SED) Conference, the University of Toronto, the University of Melbourne, the University of Virginia, the University of Stavanger, the 2023 Bank of Canada Graduate Student Paper Award, the 2023 Canadian Resource and Environmental Economics Association Conference (CREEA), the 2023 EEA-ESEM Congress, the 2023 European Association for Research in Industrial Economics (EARIE) Conference, 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@uchicago.edu Website: https://sites.google.com/view/emmanuelmurrayleclair/home/

1 Introduction

The reliance on fossil fuels in many manufacturing processes has profound environmental repercussions. The release of greenhouse gases from fossil fuel combustion is the largest contributor to climate change, accounting for 75% of total greenhouse gas emissions (UN Climate, 2023). Market-based policies such as carbon taxes and tradable permit systems have been proposed to induce firms to internalize the social cost of fuel combustion. One of the main objectives of these policies is to incentivize firms to adopt cleaner fuels, thereby reducing emissions while maintaining stable levels of production. The economic costs of such policies, such as loss of output and higher consumer prices, thus crucially depend on the ability and incentives of firms to substitute between fuels.

Fundamental economic forces that drive firms' decisions to use different fuels are ubiquitous. They reflect disparities in prices, technology, and access to critical infrastructures like transportation networks, all of which vary across space and over time (Scott, 2021; Collard-Wexler and De Loecker, 2015). Yet, little is known about how these forces interact with each other, partly because they reflect unobserved heterogeneity and forward-looking considerations that are difficult to quantify. The incentives of firms to adopt a new fuel in response to a policy may vary substantially based on private information about how efficiently they expect to use this new fuel. Similarly, the ability of firms to substitute between existing fuels may be limited or enabled by specific manufacturing processes.

This paper proposes a new dynamic production model that combines multiple fundamental forces to understand how firms make fuel choices and how these choices respond to price changes. This model differs from existing production models of interfuel substitution in that it captures plants' multidimensional fuel choices, fuel productivity heterogeneity due to unobserved technology differences, and dynamic switching between fuel sets subject to fixed switching costs.[1] Leveraging the approach pioneered by Carlson, Burtraw, Cropper and Palmer (2000) and Atkinson and Luo (2023), I show that these factors provide a microfoundation for abatement costs heterogeneity that reflects the ability of firms to substitute towards cleaner fuels in the short and long run. Combined together, these factors thus generate heterogeneous and dynamic responses to price changes that have important implications for carbon policy.

While there is a significant body of work estimating production functions, understanding the heterogeneity and dynamics underlying firms' input selection when they can choose from multiple input combinations poses new challenges.[2] One of these challenges is quantifying input-specific productivity for an input a firm has never used. Leveraging minimal assumptions about optimal input choices and the technology frontier along with a detailed panel of Indian steel plants between

[1]The vast majority of papers on interfuel substitution considers firms' fuel sets to be fixed in time and does not account for fuel-specific productivity. See 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)

[2]See Demirer (2020); Gandhi, Navarro and Rivers (2020); Zhang (2019); Grieco, Li and Zhang (2016); Ackerberg, Caves and Frazer (2015); Levinsohn and Petrin (2003); Blundell and Bond (2000); Olley and Pakes (1996).

2009 and 2016, I quantify the rich heterogeneity in plants' incentives and ability to substitute between fuels. The panel features detailed information about plant-specific specific input prices and quantities along with location within 775 districts, allowing me to interpret this heterogeneity in the context of plants' proximity to key infrastructures such as coal mines and natural gas pipeline networks.

I then use the model to predict how Indian steel plants would respond to a range of different carbon taxes over an horizon of 40 years. India is the second largest steel producer, and steel is one of India's most polluting industries, with coal accounting for nearly 70% of its energy sources. I perform counterfactual policy simulations by imposing a carbon tax levied on fossil fuels, varying the level of the tax. I find that cutting emissions by 50% relative to a no-tax scenario entails a reduction of output by 6.5%. In contrast, under an economy in which all fuels are assumed to be equally productive across firms, as is standard in the production function and abatement costs literature (Atkinson and Luo, 2023; Hawkins-Pierot and Wagner, 2022; Shapiro and Walker, 2018; Fowlie, Reguant and Ryan, 2016; Carlson, Burtraw, Cropper and Palmer, 2000), I find that obtaining the same 50% reduction in emissions leads to a 12% reduction in output, thus overestimating the output costs of such a policy by almost 100%.

Multidimensional heterogeneity significantly reduces the predicted economic cost associated with reducing emissions because it improves the ability of the carbon price to target high-emission plants. Plants that are more productive at dirty fuels relative to cleaner fuels have a higher marginal abatement cost because they are less willing to substitute away from their most productive fuels. Consequently, they face a larger increase in their marginal cost, pass a larger portion of the tax to consumers, and become less competitive relative to plants that are more productive at using cleaner fuels. My results thus provide quantitative support for the long-established idea in environmental economics that firm-level heterogeneity in abatement costs improves the effectiveness of market-based mechanisms relative to command-and-control regulation (Goulder and Parry, 2008).

In the Indian steel context, the carbon tax generates a significant reallocation of output from plants using large coal-powered blast furnaces in Eastern India to cleaner plants in Western India. This reallocation of output is the main channel to reduce emissions. With a carbon price equivalent to $25 U.S. dollars per ton of carbon, 58% of emissions reduction come from output reallocation, 24% from an aggregate reduction in output and only 18% from plant-level fuel substitution. Large degrees of fuel specialization due to fuel productivity coupled with high fuel switching costs explain the lack of fuel substitution, capturing significant technology lock-in.

To tackle this technological lock-in, I study a subsidy that reduces the fixed cost of natural gas adoption and find that it is not a cost-effective tool to reduce emissions. I find that 90% of the beneficiaries from the subsidy are inframarginal plants who would have switched to natural gas regardless, which significantly increases the cost of incentivizing a meaningful change in natural gas adoption. I estimate the cost to increase natural gas adoption by five percentage points at the equivalent of \$2.793 billion U.S. dollars over 40 years, amounting to 12% of the industry's profits in the same periods.

Quantifying the role of fossil fuels in production in this context requires overcoming two important measurement issues. First, the energy plants use in production is unobserved because it depends on how plants use fuels. That is, the energy service that a plant receives is different from the physical quantity of fuels it uses, measured in common heating potential units. The wedge between a fuel's heating potential and the energy service it provides reflects the fuel's productivity. Second, plants choose fuel sets on the basis of unobserved heterogeneity, such as how productive they would be at using alternative fuel combinations. For instance, a plant can choose to use coal because it anticipates high coal productivity and low gas productivity. However, the plant's gas productivity remains unobserved to the researcher.

I address these measurement issues in three steps. First, I identify the latent quantity and price of energy services by adapting the methods of Grieco, Li and Zhang (2016). This method relies on optimality conditions from profit maximization to map observed relative input spending to unobserved relative input quantities. Second, I estimate the function that maps fuels to energy services following Zhang (2019) and Blundell and Bond (1998, 2000). This allows me to recover the distribution of fuel productivity across plants. Third, I adapt Arcidiacono and Jones (2003) to jointly estimate fixed switching costs and the distribution of fuel productivity for unused fuels. Using this three-step approach, I recover all production function parameters, the distribution of fuel productivity, and switching costs between fuel sets. Lastly, I estimate the elasticity of substitution between output varieties following Ganapati et al. (2020). These estimates allow me to conduct policy counterfactuals that affect plants' fuel choices.

The production model that I estimate is consistent with recent evidence suggestive of heterogeneity in fuel productivity and high fixed costs of fuel adoption. Lyubich, Shapiro and Walker (2018) find that firms vary substantially in energy and CO2 productivity. These disparities in productivity are due to varying heat efficiency that different fuel-burning technologies provide (Allcott and Greenstone, 2012), energy retrofit efforts to curb energy waste (Christensen, Francisco and Myers, 2022), unobserved fuel quality (e.g. anthracite vs. bituminous coal), and intangible factors such as the ability of workers to use different fuels and management practices (Gosnell, List and Metcalfe, 2020). As in Scott (2021), I find significant fixed costs and time commitments associated with adopting natural gas. These costs encompass technological adaptations, new storage facilities, and transportation infrastructure, all microfoundations that my approach captures.

Heterogeneity in fuel-specific productivity also provides a microfoundation for abatement cost heterogeneity because it affects plants' willingness to substitute between fuels. I thus contribute to a longstanding literature that estimates marginal abatement costs (Atkinson and Luo, 2023; Shapiro and Walker, 2018; Culler and Mansur, 2017; Carlson, Burtraw, Cropper and Palmer, 2000). These fuel-specific productivity differences, along with dynamic switching between fuel sets, also provide a cautionary tale against aggregate production functions to study fuel substitution, common in the integrated assessment literature (Miftakhova and Renoir, 2021; Golosov, Hassler, Krusell and Tsyvinsky, 2014), because such production functions will not be invariant to policy changes.

The production model also reflects important channels of firm responses to changes in fuel prices. Numerous empirical studies have shown that firms respond to changes in fuel cost by substituting across fuels (Alpino, Citino and Frigo, 2023; Ahmadi and Yamazaki, 2020; Andersson, 2019), but also by passing on the cost increase to consumers, thereby reducing output (Fontagné, Martin and Orefice, 2023; Ganapati, Shapiro and Walker, 2020; Gittens, 2020). While both types of responses can reduce emissions, their welfare implications are unequal and my model provides a cohesive framework to aggregate these responses and relate them to marginal abatement costs.

To estimate this model, I also contribute to the literature on production function estimation in industrial organization[3]. I make a methodological contribution by showing how to identify and estimate a dynamic production function with input-augmenting productivity and dynamic input selection. I solve this problem by drawing from methods in the dynamic discrete choice literature with unobserved heterogeneity (Arcidiacono and Jones, 2003; Arcidiacono and Miller, 2011).

Within the production function estimation literature, my paper most closely relates to Collard-Wexler and De Loecker (2015), who study technological change in the U.S. Steel Industry (particularly the introduction of electric arc furnaces) and Hawkins-Pierot and Wagner (2022), who estimate the energy productivity of manufacturing plants and its implication for technology lock-in. I complement the former by emphasizing the role of fuels and emissions as part of this technological change. I complement the latter by decomposing energy productivity into the relative productivity of different fuels, and I show that this distinction is crucial to understanding the heterogeneous impact of a carbon tax.

The rest of the paper is structured as follows: Section 2 presents an overview of the Indian steel dataset. Section 3 presents some key evidence of plant-level decisions that motivate modeling choices. Section 4 presents the model in detail. Section 5 presents identification results for the model. Section 6 presents the main estimation results. Finally, Section 7 presents the results of the counterfactual experiments.

[3]See footnote[2] for details on this literature.

2 Data

I use longitudinal data on prices and quantities of all inputs and outputs from Indian steel establishments. Additionally, I observe plants' locations across 700 districts, which I link to data on India's natural gas pipeline network. The panel comes from the Indian Survey of Industries (ASI) and covers 2009-2016. It is a restricted-use dataset that covers all manufacturing establishments with at least 100 workers and a representative sample of establishments with fewer than 100 workers. The sample is stratified at various levels, including number of workers and location. See Appendix A.1 for details. The ASI contains information on prices and quantities of Coal, Oil, Electricity, and Natural Gas, which I convert to million British thermal units (mmBtu) following standard practices by the U.S. Environmental Protection Agency (EPA, 2022). I remove the 1% left and right tails of plant-level inputs and output due to the presence of outliers in the ASI.[4]

To convert nominal into real units, I follow Harrison, Hyman, Martin and Nataraj (2016) by deflating output with industry-specific wholesale price indices (WPI), labor with the consumer price index (CPI), intermediate materials with the aggregate wholesale price index, labor with the consumer price index (CPI), and capital stock with an India-specific capital deflator from the Penn World Table (Feenstra, Inklaar and Timmer, 2015).

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}). During combustion, each mmBtu of fuel releases some quantity of carbon dioxide CO_2, methane CH_4, and nitrous oxide N_2O into the atmosphere, which varies by industry based on standard practices in India (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.2).

3 Facts about Emissions and Fuels in India

Using this data, I highlight facts about fuel usage and carbon emissions that motivate my choice of India's Steel sector to conduct this analysis and influence modeling decisions. Many of these facts are not specific to either India or steel.

Fact 1: High Pollution Levels from Indian Steel Establishments

In Table 1, I show that total annual greenhouse gas emissions from Indian Steel plants average 25 million tons of CO_{2e}, accounting for 31% of annual emissions in Indian manufacturing (Dhar, Pathak and Shukla, 2020). This high emission level is attributed to the sizeable aggregate share of coal as part of the energy mix, averaging 70%. This share is significantly larger than in other

[4] The ASI dataset is known to have such outliers, typically due to reporting errors that are inconsistent with a wide range of official statistics (Bollard, Klenow and Sharma, 2013).

Indian manufacturing industries and larger than in Steel manufacturing abroad. Indeed, switching from coal to gas has contributed significantly to the manufacturing clean-up in developed economies (Rehfeldt, Fleiter, Herbst and Eidelloth, 2020).

Industry	Average Annual Revenue by Plant (Million USD)	Annual Average Emissions (Thousand tons CO_{2e})	Aggregate Energy Input Share	Aggregate Coal Fuel Share
Steel	19.41	29.34	0.13	0.72
Other	6.18	8.52	0.13	0.37

Table 1: Descriptive Statistics for Steel Manufacturing (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 Steel Establishments Use Different Fuel Sets

Steel-producing plants use different fuel sets, and most fuel sets include oil and electricity. Most of the variation in fuel sets comes from whether plants use coal or natural gas. see Table 2. There are multiple reasons for this heterogeneity. Plants can use different fuel-burning technologies during the steelmaking process. To turn iron ore into iron, blast furnaces use coke (coal), whereas direct reduced iron furnaces use either coal or natural gas. To turn iron into primary steel, basic oxygen furnaces can use various fuel combinations, whereas electric arc furnaces discharge electricity through an electric arc.

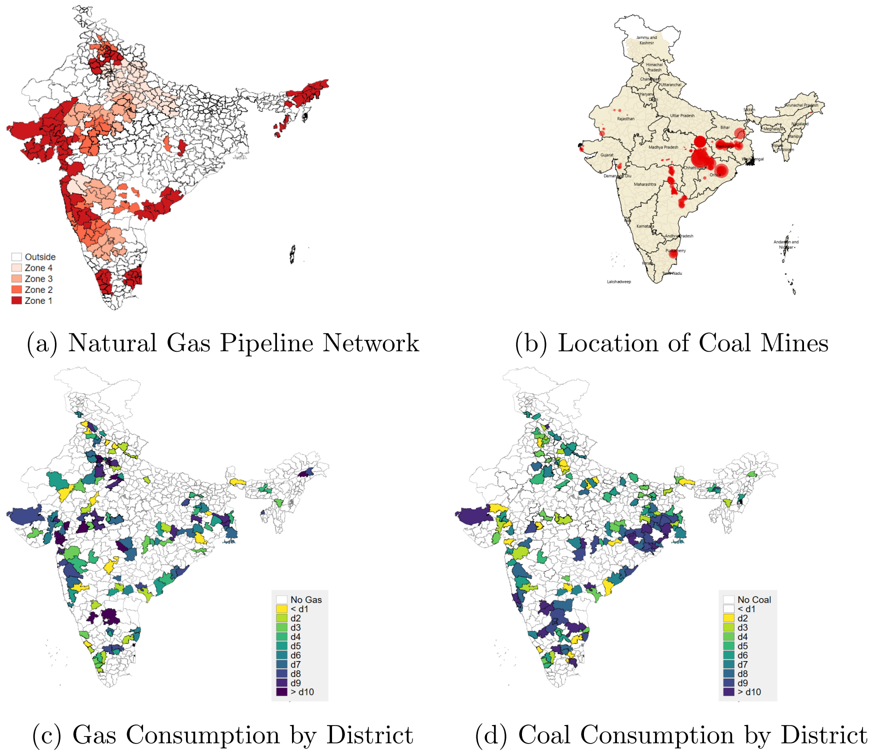
	Percentage of Plants	Output Share
Oil, Electricity	51.3	44.9
Oil, Electricity, Coal	19.3	21.2
Oil, Electricity, Gas	10.8	25.7
Oil, Electricity, Coal, Gas	7.4	3.4
Other	11.1	4.8

Table 2: Distribution of Fuel Sets Across Steel Plants

Notes: this table shows the distribution of fuel sets across plants. The “Other” category comprises any other combinations of the same four fuels. Variation in fuel sets is not driven by variation in the variety of steel produced. In Appendix A.3, I show that a similar distribution exist within different varieties of steel produced.

Many reasons explain why plants use different fuel sets, some of which have geographical underpinnings. In Figure 1, I document a concentration of coal usage in Eastern and Southern India and a concentration of natural gas usage in Eastern India. Many states in Eastern India form a region colloquially known as the “Steel Belt”, due to the prevalence of coal and iron ore mines. Plants with coal-powered Blast furnaces tend to be located near these mines. Moreover, this region also lacks critical infrastructure, such as natural gas pipelines for plants to adopt technologies that rely on cleaner fuels.

This heterogeneity in fuel sets has two other noteworthy implications. First, burning coal releases more CO_{2e} than burning natural gas. Second, plants with more fuels have access to additional substitution margins, allowing them to better hedge against fuel price volatility owing to global



(a) Natural Gas Pipeline Network (b) Location of Coal Mines

(c) Gas Consumption by District (d) Coal Consumption by District

Figure 1: Spatial Variation in Key Infrastructures and Fuel Usage

Notes: the maps on the bottom show the distribution of (log) natural gas and coal consumption across Indian districts. Each shade corresponds to a decile. Darker areas correspond to districts with more fuel consumption. The maps on the top show the distribution of key infrastructures: natural gas pipelines (left) and coal mines (right). Natural gas pipeline transportation tariffs are organized by zones, where each zone corresponds to 250km segments along a pipeline. Zone 1 is the closest to the source of the pipeline. Data on coal mines was taken from Pai and Zerriffi (2021).

supply and demand fluctuations. To see this, I provide correlational evidence that plants' relative fuel quantities respond to changes in relative fuel quantities with the following specifications:

		$\Delta \ln p$		
$\Delta \ln e$	Coal/Oil	-1.12***	-1.25***	-1.23***
		(0.13)	(0.14)	(0.14)
	Coal/Elec	-0.90***	-0.92***	-0.89***
		(0.07)	(0.07)	(0.07)
	Coal/Gas	-1.11***	-1.1***	-1.10***
		(0.08)	(0.10)	(0.08)
	Gas/Oil	-0.96***	-0.96***	-1.00***
		(0.05)	(0.08)	(0.08)
	Gas/Elec	-0.95***	-0.96***	-0.97***
		(0.05)	(0.05)	(0.05)
	Oil/Elec	-0.71***	-0.67***	-0.96***
		(0.10)	(0.16)	(0.16)
Control for Δ capital		Y	Y	Y
Year FE			Y	Y
Controls for Δ labor and materials				Y

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

Table 3: Preliminary Evidence of Fuel Substitution

Notes: this table was constructed from the following regression specification of changes in relative log fuel prices $\Delta \ln p$ on changes in relative log fuel quantities $\Delta \ln e$. For any two fuels $f \neq j$: $\Delta \ln(e_{fit}/e_{jit}) = \beta_0 + \beta_1 \Delta \ln(p_{fit}/p_{jit}) + X_{it}^t \beta + u_{it}$. Where i indexes plants and t indexes years.

Fact 3: Indian Steel Establishments Often Switch Between Fuel Sets

Not only is there heterogeneity in fuel sets but also a significant prevalence of switching between fuel sets, which occurs when a plant uses a different combination of fuels between two years (e.g. from oil and electricity to oil, electricity, and gas). On average, 15% of plants add a new fuel, and 15% drop an existing fuel from their set every year. Moreover, 40% of unique plants add and drop fuel at least once in the sample. Steel plants typically switch fuels by attaching different burners to existing furnaces. There are many reasons why plants want to switch. The development of new technologies may increase the heat efficiency of some fuels. Large and persistent fuel price shocks incentivize plants to readjust their fuel mix. Expanding transportation infrastructures, particularly pipeline networks, decreases fixed costs and eases access to new fuels (Scott, 2021).

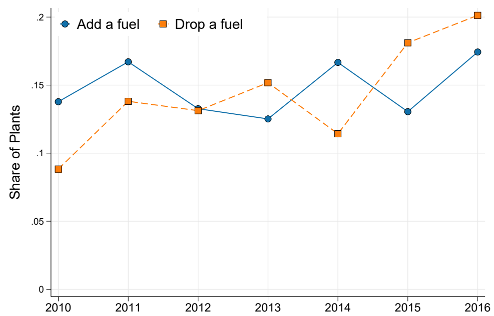


Figure 2: Fuel Set Switching Across Years

Notes: this figure shows the fraction of plants that add a new fuel or drop an existing fuel between year t and $t-1$, starting from 2010 all the way until 2016.

	Adds New Fuel	Drop Existing Fuel	Both Add and Drop
Yes (%)	39.4	39.6	26.0

Table 4: Fraction of Unique Plants that Add and Drop a Fuel At least Once

Notes: this table shows the fraction of unique plants that add a fuel at least once in the sample, and similarly for plants that drop a fuel at least once. This is an underestimate of the prevalence of fuel switching because plants are only observed between 2009 and 2016.

4 Model

Consistent with these facts, I develop and estimate a rich dynamic production model to quantify plants' fuel choices. 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 an outer nest of production. Plants can choose to change fuel sets across periods in a dynamic discrete choice framework. There are fixed costs for adding new fuels and salvage values from dropping existing ones. I first present the production structure 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.

4.1 Production Model

There are two levels of production, which correspond to two nests. The outer nest is a CES production function and features Hicks-neutral productivity z_{it}, labor L_{it}, capital K_{it}, intermediate inputs M_{it}, and energy E_{it}. Following Grieco et al. (2016), the production function is explicitly normalized around the geometric mean of each variable $\overline{X} = \Big(\prod_{i=1}^{n}\prod_{t=1}^{T} X_{it}\Big)^{\frac{1}{nT}}$.[5]

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

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 include energy, E_{it}. Given the current fuel set $\mathcal{F}_{it} \subseteq \mathbb{F} = \{\text{oil, gas, coal, elec}\}$, plants combine all fuels available to produce a quantity of energy E_{it} in the inner nest of production:

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

e_{fit} refers to the quantity of fuel f for plant i in year t. p_{fit} and ψ_{fit} are the corresponding fuel price and productivity, respectively. The fuel-specific productivity terms are novel; they allow for flexible variation in input usage and heterogeneity in fuel substitution. Plants specialize in fuels that they can use more efficiently, which means that different plants are affected differently by changes in fuel prices. This heterogeneity is especially relevant in the context of a carbon tax, which raises prices of dirty fuels more than clean fuels. For example, plants specializing in dirty fuels such as coal will bear a disproportionate share of the tax burden. While initially unobserved, I recover fuel productivity for each plant each year by exploiting profit maximization.

Moreover, allowing plants to have different fuel sets $\mathcal{F}_{it}$ and to switch between them is also novel; plants with a larger fuel set have more substitution possibilities when facing changes in fuel prices. This option value creates another layer of heterogeneity in response to carbon taxation. Larger and more productive plants with a larger fuel set will have an easier time substituting out of a carbon tax than smaller plants with smaller fuel sets.

These two novel features are a significant departure from the literature, where most previous

[5] All CES functions are either implicitly or explicitly normalized around a point (León-Ledesma, McAdam and Willman, 2010). I chose the geometric mean as a normalization point to be consistent with the literature.

papers that estimate a production function with fuels do not allow for fuel-specific productivity and do not allow for fuel sets to vary within the same production function (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.

The elasticity of substitution between fuels λ plays a crucial role in this model. It determines the option value a plant gets by expanding its fuel set $\mathcal{F}_{it}$. As long as fuels are gross substitutes ($\lambda > 1$), there is an option value to have more fuels due to additional substitution margins. However, the lower λ is, the larger the option value. 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[6]. In Online Appendix B.1, I show that this option value helps plants hedge against fuel price shocks and quantity shortages. 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.*

On the demand side, there is a representative consumer with quasi-linear utility over the total output produced Y_t and an outside good Y_{0t}. Steel consumption is widespread in India, and the majority of demand comes from housing construction, infrastructure, and automobiles. Total output is produced by aggregating all the varieties with standard Dixit-Stiglitz preferences. Steel plants produce a wide variety of steel products. There are 404 varieties produced by plants in the ASI, none of which has a disproportionate market share. *Ferrous products from direct reduction of iron ore* is the most common variety with a 5.5% market share. See the online appendix for details. Given a mass of N_t operating plants, income I_t, and an aggregate demand shock e^{Γ_t}, the representative consumer solves:

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

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

[6] This option value is similar to the concept of gains from variety in the trade literature investigating the composition of intermediate inputs (Ramanarayanan, 2020; Goldberg, Khandelwal, Pavcnik and Topalova, 2010; Kasahara and Rodrigue, 2008; Broda and Weinstein, 2006; Romer, 1990; Ethier, 1982)

Helpman and Itskhoki (2010), I restrict $\theta < \frac{\rho-1}{\rho}$, which ensures that output varieties are more substitutable between each other than with the outside good. These quasi-linear CES preferences were first proposed by Helpman and Itskhoki (2010) and provide analytical convenience for welfare evaluation. Quasi-linear preferences are standard in the literature on externality taxation (Fowlie et al., 2016). In this particular instance, it allows me to use the social cost of carbon (SCC) to quantify externality damages, expressed as the net present value of future climate damages from emissions.[7] 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 sloping demand for each variety Y_{it}, which I augment with an idiosyncratic ex-post 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. In Appendix C.1, I show robustness checks on this demand specification by allowing for nested CES demand across various categories of steel varieties.

4.2.1 Static Profit Maximization

Given a set of fuels $\mathcal{F}_{it} \subseteq \mathbb{F}$, technological constraints, inverse demand, and all input prices, the plant chooses input quantities that maximize static profits.[8] 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.

$$\max_{K_{it}, M_{it}, L_{it}, \{e_{fit}\}_{f\in\mathcal{F}_{it}}} \left\{ P_{it}(Y_{it})Y_{it} - w_t L_{it} - r_{kt}K_{it} - p_{mit}M_{it} - \sum_{f\in\mathcal{F}_{it}} p_{fit}e_{fit} \right\}$$

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

$$P_{it}(Y_{it}) = \left(\frac{e^{\tilde{\Gamma}_t}}{N_t Y_{it}}\right)^{\frac{1}{\rho}} P_t^{\frac{1+\rho(\theta-1)}{(\theta-1)\rho}}$$

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

[8] I derive the decision of plants under the assumption that plants flexibly rent capital with a 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 nested structure of production is such that it can be expressed in two stages:

1. Plants choose fuel quantities to minimize the cost of producing energy (inner nest):

Given a fuel set $\mathcal{F}_{it}$ and fuel prices, plants find the combination of fuels that minimizes the cost of producing a given unit of energy. Fuel prices in mmBtu are observed and vary across plants and years:

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

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

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

Where the unobserved price of realized energy $\tilde{p}_{E_{it}}$ corresponds to a CES price index in fuel prices over productivity. Constant returns in the energy production function imply 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. Plants choose inputs to maximize static 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 the quantity of energy because p_{Eit} is only a function of the optimal *relative* allocation of fuels, not the scale of energy. 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 fuel $\{\psi_{fit}\}_{f\in\mathcal{F}_{it}}$, and all input prices $\{w_{it}, r_{kit}, p_{mit}, \{p_{fit}\}_{f\in\mathcal{F}_{it}}\}$. Together with the years 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 expectations over the evolution of state variables and choose a fuel set for the 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. I allow fuel set switching 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[9]:

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

The switching cost function $k(\mathcal{F}' \mid \mathcal{F}_{it}, d_i)$ in Table 5 is composed of two types of arguments. First, there are fixed costs of adding a fuel κ_f. Second, there are salvage values that plants obtain from dropping a fuel γ_f. Salvage values are not restricted to be positive. Since 90% of plants in the dataset always use electricity and oil, I assume that the choice set of plants is as follows, where e = electricity, o = oil, g = gas, c = coal: $\mathbb{F} = \Big\{(oe); (oge); (oce); (ogce)\Big\}$ and restrict the sample accordingly. In the next Section, I show how this model can be estimated.

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

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

Notes: rows correspond to fuel sets today $\mathcal{F}$, whereas columns correspond to fuel sets next period $\mathcal{F}'$. I allow fixed costs for natural gas to vary by plants' proximity to the natural gas pipeline network in a binary fashion, where $d = 0$ if plants have no access to pipelines and $d = 1$ if plants have access to pipelines. I define plants as 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.

[9] Plant size is endogenous, but a Ceteris paribus increase in z_{it} increases the scale of a plant's operation. Scott (2021) shows that proximity to the natural gas pipeline network decreases the fixed cost of adding natural gas. Plants too far from the pipeline network can use liquified natural gas (LNG) but need access to a gasification terminal, which can be very costly. In Online Appendix A.3.3, I show that plants that experienced an expansion of the pipeline network in their district between 2008 and 2016 were more likely to add natural gas to their mix.

5 Identification and Estimation

The model is estimated in four steps using a novel combination of methods. First, I use the cost-shifter of Ganapati et al. (2020) as an instrument to estimate demand. Second, I adapt the method of Grieco et al. (2016) in order to estimate the outer production function in the presence of an unobserved input (energy). Third, I estimate the energy production function following recent developments in production function estimation with input-augmenting productivity (Demirer, 2020; Zhang, 2019; Doraszelski and Jaumandreu, 2018), combined with dynamic panel methods (Blundell and Bond, 1998, 2000, 2023). Fourth, I estimate fixed costs in a dynamic discrete choice framework in the presence of unobserved heterogeneity following Arcidiacono and Jones (2003), allowing me to capture systematic differences in fuel productivity for fuels that plants are not currently using. Splitting the estimation into four subsequent stages allows for tractability and does not necessitate that I impose subsequent assumptions on each previous stage. As such, while capital is assumed to be rented flexibly to solve firms' dynamic discrete choice problem, I do not make any assumption on capital when estimating the production function.

5.1 Estimating Elasticity of Demand

I estimate demand from observed output prices and quantities using the demand equation (4).[10]

$$\ln Y_{it} = \Lambda_t - \rho \ln P_{it} + \epsilon_{it}, \tag{7}$$

where $\Lambda_t = \tilde{\Gamma}_t + \ln\left(\frac{1}{N_t}\right) + \frac{\rho(1-\theta)-1}{1-\theta} \ln P_t$ contains both the unobserved aggregate output price index P_t and aggregate demand shocks $\tilde{\Gamma}_t$. 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 a pre-shock variation in exposure to aggregate fuel prices by Indian States (the share):

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

$\overline{p}_{-s,t,f}$ 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

[10] In Appendix C.1, I explore more flexible demand specifications. Since plants produce various steel varieties that can be categorized, I allow for a nested CES demand specification allowing for different elasticities of substitution within and across categories. I find that results are quantitatively very similar to the simple CES specification.

demand and supply variation induced by geopolitical turmoil, aggregate technological evolution, and growth. $\sigma_{s,2008,f}$ is the pre-sample aggregate share of fuel f used to generate electricity in state s. Variation in the price of a fuel will induce more variation in electricity prices in states that use more of that fuel to generate electricity. This creates exogenous variation in 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 (7) and is endogenously determined by the elasticity of demand ρ. I first estimate demand using year dummies Λ_t and then solve for the price index ex-post given the estimate of $\hat{\rho}$, observed output prices P_{it} and the number of plants N_t, $P_t = \left(\frac{1}{N_t}\sum_{i=1:N_t} P_{it}^{1-\hat{\rho}}\right)^{1/(1-\hat{\rho})}$. I then separately recover the elasticity of the outside good θ from the aggregate demand shifter $\tilde{\Gamma}_t$ in a simple time series regression of the year dummies Λ_t on the output price index and a constant.

5.2 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 the heating potential of fuels, energy is the output of combining different fuels, which is unobserved. Under the assumption that plants are price-takers in the input market, I adapt the estimation method proposed by Grieco et al. (2016) to uniquely recover the price and quantity of energy when at least one flexible input (labor) is observed.[11] The key to this method relies on using relative first-order conditions to map observed expenditure shares to unobserved input quantity shares. To see this, one can rearrange the ratio of first-order conditions for labor and energy from profit maximization in equation 6:

$$\underbrace{\frac{w_{it}L_{it}}{p_{E_{it}}E_{it}}}_{\text{Expenditure ratio}} = \frac{\alpha_L}{\alpha_E}\Bigg(\underbrace{\frac{L_{it}/\overline{L}}{E_{it}/\overline{E}}}_{\text{Quantity ratio}}\Bigg)^{(\sigma-1)/\sigma} \tag{8}$$

Given production function parameters, $\frac{E_{it}}{\overline{E}}$ can be recovered from (8) 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 the quantity of labor. Identification 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

[11] 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.

$\sigma = 1$ (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.

$$\frac{E_{it}}{\overline{E}} = \Big(\frac{p_{eit}E_{it}}{w_{it}L_{it}}\Big)^{\frac{\sigma}{\sigma-1}}\Big(\frac{\alpha_L}{\alpha_E}\Big)^{\frac{\sigma}{\sigma-1}}\frac{L_{it}}{\overline{L}} \tag{9}$$

Thus, in this setting, one can identify production parameters by replacing $\tilde{E}_{it}$ for (9) in the production function and exploiting the labor first-order condition 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). I also use the same method to control for unobserved price dispersion in the bundle of material inputs.

The main dependent variable is revenues, where $e^{u_{it}}$ is an unobserved iid shock that 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.2. Taking logs of revenues yields the main estimating equation:

$$\ln R_{it} = \ln\frac{\hat{\rho}}{\hat{\rho}-1} + \ln\frac{1}{\eta} + \ln\left[w_{it}L_{it}\Big(1+\frac{\alpha_K}{\alpha_L}\Big(\frac{K_{it}/\overline{K}}{L_{it}/\overline{L}}\Big)^{\frac{\sigma-1}{\sigma}}\Big) + p_{mit}M_{it} + p_{E_{it}}E_{it}\right] + u_{it} \tag{10}$$

The main parameters of interest are the elasticity of substitution (σ) and the returns to scale (η) in (10). The former is identified from observed variation in the capital-to-labor ratio. The latter is identified because the elasticity of demand was estimated separately in the previous stage. Lastly, since $\tilde{E}_{it}$ and $\tilde{M}_{it}$ were factored out of the production function, the main estimating equation (10) 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 (8) for energy and labor and likewise for materials and labor.[12]

$$\begin{aligned} &\overline{wL}/\overline{p_E E} = \frac{\alpha_L}{\alpha_E}; \qquad \overline{wL}/\overline{p_m M} = \frac{\alpha_M}{\alpha_E} \\ &\alpha_K + \alpha_L + \alpha_M + \alpha_E = 1 \end{aligned} \tag{11}$$

Then, I estimate (10) subject to (11) with non-linear least squares.[13]

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

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

5.3 Identification of inner production function for energy

The energy production function in equation (2) can be rewritten by factoring 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 stage. I also observe fuel quantities, $\{e_{fit}\}_{f \in \mathcal{F}_{it}}$, and fuel prices: $\{p_{fit} = \frac{s_{fit}}{e_{fit}}\}_{f \in \mathcal{F}_{it}}$. I show how to recover the elasticity of substitution λ, and all productivity terms for fuels that plants are using $\{\psi_{fit}\}_{f \in \mathcal{F}_{it}}$. 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 effectively combines the dynamic panel approach of Blundell and Bond (2000, 1998) with the method proposed by Zhang (2019).

Relative first-order conditions from the cost-minimization in (5) identify the productivity of fuel f relative to electricity 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 straightforward: relative fuel productivities equate relative fuel prices 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 substituting back the implied relative fuel productivity terms (13) into the energy production function (12) and rearranging:

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

Where $s_{fit} \equiv p_{fit} e_{fit}$ is spending on fuel f. The intuition underlying equation 14 is fairly straightforward. The left-hand side is the value added of an additional unit of electricity in terms of energy, while the right-hand side is the contribution of electricity productivity and relative spending on other fuels to that value added. Naturally, higher electricity productivity increases the value added of electricity, and higher spending on other fuels also increases the quantity of energy produced for a given unit of electricity. 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 plants when choosing fuel quantities. To deal with this issue, I assume that the productivity of electricity follows an AR(1) Markov process with year and plant fixed effects.[14]

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

I then take the 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 the productivity of electricity at time t, it is uncorrelated with choices made at time $t-1$:

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

There are two main endogeneity concerns in this model. First, the lagged value added of electricity and the lagged relative spending on other fuels are correlated with the plant fixed effect μ_i^*, which biases the persistence of electricity productivity ρ_{ψ_e}. This is the standard concern in the dynamic panel literature. Second, contemporaneous relative spending on other fuels is 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. I use the system GMM approach, which combines both level and difference moment conditions as follows:

[14] 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 variable observed in the data corresponds to the unconditional average of this process. This means that even though the model is estimated from a short panel (between 2 and 8 years, depending on the plant), 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 follow an AR(1) process rather than electricity productivity itself.

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

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 persistently affecting fuel prices in the market. Lastly, I get standard errors on the elasticity of substitution using the delta method.

5.4 Identification and Estimation of Fixed Fuel Switching Costs

Each plant has 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 start from the recursive formulation of the problem. The plant chooses a fuel set next period $\mathcal{F}'$ to maximize the net present value of lifetime profits:

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

Where the fuel set switching cost function, $\mathcal{K}(\mathcal{F}' \mid \mathcal{F}_{it}, s_{it})$, was defined in Table 5. σ_ϵ is a parameter that maps units of the fixed cost shocks to units of profits (dollars). From now on, I define the parameters governing the switching cost function $\theta_1 = \{\kappa_{g1}, \kappa_{g0}, \kappa_c, \gamma_g, \gamma_c\}$ 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. Cost shocks are assumed to be iid and come from a standardized Type 1 Extreme value $\epsilon_{\mathcal{F}'it} \sim$ *Gumbel*$(0, 1)$. This allows me to analytically integrate these shocks and work with the expected value function, $W(s_{it}, \mathcal{F}_{it}) = \mathbb{E}(V(s_{it}, \epsilon_{it}, \mathcal{F}_{it}))$:

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

Where γ 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. Next, I discuss the evolution of each state variable.

$$Pr(\mathcal{F}' \mid \mathcal{F}_{it}, s_{it}; \theta_1, \theta_2) = \frac{exp\Big(\upsilon_{\mathcal{F}'}(s_{it}, \mathcal{F}_{it}; \theta_1, \theta_2)\Big)^{1/\sigma_\epsilon}}{\sum_{\mathcal{F} \in \mathbb{F}} \exp\Big(\upsilon_{\mathcal{F}}(s_{it}, \mathcal{F}_{it}; \theta_1, \theta_2)\Big)^{1/\sigma_\epsilon}}$$

Plants take expectation over all productivity terms, fuel prices and material prices. I separate state variables 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 $(\psi_{cit}, \psi_{git}, p_{cit}, p_{git})$. I assume that plants do not take expectation over the rental rate of capital and aggregate wages to reduce the computational burden of this problem.

Evolution of state variables

All state variables follow a persistent AR(1) process with time (t) and plant (i) fixed effects. To reduce the state space, I assume that the persistence for fuel prices and productivity is the same $\rho_{p_f} = \rho_{\psi_f} = \rho_f$. This assumption allows me to define fuel productivity over prices as a single state variable, which always enters together in plants' profit function p_{fit}/ψ_{fit} through the energy price index. I refer to plant fixed effects in fuel productivity/prices as *fuel comparative advantage*, denoted by μ_i^f.

$$\mathbb{E}\left[\ln(\psi_{fit+1}/p_{fit+1}) \mid \mathcal{I}_{it}\right] = (1-\rho_f)(\mu_0^f + \mu_t^f + \mu_i^f) + \rho_f \ln(\psi_{fit}/p_{fit}) \tag{18}$$

In this context, A positive (log) fuel price shock is isomorphic to a negative (log) fuel productivity shock, and vice versa. I also assume a similar process for hicks-neutral productivity $\ln z_{it}$ and for the (log) price of materials $\ln p_{mit}$. These two state variables, together with the price/productivity of oil and electricity were recovered in previous sections for all plants in all years. Their Markovian processes can be estimated directly using the system GMM approach of Blundell and Bond (2000). Following Bonhomme and Manresa (2015), I reduce the dimension of plant fixed effects $\mu_{\mathbf{i}} = \{\mu_i^z, \mu_i^m, \mu_i^o, \mu_i^e\}$ by grouping plants using K-means clustering.

For coal and natural gas, the evolution of the price/productivity process depends on whether plants are initially using coal/gas. If plants are not using coal/gas at t, they start at the initial condition (equivalent to $t = 0$) and take expectation over idiosyncratic shocks. $\forall f \in \{coal, gas\}$:

$$\mathbb{E}\left[\ln(\psi_{fit+1}/p_{fit+1}) \mid f \notin \mathcal{F}_{it}\right] = \mu_0^f + \mu_t^f + \mu_i^f$$

Estimating the distribution of plant-specific comparative advantages for coal and gas only from plants who currently use gas or coal is likely to be biased due to selection. Recovering the distribution of μ_i^f from selected plants may not reflect the distribution across all plants, which would bias fixed costs estimates. In the next section, I show how to recover the distribution of these fuel comparative advantages for plants that are not currently using gas or coal, jointly with fixed costs following the approach of Arcidiacono and Jones (2003). Using this approach, I recover the distribution of fuel comparative advantages that is most likely to rationalize observed fuel set choices, under the assumption that plants know about their comparative advantage and use that information to make decisions. Lastly, I allow all shocks to state variables to be arbitrarily correlated in a multivariate normal distribution with mean zero and a positive semi-definite covariance matrix $\mathbf{\Sigma}$. I then discretize the entire state space following Farmer and Akira Toda (2017).

$$\left(\epsilon_{it}^o, \epsilon_{it}^e, \epsilon_{it}^g, \epsilon_{it}^c, \epsilon_{zit}, \epsilon_{mit}\right) \equiv \epsilon_{\mathbf{it}} \sim \mathcal{N}(\mathbf{0}, \mathbf{\Sigma})$$

Estimating Fixed Costs and Fuel Comparative Advantages

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). I assume that the distribution of comparative advantages comes from 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 $(\tilde{\mu}_f, \tilde{\sigma}^2_{\mu_f})$:

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

Where $\pi_{fk}^0 = Pr(\mu_{fk})$ is the unconditional probability of being type k, where types refer to support points of the fuel comparative advantage distribution, and $\sum_k \pi_{fk}^0 = 1$. In this context, external estimation of parameters governing the distribution of random effects from a selected sample of plants that 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 a larger comparative advantage for 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 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 $\pi_{kc} \in \Pi_c$:

$$\ln \mathcal{L}(\mathcal{F}, s \mid \theta_1, \theta_2) = \sum_{i=1}^{n} \ln \left[\sum_k \pi_k \left[\prod_{t=1}^{T} Pr(\mathcal{F}_{it+1} \mid \mathcal{F}_{it}, s_{it}, \mu_i = \mu_k; \theta_1, \theta_2) \right] \right] + \sum_{i=1}^{n} \sum_{t=1}^{T} \ln f(s_{it} \mid s_{it-1}; \theta_2) \tag{19}$$

In particular, the likelihood in (19) assumes that the state transitions are independent of the distribution of comparative advantages for coal and gas.[15] This is possible if the parameter estimates $\hat{\theta}_2$ are unbiased from selected data. In Online Appendix C.2, I show Monte-Carlo simulation results 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) provides a method to recover the unselected distribution by sequentially iterating 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. Following Bayes' 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:

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

Where $\mathcal{F}_i$ is the sequence of fuel set choices I observe establishment i making. Using Bayes' rule, the conditional probability 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:

$$\rho(\mu_k \mid \mathcal{F}_i, s_i; \theta_1, \theta_2, \hat{\pi}) = \frac{\hat{\pi}_k \left[\prod_{t=1}^{T} \left[\prod_{\mathcal{F} \subseteq \mathbb{F}} \left[Pr(\mathcal{F}_{it} \mid s_{it}, \mu_i = \mu_k; \theta_1, \theta_2) \right]^{\mathbb{I}(\mathcal{F}_{it} = \mathcal{F})} \right] \right]}{\sum_k \hat{\pi}_k \left[\prod_{t=1}^{T} \left[\prod_{\mathcal{F} \subseteq \mathbb{F}} \left[Pr(\mathcal{F}_{it} \mid s_{it}, \mu_i = \mu_k; \theta_1, \theta_2) \right]^{\mathbb{I}(\mathcal{F}_{it} = \mathcal{F})} \right] \right]} \tag{20}$$

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

The idea underlying the EM algorithm is to iteratively estimate fixed cost 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 (20), which are used to update the unconditional distribution of comparative – E step, and repeat this procedure until the likelihood in (19) is minimized. Details of the algorithm can be found in Appendix C.4.

6 Estimation Results – Steel Manufacturing

Outer Production function estimation results

Estimates of the outer production function parameters can be found in Table 6. The average output and revenue elasticities with respect to intermediate materials are much larger than those with respect to other inputs and are consistent with the literature (Gandhi et al., 2020; Grieco et al., 2016; Doraszelski and Jaumandreu, 2013, 2018). This is primarily due to the importance of iron ore in steel production. Average output and revenue elasticities are considerably larger for energy than labor and capital due to the large quantities of fuels required to produce steel. 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 9, which I use to estimate the energy production function.

Table 6: Outer Production Function Estimation

Production and Demand Parameters		Average Output Elasticities		Average Revenue Elasticities
Elasticity of substitution $\hat{\sigma}$	1.87 [1.400,3.883]	Labor	0.039 [0.036,0.046]	0.030 [0.029,0.031]
Returns to scale $\hat{\eta}$	1.19 [1.098,1.409]	Capital	0.021 [0.012,0.032]	0.016 [0.009,0.024]
Elasticity of demand $\hat{\rho}$	4.20 [2.788,5.675]	Materials	0.978 [0.902,1.160]	0.745 [0.742,0.749]
"Effective" returns to scale	0.91 [0.89,0.91]	Energy	0.150 [0.139,0.178]	0.115 [0.113,0.117]
Elasticity of outside good $\hat{\theta}$	0.66 [0.559,0.713]			
Observations	8,554			

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 $x_{jit} \in \{L_{it}, K_{it}, M_{it}, E_{it}\}$. "Effective" returns to scale capture the net curvature in the production function, taking into account downward slopping demand.

Energy production function estimation results

Turning to the energy production function, results indicate that the elasticity of substitution between fuels $\hat{\lambda}$ is larger than the elasticity of substitution between energy and non-energy inputs $\hat{\sigma}$ from Table 6. This is important because the larger the elasticity of substitution between fuels, the

Table 7: Estimates of Energy Production Function

	Steel
Elasticity of substitution $\hat{\lambda}$	2.367^{***}
	(0.269)
Persistence of electricity productivity $\hat{\rho}_{\psi_e}$	0.659^{***}
	(0.117)
Observations	3,482

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

Notes: 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 is lower than in the outer production function. 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).

larger the aggregate gains from carbon taxation (Acemoglu, Aghion, Bursztyn and Hemous, 2012). More substitution possibilities mean more emission reduction can be achieved by substituting away from polluting fuels rather than by reducing output, which is a key trade-off in evaluating carbon policy. Next, I discuss estimates of fixed costs and study the implication of these results for fuel substitution in the short and long run.

Estimation of Fixed Costs

Fixed costs are reported in Table 8. 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. These costs are substantial, ranging from 40 to 60 million dollars, and align well with the upper echelon of existing accounting estimates.[16]

Coal adoption ranges from 40% cheaper to 10% more expensive than gas adoption. This is because plants without access to high-pressure natural gas pipelines incur 50% higher adoption costs due to the need for costly alternative transportation methods, such as liquefied natural gas (LNG). This effect of pipeline accessibility is consistent with findings from Scott (2021) in his study of U.S. power plants. The observed salvage values for coal and natural gas are multiple orders of magnitude lower than fixed costs. While fixed costs are nominally very large, the role of plant size is relatively small, as raising productivity by 1% only leads to an $1,100 increase in fixed costs and a $10,040 decrease in salvage values. Importantly, the combination of substantial fixed costs and

[16]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 a coal-powered blast furnace, to direct reduced iron, typically produced with gas or coal-powered oxygen furnaces would historically cost upwards of USD 70 millions Miller (1976).

		Fixed Costs (Million USD)	Salvage Values (Million USD)
Natural Gas	*Pipeline Access*	39.5	19.3
	No Pipeline Access	59.9	
	Coal	43.6	9.2
Total Factor Productivity (100 % increase)		0.11	-1.04
Observations		2,460	

Table 8: 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 removed the last year of observation since I don't observe subsequent fuel set choices.

relatively low salvage values likely contributes to situations of technology lock-in, which I discuss in the next section.

6.1 Selection Bias in Fuel Productivity – Evidence of Technology Lock-in

The problem of technology lock-in is pervasive, as the Indian Ministry of Steel 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)

To understand factors that prevent this transition, I revisit the distribution of fuel productivity by taking into account selection bias in the distribution of fuel comparative advantages. I find significant evidence of selection bias for both coal and natural gas. Indeed, plants that do not use natural gas would be 30% less productive at it relative to plants using natural gas, whereas this effect goes up to 80% for coal. Combined with high fixed costs, this productivity gap undermines switching from coal to natural gas and exacerbates technology lock-in. This is because plants that do not currently use natural gas have less to gain from paying the fixed costs, whereas plants that currently use coal have little to gain from dropping coal.

In Online Appendix D.1-D.2, I further strengthen this argument by showing that plants with more fuels in their set tend to face a lower marginal cost of energy. This difference is largely explained by existing variation in fuel productivity rather than the additional substitution margin that a new fuel provides. In such a context, plants have little incentive to pay the large fixed costs required to add natural gas.

Spatial differences also explain much of the technological lock-in in the industry. As discussed previously with Figure 1, may plants using large coal-powered blast furnaces are located in the "Steel Belt" near coal and iron ore mines in Eastern India. At the same time, the natural gas

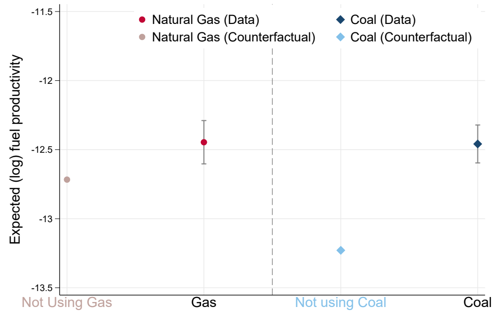
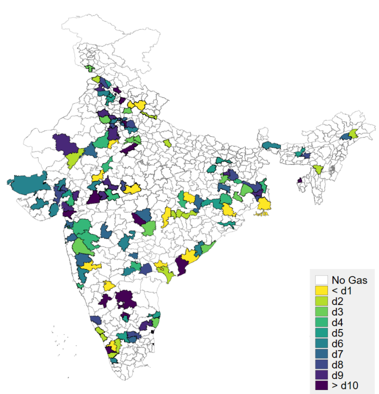


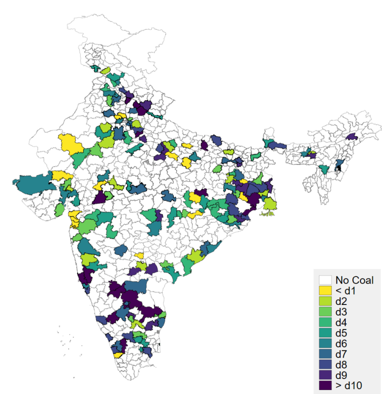
Figure 3: Distribution of fuel productivity – Including counterfactual fuel sets

Notes: This figure shows the distribution of fuel productivity per mmBtu ($\ln \psi_{fit}/\bar{e}_f$). For plants using each fuel, it includes the 95% confidence intervals around the mean. For plants not using each fuel, I compute the distribution of fuel productivity by simulating draws from the estimated distribution of unobserved heterogeneity (*comparative advantages*) in the dynamic discrete choice model, using the conditional probability distribution $\rho(\mu_k \mid \mathcal{F}_i, s_i; \hat{\theta}_1, \hat{\theta}_2, \hat{\pi})$.

pipeline network is developed in Western India but undedeveloped in Eastern India. While Table 8 showed that the direct cost of natural gas increases by 50% without pipeline access, Figure 4 suggests that the opportunity cost for large coal users in Eastern and Southern India would also be very large due to their high comparative advantage at using coal relative to natural gas. In the next section, I discuss the critical role that technology lock-in plays in understanding the effect of a carbon tax.



(a) Natural Gas



(b) Coal

Figure 4: Spatial Distribution of Estimated Fuel Productivity

Notes: the figures plot the distribution of (log) productivity for coal and natural gas across Indian districts. Each shade corresponds to a decile. Darker shades are regions in which plants are more productive at using each fuel. This distribution only includes plants using each fuel and excludes counterfactual productivity estimates based on the distribution of comparative advantages.

Model fit

Overall, the estimates of switching costs allow the model to predict quite well the empirical distribution of fuel set choices and the observed transition patterns between fuel sets. The model does slightly worse at predicting the transitions for plants that start with all four fuels because it only represents 8% of the sample. The blue bars (model) are constructed in all figures below by adding the predicted probability that each plant uses each fuel set, integrated over the conditional distribution of comparative advantages. Details are given in Online Appendix D.1.

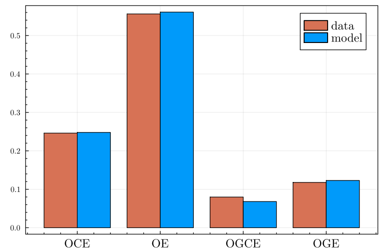
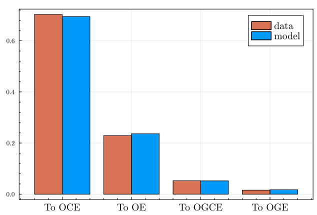
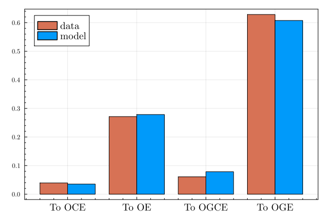


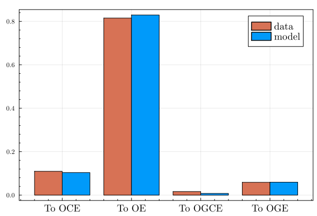
Figure 5: Unconditional distribution of fuel sets, model vs. data ($N = 2,460$)



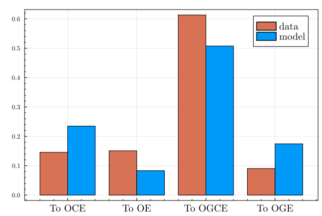
From Oil, Coal, Electricity (OCE) – $N = 584$



From Oil, Gas, Electricity (OGE) – $N = 294$



From Oil, Electricity (OE) – $N = 1,371$



From Oil, Gas, Coal, Electricity (OGCE) – $N = 211$

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

7 Externality Mitigation Policies

In this section, I study the effectiveness of various policies in mitigating externality damages from fuel combustion to improve social welfare. I detail how externality damages are constructed and perform two counterfactual policy experiments. First, I quantify the trade-off between emission reduction and output for various levels of fossil fuel taxes, where the tax rate is proportional to each fuel's emission intensity (carbon tax). I compare this trade-off with an economy without heterogeneity in fuel productivity. Second, I discuss the pervasiveness of technology lock-in in this economy and evaluate a potential solution that uses proceeds from the carbon tax to finance a subsidy that reduces the fixed cost of natural gas adoption.

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 U.S. EPA. Then, each unit of fuel f's potential energy 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 intensities calculated using the global warming potential (GWP) method detailed in Appendix A.2. 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 intensities define the relative tax rate between different fuels. See table 9.

Fuel	Average Price (Rupees/mmBtu)	Emission Factor (Kg CO_{2e}/mmBtu)
Coal	262	100
Oil	665	82
Electricity	1,681	65
Natural Gas	1,307	60

Table 9: Example of Average Fuel Prices With and Without Carbon Tax

Notes: These prices are averaged across all sample years and all plants. Natural gas is slightly less polluting than electricity. This is because the vast majority of Indian electricity is generated with either coal or renewables such as hydro-electricity.

Since the carbon tax is a per-unit tax on fossil fuels $p_{fit} + \tau_f$, I separate the evolution of fuel prices/productivity of Equation 18 into two separate processes: one for prices and one for productivity separately. In practice, from the discrete grid that form the Markov chain describing the evolution of fuel price/productivity, I construct two separate grids, one for prices and another productivity. Grid points are found by matching moments of the discrete Markov chain for fuel prices and productivity with moments from observed fuel prices in the data and estimated fuel productivity. Details can be found in Appendix D.1. Intuitively, this is akin to decomposing the shocks to fuel prices/productivity into a shock to prices and a shock to productivity.

The Equimarginal Principle with Multidimensional Heterogeneity

While both multidimensional heterogeneity in fuel productivity and dynamic fuel switching complicates the graphical analysis of the equilibrium under a carbon tax, the equimarginal principle still holds. A carbon tax levied on fossil fuels not only equalizes marginal abatement costs (expressed in forgone profit) across plants, it also equalizes marginal abatement costs across fuels within a plant. This can be seen by rearranging first-order conditions of the plant's static profit maximization, where the carbon tax τ is levied on firms' total emissions $CO_{2e,it} = \sum_{f \in \mathcal{F}_{it}} \gamma_f e_{fit}$:

$$\left[\frac{\partial Y_{it}}{\partial e_{fit}}\Big(P(Y_{it}) + Y_{it}P'(Y_{it})\Big) - p_{fit}\right]\Big/\gamma_f = \tau \quad \forall f \in \mathcal{F}_{it} \tag{21}$$

The left-hand side is plant i's marginal cost of abating one ton of CO_{2e} using fuel f (in terms of forgone profits), whereas the right-hand side is the carbon tax rate. This equation is key because it provides a micro foundation for marginal abatement costs. In this highly heterogeneous context, firms have multiple ways to equalize their marginal abatement costs. In the short run, they can vary the quantity of each fuel in their set or their output. In the long run, they can pay fixed switching costs to adopt new fuels or drop existing ones. Critically, because fuel marginal products $\frac{\partial Y_{it}}{\partial e_{fit}}$ vary across fuels and across plants, plants will make different decisions to equalize their marginal abatement costs. This is the benefit of a market-based policy. Despite facing the same tax schedule, plants and may be more or less targeted by the tax based on input prices they face, their fuel sets, how productive their are overall, and how productive they are at using different fuels.

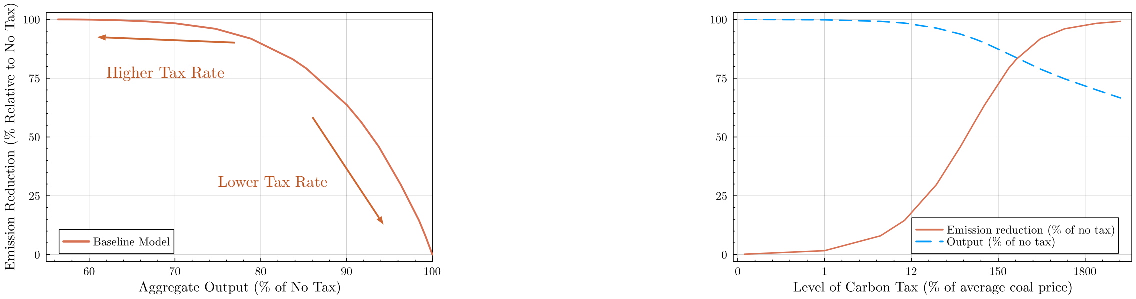
7.1 Carbon Tax and the Trade-off between Output and Emission Reduction

In Figure 7, 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. I simulate the economy with and without the carbon tax for 40 years starting from 2016, and look at the net present value (NPV) of outcomes along the entire path. For $X = \{\text{aggregate output, aggregate emissions}\}$ and a given carbon tax rate τ: $\mathbb{E}(X(\tau)) \approx \frac{1}{S}\sum_{s=1}^{S}\sum_{t=0}^{40}\beta^t X_{ts}(\tau)$

As the level of the tax approaches zero, the economy converges to the no-tax economy with 100% of output and 0% of emission reduction. As the tax level increases, emissions decrease but so does aggregate output. The production frontier is concave because of the increasing marginal cost of reducing aggregate emissions, consistent with previous findings by Fowlie et al. (2016). Fuel substitution (and more generally input substitution) is more effective initially, where much emission reduction can be achieved by substituting coal with 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, decreasing aggregate output because marginal plants have already switched towards cleaner fuels.

7.1.1 The Role of Heterogeneity in Fuel Productivity

I compare in Figure 8 what happens when removing heterogeneity in fuel productivity from the economy to highlight its role in the trade-off. 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. Details on the estimation of this production function are in Appendix D.2. Importantly,



(a) Trade-off between Output and Emission Reduction

(b) Emission Reduction and Output Across Tax Levels

Figure 7: 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 carbon tax levels, ranging from 0 (no tax) to approximately infinity. Linear interpolation is assumed for the trade-off between each tax level. As the level of the tax approaches infinity, the 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, and plants switch toward non-energy inputs.

this production function matches average fuel quantities and aggregate emissions levels but misses the heterogeneity in fuel shares across plants: $E_{it} = \psi_{Eit}\left(\sum_{f\in\mathcal{F}_{it}} \beta_f e_{fit}^{\frac{\lambda-1}{\lambda}}\right)^{\frac{\lambda}{\lambda-1}}$

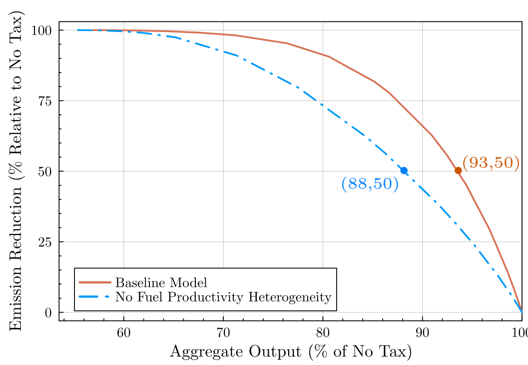


Figure 8: Comparison of Trade-off Across Model Specification

Notes: This production frontier compares the baseline economy (orange) with an economy that assumes all fuels are equally productive (blue).

The blue production frontier in Figure 8 corresponds to this restricted economy and shrinks inwards compared to the baseline economy. The restricted economy operates at 88% of no-tax output when reducing emissions by 50%, with an implied elasticity between emission reduction and output of 4.17. This represents an economically significant aggregate output loss of 50% when compared to the 93.5% of no-tax output in the baseline economy, with an implied elasticity of 7.7. Heterogeneity in fuel productivity diminishes how much output must decrease to achieve any reduction in emissions. To understand this result better, I decompose the total changes in emissions from the policy into three channels, following Levinson (2015). First, the scale channel captures variation in aggregate output. Second, the composition channel captures output reallocation across plants. Third, the technique channel captures plant-level improvements in emissions intensity, here within plant fuel substitution.

(a) Baseline

(b) Equally Productive Fuels

Figure 9: Decomposition of Emissions Reduction Under Different Levels of Carbon Tax

Notes: this figure decomposes emissions reduction into three channels following the decomposition method of Levinson (2015) for different levels of the carbon tax. While the level of the carbon tax is expressed in U.S. dollars, it does not account for purchase power parity (PPP) differences between India and the United States. For example, the 40 USD carbon tax is closer to 150 when accounting for PPP differences.

As shown in Figure 9, introducing heterogeneity in fuel productivity creates two countervailing effects. On the one hand, as long as fuels are gross substitutes ($\lambda > 1$), within-plant variation in fuel productivity induces plants to specialize in different fuels, increasing the opportunity cost of fuel substitution. This leads to technology lock-in, which is more pronounced for plants relatively more productive at dirty fuels like coal. The higher fuel productivity varies *within* plants, the higher this technology lock-in, and less emissions reduction comes from the technique channel. Substantive reasons underlying technology lock-in were discussed in Section 6.1. On the other hand, variation in fuel productivity *across* plants creates heterogeneity in exposure to the carbon tax. Plants specializing in dirty fuels become more exposed, and output reallocates towards plants specializing in cleaner fuels.

To better understand this last point, note that the carbon tax increases the relative price of dirtier fuels. Plants that are more productive at using dirty fuels face a higher marginal cost increase from the tax. As a result, these plants pass on a larger portion of the tax to consumers, making them less competitive. This variation in pass-through leads to a reallocation of output from dirtier plants to cleaner ones. For derivations, see Appendix D.2.1. In the Indian Steel context, the emissions reduction gains from output reallocation outweigh the losses from the technique channel. Therefore, introducing heterogeneity in fuel productivity mitigates the tradeoff between aggregate output and aggregate emissions reduction.[17]

[17]Importantly, this is not imposed by the model. The opposite outcome is also possible. Consider an economy in which all firms use the same fuels and have the same fuel productivity. All firms are equally exposed to the tax, shutting down the output reallocation channel. However, there is still technology lock-in because firms specialize in their most productive fuels. In such an economy, fuel productivity worsens the trade-off between aggregate output

To benchmark this result with the literature, 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 establishment entry/exit and dynamic investments in output capacity. However, they do not allow for input substitution. I show in Table 11 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).[18] Comparing this to the more flexible economy in which plants can substitute at both margins in Figure 10 sheds light on the critical role that input substitution plays in mitigating the loss of output for any emission reduction target.

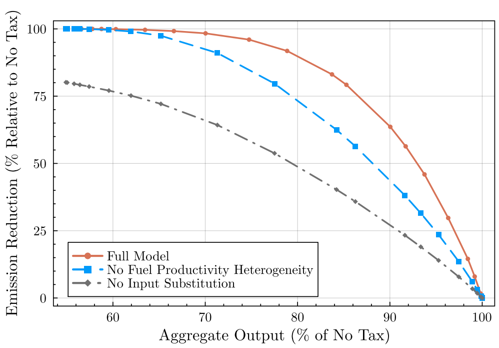


Table 10: Comparison of Trade-off Including No Input Substitution

	Average Elasticity $\frac{\%\Delta CO_{2e}}{\%\Delta Y}$
Baseline Economy	5.46
No Fuel Productivity Heterogeneity	3.89
No Input Substitution	2.48
Fowlie et al. (2016) – U.S. Cement	1.04

Table 11: Comparison of Average Elasticity

Notes: The average elasticity of 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 prices. I specifically approximate their *Auctioning* policy, which is isomorphic to a carbon tax in my context.

In the comparison with Fowlie et al. (2016), one potential concern is that I overestimate the extent of output reallocation because I do not capture capacity constraints in output, a feature that is excluded from the ASI dataset. I show that the results presented here hold even when imposing strict capacity constraints. In particular, I impose that plant-level output must be no greater under a carbon tax than under a no-tax economy. Since the majority of plants that become more competitive in the aftermath of the tax still produce below their no-tax output, they do not face their capacity constraints, and results remain quantitatively similar. That is, gains from output reallocation largely dominate losses from the technique effect. See Figure 10. This is because the tax increases marginal costs for all plants, pushing down output.

and aggregate emissions reduction. The same would be true if, instead of shutting down across firms heterogeneity, we assumed that output varieties were perfect complements.

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

(a) Baseline

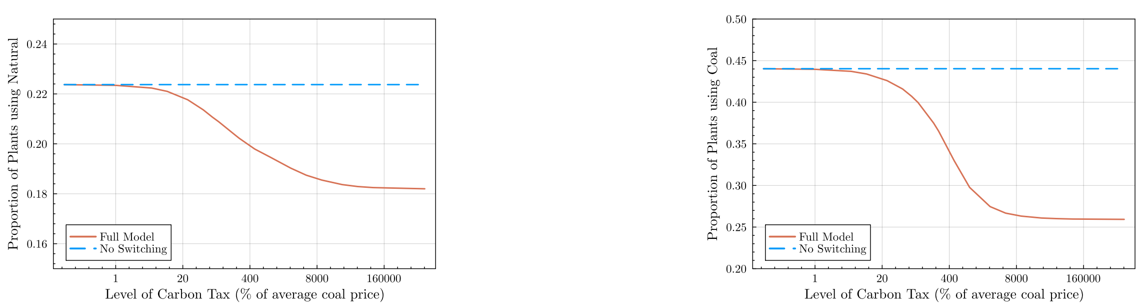
(b) Equally Productive Fuels

Figure 10: Decomposition of Emissions Reduction — Strict Capacity Constraint

Notes: this figure decomposes emissions reduction into three channels following the decomposition method of Levinson (2015) for different levels of the carbon tax. The only difference with Figure 9 is that I impose that plants cannot increase their level of output following the imposition of a carbon tax.

7.1.2 Ineffectiveness of Carbon Tax at the Extensive Margin

While a carbon tax can cost-effectively reduce emissions through its effect on fuel substitution at the intensive margin and reallocating output from high-emission to low-emission plants, its effect on the transition from coal to natural gas at the extensive margin is more limited. Indeed, any level of the carbon tax leads to a net decrease in the fraction of plants using coal or natural gas. The reduction in coal is relatively small, as it would take a carbon tax that raises the price of coal by 400% to incentivize a 10% decrease in the fraction of plants that use coal. Aside from coal being initially cheap relative to other fuels, this effect is primarily due to the option value that coal provides. Plants would rather reduce their coal consumption at the intensive margin but keep the option of using coal for the additional substitution margin it provides.



(a) Aggregate Share of Gas Users

(b) Aggregate Share of Coal Users

Figure 11: Natural Gas and Coal Take-up Across Levels of Carbon Tax

The lack of natural gas adoption can be partially explained by a combination of facts: the carbon tax also raises the price of natural gas, the fixed costs of adoption are economically high—especially for plants away from the pipeline network—and plants that are not currently using natural gas would be, on average, 30% less productive at it compared to those who already use natural gas. In the next section, I explore the effectiveness of a subsidy to incentivize natural gas take-up.

7.2 Alleviating Technology Lock-in — Subsidizing Natural Gas Adoption

To complement the carbon tax, I investigate how proceeds from the tax can be used to finance a subsidy to the fixed cost of natural gas to alleviate technology lock-in and increase natural gas take-up. I do these experiments jointly with a carbon tax, and I narrow down on a subsidy covering 10% of the fixed cost of natural gas, which can be fully financed by a carbon tax.[19]

To choose a representative social cost of carbon (SCC), I first set the social discount rate to 3% ($\beta = 0.97$) to match India's average real interest rate during the sample 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. 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 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 the 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 suffers externality damages from pollution in dollars from the social cost of carbon. To include a subsidy towards natural gas adoption, I assume that the subsidy is financed by government revenue from the carbon tax and that every plant faces the same permanent subsidy amount of s. In this context, producer surplus is the sum of total profits net of subsidized fixed costs, and net government revenue is total tax revenues minus subsidy paid out.

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

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

Note that externality damages cancel out with tax revenue, and the subsidy cancels out because

[19] I also investigate what happens under subsidies ranging from 0% to 100% in the **online appendix X**.

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:

$$w_t(\tau, s) = \underbrace{\nu_t(\tau, s)}_{\text{consumer surplus}} + \underbrace{\sum_{i=1}^{N} \pi_{it}(\tau, s)}_{\text{variable profits}} - \underbrace{\sum_{i}^{N} \Big(\sum_{\mathcal{F}' \subseteq \mathbb{F}} \mathcal{K}(\mathcal{F}' \mid \mathcal{F}_{it}) \mathbf{I}(\mathcal{F}_{it+1} = \mathcal{F}' \mid \tau, s) \Big)}_{\text{total fixed costs}} \tag{22}$$

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

	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 Net Fixed Costs (*Paid by plants + subsidy*)	-18.633	-18.601	31.61
Externality Damages/Tax Revenue	2.64	2.65	9.78

Table 12: Decomposition of Welfare Effects – Carbon Tax with and without Subsidy

Notes: All welfare components are reported by their net present value (NPV) over a 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 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 $Welfare = ConsumerSurplus + VariableProfit - TotalFixedCost$. Net total fixed costs can be negative because of salvage values firms can get from dropping fuels.

	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 13: Total Subsidy paid

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

Welfare results are shown in Table 12 and 13. In net, there is a positive but small welfare effect from the subsidy relative to a regime with only a carbon tax. These small welfare effects reflect important variations within each welfare component. First, the fraction of plants adopting natural gas goes up by five percentage points, but plants do not drop more polluting fuels. The option value of natural gas reduces marginal costs for newly adopting plants, pushing down output prices. Lower marginal costs increase output, variable profits, and consumer surplus by 19 and 13 million dollars, respectively. However, emissions also increase by an equivalent of 9.8 million dollars in externality damages, as the income effect dominates the substitution effect. Details underlying are the income and substitution effects are discussed at length in **online appendix X**.

Second, while the subsidy's cost in the government budget gets canceled out with the subsidy's revenue in plants' profit, 99% of these transfers go to inframarginal plants that would've adopted gas in the absence of a subsidy. This can be seen from the increase in total fixed costs paid in the economy after the subsidy is introduced (31.61 million dollars), which is considerably lower than the total subsidy dispensed by the government (2.7 billion dollars).

This prevalence of inframarginal plants coupled with a small welfare effect raises the question of whether the government could find more profitable avenues to alleviate technology lock-in. For example, it could invest in energy efficiency training programs to increase energy and fuel productivity or carbon capture technologies that reduce ex-post emissions. Alternatively, following the footsteps of Belgium, Austria, and others, the government could outright phase out/ban coal. This paper is well-equipped to handle such a policy. While outside the scope of this paper, this is an interesting avenue for future research.

8 Conclusion

In conclusion, I develop a rich dynamic production model to study fuel substitution from manufacturing establishments. It includes switching between fuel sets at a cost 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 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. I then perform various counterfactual policy experiments to reduce emissions at the lowest cost possible, including a carbon tax and a carbon tax with a subsidy towards adopting cleaner fuels.

Moreover, 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 multiple layers of heterogeneity. As a result, high-emission plants become relatively less competitive, reallocating output towards low-emission plants. This considerably reduces the overall economic cost of reducing emissions. However, more than a carbon tax is needed to increase adoption of cleaner fuels such as natural gas. For this reason, I show how proceeds from the carbon tax can be used to subsidize the fixed cost of natural gas adoption. There is a small but positive welfare effect, unexpectedly through a larger private surplus (producer and consumer) at the expense of higher emissions. This is due to the option value that an additional fuel provides, which lowers production costs. However, the welfare effects of the subsidy are minor compared to its cost. Overall, these results highlight the importance of producer heterogeneity and inter-temporal decisions when quantifying the impact of carbon policy.

References

Acemoglu, Daron, 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.

Ahmadi, Younes and Akio Yamazaki, "The Effectiveness of Revenue-Neutral Carbon Taxes," 2020.

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

Alpino, Matteo, Luca Citino, and Annalisa Frigo, "The effects of the 2021 energy crisis on medium-sized and large industrial firms: evidence from Italy," Questioni di Economia e Finanza (Occasional Papers) 776, Bank of Italy, Economic Research and International Relations Area June 2023.

Andersson, Julius J., "Carbon Taxes and CO2 Emissions: Sweden as a Case Study," *American Economic Journal: Economic Policy*, 2019, *11* (4), 1–30.

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.

Atkinson, Scott E. and Rong Luo, "Estimation of Production Technologies with Output and Environmental Constraints," *International Economic Review*, 2023, *0* (0).

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.

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

Bonhomme, Stéphane and Elena Manresa, "Grouped Patterns of Heterogeneity in Panel Data," *Econometrica*, 2015, *83* (3), 1147–1184.

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

Carlson, Curtis, Dallas Burtraw, Maureen Cropper, and Karen L. Palmer, "Sulfur Dioxide Control by Electric Utilities: What Are the Gains from Trade?," *Journal of Policial Economy*, 2000, *108* (6).

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.

Collard-Wexler, Allan and Jan De Loecker, "Reallocation and Technology: Evidence from the US Steel Industry," *American Economic Review*, 2015, *105.*

Culler, Joseph A. and Eric T. Mansur, "Inferring Carbon Abatement Costs in Electricity Markets: A Revealed Preference Approach Using the Shale Revolution," *American Economic Journal: Economic Policy*, 2017, *9* (3).

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

Dhar, Subash, Minal Pathak, and Priyadarshi R. Shukla, "Transformation of India's steel and cement industry in a sustainable 1.5°C world," *Energy Policy*, 2020, *137.*

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.

Farmer, Leland E. and Alexis Akira Toda, "Discretizing nonlinear, non-Gaussian Markov processes with exact conditional moments," *Quantitative Economics*, 2017, *8.*

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.

United Nations — Climate Action

United Nations — Climate Action, "Causes and Effects of Climate Change," 2023. Retrieved from https://www.un.org/en/climatechange/science/causes-effects-climate-change.

Fontagné, Lionel, Philippe Martin, and Gianluca Orefice, "The Many Channels of Firm's Adjustment to Energy Shocks: Evidence from France," Working Paper 929, Banque de France Eurosystème 2023.

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

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.

Gittens, Jarone, "The Impact of Carbon Taxes on Market Structure: Evidence from British Columbia's Retail Gasoline Industry," Technical Report, Department of Economics, Queen's University 2020.

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.

Gosnell, Greer K., John List, and Robert D. Metcalfe, "The Impact of Management Practices on Employee Productivity: A Field Experiment with Airline Captains," *Journal of Political Economy*, 2020, *128* (4).

Goulder, Lawrence H. and Ian W. H. Parry, "Instrument Choice in Environmental Policy," *Review of Environmental Economics and Policy*, 2008, *2* (2).

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.

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.

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.

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.

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.

Levinson, Arik, "A Direct Estimate of the Technique Effect: Changes in the Pollution Intensity of US Manufacturing, 1990–2008," *Journal of the Association of Environmental and Resource Economists*, 2015, *2* (1).

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.

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.

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

Pai, Sandeep and Hisham Zerriffi, "A novel dataset for analysing sub-national socioeconomic developments in the Indian coal industry," *IOPSciNotes*, 2021, *2.*

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

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).

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.

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.

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

Appendices

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 out 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. Strata were constructed by state/industry pairs for the remaining plants, 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 were part of the census: Arunachal Pradesh, Manipur, Meghalaya, Nagaland, Sikkim, Tripura, and Andaman Nicobar Islands. Lastly, the census included plants with more than 100 workers in all other states.

A.2 Calculating 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}) 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:

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

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

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 are converted into carbon dioxide equivalent CO_2e using the Global Warming Potential (GWP) method as follows:

$$CO_{2e} = \underbrace{GWP_{CO_2}}_{=1} * CO_2 + GWP_{ch4} * CH_4 + GWP_{n2o} * N_2O$$

From the calculations above, I can define fuel-specific emission factors that will be used to directly convert fuels to CO_{2e} (or GHG). For fuel f in industry j (excluding electricity),

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

Calculations of emissions from electricity are done slightly differently than from fossil fuels because emissions come from production rather than end usage of electricity. Figure 12 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.

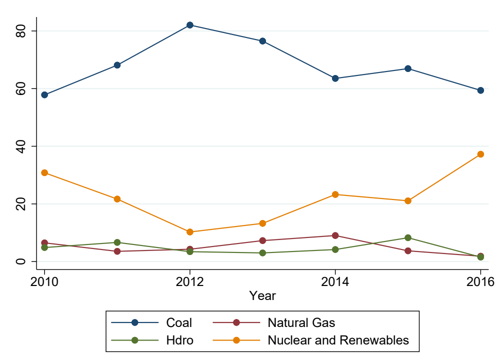


Figure 12: Annual Indian Electricity Generation by Source (% of Total)
Source: International Energy Agency (IEA)

To construct measures of emissions from electricity, I 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 across the country, then

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

Where γ_{fj} is the emission intensity of fuel f and was defined above.Total GHG emissions for plant i in industry j and year t is then 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 the average over all petroleum 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 14: 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 15: Emission factors from Electricity

A.3 Additional Evidence

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:

Steel Product \ Fuel Set	oil, elec	oil, elec, coal	oil, elec, gas	oil, elec, gas, coal	other	*Total*
Pig iron	50.46	23.39	8.29	5.71	12.15	100
Direct reduced iron	62.42	22.53	4.89	2.31	7.85	100
Ingots	54.12	17.51	9.23	6.76	12.38	100
Ferro-alloy	51.42	24.05	4.91	4.11	15.51	100
Hot and cold-rolled steel	38.32	21.14	15.00	14.80	10.75	100
Tubes	67.97	4.74	15.32	4.46	7.52	100
Wires	70.04	4.61	9.93	1.42	14.01	100
Other	47.59	19.79	10.35	9.96	12.30	100

Table 16: Distribution of Fuels Sets by Steel Variety

$$\begin{aligned} \max_{\{Y_{it}\}_{i=1}^{N_t}, Y_{0t}} \quad & \mathbb{U} = Y_{0t} + \frac{e^{\Gamma_t}}{\theta}\left(\frac{1}{N_t}\int_{\Omega_i} (N_t Y_{it})^{\frac{\rho-1}{\rho}} di\right)^{\frac{\theta\rho}{\rho-1}} \\ s.t. \quad & Y_{0t} + \int_{\Omega_i} P_{it} Y_{it} di \leq I_t \end{aligned} \tag{23}$$

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

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

The 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 $Y_{0t}(P_t) = I_t = P_t Y_t(P_t) = I_t - e^{\Gamma_t \frac{1}{1-\theta}} P_t^{\frac{-\theta}{1-\theta}}$. 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. Consumer surplus is decreasing in the aggregate price index, keeping income constant. Then, the representative consumer chooses which varieties to allocate for a given quantitiy 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:

$$Y_{it}(Y_t) = \frac{Y_t}{N_t}\left(\frac{P_{it}}{P_t}\right)^{-\rho} \tag{24}$$

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 and Estimation

C.1 Alternative Demand Specification

As a more flexible alternative to a simple CES demand function, I consider a nested CES demand specification with three nests and an outside option: (1) ρ; across firms within products, (2) γ; across products within product category, and (3) ϵ; across product category. Steel products and product categories are classified as follows, directly mapping into the Indian National Industry Classification.

Product	**Product Category**	**Share of output**
Pig Iron	Iron	0.1
Direct reduced iron	Iron	0.07
Ingots	Semi-finished	0.09
Ferro-alloy	Semi-finished	0.05
Hot and cold-rolled steel	Semi-finished	0.45
Tubes	Finished	0.03
Railway tracks	Finished	0.003
Wires	Finished	0.2

Table 17: Categorization of Steel Products

The estimating equation then includes the own prices, the aggregate prices of firms producing the same products, the same product category, and any steel product:

$$\ln Y_{it} = \beta_0 + \frac{\epsilon(\theta-1)+1}{\theta-1}\underbrace{\ln P_t}_{\text{Aggregate price}} + (\gamma-\epsilon)\underbrace{\ln P_{jt}}_{\text{Same Product Cateogry}} + (\rho-\gamma)\underbrace{\ln P_{sjt}}_{\text{Same product}} - \rho \underbrace{\ln P_{it}}_{\text{Own price}} + u_{it}$$

Elasticity of Substitution	Estimate (s.e.)	
Across product category ϵ	3.3*	(1.4)
Across products within category γ	3.8***	(1.1)
Across plants within product ρ	4.0***	(1.2)
Observations	8,517	

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

Table 18: Nested CES Demand Estimates

Notes: Aggregate price indices P_t, P_{jt}, P_{sjt} are constructed for each guess of parameter estimates $\{\rho^o, \gamma^o, \epsilon^o\}$. Under monopolistic competition, individual firms' demand shocks u_{it} do not affect these aggregate price indices. Own prices P_{it}, however, are instrumented with the shift-share cost-shifter of Ganapati et al. (2020).

C.2 Derivation of Estimating Equation for Outer Production Function

Production function:

$$\frac{Y_{it}}{\overline{Y}} = e^{\omega_{it}}\left(\alpha_k\left(\frac{K_{it}}{\overline{K}}\right)^{\frac{\sigma-1}{\sigma}} + \alpha_l\left(\frac{L_{it}}{\overline{L}}\right)^{\frac{\sigma-1}{\sigma}} + \alpha_m\left(\frac{M_{it}}{\overline{M}}\right)^{\frac{\sigma-1}{\sigma}} + \alpha_e\left(\frac{E_{it}}{\overline{E}}\right)^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\eta\sigma}{\sigma-1}} \tag{25}$$

$$= e^{\omega_{it}}\left(\alpha_k\tilde{K}_{it}^{\frac{\sigma-1}{\sigma}} + \alpha_l\tilde{L}_{it}^{\frac{\sigma-1}{\sigma}} + \alpha_m\tilde{M}_{it}^{\frac{\sigma-1}{\sigma}} + \alpha_e\tilde{E}_{it}^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\eta\sigma}{\sigma-1}} \tag{26}$$

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

Assumption 2. *L_{it}, M_{it}, E_{it} are flexible inputs*

Assumption 3. *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:

$$\max_{L_{it}, M_{it}, E_{it}} \left\{P_{it}(Y_{it})Y_{it} - p_{Mit}M_{it} - p_{Eit}E_{it} - w_t L_{it}\right\}$$

$$s.t. \quad Y_{it} = \overline{Y}e^{\omega_{it}}\left(\alpha_K\tilde{K}_{it}^{\frac{\sigma-1}{\sigma}} + \alpha_L\tilde{L}_{it}^{\frac{\sigma-1}{\sigma}} + \alpha_M\tilde{M}_{it}^{\frac{\sigma-1}{\sigma}} + \alpha_E\tilde{E}_{it}^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\eta\sigma}{\sigma-1}}$$

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

First-order conditions:

M_{it}/L_{it}:

$$\frac{M_{it}}{\overline{M}} = \Big(\frac{\alpha_L}{\alpha_M}\frac{S_{Mit}}{S_{Lit}}\Big)^{\frac{\sigma}{\sigma-1}}\frac{L_{it}}{\overline{L}} \tag{27}$$

E_{it}/L_{it}:

$$\frac{E_{it}}{\overline{E}} = \Big(\frac{\alpha_L}{\alpha_E}\frac{S_{Eit}}{S_{Lit}}\Big)^{\frac{\sigma}{\sigma-1}}\frac{L_{it}}{\overline{L}} \tag{28}$$

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 $e^{\omega_{it}}$:

$$e^{\omega_{it}\frac{\rho-1}{\rho}} = \overline{Y}^{\frac{\rho-1}{\rho}} \frac{\rho}{\rho-1}\frac{1}{\eta}\Big(\frac{N_t}{e^{\Gamma_t}}\Big)^{\frac{1}{\rho}} P_t^{\frac{1-\rho(1-\theta)}{(1-\theta)\rho}} \frac{S_{L_{it}}}{\alpha_L L_{it}^{\frac{\sigma-1}{\sigma}}} ces_{it}^{\frac{\eta\sigma-\rho[\sigma(\eta-1)+1]}{(\sigma-1)\rho}} \tag{29}$$

Plug (29) 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 (27) and (28) into the previous equation:

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

Estimating Equation:

$$
\ln R_{it} = \ln\frac{\rho}{\rho-1}+\ln\frac{1}{\eta}+\ln\left(S_{Lit}\left(1+\frac{\alpha_k}{\alpha_L}\left(\frac{\tilde{K}_{it}}{\tilde{L}_{it}}\right)^{\frac{\sigma-1}{\sigma}}\right)+S_{Mit}+S_{Eit}\right)+u_{it} \tag{30}
$$

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 $s, \mathcal{F}$.

$$
W^n(s,\mathcal{F}) = \gamma+\sigma_\epsilon\log\left(\sum_{\mathcal{F}'\in\mathbb{F}}\exp\left(\pi(s,\mathcal{F})+\Phi(\mathcal{F}'\mid\mathcal{F})+\beta\int W^n(s',\mathcal{F}')dF(s'\mid s)\right)^{1/\sigma_\epsilon}\right)
$$

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 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,

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

Where t corresponds to the year of observation and d is an indicator for access to a 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 with correlated errors, I use the approach of Farmer and Akira Toda (2017) to discretize all the state variables into a single grid. Let M be the number of points on each grid. Using this discretization process, I can then represent the value function as a block matrix $\vec{W}^n$ containing all state combinations. 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 the computational burden, I iterate over equation (31) by parallelizing across all possible combinations of starting states using graphics processing units (GPU) Arrays with CUDA. Computational gains using GPU Arrays are significant over standard CPU parallelization. Detailed Julia code is available on my GitHub.

C.4 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.

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

$$\equiv \arg\max_{\theta_1} \sum_{i=1}^{N}\sum_{t=1}^{T}\sum_k \rho(\mu_k \mid \mathcal{F}_i, s_i; \hat{\theta}_1, \hat{\theta}_2, \hat{\pi}) \ln Pr(\mathcal{F}_{it+1} \mid \mathcal{F}_{it}, s_{it}, \mu_i = \mu_k; \theta_1, \hat{\theta}_2)$$

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 combinations 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.

$$\mathrm{W}(\mathrm{s},\mathrm{F},\mu_k; \theta_1^0, \hat{\theta}_2) = \gamma + \sigma_\epsilon \ln\left(\sum_{\mathcal{F}' \in \mathbb{F}} \exp\left(\pi(s,\mathcal{F}) + \mathcal{K}(\mathcal{F}' \mid \mathcal{F}, s; \theta_1^0) + \beta \int W(s', \mathcal{F}', \mu_k; \theta_1^0, \hat{\theta}_2) dF(s' \mid s; \hat{\theta}_2)\right)^{1/\sigma_\epsilon}\right)$$

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 Discretizing the Process for Fuel Prices and Productivity Separately

The problem to solve is that I need to separate fuel prices from fuel productivity when studying the impact of a per-unit carbon tax levied on fossil fuels $(p_{fit} + \tau_f)$ because $\frac{p_{fit}+\tau_f}{\psi_{fit}} = \frac{p_{fit}}{\psi_{fit}} + \frac{\tau_f}{\psi_{fit}}$

Initially, the model is estimated with a process for the log of fuel prices/productivity from Equation 18, which I discretize into a Markov Chain. The Markov chain is a sequence of fuel prices/productivity realizations $\ln \frac{p_{f1}}{\psi_{f1}}, \ln \frac{p_{f2}}{\psi_{f2}}, \ln \frac{p_{f3}}{\psi_{f3}}, ...$ such that

$$
\begin{gathered}
Pr\Big(\ln \frac{p_{ft+1}}{\psi_{ft+1}} \Big| \ln \frac{p_{ft}}{\psi_{ft}}, \ln \frac{p_{ft-1}}{\psi_{ft-1}}, \ln \frac{p_{ft-2}}{\psi_{ft-2}}, ... \Big) = Pr\Big(\ln \frac{p_{ft+1}}{\psi_{ft+1}} \Big| \ln \frac{p_{ft}}{\psi_{ft}} \Big) \\
Pr\Big(\ln \frac{p_{ft+1}}{\psi_{ft+1}} \Big| \ln \frac{p_{ft}}{\psi_{ft}} \Big) > 0 \quad \forall \quad k
\end{gathered}
$$

From this Markov chain for fuel price/productivity, I create two Markov Chains: one for prices $\ln p_{f1}, \ln p_{f2}, ...$ and one for productivity $\ln \psi_{f1}, \ln \psi_{f2}, ...$ such that

$$
\begin{gathered}
\ln p_{ft} := \ln \frac{p_{ft}}{\psi_{ft}} + \ln \psi_{ft} \\
Pr\Big(\ln p_{ft+1} \mid \ln p_{ft} \Big) = Pr\Big(\ln \psi_{ft+1} \mid \ln \psi_{ft} \Big) = Pr\Big(\ln \frac{p_{ft+1}}{\psi_{ft+1}} \Big| \ln \frac{p_{ft}}{\psi_{ft}} \Big)
\end{gathered}
$$

To find the grid points that form the grid for fuel prices and fuel productivity, I match the moments of the newly constructed Markov chains with moments from the distribution of fuel prices, fuel productivity and other state variables in the data. All the moments I use include the variance of fuel prices, the variance of fuel productivity, the covariance between fuel prices and fuel productivity, the covariance between fuel prices and all other states, and the covariance between fuel productivity and all other states.

D.2 Energy Production Function with Energy Productivity – Identification and Results

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 19: 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	3,459	

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

D.2.1 Elasticity of the Price of Energy with Respect to Relative Fuel Prices

Here, I show that since the carbon tax effectively increases the relative price of dirtier fuels, plants that are more productive at using dirty fuels are more exposed to the carbon tax. To see this, let $\tilde{p}_{cit} = p_{cit}/p_{git}$ be the price of coal relative to gas and likewise for relative fuel productivity $\tilde{\psi}_{cit} = \psi_{cit}/\psi_{git}$. Then,

$$\frac{\partial \ln p_{E_{it}}}{\partial \ln \tilde{p}_{cit}} = \frac{\left(\tilde{p}_{cit}/\tilde{\psi}_{cit}\right)^{1-\lambda}}{\sum_{f\in\mathcal{F}_{it}}\left(\tilde{p}_{fit}/\tilde{\psi}_{fit}\right)^{1-\lambda}} = \frac{p_{cit}e_{cit}}{\sum_{f\in\mathcal{F}_{it}} p_{fit}e_{fit}} \tag{32}$$

Under cost-minimization, the elasticity of the marginal cost of energy with respect to the relative price of any fuel (e.g. coal relative to gas) is just the plant-specific spending share of that fuel relative to all fuels. This is an application of the Envelope Theorem. Indeed, The price of energy in the fully flexible model is as follows:

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

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 $\tilde{\psi}_{fit}$

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

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

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

Moreover, this elasticity is equal to the spending share of coal relative to all other fuels. To see this, relative first-order conditions of the cost-minimization problem in (5) for two fuels (c, g) are:

$$
\begin{aligned}
\frac{p_{cit}}{p_{git}} &= \left(\frac{\psi_{cit} e_{cit}}{\psi_{git} e_{git}}\right)^{-\frac{1}{\lambda}} \frac{\psi_{cit}}{\psi_{git}} \\
\frac{e_{cit}}{e_{git}} &= \left(\frac{p_{git}}{p_{cit}}\right)^{\lambda} \left(\frac{\psi_{cit}}{\psi_{git}}\right)^{\lambda-1}
\end{aligned}
$$

Multiplying both sides by relative prices, this yields:

$$
\frac{p_{cit} e_{cit}}{p_{git} e_{git}} = \frac{(p_{cit}/\psi_{cit})^{1-\lambda}}{(p_{git}/\psi_{git})^{1-\lambda}}
$$

Summing across all relative fuel spending shares yields the elasticity:

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

Most importantly, this elasticity is increasing in relative fuel productivity. This means that conditional on fuel prices and fuel set, plants that are more productive at using coal spend more on coal and are more sensitive to relative changes in the price of coal:

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

In contrast, in the economy 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 sets.