Balancing Production and Carbon Emissions with Fuel Substitution*

Emmanuel Murray Leclair†

Department of Economics, Western University

This Version: *September 27, 2023*

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

Abstract

I evaluate the role of fuel substitution in shaping the trade-off between output and emission reduction, central to debates on carbon policy. To do so, I develop a dynamic production model of firms with fuel choices that reflect inter-temporal substitution between fuels and fuel productivity heterogeneity. I estimate the model using a panel of steel manufacturing plants from the Indian Survey of Industries (2009-2016) and examine plants' responses to carbon taxation via fuel-specific tax rates, that affect input choices. Fuel substitution reduces the estimated output costs from carbon taxation. Heterogeneity in fuel productivity creates variation in exposure to the tax, with more polluting plants being more exposed. As these plants become less competitive, production reallocates towards cleaner plants, mitigating output losses from emission reduction. However, "technology lock-in" prevents the widespread adoption of "clean" fuels. Adoption entails high fixed costs and fuel productivity varies across plants. Many plants are too inefficient to pay these costs and thus do not transition away from "dirty" fuels. Using proceeds from carbon taxation to subsidize the adoption of cleaner fuels only minimally improves welfare due to this technology lock-in.

*I would like to thank my advisors David Rivers, Salvador Navarro and Sergio Ocampo Diaz for their incredible support throughout the elaboration of this project. I also want to thank Rory McGee, Michael Sullivan, Nouri Najjar, Nirav Mehta, Charles Saunders, Martin O'Connell, Matt Gordon, Yanyou Chen, Richard Sweeney, Pascual Restrepo, Elie Tamer, Yongseok Shin, and Andrew McCallum for their insightful comments. I also want to thank participants that gave me valuable feedback at 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@uwo.ca Website: https://sites.google.com/view/emmanuelmurrayleclair/home/

1 Introduction

The reliance on fossil fuels of 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 and 90% of carbon dioxide emissions (UN Climate, 2023). Manufacturing activities account for 37% of global greenhouse gas emissions (Worrell, Bernstein, Roy, Price and Harnisch, 2009). Notably, the degree of pollution varies among fossil fuels, with coal, for instance, being twice as polluting as natural gas per British thermal unit (Btu) (EPA, 2023). These disparities are particularly relevant in developing economies such as India, which currently accounts for 7% of global emissions and 12% of coal consumption (Ritchie, Roser and Rosado, 2022).

Reducing greenhouse gas emissions entails substituting dirty fuels for clean ones (Joskow and Mishkin, 1977; Atkinson and Halvorsen, 1976) or decreasing output (Ganapati, Shapiro and Walker, 2020). This trade-off between emissions reduction and output is particularly biting in developing economies, where economic growth concerns have arguably impeded the implementation of such policies (Jack, 2017). In this context, the substitutability of various fuels and the barriers preventing substitution become important determinants of the welfare trade-offs of emission-reduction policies.

To study the trade-off between output and emission reduction, I develop a dynamic production model with multidimensional fuel choices and heterogeneity in fuel productivity. I take into account heterogeneity in plants' ability to use different fuels combined with the fixed costs of adopting new fuels, adding to previous work on fuel substitution (Ganapati et al., 2020; Hyland and Haller, 2018; Wang and Lin, 2017; Ma et al., 2008; Cho et al., 2004; Pindyck, 1979). Heterogeneity in fuel productivity enter in the production function and creates variation in the desirability to substitute across fuels. Fixed costs enter in a dynamic discrete choice model of fuel set choices and create a trade-off between contemporaneous profits and future marginal costs due to the substitution margin that a new fuel provides (Broda and Weinstein, 2006). I then estimate the model on a panel of Indian steel manufacturing establishments.

I characterize the trade-off between emission reduction and output by imposing different levels of a carbon tax levied on fossil fuels. The relative tax rate on each fuel equals its marginal externality damage. Carbon taxes are more effective at reducing emissions when allowing for heterogeneity in fuel productivity. Plants specializing in highly polluting fuels due to productivity advantages are more exposed to the tax, facing larger increases in marginal costs. As these plants pass the increase in costs to consumers, they become less competitive, and output reallocates towards cleaner, less exposed, plants. While the trade-off is non-linear, reducing emissions by 50% entails a reduction of output by 6.5% (relative to a no-tax economy), compared to a 12% reduction without this composition effect. This composition effect hinges on the pass-through rate of carbon tax, which I estimate following Ganapati et al. (2020).

However, difficulties in transitioning out of old technologies, known as "technology lock-in", limit the adoption of cleaner fuels and increase the trade-off between emission reduction and output. Fixed adoption costs, together with heterogeneity in fuel productivity, generate this lock-in. The most polluting plants, currently using coal, would be too unproductive at using natural gas to justify the fixed cost of adoption. A carbon tax does not address this imbalance and therefore does not lead to more plants adopting natural gas. Subsidizing the adoption of natural gas increases output but comes at the cost of higher emissions due to the substitution margin that natural gas provides to adopting plants. If the subsidy was financed with proceeds from a carbon tax, the net effect would be a small welfare improvement of 0.003% relative to a regime where carbon tax revenues are transfered back to consumers.

Quantifying the role of fossil fuels in production requires overcoming two important measurement issues coming from heterogeneity in fuel productivity. First, the energy that 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 quantity of fuels (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 exclusively because it anticipates high coal productivity and low gas productivity. However, the plant's gas productivity remains unobserved to the researcher.

I address these issues in three steps. First, I identify the quantity and price of energy services by adapting the method of Grieco, Li and Zhang (2016) and Ganapati, Shapiro and Walker (2020). This method relies on optimality conditions from profit maximization to map observed relative input spending to unobserved relative input quantities. Second, I estimate the function that maps fuels to 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. These estimates allow me to conduct policy counterfactuals that affect plants' fuel choices.

I estimate this model for the Indian steel industry between 2009 and 2016 using a panel of establishments from the Indian Survey of Industries (ASI). Steel manufacturing is one of India's most polluting industries, with coal accounting for nearly 70% of its energy sources. Moreover, fuel substitution, through the retrofitting of existing plants, is the focal point of decarbonization efforts in the industry (Lei, Wang, Yu, Ma, Zhao, Cui, Meng, Tao and Guan, 2023). The panel features quantities and prices of disaggregated inputs that plants purchase and outputs that plants manufacture, as well as plants' locations into 775 districts, which I map to the Indian network of natural gas pipelines. This information allows fixed costs of natural gas adoption to vary with plants' proximity to pipelines.

Indian steel plants use different fuel sets. While 90% of plants always use oil and electricity, 28% of plants also use coal, 18% of plants also use gas, and 8% of plants use both coal and gas. Moreover, they often switch between fuel sets. On average, 15% of plants add a new fuel and 15% of plants drop an existing fuel every year. 40% of unique plants add a new fuel at least once, and likewise for plants dropping an existing fuel.

I find a higher elasticity of substitution among fuels than between fuels and non-energy inputs. This is important because the more substitutable fuels are, the more cost-effective a carbon tax is at reducing emissions. I also find a combination of large and persistent heterogeneity in fuel productivity with economically high fixed costs of natural gas adoption, averaging between 28 and 40 million U.S. dollars. Plants that do not utilize natural gas would be 30% less productive at using gas compared to plants who use natural gas. This disparity is accentuated by natural gas prices that are five times coal prices. These interrelated factors generate technological lock-in and help explaining the prevalent use of coal within the Indian steel industry.

Literature and Contribution

I estimate a production model 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 CO_2 productivity. These disparities in productivity are due to varying heat efficiency that different fuel-burning technologies provide (Allcott and Greenstone, 2012) and energy retrofit efforts to curb energy waste (Christensen et al., 2022). 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.

In estimating my model, I contribute to the literature on production function estimation (Olley and Pakes, 1996; Blundell and Bond, 2000; Levinsohn and Petrin, 2003; Ackerberg, Caves and Frazer, 2015; Grieco, Li and Zhang, 2016; Zhang, 2019; Gandhi, Navarro and Rivers, 2020; Demirer, 2020). I make a methodological contribution by showing how to identify and estimate a dynamic production function with input-augmenting productivity and dynamic selection on unobservables from plants' fuel choices. 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).

My model with multidimensional fuel choice relaxes the canonical asusmption of a pollution function that underlie a uni-dimensional choice of pollution abatement that has been a staple in this literature (Copeland and Taylor, 2004; Shapiro and Walker, 2018). This allows me to revisit the effects of environmental policies on plant-level pollution and their optimal design.

My results on the heterogeneity of energy productivity across fuels and manufacturing plants is related to Hawkins-Pierot and Wagner (2022) who estimate the energy productivity of manufacturing plants and its implication for technology lock-in. My paper decomposes 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. These results also contribute to the literature on energy efficiency in the residential sector (Fowlie and Meeks, 2021; Chan and Gillingham, 2015), power generation (Cicala, 2022; Davis and Wolfram, 2012; Fabrizio, Rose and Wolfram, 2007), and technology adoption in the industrial sector (Gerarden, Newell and Stavins, 2017; Allcott and Greenstone, 2012)[1].

2 Data

I use longitudinal data on prices and quantities of all inputs and outputs from Indian steel establishments, 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. More details on sampling rules, including changes over time, can be found in Appendix A.1. The ASI contains information on prices and quantities of Coal, Oil, Electricity, and Natural Gas, which I convert to million British thermal unit (mmBtu) using standard scien-

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

tific calculations from the U.S. Environmental Protection Agency (EPA, 2022). I follow standard practices by removing the 1% left and right tails of plant-level inputs and output by industry.[2]

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.

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 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).

Fact 2: Indian Steel Establishments Use Different Fuel Sets

Steel-producing plants use different fuel sets, and the vast majority of fuel sets include both oil and electricity. Most of the variation in fuel sets thus comes from whether plants use coal, natural

[2]Such outliers are typically due to reporting errors and are inconsistent with a wide range of official statistics (Bollard, Klenow and Sharma, 2013).

Industry	Annual Average Number of Plants	Average Annual Revenue by Plant (Million USD)	Annual Average Emissions (Thousand tons CO_{2e})	Aggregate Energy Input Share	Aggregate Coal Fuel Share
Steel	1,077	19.41	29.34	0.13	0.72
Other	33,726	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.

gas, neither, or both. see Table 3. There are multiple reasons for this heterogeneity. For example, plants can use different type of furnaces to turn iron ore into steel. Blast furnaces combined with basic oxygen furnaces rely on coke (coal) as a primary fuel; electric arc furnaces can use any fuel combinations to generate high-power electricity, which is discharged through an electric arc to melt either steel scrap or sponge iron. Sponge iron is created from iron ore in direct reduced iron furnaces powered bys natural gas.

	Percentage
Oil, Electricity	51.6
Oil, Electricity, **Coal**	18.8
Oil, Electricity, **Gas**	10.1
Oil, Electricity, **Coal**, **Gas**	7.9
Other	11.5

Table 2: Distribution of Fuel Sets Across Steel Plants

Notes: I define fuel sets as any combination of fuels that a plant uses in a given year. The *Other* category comprises of any other mix of these four fuels.

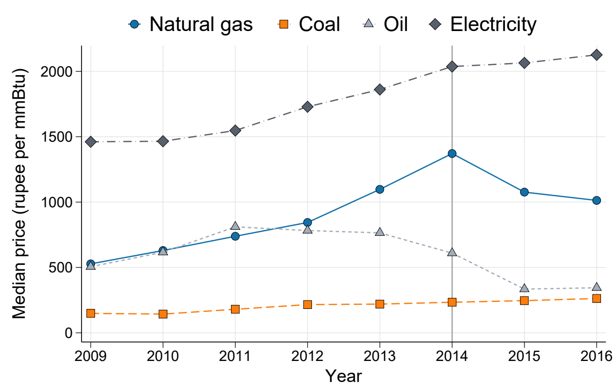


Figure 1: Median Fuel Prices (rupees/mmBtu)

This heterogeneity in fuel sets has two noteworthy implications. First, burning coal releases more CO_{2e} than burning natural gas. Second, plants with more fuels have access to additional margins of substitution, which is rellevant because many fossil fuels are susceptible to price volatility owing to global supply and demand fluctuations. In 2014, for instance, oil and natural gas prices fells due to booming U.S. shale oil production and excess supply from emerging market economies (EMEs) such as Saudi Arabia. See Figure 4.

Fuel substitution can serve as a means of adjustment for plants to insure themselves against fuel price variation. Both the quantity and fuel share of coal increased prior to 2014. However, coal usage started to decline after 2014 following the crash in the price of oil and natural gas. See Figure 2 and 3. Electricity prices have been steadily increasing in India, and the share of electricity in steel plants' energy mix has been steadily decreasing.

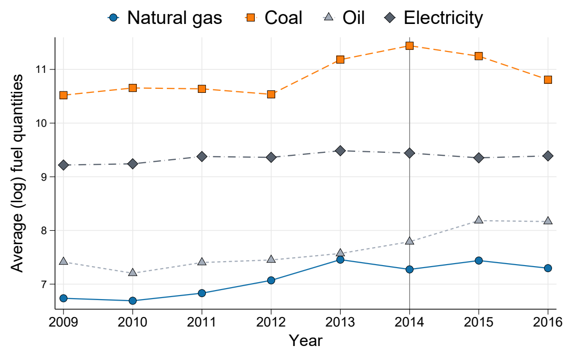


Figure 2: (Average (log) Fuel Quantities

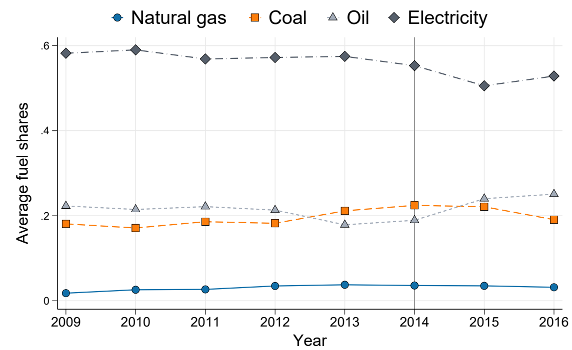


Figure 3: Average Fuel Shares

Notes: these figures show the average quantities of each fuel (in log) and the average within-plant fuel share, respectively. Fuel quantities are measured in mmBtu for each fuel and these figures are constructed from steel-producing plants only.

Fact 3: Indian Steel Establishments Often Switch Between Fuel Sets

Not only is there heterogeneity in fuel sets, but there is also a significant prevalence of switching between fuel sets, occuring when a plant uses different fuels between two years. On average, 15% of plants add a new fuels and 15% of plants drop an existing fuel from there set every year. Moreover, 42% of unique plants add and drop a fuel at least once in the sample. There are many reasons why plants switch between fuel sets. The development of new technologies may increase the productivity of some fuels. Electric arc furnaces are more efficient at using the heating potential of underlying fuels than blast furnaces (Worrell, Bernstein, Roy, Price and Harnisch, 2009). Large and persistent fuel price shocks incentivize plants to readjust their input mix, and expanding transportation infrastructures, particularly pipeline networks, decreases fixed costs and eases access to new fuels (Scott, 2021). The following fact provides further clarity to understand heterogeneity in fuel sets.

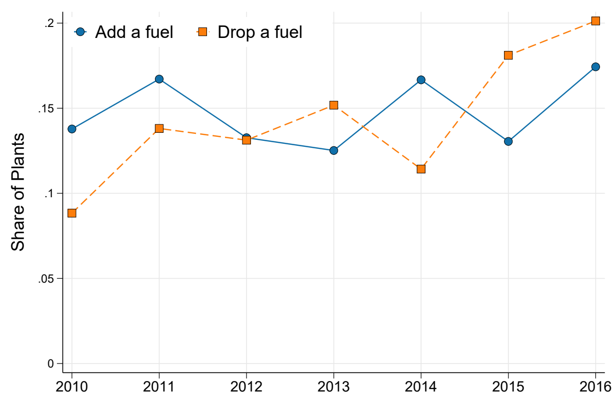


Figure 4: Fuel Set Switching Across Years

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.

	Add New Fuel	Drop Existing Fuel
Yes	0.42	0.42

Table 3: 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 samepl, 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.

Fact 4: Larger Fuel Set is Associated with Higher Productivity

Adopting a new fuel is associated with substantial fixed costs – acquiring and installing new equipment such as furnaces, transportation infrastructure, and storage facilities. In the presence of fixed costs, it is natural to expect larger and more productive plants to use more fuels because they have marginally more to gain from fixed investments. This is what I find. As plants produce more output per worker, which can be interpreted as a proxy for productivity, they tend to have a larger fuel sets. See Figure 5

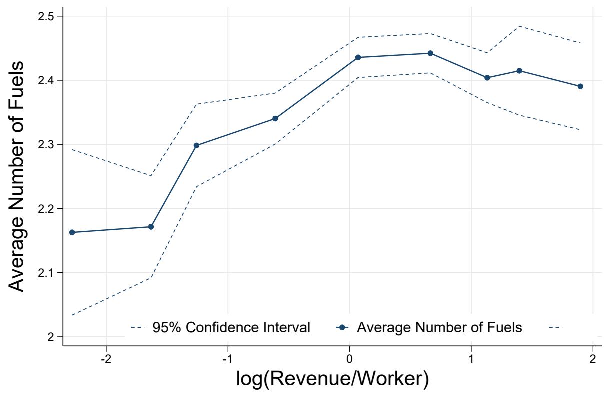


Figure 5: Number of Fuels, by Revenue per Worker

Notes: This figure reports the average number of fuels across percentiles of the residualized revenue per worker distribution. Revenue per worker is residualized by taking out year-fixed effects.

Increasing Revenue/Worker by 1%		
1 Fuel	-0.009***	(0.001)
2 Fuels	-0.025***	(0.004)
3 Fuels	0.023***	(0.004)
4 Fuels	0.010***	(0.002)
Year Fixed Effects	Yes	
N	8,583	

Table 4: Relationship between Revenue/Worker and Number of Fuels

Notes: this table reports marginal effects of an ordered logit regression of the number of fuels against log(Revenue/Productivity). The point estimates are the percentage increase (decrease) in the probability of having each number of fuels for a 1% increase in Revenue per Worker.

In Table 4, I show that a 1% increase in Revenue per Worker decreases the probability of having one or two fuels, and increases the pobability of having 3 or 4 fuels. A positive gradient between the number of fuels and productivity coupled with high fixed costs of new fuel adoption may lock unproductive plants into their current fuel set. A similar technological lock-in has been previously documented in manufacturing by Hawkins-Pierot and Wagner (2022) and decreases the effectiveness of policy at incentivizing switching to cleaner fuels.

4 Model

Consistent with the evidence provided so far, I develop and estimate a rich dynamic production model that allows me to quantify establishments' 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 the outer nest of production. Plants can choose to change fuel sets across periods in a dynamic discrete choice framework. There are fixed costs of adding new fuels and salvage values from dropping existing fuels. I first present the structure of production for a given plant in a static setting and then consider inter-temporal decisions. Throughout the exposition, subscript i refers to a plant, and t refers to a year.

4.1 Production Model

There are two levels of production, which correspond to two nests. The outer nest is a standard CES production function and features Hicks-neutral productivity z_{it}, labor L_{it}, 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 $\overline{X} = \left(\prod_{i=1}^{n} \prod_{t=1}^{T} X_{it} \right)^{\frac{1}{nT}}$.[3]

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

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

Where $\sigma \geq 0$ is the elasticity of substitution between inputs, and $\eta > 0$ is the returns to scale. In the outer nest, plants choose input quantities given input prices, which 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:

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

e_{fit} refers to 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. Plants specializing in dirty fuels such as coal, for example, will bear a disproportionate share of the tax burden. While initially unobserved, I recover fuel productivity for each plant in each year by exploiting profit maximization.

[3]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.

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 takes the form of an option value, which is also relevant in the context of a carbon tax. 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 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 that 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[4]. In Online Appendix B.1, I develop this option value more formally. 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.*

I make this assumption for two reasons. First, the Indian steel industry is very competitive. It has the most plants out of all indian heavy manufacturing industries, and the industry face fierce competition from China, which is the largest steel producer in the world. Second, steel plants produce a wide variety of steel products. There are 404 varieties produced by plants in the ASI. *Ferrous products from direct reduction of iron ore* is the most common variety with a 5.5% market share. See Online Appendix A.3.7.

In each industry, there is a representative consumer with quasi-linear utility over the total output

[4]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)

produced in a given period Y_t and an outside good Y_{0t}. The consumption of steel 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 across varieties. Given a mass of N_t operating plants, income I_t, and an aggregate demand shock e^{Γ_t}, the representative consumer solves:

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

Where $\rho > 1$ is the elasticity of substitution between varieties, and $\theta \in (0,1)$ indexes the substitution between consumption of the differentiated varieties and the outside good. Following Helpman and Itskhoki (2010), I restrict $\theta < \frac{\rho-1}{\rho}$, which ensures that output varieties are more substitutable between each 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) and allow researchers to use the social cost of carbon (SCC), which expresses the net present value of expected future damages from carbon emissions in dollars.[5] Externality damages thus 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 ex-post idiosyncratic demand shock $e^{\epsilon_{it}}$:

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

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

Plants choose inputs to maximize profits

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

Given a set of fuels $\mathcal{F}_{it} \subseteq \mathbb{F}$, technological constraints, inverse demand, and all input prices, the plant's problem is a static profit maximization.[6] 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.

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

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

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

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 year:

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

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

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

Where the unobserved price of realized energy $\tilde{p}_{E_{it}}$ corresponds to a CES price index in fuel prices over productivity. Constant returns in the energy production function imply that the marginal cost

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

of realized energy is the price of realized energy and is constant $MC(\tilde{E}_{it}) = p_{\tilde{E}_{it}}$.

2. Plants choose inputs other than fuels to maximize profit (outer nest):

Given a cost-minimizing allocation of fuels that produce a quantity of energy, plants pay a price p_{Eit} for each unit of energy. They take this price as given when choosing 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 $\pi(\mathbf{s}_{it}, \mathcal{F}_{it})$.

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

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:

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

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[7]:

[7]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 who experienced an expansion of the pipeline network in their district between 2008 and 2016 were more likely to add natural gas to their mix.

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

The switching cost function $k(\mathcal{F}' \mid \mathcal{F}_{it}, d_{it})$ is composed of two types of arguments. First, there are fixed costs of adding a fuel κ_f. Second, there are salvage values of dropping a fuel γ_f that plants obtain by selling old capital. 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: $\mathbb{F} = \Big\{(oe); (oge); (oce); (ogce)\Big\}$ In the next Section, I show how this model can be estimated.

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

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

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

5 Identification of the Production Function

The model is estimated in three steps using a novel combination of methods. First, 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). I then estimate the energy production function following recent developments in production function estimation with input-augmenting productivity (Demirer, 2020; Zhang, 2019), combined with dynamic panel techniques (Blundell and Bond, 1998, 2000, 2023). Lastly, I estimate fixed costs in a dynamic discrete choice framework in the presence of unobserved heterogeneity following Arcidiacono and Jones (2003), which allows me to capture systematic differences in fuel productivity for fuels that plants are not currently using.

5.1 Identification of outer production function

In the outer nest, the main unobserved quantity that departs from standard models is realized energy $\tilde{E}_{it}$. In contrast to the heating potential of fuels, energy is the output of combining different fuels, which is unobserved. I adapt the estimation method proposed by Grieco et al. (2016) to uniquely recover the price and quantity of energy when other flexible inputs are observed under the assumption that plants are price-takers in the input market.[8] 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 look at the ratio of first-order conditions for labor and energy from profit maximization in equation 6, and rearranging:

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

Given production function parameters, $\frac{E_{it}}{\overline{E}}$ can be recovered from (7) because I observe expenditures for both inputs (recalling that energy expenditure is the sum of fuel expenditures from the energy production function: $p_{E_{it}}E_{it} = \sum_{f\in\mathcal{F}_{it}} p_{fit}e_{fit}$) and I observe quantity of labor. 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 $\sigma = 1$ (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.

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

In this setting, one can identify production parameters by replacing $\tilde{E}_{it}$ for (8) in the production function and exploiting first-order conditions to control for the transmission bias from unobserved hicks-neutral productivity z_{it} to observed inputs, a method that is also used by Doraszelski and Jaumandreu (2013, 2018). I also use the same method to control for unobserved price dispersion in the bundle of material inputs:

[8]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.

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

The main dependent variable is revenues, where $e^{u_{it}}$ is an unobserved iid shock 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.1. Taking logs of revenues yields the main estimating equation:

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

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

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

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

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

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

5.1.1 Estimating Elasticity of Demand

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

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

where $\Lambda_t = \tilde{\Gamma}_t + \ln\left(\frac{1}{N_t}\right) + \frac{\rho(1-\theta)-1}{1-\theta} \ln P_t$ 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 variation in demand and supply 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 is going to 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 (11) 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 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 step, fuel quantities, $\{e_{fit}\}_{f\in\mathcal{F}_{it}}$, and fuel prices: $p_{\hat{E}_{it}} = \frac{S_{Eit}}{\hat{E}_{it}}, \{p_{fit} = \frac{s_{fit}}{e_{fit}}\}_{f\in\mathcal{F}_{it}}$. I show how to recover the elasticity of substitution λ, and all productivity terms ψ_{fit}. To do so, I rely on optimality conditions from the energy cost-minimization problem coupled with a Markovian assumption on the productivity of electricity. This effectively combines the dynamic panel approach of Blundell and Bond (2000, 1998) with the method proposed by Zhang (2019). As a reminder, the energy cost-minimization problem of the plant is as follows:

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

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

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

The intuition underlying equation (13) is that relative fuel 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.

At this point, the only unobservable left in the energy production function is the productivity of electricity, which is correlated with current period quantities and spending on fuels since it is assumed to be known to the plant at the time of choosing fuel quantities. To deal with this issue, I assume that the productivity of electricity follows an AR(1) Markov process with year-fixed effects.[11] Moreover, I allow for plant-specific fixed effects in the productivity of electricity $\mu_i^{\psi_e}$

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

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

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

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 $t-1$:

[11] The choice of these modified AR(1) processes where the mean is normalized by the persistence are standard in the dynamic panel literature with short panels (Blundell and Bond, 2023). It ensures that the average of each state variables observed in the data corresponds to the unconditional average of this process. This means that many even though the model is estimated from a short panel (between 2 and 8 years), forward simulations multiple years ahead will match the support of the data. It is equivalent to the assumption that the residuals of the productivity distribution follows an AR(1) process, rather than electricity productivity itself.

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

There are two main endogeneity concerns in this model. First, the lagged value added of energy 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 conditions. I use the system GMM approach, which combines both level and difference moment conditions as follows:

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

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

6 Identification and Estimation of Fixed Fuel Switching Costs

Each plant 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:

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

Where the fuel set switching cost function, $\mathcal{K}(\mathcal{F}' \mid \mathcal{F}_{it}, s_{it})$, was defined in Table 5. It is a function of productivity z_{it} and access to natural gas pipelines. σ_ϵ is a parameter that maps units of profits (dollars) to units of the fixed cost shocks.[12] From now on, I define the parameters governing the switching cost function $\theta_1 = \{\kappa_{g1}, \kappa_{g0}, \kappa_c, \gamma_{g1}, \gamma_{g0}, \gamma_c\}$ for coal c and gas g, and θ_2 the parameters underlying the evolution of state variables. I use κ_{g1} to denote the fixed cost of adding natural gas for plants that are located in a district near the pipeline network and κ_{g0} for plants that are located in a district that isn't immediately adjacent to the pipeline network, and likewise for salvage values. I make the assumption that cost shocks are iid and come from a standardized Type 1 Extreme value $\epsilon_{\mathcal{F}'it} \sim Gumbel(0,1)$. This allows me to analytically integrate out these shocks and work with the expected value function, $W(s_{it}, \mathcal{F}_{it}) = \mathbb{E}(V(s_{it}, \epsilon_{it}, \mathcal{F}_{it}))$:

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

Where $\gamma \approx 0.5772$ is the Euler–Mascheroni constant and $\upsilon_{\mathcal{F}'}(s_{it}, \mathcal{F}_{it})$ is the choice-specific value function. Then, the probability of choosing fuel $\mathcal{F}'$ has a logit formulation, which simplifies the likelihood. Note that this probability is implicitly a function of both θ_1 and θ_2. Next, I discuss the evolution of each state variable.

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

Plants take expectation over all productivity terms, fuel prices and materical prices, which I separate 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.

[12] An equivalent approach would be to map units of the fixed cost shocks to units of profits (dollars). Once σ_ϵ is known, I can always switch between dollars and units of the gumbel distribution.

6.1 Evolution of non-selected state variables

The price/productivity process of both electricity and oil follows a persistent AR(1) process with time (t) and region (r) fixed effects. I model fuel prices and productivity as a single state variable because they always enter together in plants' profit function p_{fit}/ψ_{fit}[13]. $\forall f = \{elec, oil\}$:[14]

$$\ln(p_{fit}/\psi_{fit}) = (1-\rho_f)(\mu_0^f + \mu_r^f) + \mu_t^f - \rho_f \mu_{t-1}^f + \rho_f \ln(p_{fit-1}/\psi_{fit-1}) + \epsilon_{it}^f$$

The region dummies μ_r^f (Northern, North-East, Eastern, Center, Western, Southern) capture the bulk of spatial variation in fuel prices and technology. I also assume a similar process for hicks-neutral productivity $\ln z_{it} = (1-\rho_z)(\mu_0^z + \mu_r^z) + \mu_t^z - \rho_z \mu_{zt-1} + \rho_z \ln z_{it-1} + \epsilon_{zit}$, and the (log) price of materials $\ln p_{mit}$.

6.2 Evolution of selected state variables – systematic heterogeneity

Selected state variables refer to price and productivity pertaining to specific fuels that are only observed when a plant uses that fuel. The transition for these state variables follows a Markovian process which depends on whether plant i is initially using fuel f or not. If the plant isn't initially using fuel f, it has an initial condition $\mu_0^f + \mu_t^f + \mu_r^f + \mu_i^f$ and takes expectation over idiosyncratic shocks ϵ_{it}^f for fuel price/productivity. $\forall f = \{gas, coal\}$:

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

If the plant is initially using fuel f, the transition of fuel price/productivity follows an AR(1) process similar to non-selected state variables:

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

[13] A positive shock to (log) fuel prices is the same as negative shock to (log) fuel productivity, and vice versa. However, I model fuel prices and productivity separately when studying the carbon tax in the counterfactual section because they have a differential effect on production decisions (fuel quantities, energy quantity).

[14] I allow for the shock to productivity to include a plant fixed effect $\mu_i^{\psi_f}$ when estimating parameters of the productivity transition process to get a consistent estimate of the auto-correlation parameters (Blundell and Bond, 1998). However, I do not currently separate this plant fixed-effect from the shocks to ϵ_{it}^f when simulating the choice probability to reduce dimensionality of the dynamic discrete choice problem.

I allow all shocks to productivity and prices to be arbitrarily correlated in a multivariate normal distribution with mean zero and a positive semi-definite covariance matrix $\mathbf{\Sigma}$:

$$\left(\epsilon^o_{it},\epsilon^e_{it},\epsilon^g_{it},\epsilon^c_{it},\epsilon_{zit},\epsilon_{mit}\right)\equiv\epsilon_{\mathbf{it}}\sim\mathcal{N}(\mathbf{0},\mathbf{\Sigma})$$

I allow for systematic differences in the productivity of coal and gas across plants, $\mu_i^{\psi_f}$, which I call *fuel comparative advantage*. These comparative advantages, which are assumd to be known to plants when making decisions but unknown to the research, underlie a central econometric issue of selection on unobservables. This selection creates a non-trivial challenge to both the estimation of fixed costs and counterfactual policy experiment. Indeed, recovering the distribution of $\mu_i^{\psi_f}$ only from plants that use coal and gas in a given year may not reflect the true distribution, which would bias counterfactual fuel choice predictions under alternative policy regimes.

6.3 Identification of 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 a finite mixtures with $K=3$ points of support for each fuel. I parameterize the initial guess of the mean and variance of the finite mixture to the mean and variance of the empirical (selected) distribution $(\tilde{\mu}_f,\tilde{\sigma}^2_{\mu_f})$:

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

Where $\pi^0_{fk}=Pr(\mu_{fk})$ is the unconditional probability of being type k, where types refer to support points of the fuel comparative advantage distribution, and $\sum_k\pi^0_{fk}=1$. In this context, external estimation of parameters governing the distribution of random effects from a selected sample of plants 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 over the unconditional distribution of comparative advantages using the full information (log) likelihood. Assuming there is only one finite mixture over both coal and gas for notation convenience, and where the distribution of comparative advantages are independent across fuels such that $\pi_k\in\Pi=vec(\Pi_g\otimes\Pi_c)$, where $\pi_{kg}\in\Pi_g$ and $\pi_{kc}\in\Pi_c$:

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

In particular, the likelihood in (18) assumes that the state transitions are independent of the distribution of comparative advantages for coal and gas.[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 that are consistent with this assumption. Initially, the true probability weights π_k over the support of the finite mixture are unknown due to selection, but Arcidiacono and Jones (2003); Arcidiacono and Miller (2011) 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 choices that we 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:

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

[15] This assumption isn't necessary but it simplify 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 (19), which are used to then update the unconditional distribution of comparative – E step, and repeat this procedure until the likelihood in (18) is minimized. Details of the algorithm can be found in Appendix C.3.

7 Estimation Results – Steel Manufacturing

7.0.1 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 for 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 8, which I use to estimate the energy production function.

Table 6: Production Function Estimation (Steel Manufacturing)

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

Bootstrap 95% confidence interval in brackets (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 $x_{jit} \in \{L_{it}, K_{it}, M_{it}, E_{it}\}$

7.0.2 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 $\hat{\sigma}$

Table 7: Estimates of Energy Production Function

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

Standard errors in parentheses

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

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

from Table 6. This is important because the larger the elasticity of substitution between fuels, the larger the aggregate gains from carbon taxation (Acemoglu, Aghion, Bursztyn and Hemous, 2012). More substitution possibilities mean that 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 construct estimates of the fuel-specific productivity for each plant in each year $\hat{\psi}_{fit}$, and I discuss its implications.

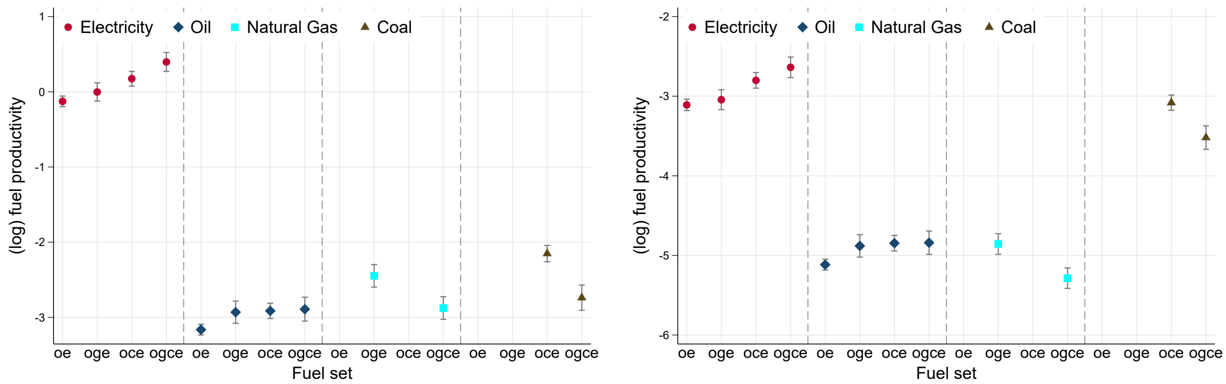
Heterogeneity in fuel productivity across fuel sets

Results are summarized in Figure 6. There are a few takeaways. First, electricity is by far the most productive fuel, averaging between two and three times the productivity of other fuels. When comparing coal and natural gas, 1 mmBtu of coal is, on average, 30% more productive than 1 mmBtu of gas. However, this gap expands to 185% when looking at productivity per dollar because coal is significantly cheaper – averaging one-fifth of the price of natural gas. This can explain the prevalence of coal in the Indian steel industry.

7.0.3 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. On average, these costs are substantial, ranging from 28 to 40 million dollars, and align well with the upper echelon of existing accounting estimates.[16]

[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:



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

Figure 6: Mean (log) fuel productivity among plants using different fuel sets

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

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

Table 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.

Coal adoption tends to be 30% cheaper than gas adoption. Moreover, plants without access to high-pressure natural gas pipelines incur 40% higher costs of adoption due to the need for alternative transportation methods, such as liquefied natural gas (LNG), which can be costly. 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 both coal and natural gas are significantly lower, ranging from 57% to 71% below the 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 $\$8,200$ increase in fixed costs and a $\$2,500$ increase in salvage values. Importantly, the combination of substantial fixed costs and relatively low salvage values likely contributes to situations of technology lock-in, which I discuss in the next section.

7.1 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 using natural gas than plants that do, whereas this effect goes up to 80% for coal. Combined with high fixed costs, this gap in productivity 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 tend to face a lower marginal cost of energy. This difference is largely explained by existing variation in fuel productivity, rather than by the additional substitution margin that a new fuel provides. In such a context, there are little incentives for plants to pay the large fixed costs required to add natural gas.

---

alibaba's listings https://www.alibaba.com/product-detail/WONDERY-Custom-Made-Siemens-PLC-Industrial_1600732474634.html), whereas switching from pig iron, typically produced with coal-powered blast furnace, to direct reduced iron, typically produced with gas-powered oxygen furnaces would historically cost upwards of USD 70 millions Miller (1976).

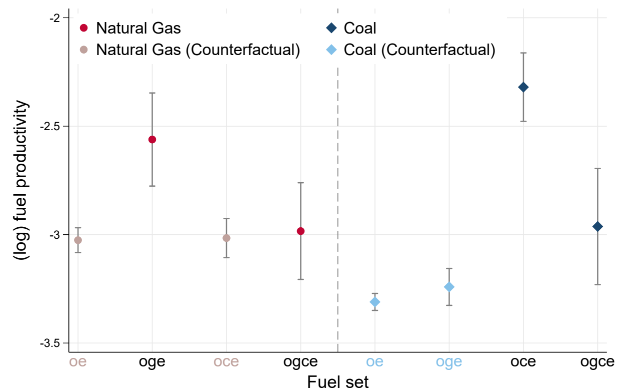


Figure 7: Distribution of fuel productivity – Including counterfactual fuel sets

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

Model fit

Overall, the estimates of switching costs 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. In all figures below, the blue bars (model) are constructed by adding the predicted probability that each plant uses each fuel set, integrated over the conditional distribution of comparative advantages. Details in Online Appendix D.1.

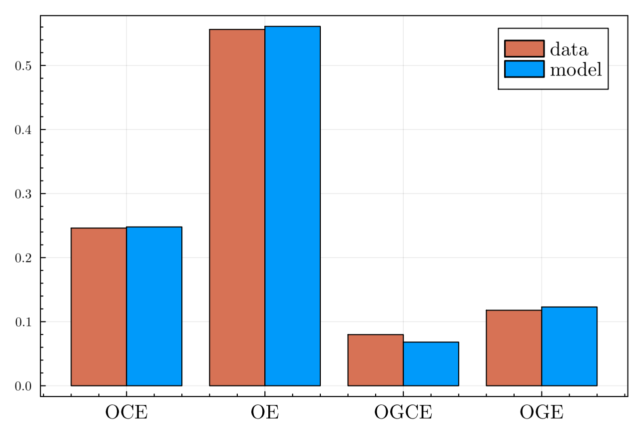


Figure 8: Unconditional distribution of fuel sets, model vs. data ($N = 2,393$)

I also look at the relationship between plant size (proxied by Hicks-neutral productivity) and fuel switching, and find a positive relationship between the probability that plant adds a fuel at $t+1$ and its productivity at t, both in the data and in the model. See Figure 10.

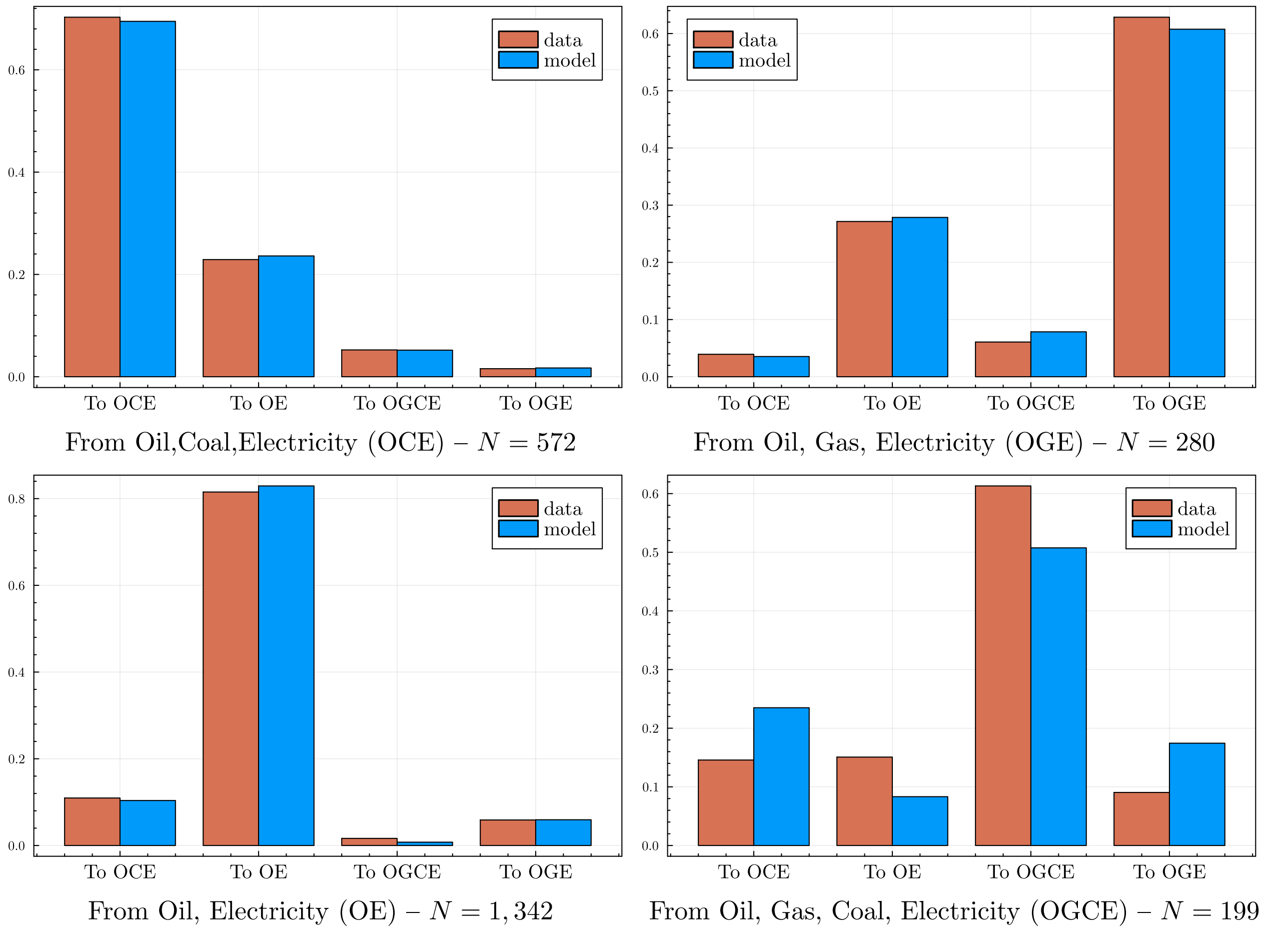


Figure 9: Conditional distribution of fuel sets (transition), model vs. data

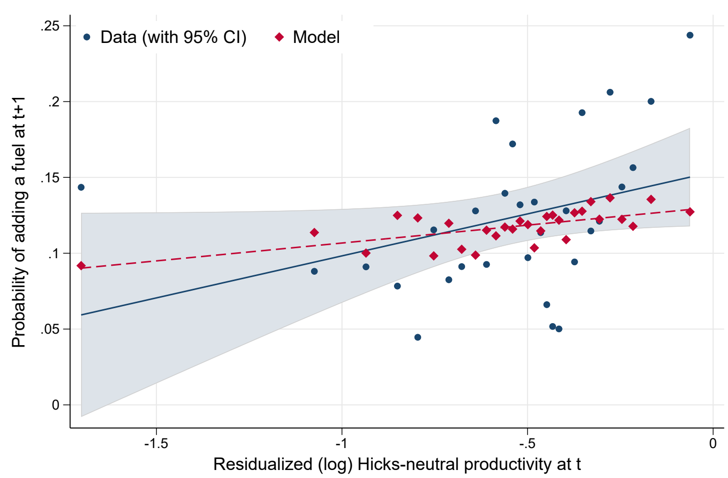


Figure 10: Adding a fuel at $t+1$ against Hicks-neutral productivity at t

This figure is a binned scatter plot projecting fuel switching $t+1$ against Hicks-neutral productivity at t. For the data, I do a linear projection of whether plants added a fuel between t and $t+1$ against productivity. For the model, I do the same, replacing choices made by the predicted choice probability. I control for both year fixed effects and input prices in the model and in the data.

8 Externality Mitigation - Steel Manufacturing

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

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

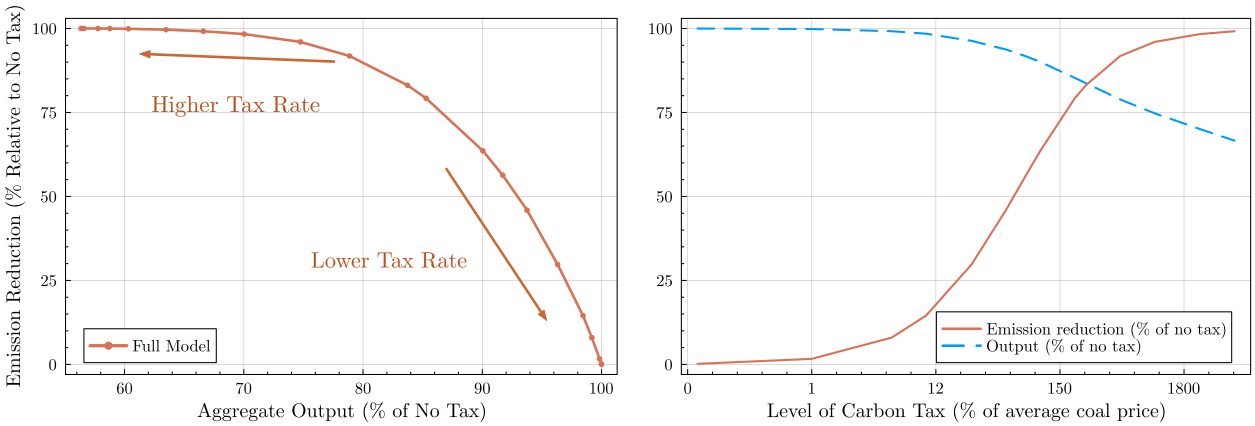
Table 9: Example of Average Fuel Prices With and Without Carbon Tax

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

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

In Figure 11, 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 and look at the net present value (NPV) of outcomes along the entire path. As the level of the tax approaches zero, the model converges to the no-tax economy with 100% of output and 0% of emission reduction. As the level of the tax increases, emissions decreases but so does aggregate output.

The production frontier is concave because of 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 substituted towards cleaner fuels.



(a) Trade-off between Output and Emission Reduction (b) Emission Reduction and Output Across Tax Levels

Figure 11: Production Frontier in Output and Emission Reduction for Various Carbon Tax Rates

Notes: This production frontier was constructed by simulating the economy under 21 different levels of the carbon tax, ranging from 0 (no tax) to approximately infinity. Linear interpolation is assumed for the trade-off 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, but plants always use some positive amount of fuel.

8.1.1 The Role of Heterogeneity in Fuel Productivity – Intensive Margin

I compare in Figure 12 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.1, and follow the dynamic panel approach, similar to the energy production function with heterogeneity in fuel productivity. Notably, estimating this production function matches average fuel quantities and aggregate emissions levels, but misses the heterogeneity in fuel shares across plants.

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

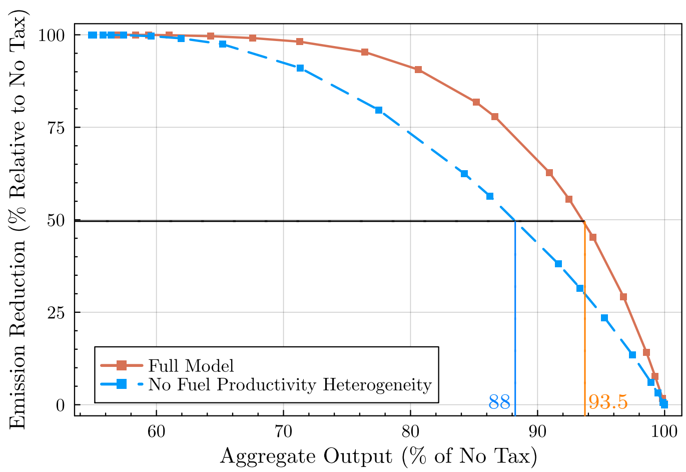


Figure 12: Comparison of Trade-off Across Model Specification

Notes: This production frontier highlights the role of the intensive margin only. It compares the trade-off across contemporaneous taxes, excluding forward simulations and switching between fuel sets (extensive margin). The comparison of trade-offs is the same when allowing for the extensive margin to adjust in a forward simulation because heterogeneity in fuel productivity affects contemporaneous reallocation across plants.

The blue production frontier in figure 12 corresponds to this economy and shrinks inwards when compared to the baseline economy. To reduce emissions by 50%, the restricted economy operates at 88% of no-tax output, with an implied elasticity between emission reduction and output of 4.17 – compared to 93.5% of no-tax output in the flexible economy, with an implied elasticity of 7.7. Allowing for heterogeneity in fuel productivity diminishes how much output must decrease to achieve any reduction in emissions because it increases the aggregate elasticity of substitution between fuels by reallocating output from dirty to clean plants. Two channels explain this reallocation.

First, conditional on fuel prices and fuel sets, the elasticity of the marginal cost of energy with respect to relative fuel prices is not constant across plants. For example, as the price of coal increases relative to the price of gas, plants that are more productive at using coal relative to gas face a larger percentage increase in their cost of energy. This is because larger coal productivity induces specialization in coal, making them more exposed to the relative price increase, as long as fuels are gross substitutes $\lambda > 1$. Since the carbon tax effectively increases the relative price of polluting fuels, plants that are more productive at using polluting fuels are especially hurt by 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,

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

Under cost-minimization, the elasticity of the marginal cost of energy with respect to the relative price of any fuel (here 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. Details in Appendix D.1.1. 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:

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

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.

Second, this heterogeneous increase in marginal energy costs makes polluting plants less competitive than cleaner plants, consistent with the aggregation result by Oberfield and Raval (2021). The tax thus induces a reallocation of output from more polluting to less polluting plants, which increases aggregate fuel substitution and diminishes how much aggregate output must be reduced to achieve any emission reduction target[17]. Moreover, this reallocation channel is a function of the elasticity of demand and returns to scale. Indeed, as the elasticity of demand increases and different output varieties become more substitutable with each other, any variation in relative marginal costs across plants will lead to a larger reallocation of output. In Appendix D.3, I confirm this

[17]Note that the correlation between fuel productivity and total factor productivity (TFP) also matters for this result. Indeed, if plants that are more affected by the carbon tax were also initially more productive overall, this reallocation effect may reduce aggregate TFP and aggregate output. However, I show in Appendix D.2 that the opposite is true. Plants with higher fuel productivity tend to be less productive overall.

intuition by showing that the difference between both production frontiers expands as the elasticity of demand increases. In summary, not allowing for this rich heterogeneity between fuel productivity shuts down the reallocation of output from high-emission to low-emission plants and decreases the effectiveness of a carbon tax.

To benchmark this result with the literature, I do two exercises. First, I compare the aggregate trade-off between output and emissions with Fowlie et al. (2016), who conduct similar policy exercises for U.S. cement plants. Crucially, their margin of interest is 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.

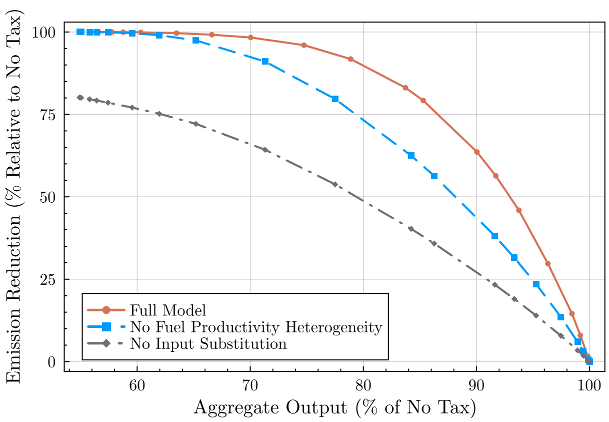


Table 10: Comparison of Trade-off Including No Input Substitution

	Average Elasticity $\frac{\%\Delta CO_{2e}}{\%\Delta Y}$
Full Model	5.46
No Fuel Producitivty 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.

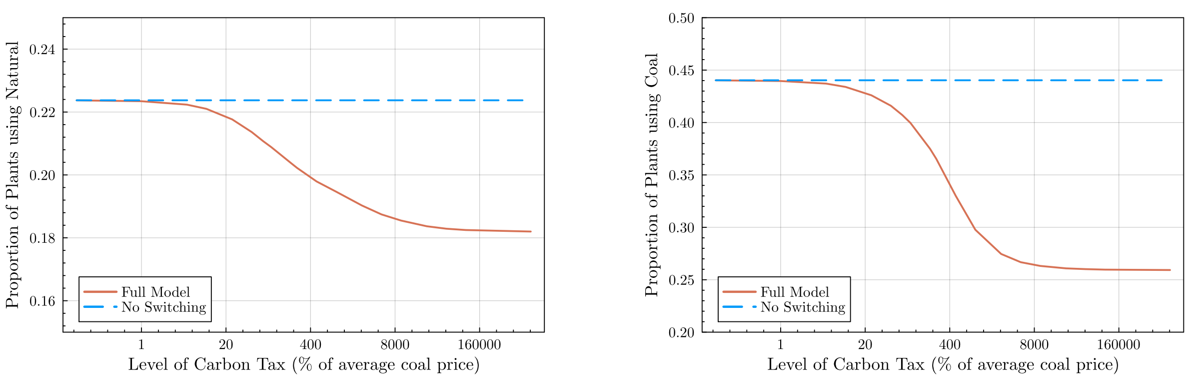
Second, I show that heterogeneity in fuel productivity and inter-temporal switching between fuel sets also serve as a cautionary tale for larger-scale climate models in which an aggregate production function in different fossil fuels is typically assumed as part of a larger integrated assessment model (IAM). Such models study the relationship between climate change and economic growth. For example, Golosov et al. (2014) and Miftakhova and Renoir (2021) postulate an aggregate CES production for composite energy that combines different fuels. In Online Appendix D.3, I show that such an aggregate CES production function can be micro-founded from an economy in which

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

many plants have an energy production function without heterogeneity in fuel productivity and where plants have access to different fuel sets but cannot switch between them. Given the results of this section, the use of such a production function may understate the extent of fuel substitution as a response to policy because it does not capture the reallocation mechanism induced by heterogeneity in fuel productivity and the inter-temporal substitution between fuel sets.

8.1.2 Technology Lock-in – Ineffectiveness of Carbon Tax at the Extensive Margin

While a carbon tax can cost-effectively reduce emissions through its both effect on fuel substitution at the intensive margin and by 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 that use either 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 13: Natural Gas and Coal Take-up Across Levels of Carbon Tax

Overall, the lack of natural gas adoption indicates that the carbon tax is ineffective at incentivizing plants to overcome technology lock-in. This can be partially explained by a combination of the facts that the carbon tax also raises the price of natural gas, fixed costs of adoption are economically high – especially for plants away from the pipeline network, and plants who are not currently using natural gas would be on average 30% less productive at using natural gas compared to those who already use it. In the next section, I explore the effectiveness of a subsidy to incentivize natural gas take-up. In Figure 14, I show that if fixed costs were removed from the economy, every plant would adopt all fuels, which would reduce the trade-off between outpout and emission reduction

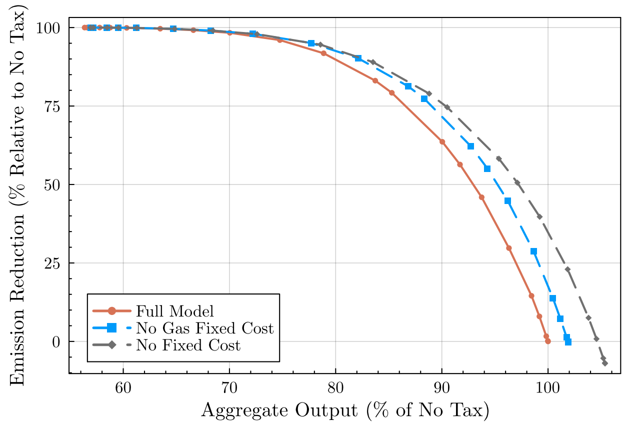


Figure 14: Comparison of Trade-off When Removing Fixed Costs

At the baseline, both output and emissions would be greater in the economy because of the option value that more fuels provide. As the carbon tax rate increases, more emissions emissions can be reduced through fuel substitution because of the additional substitution margin that these new fuels provide. This is true both when all fixed costs are removed and when only natural gas fixed costs are removed.

8.2 Alleviating Technology Lock-in - Combining the Carbon Tax with Natural Gas Subsidy

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. Below I show results for various permanent subsidies, ranging from 0% to 100% of the average fixed cost of natural gas. I do these experiments jointly with a carbon tax. To choose a representative social cost of carbon (SCC), I first set the social discount rate to 3% ($\beta = 0.97$) to match India's average real interest rate during the sampled period. Then, following the most recent estimates from the Inter-agency Working Group on the Social Cost of Carbon (IWG, 2021), I set the SCC to the 2020 estimates for a social discount rate of 3% at USD $51/tCO_{2e}$. This SCC corresponds to a mid-range estimate in the literature. In Figure 15, I show what happens to carbon tax revenues/externality damages and total subsidy paid out as the subsidy rate increases

First, up to 10% of the subsidy can be fully financed by a carbon tax in expectation. More importantly, carbon tax revenues monotonically increase as the subsidy rate increases. I show this more clearly in Figure 16, where I compare the evolution of the tax revenues with the fraction of plants that use natural gas and coal.

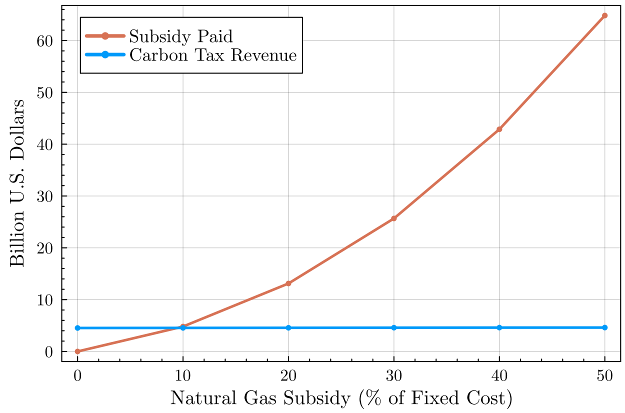
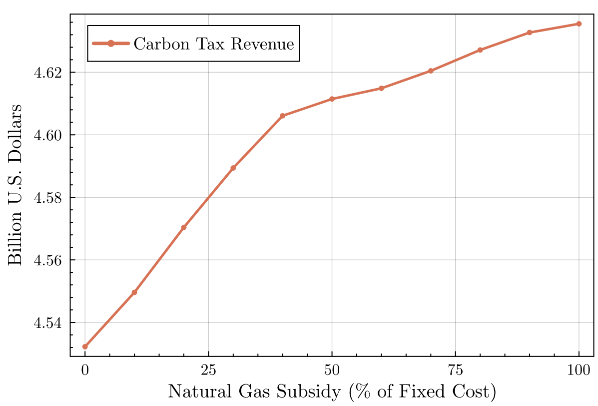
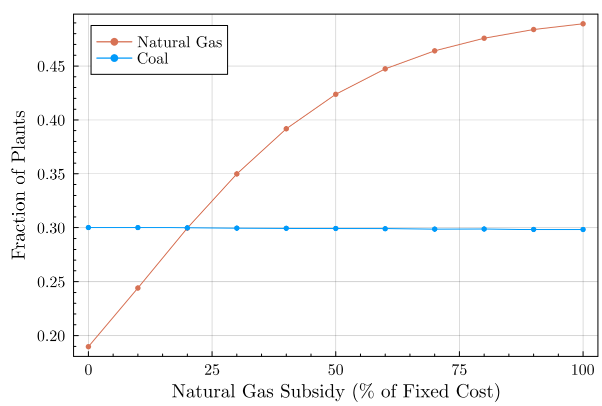


Figure 15: Carbon tax revenue and subsidy paid out along subsidy rate

Notes: This figure was calculated by simulating the expected total tax revenues and subsidy paid out to plants for a horizon of 40 years. Note that carbon tax revenues correspond to externality damages since carbon tax rates on each fuel are equal to marginal externality damages. As such, the evolution of the carbon tax revenues is indicative of emissions.



(a) Carbon Tax Revenues/Externality Damages



(b) Fraction of Plants who use Natural Gas/Coal

Figure 16: Comparison of Selected Outcomes Along Subsidy Rate

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

These results are important. As the subsidy rate increases, more plants add natural gas, but the fraction of plants that use coal remains almost the same. This isn't surprising because coal is still significantly cheaper than gas, and the salvage value of coal is much lower than gas. In this context, it makes more sense for plants to keep coal for the option value it provides. These results create two countervailing effects on carbon tax revenues/externality damages. On the one hand, plants that add natural gas substitute away from more polluting fuels such as oil, electricity, and coal. This substitution effect reduces tax revenues and externality damages. On the other hand, as plants add natural gas, they have more fuels available, increasing their option value and decreasing the price of energy. The net effect is a decrease in marginal costs of production and an increase in output, which increases all input demand. Thus, even with substitution towards natural gas, the input demand for other fuels such as coal, oil, and electricity goes up, which increases tax revenues and externality damages. In Figure 18, I do a full Shapley decomposition of the change in tax revenues as the subsidy rate increases between these two channels. Details of the Shapley decomposition can be found in Appendix D.0.1.

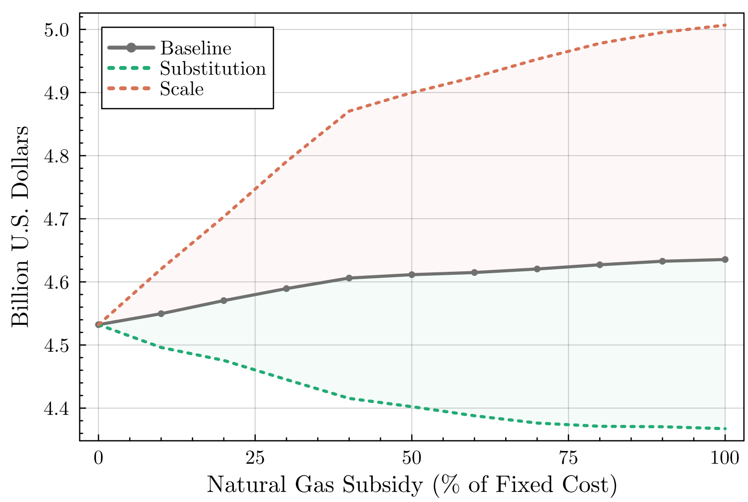


Figure 18: Shapley Decomposition of Changes in Tax Revenues

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

Welfare Analysis of a 10% Subsidy

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

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

Where consumer surplus is 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 fixed-cost subsidy towards natural gas adoption, I make some simplifying assumptions for tractability. I assume that the subsidy is financed by government revenue from the carbon tax and that every plant faces the same permanent subsidy amount s. In this context, producer surplus is the sum of total profits net of subsidized fixed costs, and net government revenue is total tax revenues minus subsidy paid out.

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

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

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

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

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

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

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

To understand how small the welfare effects are, I compare in Table 13 how much of the total fixed costs are financed by the subsidy. While 2.7 billion dollars go towards adopting natural gas, the fraction of plants that use natural gas only goes up by 20% from 0.18 to 0.24, while variable profits and consumer surplus jointly increase by 31 million dollars. Hence, private gains from the subsidy are only 1.1% of the policy's cost.

There are a few reasons explaining this small effect. First, by virtue of being a universal subsidy,

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

Table 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 exactly cancel out in the welfare function. The subsidy also cancels out because it is simply a transfer from net government revenue towards producer surplus. As a result, variable profits, consumer surplus, and total fixed costs are the remaining components in the welfare function such that $Welfare = ConsumerSurplus + VariableProfit - TotalFixedCost$

	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.

the government effectively finances the adoption of natural gas for plants that would have still adopted natural gas in the absence of the subsidy. This can be seen from the increase in total fixed costs paid in the economy by 31.61 million dollars in Table 12 after the policy, which is considerably lower than then total amount paid by the subsidy (2.7 billion). Second, plants at the margin who are actually incentivized to adopt natural gas in the aftermath of the policy are, on average, 30% less productive at using natural gas than plants that already use natural gas. See Figure 7. At the same time, natural gas is, on average, less productive per dollar invested into it than any other fuel. See Figure 6b.

This small welfare effect raises the question of whether the government could find more profitable avenues to invest proceeds from the carbon tax 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 emissions ex-post. While outside the scope of this paper, this is an interesting avenue for future research.

9 Conclusion

In this paper, 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 aimed at reducing emissions at the lowest cost possible, including a carbon tax and a carbon tax with a subsidy towards adopting cleaner fuels.

I show that novel features of this model have important quantitative implications for the scope of these policies. Indeed, carbon taxation is much more targeted towards high-emission plants than previously thought due to 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 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.

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

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

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

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

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.

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

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

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

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

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

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

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

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

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.

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

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

Environmental Protection Agency - Center for Corporate Climate Leadership

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

U.S. Environmental Protection Agency

U.S. Environmental Protection Agency, "Carbon Dioxide Emissions Coefficients – Methodology," 2023. https://www.eia.gov/environment/emissions/includes/methodology.php.

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.

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

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

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

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

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

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

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

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.

Jack, B. Kelsey, "Environmental economics in developing countries: An introduction to the special issue," *Journal of Environmental Economics and Management*, 2017, *86*, 1–7.

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.

Lei, Tianyang, Daoping Wang, Xiang Yu, Shijun Ma, Weichen Zhao, Can Cui, Jing

Meng, Shu Tao, and Dabo Guan**, "Global iron and steel plant CO_2 emissions and carbon-neutrality pathways," *Nature*, 2023.

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.

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.

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

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

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

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.

Ritchie, Hannah, Max Roser, and Pablo Rosado, "CO and Greenhouse Gas Emissions," *Our World in Data*, 2022. https://ourworldindata.org/co2-and-greenhouse-gas-emissions.

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.

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

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

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

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

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

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 the survey by the Government of India's

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

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

A.2 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 is 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 which 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 is done slightly differently than from fossil fuels because emissions comes from production rather than end usage of electricity. Figure 19 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.

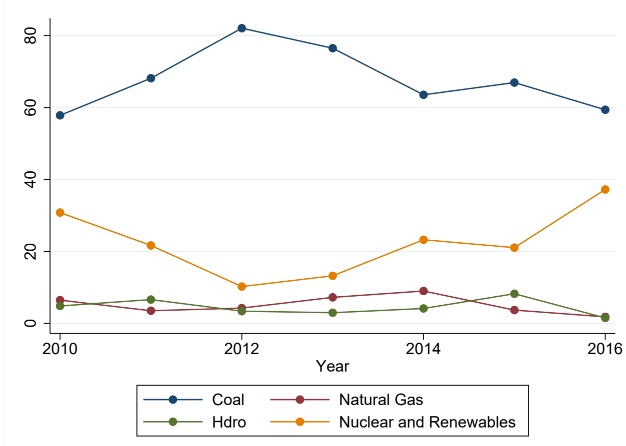


Figure 19: Annual Indian Electricity Generation by Source (% of Total)
Source: International Energy Agency (IEA)

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

B Model

B.1 Closing the Model: Aggregation details

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

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

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

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

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. Keeping income constant, consumer surplus is decreasing in the aggregate price index. Then, the representative consumer chooses which varieties to allocate for a given quantities of good Y_t by minimizing the cost of different varieties:

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

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

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

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

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

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

C Identification

C.1 Derivation of Estimating Equation for Outer Production Function

Production function:

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

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

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}:

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

E_{it}/L_{it}:

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

L_{it}:

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

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

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

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

Plug (28) into revenue equation:

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

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

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

Estimating Equation:

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

C.2 Computational Details on Solving the Dynamic Discrete Choice Model

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

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

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

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

Where t corresponds to year of observation and d is an indicator for access to natural gas pipeline.

I approximate this high dimensional expected value function by discretizing the state space and the underlying Markov process. Since most state variables are highly persistent AR(1) processes, I use Rouwenhorst (1995) to discretize the process. Let M be the number of points on each grid. I am currently using $M = 4$ which gives me $4^6 = 4,096$ grid points for the Markovian state variables. When adding the 6 years of observations between 2010 and 2015 as well as the access to pipeline indicator, I get $(4^6) * 6 * 2 = 49,152$. However, the curse of dimensionality really starts to kick in when I add the distribution of comparative advantages for gas and coal (see later sections). With three possible values for gas and coal, this gives me 9 possible combination of comparative advantages. Ultimately, I am left with $(4^6)*6*2*9 = 442,368$ grid points. Using this discretization process, I can then represent the value function as a block matrix $\vec{W}^n$ that contains all combinations of states. Let S be the set of all state variable combinations, $\Gamma(s' \mid s)$ be the vector of all state transition probabilities when starting at state s (in vectorized form), $\mathbf{\Pi}$ be the vector of all possible profit combinations, $\vec{\mathcal{K}}$ be the vector of all possible fuel set switching costs. Then

$$\vec{W} \approx \gamma + \log\left(\sum_{\mathcal{F}' \in \mathbb{F}} \exp\left(\mathbf{\Pi} + \vec{\mathcal{K}}(\mathcal{F}') + \beta\Big[\bigotimes_{s \in S} \Gamma(s' \mid s)\Big]^T \vec{W}\right)\right) \tag{30}$$

Lastly, to reduce computational burden, I iterate over equation (30) by paralellizing across all possible combination of starting states using graphics processing units (GPU) Arrays with CUDA. Computational gains using GPU Arrays are significant order of magnitude over standard CPU parallelization. Detailed Julia code is available on my Github.

C.3 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.2.

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

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

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

Estimation then proceeds iteratively as follows:

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

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

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

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

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

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

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

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

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

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

D Counterfactuals

D.0.1 Shapley-Owen-Shorrocks Decomposition of the Carbon Tax Revenues Along Subsidy Rate

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

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

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

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

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

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

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

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

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

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

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

D.1 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:

D.1.1 Elasticity of the Price of Energy with Respect to Relative Fuel Prices

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

Table 16: Estimates of Energy Production Function with Energy Productivity

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

Standard errors in parentheses

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

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

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

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

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:

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

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:

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

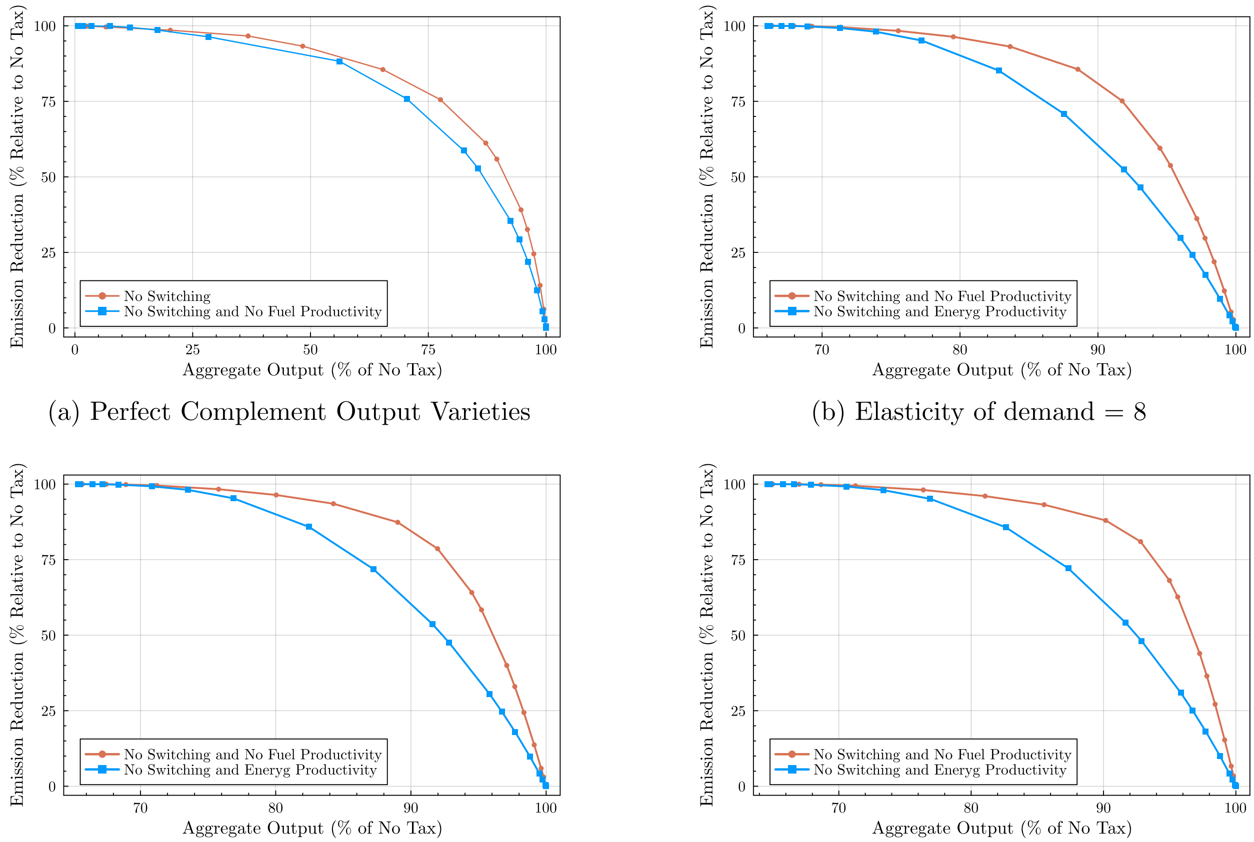
D.2 Correlation between Fuel Productivity and TFP

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

Table 17: Correlation Matrix of Fuel Productivity and Total Factor Productivity

D.3 Trade-off Across Values of Demand Elasticity

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



(a) Perfect Complement Output Varieties

(b) Elasticity of demand = 8

(c) Elasticity of demand = 10

(d) Elasticity of demand = 12

Figure 20: Graphs comparing and contrasting the difference in the aggregate trade-off between emission reduction and output, across different elasticities of demand.

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

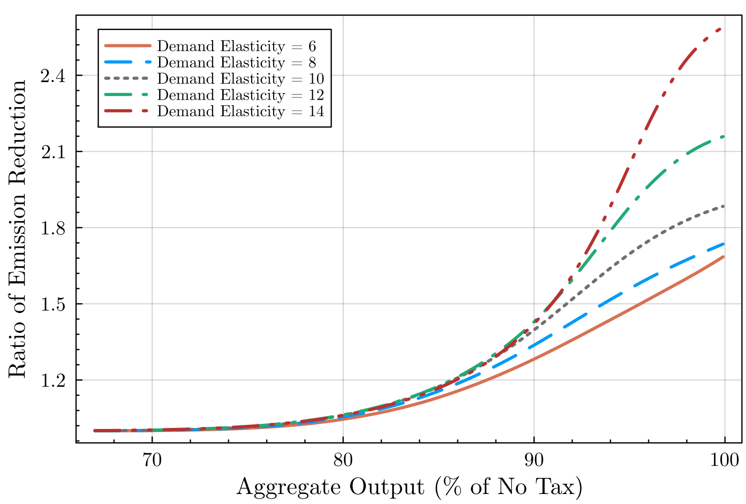


Figure 21: Comparison of the gap between the economy with fuel productivity and the economy with energy productivity.