Carbon Pricing and Fuel Switching by Firms: Theory and Evidence*

Maija Kaartinen† and Ossian Prane‡

November 14, 2024

(most recent version here)

Abstract

What determines the effectiveness of carbon pricing in reducing CO2 emissions? Prior research points to the importance of firm-level technology adjustments but provides limited evidence on the specific technologies involved. We develop and estimate a model of heterogeneous firms' fuel choices to quantify the importance of fuel switching. The model explains three empirical findings from Swedish microdata: (1) fuel choices vary both across and within industries, (2) fuel switching accounts for a large share of the decline in manufacturing emissions between 2004 and 2020, and (3) higher carbon taxes have led firms to switch away from fossil fuels toward electricity and biofuels. Counterfactual analysis suggests that carbon pricing is effective in reducing manufacturing emissions, with fuel switching explaining roughly half of this effect. Market share reallocation toward cleaner firms further reduces emissions but increases the adverse impact on output. The cost of carbon pricing varies considerably across industries, due to differences in the firms' abilities to replace fossil fuels with efficient alternatives.

*Ossian Prane is grateful to his advisors Rikard Forslid and Horng Chern Wong for their guidance, encouragement, and support throughout this project, Pol Antràs for his helpful comments at the project's early stages, and Anders Akerman, John Hassler and Anna Tompsett for valuable feedback at the later stages of the project. We also thank participants in conferences and seminars at Aarhus University, Stockholm University, and the Institute for International Economic Studies.

†Institute for International Economic Studies, Stockholm University. E-mail: maija.kaartinen@iies.su.se

‡Department of Economics, Stockholm University. E-mail: ossian.prane@su.se

1 Introduction

The manufacturing sector accounts for a quarter of global energy-related CO2 emissions. Reducing the use of fossil fuels in this sector is therefore critical to limit global warming (International Energy Agency, 2023). While carbon pricing is widely considered the most cost-effective policy to do this, its effects on emissions and output depend on two key responses: (1) firm-level adjustments of the production process and (2) market share reallocation toward cleaner and more productive firms. In the absence of these responses, carbon pricing will reduce emissions primarily via lower production volumes (Copeland, Shapiro, and Taylor, 2022). Prior research offers limited quantitative evidence on how firm-level adjustments and market share reallocation matter for the aggregate effects of carbon pricing. A potentially important factor shaping both of these responses is firms' fuel choices, which are directly linked to their CO2 emissions. Studying these choices has been challenging for two reasons: a lack of comprehensive data on firm-level fuel use and a lack of structural frameworks that explicitly capture these choices.

In this paper, we examine the importance of firms' fuel choices for the aggregate effects of carbon pricing in the manufacturing sector. We overcome previous challenges by leveraging Swedish administrative data on firm-level fuel use and by developing and estimating a structural model of heterogeneous firms' fuel choices. The Swedish data cover the entire manufacturing sector from 2004 to 2020 and provide detailed information on the types of fuels used in production. The model captures fuel choices at both the extensive margin (i.e., which fuels to use) and the intensive margin (i.e., how much of each fuel to use), and how these decisions are influenced by industry, firm, and fuel characteristics. The aggregate effects of carbon pricing depend on firm-level adjustments via fuel switching and reallocation of production via heterogeneity in firms' fuel choices.

The paper presents three complementary analyses. We first perform a decomposition analysis to show that firm-level fuel switching explains a large part of the aggregate decline in fossil CO2 emissions in the Swedish manufacturing sector. We then conduct a reduced-form analysis, exploiting firm-level variation in the exposure to the Swedish carbon tax, to show that firms have switched away from fossil fuels in response to higher carbon prices. Finally, we use our model to analyze how an increase in the carbon tax affects aggregate emissions and output. The analysis shows that fuel switching explains about half of the aggregate effect on emissions and that the negative impact on output is amplified by market share reallocation away from the most productive firms. We also find large heterogeneity in the effects across industries, and that this heterogeneity is primarily driven by differences in the firms' abilities to substitute fossil fuels with efficient alternatives.

The first contribution of the paper is to demonstrate the importance of fuel choices for the evolution of manufacturing emissions in Sweden. We derive a new statistical decomposition that decomposes changes in aggregate emissions into three margins of adjustment: (1) changes in aggregate energy use, (2) reallocation of energy use across firms, and (3) within-firm fuel switching. Emissions from fossil fuels have declined by 38% in the Swedish manufacturing sector between 2004 and 2020, and each margin explains about one-third of this decline. However, the relative importance of these margins varies across industries. In the basic metals industry, emissions have declined by 25% and this decline is almost entirely explained by lower aggregate energy use. In the paper and pulp industry, emissions have declined by 75% and this decline is primarily due to within-firm fuel switching. The large heterogeneity across industries suggests that firms' abilities to switch away from fossil fuels depend on industry-specific characteristics.

The second contribution of the paper is to provide empirical evidence that firms' fuel choices respond to higher carbon prices. We estimate the effect of carbon prices by exploiting shocks in the Swedish carbon tax. Sweden has one of the highest carbon taxes in the world and it has increased by more than 500% between 2004 and 2020. We use two complementary approaches to identify the effect of the carbon tax. The first approach is to construct a shift-share instrument that captures differences in the firms' exposure to the carbon tax based on fuel use in a pre-sample year. Intuitively, firms that use large shares of fossil fuels are more exposed to changes in the carbon tax than firms that use large shares of biofuels or electricity. The second approach is a matched difference-in-differences design in which firms covered by the EU ETS act as a control group. We use this second approach to identify the effect of a sharp increase in the Swedish carbon tax in 2015.

The two approaches yield consistent results and suggest that firms respond to higher carbon taxes by substituting away from fossil fuels toward electricity and biofuels. We find significant effects at both the intensive and extensive margins of the firms' fuel decisions. Firms respond by using relatively less fossil fuels in their fuel mix and part of this effect is explained by firms dropping fossil fuels completely. The effects are sizeable and imply that increases in the Swedish carbon tax account for 67% of the average reduction in the firms' fossil fuel reliance over the sample period. However, to fully understand the effects of carbon pricing on aggregate emissions, we also need to account for changes in the allocation and scale of production across firms.

The third contribution of the paper is to develop a structural model that quantifies how fuel choices matter for the aggregate effects of carbon pricing. We use modeling tools from the quantitative trade literature, and in particular Antras, Fort, and Tintelnot (2017), to embed heterogeneous firms' fuel decisions in a quantifiable general equilibrium framework. Our model comprises a set of manufacturing industries and an energy sector producing various fuels. For simplicity, we consider a closed economy and abstract away from long-run technical change. Nevertheless, the model captures the effects of carbon pricing on all three margins highlighted by the statistical decomposition: changes in aggregate energy use, reallocation of energy use across firms, and within-firm fuel switching. Our model also accounts for key patterns of fuel consumption in the data: heterogeneity in fuel use across industries, heterogeneity in fuel use across the firm size distribution, and heterogeneity in the relative importance of fuels within firms.

The firms in our model require energy for each of their production stages and make two key decisions. The first decision is to choose the optimal set of fuels to use in production by investing in fuel-specific technology. The second decision is to choose which fuel in this set to use for each production stage. The value of investing in a specific fuel depends on the fuel's production potential, which is a function of its price and technological state, as well as how it is combined with other fuels. Fuels with higher production potentials are more likely to be the lowest-cost alternative in each production stage and are therefore used to a larger extent once the firms have invested in them. More productive firms can spread the investment costs over larger production volumes and therefore tend to use more fuels in production. We allow the fuels' technological states to be industry-specific to account for differences in fuel use across industries.

The model shows that carbon pricing affects the firms' fuel decisions at both the extensive and intensive margins by lowering the production potentials of fossil fuels. A higher carbon price reduces the share of production stages in which fossil fuels are the lowest-cost alternative, leading firms to shift away from fossil fuels at the intensive margin. As the value of fossil fuels in the production process decreases, the profitability of incurring the fixed costs of these fuels also declines. Thus, firms may find it optimal to drop them entirely from their fuel set. Four parameters in particular govern the strength of these effects: the elasticity of substitution between fuels, the price elasticity of demand, the fuel-specific investment costs, and the technological state of each fuel.

We use our firm-level data to estimate the parameters of the model. The estimation proceeds in four steps. In the first step, we estimate each fuel's production potential with log-linear regressions. The results show that the production potentials, and therefore the technological states of each fuel, vary significantly across industries. In the second and third steps, we estimate the elasticity of substitution between fuels and the price elasticity of demand using microfounded equations. These estimates also exhibit considerable heterogeneity across industries. In the last step, we solve the firms' combinatorial discrete choice problem using the approach of Arkolakis, Eckert, and Shi (2023), and estimate the fuel-specific investment costs by the simulated method of moments. Arkolakis et al. (2023)'s approach reduces the computational burden of solving the firms' problem and enables us to analyze the choice between all possible fuel sets. Previous studies on fuel substitution either assume that the firms' fuel sets are fixed or considerably limit the number of fuels to choose from (Hyland and Haller, 2018; Leclair, 2024).

The estimated model performs well at reproducing the pattern of fuel consumption in the data. The baseline equilibrium describes fuel consumption in 2004 and is estimated conditional on the Swedish carbon tax in that year. In the counterfactual analysis, we simulate ten incremental increases in the carbon tax, such that the total increase matches the actual change in the Swedish carbon tax between 2004 and 2020. This corresponds to a rise in the carbon tax by more than 500%. All parameters, including the fuels' technological states and the firms' core productivity levels, are kept fixed at their baseline values. Thus, the counterfactual experiment captures the equilibrium effects of carbon pricing in an environment in which the technology is stable and does not consider the effects on long-run growth and innovation (Acemoglu, Aghion, Bursztyn, and Hemous, 2012; Golosov, Hassler, Krusell, and Tsyvinski, 2014).

The increase in the carbon tax to the 2020 level reduces fossil CO2 emissions in the manufacturing sector by 55%. Of this reduction, 30% is due to a decrease in aggregate energy use, 20% results from energy reallocation toward initially cleaner firms, and 50% is explained by firms switching away from fossil fuels. The fuel switching and reallocation responses cause a surge in the aggregate demand for electricity and biofuels. Aggregate output decreases by 16%, and this effect is reinforced by market share reallocation away from more productive firms. In the baseline equilibrium, more productive firms tend to rely more on fossil fuels in their production and are therefore disproportionately affected by the carbon tax increase. This reallocation effect highlights a trade-off between economic and environmental efficiency in implementing carbon pricing policies.

We also find that the effects on emissions and output vary substantially across industries. The key factors driving this heterogeneity are the elasticity of substitution between fuels and the relative efficiency of renewable energy sources in production. In industries where firms can easily substitute between fuels and efficient alternatives to fossil fuels are available, carbon pricing effectively reduces emissions at a relatively low cost to production. However, in industries where fuel substitution is difficult or renewable energy sources are less efficient, the effectiveness of carbon pricing is significantly lower. These hard-to-abate industries include the cement industry and the basic metals industry.

We check the validity of the counterfactual predictions in two ways. We first compare the predicted average fuel switching response with our reduced-form estimates. The estimated response in the reduced-form analysis is slightly lower than, but not statistically different from, our model prediction. We also compare the predicted patterns of emissions reductions across industries with the actual patterns in our decomposition analysis. In the basic metals industry, the predicted decline in fossil CO2 emissions is primarily driven by a reduction in aggregate energy use. In the paper and pulp industry, the predicted decline is largely driven by fuel switching. These patterns are strikingly similar to what we observe in the data between 2004 and 2020. The fact that the model reproduces these patterns speaks to the empirical relevance of the counterfactual predictions.

The findings in this paper offer important insights for policy. We show that some industries are more adversely affected by carbon pricing due to difficulties in replacing fossil fuels in production. Targeting these industries with complementary policies, such as subsidies to R&D and carbon border adjustment mechanisms, could be important to mitigate output losses and reduce the risk of carbon leakage. Second, we show that carbon pricing leads to a surge in the aggregate demand for electricity and biofuels. Public investments in the infrastructure of renewable energy production might therefore be critical to facilitate the full potential of carbon pricing. Without such investments, capacity constraints in renewable energy production could dampen the fuel-switching response.

Our decomposition analysis contributes to a literature using statistical decompositions to understand changes in aggregate emissions (Grossman and Krueger, 1991; Levinson, 2009; Shapiro and Walker, 2018; Copeland, Shapiro, and Taylor, 2022). These studies find that the key driver behind observed emissions reductions is within-industry changes in emission intensity. We derive a new decomposition that identifies the importance of firms' fuel choices and show that these have been an important factor for the decline in emissions in the Swedish manufacturing sector.

Our reduced-form analysis of the Swedish carbon tax contributes to a growing literature evaluating the effects of environmental policies on firm-level CO2 emissions (Colmer, Martin, Muûls, and Wagner, 2023; Dechezleprêtre, Nachtigall, and Venmans, 2023; Martinsson, Sajtos, Strömberg, and Thomann, 2024), economic performance (Colmer, Martin, Muûls, and Wagner, 2023; Dechezleprêtre, Nachtigall, and Venmans, 2023), abatement investments (Colmer, Martin, Muûls, and Wagner, 2023), and carbon leakage (Dechezleprêtre, Gennaioli, Martin, Muûls, and Stoerk, 2022; Colmer, Martin, Muûls, and Wagner, 2023). Most of these studies focus on the first two trading phases of the EU ETS during which the carbon price was relatively low. Colmer, Martin, Muûls, and Wagner (2023) investigate fuel-mix responses to the EU ETS and find no significant effects. In our setting, the carbon price shocks are substantially larger, and we find that firms respond to these shocks by adjusting their fuel mix.

Our structural analysis contributes to an empirical literature on interfuel substitution. Much of this literature applies cost function estimation techniques developed in Fuss (1977) and Pindyck (1979) to estimate fuel elasticities of substitution using both aggregate data (Serletis, Timilsina, and Vasetsky, 2010a,b; Steinbuks, 2012) and firm-level data

(Woodland, 1993; Bousquet and Ivaldi, 1998; Bousquet and Ladoux, 2006; Bardazzi, Oropallo, and Pazienza, 2015; Hyland and Haller, 2018). More recent papers in this literature estimate elasticities of substitution between clean and dirty fuels based on CES production functions (Papageorgiou, Saam, and Schulte, 2017; Jo, 2023). We contribute to this literature by providing quantitative evidence on the aggregate implications of fuel substitution for carbon pricing. Our model also highlights the various mechanisms through which firms' fuel choices affect aggregate emissions.

Our structural analysis also contributes to a literature using heterogeneous firm models to assess the impacts of environmental policies on emissions. Shapiro and Walker (2018) develop a model to analyze the effects of international trade, productivity growth, and environmental regulation on US emissions. Fowlie, Reguant, and Ryan (2016) develop a model to analyze the effects of carbon pricing in the US cement industry. However, neither Shapiro and Walker (2018) nor Fowlie, Reguant, and Ryan (2016) allow environmental policy to affect emissions via changes in the firms' fuel mix. Instead, firms adjust to regulation by either increasing abatement efforts more generally or by investing in more energy-efficient technology. We develop a model that explicitly captures changes in the firms' fuel mix at both the extensive and the intensive margin. Firms can abate emissions by using cleaner fuels in their production and energy efficiency is an outcome of these fuel choices. Similar adjustments are found in contemporaneous work by Leclair (2024), who analyzes the effects of carbon pricing in the Indian steel industry. In contrast to this paper, we provide evidence of the general equilibrium effects of carbon pricing and the heterogeneous responses across industries for the entire manufacturing sector.

The paper proceeds as follows. Section 2 describes the data and documents stylized facts about fuel consumption in Swedish manufacturing. Section 3 presents the statistical decomposition of aggregate emissions and the reduced form analysis of the Swedish carbon tax. Section 4 introduces the model and Section 5 describes how we estimate the parameters of the model. Section 6 presents the counterfactual analysis. Section 7 concludes.

2 Data

2.1 Data sources and sample

We use microdata on energy use and business statistics provided by Statistics Sweden. Our main data source is Statistics Sweden's annual survey Energy use in manufacturing industry (ISEN) between 2004 and 2020. The data set covers plants with 10 or more employees in the manufacturing sector in Sweden.[1] The data set covers all energy used by these plants and reports energy use by fuel type. There are 19 fuels including electricity and these fuels can be either purchased or self-produced.[2] We compute the plants' CO2 emissions by combining this data with CO2 emission factors from the Swedish

[1]Industries 10-33 in the NACE Rev. 2. classification.

[2]Automatic and manual checks are made by Statistics Sweden to ensure high quality of survey responses. Automatic checks involve comparisons with the previous year's responses and cross-checking with other sources of energy statistics. The response rate is around 80-90% for the years 2004-2020.

Table 1: Ranking of energy sources by firm-year observations, energy and CO2 emissions

Type		Rank by Firm-years	Rank by Energy	Rank by CO2	Percent of total Firm-years	Percent of total Energy	Percent of total CO2
e	Electricity	1	1	18	100	34.85	0
f	Diesel	2	13	11	43.77	0.82	0.91
f	Light fuel oil	3	12	10	23.36	1.06	1.21
f	Gasoline	4	14	14	20.69	0.07	0.07
f	Liquefied petroleum gas	5	7	7	9.23	3.20	3.21
b	Wood fuel	6	3	2	8.05	10.66	17.25
f	Natural gas	7	8	8	5.43	2.75	2.41
f	Heavy fuel oils	8	6	6	4.19	3.66	4.30
b	Black liquor	9	2	1	0.40	27.42	44.38
b	Tall oil	10	9	9	0.27	1.55	1.82
f	Coke	11	4	3	0.24	7.71	12.71
f	Coal	12	5	4	0.18	3.76	5.39
f	Bio gas	13	15	15	0.14	0.05	0.07
b	Coal gas	14	17	16	0.13	0.01	0.01
f	Peat	15	16	13	0.05	0.05	0.08
f	Coke oven gas	16	10	12	0.05	1.24	0.86
f	Kerosene	17	18	17	0.03	0.00	0.00
f	Blast furnace gas	18	11	5	0.03	1.16	5.33
Total					90,666	8.38	543.30

Notes: The sample contains 90,666 firm-year observations over the years 2004-2020. Total energy is in million tera joules (TJ). Total CO2 emissions are in million tonnes. We divide the energy sources into three types: electricity (e), fossil fuels (f), and biofuels (b).

Environmental Protection Agency.[3] We then match our plant-level data with firm-level operating data from Statistics Sweden's Structural Business Statistics (FEK). This data set is based on financial statements collected by the Swedish Tax Authority and provides data on sales, investments, capital, the number of employees, wages, and material inputs.

We also use data on fuel prices and carbon prices in the Swedish manufacturing sector. Fuel prices are calculated using aggregate data from Statistics Sweden on the quantity and purchase value of fuels in the Swedish manufacturing sector. CO2 tax rates and rules for exemptions are provided by Martinsson et al. (2024).[4] Between 2004 and 2015, firms with CO2 tax bills relative to sales above a specific cutoff paid 75% less in tax on emissions above this cutoff. We identify the firms that were granted exemptions by aggregating our plant-level emissions data to the firm level. In 2005, the EU Emissions Trading System (EU ETS) was introduced. Our measure of the carbon price in the EU ETS is the average annual spot price of EU Allowances (EUAs) obtained from the European Environment Agency. We identify the firms that are covered by the EU ETS using Statistics Sweden's Firm register and individual databases (FRIDA).

[3]Calculating CO2 emissions based on fuel use is accurate in the absence of end-of-pipe abatement technology such as Carbon Capture and Storage (CCS).

[4]Martinsson et al. (2024) provide CO2 tax rates per kg CO2 between 1990 and 2017. For the years 2018-2020, we use CO2 tax rates for liquefied petroleum gas (LPG) per tonne LPG from the Swedish Tax Authority and convert them into tax rates per kg CO2. The CO2 tax rate per kg CO2 is uniform across fuels so these calculations yield tax rates for all taxed fuels.

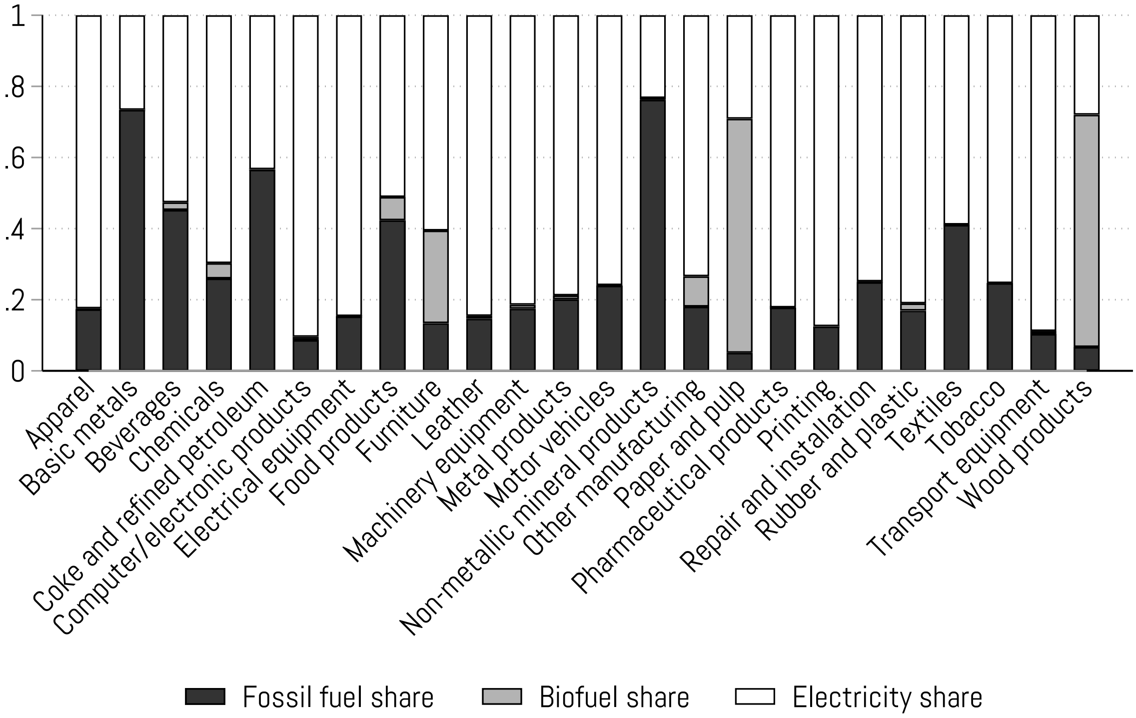

Figure 1: Fuel shares by industry

Notes: The sample contains 90,666 firm-year observations over the years 2004-2020. The fuel shares are calculated as the total use of fossil fuels, biofuels, and electricity over the total use of energy by 2-digit NACE industries 10-33.

We also use data on innovations related to CO2 mitigation in the manufacturing sector to construct a proxy for green technology trends. We collect this data from the US Patent Office (USPTO) which documents all patents secured in the US. The data set contains information on patent names, patent codes, and grant dates for patents granted by the USPTO since 1976, as well as citations made to each patent in other patent applications and by citation examiners and third parties. We focus on patents that fall under at least one of the following subcategories: Climate change mitigation technologies in the production process for final industrial or consumer products (Y02P70), Climate change mitigation technologies for sector-wide applications (Y02P80), and Enabling technologies with a potential contribution to greenhouse gas emissions mitigation (Y02P90).[5]

Our base sample contains 90,666 firm-year observations including 12,484 unique firms and 13,118 unique plants over the years 2004-2020. The sample is confined to firms in the manufacturing sector for which we have data on fuel use. We drop energy from waste which accounts for 0.07% of total energy use since we do not have data on prices for this fuel. We also limit the sample to firms that report positive sales and positive electricity use.[6] Table 1 provides information on the use of each fuel in our base sample. The

[5]In addition to these general categories, we also use patents related to CO2 mitigation within the following industry-specific codes Y02P10, Y02P20, Y02P30, Y02P40 and Y02P70/62 for firms active in the basic metals, chemical, coke and refined petroleum, non-metallic mineral products, and textile industries, respectively.

[6]Missing data on electricity use may arise if electricity is included in the firms' rent. The share of observations with missing data on electricity is 2.5%.

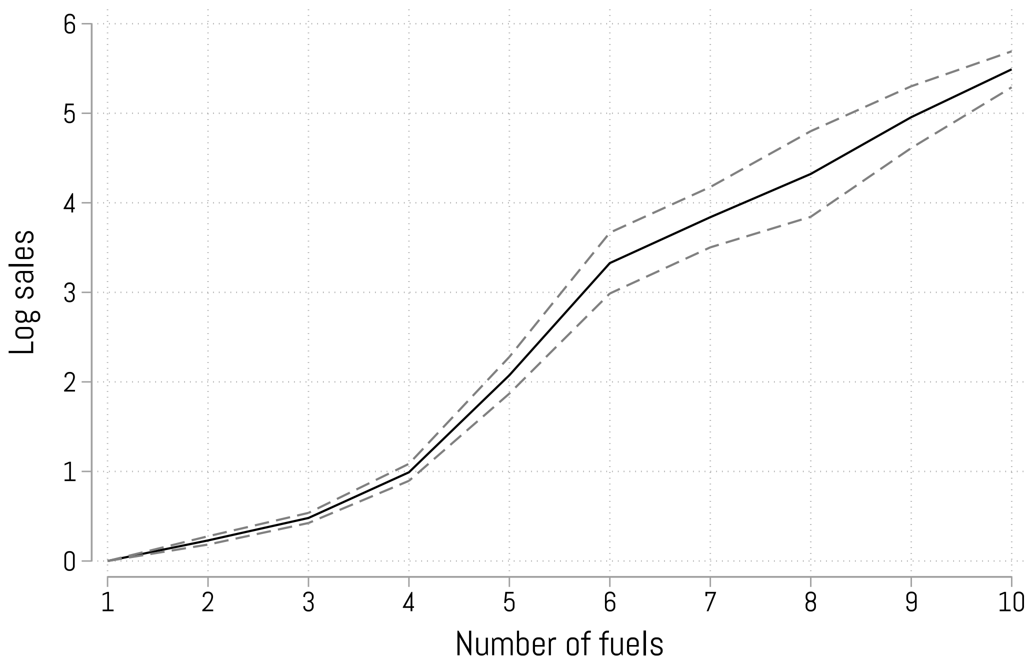

Figure 2: Firm size and number of fuels

Notes: The sample contains 90,666 firm-year observations over the years 2004-2020. The figure plots estimates and 95% confidence intervals from a regression of log sales on dummies indicating the number of fuels used by firms. The omitted category is firms that only use electricity. The regression includes 2-digit industry fixed effects and year fixed effects. Sales are deflated by the producer price index in Swedish manufacturing. Standard errors are clustered by firms.

table ranks the fuels based on the number of firms that use them, the amount of energy they provide, and the amount of CO2 emissions they emit. The use of electricity does not generate CO2 emissions although production of it may. If a firm uses other fuels to generate electricity, then the associated CO2 emissions are attributed to those other fuels and not electricity. Fuels that are used by a large number of firms are not necessarily the fuels that provide the most energy, suggesting that large and energy-intensive firms use different fuels than the majority of firms.

2.2 Three stylized facts about fuel consumption

We document three stylized facts to better understand what characterizes fuel consumption in the Swedish manufacturing sector.

Fact 1. *Patterns of fuel consumption differ substantially across manufacturing industries.*

Figure 1 plots the energy shares of fossil fuels, biofuels, and electricity in each 2-digit industry in our sample of manufacturing firms. While most industries primarily consume energy from electricity, some industries rely more on fossil fuels or biofuels in their production. This fact partly explains why some fuels in Table 1 rank high in terms of energy but low in terms of the number of firms that use them. For example, the second highest-ranked fuel in terms of energy, black liquor, accounts for a large share of energy in the most energy-intensive industry, paper and pulp, but is not used much in other

Table 2: Summary statistics on firms' fuel-mix

	Mean	Min.	25th percentile	Median	75th percentile	Max.	Observations
Number of fuels	2.16	1	1	2	3	12	90,666
Top 1 fuel share	0.74	0.25	0.59	0.74	0.90	0.99	60,354
Top 1 and 2 fuel shares	0.92	0.48	0.88	0.95	0.98	0.99	30,756

Notes: The sample contains 90,666 firm-year observations over the years 2004-2020. Summary statistics over top 1 and top 2 fuel shares are calculated for firm-years with more than 1 and 2 fuels, respectively.

industries. A similar argument goes for coke and coal and the basic metals industry. Fact 1 suggests that the suitability of using different fuels varies across industries.

Fact 2. *Larger firms use more fuels and are more likely to add new fuels to their fuel-mix.*

Figure 2 shows that the size of firms tends to increase in the number of fuels. The figure plots the point estimates and the confidence intervals from a regression of log sales on dummy variables indicating the firms' number of fuels. The regression includes 2-digit industry fixed effects to capture within-industry variation in the number of fuels. Firms that use two fuels are on average about 23% larger than firms that only use electricity, firms that use 5 fuels are on average 2 log points larger, and firms that use 10 fuels are on average more than 5 log points larger. Conditional on the number of fuels in the firms' fuel set, Table A1 in Appendix A shows that larger firms are also more likely to add fuels to their fuel set and less likely to drop fuels from their fuel set compared to smaller firms.

Fact 2 is indicative of fixed costs being associated with the use of fuels in production. Manufacturing firms use fuels in capital equipment such as furnaces, boilers, and electrical heaters. Installation of such capital may involve substantial investment costs that prevent firms from using the most efficient fuels for each production process. Larger firms can spread the fixed costs over larger production volumes and are therefore more likely to invest in fuel-specific capital equipment as long as this equipment has the potential to cut marginal cost.

Fact 3. *Firms using multiple fuels concentrate energy consumption in just a few of them.*

Table 2 shows that the average firm in our sample uses about two fuels. About one-third relies solely on electricity, while one-third uses two fuels and one-third uses three or more fuels. The average firm using at least two fuels concentrates 74% of its energy consumption in the top 1 fuel while the average firm using at least three fuels concentrates 92% of its energy consumption in the top 2 fuels. Hence, some fuels are more important in the firms' production process than others, suggesting that fuels differ in their applicability and efficiency across different tasks.

3 Empirical analysis

In this section, we analyze how observed changes in manufacturing CO2 emissions depend on firms' fuel decisions and how these fuel decisions in turn have responded to changes in the Swedish carbon tax.

3.1 A statistical decomposition

Aggregate manufacturing CO2 emissions can be written as,

$$E = Z \sum_i z_i e_i, \tag{1}$$

where Z is aggregate energy use, z_i is plant i's share of the aggregate energy use, and e_i is plant i's CO2 emissions per energy unit. If we take logs of (1) and differentiate with respect to (Z, z_i, e_i), we can decompose changes in manufacturing emissions into three components:

$$\frac{dE}{E} = \underbrace{\frac{dZ}{Z}}_{\text{Energy scale}} + \underbrace{\sum_i \frac{E_i}{E}\frac{dz_i}{z_i}}_{\text{Energy composition}} + \underbrace{\sum_i \frac{E_i}{E}\frac{de_i}{e_i}}_{\text{Fuel switching}}. \tag{2}$$

The first component is a scale effect that goes via changes in aggregate energy use, the second component is a composition effect that goes via changes in the composition of energy use across plants, and the third component is a technique effect that goes via changes in the fuel mix of individual plants.[7]

We refer to the technique effect as the fuel-switching component. Changes in e_i represent both extensive and intensive margin changes in the fuel mix of firms. A firm can reduce its CO2 emissions per energy unit by replacing a dirty fuel with a clean fuel or by shifting energy consumption towards relatively clean fuels. Since end-of-pipe emission control technologies are not yet viable (e.g. carbon capture and storage), changes in e_i can only come from changes in the plant's fuel mix. The effect of improvements in overall energy efficiency is captured by the scale component and not the fuel-switching component. The benefit of our decomposition relative to the previous literature (Grossman and Krueger, 1991; Levinson, 2009; Shapiro and Walker, 2018) is that we can quantify the importance of changes in the firms' fuel mix for aggregate emissions.

We take the decomposition to the data by following the approach in Shapiro and Walker (2018). In particular, for each year 2004-2020, we compute aggregate energy use, Z, the plants' energy shares, z_i, and the plants' fossil CO2 per energy unit, e_i. We then project how manufacturing emissions would have evolved if the z_i's and the e_i's were kept constant at their 2004 levels and compare these counterfactual emission paths with the actual evolution of emissions. Comparing the three scenarios allows us to quantify the relative importance of the different components. This approach requires a balanced panel of plants and can therefore not capture the effect of plant-level entry and exit.[8] However, we can assess the importance of entry and exit by comparing the evolution of aggregate emissions in the balanced panel of plants with that of the unbalanced panel.

The result is presented in Figure 3. Panel A shows the decomposition of fossil fuel emissions in the entire manufacturing sector. Emissions in each year are expressed relative

[7] See Appendix B for a derivation of equation (2).

[8] The balanced panel of plants covers 94% of total CO2 emissions and 92% of total fossil CO2 emissions between 2004 and 2020 in our full sample of plants.

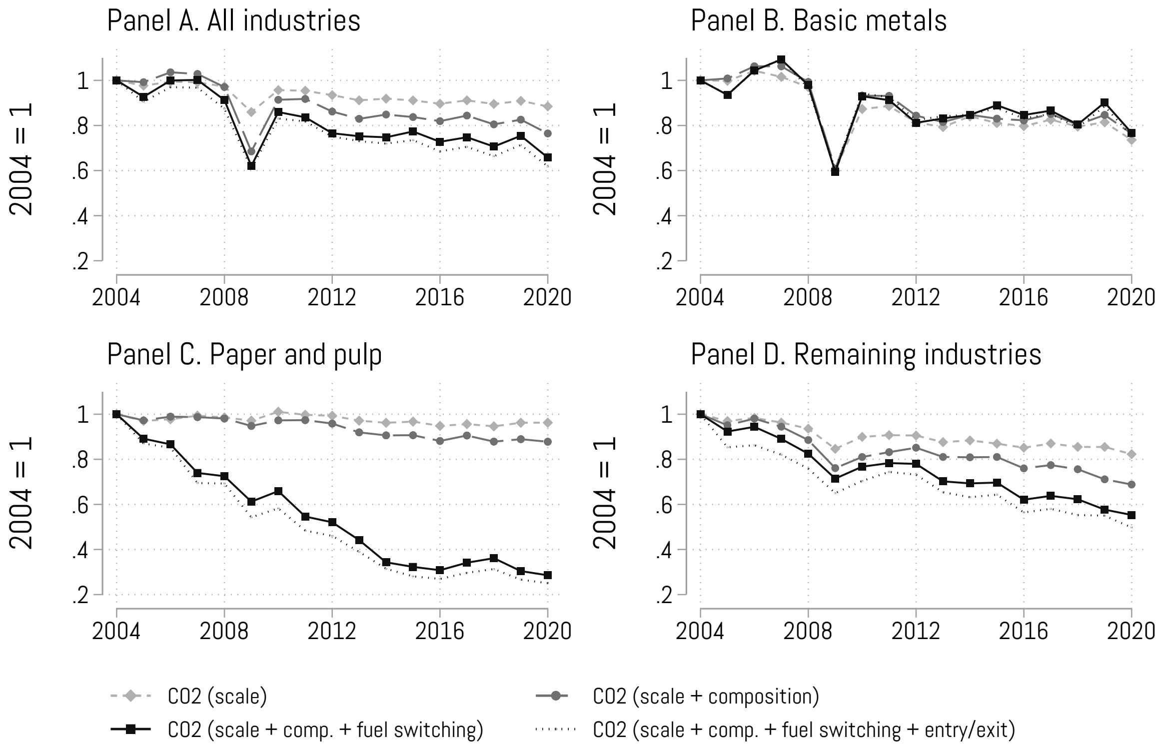

Figure 3: Decomposition of aggregate fossil CO2 emissions

Notes: The decomposition is based on plant-year observations in our base sample of 90,666 firm-year observations over the years 2004-2020. The (scale), (scale + composition), and (scale + comp. + fuel switching) lines are based on balanced samples while the (scale + comp. + fuel switching + entry/exit) lines are based on unbalanced samples. The balanced and unbalanced samples in Panel A contain 39,236 and 103,543 plant-year observations, respectively. The balanced and unbalanced samples in Panel B contain 1,666 and 2,977 plant-year observations, respectively. The balanced and unbalanced samples in Panel C contain 1,887 and 3,075 plant-year observations, respectively. The balanced and unbalanced samples in Panel D contain 35,683 and 97,491 plant-year observations, respectively. The balanced sample of plants in Panel A covers 92% of total fossil CO2 emissions between 2004 and 2020 in the full sample of plants.

to emissions in 2004.[9] The light gray line shows the evolution of emissions if we fix the energy composition and the plants' fuel mix at 2004 levels. This line illustrates the importance of changes in aggregate energy use. The dark gray line represents the evolution of emissions if we only fix the plants' fuel mix at 2004 levels. The difference between this line and the light gray line represents the importance of changes in the energy composition. The black line plots the actual evolution of emissions in the balanced panel of plants. The difference between this line and the dark gray line represents the importance of plant-level fuel switching. Finally, the dotted line shows the evolution of emissions in the unbalanced panel, so the difference between this line and the black line represents the importance of entry and exit in explaining changes in emissions.

Panel A shows that annual manufacturing emissions from fossil fuels have declined by about 38% between 2004 and 2020. This represents a reduction of 5.7 million tonnes of CO2 in the year 2020 relative to 2004. The scale, composition, and fuel-switching components each account for about 30% of this decline. The remaining 10% are accounted

[9]Figure B1 presents the same graphs as in Figure 3 but in level changes instead of relative changes.

for by plant entry and exit. Hence, the observed decline is explained by a combination of a reduction in aggregate energy use, a reallocation of energy towards plants that use cleaner fuels, and changes in the plants' fuel mix away from dirty fossil fuels. The entry and exit effect shows that entrants tend to use cleaner fuels than exiting plants.

Figure 3 also shows the decomposition for the basic metals industry, the paper and pulp industry, and the remaining industries, separately. We single out the basic metals industry and the paper and pulp industry since these are by far the two most energy-intensive industries in Swedish manufacturing.[10] Panel B shows that fossil CO2 emissions in the basic metals industry have declined by 25% between 2004 and 2020 and that this decline is almost exclusively explained by a reduction in aggregate energy consumption. The other components have not played significant roles in this industry. Panel C shows the decomposition for the paper and pulp industry. This industry has experienced a 75% decline in fossil CO2 emissions over the sample period and 80% of this decline is due to within-firm fuel switching. Changes in aggregate energy use and the composition of energy use across plants account for 5% and 11%, respectively. Lastly, Panel D shows that fossil CO2 emissions have declined by 50% in the remaining industries and that the scale, composition, and fuel-switching components explain about 30% each.

The cross-industry heterogeneity in what drives changes in aggregate emissions may reflect differences in the elasticity of substitution between fuels, productivity growth, and industry-specific market characteristics. It may also reflect differences in the availability and suitability of non-fossil fuels in production. An important limitation of the statistical decomposition is that it is uninformative about what causes the different components to change and how these components may interact with each other. In the next section, we assess the role of the Swedish carbon tax as a potential cause of firm-level fuel switching. We then build and estimate a model that can account for each of the four components affecting aggregate emissions.

3.2 The Swedish carbon tax and fuel switching

3.2.1 Institutional background

The CO2 tax is Sweden's main policy tool to reduce CO2 emissions and is levied on fossil fuels used for fuel combustion in vehicles and in industrial processes.[11] The tax rate per tonne CO2 is the same across fuels but the tax rate per gigajoule differs depending on the fuels' carbon content. In 2005, the EU Emissions Trading System (EU ETS) was introduced and firms with installations covered by this cap and trade scheme initially had to pay both the national carbon tax and surrender carbon credits (EUAs) for each tonne CO2 emitted. The EUAs are allocated to firms either for free or via auctioning and are then traded among firms at market price. From 2011 and onwards, installations in the EU ETS were completely exempted from the national CO2 tax. Firms outside the EU ETS could also be granted exemptions from the national CO2 tax. In particular, up until 2015, the most carbon-intensive firms faced a 75% reduction in their marginal tax rate.[12]

[10] The basic metals industry accounts for 20% and the paper and pulp industry for 54% of total energy use in our sample.

[11] The fossil fuels peat, coke oven gas, and blast furnace gas are exempted from the Swedish carbon tax.

[12] Exemptions are based on how much a firm pays in carbon tax relative to its sales. Between 2004-2010, firms were granted exemptions if the CO2 tax bill exceeded 0.8% of their sales. Between 2011-2014, this

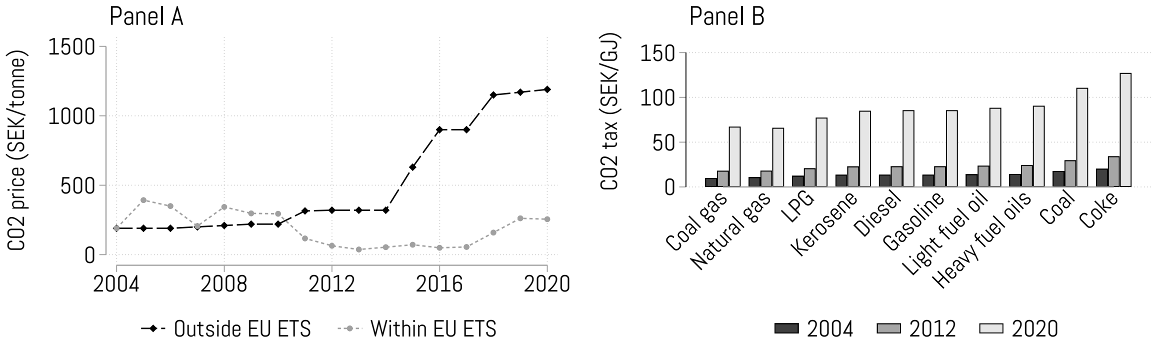

Figure 4: Carbon price in the Swedish manufacturing sector

Notes: Panel A shows the national CO2 tax for manufacturing firms outside the EU ETS and the price of carbon for firms within the EU ETS between 2004 and 2020. The carbon price for firms within EU ETS is calculated as the sum of the spot prices of EUAs and the national carbon price they face. Panel B shows the national CO2 tax per gigajoule for fossil fuels subject to the tax in 2004, 2012, and 2020.

The firms that have been granted exemptions from the national carbon tax only comprise 3% of the full sample of firms. Nonetheless, the EU ETS firms in our sample account for 61-86% of total CO2 emissions each year.

Panel A of Figure 4 shows the carbon tax per tonne CO2 in each year between 2004 and 2020. It also shows the carbon price for firms covered by the EU ETS, which is measured as the average spot price of EUAs plus the national carbon tax they face. The CO2 tax rate was relatively stable up until 2011 when it jumped from 200 SEK to 315 SEK. It was then flat for a few years until 2015 when it started to increase sharply. In total, the CO2 tax rate increased by more than 500% over the sample period. Panel B shows how the CO2 tax rate per gigajoule has changed for the different fossil fuels subject to the tax. The proportional change is the same across the fuels but the increase in levels is larger for fuels with higher carbon contents.

Two concerns about estimating the effect of changes in the carbon tax are anticipation effects and other coinciding policy changes. Changes in the tax rate are disclosed by the Swedish government during the budget process each year. Thus, firms are generally informed about a new tax rate only a few months in advance (Martinsson et al., 2024). However, in 2009, the Swedish Parliament adopted a reform package (Government Bill 2009/10:41) which included the tax increases in 2011 and 2015 as well as the 2011 tax exemption for EU ETS firms. This could have triggered anticipation effects which would attenuate the contemporaneous effect of changes in the carbon tax. However, we check for this and do not find any evidence of anticipation effects. Moreover, based on policy documents, the only coinciding policy change of relevance is an increase in the energy tax in 2011. At this point, the energy tax on fossil fuels went from zero to the EU's minimum level for all manufacturing firms in our sample (Hammar and Åkerfeldt, 2011). We address this by controlling for fuel prices that include energy taxes.

threshold was increased to 1.2%.

3.2.2 Empirical strategy

We use two complementary approaches to identify the effect of the Swedish CO2 tax on firm-level fuel switching. The first is a shift-share instrumental variable (IV) approach where we exploit that there is large variation in the exposure to changes in the carbon tax across firms in our sample. A firm that uses a large share of fossil fuels is more exposed to changes in the tax than a firm that uses a large share of biofuels or electricity.

We measure firm-level tax exposure as the predicted average carbon tax per gigajoule,

$$\bar{\tau}_{it}^{pre} = \sum_{j} s_{ji}^{pre} \tau_{jt}, \tag{3}$$

where s_{ji}^{pre} is the share of fuel j in firm i's fuel mix in a pre-sample year, and τ_{jt} is the tax per gigajoule of fuel j in year t.[13] We then use this as an instrument to estimate the fuel switching response to changes in the firms' average carbon tax, $\bar{\tau}_{it} = \sum_j s_{jit}\tau_{jt}$. In particular, we estimate regressions of the following form,

$$y_{it} = \beta\bar{\tau}_{it} + \gamma_i + \gamma_{tk} + \delta' X_{it} + \varepsilon_{it}. \tag{4}$$

where y_{it} is an outcome variable related to the firms' fuel mix. The main outcome variable of interest is the firms' fossil fuel emissions per energy unit, which corresponds to the fuel-switching component in the statistical decomposition. However, we also estimate the effects on the firms' fossil fuel shares, electricity shares, and biofuel shares, as well as the number of fossil fuels used by the firms. We include firm fixed effects to measure the effect of within-firm changes in the average carbon tax. Without firm fixed effects, the estimate would be driven by variation in the level of tax exposure across firms, which is positively correlated with the outcome variables. We also include industry-year fixed effects to control for industry-wide trends (e.g. industry-specific technological changes that affect firms' fuel choices in the same way). We instrument for $\bar{\tau}_{it}$ with $\bar{\tau}_{it}^{pre}$ since the contemporaneous fuel shares are endogenous and correlated with the outcome variables by construction.

The key identifying assumption is that firms in the same 2-digit industry with different exposure to the carbon tax according to (3) would exhibit parallel trends in the outcome variables in the absence of any changes in the tax. This is a strong assumption since firms with different pre-sample fuel shares may be differentially exposed to other time-varying variables affecting the firms' fuel decisions. In particular, there are two main threats to this assumption. The first is changes in fuel prices. We therefore include the following two variables, $\bar{p}_{it}^{pre} = \sum_j s_{ji}^{pre} p_{jt}$ and $\bar{q}_{it}^{pre} = \sum_h s_{hi}^{pre} q_{ht}$, where p_{jt} denotes the price of fossil fuel j (net of the carbon tax) and q_{ht} denotes the price of non-fossil fuel h, and include them in the set of control variables X_{it}. The second threat is fuel-biased technological change. This threat is harder to address since we do not directly observe technological change. However, we can construct a proxy by using our data on patents related to innovations in climate change mitigation technologies in the production and processing of industrial goods. The proxy we use is a weighted sum of patents in each year of our sample period, where the weights are the patents' average number of citations per year. We then interact this proxy with the firms' pre-sample fossil fuel shares to capture that firms that use

[13] We define the pre-sample year for each firm to be the first year we observe the firm in our base sample and then drop this observation from the estimation sample.

larger shares of fossil fuels may be differentially exposed to technological change related to CO2 mitigation.

As a complement to the IV strategy, we apply a matched difference-in-differences approach where we exploit the sharp increase in the CO2 tax in 2015 and the fact that EU ETS firms were exempted from the tax as of 2011. We use the following dynamic specification to estimate the effect of the carbon tax,

$$y_{it} = \sum_{\substack{s=2011, \\ s \neq 2014}}^{2020} \beta_s D_{it}^s + \gamma_i + \gamma_{tk} + \delta' X_{it} + \varepsilon_{it}, \tag{5}$$

where D_{it}^s is an indicator equal to one if firm i is subject to the national carbon tax and $s = t$. The parameters of interest, β_s, capture the average fuel switching response in year s among firms subject to the tax relative to firms covered by the EU ETS. We confine the sample period to 2011-2020 and perform a one-to-one nearest-neighbor matching procedure where we match EU ETS firms with non-EU ETS firms. We match on the firms' carbon emission intensities in 2011 and 2-digit industry. This results in a perfect match on industry and a good match on carbon intensity when we allow the max difference to be 0.01. The purpose of the matching procedure is to ensure that we compare firms that are likely to have parallel trends in the outcome variables in the absence of any changes in the carbon price. Since the carbon price was flat for both groups several years before 2015, we can evaluate the plausibility of this assumption by checking for pre-trends. A downside with the difference-in-differences approach is the small number of EU ETS firms in our sample.[14] After the matching procedure, we end up with a balanced panel of 49 firms in the treatment group and 49 firms in the control group over the years 2011-2020.

Figure B2 in Appendix B shows the balance between the treatment and control groups before and after the matching procedure with respect to both targeted and untargeted variables. Although the matched sample is much more balanced overall, there are still differences in the firms' fuel mix across the two groups. To control for differential exposure to fuel prices and technical change, we therefore include the same set of control variables in the difference-in-differences specification as in the shift-share specification.

3.2.3 Results

The IV estimates of equation (4) are presented in Table 3. The estimation sample is confined to firms with a non-zero share of taxed fossil fuels in the pre-sample year. We also drop EU ETS firms and firms that are subject to tax exemptions to avoid variation in the tax rate that is due to sorting.[15] Table B2 in Appendix B shows that the estimates are not statistically different when we include the full set of firms. Column (1) of Table 3 shows the estimated effect of the average tax rate on the firms' fossil CO2 emissions per gigajoule. The effect is negative and highly significant. For each SEK increase in the average carbon tax, firms emit 0.82 kg CO2 less per gigajoule on average. Is this effect large or small? In a simple back-of-the-envelope calculation, the estimate predicts that 67% of the average decline in fossil CO2 per gigajoule is caused by the increase in

[14] There are around 100 firms covered by the EU ETS in each year of our base sample.

[15] Economic theory predicts a gap in the density around the CO2 tax-to-sales threshold (Saez, 2010). When we explore this, such a gap is not visible, but we exclude observations above the threshold to avoid potential endogenous variation in the tax rate.

Table 3: The shift-share IV analysis

	Fossil CO2 / Energy (1)	Fossil fuel share (2)	Electricity share (3)	Biofuel share (4)	# fossil fuels (5)
CO2 tax	-0.832***	-0.011***	0.006***	0.005***	-0.009**
	(0.117)	(0.002)	(0.001)	(0.001)	(0.004)
Fossil fuel price	-0.013	-0.000	0.000	0.000*	-0.000
	(0.009)	(0.000)	(0.000)	(0.000)	(0.000)
Non-fossil fuel price	0.077***	0.001***	-0.002***	0.001***	0.001
	(0.015)	(0.000)	(0.000)	(0.000)	(0.001)
CO2 mitigation patents	0.004***	0.000***	-0.000***	-0.000	0.000*
	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)
Observations	48,715	48,715	48,715	48,715	48,715
First-stage F-statistic	197	197	197	197	197
Firm FE	Yes	Yes	Yes	Yes	Yes
Industry-year FE	Yes	Yes	Yes	Yes	Yes

Notes: The base sample is an unbalanced panel of 90,666 firm-year observations over the years 2004-2020. The estimation sample of the shift-share IV analysis is confined to firms that have not been part of the EU ETS nor faced any tax exemptions over the sample period. It is also confined to firms that use non-zero taxed fuels in the pre-sample year, implying that the instrument is non-zero for all observations in the estimation sample. Finally, the estimation sample drops the first (pre-sample) year for all observations. The inclusion of firm fixed effects also removes singletons from the estimation sample. Standard errors are clustered at the firm level. The first stage F-statistic is the Kleibergen-Paap F-statistic.

the carbon tax over the sample period.[16] In columns (2), (3), and (4), we also present the estimated effects on the firms' fossil fuel shares, electricity shares, and biofuel shares, respectively. As expected, the fossil fuel share decreases with the tax while the electricity share and the biofuel share increase with the tax. The last column shows that there is also a significant effect on the extensive margin of the firms' fuel mix. Firms respond to a higher carbon tax by reducing the number of fossil fuels they use in production.

Figure 5 illustrates the first-stage and reduced-form relationships behind the IV estimate in column (1) of Table 3. All variables are residualized by the fixed effects and control variables. The binned scatter plots show that the relationships are well approximated by our linear model and that our estimates are not driven by outliers in the data. Table B1 in Appendix B shows the point estimates for the first stage and reduced form regressions with respect to all outcomes.

The matched difference-in-differences estimates are presented in Figure 6. Panel A plots the average change in the predicted CO2 price per gigajoule of the non-EU ETS firms relative to the EU ETS firms. The predicted CO2 price is based on the firms' fuel shares in 2010 and we see that this variable began to increase sharply in 2015 among firms subject to the national carbon tax. Panels B-D plot the estimated effects of this increase on the firms' fossil CO2 emissions per gigajoule, fossil fuel shares, and electricity shares, respectively. There is a clear response in each of the three outcome variables and no signs of anticipation effects prior to 2015. The results suggest that firms subject to the Swedish carbon tax responded to the tax increase by switching away from fossil fuels

[16] The average decline in the firms' fossil CO2 per gigajoule is 10.04 between 2005 and 2020 while the average increase in their CO2 tax per gigajoule is 8.14. The estimate of -0.832 suggests that the change in the carbon tax has reduced the firms' fossil CO2 per gigajoule by 6.77 on average.

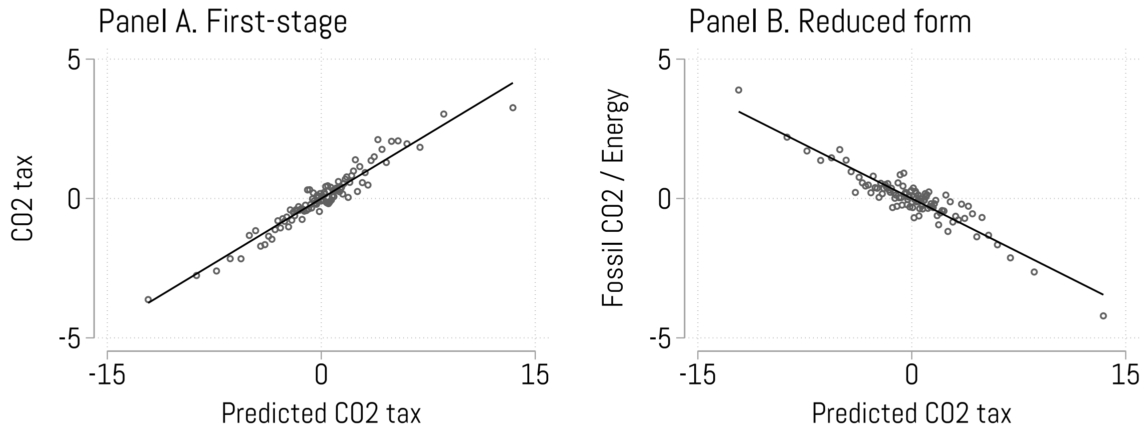

Figure 5: Illustration of the first-stage and reduced form regressions

Notes: The base sample is an unbalanced panel of 90,666 firm-year observations over the years 2004-2020. The estimation sample of the shift-share IV analysis comprises 48,715 observations and is confined to firms that have not been part of the EU ETS nor faced any tax exemptions over the sample period. It is also confined to firms that use non-zero taxed fuels in the pre-sample year, implying that the instrument is non-zero for all observations in the estimation sample. Finally, the estimation sample drops the first (pre-sample) year for all observations. The inclusion of firm fixed effects also removes singletons from the estimation sample. The binned scatter plots are constructed by first residualizing the variables by the fixed effects and control variables, then grouping the x-axis variables into equal-sized bins, and finally computing the mean of the x-axis and y-axis variables within each bin. The linear lines are constructed using OLS.

towards electricity. In Figure B3 in Appendix B, we show that there is no significant effect on the firms' biofuel share or number of fossil fuels in the matched difference-in-differences sample.

3.2.4 Robustness checks

In Appendix B, we perform several robustness checks to challenge our results. Table B3 shows a placebo test where we estimate the reduced form effects of the predicted CO2 tax on the sample of EU ETS firms between 2011 and 2020. This sample was exempted from the national carbon tax and should therefore not respond to changes in the instrument. The point estimates are statistically insignificant and close to zero. Figure B4 shows that our matched difference-in-differences results are insensitive to the exclusion of control variables and Figure B5 shows that the point estimates are not driven by any single pair of firms in the matched sample. Finally, in Figure B6, we perform a placebo test where we randomize treatment status 200 times among the firms in our matched sample such that 50% are always treated. The placebo test evaluates the likelihood of the observed effects under the null hypothesis that treatment has no effect. The placebo estimates are concentrated around zero and the actual treatment effects are statistically different from the placebo estimates at the 1% level as of 2017. Hence, we reject the null of no treatment effect.

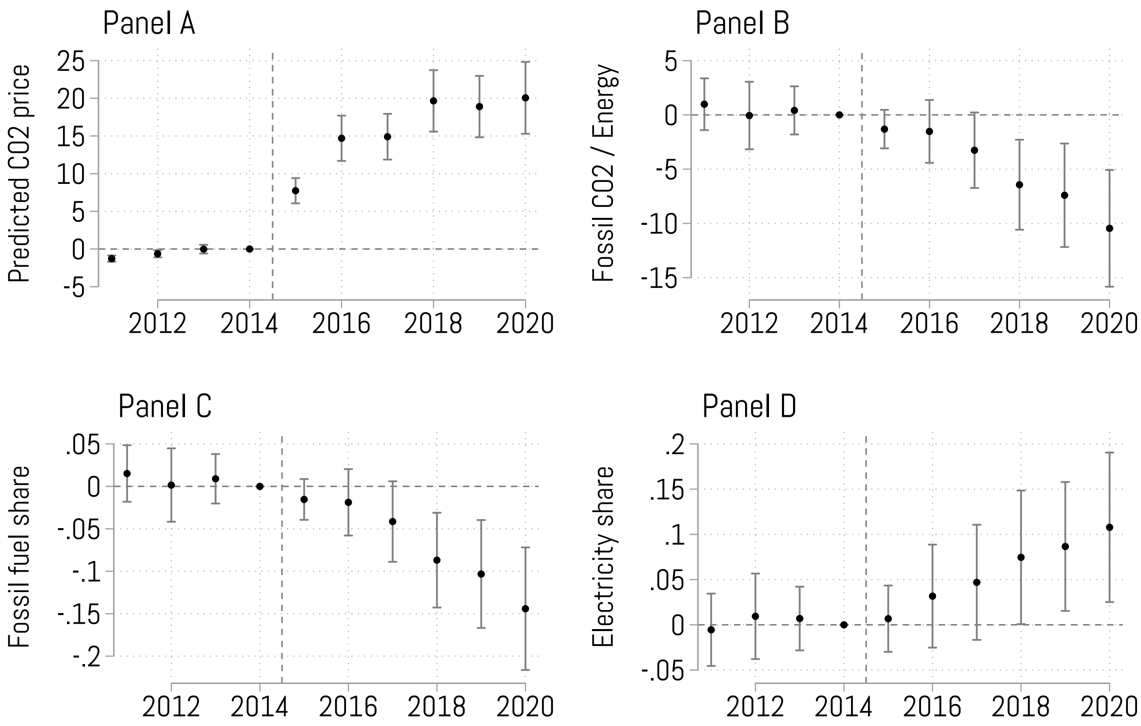

Figure 6: The matched difference-in-differences analysis

Notes: The matched difference-in-differences estimation sample contains 980 firm-year observations over the years 2011-2020. The figure shows the point estimates and 95% confidence intervals when we run the regression in (5). Standard errors are clustered at the firm level and do not account for potential correlation in the residuals within the treatment and control groups. However, we compute the intraclass correlation and find that it is practically zero.

4 Model

We draw on the quantitative trade literature, and in particular Antras et al. (2017), to develop a model of manufacturing firms' fuel decisions that can account for the empirical facts documented in the data: (i) heterogeneity in fuel consumption across industries, (ii) heterogeneity in fuel consumption across the firm size distribution, and (iii) heterogeneity in fuel consumption across fuels within firms.

The model is quantifiable and can be used to perform counterfactual experiments. We are interested in the general equilibrium effects of changes in the carbon price on firms' fuel decisions and how these matter for the aggregate effects on CO2 emissions and output. The model allows carbon prices to affect aggregate emissions via all four channels highlighted in the statistical decomposition: (i) aggregate energy use, (ii) reallocation of energy use across firms, (iii) firm-level fuel switching, and (iv) entry and exit.[17]

[17]Appendix C provides derivations of key equations in this section.

4.1 Setup

Consider a closed economy with one manufacturing sector and one energy sector. The energy sector uses labor to produce different fuels, collected in the set J, and is characterized by perfect competition and constant returns to scale. The unit labor requirement differs across fuels and the fuels are sold at marginal cost, $r_j = d_j w$, where d_j is the unit labor requirement for fuel $j \in J$ and w is the wage rate.

The manufacturing sector consists of K manufacturing industries and each of these industries is characterized by monopolistic competition. Following Melitz (2003), there is a competitive pool of potential entrants into each industry and firms that enter must incur a sunk entry cost equal to f_{ek} units of labor. Once this entry cost is paid, the firms learn their productivity and decide whether or not to stay in the industry and produce. The manufacturing firms that stay use both labor and fuels in the production of differentiated final goods. Some fuels emit CO2 emissions and are taxed by the government and the tax income is transferred to consumers in the form of a lump sum payment. The tax rate per energy unit, τ_j, differs across fuels depending on their carbon content such that the tax rate per CO2 is uniform across fuels.

The labor market is assumed to be perfectly competitive and there is a mass $\bar{L}$ of consumers that inelastically supply one unit of labor to firms. By treating labor as the numéraire, the fuel prices are pinned down by the fuels' labor unit requirements. The consumers spend their income on manufactured goods.

4.2 Preferences

The preferences of a representative consumer over goods produced in the manufacturing sector are described by the utility function:

$$U = \prod_{k=1}^{K} Q_k^{\beta_k} - g(\mathrm{CO2}), \tag{6}$$

where Q_k denotes consumption of varieties produced in industry k and $g(\mathrm{CO2})$ represents the disutility from aggregate CO2 emissions in the economy. The Cobb-Douglas structure of the utility function entails that consumers spend a constant share, β_k, of their income on varieties from industry k. Preferences over these varieties are assumed to take the constant elasticity of substitution (CES) Dixit and Stiglitz (1977) form:

$$Q_k = \left(\int_{\omega \in \Omega_k} q_k(\omega)^{\frac{\sigma_k - 1}{\sigma_k}} d\omega \right)^{\frac{\sigma_k}{\sigma_k - 1}}, \quad \sigma_k > 1, \tag{7}$$

where Ω_k is the set of varieties in industry k available to consumers and σ_k is the elasticity of substitution across these varieties.

The market demand for variety ω in industry k is obtained by solving the consumer's utility maximization problem and is given by,[18]

$$q_k(\omega) = E_k P_k^{\sigma_k - 1} p_k(\omega)^{-\sigma_k}, \tag{8}$$

[18]To obtain market demand, we replace the consumer's budget constraint with the aggregate budget constraint across consumers.

where E_k is aggregate spending on goods in industry k and P_k is the ideal price index in industry k,

$$P_k = \left(\int_{\omega \in \Omega_k} p_k(\omega)^{1-\sigma_k} d\omega \right)^{\frac{1}{1-\sigma_k}}. \tag{9}$$

The market demand for variety ω depends on an industry demand index $A_k = E_k P_k^{\sigma_k - 1}$ and the price of variety ω. Since we have a continuum of firms in each industry, the firm producing ω has zero mass relative to the industry as a whole and therefore takes A_k as given (Melitz and Redding, 2014). The demand elasticity σ_k governs how sensitive consumers are to price changes.

4.3 Technology

Each manufacturing firm in industry k produces a single differentiated variety ω by assembling intermediate inputs v according to the following production function,

$$q_k(\varphi) = \varphi \left[\int_0^1 e(v)^{\frac{\rho_k - 1}{\rho_k}} dv \right]^{\frac{\rho_k}{\rho_k - 1}}, \tag{10}$$

where φ denotes the firm's core productivity, which is unique for each firm, and thus can be used to index firms instead of ω; $e(v)$ denotes the quantity used of input v; and ρ_k denotes a constant elasticity of substitution between inputs.

A firm φ produces each intermediate input $v \in [0, 1]$ in-house by using a single fuel $j \in J$ according to the following constant returns to scale technology,

$$e(v) = \frac{x_j(v)}{a_j(v, \varphi)}, \tag{11}$$

where $x_j(v)$ denotes quantity of fuel j and $a_j(v, \varphi)$ denotes the unit energy requirement of fuel j associated with firm φ and intermediate input v. We interpret $a_j(v, \varphi)$ as a measure of how suitable fuel j is for firm φ in the production of input v.

To use fuels in production, firms must invest in fuel-specific technology, which we model as fuel-specific fixed costs, f_{jk}, measured in terms of labor units. That is, a firm φ needs to hire f_{jk} units of labor to install the technology associated with fuel j in industry k in order to use it for production. Once this technology is in place, the firm can use fuel j for any input v and buys the fuel from the energy sector at price $r_j + \tau_j$. We denote the set of fuels for which firm φ has incurred the fixed costs by $\mathcal{J}(\varphi) \subseteq J$.

Note that we have two sources of heterogeneity among firms in a given industry. The first is the firms' core productivity, φ, and the second is the vectors $\{a_j(v, \varphi)\}_{v \in [0,1]}$ of input efficiencies for each fuel j. That is, firms differ in their productivity in each of the two production stages: the production of inputs and the assembly of inputs into a final good.

As in Antras et al. (2017), we assume that both of these sources of heterogeneity are exogenously determined. In particular, firms draw their core productivity from a Pareto distribution:

$$F_k(\varphi) = 1 - \left(\frac{\underline{\varphi}_k}{\varphi} \right)^{\alpha_k}. \tag{12}$$

The lower bound of the productivity draws is given by $\underline{\varphi}_k > 0$ while the dispersion of the productivity draws is governed by the shape parameter $\alpha_k > 0$. A lower α_k implies a higher degree of productivity dispersion across firms.[19] Moreover, the input efficiencies, $1/a_j(v,\varphi)$, for each firm are drawn from a Fréchet distribution:

$$G_k(a) = e^{-T_{jk}a^{-\theta_k}}, \tag{13}$$

where each draw is assumed independent across inputs and fuels. The scale parameter $T_{jk} > 0$ represents fuel j's state of technology in industry k and the shape parameter $\theta_k > 0$ reflects the amount of variation in $1/a_j(v,\varphi)$ across inputs. A large T_{jk} means that fuel j is well suited for the production of inputs in industry k while a large θ_k implies low variation in the efficiency of fuel j across inputs.

4.4 The firm's problem

We now describe the optimal behavior of a manufacturing firm in this model. The firm's problem consists of two stages. In the first stage, the firm chooses the optimal set of fuels to use in production. In the second step, the firm decides which fuel in this set to use in the production of each input and sets the price of its final good to maximize variable profit. We begin with the second step before we turn to the choice of the optimal fuel set.

4.4.1 The firm's behavior after the set of fuels is chosen

A firm in industry k with productivity φ and fuel set $\mathcal{J}(\varphi)$ will choose to produce each input v at the lowest possible unit cost. The lowest unit cost to produce input v is

$$z(v,\varphi;\mathcal{J}(\varphi)) = \min_{j\in\mathcal{J}(\varphi)} \Big\{ z_j(v,\varphi) = (r_j + \tau_j)a_j(v,\varphi) \Big\}. \tag{14}$$

The probability that a fuel $j \in \mathcal{J}(\varphi)$ is the lowest cost alternative depends on the fuels' distributions of unit costs across inputs. We obtain these distributions by substituting $z_j(v,\varphi) = (r_j + \tau_j)a_j(v,\varphi)$ into the distributions of input efficiencies in (13):

$$G_{jk}(z) = \mathbb{P}\Big[z_j(v,\varphi) \le z \Big] = 1 - \exp\Big[-T_{jk}\Big(\frac{r_j+\tau_j}{z}\Big)^{-\theta_k} \Big]. \tag{15}$$

The probability that a fuel $j \in \mathcal{J}(\varphi)$ is the lowest cost alternative to produce input v is then given by,

$$\Psi_{jk}(\varphi;\mathcal{J}(\varphi)) = \int_0^\infty \prod_{\substack{s\in\mathcal{J}(\varphi),\\ s\neq j}} \Big[1 - G_{sk}(z) \Big] dG_{jk}(z) = \frac{T_{jk}(r_j+\tau_j)^{-\theta_k}}{\sum_{s\in\mathcal{J}(\varphi)} T_{sk}(r_s+\tau_s)^{-\theta_k}}. \tag{16}$$

This probability is also the share of inputs $v \in [0,1]$ that is produced using fuel j. Intuitively, fuels that are well suited for production in industry k (high T_{jk}) and are cheap to purchase (low $r_j + \tau_j$) will be used to a greater extent in the production of a

[19] There are two main reasons to assume a Pareto distribution. First, it provides a good fit to the distribution of firm size observed in the data (Axtell, 2001). Second, it gives rise to closed-form solutions for key endogenous variables in this type of model, e.g. the productivity cutoffs determining whether an entrant chooses to stay and produce (Melitz and Redding, 2014).

firm's output. The parameter θ_k governs how sensitive this share is to changes in the price and the carbon tax of the fuel. A higher θ_k implies a lower degree of variation in $1/a_j(v,\varphi)$ across fuels for a given input v, which means that relative fuel prices will be more important for the optimal fuel choice.

We will refer to the numerator in equation (16) as fuel j's *production potential*. Fuels with higher production potentials are used in the production of more intermediate inputs given that they are included in the firm's fuel set. The denominator in equation (16) is the sum of production potentials across the fuel set $\mathcal{J}(\varphi)$. We will refer to this term as firm φ's *production capability*.

As highlighted by Eaton and Kortum (2002), the distribution of $z(v,\varphi;\mathcal{J}(\varphi))$ when $z_j(v,\varphi)$ follows the distribution in (15) is independent of conditioning on a specific fuel j.[20] A fuel associated with better technology and lower price will be used for a wider range of inputs, exactly up to the point at which the distribution of unit costs for what that fuel produces is the same as the overall distribution of unit costs across inputs. This implies that $\Psi_{jk}(\varphi;\mathcal{J}(\varphi))$ is also equal to the spending share of fuel j and that the optimal ratio of fuel quantities between any two fuels in the fuel set is given by,

$$\frac{X_j(\varphi)}{X_s(\varphi)} = \frac{T_{jk}}{T_{sk}}\left[\frac{r_j+\tau_j}{r_s+\tau_s}\right]^{-(1+\theta_k)}, \tag{17}$$

where $X_j(\varphi)$ is the total quantity of fuel j used by the firm. The elasticity of substitution between different fuels in $\mathcal{J}(\varphi)$ is thus given by $1+\theta_k$.

The marginal cost of the firm is,

$$c_k(\varphi;\mathcal{J}(\varphi)) = \frac{1}{\varphi}Z_k(\varphi;\mathcal{J}(\varphi)), \tag{18}$$

where $Z_k(\varphi;\mathcal{J}(\varphi)) = \left[\int_0^1 z(v,\varphi;\mathcal{J}(\varphi))^{1-\rho_k}\right]^{\frac{1}{1-\rho_k}}$. Following Eaton and Kortum (2002), we can express the CES unit cost index in terms of the deeper parameters of the model,

$$Z_k(\varphi;\mathcal{J}(\varphi)) = \gamma_k\left[\sum_{j\in\mathcal{J}(\varphi)} T_{jk}(r_j+\tau_j)^{-\theta_k}\right]^{-\frac{1}{\theta_k}}, \tag{19}$$

where $\gamma_k = \left[\Gamma\left(\frac{\theta_k+1-\rho_k}{\theta_k}\right)\right]^{\frac{1}{1-\rho_k}}$ and Γ is the Gamma function.[21] Equation (19) shows that the marginal cost is directly related to the firm's production capability. The firm can reduce its marginal cost by adopting fuels with higher states of technology, T_{jk}, and lower prices, $r_j+\tau_j$. It can also reduce its marginal cost by adding new fuels to the fuel set since the added fuels will always be the cheapest alternative to produce a non-zero mass of intermediate inputs.

Under monopolistic competition, the firm maximizes profit by setting the price for its final good as a constant mark-up over marginal cost, $p_k(\varphi;\mathcal{J}(\varphi)) = \frac{\sigma_k}{\sigma_k-1}c_k(\varphi;\mathcal{J}(\varphi))$. Given this pricing rule, variable profit is given by,

$$\pi_k(\varphi;\mathcal{J}(\varphi)) = B_k c_k(\varphi;\mathcal{J}(\varphi))^{1-\sigma_k}, \tag{20}$$

[20] That is, for all fuels j, we have $\mathbb{P}\Big[z(v,\varphi;\mathcal{J}(\varphi))\le z\Big] = \mathbb{P}\Big[z(v,\varphi;\mathcal{J}(\varphi))\le z | z(v,\varphi;\mathcal{J}(\varphi)) = z_j(v,\varphi)\Big]$.

[21] For the Gamma function to be well defined, we impose the parameter restriction $\theta_k+1-\rho>0$.

where $B_k = \frac{1}{\sigma_k}\left(\frac{\sigma_k}{\sigma_k - 1}\right)^{1-\sigma_k} E_k P_k^{\sigma_k - 1}$ is an industry-specific market demand index. Since $\sigma_k > 1$, lower marginal cost translates into higher variable profits. In turn, the marginal cost depends both on the firm's core productivity and its choice of fuel set. We next characterize the optimal choice of $\mathcal{J}(\varphi)$.

4.4.2 The firm's choice of fuels

Adding fuels to the fuel set $\mathcal{J}(\varphi)$ increases the firm's variable profit by reducing its marginal cost. What prevents the firm from setting $\mathcal{J}(\varphi) = J$ and using all of the available fuels in its production process is that the firm needs to pay a fixed cost for every fuel it utilizes. The fixed cost can be interpreted as representing the cost of installing the technology associated with using a specific fuel and may vary across fuels. To choose the optimal fuel set, the firm needs to trade off the fuels' contribution to variable profit against their fixed cost. Specifically, the firm chooses the fuel set that solves the following profit maximization problem,

$$\max_{\mathcal{J} \subseteq J} \quad \Pi_k(\varphi; \mathcal{J}) = B_k \left(\frac{\varphi}{\gamma_k}\right)^{\sigma_k - 1} \left[\sum_{j \in \mathcal{J}} T_{jk}(r_j + \tau_j)^{-\theta_k}\right]^{\frac{\sigma_k - 1}{\theta_k}} - \sum_{j \in \mathcal{J}} w f_{jk}, \tag{21}$$

where $\Pi_k(\varphi; \mathcal{J})$ is firm φ's expected profit when selecting fuel set $\mathcal{J}$.[22] This problem is difficult to solve because the contribution to variable profit of using a specific fuel j generally depends on what other fuels the firm uses. The decision to use one fuel affects whether it is optimal to use another fuel. Hence, the fuel decisions are interdependent and the problem in (21) is a combinatorial discrete choice problem.

There are two sources for the interdependencies among fuels in this model. The first is that fuels compete against each other in the production of each input v. This gives rise to a negative supply-side complementarity among fuels. Intuitively, if we hold the firm's output fixed and add a new fuel to the fuel set, the firm will use less of its initial fuels since the new fuel will replace them in the production of at least some inputs. This negative complementarity is counteracted by a positive complementarity on the demand side. The added fuel will reduce the firm's marginal cost and increase its optimal scale of production. This in turn will induce the firm to use more of each fuel. The strength of the negative complementarity is governed by the elasticity of substitution among fuels, $1 + \theta_k$, while the strength of the positive complementarity is determined by the demand elasticity, σ_k.

Whether the negative or the positive complementarity dominates determines how the marginal value of a fuel depends on the other fuels used by the firm. Adopting the notation in Arkolakis et al. (2023), we can define the marginal value of fuel j as:

$$D_j \Pi_k(\varphi; \mathcal{J}) = \Pi_k(\varphi; \mathcal{J} \cup \{j\}) - \Pi_k(\varphi; \mathcal{J} \setminus \{j\}) \tag{22}$$

where D_j is the marginal value operator. We can then establish the following result.

Proposition 1. *Given any two fuel sets $\mathcal{J}_1 \subset \mathcal{J}_2 \subseteq J$ and any fuel $j \in J$:*

(i) If $1 + \theta_k > \sigma_k$, then $D_j \Pi_k(\varphi; \mathcal{J}_1) > D_j \Pi_k(\varphi; \mathcal{J}_2)$.

[22] $\Pi_k(\varphi; \mathcal{J})$ is the expected profit since the vectors of input efficiencies $1/a_j(v, \varphi)$ are realized after the firm selects $\mathcal{J}$.

(ii) If $1 + \theta_k < \sigma_k$*, then* $D_j\Pi_k(\varphi; \mathcal{J}_1) < D_j\Pi_k(\varphi; \mathcal{J}_2)$.

(iii) If $1 + \theta_k = \sigma_k$*, then* $D_j\Pi_k(\varphi; \mathcal{J}_1) = D_j\Pi_k(\varphi; \mathcal{J}_2)$.

The proof of this proposition is provided by Arkolakis et al. (2023) in a more general theoretical framework that nests our model. The analogous proof applied to our model is in Appendix C.

Proposition 1 states that if $1 + \theta_k > \sigma$, the marginal value of a fuel is monotonically decreasing as new fuels are added to the fuel set. In other words, the positive contribution to variable profits of using fuel j decreases if the firm's fuel set expands since fuels are relatively substitutable and demand is relatively inelastic. On the other hand, if $1+\theta_k < \sigma$, the marginal value of a fuel is monotonically increasing as new fuels are added. The positive contribution to variable profits of using fuel j is increasing if the fuel set expands since the scale effect from lower marginal cost dominates the substitution effect. Lastly, if $1 + \theta_k = \sigma_k$, the two forces cancel each other out and the contribution of fuel j is always equal to $B_k(\varphi/\gamma_k)^{\sigma_k-1}T_{jk}(r_j + \tau_j)^{-\theta_k}$ irrespective of what other fuels are included in the fuel set.

The policy function $\mathcal{J}(\varphi) = \arg\max_{\mathcal{J}} \Pi_k(\varphi; \mathcal{J})$, mapping firm productivity into optimal fuel sets, also depends on the relative strength of the positive and negative complementarities between fuels. If $1 + \theta_k > \sigma$, there is no necessary relationship between firm productivity and the size of the fuel set. If $1 + \theta_k \leq \sigma$, high productivity firms always use weakly larger fuel sets than low productivity firms. To gain intuition for this result, suppose we have two firms with productivity levels $\varphi_1 < \varphi_2$. The low productivity firm φ_1 finds it optimal to use electricity, heating oil, and coal to produce its inputs. This implies that all other fuels have negative marginal values given this fuel set and productivity φ_1. However, the marginal value of each fuel increases with productivity so it might be profitable for firm φ_2 to add fuels to this fuel set. Suppose it is profitable to add a biofuel. If $1 + \theta_k > \sigma$, adding the biofuel decreases the marginal values of the initial fuels. If the marginal values of heating oil and coal become negative, it might be optimal for firm φ_2 to drop these fuels and only use electricity and the biofuel. This could never happen if $1 + \theta_k \leq \sigma$. In this case, the marginal values of the initial fuels would only increase by adding the biofuel. Hence, it could never be profitable for firm φ_2 to drop fuels that firm φ_1 finds optimal to use. That is, $\mathcal{J}(\varphi_1) \subseteq \mathcal{J}(\varphi_2)$ if $1 + \theta_k \leq \sigma$.

Regardless of whether $1 + \theta_k > \sigma$ or $1 + \theta_k \leq \sigma$, high-productivity firms choose fuel sets with weakly higher production capabilities than low-productivity firms. That is,

$$\left[\sum_{j\in\mathcal{J}(\varphi_1)} T_{jk}(r_j + \tau_j)^{-\theta_k}\right]^{\frac{\sigma_k-1}{\theta_k}} \leq \left[\sum_{j\in\mathcal{J}(\varphi_2)} T_{jk}(r_j + \tau_j)^{-\theta_k}\right]^{\frac{\sigma_k-1}{\theta_k}}, \quad \varphi_1 < \varphi_2. \tag{23}$$

Hence, the model predicts that the cost advantage of more productive firms is reinforced by their ability to choose more efficient fuel sets. In Section 5, we show that this prediction is sometimes violated in the data, and extend the model by allowing the fuel-specific fixed investment costs to vary across firms to account for this violation.

We can solve the problem in (21) by implementing the solution algorithm to combinatorial discrete choice problems developed by Arkolakis et al. (2023). This approach exploits that the marginal value functions, $D_j\Pi_k(\varphi; \mathcal{J})$, are monotonic functions of $\mathcal{J}$, to reduce the dimensionality of the choice space. Without such an algorithm, it would be infeasible to solve the problem in (21) since the alternative is to evaluate $\Pi_k(\varphi; \mathcal{J})$ for each possible combination of fuels. In our empirical setting, where firms can choose among 18 fuels, this would amount to performing $2^{18} = 262,144$ calculations per firm. We describe Arkolakis et al. (2023)'s solution algorithm in more detail in Section 5.

4.5 General equilibrium

Let us now characterize the general equilibrium of the model. Labor is the only factor of production and we assume that workers can move freely across the manufacturing industries and the energy sector. This ensures that the wage rate is the same in all industries and sectors. Furthermore, we choose labor as the numéraire such that $w = 1$. This entails that the fuel price vector is pinned down by the unit labor requirements in the energy sector, $\{r_j = d_j\}_{j \in J}$. General equilibrium then follows from equilibrium in each manufacturing industry k. The variables characterizing these equilibria are total consumer expenditures, E_k, the survival productivity cutoff, $\tilde{\varphi}_k$, the market demand index, B_k, the mass of entrants, M_k^E, the aggregate quantity used of each fuel j, X_j^k, and the aggregate labor employed both directly and indirectly via the energy market, L_k.

The Cobb-Douglas preferences in (6) imply that consumers spend a constant share β_k of their income on goods produced in industry k. Total income is given by $w\bar{L} + \sum_k \sum_{j \in J} \tau_j X_j^k$, where the first term represents total labor income and the second term represents total tax income. We can thus express total expenditures on goods produced by industry k as,

$$E_k = \beta_k \Big(w\bar{L} + \sum_k \sum_{j \in J} \tau_j X_j^k \Big). \tag{24}$$

Note that the total supply of labor $\bar{L}$ is exogenous while X_j^k is an endogenous variable that depends on other equilibrium variables. Also note that E_k is equal to aggregate industry revenue, R_k, in a closed economy.

Since production involves fixed costs, some entrants with low productivity will not find it profitable to produce. The survival productivity cutoff $\tilde{\varphi}_k$ is defined as the productivity level at which a firm in industry k makes zero profits,

$$B_k \Big(\frac{\tilde{\varphi}_k}{\gamma_k} \Big)^{\sigma_k - 1} \Big[\sum_{j \in \mathcal{J}(\tilde{\varphi}_k)} T_{jk} (r_j + \tau_j)^{-\theta_k} \Big]^{\frac{\sigma_k - 1}{\theta_k}} - \sum_{j \in \mathcal{J}(\tilde{\varphi}_k)} w f_{jk} = 0. \tag{25}$$

Entrants with productivity below $\tilde{\varphi}_k$ exit immediately. This zero-profit condition implies that the market demand index B_k can be expressed as a monotonically decreasing function of $\tilde{\varphi}_k$. By combining this with the free-entry condition, stating that the equilibrium mass of entrants must be such that the expected profit before entering is equal to the sunk entry cost, we obtain

$$\int_{\tilde{\varphi}_k}^{\infty} \Big[B_k(\tilde{\varphi}_k) \Big(\frac{\varphi}{\gamma_k} \Big)^{\sigma_k - 1} \Big[\sum_{j \in \mathcal{J}(\varphi)} T_{jk} (r_j + \tau_j)^{-\theta_k} \Big]^{\frac{\sigma_k - 1}{\theta_k}} - \sum_{j \in \mathcal{J}(\varphi)} w f_{jk} \Big] dF_k(\varphi) = w f_{ek}. \tag{26}$$

This equation identifies a unique equilibrium productivity cutoff $\tilde{\varphi}_k$ (see Antras et al. (2017) for a proof). With this in hand, the equilibrium market demand index B_k follows directly from (25). Moreover, the equilibrium productivity distribution among producing firms is then given by $F_k(\varphi)/[1 - F_k(\tilde{\varphi}_k]$.

The mass of producing firms in industry k is $M_k = R_k/\bar{R}_k$ where $\bar{R}_k$ denotes average firm revenue. Together with (24) and (26) and the fact that $M_k^E = M_k/[1 - F_k(\tilde{\varphi}_k]$, this implies that the mass of entrants in equilibrium can be expressed as,

$$M_k^E = \frac{\beta_k\left(w\bar{L} + \sum_k \sum_{j\in J} \tau_j X_j^k\right)}{\sigma_k w\left[f_{ek} + \int_{\tilde{\varphi}_k}^{\infty} \sum_{j\in\mathcal{J}(\varphi)} f_{jk} dF_k(\varphi)\right]}. \tag{27}$$

The industry equilibrium demand for fuel j, X_j^k, depends on the equilibrium variables B_k and M_k^E. To see this, note that the firm-level equilibrium demand for fuel $j \in \mathcal{J}(\varphi)$ is given by,

$$X_j(\varphi) = T_{jk}(r_j + \tau_j)^{-(1+\theta_k)}(\sigma_k - 1)B_k\left(\frac{\varphi}{\gamma_k}\right)^{\sigma_k - 1}\left[\sum_{j\in\mathcal{J}(\varphi)} T_{jk}(r_j + \tau_j)^{-\theta_k}\right]^{\frac{\sigma_k - \theta_k - 1}{\theta_k}}, \tag{28}$$

which is a share $(\sigma_k - 1)\Psi_{jk}(\varphi; \mathcal{J}(\varphi))$ of the firms' deflated operating profit $\pi_k(\varphi; \mathcal{J}(\varphi))/(r_j + \tau_j)$. It follows that aggregate equilibrium demand for fuel j is given by,

$$X_j^k = M_k^E \int_{\tilde{\varphi}_k}^{\infty} X_j(\varphi) dF_k(\varphi). \tag{29}$$

Finally, the labor market must clear in general equilibrium. Note that an industry employs labor directly to pay for the fixed costs of entry and the fixed costs of installing fuel-specific technology. However, it also requires labor to produce the fuels it purchases from the energy sector. The labor market clearing condition in industry k can be expressed as,

$$L_k = \frac{E_k - \sum_{j\in J} \tau_j X_j^k}{w}. \tag{30}$$

This condition states that the industry uses its revenues to pay for aggregate labor and carbon tax costs. Substituting for E_k using (24) shows that (30) implies labor market clearing in the whole economy, $\bar{L} = \sum_k L_k$.

We can define the general equilibrium in this model as follows:

Definition 1. *General equilibrium is a set of policy functions $\mathcal{J}(\varphi)$ and aggregates $\{w, r_j, E_k, \tilde{\varphi}_k, B_k, M_k^E, X_j^k, L_k\}$ such that:*

1. *given the aggregates, the policy functions solve the problem in (21), and*
2. *given the policy functions, the aggregates satisfy $w = 1$, $r_j = d_j\ \forall j \in J$, and equations (24), (25), (26), (27), (29), and (30).*

5 Structural estimation

We use our firm-level data on fuel use to estimate the parameters of the model. The goal is to have a fully estimated model that can explain the patterns of fuel consumption observed in the data and to then use this model to quantify the effects of carbon pricing.

We focus the estimation such that the model matches fuel consumption in 2004. We do this for two reasons. First, 2004 is the only year in our sample in which none of the firms are subject to the EU ETS which simplifies the estimation of some of the parameters. Second, it allows us to compare the counterfactual emission path when increasing the CO2 price with the actual path of emissions between 2004 and 2020. Such comparison is valuable for assessing the empirical validity of the model.

Our estimation procedure follows the steps laid out in Antras et al. (2017). We first estimate each fuel's production potential, $T_{jk}(r_j + \tau_j)^{-\theta_k}$, the fuel elasticity of substitution, $1 + \theta_k$, and the elasticity of demand, σ_k, by using the full panel of firms between 2004 and 2020. We allow the production potentials to vary by industry-year and the two elasticities to vary by industry. The rest of the parameters are estimated using the simulated method of moments where we target industry-specific moments in 2004. In particular, this step provides estimates of the market demand index, B_k, the distribution of fuel-specific investment costs, f_{kj}, and the Pareto shape parameter, α_k. The parameters are estimated by industry to enable the model to account for the observed cross-industry heterogeneity in fuel consumption. To ensure a reasonable amount of statistical power, we confine the structural analysis to 2-digit industries with more than 30 observations in 2004.[23]

5.1 Estimating fuels' production potentials

The fuels' production potentials measure to what extent the fuels' are used in the firms' production process given that the firms choose to invest in them. A fuel with a high production potential is used for a larger part of the production process than a fuel with a low production potential. Equation (16) implies that spending on fuel j relative to electricity is equal to the production potential of fuel j relative to electricity. We use this to specify the following estimating equation,

$$\log \frac{(r_{jt} + \tau_{jt})X_{ijt}}{(r_{et} + \tau_{et})X_{iet}} = \log \frac{T_{jkt}(r_{jt} + \tau_{jt})^{-\theta_k}}{T_{ekt}(r_{et} + \tau_{et})^{-\theta_k}} + \varepsilon_{ijt}, \tag{31}$$

where subscripts i, t and e denote firms, years and electricity, respectively, and ε_{ijt} is an error term.

We estimate equation (31) for each 2-digit industry separately with OLS and fuel-year fixed effects. The fixed effects equal the average log spending on fuel j relative to electricity among firms that use fuel j within industry-years. The true average is identified by this estimator as long as the error term is not correlated with the firms' selection of fuels. We interpret ε_{ijt} as representing measurement error since the model implies (31) exactly. Moreover, since we can only identify production potentials for fuels that are used in a given industry-year, we assume that $T_{jk}(r_j + \tau_j)^{-\theta_k}$ approaches zero for fuels that are not used and exclude them from the structural analysis.[24]

Figure 7 plots the 2004 estimates of the fuels' log production potentials relative to electricity. The hollow squares represent electricity, the gray circles represent fossil fuels, and the black triangles represent biofuels. The plot shows that in all industries except the

[23]The dropped industries are Beverages (NACE 11), Tobacco (NACE 12), Wearing apparel (NACE 14), Leather products (NACE 15), and Coke and refined petroleum products (NACE 19).

[24]We also exclude fuels that are used by a single firm within industry-years.

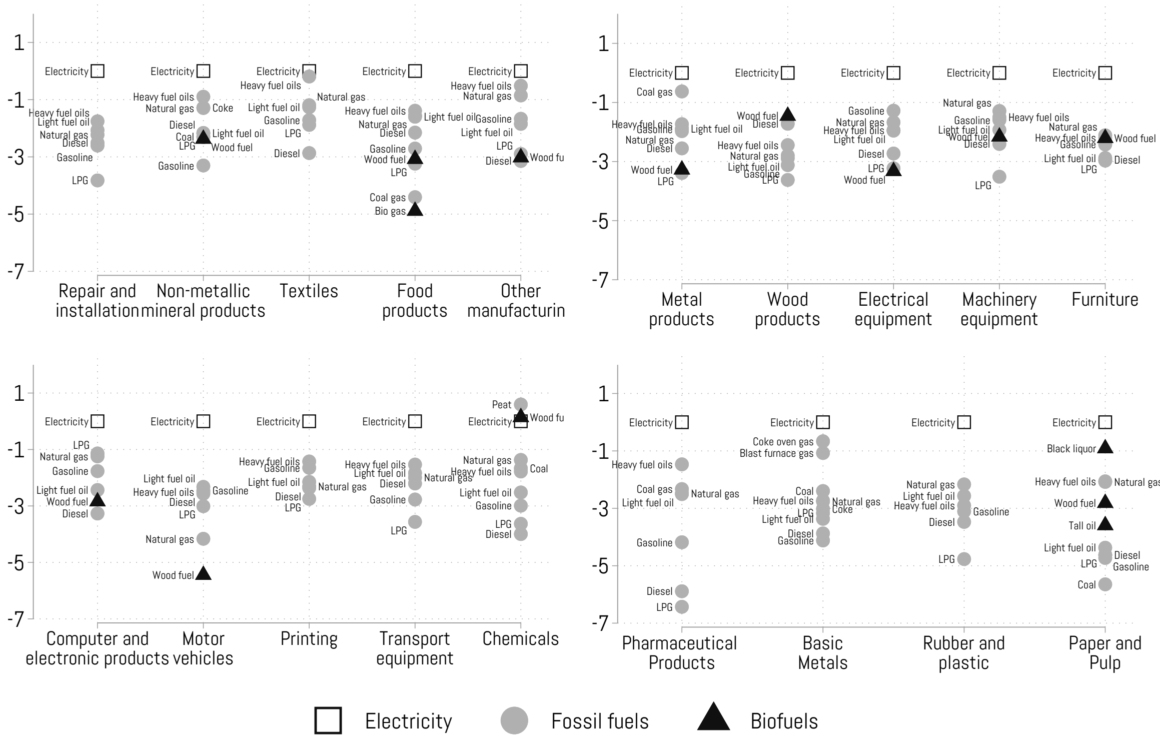

Figure 7: Log production potentials by industry

Notes: The figure shows the estimated 2004 log production potentials relative to electricity by industry. The base sample is an unbalanced panel of 90,666 firm-year observations over the years 2004-2020. The production potentials are estimated at the fuel-firm-year level and the base sample at this level contains 1,631,988 observations. We estimate equation (31) via OLS and fuel-year fixed effects for each 2-digit industry separately. We drop industries with less than 30 observations in 2004 and fuels used by a single firm within industry-years.

chemical industry, electricity is the energy source with the highest production potential. Moreover, the production potentials of fossil fuels and biofuels vary substantially across industries which is consistent with the great heterogeneity in the relative consumption of electricity, fossil fuels, and biofuels documented in Figure 1. Interpreted through the lens of the model, this heterogeneity origins from industry differences in the fuels' relative states of technology, T_k/T_e, and the firms' responsiveness to the relative fuel prices, θ_k.[25]

We can now test the model prediction in equation (23) that larger firms choose weakly better fuel sets, or equivalently, have weakly higher production capabilities, $\sum_j T_{jkt}(r_{jt} + \tau_{jt})^{-\theta_k}$, than smaller firms. Figure 8 plots the average share of firms across industries with sales and production capabilities either above or below the industry median. The model predicts that all firms with sales above the median also have production capabilities above the median and vice versa. The figure shows that this prediction is violated in the data since a non-negligible share of the firms falls outside the categories "Above, above" and "Below, below". To address this, we extend the model such that it can explain this pattern by allowing the fixed investment costs to vary across firms. A larger firm could have a lower production capability than a smaller firm if the fixed investment

[25]In Table D1 in Appendix D, we present the point estimates and standard errors of the 2004 production potentials illustrated in Figure 7.

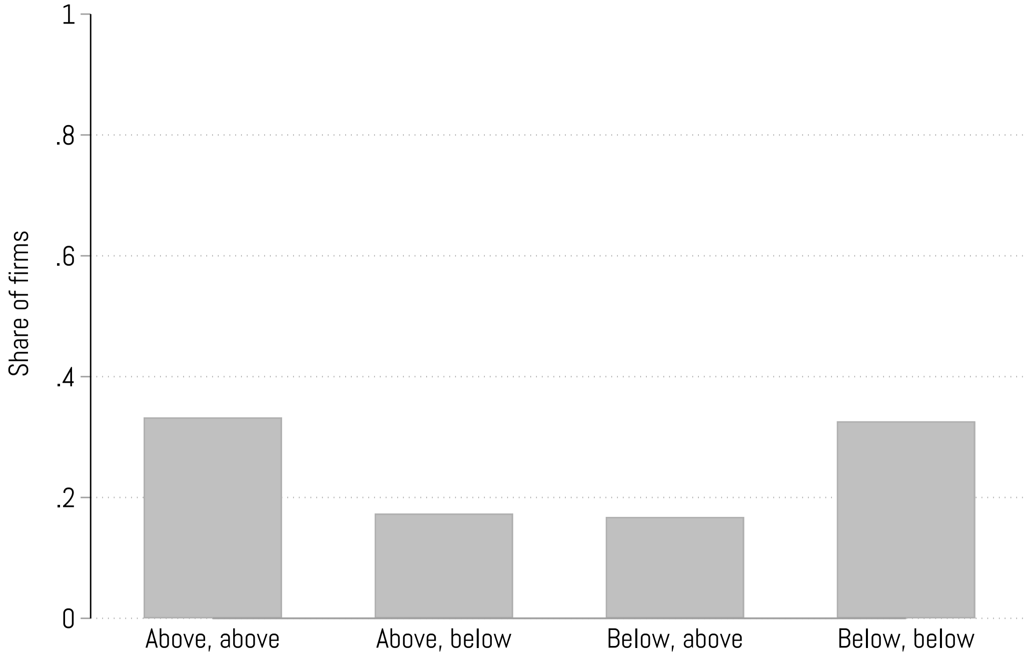

Figure 8: Average shares of firms above/below median sales and production capability

Notes: The figure shows 2004 averages across industries of the shares of firms with sales and production capability above/below the industry medians. Specifically, the "Above, above" bar shows the average share of firms with sales above the median and production capability above the median, the "Above, below" bar shows the average share of firms with sales above the median and production capability below the median, and so on. We drop industries with less than 30 observations and fuels used by a single firm within industry-years.

costs are higher for the larger firm. We discuss in Section 5.4 how we implement this extension in the simulation of the model.

5.2 Estimating elasticities of substitution between fuels

To estimate θ_k, we express the estimated production potentials as,

$$\log\left[\frac{\widehat{T_{jkt}(r_{jt}+\tau_{jt})^{-\theta_k}}}{T_{ekt}(r_{et}+\tau_{et})^{-\theta_k}}\right] = \log\frac{T_{jkt}}{T_{ekt}} - \theta_k \log\left[\frac{r_{jt}+\tau_{jt}}{r_{et}+\tau_{et}}\right] + \eta_{jkt}, \tag{32}$$

where η_{jkt} is an error term. A problem we face in estimating this equation is that the fuels' relative states of technology, $\log(T_{jkt}/T_{jkt})$, are unobserved. Leaving out this term from the estimating equation may yield biased estimates of θ_k if the states of technology are correlated with fuel prices. For example, an increase in the state of technology of fuel j in a given industry may increase the price of this fuel via higher demand. In this case, the OLS estimates of θ_k are downward biased.

To deal with this issue, we use an instrumental variable (IV) approach where we instrument for the relative fuel prices using the fuels' carbon taxes, τ_{jt}. The carbon taxes are determined by the carbon contents of the fuels and not by fuel-specific technology. Hence, a change in the technology of a fuel will not affect the carbon tax as long as it does not influence the Swedish government to set a new tax level. Carbon taxes could be correlated with the fuels' states of technology if changes in the carbon taxes affect investments in technological development. However, the technology frontier is likely to be determined outside of Sweden and thus not respond to changes in the Swedish carbon tax.

We implement the IV approach by first estimating the first stage between the log relative prices and the log carbon taxes. We then estimate the reduced form relationship between the log production potentials and the log carbon taxes for each industry separately.[26] The IV estimate of θ_k is defined by the ratio between the reduced form estimate and the first stage estimate. We use bootstrapping with replacement to compute standard errors. In a robustness check, we include the citation-weighted sum of patents related to climate change mitigation technologies used in Section 3. This controls for technological development that is arguably more likely to be correlated with carbon taxes.

Figure 9 plots our estimates of the elasticities of substitution $1 + \theta_k$.[27] The estimates range between 1.39 in the repair and installation industry and 4.40 in the paper and pulp industry. The high elasticity in the paper and pulp industry could explain why firms in this industry have been so successful in switching away from fossil fuels over the sample period (see Section 3.1). However, the elasticity is also high in the basic metals industry and firms in this industry have not reduced their use of fossil fuels. This can be explained by the fact that there are not as many feasible alternatives to fossil fuels in the production of basic metals compared to paper and pulp. Our estimated elasticities of substitution are within the range of previous estimates in the literature. Papageorgiou et al. (2017) find that the elasticity of substitution between clean and dirty energy inputs is between 1.38 and 2.86 in the non-energy sector. Jo (2023) finds that it is 3.32 in the manufacturing sector as a whole but that there is considerable heterogeneity across 2-digit industries with elasticities ranging from 2.44 to 5.12. Figure D1 in Appendix D shows that our estimates are robust to controlling for technical change related to CO2 mitigation.

5.3 Estimating elasticities of demand

To estimate the elasticities of demand, we use the following relationship implied by the model,

$$\frac{R_{ikt}}{C_{ikt}} = \frac{\sigma_k}{1 - \sigma_k}, \tag{33}$$

where R_{ikt} is firm-level revenue and C_{ikt} is firm-level variable costs. In particular, we compute median revenues over variable costs within each industry and then use equation (33) to obtain an estimate of σ_k. Firm-level revenue is directly observed in the data and firm-level variable cost is calculated as the sum of expenditures on materials, wages, energy inputs, and variable capital costs.[28] We plot the estimates in Figure 9 together with bootstrapped standard errors. The estimates range between 2.14 and 3.85 and are well within the interval of previous estimates of demand elasticities in the manufacturing sector (Broda and Weinstein, 2006; De Loecker, 2011).

[26] The fuels' production potentials are calculated based on averages across firms. We therefore weigh each observation in the reduced form regression by the number of firms used to calculate those averages. This results in larger estimates of θ_k compared to an unweighted reduced form regression (see Table D2 in Appendix D).

[27] Tables D2 and D3 in Appendix D present the first stage and reduced form estimates for each industry.

[28] In calculating capital expenditures, we assume a real interest rate of 2% on the firms' value of tangible assets.

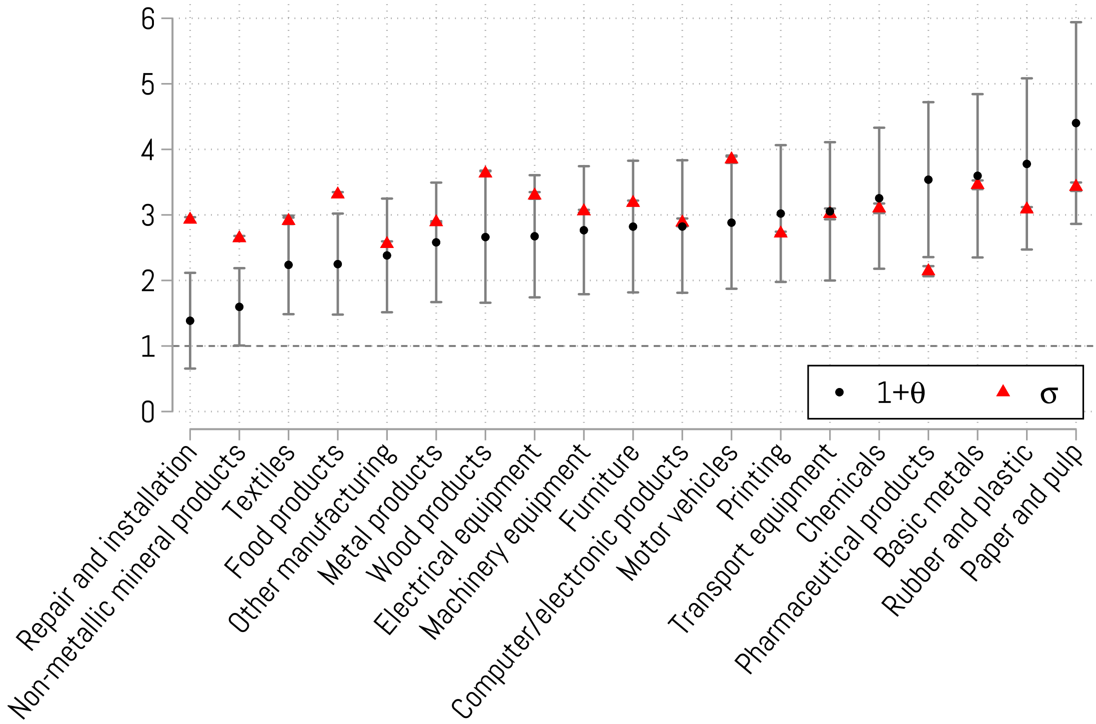

Figure 9: Elasticity of substitution between fuels and elasticity of demand by industry

Notes: The elasticities of substitution are estimated at the fuel-year level. There are 306 fuel-years in our base sample. The number of fuel-year observations underlying the estimates varies by 2-digit industry and is reported in Table D3. We estimate equation (32) in a two-sample IV procedure. We first estimate the first-stage relationship between the log relative price and the log carbon tax for all industries. We then estimate the reduced form relationship between the log relative production potentials and the log carbon tax for each industry separately. Standard errors are bootstrapped and based on 200 replications with replacement. The elasticities of demand are computed from the median revenue over total variable cost within 2-digit industries in our base sample. Standard errors are bootstrapped and based on 200 replications with replacement.

As we discuss in Section 4, the relative magnitude of $1+\theta_k$ and σ_k governs whether the benefit of using a fuel increases or decreases with other fuels. If $1+\theta_k > \sigma$, investing in new fuels decreases the marginal value of existing fuels in the firms' fuel set and vice versa. Figure 9 shows that we have $1+\theta_k > \sigma$ in seven industries and $1+\theta_k < \sigma$ in twelve industries. In many industries, however, the two elasticities lie close to each other, suggesting that the decision to use one fuel is more or less independent of the decision to use other fuels. Whether or not the elasticity of substitution is greater than the elasticity of demand matters for how we apply Arkolakis et al. (2023)'s solution algorithm. However, it also matters for the effect of carbon pricing. If $1+\theta_k > \sigma$, a reduction in the production potentials of fossil fuels due to higher carbon taxes increases the benefit of using alternative fuels. If $1+\theta_k < \sigma$, the opposite is the case. Hence, industries with strong complementarities among fuels (i.e. high σ_k relative to $1+\theta_k$) are likely to suffer more in terms of output from an increase in the tax than industries with weak complementarities among fuels.

5.4 Estimating fuel-specific investment costs

In the simulation of the model, we allow the fuel-specific investment costs to vary across firms to account for the fact that the distribution of sales and production capabilities do not perfectly overlap in the data. In particular, we assume that the fixed costs are drawn from a lognormal distribution such that,

$$f_{ijk} = e^{\mu_{jk}+\mu_k^{disp} z_{ijk}}, \qquad \forall j \neq \text{electricity}, \tag{34}$$

where z_{ijk} is a random draw from a standard normal distribution. The scale parameter, μ_{jk}, is the mean of $\log f_{ijk}$ while the dispersion parameter, μ_k^{disp}, is the standard deviation. The fuel-specific scale parameter is consistent with the large heterogeneity in the number of firms that use different fuels documented in Table 1. The dispersion parameter is constant across fuels but can vary across industries. Greater dispersion weakens the link between firm size and production capability. We cannot identify the fixed cost of electricity, since all firms use it, and set it to zero. This implies that the mass of active firms, M_k, equals the mass of entrants, M_k^E, in the simulation. Modeling fixed costs as drawn from a lognormal distribution is standard in quantitative models of firm behavior and ensures a positive support for the fixed costs (Eaton et al., 2011; Tintelnot, 2017).

We use the simulated method of moments to estimate the following set of parameters for each industry separately,

$$\Theta_k = \{\tilde{B}_k, \{\mu_{jk}\}, \mu_k^{disp}, \alpha_k\}, \tag{35}$$

where $\tilde{B}_k = B_k/\gamma_k^{\sigma_k-1}$ is the market demand index normalized by the scalar $\gamma_k^{\sigma_k-1}$, and α_k is the Pareto shape parameter of the productivity distribution in equation (12). We follow Melitz and Redding (2015) and normalize the lower bound of this distribution such that $\underline{\varphi}_k = 1$.

The estimation proceeds as follows. We first simulate 21,600 firms following the simulation procedure in Antras et al. (2017). Each firm is characterized by a productivity draw from a uniform distribution and a set of fixed costs draws from a standard normal distribution. The set of scale parameters, $\{\mu_{jk}\}$, differ across industries since we assume that only fuels that are actually used in an industry are available to firms. For a given guess of the set of parameters, Θ_k, we invert the distribution of productivity draws into a Pareto distribution and the distributions of fixed cost draws into lognormal distributions. We then solve the combinatorial discrete choice problem in (21) for each simulated firm by the method developed in Arkolakis et al. (2023). Finally, we compute a vector of aggregate moments $\hat{m}(\Theta)$ from the simulated firms and compare them with the corresponding moments m in the actual data. The procedure is iterated over different Θ to find the best match between the simulated and actual moments.

Arkolakis et al. (2023)'s method to solve the firms' problem can be used both when fuels are substitutes $(1 + \theta_k > \sigma_k)$ and complements $(1 + \theta_k < \sigma_k)$ at the extensive margin.[29] The method is a set-valued mapping applied to the firms' choice space that iteratively eliminates non-optimal fuel sets from the choice space. Specifically, the procedure successively updates two bounding sets $\underline{\mathcal{J}}$ and $\overline{\mathcal{J}}$, where $\underline{\mathcal{J}}$ includes all fuels

[29] An alternative solution method to combinatorial discrete choice problems is developed in Jia (2008). However, in our case, this method is only applicable in industries where fuels are complements at the extensive margin.

that must be part of the optimal fuel set $\mathcal{J}^*$ and $\overline{\mathcal{J}}$ excludes all fuels that cannot be part of the optimal fuel set. The procedure converges to a fixed point $[\underline{\mathcal{J}}^F, \overline{\mathcal{J}}^F]$ where $\underline{\mathcal{J}}^F \subseteq \mathcal{J}^* \subseteq \overline{\mathcal{J}}^F$. If $\underline{\mathcal{J}}^F = \overline{\mathcal{J}}^F$, then the procedure has identified the fuel set that solves the firm's profit maximization problem. If $\underline{\mathcal{J}}^F \neq \overline{\mathcal{J}}^F$, then the optimal fuel set can be identified by comparing the firm's profit from choosing among all the fuel combinations that are not excluded by $[\underline{\mathcal{J}}^F, \overline{\mathcal{J}}^F]$. Applying Arkolakis et al. (2023)'s method significantly reduces the computational burden of solving the firms' problem.[30]

We now describe the moments we use to estimate Θ_k. The first set of moments is the share of firms with electricity spending and total fuel spending below the median of these two variables in the data. We use these moments to identify the scale parameter $\tilde{B}_k$. The second set of moments is the share of firms that use each fuel. We use these moments to identify the scale parameters in the fixed costs distribution, $\{\mu_{jk}\}$. The third set of moments contains the share of firms with both sales and production capability below the median, the share of firms with both sales and production capability above the median, and the share of firms with sales below the median and production capability above the median. We use these three moments to identify the dispersion parameter in the fixed costs distribution, μ_k^{disp}. The last moment is the standard deviation in fuel spending and we use this moment to identify the shape parameter in the productivity distribution, α_k. We have selected the moments so that the model matches the pattern of fuel consumption across firms, and not the sales pattern, since our primary interest lies in understanding how fuel consumption and CO2 emissions are affected by carbon taxes.[31]

For each moment r, we define the percent deviation between the moment in the simulated data and the actual data,

$$\hat{d}_r(\Theta_k) = \frac{\hat{m}_r(\Theta_k) - m_r}{m_r}, \tag{36}$$

and denote the vector of deviations as $\hat{d}(\Theta_k)$. The deviations are in percent since the moments are measured in different units. The simulated method of moment estimator is based on the moment condition that $\mathbb{E}[\hat{d}(\Theta_k^*)] = 0$ where Θ_k^* is the true value of Θ_k. The vector of estimates is given by,

$$\hat{\Theta}_k = \arg\min_{\Theta_k} [\hat{d}_r(\Theta_k)' W \hat{d}_r(\Theta_k)] \tag{37}$$

where W is a weighting matrix. We choose the identity matrix as the weighting matrix which means that each moment receives the same weight in the estimation.

Figure 10 plots our estimates of the scale parameter $\tilde{B}_k$, the fixed cost dispersion parameter μ_k^{disp}, and the Pareto shape parameter α_k, by industry. A greater $\tilde{B}_k$ implies that firms use more energy on average. A greater μ_k^{disp} implies more variability in the fixed costs across firms and consequently a weaker link between firm size and production capability. Finally, a smaller α_k implies higher variance in the energy distribution across firms. Our estimates of α_k range between 1.18 and 3.69 with a mean of 2.75 across industries. This can be compared with Shapiro and Walker (2018)'s estimates of the

[30] In the case where $\underline{\mathcal{J}}^F \neq \overline{\mathcal{J}}^F$, we further speed up the solution procedure by applying Arkolakis et al. (2023)'s branching procedure.

[31] Achieving a good match on both the pattern of fuel consumption and sales would require extending the model to include capital, labor, and materials in the production function.

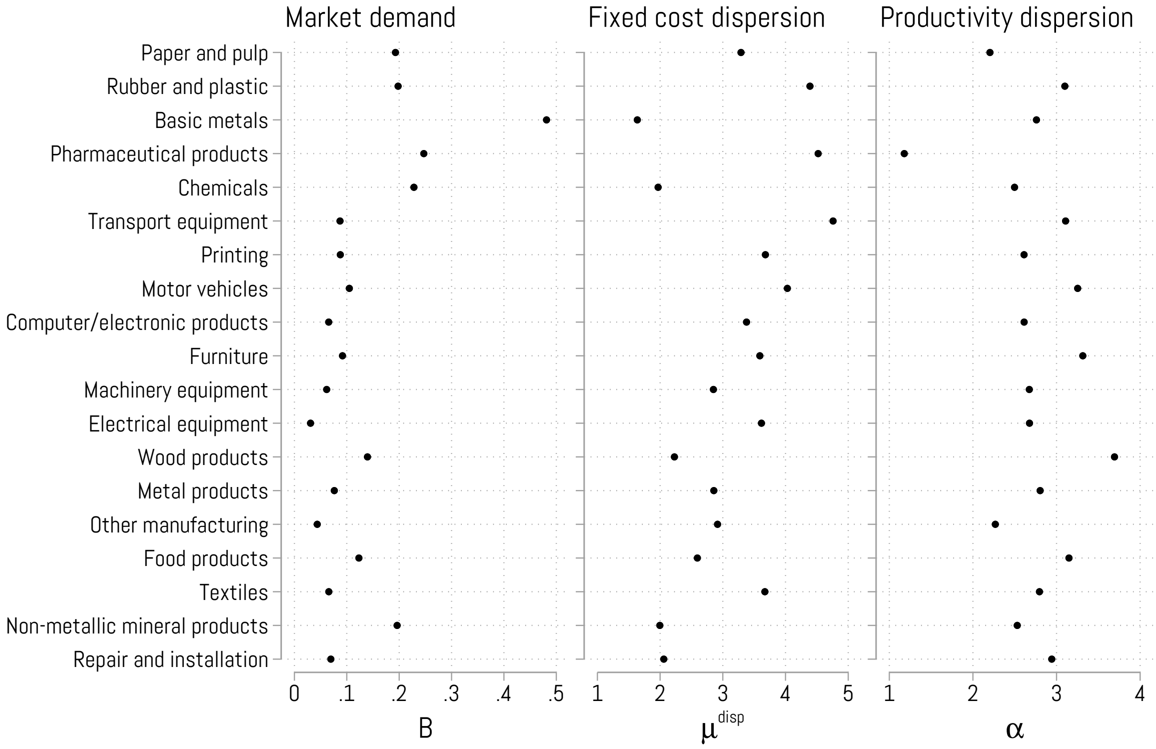

Figure 10: Market demand, Fixed cost dispersion, and Pareto shape parameters

Notes: The figure shows the simulated method of moments estimates of the market demand index, $\tilde{B}_k$, the fixed cost dispersion parameters, μ_k^{disp}, and the Pareto shape parameters, α_k, by 2-digit industries. The simulated method of moments estimator is based on 21,600 simulated firms within each 2-digit industry and 2004 data moments in our base sample.

Pareto shape parameter based on US data. They find that it is equal to 5.14 on average across industries, however, both the estimation and the interpretation of the estimates differ from each other. They estimate α_k by regressing the log sales rank of firms on the firms' log sales, exploiting the fact that the domestic sales distribution is Pareto with shape parameter $\alpha_k/(1-\sigma_k)$ in a standard Melitz (2003)-model. In our model, this fact is no longer true, since the sales distribution also depends on the firms' selection of fuel sets (Di Giovanni et al., 2011). To come around this, we instead estimate α_k internally via the simulated method of moments. The interpretation is also different since we estimate α_k such that it is consistent with the distribution of fuel consumption in the data and not the sales distribution. If we instead targeted the sales distribution, our estimates of α_k would be even lower, and the model would overestimate the variance in fuel consumption across firms.

Figure 11 plots the estimates of the fixed cost scale parameters, μ_{jk}. These estimates represent the fuels' log median fixed cost within industries. Panel A shows the relationship between these estimates and the share of firms that use each fuel in a given industry. The relationship is negative. Fewer firms use fuels associated with high investment costs. Panel B shows the relationship between the median fixed costs and the fuels' production potentials. The correlation is slightly positive, suggesting that fuels with high production potentials tend to be more costly to invest in. In terms of magnitude, the estimates suggest that the median fixed cost of investing in fuel-specific technology ranges between

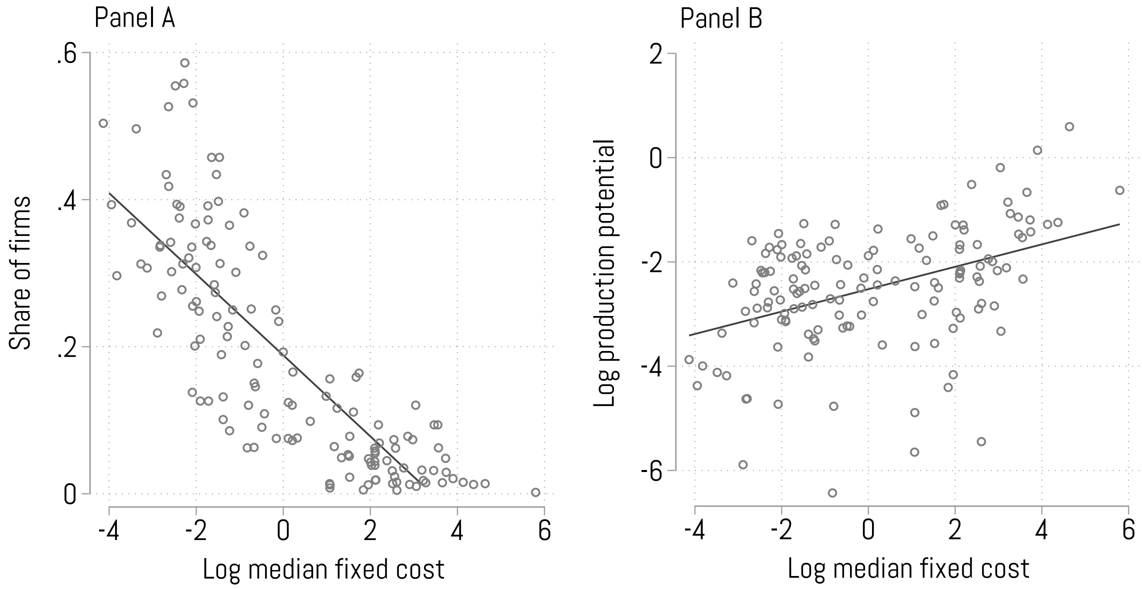

Figure 11: Log median fixed cost, share of firms, and log production potential

Notes: The figure plots the simulated method of moments estimates of the fuel-specific fixed cost scale parameters, μ_{jk}^{disp}, against the share of firms that use each fuel in the data and the fuels' log production potentials. The observations are at the industry-fuel level. Fixed costs are expressed in millions SEK. The simulated method of moments estimator is based on 21,600 simulated firms within each 2-digit industry and 2004 data moments in our base sample.

16,000 SEK and 329 million SEK with an average of 9.8 million SEK.[32]

5.5 Model fit

We now evaluate how well the estimated model can replicate both targeted and untargeted moments in the actual data. We start with the moments we target in the simulated method of moments estimator. Panel A in Figure 12 shows that the model fits the share of firms that use each fuel almost perfectly. The correlation between the simulated shares and the actual shares is 0.994. The feature of the model that generates this fit is the fuel-specific fixed costs. Figure 13 shows that the estimated model manages to capture the other targeted moments as well. The figure displays the average of each moment across industries in both the simulated and the actual data. In the upper-left graph, we see that the introduction of firm-specific fixed costs successfully breaks the monotonic relationship between sales and production capability originally predicted by the model. The upper-right graph shows that the estimation also generates a good fit with respect to the median spending on electricity and total energy. Finally, in the lower-left graph, we see that the estimated model is able to replicate the standard deviation in energy

[32]These magnitudes should be interpreted with caution since we do not target the sales distribution in our estimation.

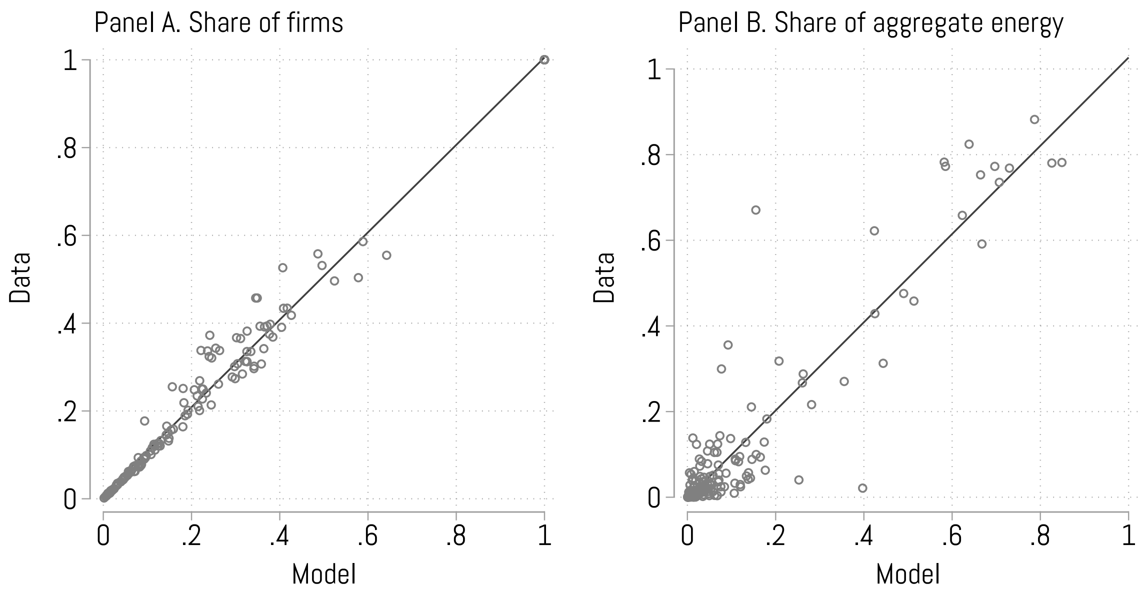

Figure 12: Fit between model and data

Notes: Panel A plots the share of firms that use each fuel in the simulated data and the actual data in 2004. Panel B plots the share of energy that is consumed from each fuel in the simulated data and the actual data. The observations are at the industry-fuel level. The simulated method of moments estimator is based on 21,600 simulated firms within each 2-digit industry and 2004 data moments in our base sample.

spending across firms.

We next turn to three sets of moments that we do not target in the structural estimation. The first is the share of aggregate energy that is consumed from each fuel. The fit of these moments is displayed in Panel B of Figure 12. The fit is reasonably good with a correlation of 0.928 between the simulated shares and the actual shares. The second set of moments is the average production capability in each quarter of the sales distribution. The lower-right graph in Figure 13 shows that the model fits these moments well and captures the pattern in the data that larger firms tend to choose better fuel sets. Finally, we check how the model matches median CO2 emissions across industries. On average, the model prediction is only 6% below the actual median. Together, these checks suggest that our estimated model does a good job of describing the pattern of fuel consumption in the data.

6 Counterfactual analysis

We use the estimated model to analyze the general equilibrium effects of higher carbon prices on fuel switching and how these effects matter for the aggregate effects on CO2 emissions and output. The model is estimated based on 2004 data and therefore describes

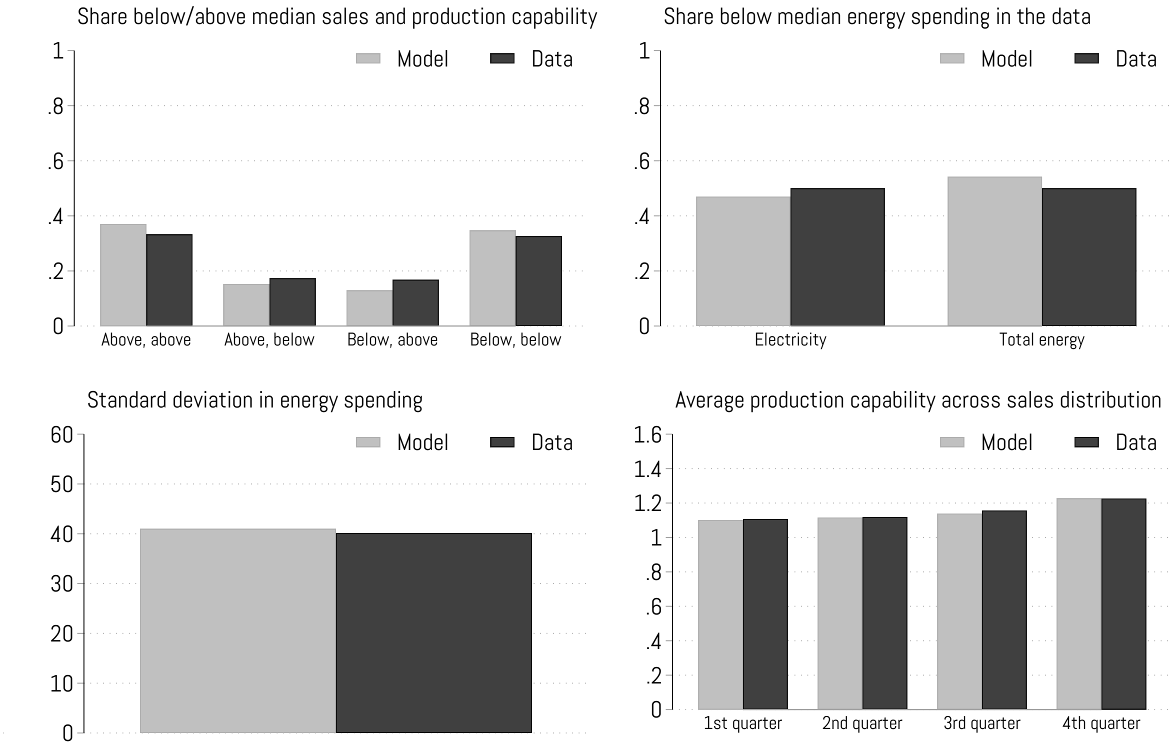

Figure 13: Fit between model and data

Notes: The upper-left graph plots the average shares of firms above/below median sales and median production capability across industries in the simulated data and the actual data in 2004. The upper-right graph plots the average shares of firms below median electricity and median total firm energy expenditures in the data across industries in the simulated data and the actual data in 2004. The lower-left graph shows the average standard deviation in total firm energy expenditures across industries in the simulated data and the actual data in 2004. The lower-right graph shows the average mean production capability within the 1st, 2nd, 3rd, and 4th quarters of the sales distribution across industries in the simulated data and the actual data in 2004.

fuel consumption in the Swedish manufacturing sector conditional on the carbon tax in 2004. In the counterfactual analysis, we consider 10 increases in the level of the Swedish carbon tax. The total tax increase equals the actual change in the Swedish carbon tax between 2004 and 2020. Parameters not affected by the tax are kept constant at their baseline levels. The experiment thus evaluates how fossil CO2 emissions would have changed in 2004 if the carbon tax increased to the level of 2020. The actual evolution of emissions between 2004 and 2020 displayed in Figure 3 is likely to have been affected by other factors such as technological change and fuel prices. Another important difference is that a large part of actual CO2 emissions were covered by the EU ETS from 2005 and onwards, while all emissions are covered by the CO2 tax in the counterfactual simulations. Despite this, the actual evolution of emissions still serves as a meaningful point of reference when evaluating the counterfactual predictions.

6.1 Computing baseline and counterfactual general equilibria

The estimates in Section 5 allow us to solve for the remaining parameters and aggregates characterizing the baseline equilibrium in the simulated model. We use the free-entry condition in (26) to derive the fixed entry cost, f_{ek}, in each industry. We then use equation (27), repeated here for convenience, to solve for the mass of firms:

$$M_k^E = \frac{\beta_k\Big(w\bar{L} + \sum_k \sum_{j\in J} \tau_j X_j^k\Big)}{\sigma_k w\Big[f_{ek} + \int_{\tilde{\varphi}_k}^{\infty} \sum_{j\in\mathcal{J}(\varphi)} f_{jk} dF_k(\varphi)\Big]}. \tag{38}$$

where we use the sum of revenues within each 2-digit industry in 2004 in our base sample to compute the numerator and the Cobb-Douglas expenditure shares β_k.[33] With M_k^E in hand, we can compute aggregate outcomes in the baseline equilibrium, including aggregate tax incomes, $\sum_k \sum_{j\in J} \tau_j X_j^k$, which we use to solve for the constant part, $w\bar{L}$, in aggregate income.

The counterfactual equilibria are computed in six steps. First, we update the production potential for each fuel that is affected by the tax. Denoting the baseline production potential of fuel j by ξ_j, the counterfactual production potential is given by,

$$\xi_j^c = \left(\frac{r_j + \tau_j + \Delta\tau_j}{r_j + \tau_j}\right)^{-\theta_k} \xi_j. \tag{39}$$

Second, we numerically solve for a new demand index, $\tilde{B}_k$, such that the free-entry condition is satisfied. Third, we solve the firms' problem in each industry given the new production potentials and market demand indices. Fourth, we update the mass of firms in each industry using (38). Fifth, we compute aggregate tax income across industries. Finally, we iterate steps four and five until the vector $\{M_k^E\}_k$ converges to a fixed point.

6.2 Counterfactual effects

Figure 14 shows the counterfactual effects of increasing the carbon tax on aggregate emissions and output.[34] The effect on emissions is decomposed into an energy scale effect, an energy composition effect, a fuel switching effect, and a firm entry and exit effect.[35] Panel A documents that sector-wide emissions decline by 56% when we increase the carbon tax to the 2020 level. About 30% of this decline is accounted for by a decrease in aggregate energy use, about 20% is accounted for by a reallocation of energy use across firms, and about 50% is accounted for by within-firm fuel switching. Firm entry and exit do not play an important role. The effect on aggregate output serves as a measure of the cost of increasing the tax. In the manufacturing sector as a whole, aggregate output decreases by about 15% when we increase the tax to the 2020 level. We do not attempt to weigh the

[33]Figure E1 in Appendix E displays our estimates of the fixed entry costs and consumer expenditure shares by industry.

[34]Just as with the actual carbon tax, the fossil fuels peat, coke oven gas, and blast furnace gas are exempted from the counterfactual tax increase.

[35]We use the following expression to decompose aggregate industry emissions in the model:

$$\mathrm{CO2}_k = M_k^E X_k\Big(\sum_i x_i e_i \delta_i\Big), \tag{40}$$

where X_k is industry energy use (scale), x_i is the share of energy used by firm i (composition), e_i is fossil emissions over energy use (fuel switching), and M_k^E is the mass of firms (entry/exit). δ_i is the weight assigned to firm i in the sampling of the simulated data.

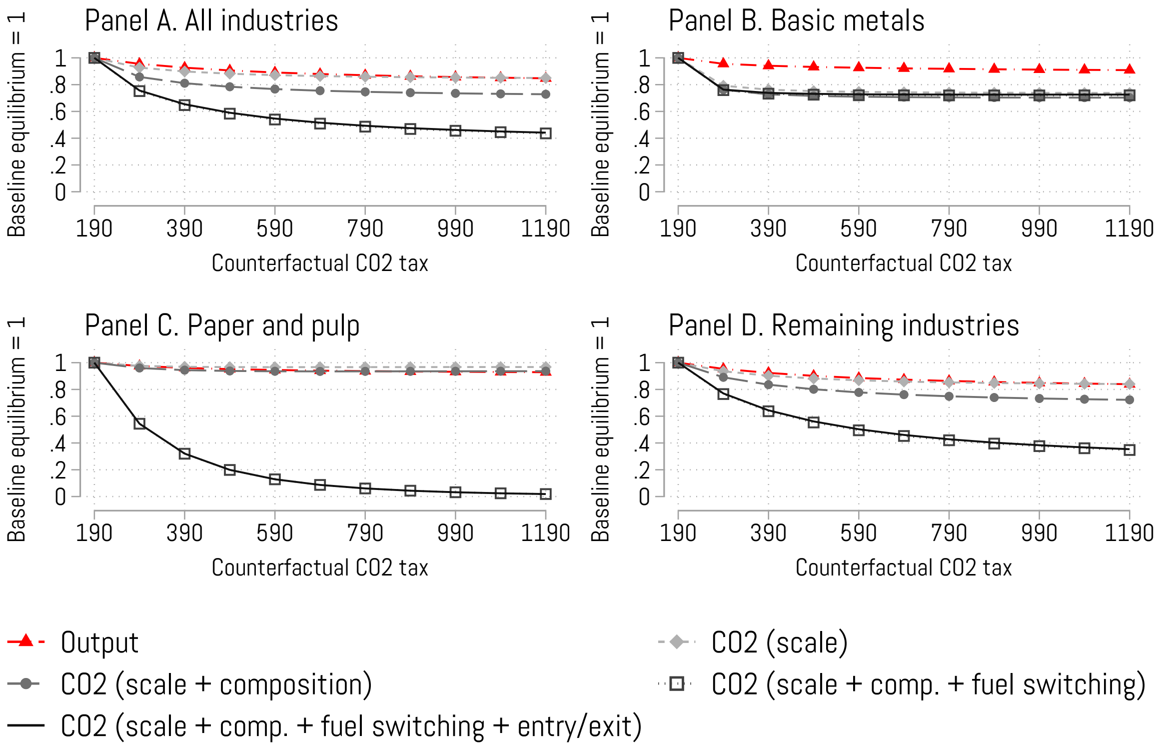

Figure 14: Decomposition of the counterfactual effects on fossil CO2 emissions

Notes: The figure shows the general equilibrium effects of successively increasing the carbon tax (SEK / tonne CO2) to the 2020 level relative to the baseline equilibrium. The decomposition of the effect on CO2 is based on equation (40). The scale effect represents the change in CO2 if we let aggregate energy use, X_k, change. The scale + composition effect represents the change in CO2 if we let aggregate energy use, X_k, and the firms' energy shares, x_i, change. The scale + comp. + fuel switching effect represents the change in CO2 if we let aggregate energy use, X_k, the firms' energy shares, x_i, and the firms' fossil CO2 per energy, e_i, change. Finally, the scale + comp. + fuel switching effect + entry/exit effect represents the change in CO2 if we let all variables change.

benefit of lower emissions against the cost of lower output. Such normative analysis falls outside the scope of this paper and would require that we specify how aggregate CO2 emissions enter the utility function. However, we use the effect on output to compare the relative cost of the tax across industries and model specifications.

The predicted change in sector-wide emissions resembles the actual change in sector-wide emissions between 2004 and 2020 in Figure 3. However, there are some differences worth highlighting. First, the predicted decline in emissions is larger than the actual decline of 38%. This can be explained by the fact that a large part of actual emissions were covered by the EU ETS after 2004 and subject to a smaller increase in the price of carbon over the sample period. Second, the importance of compositional changes is somewhat smaller, and the importance of fuel switching is somewhat larger, in the counterfactual decline of emissions compared to the actual decline. This does not necessarily mean that the model over-predicts the fuel-switching response of firms. It could also be explained by other time-varying factors that we do not consider in the counterfactual analysis, such as uneven productivity growth across clean and dirty firms.

Figure 14 also shows the decomposition of the counterfactual effects in the basic metals industry, the paper and pulp industry, and the remaining industries, separately. The pattern is strikingly similar to the pattern in actual emissions. The decline in emissions in the basic metals industry is almost exclusively explained by a reduction in aggregate energy while the decline in the paper and pulp industry is mostly due to fuel switching. The fact that the model is able to predict the pattern of emission reductions in these two industries supports the validity of the model. Moreover, in the basic metals industry, the counterfactual decline almost exactly matches the actual decline of 25%. Most of this effect occurs at low levels of the carbon tax. In the paper and pulp industry, the counterfactual decline is 98% while the actual decline is 75%. Again, this can be explained by the fact that the dirtiest firms in this industry are covered by the EU ETS and did not see the price of carbon exceed 500 SEK/tonne during the sample period. The analysis predicts that increasing the carbon price above this level in the paper and pulp industry is effective in reducing emissions further. Perhaps surprisingly, the cost of the tax increase in terms of output is only slightly higher in the basic metals industry (9% decline) relative to the paper and pulp industry (7% decline). The model sheds light on why this is the case. Even though fuel switching does not have an effect on aggregate fossil fuel emissions in the basic metals industry, firms do adapt. In particular, they increase their relative use of coke oven gas and blast furnace gas, two fossil fuels with high carbon contents that are exempted from the Swedish carbon tax. These two fuels also have high production potentials which enables the firms to keep down marginal cost. Panel D shows that the loss in output in the remaining industries is higher (16%). The counterfactual decline in emissions is 65% in the remaining industries while it is 50% in the actual data.

To understand what drives our results, we can unpack the aggregate effects on output and emissions further. In the upper-left graph of Figure 15, we display the average growth in output across the productivity distribution when the carbon tax increases to the 2020 level. The graph shows that there is a substantial degree of reallocation of output behind the aggregate decline output. Low productivity firms benefit from the tax hike and increase their output while firms above the 80th percentile reduce their output. Firms in the 99th percentile are the biggest losers and shrink their output by 16% on average. The reason why more productive firms are hit harder is that they tend to use more fuels in general, and more fossil fuels in particular, compared to less productive firms. It may seem puzzling that a large share of firms actually increase their output as a response to the carbon tax. This is a general equilibrium effect and is due to an increase in the industry price indices. In partial equilibrium, holding constant the market demand indices B_k and the mass of firms M_k, all percentiles experience a decline in output.

The rest of the graphs in Figure 15 unpack the components driving the decline in sector-wide emissions. The scale component captures the change in aggregate energy use. As displayed in the upper-right graph, the 15% decline in total energy is driven by a 68% decline in the demand for fossil fuels. The demand for electricity and biofuels increases by 13% and 14%, respectively. The composition component captures changes in the firms' shares of aggregate energy. In the lower-left graph, we display the average change in these shares across the productivity distribution. Aggregate emissions decline via this channel due to reallocation of energy towards less productive firms. The intuition is that these firms rely less on fossil fuels. Finally, in the lower-right graph, we investigate whether the fuel-switching response differs across the productivity distribution. The results show that the response is similar across firms with different productivity levels but that high-productivity firms are slightly less responsive in proportional terms. The reason

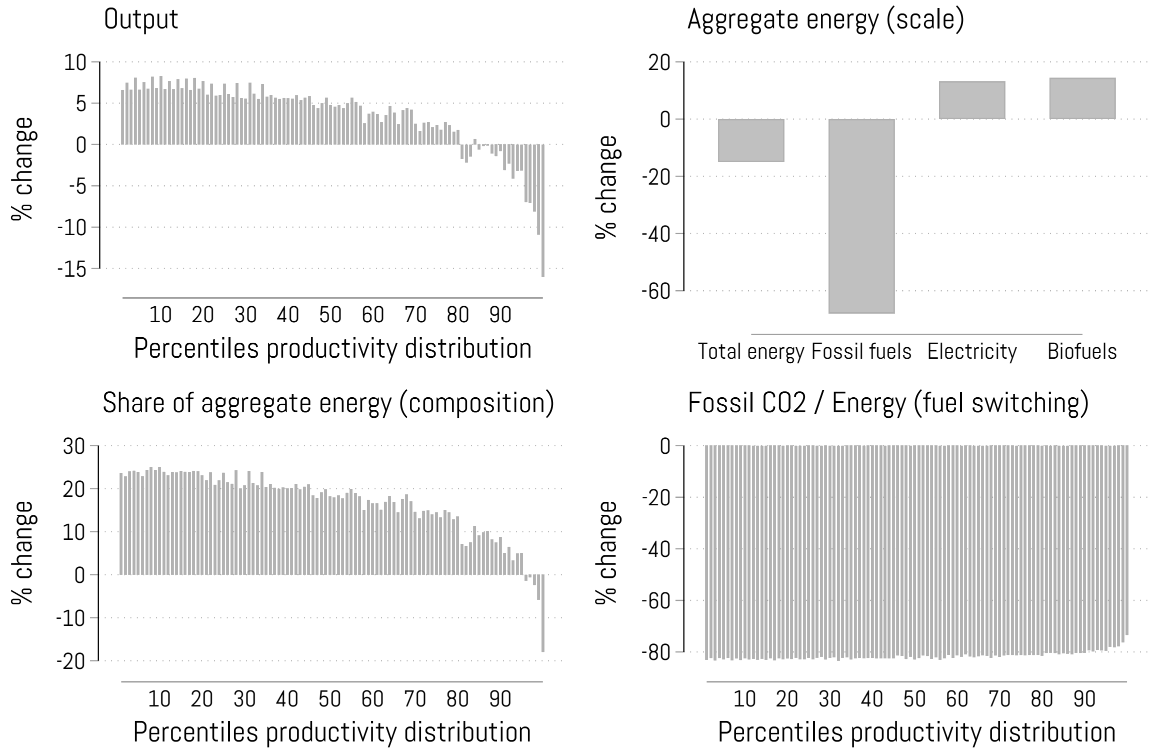

Figure 15: Unpacking the aggregate effects

Notes: The figure shows the general equilibrium effects of increasing the carbon tax to the 2020 level relative to the baseline equilibrium. The graphs show the average percentage change within each percentile of the productivity distribution for the entire manufacturing sector.

for this is that very large firms still find it profitable to bear the fixed costs of fossil fuels even after the tax hike.

As a validity check for the counterfactual fuel-switching response, we assess the average firm-level effects and compare them with our reduced-form evidence. The results are displayed in Figure 16. We express the average effects in differences between the counterfactual and baseline equilibria so that they are in the same units as the reduced-form estimates. Panel B shows that the predicted average effect on fossil CO2 per gigajoule is -7.14 when we increase the tax to the 2020 level. This is lower than the 2020 point estimate of -10.45 in the difference-in-differences analysis. The fuel switching response per unit increase in the average carbon price is slightly higher in the counterfactual analysis (-0.80) compared to the difference-in-differences analysis (-0.52).[36] However, the effects are not statistically different from each other which speaks to the empirical relevance of our counterfactual predictions. Moreover, for an increase in the carbon price similar to that in the difference-in-difference analysis, the difference in the fuel switching response would diminish since the counterfactual effect tapers off, suggesting that it becomes harder and harder to abate emissions.

Panel C in Figure 16 displays the effect on the share of electricity, biofuels, and fossil fuels in the firms' fuel mix. The effects are similar to the point estimates in the

[36]We calculate these effects as the ratio between the 2020 effects on fossil CO2 per gigajoule and the predicted CO2 price in Figure 16 and Figure 6, respectively.

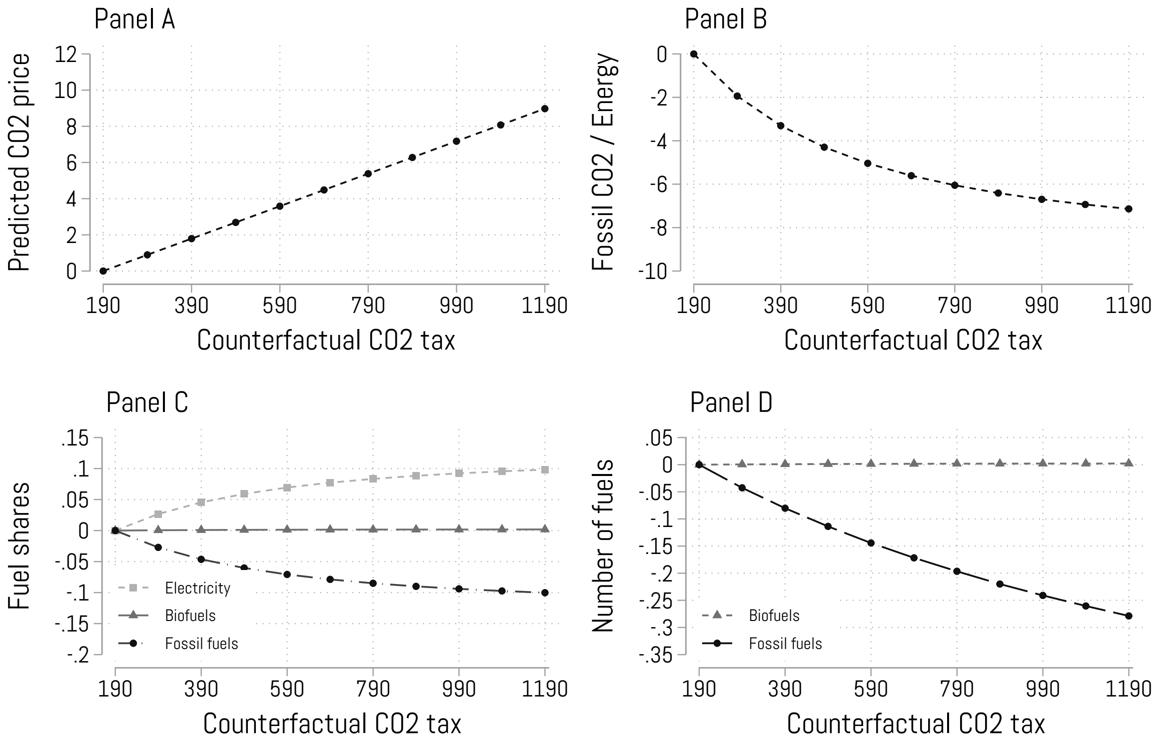

Figure 16: Counterfactual average firm-level effects

Notes: The figure shows the general equilibrium effects of successively increasing the carbon tax (SEK / tonne CO2) to the 2020 level relative to the baseline equilibrium. The firm-level averages are first computed at the industry level and then weighted by the mass of firms in each industry. The effects are expressed in differences between the counterfactual equilibrium and the baseline equilibrium.

difference-in-difference analysis. On average, firms respond to higher carbon taxes by replacing fossil fuels with electricity. Consistent with the reduced-form evidence in Table 3, the model also predicts that firms respond at the extensive margin by dropping fossil fuels.

6.3 Heterogeneous effects across industries

The counterfactual effects on emissions and output vary substantially across industries. To understand what drives this heterogeneity, Figure 17 shows the effects on output and CO2 emissions by industry under different model specifications. The upper-left graph shows the effects in our baseline model where heterogeneity may come from differences in any of the industry-specific parameters: the elasticity of substitution between fuels, the price elasticity of demand, the fuels' production potentials, the market demand index, and the parameters governing the fixed cost distribution and the productivity distribution.

The rest of the graphs in Figure 17 show how the effects are altered when we remove these sources of heterogeneity. The upper-right graph shows the effects when we impose that the elasticity of substitution and the elasticity of demand are common across industries. The lower-left graph shows the effects when we in addition impose that the fuels' production potentials are common across industries. Finally, the lower-right graph shows the effects when all parameters are common across industries and the only source

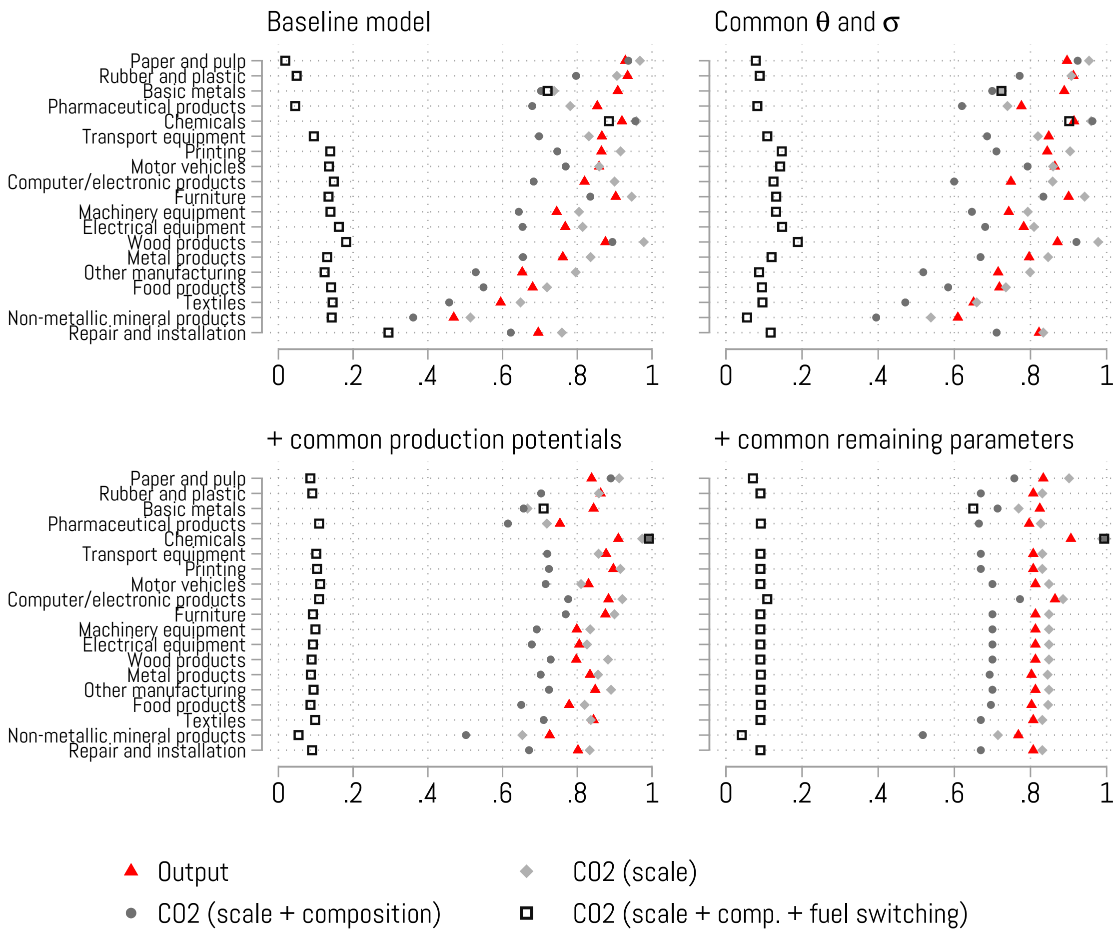

Figure 17: Effects of higher carbon tax under different model specifications by industry

Notes: The figure shows the general equilibrium effects of increasing the carbon tax to the 2020 level relative to the baseline equilibrium. Panel A shows the effects on aggregate output and fossil CO2 emissions by industry in the baseline model. Panel B shows the effects in a model where we set $1 + \theta_k$ and σ_k to the average values of these parameters across industries. Panel C shows the effects when we also set the fuels' production potentials to the average values across industries, but let the availability of fuels to differ across industries. Panel D shows the effects when all industry-specific parameters are set to the average values across industries, and the only source of heterogeneity is the availability of fuels.

of heterogeneity is the set of available fuels to choose from.

In the baseline model, all industries except basic metals, chemicals, and repair and installation reduce their emissions by over 80% when the carbon tax increases to the 2020 level. However, the decline in output associated with these emissions reductions varies significantly across industries. The two industries that are most adversely affected are the textiles industry and the non-metallic mineral products industry with over 40% reductions in output. We have sorted the industries by their elasticities of substitution between fuels. Inspection reveals that carbon pricing tends to be less effective and more costly in industries with lower elasticities of substitution.

When we remove differences in the elasticity of substitution, the effects on emissions and output become more similar across industries, though considerable heterogeneity persists. When we in addition remove differences in the fuels' production potentials across

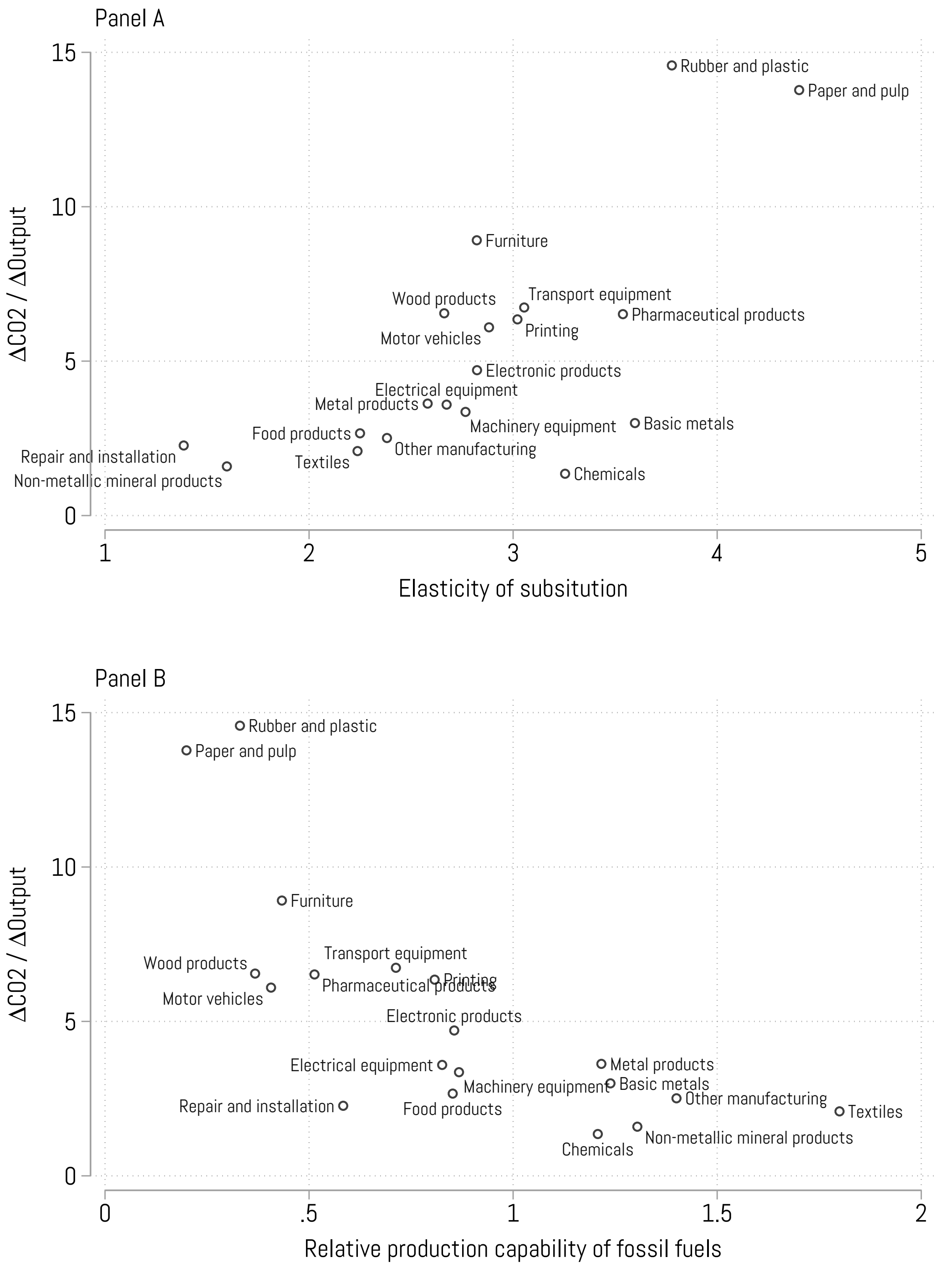

Figure 18: Determinants of the effectiveness of carbon pricing

Notes: The figure shows the general equilibrium effects of increasing the carbon tax to the 2020 level relative to the baseline equilibrium. On the y-axis, we have the percentage change in fossil CO2 over the percentage change in output. On the x-axis in Panel A, we have the elasticity of substitution between fuels. On the x-axis in Panel B, we have the sum of production potentials among fossil fuels over the sum of production potentials among non-fossil fuels.

industries, the effects on emissions and output converge considerably, with only modest heterogeneity remaining. Finally, in the model where only the sets of available fuels differ across industries, the effects on emissions and output are almost identical across industries. However, carbon pricing remains less effective in the basic metals industry and the chemicals industry which can be explained by the availability of non-taxed fossil fuels in these industries.[37]

Figure 17 shows that the main factors explaining heterogeneous responses across industries are the elasticities of substitution between fuels and the fuels' production potentials. In Figure 18, we relate these two factors to the efficacy of carbon pricing, which we measure as the percentage decline in industry emissions per percentage decline in industry output. Panel A shows that the efficacy of carbon pricing increases in the elasticity of substitution while Panel B shows that the efficacy decreases in the relative efficiency of fossil fuels. The efficacy is high in industries such as the paper and pulp industry where it is easy to substitute between fuels and where efficient alternatives to fossil fuels exist. The efficacy is significantly lower in industries such as the basic metals, cement, and chemicals industries where either fuel substitution is difficult or renewable energy sources are inefficient.

7 Conclusion

The effects of carbon pricing on emissions and output depend on firms' abilities to replace fossil fuels in production and on market share reallocation toward firms that use cleaner fuels. We use detailed data on firm-level fuel use in the Swedish manufacturing sector and develop a structural model of firms' fuel choices to quantify these transmission channels.

We first show in a statistical decomposition that both firm-level fuel switching and reallocation toward firms that use cleaner fuels have been important for the decline in emissions in the Swedish manufacturing sector. We then show in a reduced-form analysis of the Swedish carbon tax, which has increased by more than 500% over the last two decades, that firms have switched away from fossil fuels in response to higher carbon prices. Finally, we estimate our model and quantify the equilibrium effects of carbon pricing on emissions and output, accounting for firm-level fuel adjustments and market share reallocation, while holding the economy's overall technology constant.

We find that higher carbon prices effectively reduce manufacturing emissions and that fuel switching explains about half of this effect. We also find that higher carbon prices reduce emissions by market share reallocation toward cleaner firms. However, this reallocation amplifies the adverse effect on output. More productive firms tend to rely more heavily on fossil fuels and therefore experience relatively larger increases in their marginal costs. This result highlights that gains in environmental efficiency may come at a cost to economic efficiency. Finally, we find that the effects of higher carbon prices on output vary significantly across industries. This heterogeneity is driven by technological differences in the efficiency of non-fossil fuels and the overall substitutability between fuels. In industries such as paper and pulp, where firms can replace fossil fuels with efficient biofuels, output losses are limited. In industries such as non-metallic minerals, where the

[37]The fossil fuels coke oven gas, blast furnace, and peat are exempted from the tax. Figure 7 shows that the two former are available in the basic metals industry and the latter is available in the chemicals industry.

advantage of using fossil fuels is relatively large, output losses are substantial.

Our findings have policy relevance and point to the importance of tailoring carbon policies to the specific conditions of each industry. In industries where it is difficult to replace fossil fuels in production, carbon pricing should be implemented with care to minimize output losses and reduce the risk of carbon leakage. Clean technology subsidies, R&D investments, and carbon border adjustments are particularly valuable in such industries. More broadly, our results show that carbon pricing leads to a surge in the aggregate demand for electricity and biofuels. It may therefore be critical to accompany carbon pricing with public investments in renewable energy production to avoid capacity constraints and facilitate fuel switching.

References

D. Acemoglu, P. Aghion, L. Bursztyn, and D. Hemous. The environment and directed technical change. *American Economic Review*, 102(1):131–166, 2012.

P. Antras, T. C. Fort, and F. Tintelnot. The margins of global sourcing: Theory and evidence from US firms. *American Economic Review*, 107(9):2514–2564, 2017.

C. Arkolakis, F. Eckert, and R. Shi. Combinatorial discrete choice: A quantitative model of multinational location decisions. Technical report, National Bureau of Economic Research, 2023.

R. L. Axtell. Zipf Distribution of U.S. Firm Sizes. *Science*, 293(5536):1818–1820, 2001.

R. Bardazzi, F. Oropallo, and M. G. Pazienza. Do manufacturing firms react to energy prices? Evidence from Italy. *Energy Economics*, 49:168–181, 2015.

A. Bousquet and M. Ivaldi. An individual choice model of energy mix. *Resource and Energy Economics*, 20(3):263–286, 1998.

A. Bousquet and N. Ladoux. Flexible versus designated technologies and interfuel substitution. *Energy Economics*, 28(4):426–443, 2006.

C. Broda and D. E. Weinstein. Globalization and the gains from variety. *The Quarterly Journal of Economics*, 121(2):541–585, 2006.

J. Colmer, R. Martin, M. Muûls, and U. J. Wagner. Does pricing carbon mitigate climate change? Firm-level evidence from the European Union emissions trading scheme. Discussion Paper 1728, Centre for Economic Performance, 2023.

B. R. Copeland, J. S. Shapiro, and S. M. Taylor. Chapter 2 - Globalization and the environment. In G. Gopinath, E. Helpman, and K. Rogoff, editors, *Handbook of International Economics: International Trade*, volume 5, pages 61–146. Elsevier, 2022.

J. De Loecker. Product differentiation, Multiproduct Firms, and Estimating the Impact of Trade Liberalization on Productivity. *Econometrica*, 79(5):1407–1451, 2011.

A. Dechezleprêtre, C. Gennaioli, R. Martin, M. Muûls, and T. Stoerk. Searching for carbon leaks in multinational companies. *Journal of Environmental Economics and Management*, 112:102601, 2022.

A. Dechezleprêtre, D. Nachtigall, and F. Venmans. The joint impact of the European Union emissions trading system on carbon emissions and economic performance. *Journal of Environmental Economics and Management*, 118:102758, 2023.

J. Di Giovanni, A. A. Levchenko, and R. Ranciere. Power laws in firm size and openness to trade: Measurement and implications. *Journal of International Economics*, 85(1): 42–52, 2011.

A. K. Dixit and J. E. Stiglitz. Monopolistic competition and optimum product diversity. *American Economic Review*, 67(3):297–308, 1977.

J. Eaton and S. Kortum. Technology, geography, and trade. *Econometrica*, 70(5):1741–1779, 2002.

J. Eaton, S. Kortum, and F. Kramarz. An anatomy of international trade: Evidence from french firms. *Econometrica*, 79(5):1453–1498, 2011.

M. Fowlie, M. Reguant, and S. P. Ryan. Market-based emissions regulation and industry dynamics. *Journal of Political Economy*, 124(1):249–302, 2016.

M. A. Fuss. The demand for energy in canadian manufacturing: An example of the estimation of production structures with many inputs. *Journal of Econometrics*, 5(1): 89–116, 1977.

M. Golosov, J. Hassler, P. Krusell, and A. Tsyvinski. Optimal taxes on fossil fuel in general equilibrium. *Econometrica*, 82(1):41–88, 2014.

G. M. Grossman and A. B. Krueger. Environmental impacts of a North American Free Trade Agreement. Working Paper 3914, NBER, 1991.

H. Hammar and S. Åkerfeldt. CO2 Taxation in Sweden. 20 Years of Experience and Looking Ahead. Technical report, Sweden's Ministry of Finance, 2011.

M. Hyland and S. Haller. Firm-level Estimates of Fuel Substitution: An Application to Carbon Pricing. *The Energy Journal*, 39(6):71–98, 2018.

International Energy Agency. Industry. https://www.iea.org/energy-system/industry, 2023. [Online; accessed 23-May-2024].

P. Jia. What happens when Wal-Mart comes to town: An empirical analysis of the discount retailing industry. *Econometrica*, 76(6):1263–1316, 2008.

A. Jo. Substitution between Clean and Dirty Energy with Biased Technical Change. *Available at SSRN 4878270*, 2023.

E. M. Leclair. Balancing Production and Carbon Emissions with Fuel Substitution, 2024.

A. Levinson. Technology, International Trade, and Pollution from US Manufacturing. *American Economic Review*, 99(5):2177–92, 2009.

G. Martinsson, L. Sajtos, P. Strömberg, and C. Thomann. The Effect of Carbon Pricing on Firm Emissions: Evidence from the Swedish CO2 Tax. *The Review of Financial Studies*, 2024.

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

M. J. Melitz and S. J. Redding. Heterogeneous firms and trade. *Handbook of International Economics*, 4:1–54, 2014.

M. J. Melitz and S. J. Redding. New trade models, new welfare implications. *American Economic Review*, 105(3):1105–1146, 2015.

C. Papageorgiou, M. Saam, and P. Schulte. Substitution between clean and dirty energy inputs: A macroeconomic perspective. *The Review of Economics and Statistics*, 99(2): 281–290, 2017.

R. S. Pindyck. Interfuel substitution and the industrial demand for energy: an international comparison. *The Review of Economics and Statistics*, pages 169–179, 1979.

E. Saez. Do taxpayers bunch at kink points? *American Economic Journal: Economic Policy*, 2(3):180–212, 2010.

A. Serletis, G. R. Timilsina, and O. Vasetsky. Interfuel substitution in the United States. *Energy Economics*, 32(3):737–745, 2010a.

A. Serletis, G. R. Timilsina, and O. Vasetsky. International evidence on sectoral interfuel substitution. *The Energy Journal*, 31(4):1–30, 2010b.

J. S. Shapiro and R. Walker. Why is Pollution from US Manufacturing Declining? The Roles of Environmental Regulation, Productivity, and Trade. *American Economic Review*, 108(12):3814–54, 2018.

J. Steinbuks. Interfuel substitution and energy use in the UK manufacturing sector. *The Energy Journal*, 33(1):1–30, 2012.

F. Tintelnot. Global production with export platforms. *The Quarterly Journal of Economics*, 132(1):157–209, 2017.

A. D. Woodland. A micro-econometric analysis of the industrial demand for energy in NSW. *The Energy Journal*, 14(2):57–89, 1993.

A Appendix: Data

A.1 Additional tables and figures

Table A1: Firm size and fuel-switching

Dependent variable: log sales	(1)
Added fuel	0.26***
	(0.02)
Dropped fuel	-0.20***
	(0.02)
Lagged number of fuels	0.41***
	(0.02)
Observations	90,666

Notes: The sample contains 90,666 firm-year observations over the years 2004-2020. All regressions include 2-digit industry fixed effects and year fixed effects. Standard errors are clustered by firms.

B Appendix: Empirical analysis

B.1 Derivation of statistical decomposition

Assume a fixed set of plants. Aggregate emissions is then equal to,

$$E = \sum_i E_i = Z \sum_i \frac{Z_i}{Z} \frac{E_i}{Z_i} = Z \sum_i z_i e_i,$$

where E is aggregate emissions, E_i is emissions from plant i, Z is aggregate energy and Z_i is energy from plant i. Taking logs of this expression gives,

$$\log E = \log Z + \log \left[\sum_i z_i e_i \right].$$

The partial derivatives with respect to the arguments (Z, z_i, e_i) are,

$$\begin{aligned} \frac{\partial \log E}{\partial Z} &= \frac{1}{Z} \\ \frac{\partial \log E}{\partial z_i} &= \frac{e_i}{\sum_i z_i e_i} \\ \frac{\partial \log E}{\partial e_i} &= \frac{z_i}{\sum_i z_i e_i}. \end{aligned}$$

Total differentiation with respect to the arguments (Z, z_i, e_i) yields,

$$d \log E = \frac{\partial \log E}{\partial Z} dZ + \sum_i \frac{\partial \log E}{\partial z_i} dz_i + \sum_i \frac{\partial \log E}{\partial e_i} de_i$$

Substituting in the partial derivatives gives,

$$d \log E = \frac{dZ}{Z} + \sum_i \frac{e_i}{\sum_i z_i e_i} dz_i + \sum_i \frac{z_i}{\sum_i z_i e_i} de_i.$$

Rearranging gives the decomposition in (2):

$$\begin{aligned} d \log E &= \frac{dZ}{Z} + \sum_i \frac{\frac{E_i}{Z_i}}{\frac{E}{Z}} dz_i + \sum_i \frac{\frac{Z_i}{Z}}{\frac{E}{Z}} de_i \\ &= \frac{dZ}{Z} + \sum_i \frac{E_i}{E} \frac{dz_i}{z_i} + \sum_i \frac{Z_i}{E} de_i \\ &= \frac{dZ}{Z} + \sum_i \frac{E_i}{E} \frac{dz_i}{z_i} + \sum_i \frac{Z_i}{E} \frac{E_i}{Z_i} \frac{de_i}{e_i} \\ &= \frac{dZ}{Z} + \sum_i \frac{E_i}{E} \frac{dz_i}{z_i} + \sum_i \frac{E_i}{E} \frac{de_i}{e_i}. \end{aligned}$$

B.2 Statistical decomposition: additional figures

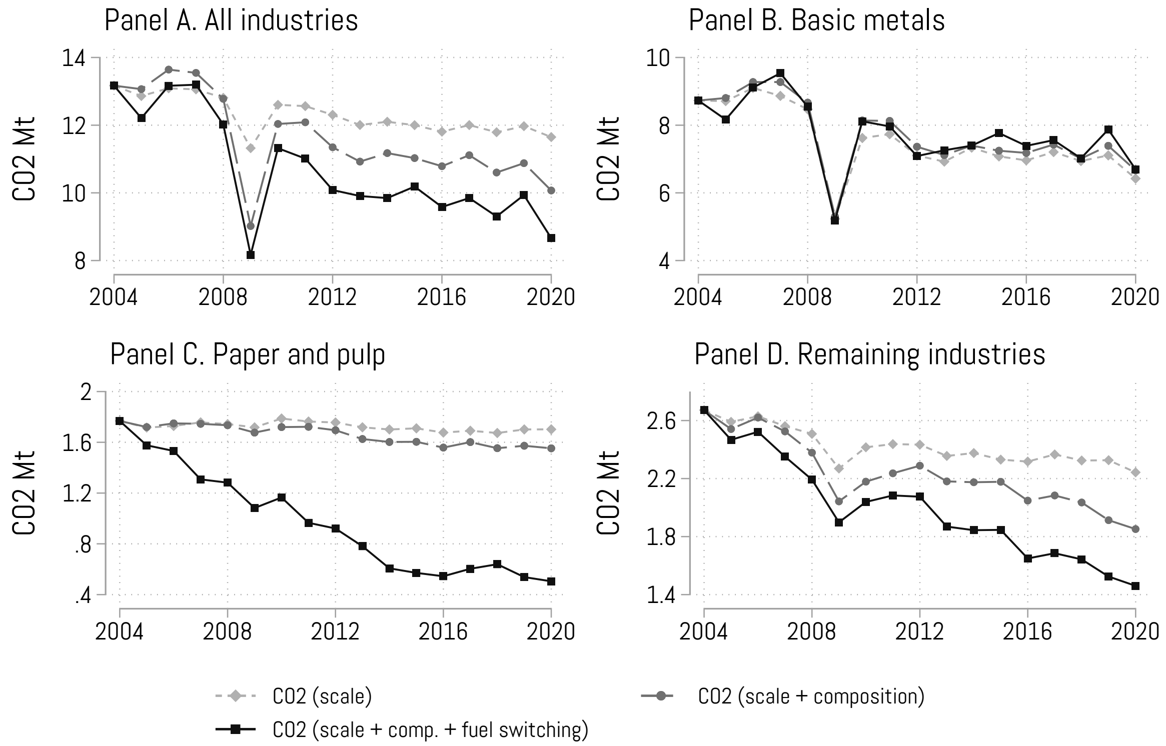

Figure B1: Decomposition of aggregate fossil CO2 emissions

Notes: The decomposition is based on plant-year observations in our base sample of 90,666 firm-year observations over the years 2004-2020. The (scale), (scale + composition), and (scale + comp. + fuel switching) lines are based on balanced samples. The balanced sample in Panel A contain 39,236 plant-year observations. The balanced sample in Panel B contains 1,666 plant-year observations. The balanced sample in Panel C contain 1,887 plant-year observations. The balanced sample in Panel D contain 35,683 plant-year observations. The balanced sample of plants in Panel A covers 92% of total fossil CO2 emissions between 2004 and 2020 in the full sample of plants.

B.3 The Swedish carbon tax and fuel switching: additional results and robustness checks

B.3.1 Shift-share IV analysis

B.3.2 Matched difference-in-differences analysis

Table B1: Shift-share IV analysis: First stage and reduced form

	CO2 tax (1)	Fossil CO2 / Energy (2)	Fossil fuel share (3)	Electricity share (4)	Biofuel share (5)	# fossil fuels (6)
Predicted CO2 tax	-0.309***	-0.257***	-0.003***	0.002***	0.002***	-0.003**
	(0.022)	(0.021)	(0.001)	(0.001)	(0.000)	(0.001)
Fossil fuel price	-0.002	-0.011**	-0.000*	0.000	0.000**	-0.000
	(0.005)	(0.006)	(0.000)	(0.000)	(0.000)	(0.000)
Non-fossil fuel price	0.012	0.067***	0.001***	-0.002***	0.001	0.001
	(0.015)	(0.010)	(0.000)	(0.000)	(0.001)	(0.001)
CO2 mitigation patents	0.004***	0.001	-0.000	-0.000	0.000	0.000
	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)
Observations	48,715	48,715	48,715	48,715	48,715	48,715
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry-year FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The base sample is an unbalanced panel of 90,666 firm-year observations over the years 2004-2020. The estimation sample of the shift-share IV analysis is confined to firms that have not been part of the EU ETS nor faced any tax exemptions over the sample period. It is also confined to firms that use non-zero taxed fuels in the pre-sample year, implying that the instrument is non-zero for all observations in the estimation sample. Finally, the estimation sample drops the first (pre-sample) year for all observations. The inclusion of firm fixed effects also removes singletons from the estimation sample. Standard errors are clustered at the firm-level.

Table B2: Shift-share IV analysis: Full sample

	Fossil CO2 / Energy (1)	Fossil fuel share (2)	Electricity share (3)	Biofuel share (4)	# fossil fuels (5)
CO2 tax	-0.863***	-0.012***	0.008***	0.004***	-0.022***
	(0.090)	(0.001)	(0.001)	(0.001)	(0.004)
Fossil fuel price	-0.029***	-0.000***	0.000***	0.000***	-0.001***
	(0.009)	(0.000)	(0.000)	(0.000)	(0.000)
Non-fossil fuel price	0.104***	0.001***	-0.002***	0.001***	0.004***
	(0.010)	(0.000)	(0.000)	(0.000)	(0.001)
CO2 mitigation patents	0.002	0.000	-0.000***	-0.000*	0.000
	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)
Observations	77,095	77,095	77,095	77,095	77,095
First-stage F-statistic	326	326	326	326	326
Firm FE	Yes	Yes	Yes	Yes	Yes
Industry-year FE	Yes	Yes	Yes	Yes	Yes

Notes: The base sample is an unbalanced panel of 90,666 firm-year observations over the years 2004-2020. The estimation sample in this table drops the first (pre-sample) year for all observations. The inclusion of firm fixed effects also removes singletons from the estimation sample. Standard errors are clustered at the firm-level. The first stage F-statistic is the Kleibergen-Paap F-statistic.

Table B3: Shift-share IV analysis: Placebo effect on EU ETS firms post 2010

	Fossil CO2 / Energy (1)	Fossil fuel share (2)	Electricity share (3)	Biofuel share (4)	# fossil fuels (5)
Predicted CO2 tax	-0.047	-0.001	0.000	0.000	-0.000
	(0.047)	(0.001)	(0.001)	(0.001)	(0.007)
Fossil fuel price	0.016	-0.000	0.000	-0.000	-0.001
	(0.021)	(0.000)	(0.000)	(0.000)	(0.002)
Non-fossil fuel price	-0.024	0.0010	-0.000	0.001	0.004
	(0.027)	(0.000)	(0.000)	(0.001)	(0.004)
CO2 mitigation patents	0.001	0.000	-0.000	-0.000*	0.000
	(0.002)	(0.000)	(0.000)	(0.000)	(0.000)
Observations	867	867	867	867	867
Firm FE	Yes	Yes	Yes	Yes	Yes
Industry-year FE	Yes	Yes	Yes	Yes	Yes

Notes: The base sample is an unbalanced panel of 90,666 firm-year observations over the years 2004-2020. The estimation sample cover the sample period 2011-2020 and only contain firms covered by the EU ETS. The inclusion of firm fixed effects also removes singletons from the estimation sample. EU ETS firms were completely exempted from the national CO2 tax. The predicted CO2 tax is based on the firms fuel shares in 2010 and the fuels' national carbon tax per gigajoule between 2011 and 2020. Standard errors are clustered at the firm-level.

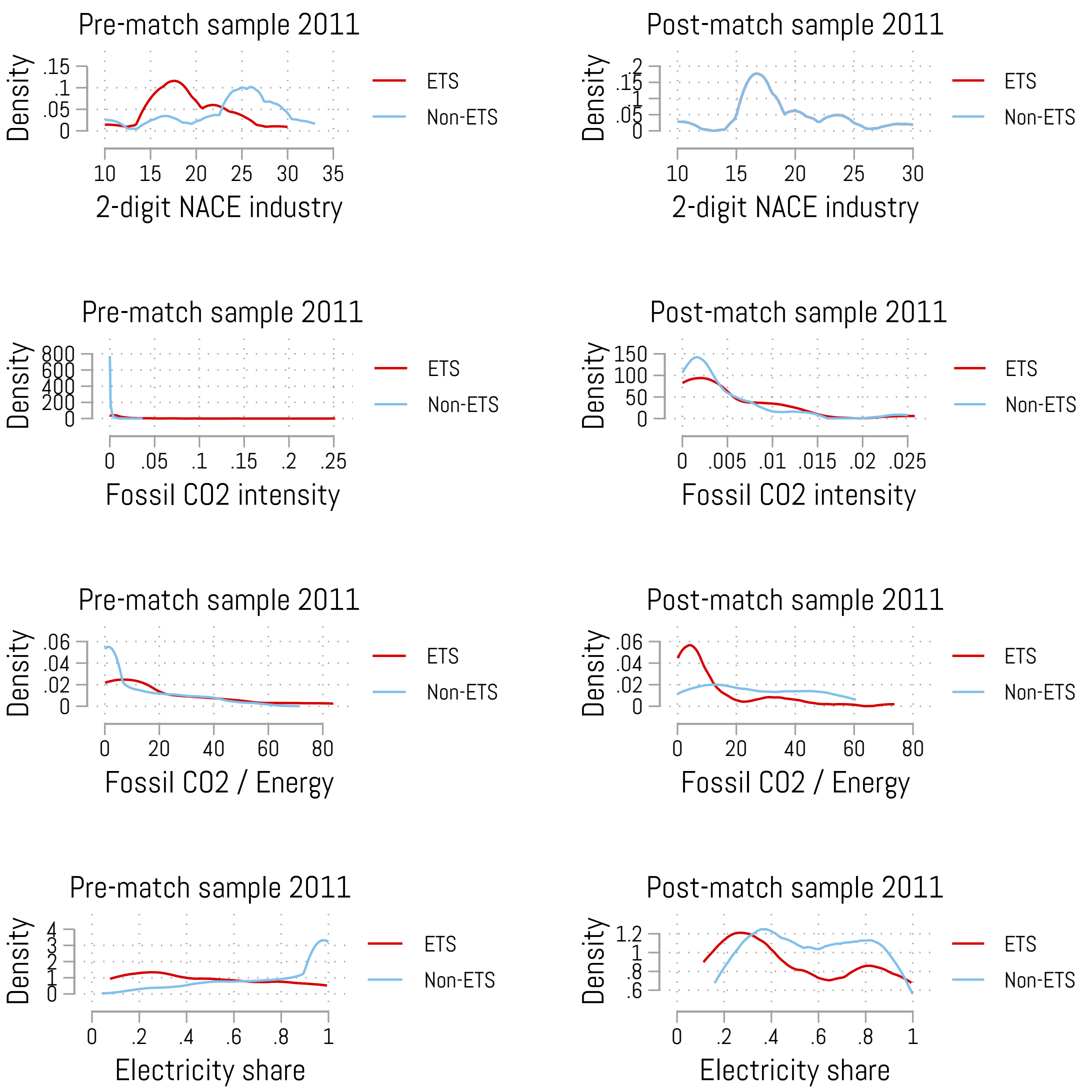

Figure B2: Matched difference-in-differences: Balance check before and after matching procedure

Notes: The pre-match sample in this figure contains a balanced panel of 27,060 observations over the years 2011-2020. The post-match sample in this figure is the matched difference-in-differences estimation sample containing 980 observations over the years 2011-2020. The figure shows the density of firms in 2011 for each variable. The 1:1 matching procedure matched EU ETS firms with non-EU ETS firms based on the firms' fossil CO2 intensity, defined as taxed CO2 over revenue, within each 2-digit NACE industry.

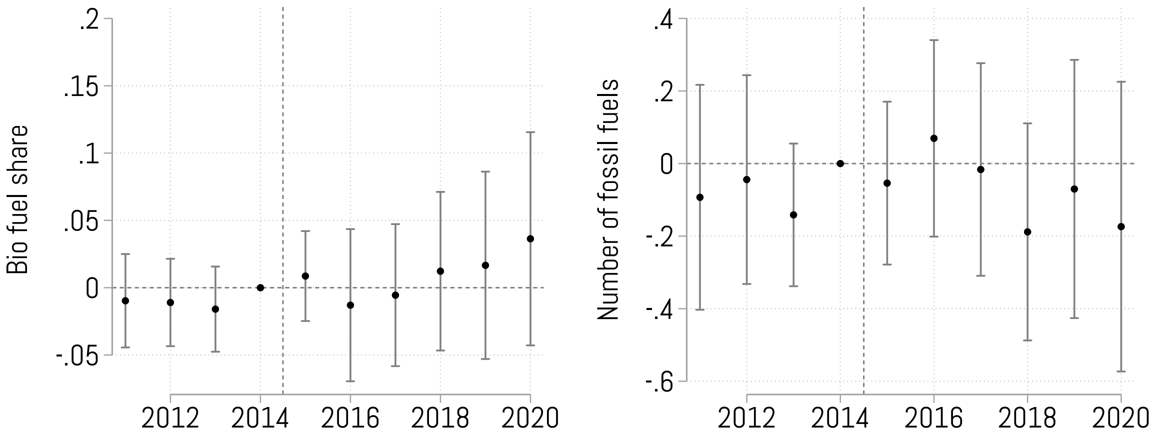

Figure B3: Matched difference-in-differences: Biofuel share and number of fossil fuels

Notes: The matched difference-in-differences estimation sample contains 980 firm-year observations over the years 2011-2020. The figure shows the point estimates and 95% confidence intervals when we run the regression in (5) on the firms' biofuel share and number of fossil fuels.

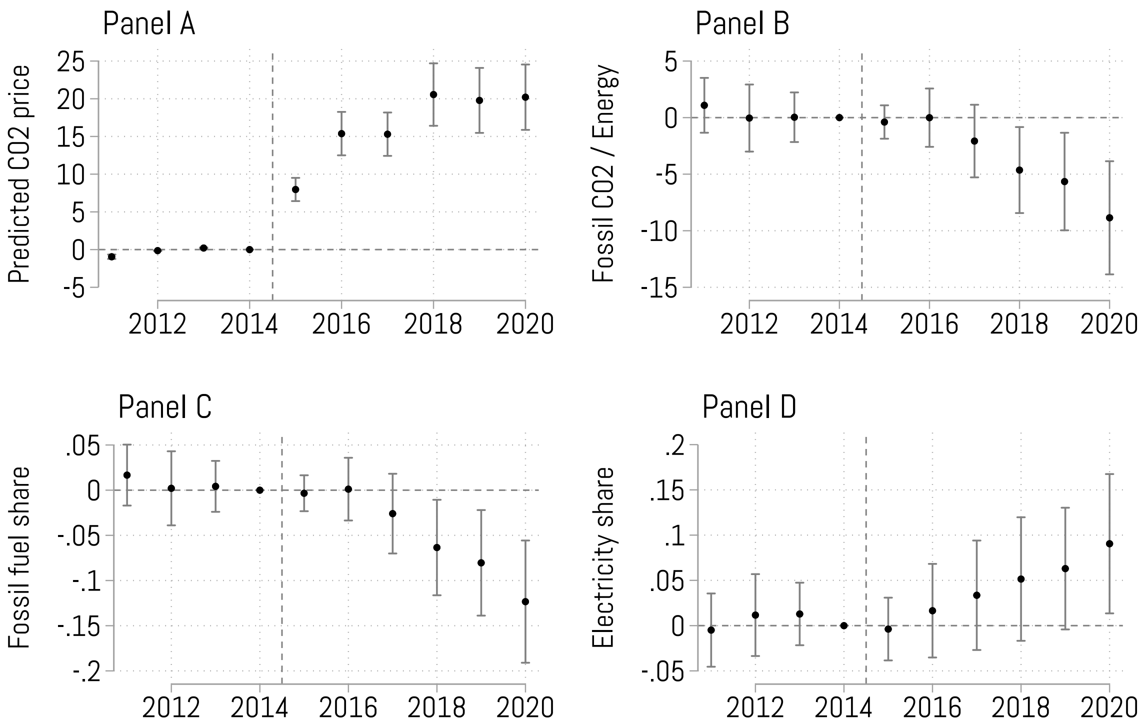

Figure B4: Matched difference-in-differences robustness check: No control variables

Notes: The matched difference-in-differences estimation sample contains 980 firm-year observations over the years 2011-2020. The figure shows the point estimates and 95% confidence intervals when we run the regression in (5) without controlling for fuel prices and our proxy for fuel-biased technical change. Standard errors are clustered by firms. Standard errors are clustered by firms.

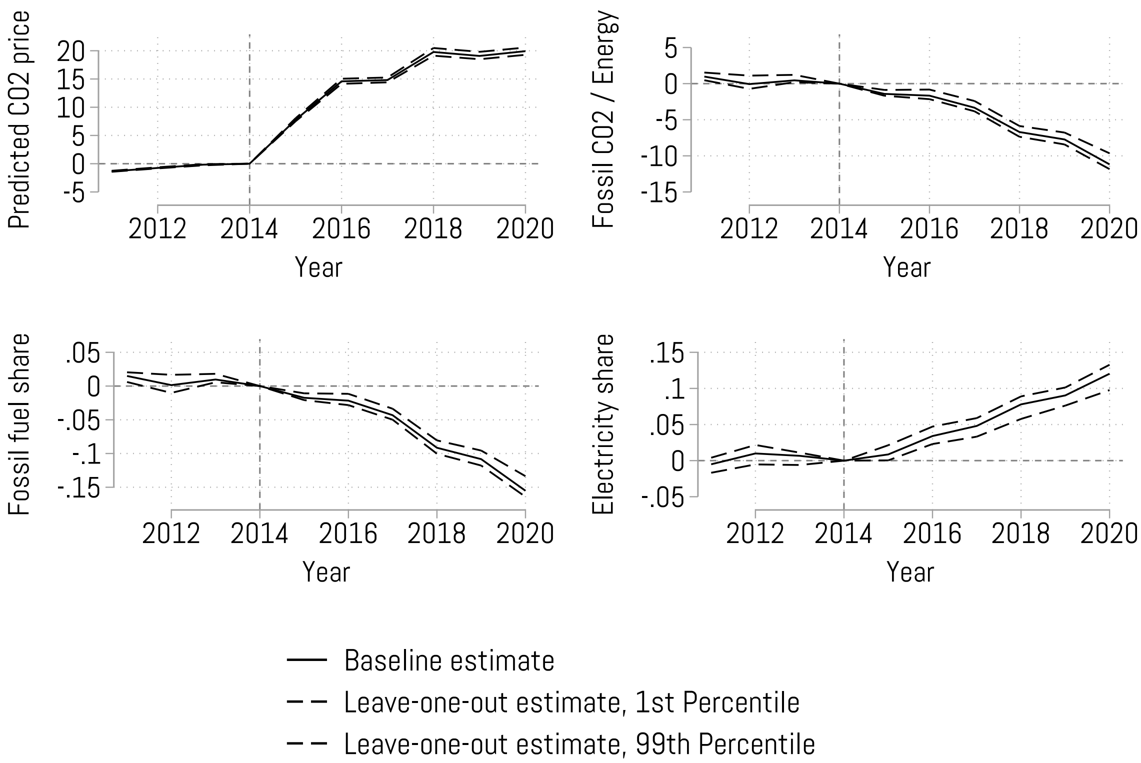

Figure B5: Matched difference-in-differences robustness check: Leave-one-match out from the estimation sample

Notes: The matched difference-in-differences estimation sample contains 980 firm-year observations over the years 2011-2020. The figure shows the baseline estimates from Figure 5 together with the 1st and 99th percentiles of the estimates in each year when we systematically leave one match out from the estimation sample. There are 49 1:1 matches between EU ETS firms and non-EU ETS firms so the 1st percentile and 99th percentile are effectively the minimum and maximum estimate in each year.

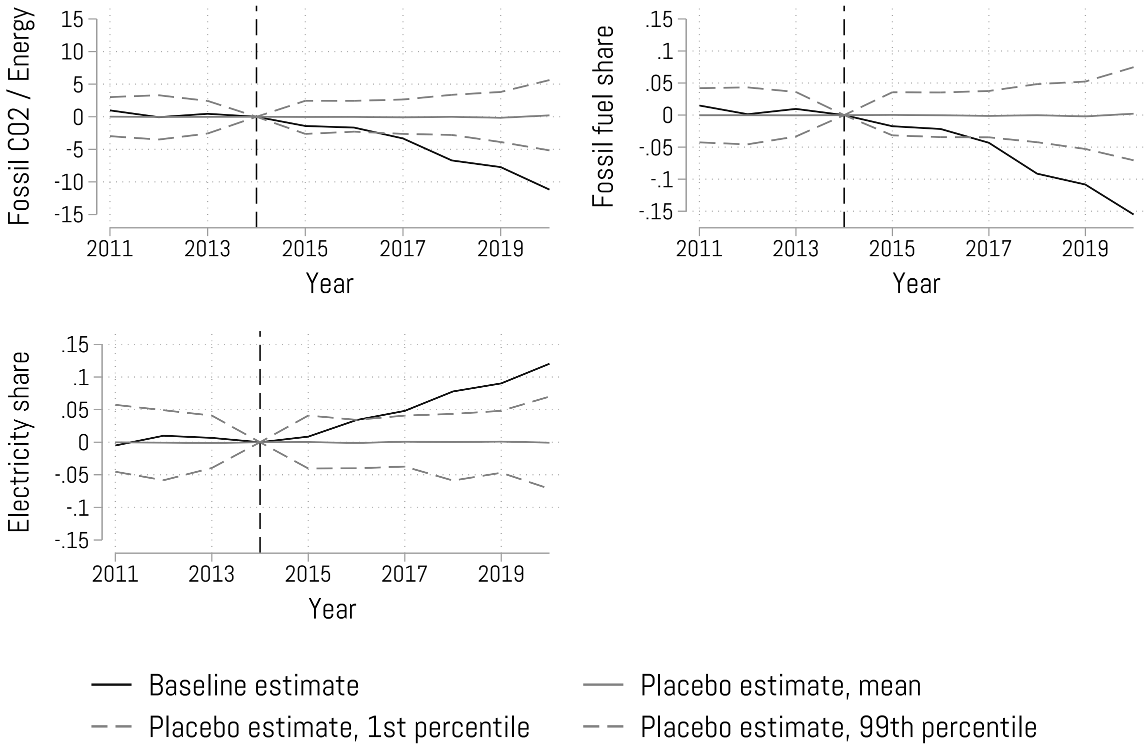

Figure B6: Matched difference-in-differences robustness check: Placebo estimates of randomized treatment

Notes: The matched difference-in-differences estimation sample contains 980 firm-year observations over the years 2011-2020. The figure shows the baseline estimates from Figure 5 together with placebo estimates when we randomize treatment status among firm-years in the estimation sample. We randomize treatment status 200 times and plot the mean, the 1st percentile, and the 99th percentile of the placebo estimates in each year.

C Appendix: Model

C.1 Preferences

C.1.1 Deriving the firm's demand function

The consumer problem is:

$$\begin{aligned} \max_{\{q_k(\omega)\}_{\Omega_k}} \quad & Q_k = \left(\int_{\omega \in \Omega_k} q_k(\omega)^{\frac{\sigma_k - 1}{\sigma_k}} d\omega \right)^{\frac{\sigma_k}{\sigma_k - 1}}, \\ s.t. \quad & E_k = \int_{\omega \in \Omega_k} p_k(\omega) q_k(\omega) d\omega, \end{aligned}$$

The Lagrangian can be written as:

$$\mathcal{L} = \int_{\omega \in \Omega_k} q_k(\omega)^{\frac{\sigma_k - 1}{\sigma_k}} d\omega + \lambda \Big[E_k - \int_{\omega \in \Omega_k} p_k(\omega) q_k(\omega) d\omega \Big]$$

The FOC is:

$$\begin{aligned} \frac{\sigma_k - 1}{\sigma_k} q_k(\omega)^{-\frac{1}{\sigma_k}} &= \lambda p_k(\omega) \\ \Longleftrightarrow q_k(\omega) &= \left[\frac{\sigma_k}{\sigma_k - 1} \right]^{-\sigma_k} \lambda^{-\sigma_k} p(\omega)^{-\sigma_k} \end{aligned}$$

Divide with the FOC for $q_k(\omega')$ to obtain,

$$\begin{aligned} & \frac{q_k(\omega)}{q_k(\omega')} = \left[\frac{p_k(\omega)}{p_k(\omega')} \right]^{-\sigma_k} \\ \Longleftrightarrow \; & p_k(\omega) q_k(\omega) = p_k(\omega')^{\sigma_k} q_k(\omega') p_k(\omega)^{1-\sigma_k} \\ \Longleftrightarrow \; & \int_{\omega \in \Omega_k} p_k(\omega) q_k(\omega) d\omega = p_k(\omega')^{\sigma_k} q_k(\omega') \int_{\omega \in \Omega_k} p_k(\omega)^{1-\sigma_k} d\omega \\ \Longleftrightarrow \; & E_k = p_k(\omega')^{\sigma_k} q_k(\omega') \int_{\omega \in \Omega_k} p_k(\omega)^{1-\sigma_k} d\omega \end{aligned}$$

It follows that,

$$\begin{aligned} q_k(\omega) &= \frac{E_k p_k(\omega)^{-\sigma_k}}{\int_{\omega \in \Omega_k} p_k(\omega)^{1-\sigma_k} d\omega} \\ &= \frac{E_k p_k(\omega)^{-\sigma_k}}{P_k^{1-\sigma_k}} \end{aligned}$$

where $P_k = \left(\int_{\omega \in \Omega_k} p_k(\omega)^{1-\sigma_k} d\omega \right)^{\frac{1}{1-\sigma_k}}$.

C.2 The firm's problem

C.2.1 Deriving the unit cost distribution for each fuel

The distribution of unit costs across the firm's inputs $v \in [0,1]$ for fuel j is derived by substituting $z_j(v,\varphi) = (r_j + \tau_j)a_j(v,\varphi)$ into the distribution of $1/a_j(v,\varphi)$:

$$\begin{aligned}
\mathbb{P}\Big[\frac{1}{a_j(v,\varphi)} \leq a\Big] &= e^{-T_{jk}a^{-\theta_k}} \\
\iff \mathbb{P}\Big[\frac{r_j+\tau_j}{z_j(v,\varphi)} \leq a\Big] &= e^{-T_{jk}a^{-\theta_k}} \\
\iff \mathbb{P}\Big[\frac{r_j+\tau_j}{a} \leq z_j(v,\varphi)\Big] &= e^{-T_{jk}a^{-\theta_k}} \\
\iff \mathbb{P}\Big[z_j(v,\varphi) \leq \frac{r_j+\tau_j}{a}\Big] &= 1 - e^{-T_{jk}a^{-\theta_k}}
\end{aligned}$$

Define $z = \frac{r_j+\tau_j}{a}$, then the distribution of unit costs for fuel j is given by,

$$G_{jk}(z) = \mathbb{P}\Big[z_j(v,\varphi) \leq z\Big] = 1 - e^{-T_{jk}\left(\frac{r_j+\tau_j}{z}\right)^{-\theta_k}}.$$

C.2.2 Deriving the minimum unit cost distribution

The distribution of the minimum unit cost, $z(v,\varphi;\mathcal{J}(\varphi)) = \min_{j\in\mathcal{J}(\varphi)}\Big\{(r_j+\tau_j)a_j(v,\varphi)\Big\}$, across the firm's inputs is derived as follows,

$$\begin{aligned}
G_k(z;\mathcal{J}(\varphi)) = \mathbb{P}\Big[z(v,\varphi;\mathcal{J}(\varphi)) \leq z\Big] &= 1 - \prod_{j\in\mathcal{J}(\varphi)}\Big[1 - \mathbb{P}\Big[z_j(v,\varphi) \leq z\Big]\Big] \\
&= 1 - \prod_{j\in\mathcal{J}(\varphi)} e^{-T_{jk}\left(\frac{r_j+\tau_j}{z}\right)^{-\theta_k}} \\
&= 1 - e^{-z^{\theta_k}\sum_{j\in\mathcal{J}(\varphi)} T_{jk}(r_j+\tau_j)^{-\theta_k}}.
\end{aligned}$$

C.2.3 Deriving the probability that fuel j is the lowest cost alternative

The probability that a fuel $j \in \mathcal{J}(\varphi)$ is the cheapest alternative to produce an input is derived as follows,

$$
\begin{aligned}
\Psi_{jk}(\varphi; \mathcal{J}(\varphi)) &= \mathbb{P}\Big[z_j(v,\varphi) = \min_{s\in\mathcal{J}(\varphi)}\{z_s(v,\varphi)\}\Big] \\
&= \int_0^\infty \prod_{s\in\mathcal{J}(\varphi), s\neq j} [1 - G_{sk}(z)] dG_{jk}(z) \\
&= \int_0^\infty \prod_{s\in\mathcal{J}(\varphi), s\neq j} e^{-T_{sk}(r_s+\tau_s)^{-\theta_k} z^{\theta_k}} dG_{jk}(z) \\
&= \int_0^\infty \prod_{s\in\mathcal{J}(\varphi), s\neq j} e^{-T_{sk}(r_s+\tau_s)^{-\theta_k} z^{\theta_k}} g_{jk}(z) dz \\
&= \int_0^\infty e^{-\sum_{s\in\mathcal{J}(\varphi), s\neq j} T_{sk}(r_s+\tau_s)^{-\theta_k} z^{\theta_k}} e^{-T_{jk}(r_j+\tau_j)^{-\theta_k} z^{\theta_k}} \theta_k z^{\theta_k - 1} T_{jk}(r_j+\tau_j)^{-\theta_k} dz \\
&= \int_0^\infty e^{-\sum_{s\in\mathcal{J}(\varphi)} T_{sk}(r_s+\tau_s)^{-\theta_k} z^{\theta_k}} \theta_k z^{\theta_k - 1} T_{jk}(r_j+\tau_j)^{-\theta_k} dz \\
&= \Big[- e^{-\sum_s T_{sk}(r_s+\tau_s)^{-\theta_k} z^{\theta_k}} \frac{T_{jk}(r_j+\tau_j)^{-\theta_k}}{\sum_{s\in\mathcal{J}(\varphi)} T_{sk}(r_s+\tau_s)^{-\theta_k}} \Big]_0^\infty \\
&= \frac{T_{jk}(r_j+\tau_j)^{-\theta_k}}{\sum_{s\in\mathcal{J}(\varphi)} T_{sk}(r_s+\tau_s)^{-\theta_k}}.
\end{aligned}
$$

C.2.4 Deriving the minimum unit cost distribution conditional on fuel j being the lowest cost alternative

The distribution of the minimum unit cost across inputs conditional on fuel j being the lowest cost alternative is given by,

$$
\mathbb{P}\Big[z(v,\varphi;\mathcal{J}(\varphi)) \leq z \Big| z_j(v,\varphi) = \min_{s\in\mathcal{J}(\varphi)}\{z_s(v,\varphi)\}\Big] = \frac{\mathbb{P}\Big[z_j(v,\varphi) \leq z, z_j(v,\varphi) \leq z_s(v,\varphi) \forall s \in \mathcal{J}(\varphi)\Big]}{\mathbb{P}\Big[z_j(v,\varphi) = \min_{s\in\mathcal{J}(\varphi)}\{z_s(v,\varphi)\}\Big]}
$$

We show above that the probability in the denominator is equal to,

$$
\mathbb{P}\Big[z_j(v,\varphi) = \min_{s\in\mathcal{J}(\varphi)}\{z_s(v,\varphi)\}\Big] = \Psi_{jk}(\varphi;\mathcal{J}(\varphi)).
$$

Suppose for simplicity that $z_j(v,\varphi)$ is a discrete random variable that can take on the values $\{1,2,3\}$. Suppose also that $z = 2$. The numerator is then,

$$
\begin{aligned}
\mathbb{P}\Big[z_j(v,\varphi) \leq z, z_j(v,\varphi) < z_s(v,\varphi) \forall s \in \mathcal{J}(\varphi)\Big] &= \mathbb{P}[z_j(v,\varphi) = 1] \prod_{s\neq j} \mathbb{P}[z_s(v,\varphi) > 1] \\
&+ \mathbb{P}[z_j(v,\varphi) = 2] \prod_{s\neq j} \mathbb{P}[z_s(v,\varphi) > 2]
\end{aligned}
$$

With a continuous random variable with lower bound 0, the corresponding expression is,

$$
\begin{aligned}
\mathbb{P}\Big[z_j(v,\varphi) \leq z, z_j(v,\varphi) < z_s(v,\varphi) \forall s \in \mathcal{J}(\varphi)\Big] &= \int_0^z g_{jk}(q) \prod_{s\in\mathcal{J}(\varphi), s\neq j} \mathbb{P}[z_s(v,\varphi) > q] dq \\
&= \int_0^z g_{jk}(q) \prod_{s\in\mathcal{J}(\varphi), s\neq j} [1 - G_{sk}(q)] dq.
\end{aligned}
$$

Now note that,

$$g_{jk}(q) = \theta_k q^{\theta_k - 1} T_{jk}(r_j + \tau_j)^{-\theta_k} \exp\Big[- T_{jk}(r_j + \tau_j)^{-\theta_k} q^{\theta_k}\Big]$$
$$1 - G_{sk}(q) = \exp\Big[- T_{sk}(r_s + \tau_s)^{-\theta_k} q^{\theta_k}\Big]$$

Substituting these into the probability above yields,

$$\begin{aligned}
&\int_0^z g_{jk}(q) \prod_{s\in\mathcal{J}(\varphi), s\neq j} [1 - G_{sk}(q)]dq \\
&= \int_0^z \theta_k q^{\theta_k-1} T_{jk}(r_j+\tau_j)^{-\theta_k} \exp\Big[- T_{jk}(r_j+\tau_j)^{-\theta_k} q^{\theta_k}\Big] \prod_{s\in\mathcal{J}(\varphi), s\neq j} \exp\Big[- T_{sk}(r_s+\tau_s)^{-\theta_k} q^{\theta_k}\Big] dq \\
&= \int_0^z \theta_k q^{\theta_k-1} T_{jk}(r_j+\tau_j)^{-\theta_k} \prod_{s\in\mathcal{J}(\varphi)} \exp\Big[- T_{sk}(r_s+\tau_s)^{-\theta_k} q^{\theta_k}\Big] dq \\
&= \int_0^z \theta_k q^{\theta_k-1} T_{jk}(r_j+\tau_j)^{-\theta_k} \exp\Big[- q^{\theta_k} \sum_{s\in\mathcal{J}(\varphi)} T_{sk}(r_s+\tau_s)^{-\theta_k}\Big] dq \\
&= \Big[- \exp\Big[- q^{\theta_k} \sum_{s\in\mathcal{J}(\varphi)} T_{sk}(r_s+\tau_s)^{-\theta_k}\Big] \frac{T_{jk}(r_j+\tau_j)^{-\theta_k}}{\sum_{s\in\mathcal{J}(\varphi)} T_{sk}(r_s+\tau_s)^{-\theta_k}}\Big]_0^z \\
&= - \exp\Big[- z^{\theta_k} \sum_{s\in\mathcal{J}(\varphi)} T_{sk}(r_s+\tau_s)^{-\theta_k}\Big] \frac{T_{jk}(r_j+\tau_j)^{-\theta_k}}{\sum_{s\in\mathcal{J}(\varphi)} T_{sk}(r_s+\tau_s)^{-\theta_k}} + \frac{T_{jk}(r_j+\tau_j)^{-\theta_k}}{\sum_{s\in\mathcal{J}(\varphi)} T_{sk}(r_s+\tau_s)^{-\theta_k}} \\
&= \frac{T_{jk}(r_j+\tau_j)^{-\theta_k}}{\sum_{s\in\mathcal{J}(\varphi)} T_{sk}(r_s+\tau_s)^{-\theta_k}} \Big[1 - \exp\Big[- z^{\theta_k} \sum_{s\in\mathcal{J}(\varphi)} T_{sk}(r_s+\tau_s)^{-\theta_k}\Big]\Big] \\
&= \Psi_{jk}(\varphi; \mathcal{J}(\varphi)) \Big[1 - \exp\Big[- z^{\theta_k} \sum_{s\in\mathcal{J}(\varphi)} T_{sk}(r_s+\tau_s)^{-\theta_k}\Big]\Big]
\end{aligned}$$

It follows that,

$$\begin{aligned}
\mathbb{P}\Big[z(v,\varphi;\mathcal{J}(\varphi)) \leq z \Big| z_j(v,\varphi) = \min_{s\in\mathcal{J}(\varphi)}\{z_s(v,\varphi)\}\Big] &= 1 - \exp\Big[- z^{\theta_k} \sum_{s\in\mathcal{J}(\varphi)} T_{sk}(r_s+\tau_s)^{-\theta_k}\Big] \\
&= G_k(z; \mathcal{J}(\varphi))
\end{aligned}$$

C.2.5 Deriving the cost function

The cost function is derived by solving the following problem,

$$\min_{\{e(v)\}_v} \int_0^1 z(v,\varphi;\mathcal{J}(\varphi)) e(v) dv$$
$$s.t. \quad q_k(\varphi) = \varphi\Big[\int_0^1 e(v)^{\frac{\rho_k-1}{\rho_k}} dv\Big]^{\frac{\rho_k}{\rho_k-1}}$$

We can solve this problem by setting up the Lagrangian,

$$\mathcal{L} = \int_0^1 z(v,\varphi;\mathcal{J}(\varphi)) e(v) dv + \lambda\Big[q_k(\varphi) - \varphi\Big[\int_0^1 e(v)^{\frac{\rho_k-1}{\rho_k}} dv\Big]^{\frac{\rho_k}{\rho_k-1}}\Big]$$

The FOCs are,

$$z(v,\varphi;\mathcal{J}(\varphi)) - \lambda\varphi e(v)^{-\frac{1}{\rho_k}} = 0$$
$$q_k(\varphi) - \varphi\Big[\int_0^1 e(v)^{\frac{\rho_k-1}{\rho_k}}dv\Big]^{\frac{\rho_k}{\rho_k-1}} = 0$$

Divide the FOC of v with the FOC of v',

$$\frac{z(v,\varphi;\mathcal{J}(\varphi))}{z(v',\varphi;\mathcal{J}(\varphi))} = \Big[\frac{e(v)}{e(v')}\Big]^{-\frac{1}{\rho_k}}$$
$$\Longleftrightarrow e(v) = \Big[\frac{z(v,\varphi;\mathcal{J}(\varphi))}{z(v',\varphi;\mathcal{J}(\varphi))}\Big]^{-\rho_k} e(v')$$

Substitute into constraint,

$$\varphi\Big[\int_0^1 \Big(\Big[\frac{z(v,\varphi;\mathcal{J}(\varphi))}{z(v',\varphi;\mathcal{J}(\varphi))}\Big]^{-\rho_k} e(v')\Big)^{\frac{\rho_k-1}{\rho_k}} dv\Big]^{\frac{\rho_k}{\rho_k-1}} = q_k(\varphi)$$
$$\Longleftrightarrow \varphi\Big[\int_0^1 \Big(\Big[\frac{z(v',\varphi;\mathcal{J}(\varphi))}{z(v,\varphi;\mathcal{J}(\varphi))}\Big]^{\rho_k-1} e(v')^{\frac{\rho_k-1}{\rho_k}}\Big)dv\Big]^{\frac{\rho_k}{\rho_k-1}} = q_k(\varphi)$$
$$\Longleftrightarrow \varphi z(v',\varphi;\mathcal{J}(\varphi))^{\rho_k} e(v')\Big[\int_0^1 z(v,\varphi;\mathcal{J}(\varphi))^{1-\rho_k}dv\Big]^{\frac{\rho_k}{\rho_k-1}} = q_k(\varphi)$$
$$\Longleftrightarrow \varphi z(v',\varphi;\mathcal{J}(\varphi))^{\rho_k} e(v') Z_k(\varphi;\mathcal{J}(\varphi))^{-\rho_k} = q_k(\varphi)$$
$$\Longleftrightarrow e(v') = \Big[\frac{Z_k(\varphi;\mathcal{J}(\varphi))}{z(v',\varphi;\mathcal{J}(\varphi))}\Big]^{\rho_k}\frac{q_k(\varphi)}{\varphi}$$

Multiply by $z(v',\varphi;\mathcal{J}(\varphi))$ and integrate over $v' \in [0,1]$ to obtain the cost function:

$$\begin{aligned}
\int_0^1 z(v',\varphi;\mathcal{J}(\varphi))e(v')dv' &= \int_0^1 Z_k(\varphi;\mathcal{J}(\varphi))^{\rho_k} z(v',\varphi;\mathcal{J}(\varphi))^{1-\rho_k}\frac{q_k(\varphi)}{\varphi}dv' \\
&= \frac{q_k(\varphi)}{\varphi} Z_k(\varphi;\mathcal{J}(\varphi))^{\rho_k}\int_0^1 z(v',\varphi;\mathcal{J}(\varphi))^{1-\rho_k}dv' \\
&= \frac{q_k(\varphi)}{\varphi}\Big[\int_0^1 z(v',\varphi;\mathcal{J}(\varphi))^{1-\rho_k}dv\Big]^{\frac{\rho_k}{1-\rho_k}}\Big[\int_0^1 z(v',\varphi;\mathcal{J}(\varphi))^{1-\rho_k}dv'\Big] \\
&= \frac{q_k(\varphi)}{\varphi}\Big[\int_0^1 z(v',\varphi;\mathcal{J}(\varphi))^{1-\rho_k}dv\Big]^{\frac{1}{1-\rho_k}}
\end{aligned}$$

The cost function is thus,

$$C_k(\varphi;\mathcal{J}(\varphi)) = \frac{q_k(\varphi)}{\varphi} Z_k(\varphi;\mathcal{J}(\varphi))$$

We can derive an expression for the unit-cost index $Z_k(\varphi;\mathcal{J}(\varphi))$ in terms of the deeper parameters of the model:

$$\begin{aligned}
Z_k(\varphi;\mathcal{J}(\varphi)) &= \Big[\int_0^1 z(v,\varphi;\mathcal{J}(\varphi))^{1-\rho_k}dv\Big]^{\frac{1}{1-\rho_k}} \\
&= \Big[\mathbb{E}[z(v,\varphi;\mathcal{J}(\varphi))^{1-\rho_k}]\Big]^{\frac{1}{1-\rho_k}} \\
&= \Big[\int_0^\infty z^{1-\rho_k} g_k(z;\mathcal{J}(\varphi))dz\Big]^{\frac{1}{1-\rho_k}}
\end{aligned}$$

Denote $\Phi_k(\mathcal{J}(\varphi)) = \sum_{j\in\mathcal{J}(\varphi)} T_{jk}(r_j+\tau_j)^{-\theta_k}$, then

$$\begin{aligned} G_k(z;\mathcal{J}(\varphi)) &= 1 - e^{-z^{\theta_k}\Phi_k} \\ \iff g_k(z;\mathcal{J}(\varphi)) &= G'_k(z;\mathcal{J}(\varphi)) = \theta_k\Phi_k z^{\theta_k-1}e^{-\Phi_k p^{\theta_k}} \end{aligned}$$

It follows that,

$$Z_k(\varphi;\mathcal{J}(\varphi)) = \left[\int_0^\infty z^{\theta_k-\rho_k}\theta_k\Phi_k e^{-\Phi_k z^{\theta_k}}dz\right]^{\frac{1}{1-\rho_k}}$$

Now define

$$\begin{aligned} x_k &= \Phi_k z^{\theta_k} \\ \iff z &= \left(\frac{x_k}{\Phi_k}\right)^{\frac{1}{\theta_k}} \\ \iff dz &= \frac{1}{\theta_k}\left(\frac{x_k}{\Phi_k}\right)^{\frac{1-\theta_k}{\theta_k}}\frac{1}{\Phi_k}dx_k \end{aligned}$$

and substitute out z and dz from the integral,

$$\begin{aligned} Z_k(\varphi;\mathcal{J}(\varphi)) &= \left[\int_0^\infty \left(\frac{x_k}{\Phi_k}\right)^{\frac{\theta_k-\rho_k}{\theta_k}}\theta_k\Phi_k e^{-x_k}\frac{1}{\theta_k}\left(\frac{x_k}{\Phi_k}\right)^{\frac{1-\theta_k}{\theta_k}}\frac{1}{\Phi_k}dx_k\right]^{\frac{1}{1-\rho_k}} \\ &= \left[\int_0^\infty \left(\frac{x_k}{\Phi_k}\right)^{\frac{1-\rho_k}{\theta_k}} e^{-x_k}dx_k\right]^{\frac{1}{1-\rho_k}} \\ &= \left(\frac{1}{\Phi_k}\right)^{\frac{1}{\theta_k}}\left[\int_0^\infty x_k^{\frac{1-\rho_k}{\theta_k}} e^{-x_k}dx_k\right]^{\frac{1}{1-\rho_k}} \end{aligned}$$

Now note that the Gamma function is defined as,

$$\Gamma(z) = \int_0^\infty t^{z-1}e^{-t}dt$$

so lets define

$$\begin{aligned} t &= x_k \\ z-1 &= \frac{1-\rho_k}{\theta_k} \end{aligned}$$

It follows that

$$\Gamma\left(\frac{\theta_k+1-\rho_k}{\theta_k}\right) = \int_0^\infty x_k^{\frac{1-\rho_k}{\theta_k}}e^{-x_k}dx_k$$

The unit-cost index is thus given by

$$Z_k(\varphi;\mathcal{J}(\varphi)) = \Phi_k(\mathcal{J}(\varphi))^{-\frac{1}{\theta_k}}\gamma_k$$

where $\gamma_k = \left[\Gamma\left(\frac{\theta_k+1-\rho_k}{\theta_k}\right)\right]^{\frac{1}{1-\rho_k}}$. It follows that the cost function is given by,

$$\begin{aligned} C_k(\varphi;\mathcal{J}(\varphi)) &= \frac{q_k(\varphi)}{\varphi}Z_k(\varphi;\mathcal{J}(\varphi)) \\ &= \frac{q_k(\varphi)}{\varphi}\Phi_k(\mathcal{J}(\varphi))^{-\frac{1}{\theta_k}}\gamma_k \end{aligned}$$

C.2.6 Deriving the profit function

The firm's operating profit maximization problem, conditioning on $\mathcal{J}(\varphi)$, is given by,

$$\begin{aligned}
&\max_{p_k(\varphi)} \pi_k(\varphi) = p_k(\varphi)q_k(\varphi) - C_k(\varphi;\mathcal{J}(\varphi)) \\
s.t. \qquad & q_k(\varphi) = E_k P_k^{\sigma_k-1} p_k(\varphi)^{-\sigma_k} \\
& C_k(\varphi;\mathcal{J}(\varphi)) = \frac{q_k(\varphi)}{\varphi}\Phi_k(\mathcal{J}(\varphi))^{-\frac{1}{\theta_k}}\gamma_k
\end{aligned}$$

Substituting the constraints into the objective function gives,

$$\max_{p_k(\varphi)} \pi_k(\varphi) = E_k P_k^{\sigma_k-1} p_k(\varphi)^{1-\sigma_k} - \frac{1}{\varphi}\Phi_k(\mathcal{J}(\varphi))^{-\frac{1}{\theta_k}}\gamma_k E_k P_k^{\sigma_k-1} p_k(\varphi)^{-\sigma_k}$$

Taking the FOC and solving for $p_k(\varphi)$ gives the optimal price:

$$\begin{aligned}
& (1-\sigma_k)E_k P_k^{\sigma_k-1} p_k(\varphi)^{-\sigma_k} + \sigma_k\frac{1}{\varphi}\Phi_k(\mathcal{J}(\varphi))^{-\frac{1}{\theta_k}}\gamma_k E_k P_k^{\sigma_k-1} p_k(\varphi)^{-\sigma_k-1} = 0 \\
\iff & \sigma_k\frac{1}{\varphi}\Phi_k(\mathcal{J}(\varphi))^{-\frac{1}{\theta_k}}\gamma_k E_k P_k^{\sigma_k-1} p_k(\varphi)^{-\sigma_k-1} = (\sigma_k-1)E_k P_k^{\sigma_k-1} p_k(\varphi)^{-\sigma_k} \\
\iff & \sigma_k\frac{1}{\varphi}\Phi_k(\mathcal{J}(\varphi))^{-\frac{1}{\theta_k}}\gamma_k p_k(\varphi)^{-\sigma_k-1} = (\sigma_k-1)p_k(\varphi)^{-\sigma_k} \\
\iff & \sigma_k\frac{1}{\varphi}\Phi_k(\mathcal{J}(\varphi))^{-\frac{1}{\theta_k}}\gamma_k p_k(\varphi)^{-1} = (\sigma_k-1) \\
\iff & \frac{\sigma_k}{\sigma_k-1}\frac{1}{\varphi}\Phi_k(\mathcal{J}(\varphi))^{-\frac{1}{\theta_k}}\gamma_k = p_k(\varphi) \\
\iff & p_k(\varphi) = \frac{\sigma_k}{\sigma_k-1}c_k(\varphi;\mathcal{J}(\varphi))
\end{aligned}$$

where $c_k(\varphi;\mathcal{J}(\varphi))$ is the marginal cost. If we insert this price back into the expression for profits we obtain the operating profit function,

$$\begin{aligned}
\pi_k(\varphi) &= E_k P_k^{\sigma_k-1} p_k(\varphi)^{1-\sigma_k} - \frac{1}{\varphi}\Phi_k(\mathcal{J}(\varphi))^{-\frac{1}{\theta_k}}\gamma_k E_k P_k^{\sigma_k-1} p_k(\varphi)^{-\sigma_k} \\
&= E_k P_k^{\sigma_k-1}\left[p_k(\varphi)^{1-\sigma_k} - \frac{1}{\varphi}\Phi_k(\mathcal{J}(\varphi))^{-\frac{1}{\theta_k}}\gamma_k p_k(\varphi)^{-\sigma_k}\right] \\
&= E_k P_k^{\sigma_k-1} p_k(\varphi)^{-\sigma_k}\left[p_k(\varphi) - \frac{1}{\varphi}\Phi_k(\mathcal{J}(\varphi))^{-\frac{1}{\theta_k}}\gamma_k\right] \\
&= E_k P_k^{\sigma_k-1} p_k(\varphi)^{-\sigma_k}\left[\frac{\sigma_k}{\sigma_k-1}c_k(\varphi;\mathcal{J}(\varphi)) - c_k(\varphi;\mathcal{J}(\varphi))\right] \\
&= E_k P_k^{\sigma_k-1} p_k(\varphi)^{-\sigma_k} c_k(\varphi;\mathcal{J}(\varphi))\left[\frac{\sigma_k}{\sigma_k-1} - 1\right] \\
&= E_k P_k^{\sigma_k-1} p_k(\varphi)^{-\sigma_k} c_k(\varphi;\mathcal{J}(\varphi))\frac{1}{\sigma_k-1} \\
&= E_k P_k^{\sigma_k-1}\left(\frac{\sigma_k}{\sigma_k-1}c_k(\varphi;\mathcal{J}(\varphi))\right)^{-\sigma_k} c_k(\varphi;\mathcal{J}(\varphi))\frac{1}{\sigma_k-1} \\
&= E_k P_k^{\sigma_k-1}\frac{1}{\sigma_k}\left(\frac{\sigma_k}{\sigma_k-1}\right)^{1-\sigma_k} c_k(\varphi;\mathcal{J}(\varphi))^{1-\sigma_k}
\end{aligned}$$

The profit function conditional on $\mathcal{J}(\varphi)$ is thus,

$$\Pi_k(\varphi; \mathcal{J}(\varphi)) = B_k c_k(\varphi; \mathcal{J}(\varphi))^{1-\sigma_k} - \sum_{j \in \mathcal{J}(\varphi)} w f_{jk}$$

where $B_k = \frac{1}{\sigma_k}\left(\frac{\sigma_k}{\sigma_k - 1}\right)^{1-\sigma_k} E_k P_k^{\sigma_k - 1}$.

C.2.7 Deriving the elasticity of substitution

The result

$$\mathbb{P}\Big[z(v,\varphi;\mathcal{J}(\varphi)) \le z \Big| z_j(v,\varphi) = \min_{s \in \mathcal{J}(\varphi)} \{z_s(v,\varphi)\}\Big] = \mathbb{P}\Big[z(v,\varphi;\mathcal{J}(\varphi)) \le z\Big]$$

implies that the distribution of optimal expenditure on each input v

$$\begin{aligned} z(v,\varphi;\mathcal{J}(\varphi))e(v,\varphi;\mathcal{J}(\varphi)) &= z(v,\varphi;\mathcal{J}(\varphi))\Big[\frac{Z_k(\varphi;\mathcal{J}(\varphi))}{z(v,\varphi;\mathcal{J}(\varphi))}\Big]^{\rho_k} \frac{q_k(\varphi)}{\varphi} \\ &= Z_k(\varphi;\mathcal{J}(\varphi))^{\rho_k} z(v,\varphi;\mathcal{J}(\varphi))^{1-\rho_k} \frac{q_k(\varphi)}{\varphi} \end{aligned}$$

is also independent of conditioning on a specific fuel j. That is,

$$\mathbb{P}\Big[z(v,\varphi;\mathcal{J}(\varphi))e(v,\varphi;\mathcal{J}(\varphi)) \le z \Big| z_j(v,\varphi) = \min_{s \in \mathcal{J}(\varphi)} \{z_s(v,\varphi)\}\Big] = \mathbb{P}\Big[z(v,\varphi;\mathcal{J}(\varphi))e(v,\varphi;\mathcal{J}(\varphi)) \le z\Big]$$

Let $\nu_j \subset [0,1]$ denote the set of inputs v for which fuel j is the lowest cost alternative. The total variable cost of production is given by,

$$\begin{aligned} C_k(\varphi;\mathcal{J}(\varphi)) &= \int_0^1 z(v,\varphi;\mathcal{J}(\varphi))e(v,\varphi;\mathcal{J}(\varphi))dv \\ &= \int_0^1 \mathbb{E}\Big[z(v,\varphi;\mathcal{J}(\varphi))e(v,\varphi;\mathcal{J}(\varphi))\Big]dv \\ &= \mathbb{E}\Big[z(v,\varphi;\mathcal{J}(\varphi))e(v,\varphi;\mathcal{J}(\varphi))\Big]\int_0^1 1dv \\ &= \mathbb{E}\Big[z(v,\varphi;\mathcal{J}(\varphi))e(v,\varphi;\mathcal{J}(\varphi))\Big] \end{aligned}$$

The variable cost of producing the inputs in $\nu_j \subset [0,1]$ is given by,

$$\begin{aligned} C_{jk}(\varphi;\mathcal{J}(\varphi)) &= \int_{\nu_j \subset [0,1]} z(v,\varphi;\mathcal{J}(\varphi))e(v,\varphi;\mathcal{J}(\varphi))dv \\ &= \int_{\nu_j \subset [0,1]} \mathbb{E}\Big[z(v,\varphi;\mathcal{J}(\varphi))e(v,\varphi;\mathcal{J}(\varphi))\Big]dv \\ &= \mathbb{E}\Big[z(v,\varphi;\mathcal{J}(\varphi))e(v,\varphi;\mathcal{J}(\varphi))\Big]\int_{\nu_j \subset [0,1]} 1dv \\ &= \mathbb{E}\Big[z(v,\varphi;\mathcal{J}(\varphi))e(v,\varphi;\mathcal{J}(\varphi))\Big]\Psi_{jk}(\varphi;\mathcal{J}(\varphi)) \end{aligned}$$

Hence, expenditure on inputs produced by fuel j as a share of total expenditures on inputs is given by,

$$\frac{C_{jk}(\varphi;\mathcal{J}(\varphi))}{C_k(\varphi;\mathcal{J}(\varphi))} = \Psi_{jk}(\varphi;\mathcal{J}(\varphi)) = \frac{T_{jk}(r_j + \tau_j)^{-\theta_k}}{\sum_{s \in \mathcal{J}(\varphi)} T_{sk}(r_s + \tau_s)^{-\theta_k}}.$$

Now note that we can write $C_{jk}(\varphi; \mathcal{J}(\varphi))$ as,

$$\begin{aligned} C_{jk}(\varphi; \mathcal{J}(\varphi)) &= \int_{\nu_j \subset [0,1]} z(v, \varphi; \mathcal{J}(\varphi)) e(v, \varphi; \mathcal{J}(\varphi)) dv \\ &= \int_{\nu_j \subset [0,1]} (r_j + \tau_j) x_j(v, \varphi; \mathcal{J}(\varphi)) dv \\ &= (r_j + \tau_j) \int_{\nu_j \subset [0,1]} x_j(v, \varphi; \mathcal{J}(\varphi)) dv \\ &= (r_j + \tau_j) X_j(v, \varphi; \mathcal{J}(\varphi)) \end{aligned}$$

where $X_j(v, \varphi; \mathcal{J}(\varphi))$ denotes total quantity of fuel j used in optimum. It follows that,

$$\begin{aligned} \frac{C_{jk}(\varphi; \mathcal{J}(\varphi))}{C_{sk}(\varphi; \mathcal{J}(\varphi))} = \frac{(r_j + \tau_j) X_j(v, \varphi; \mathcal{J}(\varphi))}{(r_s + \tau_s) X_s(v, \varphi; \mathcal{J}(\varphi))} &= \frac{T_{jk}(r_j + \tau_j)^{-\theta_k}}{T_{sk}(r_s + \tau_s)^{-\theta_k}} = \frac{\frac{C_{jk}(\varphi; \mathcal{J}(\varphi))}{C_k(\varphi; \mathcal{J}(\varphi))}}{\frac{C_{sk}(\varphi; \mathcal{J}(\varphi))}{C_k(\varphi; \mathcal{J}(\varphi))}} \\ \Longleftrightarrow \frac{X_j(v, \varphi; \mathcal{J}(\varphi))}{X_s(v, \varphi; \mathcal{J}(\varphi))} &= \frac{T_{jk}}{T_{sk}} \left[\frac{r_j + \tau_j}{r_s + \tau_s}\right]^{-(1+\theta_k)} \end{aligned}$$

Taking logs gives,

$$\log \frac{X_j(v, \varphi; \mathcal{J}(\varphi))}{X_s(v, \varphi; \mathcal{J}(\varphi))} = \log \frac{T_{jk}}{T_{sk}} - (1 + \theta_k) \log \left[\frac{r_j + \tau_j}{r_s + \tau_s}\right]$$

Hence, the elasticity of substitution among fuels in $\mathcal{J}(\varphi)$ is given by $1 + \theta_k$.

C.2.8 Proof of Proposition 1

The marginal value of fuel $s \in J$ given fuel set $\mathcal{J}_k$ is given by,

$$\begin{aligned} D_s \Pi_k(\varphi; \mathcal{J}_k) = & \left[B_k \left(\frac{\varphi}{\gamma_k} \right)^{\sigma_k - 1} \left[\sum_{j \in \mathcal{J}_k} T_{jk}(r_j + \tau_j)^{-\theta_k} \right]^{\frac{\sigma_k - 1}{\theta_k}} - \sum_{j \in \mathcal{J}_k} w f_{jk} \right] \\ & - \left[B_k \left(\frac{\varphi}{\gamma_k} \right)^{\sigma_k - 1} \left[\sum_{j \in \mathcal{J}_k \setminus \{s\}} T_{jk}(r_j + \tau_j)^{-\theta_k} \right]^{\frac{\sigma_k - 1}{\theta_k}} - \sum_{j \in \mathcal{J}_k \setminus \{s\}} w f_{jk} \right] \\ & = B_k \left(\frac{\varphi}{\gamma_k} \right)^{\sigma_k - 1} \left\{ \left[\sum_{j \in \mathcal{J}_k} T_{jk}(r_j + \tau_j)^{-\theta_k} \right]^{\frac{\sigma_k - 1}{\theta_k}} - \left[\sum_{j \in \mathcal{J}_k \setminus \{s\}} T_{jk}(r_j + \tau_j)^{-\theta_k} \right]^{\frac{\sigma_k - 1}{\theta_k}} \right\} \\ & - w f_{sk}. \end{aligned}$$

Now suppose $\mathcal{J}_1 \subset \mathcal{J}_2 \subseteq J$. It follows that,

$$\begin{aligned} D_s \Pi_k(\varphi; \mathcal{J}_1) - D_s \Pi_k(\varphi; \mathcal{J}_2) = & B_k \left(\frac{\varphi}{\gamma_k} \right)^{\sigma_k - 1} \left\{ \left[\sum_{j \in \mathcal{J}_1} T_{jk}(r_j + \tau_j)^{-\theta_k} \right]^{\frac{\sigma_k - 1}{\theta_k}} - \left[\sum_{j \in \mathcal{J}_1 \setminus \{s\}} T_{jk}(r_j + \tau_j)^{-\theta_k} \right]^{\frac{\sigma_k - 1}{\theta_k}} \right\} \\ & - B_k \left(\frac{\varphi}{\gamma_k} \right)^{\sigma_k - 1} \left\{ \left[\sum_{j \in \mathcal{J}_2} T_{jk}(r_j + \tau_j)^{-\theta_k} \right]^{\frac{\sigma_k - 1}{\theta_k}} - \left[\sum_{j \in \mathcal{J}_2 \setminus \{s\}} T_{jk}(r_j + \tau_j)^{-\theta_k} \right]^{\frac{\sigma_k - 1}{\theta_k}} \right\} \end{aligned}$$

This difference is positive (negative) if the second derivative of variable profits with respect to $\Phi_k(\mathcal{J}) = \sum_{j \in \mathcal{J}} T_{jk}(r_j + \tau_j)^{-\theta_k}$ is negative (positive), and zero if the second derivative is zero. We have,

$$\frac{\partial \pi_k(\varphi;\mathcal{J})}{\partial \Phi_k(\mathcal{J})} = B_k\Big(\frac{\varphi}{\gamma_k}\Big)^{\sigma_k-1}\frac{\sigma_k-1}{\theta_k}\Phi_k(\mathcal{J})^{\frac{\sigma-1-\theta}{\theta}} \tag{41}$$

$$\frac{\partial^2 \pi_k(\varphi;\mathcal{J})}{\partial \Phi_k(\mathcal{J})^2} = B_k\Big(\frac{\varphi}{\gamma_k}\Big)^{\sigma_k-1}\frac{\sigma_k-1}{\theta_k}\Phi_k(\mathcal{J})^{\frac{\sigma-1-\theta}{\theta}-1}\Big[\frac{\sigma-(1+\theta)}{\theta}\Big] \tag{42}$$

The second derivative is negative if $1+\theta_k > \sigma_k$, positive if $1+\theta_k < \sigma_k$, and zero if $1+\theta_k = \sigma_k$. This concludes the proof of Proposition 1.

C.3 General equilibrium

C.3.1 The productivity cutoff and industry demand index

Since production involves fixed costs, some entrants with low productivity will not find it profitable to produce. We can define the survival productivity cutoff $\tilde{\varphi}_k$ as the productivity level at which a firm in industry k makes zero profits,

$$\Pi_k(\tilde{\varphi}_k;\mathcal{J}(\tilde{\varphi}_k)) = B_k\Big(\frac{\tilde{\varphi}_k}{\gamma_k}\Big)^{\sigma_k-1}\Big[\sum_{j\in\mathcal{J}(\tilde{\varphi}_k)} T_{jk}(r_j+\tau_j)^{-\theta_k}\Big]^{\frac{\sigma_k-1}{\theta_k}} - \sum_{j\in\mathcal{J}(\tilde{\varphi}_k)} wf_{jk} = 0$$

Entrants with productivity below $\tilde{\varphi}_k$ exit immediately. Note that $\tilde{\varphi}_k$ is an equilibrium variable since it depends on the industry demand index,

$$B_k = \frac{1}{\sigma_k}\Big(\frac{\sigma_k}{\sigma_k-1}\Big)^{1-\sigma_k}E_kP_k^{\sigma_k-1}$$

as well as the fuel prices r_j and the wage rate w. The industry demand index B_k is determined by the mass of firms entering the industry. Firms will continue to enter as long as the expected profit exceeds the sunk entry cost. In equilibrium, we have

$$\mathbb{E}[\Pi_k(\varphi_k;\mathcal{J}(\varphi_k))] = \int_{\tilde{\varphi}_k}^{\infty}\Big[B_k\Big(\frac{\varphi}{\gamma_k}\Big)^{\sigma_k-1}\Big[\sum_{j\in\mathcal{J}(\varphi)} T_{jk}(r_j+\tau_j)^{-\theta_k}\Big]^{\frac{\sigma_k-1}{\theta_k}} - \sum_{j\in\mathcal{J}(\varphi)} wf_{jk}\Big]dF(\varphi) = wf_{ek}.$$

The zero-profit condition in (25) and the free-entry condition in (26) identify unique equilibrium values for $\tilde{\varphi}_k$ and B_k. To see this, first note that the zero-profit condition implies that B_k is monotonically decreasing in $\tilde{\varphi}_k$. In turn, this implies that $\mathbb{E}[\Pi_k(\varphi_k;\mathcal{J}(\varphi_k))]$ is monotonically decreasing in $\tilde{\varphi}_k$ with:

$$\lim_{\tilde{\varphi}_k\to 0}\mathbb{E}[\Pi_k(\varphi_k;\mathcal{J}(\varphi_k))] = \infty$$
$$\lim_{\tilde{\varphi}_k\to \infty}\mathbb{E}[\Pi_k(\varphi_k;\mathcal{J}(\varphi_k))] = 0$$

This ensures that if an equilibrium value of $\tilde{\varphi}_k$ exists, it must be unique. Antras et al. (2017) provide a proof of existence.

C.3.2 The productivity distribution and the mass of firms

The equilibrium productivity distribution among producing firms in industry k is determined by the equilibrium value of $\tilde{\varphi}_k$ and given by $F(\varphi)/[1-F(\tilde{\varphi}_k)]$. Given this, we can derive the average profit among producing firms in industry k from the free-entry condition:

$$\bar{\Pi}_k = \frac{\mathbb{E}[\Pi_k(\varphi_k; \mathcal{J}(\varphi_k))]}{[1 - F(\tilde{\varphi}_k)]} = \frac{wf_{ek}}{[1 - F(\tilde{\varphi}_k)]}.$$

Firm revenue is given by,

$$\begin{aligned} R_k(\varphi) &= E_k P_k^{\sigma_k - 1}\Big[\frac{\sigma_k}{\sigma_k - 1} c_k(\varphi; \mathcal{J}(\varphi))\Big]^{1-\sigma_k} \\ &= \sigma_k B_k c_k(\varphi; \mathcal{J}(\varphi))^{1-\sigma_k} \\ &= \sigma_k\Big[\Pi_k(\varphi; \mathcal{J}(\varphi)) + \sum_{j \in \mathcal{J}(\varphi)} wf_{jk}\Big]. \end{aligned}$$

Average firm revenue among producing firms in industry k is then,

$$\begin{aligned} \bar{R}_k &= \sigma_k\Big[\bar{\Pi}_k + \mathbb{E}[\sum_{j \in \mathcal{J}(\varphi)} wf_{jk}]\Big] \\ &= \sigma_k\Big[\frac{wf_{ek}}{[1 - F(\tilde{\varphi}_k)]} + \frac{1}{[1 - F(\tilde{\varphi}_k)]}\int_{\tilde{\varphi}_k}^{\infty} \sum_{j \in \mathcal{J}(\varphi)} wf_{jk} dF(\varphi)\Big] \\ &= \frac{\sigma_k w}{[1 - F(\tilde{\varphi}_k)]}\Big[f_{ek} + \int_{\tilde{\varphi}_k}^{\infty} \sum_{j \in \mathcal{J}(\varphi)} f_{jk} dF(\varphi)\Big] \end{aligned}$$

The mass of producing firms M_k is a share of the mass of entrants M_{ek} and this share is determined by the equilibrium value of $\tilde{\varphi}_k$:

$$M_k = [1 - F(\tilde{\varphi}_k)]M_{ek}$$

The equilibrium mass of entrants in each industry follows from the industries' labor market clearing conditions and the balance of payment condition below.

C.3.3 Deriving the labor market clearing condition

Let us derive the labor market clearing condition in industry k. To derive demand for labor in industry k, first note that:

$$R_k - \Pi_k = \underbrace{\sum_{j \in J}(r_j + \tau_j)X_j^k}_{\text{Fuel payments}} + \underbrace{w\frac{M_k}{[1 - F(\tilde{\varphi}_k)]}\int_{\tilde{\varphi}_k}^{\infty} \sum_{j \in \mathcal{J}(\varphi)} f_{jk} dF(\varphi)}_{\text{Fixed cost payments}}$$

Then recall that $r_j = d_j w$, where d_j is the unit labor requirement for fuel j. Hence,

$$R_k - \Pi_k = \underbrace{\sum_{j \in J}\tau_j X_j^k}_{\text{Tax payments}} + \underbrace{\sum_{j \in J} wL_j^k}_{\text{Fuel labor payments}} + \underbrace{w\frac{M_k}{[1 - F(\tilde{\varphi}_k)]}\int_{\tilde{\varphi}_k}^{\infty} \sum_{j \in \mathcal{J}(\varphi)} f_{jk} dF(\varphi)}_{\text{Fixed cost payments}}$$

which we can rewrite as

$$R_k - \Pi_k = \sum_{j \in J}\tau_j X_j^k + wL_k^{energy} + wL_k^{fixed}$$

The industry also demand labor to pay for the fixed entry cost $L_k^{entry} = M_{ek} f_{ek}$. Adding this to both sides implies that total labor demand in industry k is equal to

$$wL_k = wL_k^{energy} + wL_k^{fixed} + wL_k^{entry} = R_k - \Pi_k - \sum_{j \in J} \tau_j X_j^k + wL_k^{entry}$$

From the free-entry condition we have that

$$\Pi_k = M_k \frac{w f_{ek}}{[1 - F(\tilde{\varphi}_k)]} = w M_{ek} f_{ek} = w L_k^{entry}$$

Hence, the labor market clearing condition in industry k simplifies to,

$$L_k = \frac{R_k - \sum_{j \in J} \tau_j X_j^k}{w}$$

Summing across industries gives the labor market clearing condition in the whole manufacturing sector:

$$\bar{L} = \sum_k L_k = \sum_k \left(\frac{R_k - \sum_{j \in J} \tau_j X_j^k}{w} \right)$$

C.3.4 Balance of payment condition

General equilibrium requires that total income equals total consumer expenditures,

$$w\bar{L} + \sum_k \sum_{j \in J} \tau_j X_j^k = \sum_k E_k$$

Given that $E_k = R_k$ in a closed economy, this condition is implied by the labor market clearing condition.

C.3.5 Deriving the mass of entrants and the equilibrium price index

Cobb-Douglas preferences implies that total expenditure on industry k is a constant share of total income, $E_k = \beta_k \Big(w\bar{L} + \sum_k \sum_{j \in J} \tau_j X_j^k \Big)$. It follows that total revenue in industry k is equal to

$$R_k = \beta_k \Big(w\bar{L} + \sum_k \sum_{j \in J} \tau_j X_j^k \Big) \tag{43}$$

The mass of active firms in industry k is thus:

$$\begin{aligned} M_k = \frac{R_k}{\bar{R}_k} &= \frac{\beta_k \Big(w\bar{L} + \sum_k \sum_{j \in J} \tau_j X_j^k \Big)}{\frac{\sigma_k w}{[1 - F(\tilde{\varphi}_k)]} \Big[f_{ek} + \int_{\tilde{\varphi}_k}^{\infty} \sum_{j \in \mathcal{J}(\varphi)} f_{jk} dF(\varphi) \Big]} \\ &= [1 - F(\tilde{\varphi}_k)] \frac{\beta_k \Big(w\bar{L} + \sum_k \sum_{j \in J} \tau_j X_j^k \Big)}{\sigma_k w \Big[f_{ek} + \int_{\tilde{\varphi}_k}^{\infty} \sum_{j \in \mathcal{J}(\varphi)} f_{jk} dF(\varphi) \Big]} \end{aligned}$$

and the mass of entrants are

$$M_{ek} = \frac{\beta_k\Big(w\bar{L} + \sum_k \sum_{j\in J} \tau_j X_j^k\Big)}{\sigma_k w\Big[f_{ek} + \int_{\tilde{\varphi}_k}^{\infty} \sum_{j\in\mathcal{J}(\varphi)} f_{jk} dF(\varphi)\Big]}$$

We can now solve for the ideal consumer price index in industry k:

$$P_k^{1-\sigma_k} = M_{ek}\int_{\tilde{\varphi}_k}^{\infty} p_k(\varphi;\mathcal{J}(\varphi))^{1-\sigma_k} dG(\varphi)$$

C.3.6 Deriving the aggregate demand for fuel j

Let us first derive total expenditure on fuel j by firm φ. The cost function is,

$$C_k(\varphi;\mathcal{J}(\varphi)) = \frac{q_k(\varphi)}{\varphi}\gamma_k\Big[\sum_{j\in\mathcal{J}(\varphi)} T_{jk}(r_j+\tau_j)^{-\theta_k}\Big]^{-\frac{1}{\theta_k}}$$

The share of expenditures on fuel j is $\Psi_{jk}(\varphi;\mathcal{J}(\varphi))$. Hence, firm-level expenditures on fuel j conditional on $q_k(\varphi)$ is,

$$\begin{aligned}(r_j+\tau_j)X_j(\varphi, q_k(\varphi)) &= \Psi_{jk}(\varphi;\mathcal{J}(\varphi))C_k(\varphi;\mathcal{J}(\varphi))\\ &= \Psi_{jk}(\varphi;\mathcal{J}(\varphi))\frac{q_k(\varphi)}{\varphi}\gamma_k\Big[\sum_{j\in\mathcal{J}(\varphi)} T_{jk}(r_j+\tau_j)^{-\theta_k}\Big]^{-\frac{1}{\theta_k}}\end{aligned}$$

The optimal quantity $q_k(\varphi;\mathcal{J}(\varphi))$ is given by

$$q_k(\varphi;\mathcal{J}(\varphi)) = E_k P_k^{\sigma-1}(\frac{\sigma}{\sigma-1})^{-\sigma}\Big(\frac{1}{\varphi}\gamma_k\Big[\sum_{j\in\mathcal{J}(\varphi)} T_{jk}(r_j+\tau_j)^{-\theta_k}\Big]^{-\frac{1}{\theta_k}}\Big)^{-\sigma}$$

Combining gives

$$\begin{aligned}(r_j+\tau_j)X_j(\varphi) &= \Psi_{jk}(\varphi;\mathcal{J}(\varphi))E_k P_k^{\sigma-1}(\frac{\sigma}{\sigma-1})^{-\sigma}\Big(\frac{1}{\varphi}\gamma_k\Big[\sum_{j\in\mathcal{J}(\varphi)} T_{jk}(r_j+\tau_j)^{-\theta_k}\Big]^{-\frac{1}{\theta_k}}\Big)^{1-\sigma}\\ &= \Psi_{jk}(\varphi;\mathcal{J}(\varphi))E_k P_k^{\sigma-1}\frac{\sigma-1}{\sigma}(\frac{\sigma}{\sigma-1})^{1-\sigma}\Big(\frac{\varphi}{\gamma_k}\Big)^{\sigma-1}\Big[\sum_{j\in\mathcal{J}(\varphi)} T_{jk}(r_j+\tau_j)^{-\theta_k}\Big]^{\frac{\sigma-1}{\theta_k}}\\ &= \Psi_{jk}(\varphi;\mathcal{J}(\varphi))(\sigma-1)B_k\Big(\frac{\varphi}{\gamma_k}\Big)^{\sigma-1}\Big[\sum_{j\in\mathcal{J}(\varphi)} T_{jk}(r_j+\tau_j)^{-\theta_k}\Big]^{\frac{\sigma-1}{\theta_k}}\\ &= \frac{T_{jk}(r_j+\tau_j)^{-\theta_k}}{\sum_{s\in\mathcal{J}(\varphi)} T_{sk}(r_s+\tau_s)^{-\theta_k}}(\sigma-1)B_k\Big(\frac{\varphi}{\gamma_k}\Big)^{\sigma-1}\Big[\sum_{j\in\mathcal{J}(\varphi)} T_{jk}(r_j+\tau_j)^{-\theta_k}\Big]^{\frac{\sigma-1}{\theta_k}}\\ &= T_{jk}(r_j+\tau_j)^{-\theta_k}(\sigma-1)B_k\Big(\frac{\varphi}{\gamma_k}\Big)^{\sigma-1}\Big[\sum_{j\in\mathcal{J}(\varphi)} T_{jk}(r_j+\tau_j)^{-\theta_k}\Big]^{\frac{\sigma-\theta_k-1}{\theta_k}}\end{aligned}$$

Firm-level demand for fuel j is thus given by,

$$X_j(\varphi) = T_{jk}(r_j+\tau_j)^{-(1+\theta_k)}(\sigma-1)B_k\Big(\frac{\varphi}{\gamma_k}\Big)^{\sigma-1}\Big[\sum_{j\in\mathcal{J}(\varphi)} T_{jk}(r_j+\tau_j)^{-\theta_k}\Big]^{\frac{\sigma-\theta_k-1}{\theta_k}}$$

Aggregate demand for fuel j in industry k is then given by,

$$\begin{aligned} X_j^k &= M_{ek} \int_{\tilde{\varphi}_k}^{\infty} X_j(\varphi) dF(\varphi) \\ &= M_{ek} T_{jk} (r_j + \tau_j)^{-(1+\theta_k)} (\sigma - 1) B_k \left(\frac{1}{\gamma_k}\right)^{\sigma-1} \int_{\tilde{\varphi}_k}^{\infty} \mathbb{I}_j(\varphi) \varphi^{\sigma-1} \Big[\sum_{j \in \mathcal{J}(\varphi)} T_{jk} (r_j + \tau_j)^{-\theta_k} \Big]^{\frac{\sigma-\theta_k-1}{\theta_k}} dF(\varphi) \end{aligned}$$

Aggregate demand for fuel j in the whole manufacturing sector is given by,

$$X_j = \sum_k M_{ek} \int_{\tilde{\varphi}_k}^{\infty} X_j(\varphi) dF(\varphi).$$

D Appendix: Structural estimation

D.1 Production potentials

Industry	Light fuel oil	Heavy fuel oils	LPG	Coal gas	Natural gas	Coke oven gas	Blast furnace gas	Coal	Coke	Wood fuel	Peat	Black liquor	Tall oil	Bio gas	Gasoline	Kerosene	Diesel	Electricity
Repair and installation	-2.07 (0.18)	-1.76 (0.48)	-3.82 (0.33)	-	-2.23 (0.61)	-	-	-	-	-	-	-	-	-	-2.60 (0.18)	-	-2.51 (0.18)	0
Non-metallic mineral products	-2.18 (0.18)	-0.90 (0.34)	-2.31 (0.28)	-	-1.29 (0.45)	-	-	-2.29 (0.77)	-1.28 (1.09)	-2.37 (0.89)	-	-	-	-	-3.30 (0.27)	-	-2.16 (0.18)	0
Textiles	-1.27 (0.26)	-0.19 (0.47)	-1.88 (0.37)	-	-1.19 (0.74)	-	-	-	-	-	-	-	-	-	-1.72 (0.30)	-	-2.86 (0.33)	0
Food products	-1.59 (0.11)	-1.39 (0.27)	- 3.24 (0.21)	-4.41 (0.96)	-1.55 (0.19)	-	-	-	-	-3.08 (0.51)	-	-	-	-4.89 (0.78)	-2.71 (0.16)	-	.2.15 (0.12)	0
Other manufacturing	-1.84 (0.26)	.0.51 (0.54)	-2.90 (0.32)	-	.0.85 (0.86)	-	-	-	-	-3.02 (0.46)	-	-	-	-	-1.67 (0.23)	-	-3.14 (0.32)	0
Metal products	-1.88 (0.07)	-1.73 (0.15)	-3.39 (0.12)	-0.63 (0.88)	-1.97 (0.17)	-	-	-	-	-3.27 (0.35)	-	-	-	-	-1.91 (0.06)	-	-2.55 (0.07)	0
Wood products	-2.89 (0.11)	-2.44 (0.22)	-3.62 (0.55)	-	-2.79 (0.86)	-	-	-	-	-1.46 (0.08)	-	-	-	-	-3.12 (0.13)	-	-1.72 (0.08)	0
Electrical equipment	-1.95 (0.19)	-1.94 (0.52)	-3.23 (0.32)	-	-1.67 (0.41)	-	-	-	-	-3.33 (0.97)	-	-	-	-	-1.28 (0.17)	-	-2.73 (0.22)	0
Machinery equipment	-1.92 (0.09)	-1.50 (0.23)	-3.51 (0.18)	-	-1.29 (0.26)	-	-	-	-	-2.16 (0.39)	-	-	-	-	-1.59 (0.08)	-	-2.40 (0.10)	0
Furniture	-2.88 (0.18)	-2.20 (0.39)	-2.97 (0.48)	-	-2.11 (0.53)	-	-	-	-	-2.21 (0.15)	-	-	-	-	-2.42 (0.16)	-	-2.95 (0.16)	0
Computer and electronic products	-2.43 (0.24)	-	-1.14 (0.53)	-	-1.24 (0.83)	-	-	-	-	-2.84 (0.83)	-	-	-	-	-1.76 (0.16)	-	-3.26 (0.22)	0
Motor vehicles	-2.32 (0.15)	-2.50 (0.29)	-3.01 (0.28)	-	-4.16 (0.44)	-	-	-	-	-5.45 (0.77)	-	-	-	-	-2.45 (0.16)	-	-2.57 (0.15)	0
Printing	-2.15 (0.16)	-1.42 (0.32)	-2.75 (0.24)	-	-2.31 (0.26)	-	-	-	-	-	-	-	-	-	-1.65 (0.10)	-	-2.36 (0.17)	0
Transport equipment	-1.84 (0.28)	-1.53 (0.58)	-3.56 (0.63)	-	-1.99 (0.64)	-	-	-	-	-	-	-	-	-	-2.77 (0.32)	-	-2.20 (0.29)	0
Chemicals	-2.52 (0.23)	-1.78 (0.40)	-3.63 (0.38)	-	-1.37 (0.35)	-	-	-1.67 (1.20)	-	0.14 (0.98)	0.59 (1.20)	-	-	-	-2.99 (0.28)	-	-3.99 (0.26)	0
Pharmaceutical products	-2.50 (0.38)	-1.47 (0.61)	-6.43 (0.75)	-2.33 (0.75)	-2.47 (0.47)	-	-	-	-	-	-	-	-	-	-4.18 (0.34)	-	-5.89 (0.40)	0
Basic metals	-3.37 (0.20)	-2.73 (0.37)	-3.16 (0.20)	-	-2.76 (0.52)	-0.66 (1.16)	-1.07 (1.16)	-2.40 (0.94)	-3.02 (0.52)	-	-	-	-	-	-4.12 (0.23)	-	-3.87 (0.20)	0
Rubber and plastic	-2.57 (0.16)	-2.90 (0.31)	-4.77 (0.24)	-	-2.17 (0.31)	-	-	-	-	-	-	-	-	-	-3.11 (0.15)	-	-3.47 (0.18)	0
Paper and pulp	-4.37 (0.24)	-2.06 (0.26)	-4.62 (0.28)	-	-2.08 (0.60)	-	-	-5.65 (1.26)	-	-2.82 (0.32)	-	-0.92 (0.37)	-3.59 (0.54)	-	-4.73 (0.29)	-	-4.63 (0.26)	0

Table D1: Log production potentials by industry

Notes: The base sample is an unbalanced panel of 90,666 firm-year observations over the years 2004-2020. The production potentials are estimated at the fuel-firm-year level and the base sample at this level contains 1,631,988 observations. We estimate equation (31) via OLS and fuel-year fixed effects for each 2-digit industry separately. We drop industries with less than 30 observations in 2004 and fuels used by a single firm within industry-years.

D.2 Elasticity of substitution between fuels

D.2.1 Baseline specification

Table D2: Elasticity of substitution: first stage estimate

	$\log \frac{r_{jt}+\tau_{jt}}{r_{et}}$
$\log(1+\tau_{jt})$	0.230
	(0.037)
Observations	306

Notes: The first stage relationship is estimated at the fuel-year level and based on then 306 fuel-year observations over the years 2004-2020. Standard errors are bootstrapped and based on 200 replications with replacement.

Table D3: Elasticity of substitution: IV and reduced form estimates

	IV estimate (1)	Reduced form (2)	Observations (3)
Repair and installation	1.386 (0.373)	-0.319 (0.058)	154
Non-metallic mineral products	1.597 (0.301)	-0.368 (0.032)	191
Textiles	2.238 (0.384)	-0.515 (0.030)	129
Food products	2.250 (0.393)	-0.518 (0.029)	194
Other manufacturing	2.381 (0.443)	-0.548 (0.031)	141
Metal products	2.581 (0.465)	-0.594 (0.029)	188
Wood products	2.662 (0.512)	-0.613 (0.046)	142
Electrical equipment	2.674 (0.476)	-0.616 (0.023)	138
Machinery equipment	2.767 (0.498)	-0.637 (0.031)	149
Furniture	2.822 (0.512)	-0.650 (0.032)	140
Computer and electronic products	2.823 (0.516)	-0.650 (0.032)	129
Motor vehicles	2.882 (0.515)	-0.664 (0.028)	165
Printing	3.021 (0.533)	-0.696 (0.031)	132
Transport equipment	3.054 (0.538)	-0.703 (0.038)	135
Chemicals	3.255 (0.549)	-0.749 (0.036)	193
Pharmaceutical products	3.538 (0.603)	-0.815 (0.076)	135
Basic metals	3.597 (0.636)	-0.828 (0.038)	218
Rubber and plastic	3.778 (0.666)	-0.870 (0.034)	142
Paper and pulp	4.402 (0.785)	-1.013 (0.046)	207

Notes: The elasticities of substitution are estimated at the fuel-year level. The number of fuel-year observations underlying the estimates varies by 2-digit industry depending on what fuels are used in a given industry-year. We estimate equation (32) in a two sample IV procedure. We first estimate the first stage relationship between the log relative price and the log carbon tax for all industries. We then estimate the reduced form relationship between the log relative production potentials and the log carbon tax for each industry separately. This table shows the IV estimates and reduced form estimates by industry. Standard errors are bootstrapped and based on 200 replications with replacement.

D.2.2 Alternative specifications

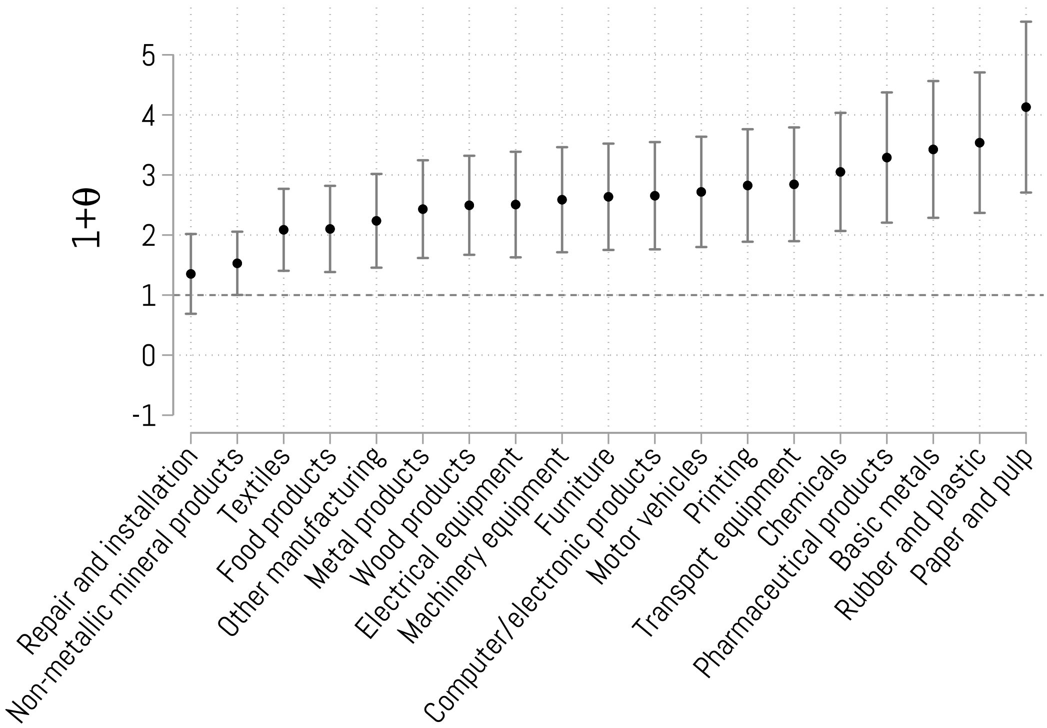

Figure D1: Elasticities of substitution: control for innovation in CO2 abatement technology

Notes: The elasticities of substitution are estimated at the fuel-year level. The number of fuel-year observations underlying the estimates varies by 2-digit industry and is reported in Table D3. We estimate equation (32) in a two sample IV procedure. We first estimate the first stage relationship between the log relative price and the log carbon tax for all industries. We then estimate the reduced form relationship between the log relative production potentials and the log carbon tax for each industry separately. This figure shows the IV estimates when we control for the sum of citation-weighted patents related to climate-change mitigation technologies in the production of goods and industrial processes. Standard errors are bootstrapped and based on 200 replications with replacement.

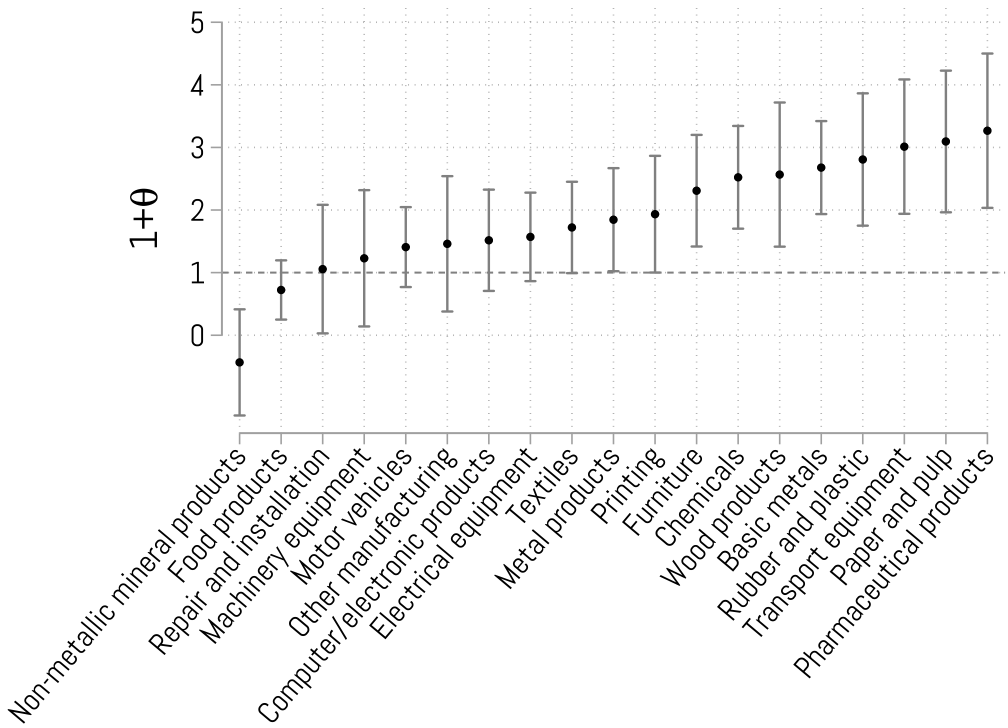

Figure D2: Elasticities of substitution: unweighted reduced form regression

Notes: The elasticities of substitution are estimated at the fuel-year level. The number of fuel-year observations underlying the estimates varies by 2-digit industry and is reported in Table D3. We estimate equation (32) in a two sample IV procedure. We first estimate the first stage relationship between the log relative price and the log carbon tax for all industries. We then estimate the reduced form relationship between the log relative production potentials and the log carbon tax for each industry separately. This figure shows the IV estimates when we do not weight the observations in the reduced form regression by the number of firms underlying the estimates of the production potentials. Standard errors are bootstrapped and based on 200 replications with replacement.

E Appendix: Counterfactual analysis

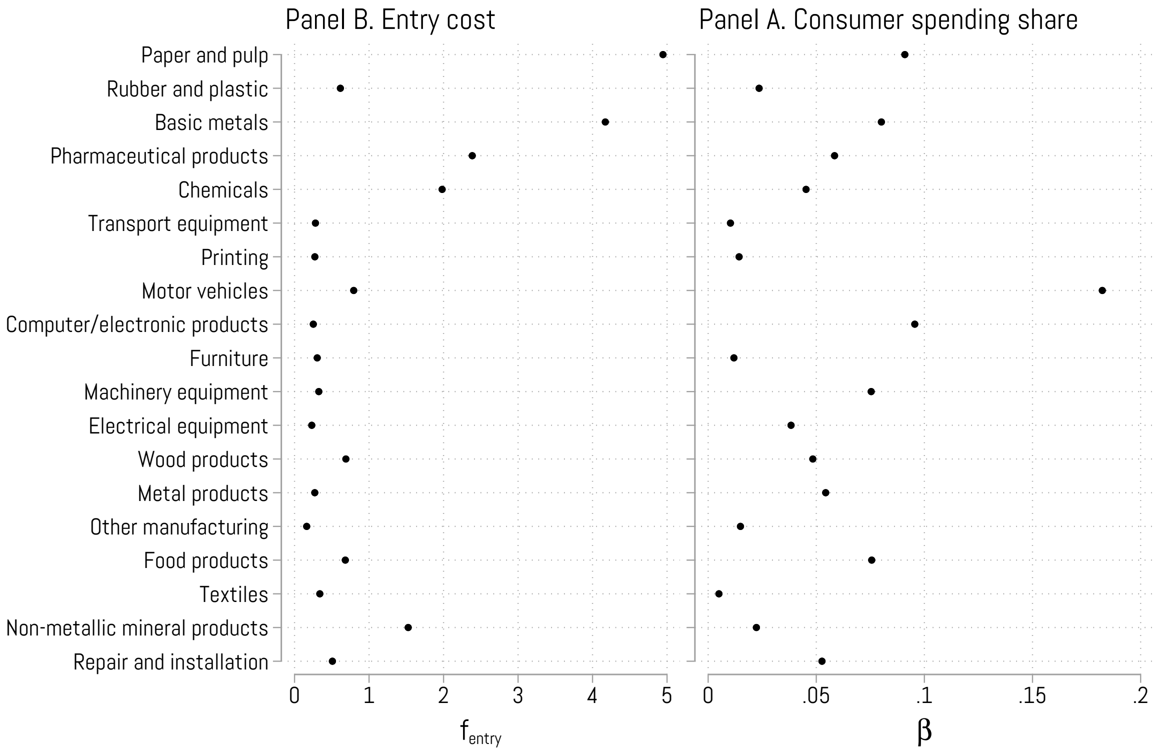

Figure E1: Market demand, Fixed cost dispersion, and Pareto shape parameters

Notes: The figure shows the estimates of the fixed entry costs, f_{ek}, and the Cobb-Douglas consumer spending shares, β_k, by 2-digit industries. The estimates are identified from the model's equilibrium conditions.