Can the plants turn green?

Thore Petersen*

Job Market Paper

31.10.2025

Click for latest version

Abstract

What is the potential for substitution from fossil fuels to clean sources of energy? Fossil fuels account for 70% of primary energy consumption in the German manufacturing sector, but there is substantial heterogeneity in the shares of fossil fuels and electricity across plants. I document that there is five times more variation in the energy mix across plants in 4-digit industries compared to within plant over time. This variation is difficult to explain with observable plant characteristics like location, industry, or products produced. I document that factor demands for fossil fuels and electricity respond differently to changes in output. The elasticity of electricity w.r.t. output is 0.62, while the elasticity of fossil fuels is only 0.16, which is inconsistent with static optimization of a homogeneous production function. These findings can be matched by adding an adjustment cost in fossil fuel use to a dynamic model of plant-level production. I show that neglecting these dynamics leads to an underestimation of the elasticity of substitution between fossil fuels and electricity with common approaches. I compare the calibrated model with adjustment costs to a version of the model without adjustment costs, and with an elasticity of substitution from the literature. The same reduction in aggregate use of fossil fuels can be achieved by a 50% lower carbon tax, and at less than one third of the cost in foregone output.

Keywords: Energy mix, energy transition, firm heterogeneity, adjustment costs, substitution.

JEL classification: D24, E23, Q40, Q41, Q42, Q43, Q48.

*Institute for International Economic Studies, Stockholm University, `thore.petersen@iies.su.se`, `thorepet.github.io`. I thank my advisors Per Krusell and Josh Weiss for a steady stream of critical questions and input. I am also grateful for comments and feedback from John Hassler, Timo Boppart, Kieran Larkin, Kurt Mitman, Brian Higgins, Yimei Zou, and Mitch Downey.

1 Introduction

Fossil fuels are both pervasive and problematic. They fuel the economy, but releasing their energy necessarily also releases carbon dioxide (CO_2), the main driver of climate change. Electricity from renewables is an alternative source of energy with no immediate CO_2 emissions. Can the economy meet its energy demands with such non-emitting sources?

Fuel requirements are typically embodied, meaning a particular piece of equipment requires a particular fuel. Gas turbines, electric motors, and petrol engines have their names for a reason. But the same process can often be performed by different equipment requiring another fuel: All these three pieces of equipment generate mechanical energy, or movement, but from different fuels. The question about the transition away from fossil fuels is thus a question about the development and adoption of technologies.

I study the use of fossil fuels and electricity in the census of German manufacturing plants. The manufacturing sector accounts for 25% of primary energy consumption in Germany, and fossil fuels comprise 70% of its consumption, see figures 1 and 2. I show that there is substantial heterogeneity in this energy mix across plants: Among the subset of plants that produce a single product (6-digit resolution), there is four times more variation across plants than within plant over time. In the full sample, among plants in the same 4-digit industry, there is almost five times more variation in the cross-section than within plant over time. This variation can not be explained by observable characteristics like location, year of entry, or the scope of production (proxied as the share of intermediate inputs in value-added). Plant size however is significantly correlated: larger plants use relatively less fossil fuels. A plant at the 80th percentile of the distribution of revenue has a 5.5 percentage points (p.p.) lower fossil fuel share in the energy mix compared to plant at the 20th percentile on average.

The factor demands for fossil fuels and electricity respond differently to changes in output. I instrument changes in physical output with plant-level demand shocks, and estimate the elasticity of energy use. Fossil fuel use is inelastic. For a 1% change in output, fossil fuel use changes by only 0.16% in the first year, and 0.24% in the second. Electricity use is more elastic. For a 1% change in output, electricity use changes by 0.62% and 0.64% in the first and second year respectively. A homothetic production function, as commonly assumed in the literature, would imply equal elasticities.

To the best of my knowledge this last finding has not previously been documented. It is inconsistent with the common assumption of the choice of energy inputs as static and without dynamic considerations. To study the implications of this finding, I develop a dynamic model of heterogeneous plants that can reproduce these empirical findings. Plants differ in their productivity and a clean energy share parameter, both of which are permanent types. They produce a homogeneous output good by combining clean and dirty energy in a constant elasticity of substitution production function with decreasing returns to scale. I add one critical ingredient: a dynamic adjustment cost for dirty energy use. This is a reduced-form representation of the fact that fossil fuel equipment is subject to technical constraints, like so-called ramp times, that make

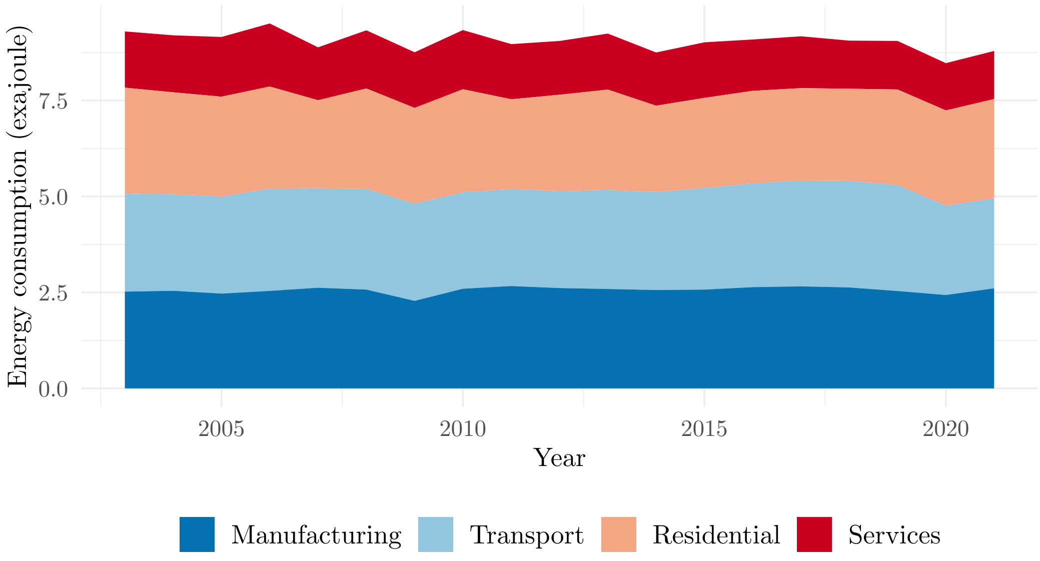


Figure 1: Primary energy consumption over time by sector in Germany. Own calculations based on Arbeitsgemeinschaft Energiebilanzen e.V. (2025).

adjustments costly. Clean energy use can be adjusted freely. There is one source of exogenous variation in the model, an idiosyncratic stochastic demand process. Plants choose their inputs to maximize the net present value of profits. The distribution of plants over types is determined by a selection mechanism at entry. A potential entrant draws a productivity and clean share parameter from independent distributions, and must pay a fixed cost to enter. The fixed cost increases with the clean share parameter, to match the empirical correlation. There are two margins of substitution between clean and dirty energy in the model: A within-plant intensive margin, where a given plant substitutes between clean and dirty energy, and a between-plant extensive margin, where plants with different clean shares are selected at entry.

I calibrate the model and can quantitatively reproduce the empirical findings. The elasticity of substitution between clean and dirty energy at the micro-level is calibrated to a value of 5.1, substantially larger than other micro-level estimates in the literature. I show that neglecting the adjustment cost leads to a downward bias in the estimate, consistent with my finding. The adjustment cost payments are small, representing less than 0.5% of a plant's total costs on average. Yet, they are sufficient to generate the observed difference in elasticity estimates.

I evaluate the relevance of the findings for climate policy by conducting two policy experiments in the calibrated model: an entry subsidy for clean plants, and a tax on dirty energy. The entry subsidy lowers the net fixed cost of entry for plants with a higher clean share parameter draw. It is not effective at increasing the aggregate clean share of energy use. The policy acts only at the entry margin by construction, but aggregate energy consumption is primarily driven by very productive firms, far above the entry cutoff. Spending one entire period's output on the entry

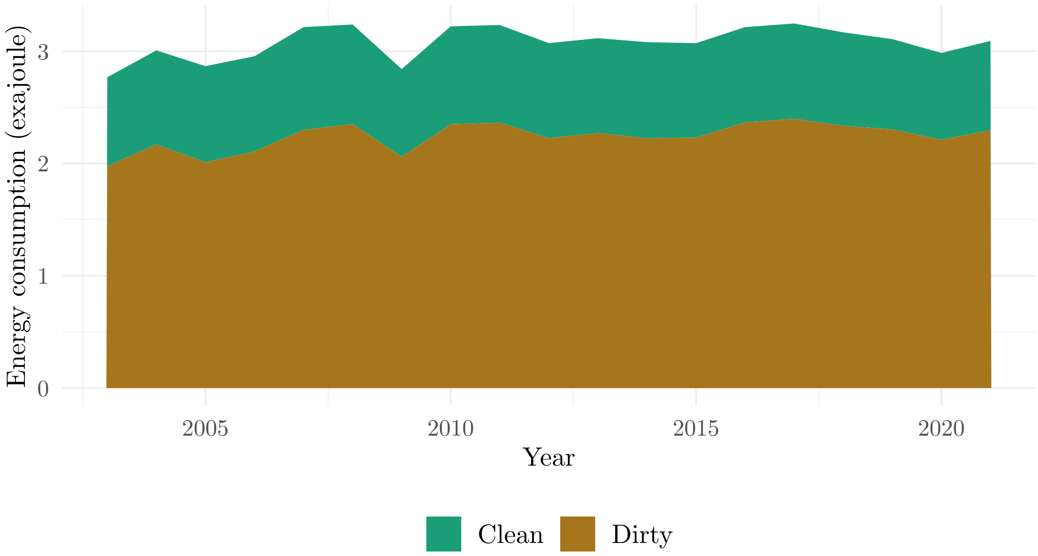


Figure 2: Primary energy consumption over time by clean and dirty sources within the manufacturing sector in Germany. Own calculations based on microdata provided by Research Data Centre of the Statistical Offices of the Federal States (2023).

subsidy increases the aggregate clean energy share by less than 0.1%. The policy is expansionary, meaning that it even increases total use of both clean and dirty energy.

A tax on dirty energy is effective at reducing dirty energy use. A 20% reduction in aggregate dirty energy use can be achieved with a permanent 7% tax on dirty energy prices. The policy strictly increases marginal costs, and thus leads to a contraction in aggregate output of 2%, while generating revenue of around 0.7% of output. I show that the effectiveness of the tax critically depends on the value of the elasticity of substitution between clean and dirty energy. In a counterfactual calibration in line with existing micro-level estimates at a value of 1.5, the tax must be 12% to achieve the same 20% reduction in dirty energy use, leading to a contraction in output of 7%. The results suggest that the plants can turn green to a larger extent than previous estimates suggest.

To map from the empirical analysis to the model, I classify energy into "clean" and "dirty". I consider all electricity consumed by a plant as clean, irrespective of how it was generated. While most electricity was generated from fossil sources during my sample, electrification is a necessary condition for many production processes to run without emissions. If a plant can run its process with electricity, it can in principle be powered by renewables. I classify all fossil fuels as dirty.

Related Literature This paper speaks to the literature on climate macroeconomics, in particular the literature on substitution between "clean" and "dirty" inputs. Seminal papers in this literature (Acemoglu et al., 2012, 2016; Golosov et al., 2014) both stress the importance of the substitutability between clean and dirty inputs, and lament the lack of empirical estimates.

The value of the elasticity of substitution determines the path and level of optimal policy, and the transition path of the economy. The transition to a clean economy is faster and cheaper, the higher the elasticity of substitution. Casey (2024) shows that the distinction between a short- and long-run elasticity is relevant for total emissions along the transition path. I contribute to the literature by providing estimates of the elasticity of substitution between clean and dirty inputs in a case for which this distinction is relatively clear: energy. By documenting the large heterogeneity in the energy mix across plants even within product, I show that there is substantial latent potential for aggregate substitution, even through mere technology adoption, rather than innovation. The within-plant elasticity can be interpreted as the short-run elasticity, while the long-run elasticity is augmented both by a demand reallocation channel between plants (Oberfield & Raval, 2021), and the selection at entry.

The paper is closely related to the literature on input substitution. A timeless focus is on the elasticity of substitution between capital and labor (Antràs, 2004; Chirinko, 2008). Diamond et al. (1978) provide an important non-identification result: in the presence of factor-augmenting technical change, the elasticity of substitution is not identified from timeseries data. A large literature devises identification strategies or structural models to estimate the substitutability between intermediate inputs (Barrot & Sauvagnat, 2016; Boehm et al., 2023; Peter & Ruane, 2025). The findings are typically a low short-run elasticity of substitution, that increases with the time horizon. A much smaller literature employs similar methods to estimate the elasticity of substitution between types of energy: Jo (2024) derives an estimation equation from a CES specification, and instruments for changes in the relative price of clean and dirty energy. Papageorgiou et al. (2017) non-linearly estimate a CES production function on sector-level data. Stern (2012) conducts a meta-analysis of estimates at different levels of aggregation. These papers report estimates of the elasticity of substitution of around 3–4 from cross-sectional variation, and around 1 from time-series variation. The different values are interpreted as a short- and long-run elasticity. Kaartinen and Prane (2024) and Leclair (2025) develop structural models and calibrate them to microdata, to study the substitution between a large set of fuels in production. Another approach is the estimation of translog cost functions (Arnberg & Bjørner, 2007; Bousquet & Ladoux, 2006; Hyland & Haller, 2018). The elasticity of substitution is implied by the estimated cross-price elasticities, but it is not typically calculated. I contribute to the literature by providing a structural estimate, and documenting possibly a large potential for substitution, given the large heterogeneity.

Lastly, there is a growing literature estimating causal effects of policy in manufacturing microdata. A focus of the literature is on the response of CO_2 emissions to carbon taxation, or emission trading systems (Andersson, 2019; Colmer et al., 2024; Gerster & Lamp, 2024; Martin et al., 2014). A robust finding is a reduction of the emissions of treated firms or plants, without adverse effects on economic activity. I contribute to the literature by studying energy use at the micro level, and add by considering the heterogeneity between plants.

Energy Primer Energy comes in different forms: thermal, mechanical, and electrical, among others. A common aggregate measure is primary energy consumption. It is defined as the heating potential or energy content of the fuels consumed, for example in units of joule (J) or watt-hours (Wh). When converting between different forms of energy there are losses. The useful energy content of a fuel is almost always lower than its heating potential. For example, a typical gas turbine that generates mechanical power from natural gas has an efficiency of around 40%: only 40% of the heating potential of the gas are converted into useful mechanical energy, the remainder is lost as waste heat. In an increasing number of applications that waste heat is recovered to increase overall efficiency to up to 80–90%. Converting electrical energy into mechanical energy is more efficient, with typical efficiencies for electric motors above 90%. Producing thermal energy from electricity using heat pumps can be done with implicit efficiencies even above 100%, since they concentrate and move heat instead of generating it. To substitute fossil fuels with electricity, only the useful energy content must be replaced, not the entire primary energy consumption. In this paper, I study primary energy consumption, which is reported in the data. With different processes, there is not a single conversion factor from primary to final energy consumption.

How is energy used in manufacturing? In the German manufacturing sector in 2023, 98.5% of fossil fuel use was for thermal applications, predominantly process heat. 67% of electricity use is for mechanical applications, mostly machine drives (for example pumps, compressors, drills, conveyors, or fans). 24% of electricity use is for thermal applications. The remaining 9% include lighting and IT systems. 76% of primary energy consumption is for thermal applications, and 21% for mechanical applications. The remainder includes lighting and IT.[1]

2 Data

2.1 Source and coverage

I use confidential microdata from the census of German manufacturing plants. The data is provided by the Research Data Center of the German statistical agency Destatis, and made available under the name AFiD (*Amtliche Firmendaten für Deutschland*, official firm data for Germany). The unit of observation is a plant (*Betrieb* in German), defined as a geographically bounded unit of production that belongs to a firm. Some variables are available only at the firm (*Unternehmen*) level, where a firm is the smallest independent legal entity required to keep accounts.

The data covers the universe of plants belonging to a firm with 20 or more employees in the manufacturing sector in Germany. I combine the modules on production, energy use, and employment, revenue, and investment (Research Data Centre of the Statistical Offices of the Federal States, 2023). The data covers the years 1995 to 2021, except for the energy module, which starts in 2003.

[1]Own calculations for the year 2023 based on Fraunhofer ISI (2025).

Sample restrictions I remove observations for which the following variables are within the top and bottom 2.5% of observations with strictly positive values by 2-digit industry: (i) dirty energy over clean energy, (ii) clean energy per worker, (iii) dirty energy per worker, and (iv) output per worker. I retain observations for which dirty energy use is zero, as some plants use only clean energy.

2.2 Variable construction

Energy The energy module records the heating potential in kilowatt-hours (kWh) of fuel consumed at the plant for 10 fuel categories. The largest fuel categories by consumption are natural gas, electricity, and coal products, accounting for 28%, 24%, and 20% of total energy use on average over time, respectively.[2] I aggregate the energy use into clean and dirty: I define clean energy as the sum of electricity and renewables, and dirty energy as the sum of fossil fuels, as well as district heat. Some plants generate electricity from fossil fuels on-site, and the data does not distinguish between the use of fossil fuel in production and for electricity generation. For those plants, I estimate the kWh of dirty energy used in production by subtracting the kWh of electricity generated on-site from the total kWh of dirty energy consumed.[3]

Output The production module records both price and quantity for each product a plant produces in a given year. This level of detail allows me to construct a plant-level price index, to generate an accurate measure of real output. I construct a Törnqvist price index at the plant level (described in Eslava et al., 2004).[4] Due to the arbitrary basis, the level of Törnqvist-deflated

[2]The remaining fuel categories and shares are: other oil products (8%), other gas products (6%), district heat (5%), renewables (3%), waste (2%) and heating oil (2%).

[3]The relation between generated electricity, $E^{c\text{ gen}}$, and dirty energy use for generation, $E^{d\text{ gen}}$, is $E^{c\text{ gen}} = \eta^{\text{elec}} E^{d\text{ gen}}$, where $\eta^{\text{elec}} \in (0,1)$ is the electrical efficiency of generation. The remaining heating potential is converted into heat in the process, $E^{d\text{ heat}} = \eta^{\text{heat}} E^{d\text{ gen}}$, with $\eta^{\text{elec}} + \eta^{\text{heat}} = 1$ by the first law of thermodynamics. Almost all plants with on-site generators employ combined heat and power generation (CHP), so use that heat in production. The adjusted dirty energy heating potential available for production is then

$$\begin{aligned} E^{d\text{ prod}} &= E^d - E^{d\text{ gen}} + E^{d\text{ heat}} = E^d - E^{d\text{ gen}} + \eta^{\text{heat}} E^{d\text{ gen}} = E^d - \left(1 - \eta^{\text{heat}}\right) E^{d\text{ gen}} \\ &= E^d - \left(1 - \eta^{\text{heat}}\right) \frac{E^{c\text{ gen}}}{\eta^{\text{elec}}} = E^d - E^{c\text{ gen}}. \end{aligned}$$

This assumes that the recovery rate of useful energy from $E^{d\text{ heat}}$ is equal to the conversion efficiency, which is the case on average. Engineering estimates for both are around 0.7–0.9, depending on exact application.

[4]The Törnqvist index is a chained price index. The change in the index P_t for a basket of goods G from period t to $t+1$ is

$$\begin{aligned} \Delta \log P_{t+1} &= \sum_{g \in G} \bar{s}_{gt+1} \Delta \log P_{gt+1}, \\ \bar{s}_{gt+1} &= \frac{s_{gt+1} + s_{gt}}{2}, \end{aligned}$$

where P_{gt} is good g's price, and s_{gt} is its quantity share of the basket in period t. The index price in period t is then

$$P_t = \exp\left(\sum_{\tau=1}^{t} \Delta \log P_\tau\right).$$

output is not comparable across plants. The index does provide a more accurate measure of the changes in real output though, compared to industry-level deflators. For a measure of output that is comparable across plants, I deflate the nominal value of production and revenue, see below.

Deflators I deflate revenue and the nominal value of production to 2015 Euro using producer price indices (PPIs) at the 2-digit industry level from the Federal Statistical Office of Germany (Destatis). I deflate expenditure on intermediate inputs to 2015 Euro using the corresponding price series at the 2-digit industry level from the EU KLEMS database (Bontadini et al., 2023).[5]

Industry and product classification The industry and product classifications change several times during the sample period. I harmonize all classifications to the versions used in the latest years in the sample. For industry, the mapping between old and new schemes is not unique for most industries. I map all 4-digit industry identifiers to WZ2008 (equivalent to NACE Rev. 2) according to the following rule. Among plants that are active before and after the change in classification, I copy the new classification to the previous years. For plants that are active only under the old classification, I assign the most common transition from the plants that are active in both classifications. For products, I map the 9-digit product identifiers to the GP2019 classification scheme (corresponding to PRODCOM up to 8 digits, with Germany-specific details at the 9th digit). At this resolution, there is a unique mapping between the old and new classifications for most products. For ambiguous cases, I choose the first product listed in the official correspondence table.

Intermediate inputs The expenditure on intermediate inputs or materials is recorded only at the firm level, and not in all years for all firms. This variable is part of the cost structure survey (*Kostenstrukturerhebung*). This survey is conducted every year for firms with 500 or more employees. Among smaller firms, a sample stratified by 4-digit industry and number of employees is drawn, such that a total of around 16,000 firms are included. Smaller firms are included every four years on average. Among multi-plant firms, I assign intermediate inputs expenditure according to the share of total production.[6]

Entry The data does not directly record the year of entry of a plant. For plants that enter after the first year of the sample, 1995, I define their year of entry as the year in which they first appear. Given the inclusion criteria, this is the actual year of entry for plants that are part of a firm with 20 or more employees. For plants belonging to smaller firms, I can only measure the year in which the firm grows to at least 20 employees.

It is implicitly normalized to the first year in which the plant is active.

[5] PPI: Destatis table number 61241-0003; intermediate input prices: EU KLEMS national accounts series II_PI.

[6] 94% of firms have one plant. 82% of plants belong to a single-plant firm. Both shares are stable over time.

Summary statistics Table 1 presents summary statistics of the sample for the main variables used in the analysis.

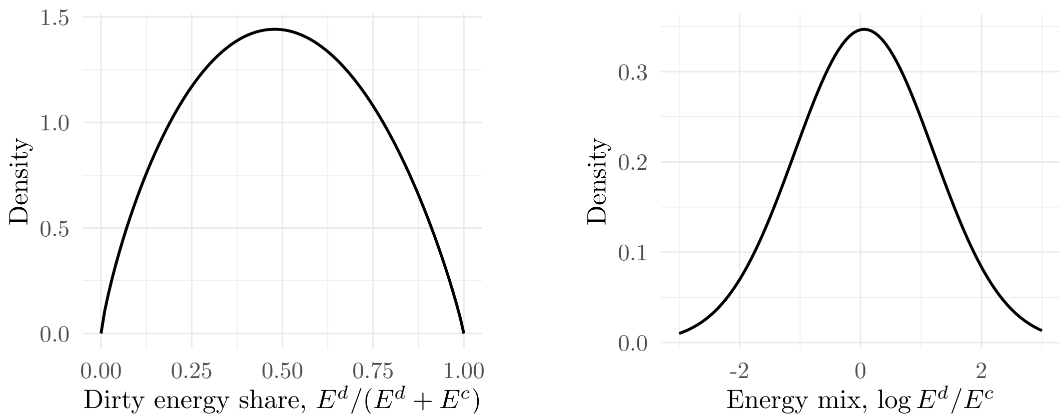
Variable	N	Mean	SD	Median	10th Pct.	90th Pct.
E^c	248590	4305614.00	2039827.00	717022.00	107978.00	8279847.00
E^d	248590	7737987.00	4540306.00	547884.00	101782.00	10426251.00
$p_y Y$	248590	30212108.00	88538350.00	9613197.00	2455404.00	66202875.00
$\log E^d/E^c$	248590	-0.09	1.20	0.06	-1.71	1.30
E. exp./VA	186771	0.06	0.06	0.04	0.01	0.13

Table 1: Summary statistics for clean and dirty energy use (E^c, E^d in kWh), revenue $p_y Y$ (in 2015 Euro), the log of the ratio of dirty to clean energy, $\log E^d/E^c$, as a measure of the energy mix, and total energy expenditure over value added (E. exp./VA, available only at the firm-level).

3 Empirical Results

3.1 Variation in the energy mix

There is substantial variation in the energy mix. I calculate the share of dirty energy in total energy use, $E^d/(E^c + E^d)$ as a measure of the energy mix. This share has a mean of 0.49 and a standard deviation of 0.23. Another measure is the log ratio of dirty to clean energy use, $\log(E^d/E^c)$, which has a mean of -0.01 and a standard deviation of 1.2. I will use this log-ratio as the primary measure for its convenient numerical properties in most subsequent analysis.



(a) Distribution of dirty energy share. (b) Distribution of energy mix.

Figure 3: Density estimates of the energy mix measures.

What drives this variation? Observable characteristics of the plants can explain only a small fraction of it. I estimate the regression

$$\log \frac{E^d_{it}}{E^c_{it}} = \text{fixed effects} + \eta_{it}, \tag{1}$$

for several sets of fixed effects: year, industry at the 2- and 4-digit level, product at the 6-digit level, district (*Landkreis*)[7], and plant. The results are reported in table 2. Table A.1 additionally reports sample size and Bayesian Information Criterion (BIC) to demonstrate that the explanatory power of plant fixed effects is not due to overfitting.

Fixed Effects	adj. R^2
Industry (2-digit)	0.14
Industry (4-digit)	0.27
Industry (4-digit) by Year	0.28
Product (6-digit)	0.35
District	0.03
District + Industry (4-digit)	0.28
Plant	0.87

Table 2: Adjusted R^2 for regressions of energy mix $\log E^d_{it}/E^c_{it}$ on different sets of fixed effects.

In the full sample, the industry a plant belongs to explains a quarter of the variation in the energy mix. Among plants that produce only a single product, the product they produce explains a little over one third of the variation in the energy mix. The district a plant is located in explains very little by itself, and adds only marginally to the explanatory power of industry. The year also explains little variation. Plant fixed effects stand out: they explain 87% of the variation in the energy mix.

Under the assumption of an additive hierarchical model ($\log E^d_{it}/E^c_{it} = \text{fe}_i + \text{fe}_{\text{group}(i)} + \varepsilon_{it}$), I can calculate the ratio of the variances from the R^2 values. The ratio of the between-plant to within-plant variance can then be calculated as

$$\frac{R^2_{\text{plant}} - R^2_{\text{group}}}{1 - R^2_{\text{plant}}},$$

where R^2_{plant} is the R^2 from plant fixed effects, and R^2_{group} is the R^2 from a grouping level (e.g., industry). Using this formula, I find that the variation in the energy mix between plants relative to within plant over time is about 4 times larger among producers of the same product, and 4.6 times larger among plants in the same 4-digit industry.

3.2 Energy mix-size correlation

One observable characteristic that can explain variation in the energy mix is the size of a plant. Larger plants have a higher share of clean energy in their energy mix. This holds within plant

[7] There are 401 *Landkreise* in Germany. The average number of plants per district is 95, the median 68.

Dependent Variables:	$\log E^d_{it}/E^c_{it}$	$\log E^d_{it}/E^c_{it}$
Model:	(1)	(2)
$\log Y_{it}$	-0.16	-0.10
	(0.008)	(0.006)
Fixed-effects		
Plant & Year	Yes	No
4-digit Industry by Year	No	Yes
Fit statistics		
Observations	248,590	248,590
R^2	0.92	0.36
within-R^2	0.013	0.03

Table 3: Conditional correlation between energy mix and plant size. Standard errors are clustered by plant and year.

over time, and in cross-section controlling for different sets of fixed effects.

I estimate the conditional correlation with the regression equation

$$\log \frac{E^d_{it}}{E^c_{it}} = \beta \log Y_{it} + \delta_{\text{fe}(i,t)} + \nu_{it}, \tag{2}$$

where $\delta_{\text{fe}(i,t)}$ are different sets of fixed effects, and Y_{it} is PPI-deflated revenue. The results are reported in table 3.

3.3 Differential dynamics of clean and dirty energy use

3.3.1 Estimation approach

To understand how clean and dirty energy enter production, I estimate their factor demand elasticities: how much does the use of each input change in response to changes in output? To estimate the factor demand elasticity, I regress the change in factor use on the change in output,

$$\Delta_k \log X_{it} = \beta_k \Delta_k \log Y_{it} + \gamma_{\iota(i,t)} + \xi_{r(i,t)} + \epsilon_{it}, \tag{3}$$

where $\Delta_k z_{it} = z_{it} - z_{it-k}$, and X_{it} are production inputs. I include fixed effects $\gamma_{\iota(i,t)}$ for 2-digit industry-by-year, and $\xi_{r(i,t)}$ for district-by-year. This equation suffers from several sources of bias: (i) simultaneity bias (Marschak & Andrews, 1944), as inputs and outputs are jointly chosen; (ii) attenuation bias, as productivity shocks introduce variation in Y_{it} without corresponding changes in X_{it}; (iii) omitted variable bias, as idiosyncratic input price shocks may change both factor demand and the optimal scale of production.

To circumvent these problems, I construct a demand shock instrument for the change in output. For each plant i and year t, I calculate the leave-one-out change in real output within the plant's 4-digit industry $\mathcal{I}_{it}$:

$$\text{shock}_{it}^k = \log \sum_{j \in \mathcal{I}_{it} \setminus \{i\}} Y_{jt} - \log \sum_{j \in \mathcal{I}_{it} \setminus \{i\}} Y_{jt-k}. \tag{4}$$

Exclusion restriction This instrument directly addresses the simultaneity and attenuation biases, as it is uncorrelated with plant-level shocks by construction. It addresses the omitted variable bias for the same reason: conditional on industry-by-year and state-by-year fixed effects, the instrument is uncorrelated with idiosyncratic input price shocks. The exclusion restriction could be violated for plants that have sufficiently high shares in either the input or output markets, such that their behavior influences prices. Concentration is low in the data in general, and I further address the issue by restricting the analysis to 4-digit industries with at least 50 plants each year.

Identifying variation The factor demand elasticity is then identified by variation at the 4-digit industry-by-year level, controlling for 2-digit industry-by-year and state-by-year fixed effects. The 2-digit industry-by-year effects control for aggregate shocks and structural trends, while the state-by-year effects absorb regional shocks.

3.3.2 Results

I estimate equation (3) using 2SLS for clean and dirty energy E^c, E^d, respectively, up to 2-year differences. The first stage results are strong, with a coefficient estimate of 0.14 in the first year, and 0.15 in the second, and F-statistics of 188 and 190, respectively. Table A.2 reports the first stage results.

The factor demand elasticity estimates are reported in Table 4. Dirty energy is not very responsive to demand shocks in general. This contrasts with the response of clean energy, which is much more elastic.

Heterogeneity and robustness The results for both the first and second stage are symmetric for positive and negative demand shocks. I estimate the regression separately by 2-digit industry, and find very similar results in all the industries that have sufficiently many plants for the first stage to be strong. The effects appear to be linear: when including a quadratic term both stages, only the linear terms are statistically significant, and are similar to the linear specification. I repeat the exercise on a subsample of plants that produce only a single product at the 6-digit level. I construct the demand shock instrument analogously at the product-by-year level. The point estimates are very similar to those in the full sample, although less precise due to the smaller sample size.

Dependent Variables: Model:	$\Delta_1 \log E^d$ (1)	$\Delta_1 \log E^c$ (2)	$\Delta_2 \log E^d$ (3)	$\Delta_2 \log E^c$ (4)
$\Delta_1 \log Y$	0.1641	0.6198		
	(0.0592)	(0.0321)		
$\Delta_2 \log Y$			0.2362	0.6411
			(0.0586)	(0.0357)
Fixed-effects				
Year by 2-digit Industry	Yes	Yes	Yes	Yes
Year by District	Yes	Yes	Yes	Yes
Fit statistics				
Observations	275,455	304,305	249,927	276,385

Table 4: Regression results for equation (3), factor demand elasticity for dirty and clean energy at 1-year and 2-year differences. Standard errors are clustered at the plant and 4-digit industry by year level. The dependent variable is the change in factor use, $\Delta_k \log X_{it}$, where k is 1 or 2 years. The key independent variable is the change in output, $\Delta_k \log Y_{it}$. Fixed effects for year by 2-digit industry and year by district are included.

4 Model

4.1 Environment and technology

I develop a dynamic partial equilibrium model of infinitely-lived plants. Plants are heterogeneous in their productivity and a clean energy share parameter, both of which are permanent types. They produce a homogeneous output good by combining clean and dirty energy. Dirty energy is subject to an adjustment cost: changing the level of dirty energy use between periods is costly. Clean energy is chosen freely. The prices of clean and dirty energy are exogenous and constant. The only source of variation to a plant are idiosyncratic demand shocks, modelled as a stochastic process for a plant's output price.

The distribution of plants is determined through an entry margin. Potential entrants draw productivity and clean share types from independent distributions. To enter, they must pay an entry cost that scales with their clean share type. A plant does enter if the expected present discounted value of profits exceeds the entry cost, given its draw of productivity and clean share type.

Plants combine clean and dirty energy, E^c and E^d, in a constant elasticity of substitution (CES) production function to produce energy services E,

$$E(E^c, E^d; b) = \left[b(E^c)^{\frac{\sigma-1}{\sigma}} + (1-b)(E^d)^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}}, \tag{5}$$

where $b \in (0,1)$ is the plant's clean energy share parameter, and $\sigma > 0$ is the elasticity of substitution between clean and dirty energy.

The final good Y is then produced with a decreasing returns to scale (DRS) production function from energy services:

$$Y(E^c, E^d; A, b) = A^{1-\alpha} E(E^c, E^d; b)^{\alpha}, \tag{6}$$

with productivity $A > 0$ and returns to scale parameter $\alpha \in (0, 1)$.

Dirty energy is subject to a quadratic adjustment cost:

$$\phi(E^d, E^d_{-1}) = \frac{\phi_1}{2} \frac{1}{E^d_{-1}} \left(E^d - E^d_{-1}\right)^2, \tag{7}$$

where E^d_{-1} is the previous period's dirty energy use, and $\phi_1 > 0$ governs the scale of the adjustment cost.

The idiosyncratic price for the final good, p_y, follows an AR(1) process in logs,

$$\log p'_y = \rho_y \log p_y + \sigma_y \epsilon', \quad \epsilon' \sim \mathcal{N}(0, 1), \tag{8}$$

with persistence $\rho_y \in (-1, 1)$ and shock standard deviation $\sigma_y > 0$. The shocks ϵ' are i.i.d. over time and across plants, and drawn from a standard normal distribution.

I consider a partial equilibrium model for the following reasons. Germany imports the vast majority of its fossil fuels (except for coal), and electricity is traded on a large European market. Plants are price-takers, and I am abstracting from equilibrium effects on energy prices.

4.2 Equilibrium

I consider a steady state partial equilibrium, with one-time entry. The equilibrium is given by a distribution of plants over types (b, A), such that all entering plants maximize their expected present discounted value of profits given the exogenous prices of clean and dirty energy, and the entry condition is satisfied.

4.3 Plant problem

Plants maximize the expected present discounted value of profits, by choosing clean and dirty energy inputs. I am solving for their value function and policy function for dirty energy use, E^d.

Within period profits are

$$\pi(E^c, E^d, E^d_{-1}, p_y; b, A) = p_y Y(E^c, E^d; A, b) - p_c E^c - p_d E^d - \phi(E^d_{-1}, E^d), \tag{9}$$

where p_c and p_d are the prices of clean and dirty energy, respectively. The Bellman equation is

$$V(E^d_{-1}, p_y; b, A) = \max_{E^c, E^d} \pi(E^c, E^d, E^d_{-1}, p_y; b, A) + \beta \mathbb{E}_{p'_y|p_y} V(E^d, p'_y; b, A), \tag{10}$$

where $\beta \in (0,1)$ is the discount factor.

4.4 Entry

There is a fixed mass of potential entrants normalized to 1. Each draws a clean share type $b \in (0,1)$, and a productivity type $A > 0$ from the independent distributions G_b and G_A.

After entering, a plant can freely choose its initial dirty energy input E^d, such that its value function after entry is

$$V_{\text{entry}}(b, A) = \max_{E^d} V(E^d, \bar{p}_y; b, A), \tag{11}$$

where $\bar{p}_y$ is the unconditional mean of the idiosyncratic price process.

The entry cost is

$$f^e(b) = f_0^e \exp(f_1^e b), \tag{12}$$

where $f_0^e > 0$ governs the scale, and f_1^e represents the semi-elasticity of the entry cost with respect to the clean share: a 1 percentage point increase in b increases the entry cost by approximately f_1^e percent.

A plant enters, if the value of entry exceeds the entry cost,

$$V_{\text{entry}}(b, A) \geq f^e(b). \tag{13}$$

4.5 Solution

The Bellman equation has the following scaling property in A (see appendix section A.1):

$$V\left(\mu E_{-1}^d, p_y; b, \mu A\right) = \mu V(E_{-1}^d, p_y; b, A), \quad \forall \mu > 0. \tag{14}$$

This implies $V(E_{-1}^d, p_y; b, A) = AV(A^{-1}E_{-1}^d, p_y; b, 1) = AV(\tilde{E}_{-1}^d, p_y; b, 1)$: the value function is linear in A when appropriately scaling E_{-1}^d. Thus, I need to solve it only for $A = 1$, and can then rescale the value and policy functions by simply multiplying by some A'.

This carries through to the entry value function, which is then $V_{\text{entry}}(b, A) = AV_{\text{entry}}(b, 1)$. With this, I define a cutoff for entry in productivity for each clean share type b, $\bar{A}(b)$:

$$\bar{A}(b) = \frac{f^e(b)}{V_{\text{entry}}(b, 1)}. \tag{15}$$

A plant with a draw (b, A) enters iff $A \geq \bar{A}(b)$.

I specify G_A as a Pareto distribution with scale parameter $A_{\min}$ and shape parameter γ. Only the relative scale between $A_{\min}$ and the scale of the entry cost f_0^e matters, so without loss of generality I normalize $A_{\min} = 1$. For now suppose that f_0^e is such that $\bar{A}(b) \geq A_{\min} = 1$ for all

b. This ensures that the selection mechanism is active for all levels of b. Then, the conditional distribution of $A|b$ among entering plants is a left-truncated version of G_A. In the case of a Pareto, the truncated distribution is also a Pareto with the same shape parameter γ and scale parameter equal to the point of truncation, $\bar{A}(b)$.

The density of the conditional distribution of entering plants' clean share types, $g_b(b|A \geq \bar{A}(b))$, is given by Bayes' rule (with a slight abuse of notation):

$$g_b(b|A \geq \bar{A}(b)) = \Pr(b|A \geq \bar{A}(b)) = \frac{\Pr(A \geq \bar{A}(b)|b)\Pr(b)}{\Pr(A \geq \bar{A}(b))}. \tag{16}$$

Since the distribution of A is Pareto with unit scale, $\Pr(A \geq \bar{A}(b)|b) = \left(\bar{A}(b)\right)^{-\gamma}$. Note that the denominator is the integral of the numerator over the support of b. It equals the share of potential entrants that do enter. The distribution of b among entering plants is then proportional to the ex-ante distribution, down-scaled by the entry cutoff at each b raised to the power of the Pareto shape parameter.

4.6 Illustration

To illustrate the interaction between the elasticity of substitution and adjustment costs, I simulate impulse response functions of energy use to a positive demand shock for different values of σ and ϕ_1. Figure 4 plots the results. The shock is a one standard deviation increase in the idiosyncratic price p_y, with persistence of 0.2.

In the case of no adjustment costs, $\phi_1 = 0$, both energy inputs respond equally to the shock, regardless of the elasticity of substitution. This is because the model effectively reduces to a sequence of static problems, and the magnitude of the response is determined by the output response to the demand shock only. The optimal ratio between both inputs is determined only by their relative prices, which remains constant. With adjustment costs, the responses differ: In general, dirty energy responds less than clean energy. For a given level of adjustment costs, a higher elasticity of substitution leads to a larger difference in responses. Vice versa, for a given elasticity of substitution, higher adjustment costs lead to a larger difference in responses. With an increasing adjustment cost parameter, the response of dirty energy becomes more muted, but also more persistent. It is this behavior that I exploit to separately identify the elasticity of substitution and adjustment costs in the calibration.

4.7 Adjustment costs in the estimation of the elasticity of substitution

Adjustment costs introduce a dynamic consideration into the plant's choice of energy inputs. Let $b = 1/2$ for simplicity, and consider the production function in equation (5)

$$E(E^c, E^d; b = 1/2) = \left[(E^c)^{\frac{\sigma-1}{\sigma}} + (E^d)^{\frac{\sigma-1}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}}.$$

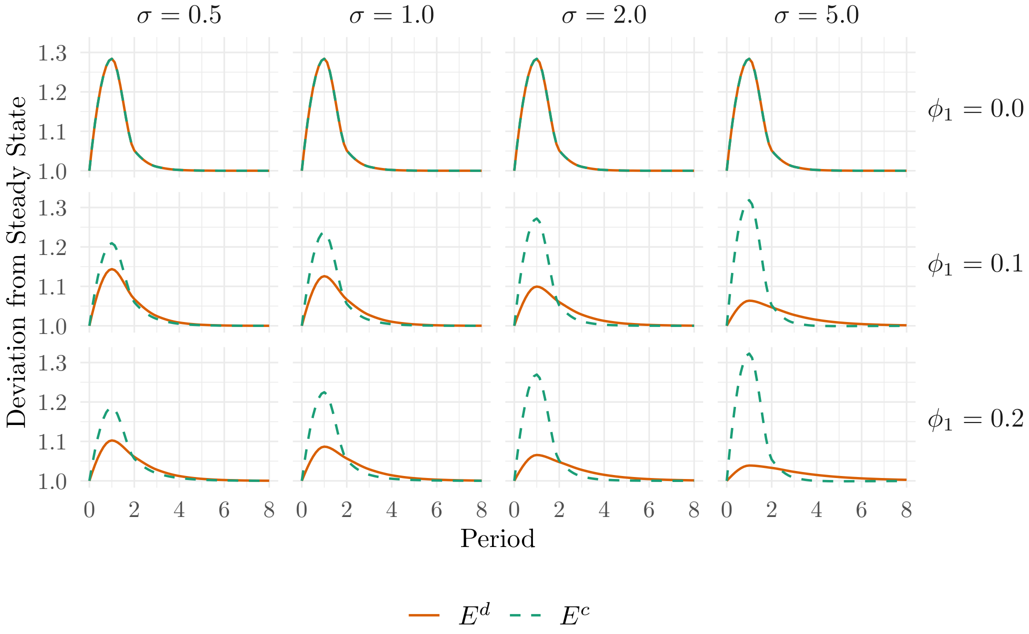


Figure 4: Smoothed impulse response functions of clean and dirty energy inputs E^c and E^d in response to a 1 standard deviation demand shock with a persistence of 0.2 in period 1. Without adjustment costs, $\phi_1 = 0$, the response of both inputs is the same: The model is effectively a sequence of static problems, and the magnitude of the response is determined by the output response to the demand shock only.

With input prices p_c and p_d, the first order conditions, and the optimal ratio of energy inputs in logs is

$$\begin{aligned} 0 &= \frac{\partial E(E^c, E^d)}{\partial E^c} - p_c, \\ 0 &= \frac{\partial E(E^c, E^d)}{\partial E^d} - p_d, \\ \log \frac{E^d}{E^c} &= \sigma \log \frac{p_c}{p_d}. \end{aligned}$$

This last equation is commonly used in the literature to estimate the elasticity of substitution between two inputs.

Consider now a simplified version of the dynamic model with adjustment costs,

$$V(E^d_{-1}, p_y) = \max_{E^c, E^d} p_y Y(E^c, E^d) - p_c E^c - p_d E^d - \phi(E^d_{-1}, E^d) + \beta \mathbb{E}_{p'_y|p_y} V(E^d, p'_y).$$

The first-order conditions for E^c and E^d are

$$
\begin{aligned}
0 &= p_y \frac{\partial Y(E^c, E^d)}{\partial E^c} - p_c, \\
0 &= p_y \frac{\partial Y(E^c, E^d)}{\partial E^d} - p_d - \frac{\partial \phi(E^d_{-1}, E^d)}{\partial E^d} + \beta \frac{\partial \mathbb{E} V(E^d, p'_y)}{\partial E^d}.
\end{aligned}
$$

Compared to the static case, there are two additional terms in the first-order condition for E^d.

Suppose for simplicity that p_y follows iid shocks, such that $\mathbb{E}V(E^d, p'_y) = \mathbb{E}V(E^d)$. Further suppose that the plant has had a series of identical shocks, such that it is at steady state with respect to its dirty energy use, $E^d_{-2} = E^d_{-1}$. If the plant is at steady state, then

$$
\left. \frac{\partial \mathbb{E} V(E^d)}{\partial E^d} \right|_{E^d = E^d_{-1}} = 0.
$$

Given concavity of the value function, any deviation $E^d \neq E^d_{-1}$ reduces the expected value. Additionally, the marginal adjustment cost is always positive for $E^d \neq E^d_{-1}$. Thus, at steady state, the plant chooses a value of E^d closer to E^d_{-1} than in the absence of adjustment costs. A plant reacts less to changing prices when there are adjustment costs.

In simulations, this seems to extend to the dynamic case with persistent shocks as well. There are no trends in the model, so on average plants are close to their steady state.

This has an important implication for the estimation of the elasticity of substitution. If there are adjustment costs and the estimation approach based on the static first-order conditions is used, the estimate will be biased downwards relative to the true elasticity of substitution. That estimate then represents a reduced-form parameter, which combines the true elasticity of substitution, the adjustment cost parameter, and the plant's expectations about the price trajectory. It is not necessarily informative about the true elasticity, which governs the potential for substitution to permanent price changes.

5 Identification and Calibration

I calibrate the parameters of the model to replicate empirical results from section 3. The scaling property of the value function allows me to separate the calibration of the dynamic within-plant part of the model, and the cross-sectional entry part.

5.1 Within-plant dynamics

Table 5 lists the parameters that govern the within-plant dynamics.

A crucial assumption is that b is fixed over time for each plant. This is necessary to identify the elasticity of substitution (Diamond et al., 1978).

Parameter	Description	Identified by
σ	Elast. of subst. between E^d, E^c	Ratio of factor demand elasticities
ϕ_1	Adj. cost parameter	Autocorrelation in $\Delta \log E^d$, $\Delta \log E^c$
ρ_y	Demand shock persistence	Autocorrelation in revenue
σ_{p_y}	Demand shock volatility	Within-firm variance in revenue
b	Energy share parameter	Distribution of energy mix
α	Returns to scale parameter	External
β	Discount factor	External
p_c, p_d	Energy prices	External

Table 5: Within-plant model parameters and identification

Elasticity of substitution and adjustment cost The difficulty is in disentangling the adjustment cost parameter and the elasticity of substitution. For transitory shocks, they have similar effects on the dynamics of dirty energy use: It could be that dirty energy is adjusted much less than clean energy either because there is a slight adjustment cost and they are highly substitutable, or the adjustment cost is very high, such that even though they are not very substitutable, it is not adjusted much.

One solution is to consider permanent shocks and long differences. Over an increasingly long horizon the adjustment cost becomes less relevant. This approach is infeasible in my case, since there is no persistent change in the relative price of clean to dirty energy over the sample period.

Instead, I separate the two by considering the dynamics of the changes of dirty energy use compared to clean energy use and output. Suppose there is some persistence in the demand shocks, $\rho_y \in (0,1)$, and that there is some adjustment cost, $\phi_1 > 0$. Then, compare the two limiting cases of σ in the CES aggregator of clean and dirty energy: In the Leontief case ($\sigma \to 0$), dirty energy use must adjust the same as clean energy use in relative terms, the plant must bear the adjustment cost if it wants to increase output. In contrast, in the perfect substitutes case ($\sigma \to \infty$), dirty energy use will not adjust at all, while clean energy use will adjust the same as output. The relationship is monotonic for intermediate values of σ; the relative response of dirty energy use to that of clean energy use and output varies between these two extremes.

Now consider the case of some $\sigma > 0$, and let the adjustment cost ϕ_1 vary. In the case of no adjustment cost, the plant will maintain the optimal energy mix, and adjust clean and dirty energy by the same relative amount. The persistence in both energy inputs is then the same as that of output. In the case of infinite adjustment cost, dirty energy use will not change at all. Dirty energy use will be perfectly persistent, while clean energy use will adjust in line with output. This relationship between the relative response of dirty energy use to that of clean energy use and output, and the level of ϕ_1 is also monotonic.

Combining these two arguments, for a given relative response of energy use to an output shock, there is a locus of combinations of σ and ϕ_1 that are consistent with it. The parameters are then separately identified from the persistence of the *changes* of the use of either energy: Conditional on the persistence of the demand shock, dirty energy use is more persistent; the plant will smooth the adjustment. In contrast, clean energy use exhibits lower persistence in its changes: it will be adjusted optimally, given the current level of dirty energy.

Demand process The demand shock parameters ρ_y and σ_y are identified from the autocorrelation and variance in plant-level revenue, conditional on the production function parameters.

Distribution of clean share parameters Assuming the plant is in a steady state, the optimal energy mix satisfies

$$\log \frac{E^d}{E^c} = \sigma \left[\log \frac{p_c}{p_d} + \log \frac{1-b}{b} \right]. \tag{17}$$

Conditional on the relative price and σ, the distribution of the clean share parameter b is identified from the distribution of the log of the energy mix in the data.

5.2 Cross-section

Table 6 lists the parameters that govern the entry and cross-sectional part of the model.

Parameter	Description	Identified by
G_A	Distr. of productivity	Observed marginal distr. of revenue
G_b	Distr. of energy share param.	Observed marginal distr. of energy mix
f_0^e	Entry cost scale	Normalized
f_1^e	Entry cost semi-elasticity w.r.t. b	Cross-sect. corr. between size and mix

Table 6: Across-plant model parameters and identification

Productivity distribution I specify G_A as Pareto with shape parameter γ and scale parameter $A_{\min} = 1$. From section 4.5 we have that a plant enters if its productivity exceeds the entry threshold $\bar{A}(b)$, given its draw for b. The distribution of entering plants' productivity is then a truncated Pareto, with shape parameter γ and scale parameter $\bar{A}(b)$.

The average revenue of a plant depends on its productivity A, as well as its clean share b, through the marginal cost. A result from the theory of regular variation is that since the between-plant differences in marginal cost from b are bounded, the tail index of the revenue distribution is the same as that of the productivity distribution. Therefore, I can calibrate γ directly to the revenue distribution, without the need to control for b.

I calibrate the shape parameter γ to match the top k sales shares of plants in the data: I residualize revenue by industry and year fixed effects, then, for $k \in \{3, 4, 5, 10, 15, 20, 50\}$ I calculate the share of the residualized revenue accounted for by the top k plants in year. The shares are stable over time, and I take the average over years as the target moments.

Clean share distribution I specify the distribution G_b as Beta with shape parameters α_b, β_b. Conditional on $\bar{A}(b)$ and γ, the ex-ante distribution G_b is identified from the ex-post distribution of clean shares, as in equation (16). I fit a Beta distribution to the observed ex-post density of the distribution of clean shares in the data, $g_b(b|A \geq \bar{A}(b))$. The ex-ante distribution is then given by rearranging equation (16):

$$g_b(b) \propto g_b(b|A \geq \bar{A}(b))\bar{A}(b)^{\gamma}. \tag{18}$$

If $G_b(b|A \geq \bar{A}(b))$ is Beta, and given the scaling term $\bar{A}(b)^{-\gamma}$, the density g_b is not exactly Beta in general. The Beta distribution is a good approximation of the observed ex-post distribution, and assuming a Beta for the ex-ante distribution leads to a simulated sample that can be well-approximated by a Beta as well.

Entry cost The size of the pool of potential entrants is not identified, and neither is the scale of the entry cost f_0^e. I normalize the size of the pool of potential entrants, and make the assumption that the scale of the entry cost is such that $\bar{A}(b) \geq 1 = A_{\min}$ for all $b \in (0, 1)$. This ensures that the selection mechanism is active for all clean share types.

The semi-elasticity of the entry cost with respect to the clean share, f_1^e, is identified from the relationship between plants' clean share and their size in the data. The target moment is the regression coefficient in equation (2), the cross-sectional energy mix-size correlation.

5.3 Target moments

Table 7 lists the empirical target moments, their standard errors, and the corresponding simulated moments. The loss function is the weighted sum of squared deviations between simulated and empirical moments, with weights given by the inverse of the squared standard errors.

Description	Value	SE	Simulated
Ratio of fact. dem. elast. dirty/clean, $\Delta t = 1$	0.26	0.05	0.42
Ratio of fact. dem. elast. dirty/clean, $\Delta t = 2$	0.38	0.05	0.48
Autoregressive coef. $\Delta \log E^d$	0.18	0.02	0.28
Autoregressive coef. $\Delta \log E^c$	0.11	0.03	0.00
Autoregressive coef. $\Delta \log p_y Y$	0.16	0.03	0.04
Std. dev. revenue	0.08	0.04	0.10

Table 7: Calibration targets, their standard errors, and simulated moments

5.3.1 Moment calculation

Ratio of factor demand elasticities The regression equations correspond to their empirical counterpart, equation (3), without the fixed effects:

$$\Delta_k \log E^d_{it} = \beta^d_k \Delta_k \log Y_{it} + \epsilon^d_{it},$$
$$\Delta_k \log E^c_{it} = \beta^c_k \Delta_k \log Y_{it} + \epsilon^c_{it}.$$

I estimate these with 2SLS like the empirical counterpart, using the true simulated demand shock as instruments for the change in output. I then take the ratio β^d_k/β^c_k as the target moment, for $k \in \{1, 2\}$.

Persistence in changes I estimate the persistence in the changes of energy use and output from the following autoregressive equations:

$$\Delta \log E^d_{it} = \rho_{E^d} \Delta \log E^d_{i,t-1} + \epsilon^d_{it},$$
$$\Delta \log E^c_{it} = \rho_{E^c} \Delta \log E^c_{i,t-1} + \epsilon^c_{it},$$
$$\Delta \log[p_{y,it} Y_{it}] = \rho_Y \Delta \log[p_{y,it} Y_{i,t-1}] + \epsilon^Y_{it}.$$

These estimation equations suffer from Nickell bias, so I instrument for the lagged difference using the level of the second lag:

$$\Delta \log E^d_{i,t-1} = \tilde{\beta}_d \log E^d_{i,t-2} + \tilde{\epsilon}^d_{it},$$

and similar for clean energy and output.

Standard deviation of revenue I calculate the standard deviation of within-plant revenue as the mean squared residual of the regression of the persistence in revenue, ϵ^Y_{it}.

Cross-sectional energy mix-size correlation I estimate the cross-sectional energy mix-size correlation from the following regression equation:

$$\log \frac{E^d_{it}}{E^c_{it}} = \beta_{\text{size}} \log\left[p_{y,it} Y_{it}\right] + \epsilon_{it}.$$

5.4 Implementation

5.4.1 Within-plant model

First, I solve the plant's dynamic programming problem by value function iteration over a three-dimensional grid of E^d_{-1}, p_y, and b.[8] The results are the value function, policy functions for E^d,

[8] I discretize the process for p_y using the Rouwenhorst method.

and optimal choices for E^c, given the policy for E^d. I simulate a panel of plants, and calculate the relevant moments from the simulated data, using the following algorithm:

1. Set the number of firms $N = 30,000$ and time periods $T = 120$ (the first 100 are burn-in), to match the empirical panel dimensions.
2. Draw N clean share types $\{b_i\}$ from the grid, such that the distribution matches the observed empirical distribution.[9]
3. Generate N discretized demand shock processes $\{p_{y,it}\}$ of length T each.
4. Initialize $\{E^d_{it=0}\}$ as the steady state level given b_i and $p_{y,it=0} = 1$ (The steady state is the fixed point in the policy function for E^d).
5. Solve the model forward for each plant i and time $t = 1, \ldots, T$:
   (a) Given $E^d_{i,t-1}$, $p_{y,it}$, and b_i, interpolate the policy functions to get optimal E^d_{it} and E^c_{it}.
   (b) Calculate output Y_{it}.
   (c) Go to the next period.
6. Discard the first 100 periods as burn-in.
7. Calculate target moments from the simulated data, as described in section 5.3.1.

To calibrate the parameters, I first run a coarse grid search over the parameter space. Then, with the best two results from the grid search as initial values, I run a simulated annealing algorithm, subject to parameter bounds. The algorithm converges to a point well within the parameter bounds.

5.4.2 Cross-sectional model

For the cross-section, I need a sample of types (b, A) from the conditional distribution of entering plants. Given the sample, I then simulate the dynamic decisions of the plant, as in the within-plant calibration.

First, I generate a sample of the conditional distribution $(b, A)|A > \bar{A}(b)$ using a two-stage accept-reject algorithm. Note that $\bar{A}(b)$ depends on the entry cost parameters. I normalize the scale f_0^e such that $\inf_b \bar{A}(b) = 1.1 > A_{\min} = 1$, to ensure that there is selection at all values of b. Then, I calibrate f_1^e to match the energy mix-size correlation in the data. Conditional on a candidate value for f_1^e, and the calibration of the within-plant parameters, I generate a sample as follows:

[9] I match the empirical distribution of $\log E^d/E^c$ as follows. Suppose there is no adjustment cost, $\phi_1 = 0$. Then, from the optimal input mix equation, $\log E^d/E^c = \sigma[\log p_c/p_d + \log(1-b)/b]$. In the data, $\log E^d_{it}/E^c_{it} \sim \mathcal{N}(0, 1.1^2)$. Assuming the relative price is constant, this implies that b follows a logit-normal distribution, with variance depending on the elasticity of substitution σ. I draw sampling weights for each grid value of b, such that the weighted distribution of b in the sample matches the implied logit-normal distribution.

1. Approximate the observed conditional distribution of b as a logit-normal distribution, as in the within-plant calibration, with mean $\log p^d/p^c$ and standard deviation $1.1/\sigma$, where σ is the calibrated elasticity of substitution. This is the proposal distribution for b.

2. Draw a value from the proposal distribution for b. Accept the draw with a probability proportional to $\bar{A}(b)^\gamma$. This gives a sample from the unconditional distribution of b.

3. Pair the value of b with a draw for A from the unconditional Pareto distribution. Accept the pair if $A \geq \bar{A}(b)$. The result is a sample from the conditional distribution $(b, A)|A \geq \bar{A}(b)$.

4. Repeat until the sample size is $N = 30,000$.

Then, given the sample of types, I simulate the dynamic decisions of each plant as in the within-plant calibration in section 5.4.1. To incorporate the productivity draws, I multiply the value and policy functions by A_i for each plant i: I adjust the state variable $E^d_{-1,it} = A_i \tilde{E}^d_{-1,it}$, the policy function $E^d_{it} = A_i \tilde{E}^d_{it}$, the optimal choice of $E^c_{it} = A_i \tilde{E}^c_{it}$, and the value function $V_{it} = A_i \tilde{V}_{it}$, where the tilde variables are from the solution for $A = 1$. Output is then $Y_{it} = A_i^{1-\alpha} E(E^c_{it}, E^d_{it}; b_i)^\alpha$.

I calculate the target moment, the energy mix-size correlation, from the simulated data as in section 5.3.1. Conditional on the calibration of the within-plant parameters, and on the directly calibrated Pareto shape parameter, I calibrate f^e_1 using a univariate bounded minimization algorithm, to minimize the distance between the simulated and empirical moments.

Finally, to calibrate the unconditional distribution of b, I parameterize a Beta distribution as follows. For each draw of b in the simulated sample, I calculate the weight $w(b) = \bar{A}(b)^\gamma$. Then, I calculate the weighted mean and variance of b in the sample, and solve for the Beta shape parameters α_b, β_b using the method of moments formulas.

5.5 Results

Table 8 presents the calibrated parameter values. The externally calibrated parameters are $\alpha = 0.65$ (which would correspond to a demand elasticity of 2.9 in a monopolistic competition model with constant returns to scale production), $\beta = 0.96$ for annual data, and energy prices $p_c = 0.5$ and $p_d = 0.18$. The level of the prices is subject to normalization, and the relative price $p_c/p_d = 2.8$ corresponds to the average aggregate relative price over the sample period.

The calibrated value for the adjustment cost parameter is such that in the simulated panel, the adjustment cost represent on average 0.1% of total costs. This small share is sufficient though to generate significant differences the response of dirty energy use, compared to clean in response to output shocks. The elasticity of substitution between clean and dirty energy is calibrated to 5.1, which is on the higher end of estimates in the literature. I show in section 4.7 that the common estimation approach underestimates the deep elasticity of substitution when there are adjustment costs. This can explain my relatively high estimate.

Parameter	Description	Calibrated value
σ	Elast. of subst. between E^d, E^c	5.10
ϕ_1	Adj. cost parameter	0.02
ρ_y	Demand shock persistence	0.01
σ_{p_y}	Demand shock volatility	0.13
G_A	Distr. of productivity	Pareto($\gamma = 1.38, A_{\min} = 1$)
G_b	Distr. of energy share param.	Beta($\alpha_b = 90, \beta_b = 32$)
f_0^e	Entry cost scale	17.55
f_1^e	Entry cost semi-elasticity w.r.t. b	0.38
α	Returns to scale parameter	0.65
β	Discount factor	0.96
p_c, p_d	Energy prices	0.5, 0.18

Table 8: Parameters and calibrated values

The calibrated value for the Pareto tail parameter is in line with estimates from the literature (it implies somewhat thinner tails than Zipf's law, as estimated in Axtell, 2001; Gabaix, 2016). The distribution of clean share parameters has a mean of 0.74, and a standard deviation of 0.04. The entry cost semi elasticity parameter is such that a 1 p.p. increase in b leads to a 0.38% increase in the entry cost.

Untargeted moments I match an untargeted moment: the within-plant over-time correlation between energy mix and revenue. The empirical correlation is -0.16, the simulated correlation is -0.26. The remaining discrepancy might be related to attenuation bias in the empirical estimate due to measurement error in revenue. This moment is closely related to the elasticity of substitution and adjustment cost parameters, providing some additional validation for their calibration.

6 Policy Experiments

I conduct two policy experiments in the parameterized model. First, a subsidy for the entry cost of clean plants. Second, a tax on dirty energy purchases, similar to a carbon tax.

6.1 Simplified aggregation

I will conduct steady-state comparisons for the policy experiments. Thus, I will abstract from the dynamics of the model, in particular the demand shocks, and focus on the steady-state aggregates.

Let $E^{d\star}(b, A)$ be the steady-state value of dirty energy consumption for a plant with type (b, A), defined as the fixed point of the policy function at the unconditional mean of the demand shock, $p_y = 1$. Let $E^{c\star}(b, A)$ be the corresponding optimal steady-state clean energy consumption. The aggregates of clean and dirty energy consumption, and output are then given by

$$\begin{aligned}
E^c_{\text{agg}} &= \int_0^1 \int_{\bar{A}(b)}^{\infty} E^{c\star}(b,A) g_b(b) g_A(A) dA db, \\
E^d_{\text{agg}} &= \int_0^1 \int_{\bar{A}(b)}^{\infty} E^{d\star}(b,A) g_b(b) g_A(A) dA db, \\
Y_{\text{agg}} &= \int_0^1 \int_{\bar{A}(b)}^{\infty} A^{1-\alpha} E(E^{c\star}(b,A), E^{d\star}(b,A))^{\alpha} g_b(b) g_A(A) dA db.
\end{aligned}$$

Given the scaling property of the value and policy functions, and the Pareto distribution of A, these simplify to (see appendix section A.2):

$$E^c_{\text{agg}} = \frac{\gamma}{\gamma-1} \int_0^1 E^{c\star}(b,1) \bar{A}(b)^{1-\gamma} dG_b(b) \tag{19}$$

$$E^d_{\text{agg}} = \frac{\gamma}{\gamma-1} \int_0^1 E^{d\star}(b,1) \bar{A}(b)^{1-\gamma} dG_b(b) \tag{20}$$

$$Y_{\text{agg}} = \frac{\gamma}{\gamma-1} \int_0^1 E\left(E^{c\star}(b,1), E^{d\star}(b,1)\right)^{\alpha} \bar{A}(b)^{1-\gamma} dG_b(b). \tag{21}$$

Recall the definition of the cutoff productivity $\bar{A}(b)$ from equation (15):

$$\bar{A}(b) = \frac{f^e(b)}{V_{\text{entry}}(b,1)}.$$

Define the mass of entering plants as

$$m = \int_0^1 \int_{\bar{A}(b)}^{\infty} g_b(b) g_A(A) dA db = \int_0^1 \bar{A}(b)^{-\gamma} dG_b(b). \tag{22}$$

6.2 Entry subsidy for clean plants

Suppose a policymaker observes the (b, A) draws of the potential entrants. They offer a multiplicative subsidy $s(b) = \exp(-s_1 b)$ for the entry of clean plants. The entry costs are then

$$f^{e,\text{sub}}(b) = f^e(b) s(b) = f^e_0 \exp[b(f^e_1 - s_1)], \tag{23}$$

where s_1 represents the semi-elasticity of the subsidy with respect to the clean share. For values of s_1 up to f^e_1, the policymaker subsidizes part of the additional entry cost for a higher clean share. For values of s_1 larger than f^e_1, the entry cost to be paid by the plant decreases in the clean share.

Under the subsidy, the entry cutoff in productivity becomes

$$\bar{A}^{\text{sub}}(b) = \frac{f^{e,\text{sub}}(b)}{V_{\text{entry}}(b,1)} = \frac{f_0^e}{V_{\text{entry}}(b,1)} \exp[b(f_1^e - s_1)]. \tag{24}$$

The subsidy influences aggregates only through changes in the entry cutoff.

The total cost of the subsidy to the policymaker is

$$C^{\text{sub}} = \int_0^1 \int_{\bar{A}^{\text{sub}}(b)}^{\infty} f^e(b)[1-s(b)] g_A(A) g_b(b) dA db = \int_0^1 f^e(b)[1-s(b)] \bar{A}^{\text{sub}}(b)^{-\gamma} dG_b(b). \tag{25}$$

Figure 5 shows the results of the entry subsidy policy experiment. I consider values $s_1 \in (0, 1.2 \times f_1^e)$: between no subsidy and a subsidy that fully reverses the clean share dependence of the entry cost, such that plants with a higher clean share have a lower entry cost.

The policy is expansionary: it increases the mass of entering plants, output, and both clean and dirty energy use. It increases the aggregate share of clean energy, but only minutely so. The policy acts only at the entry margin by construction. Given the Pareto distribution of productivity, aggregates are primarily driven by the most productive plants, which are not affected by the policy.

The policy is expensive: I plot the total cost of the subsidy as fraction of aggregate output.

6.3 Tax on dirty energy use

I consider a proportional tax τ_d on the price of dirty energy. The final price is then

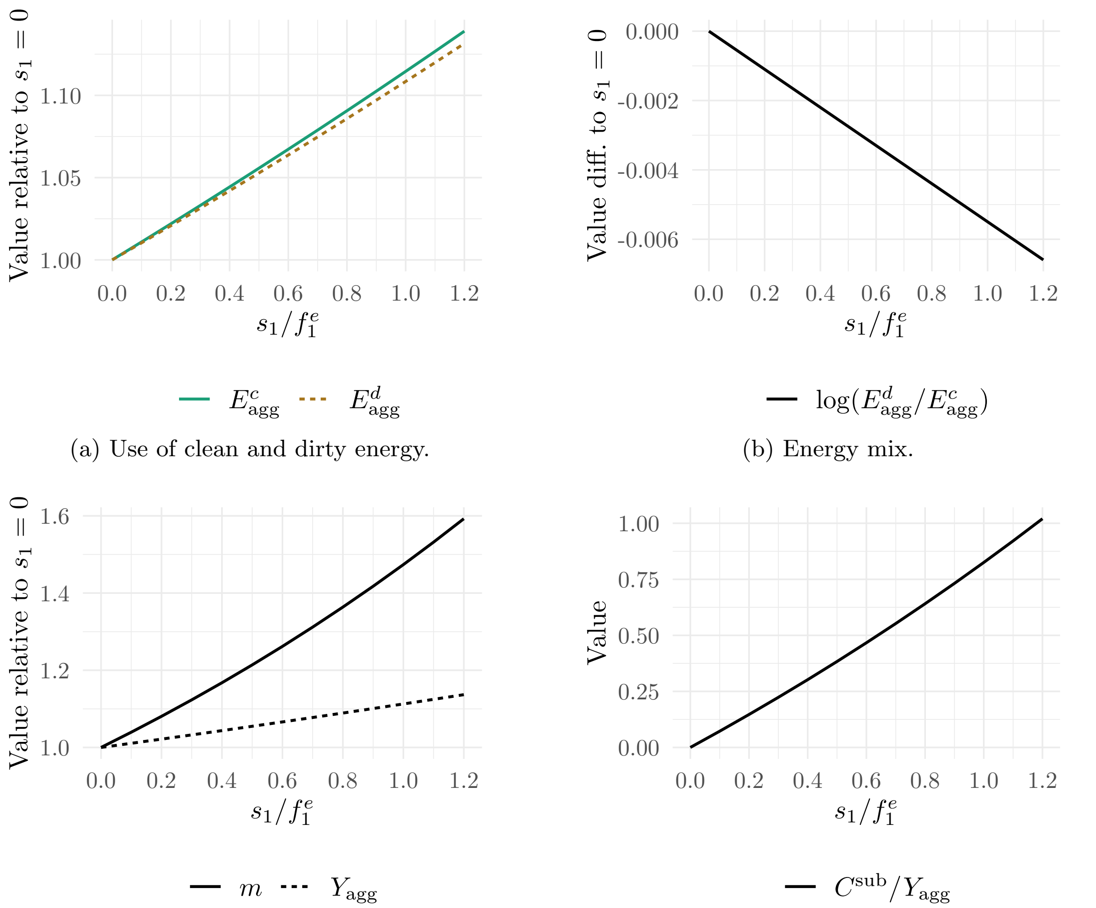
$$p_d^{\text{tax}} = p_d(1+\tau_d). \tag{26}$$

The tax is permanent and must be paid by all plants. I consider values $\tau_d \in (0,1)$, so up to a doubling of the dirty energy price.

I solve for plants' problem under the new price including the tax, and then calculate aggregates for two cases: (i) holding the distribution of plants over (b, A) fixed at the baseline equilibrium, and (ii) allowing the distribution to adjust to the equilibrium under the new prices. Case (i) can be interpreted as the short-run effects of the tax, while case (ii) represents the long-run effects with compositional changes.

The results of the tax policy experiment allowing for compositional changes are shown in figure 6. As opposed to the entry subsidy, the tax on dirty energy use is contractionary, since it strictly increases marginal cost. The policy is very effective though at reducing the aggregate use of dirty energy.

The results for the case with the fixed distribution are shown in appendix section C.1. They are similar quantitatively, the main difference being a smaller reduction in dirty energy use, and a smaller decrease in aggregate output, since the mass of firms remains constant.



(a) Use of clean and dirty energy.

(b) Energy mix.

(c) Total output and mass of entering plants.

(d) Total subsidy spending as fraction of output.

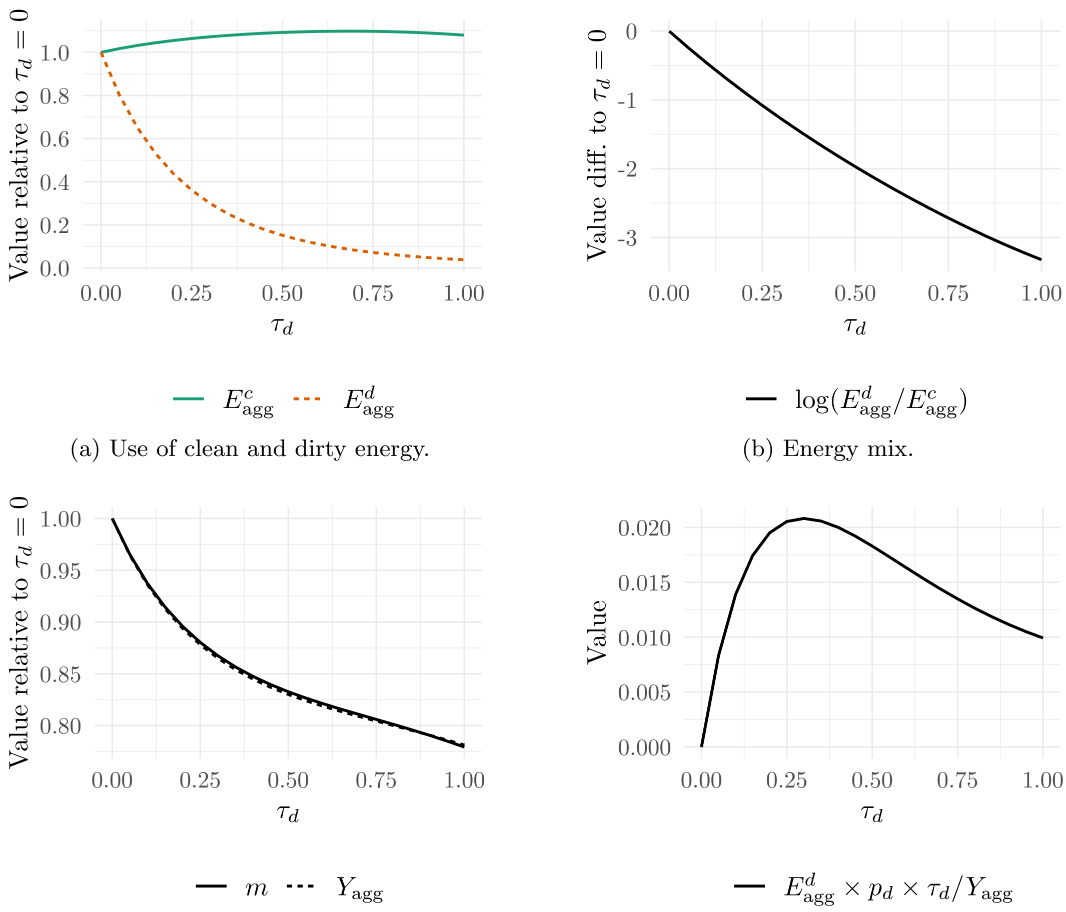
Figure 5: Results of entry subsidy policy experiment.

6.3.1 Comparison to low-σ calibration

To demonstrate the impact of the value of the elasticity of substitution between clean and dirty energy, I repeat the policy experiment in a model with a different calibration. I decrease the value of the elasticity σ from 5 to 1.5 and the value of the adjustment cost parameter ϕ_1 to 0, keeping all other parameters fixed. The value 1.5 corresponds to the upper end of estimates of the micro-level elasticity in the literature. I solve the value and policy functions for the new parameterization, and compute the new steady-state aggregates under the tax on dirty energy use allowing for compositional changes as above. Figure 7 shows the results.

The results differ substantially. In the low-elasticity case, the tax is both less effective at reducing the consumption of dirty energy, and leads to a larger contraction in output.

In the baseline calibration, a 20% reduction in aggregate dirty energy use is achieved at a tax of 7%, leading an output reduction of 2%. In the low-elasticity calibration, achieving the same



(a) Use of clean and dirty energy.

(b) Energy mix.

(c) Total output and mass of entering plants.

(d) Tax revenue as fraction of output.

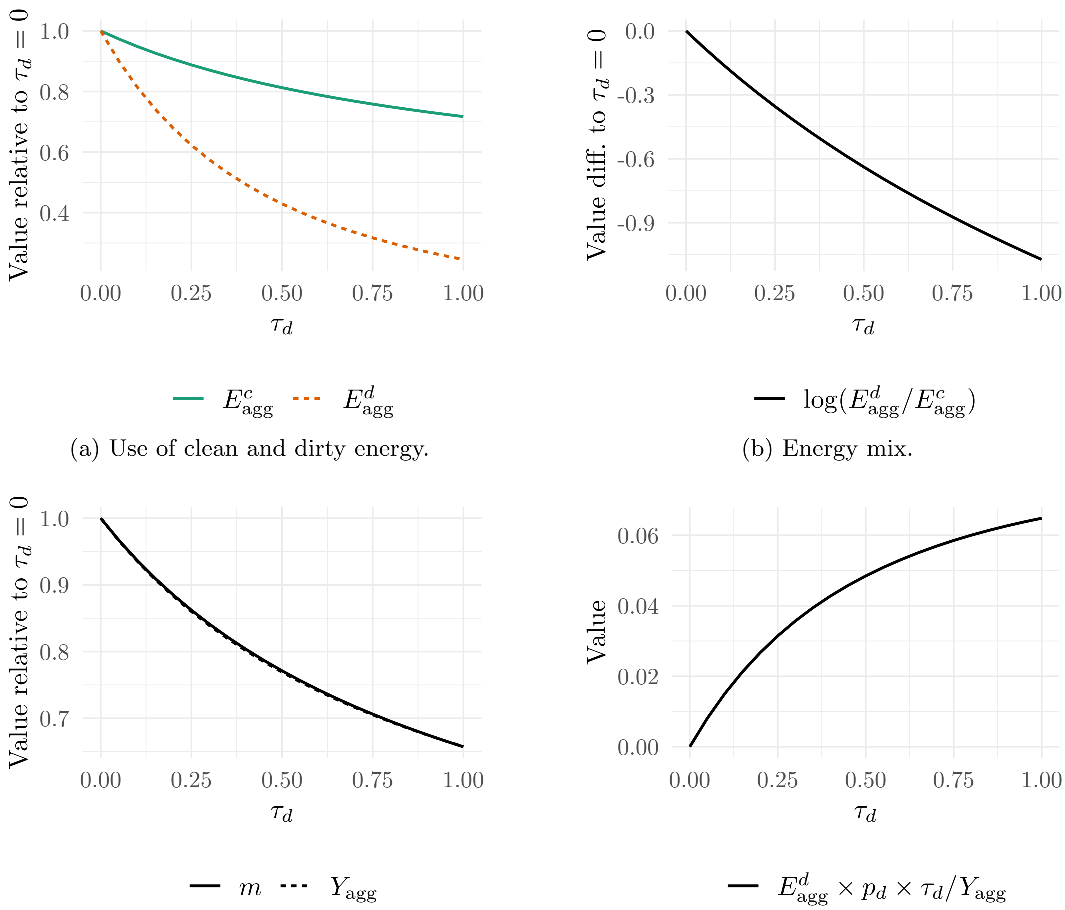
Figure 6: Results of tax on dirty energy use policy experiment, allowing for compositional changes.

20% reduction in dirty energy use requires a tax of 12%, leading to an output reduction of 7%.

7 Conclusion

In this paper, I analyze the use of clean and dirty energy in production in the German manufacturing census. I document empirically, that (i) there is substantial heterogeneity in the energy mix across plants, (ii) this heterogeneity is difficult to explain with observables, (iii) some variation can be explained by size, larger plants use a higher share of clean energy, and (iv) the factor demands for clean and dirty energy respond differently to demand shocks, dirty energy use adjusts much less than clean to the same shock to output.

Based on these findings, I develop a dynamic model of heterogeneous plants with entry. Plants differ in productivity and in their technology to combine clean and dirty energy to produce



(a) Use of clean and dirty energy.

(b) Energy mix.

(c) Total output and mass of entering plants.

(d) Tax revenue as fraction of output.

Figure 7: Results of tax on dirty energy use policy experiment with low-σ calibration, allowing for compositional changes.

output. A key ingredient of the model to match the data is an adjustment cost for dirty energy use. The model replicates both the within-plant dynamics, and the cross-sectional distribution of plants. I show that estimates of the elasticity of substitution between clean and dirty energy are downward biased if they ignore the presence of adjustment costs. My calibration of the model implies an elasticity of substitution of around 5, which is three times larger than estimates in the literature.

In the calibrated model, I study two policies: an entry subsidy for clean plants, and a tax on dirty energy use. The entry subsidy has only marginal effects: Aggregate use of clean and dirty energy is dominated by large plants, which are unaffected by the policy. A tax on dirty energy use is effective at reducing dirty energy use. I show that the effectiveness of the tax depends strongly on the elasticity of substitution between clean and dirty energy. Given the higher elasticity in my calibration, the tax has much lower contractionary effects on output for a given reduction in dirty energy use, compared to a model with an elasticity closer to estimates in the literature.

References

Acemoglu, D., Aghion, P., Bursztyn, L., & Hémous, D. (2012). The environment and directed technical change. *American Economic Review*, *102*(1), 131–66.

Acemoglu, D., Akcigit, U., Hanley, D., & Kerr, W. (2016). Transition to clean technology. *Journal of Political Economy*, *124*(1), 52–104.

Andersson, J. J. (2019). Carbon taxes and CO2 emissions: Sweden as a case study. *American Economic Journal: Economic Policy*, *11*(4), 1–30.

Antràs, P. (2004). Is the U.S. aggregate production function Cobb-Douglas? New estimates of the elasticity of substitution. *Contributions in Macroeconomics*, *4*(1), 20121005.

Arbeitsgemeinschaft Energiebilanzen e.V. (2025). *Evaluation Tables on the Energy Balance for the Federal Republic of Germany 1990 to 2024* [https://ag-energiebilanzen.de/en/data-and-facts/evaluation-tables-on-the-energy-balance/, accessed 02.10.2025].

Arnberg, S., & Bjørner, T. B. (2007). Substitution between energy, capital and labour within industrial companies: A micro panel data analysis. *Resource and Energy Economics*, *29*(2), 122–136.

Axtell, R. L. (2001). Zipf distribution of U.S. firm sizes. *Science*, *293*(5536), 1818–1820.

Barrot, J.-N., & Sauvagnat, J. (2016). Input specificity and the propagation of idiosyncratic shocks in production networks. *The Quarterly Journal of Economics*, *131*(3), 1543–1592.

Boehm, C. E., Levchenko, A. A., & Pandalai-Nayar, N. (2023). The long and short (run) of trade elasticities. *American Economic Review*, *113*(4), 861–905.

Bontadini, F., Corrado, C., Haskel, J., Iommi, M., & Jona-Lasinio, C. (2023, February). *EU-KLEMS & INTANProd: Industry productivity accounts with intangibles - sources of growth and productivity trends: Methods and main measurement challenges* (Deliverable No. D2.3.1). Luiss Lab of European Economics.

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

Casey, G. (2024). Energy efficiency and directed technical change: Implications for climate change mitigation. *The Review of Economic Studies*, *91*(1), 192–228.

Chirinko, R. S. (2008). σ: The long and short of it. *Journal of Macroeconomics*, *30*(2), 671–686.

Colmer, J., Martin, R., Muûls, M., & Wagner, U. J. (2024). Does pricing carbon mitigate climate change? Firm-level evidence from the European Union Emissions Trading System. *The Review of Economic Studies*, rdae055.

Diamond, P., McFadden, D., & Rodriguez, M. (1978). Measurement of the elasticity of factor substitution and bias of technical change. In M. Fuss & D. McFadden (Eds.), *Applications of the theory of production* (pp. 125–147, Vol. 2). Elsevier.

Eslava, M., Haltiwanger, J., Kugler, A., & Kugler, M. (2004). The effects of structural reforms on productivity and profitability enhancing reallocation: Evidence from Colombia. *Journal of Development Economics*, *75*(2), 333–371.

Fraunhofer ISI. (2025). *Erstellung von Anwendungsbilanzen für die Jahre 2021 bis 2023 für die Sektoren Industrie und GHD* [https://ag-energiebilanzen.de/wp-content/uploads/2024/11/Anwendungsbilanz_Industrie_2023_final_20250324.pdf, accessed 02.10.2025].

Gabaix, X. (2016). Power laws in economics: An introduction. *Journal of Economic Perspectives*, *30*(1), 185–206.

Gerster, A., & Lamp, S. (2024). Energy tax exemptions and industrial production. *The Economic Journal*.

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

Hyland, M., & Haller, S. (2018). Firm-level estimates of fuel substitution: An application to carbon pricing. *The Energy Journal*, *39*(6), 71–98.

Jo, A. (2024). Substitution between clean and dirty energy with biased technical change. *International Economic Review*.

Kaartinen, M., & Prane, O. (2024, November). *Carbon pricing and fuel switching by firms: Theory and evidence* (Job Market Paper). Department of Economics, Stockholm University.

Leclair, E. M. (2025, January). *Balancing production and carbon emissions with fuel substitution* (Working Paper). University of Chicago.

Marschak, J., & Andrews, W. H. (1944). Random simultaneous equations and the theory of production. *Econometrica*, *12*(3/4), 143–205.

Martin, R., de Preux, L. B., & Wagner, U. J. (2014). The impact of a carbon tax on manufacturing: Evidence from microdata. *Journal of Public Economics*, *117*, 1–14.

Oberfield, E., & Raval, D. (2021). Micro data and macro technology. *Econometrica*, *89*(2), 703–732.

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

Peter, A., & Ruane, C. (2025, May). *The aggregate importance of intermediate input substitutability* (Working Paper No. 31233). NBER.

Research Data Centre of the Statistical Offices of the Federal States. (2023). *AFiD Official Firm Data for Germany* [Modules: *Produkte* (doi:10.21242/42131.2021.00.03.1.1.0), *Energieverbrauch* (doi:10.21242/43531.2021.00.03.1.1.0), *Industriebetriebe* (doi:10.21242/42111.2021.00.01.1.1.0), and *Industrieunternehmen* (doi:10.21242/42221.2021.00.01.1.1.0)].

Stern, D. I. (2012). Interfuel substitution: A meta-analysis. *Journal of Economic Surveys*, *26*(2), 307–331.

A Derivations

A.1 Bellman equation scaling property

Static case Start with the static case, without adjustment cost and demand shocks. The problem is

$$\begin{aligned} V(A,b) &= \max_{E^c,E^d} p_y Y(E^c,E^d;A,b) - p_c E^c - p_d E^d, \\ Y(E^c,E^d;A,b) &= A^{1-\alpha} E(E^c,E^d;b)^{\alpha}, \\ E(E^c,E^d;b) &= \left[b^{\frac{1}{\sigma}} (E^c)^{\frac{\sigma-1}{\sigma}} + (1-b)^{\frac{1}{\sigma}} (E^d)^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}} N(b). \end{aligned}$$

The claim in this case is

$$V(A,b) = AV(1,b).$$

From above, we have that

$$\begin{aligned} V(b,A) = \pi^{\star}(b,A) &= A(1-\alpha) p_y^{\frac{1}{1-\alpha}} \left[\frac{\alpha}{p_E(b)} \right]^{\frac{\alpha}{1-\alpha}}, \\ p_E(b) &= \left[b^{\sigma} p_c^{1-\sigma} + (1-b)^{\sigma} p_d^{1-\sigma} \right]^{\frac{1}{1-\sigma}}. \end{aligned}$$

Clearly, this scales linearly in A. In optimum,

$$\begin{aligned} Y^{\star}(b,A) &= A \left[\frac{p_y \alpha}{p_E(b)} \right]^{\frac{\alpha}{1-\alpha}} = A^{1-\alpha} E^{\alpha}, \\ \iff E^{\star}(b,A) &= A \left[\frac{p_y \alpha}{p_E(b)} \right]^{\frac{1}{1-\alpha}}. \end{aligned}$$

Factor demand for energy services scales linearly in A. Either type of energy scales linearly in E, and thus in A.

Dynamic case The Bellman equation is

$$\begin{aligned} V(E^d_{-1}, p_y; b, A) &= \max_{E^c,E^d} p_y A^{1-\alpha} F(E^c,E^d;b) - p_c E^c - p_d E^d - \phi(E^d_{-1}, E^d) + \beta \mathbb{E}_{p'_y|p_y} V(E^d, p_y; b, A), \\ \phi(E^d_{-1}, E^d) &= \frac{\phi_1}{2E^d_{-1}} \left(E^d - E^d_{-1} \right)^2, \\ \log p'_y &= \rho \log p_y + \sigma_{p_y} \epsilon', \ \epsilon' \sim N(0,1). \end{aligned}$$

The claim is

$$V(E_{-1}^d, p_y; b, A) = AV\left(A^{-1}E_{-1}^d, p_y; b, 1\right).$$

Consider first the felicity function:

$$p_y A^{1-\alpha} F(E^c, E^d; b) - p_c E^c - p_d E^d - \frac{\phi_1}{2E_{-1}^d}\left(E^d - E_{-1}^d\right)^2.$$

If we can factor out A after adjusting the state variable E_{-1}^d by A^{-1}, we show that the felicity function scales linearly in A. So, what must μ be such that

$$\begin{aligned}
&p_y A^{1-\alpha} F(E^c, E^d; b) - p_c E^c - p_d E^d - \frac{\phi_1}{2E_{-1}^d}\left(E^d - E_{-1}^d\right)^2 \\
&= A\left[p_y F(\mu E^c, \mu E^d; b) - p_c \mu E^c - p_d \mu E^d - \frac{\phi_1}{2E_{-1}^d A^{-1}}\left(\mu E^d - E_{-1}^d A^{-1}\right)^2\right]?
\end{aligned}$$

The production function is homogeneous of degree α:

$$F(\mu E^c, \mu E^d; b) = \mu^\alpha F(E^c, E^d; b).$$

$\mu = A^{-1}$ satisfies the equation. This scales the within-period factor demands for E^c, and critically for E^d, which becomes the state variable in the next period. The scaling is consistent with the proposed adjustment of the state variable. This concludes the proof.

A.2 Aggregation

I present the aggregation for clean energy use. The steps are analogous for dirty energy use and output.

$$\begin{aligned}
E_{\text{agg}}^c &= \int_0^1 \int_{\bar{A}(b)}^\infty E^{c\star}(b, A) g_b(b) g_A(A) dA db \\
&= \int_0^1 \int_{\bar{A}(b)}^\infty A E^{c\star}(b, 1) g_b(b) g_A(A) dA db \\
&= \int_0^1 E^{c\star}(b, 1) g_b(b) \left(\int_{\bar{A}(b)}^\infty A g_A(A) dA\right) db
\end{aligned}$$

$$\begin{aligned}
\int_{\bar{A}(b)}^\infty A g_A(A) dA &= \frac{\gamma}{\gamma - 1} \bar{A}(b)^{1-\gamma} \\
\implies E_{\text{agg}}^c &= \frac{\gamma}{\gamma - 1} \int_0^1 E^{c\star}(b, 1) \bar{A}(b)^{1-\gamma} dG_b(b) \\
\implies E_{\text{agg}}^c &= \frac{\gamma}{\gamma - 1} \int_0^1 E^{c\star}(b, 1) \left(\frac{f^e(b) s(b)}{V_{\text{entry}}(b, 1)}\right)^{1-\gamma} dG_b(b)
\end{aligned}$$

B Additional Empirical Results

B.1 Variation

Sample	Fixed Effects	N	adj. R^2	BIC
Full	Industry (2)	600417	0.14	1816950
Full	Industry (2) by Year	600417	0.15	1818253
Full	Industry (4)	600417	0.27	1721316
Full	Industry (4) by Year	600417	0.28	1768868
Single Product	Product	281380	0.35	810365
Single Product	Product by Year	281380	0.37	909941
Full Geo	District (5)	600404	0.03	1893801
Full Geo	District (5) + Industry (4)	600404	0.28	1714386
Full Geo	District (5) by Year	600404	0.04	1963530
Full	Plant	600417	0.87	1537935

Table A.1: Adjusted R^2, number of observations, and BIC for regressions of energy mix $\log E^d_{it}/E^c_{it}$ on different sets of fixed effects. The BIC is only comparable within a sample due to different N. It penalizes models with more parameters, i.e., more fixed effects. In the full sample, the plant fixed effect model has the lowest BIC, indicating that its high explanatory power is not due to overfitting.

B.2 Factor demand elasticity

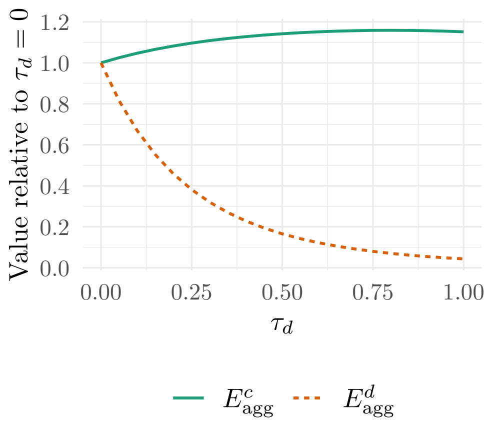
Dependent Variables:	$\Delta_1 \log Y$	$\Delta_2 \log Y$
Model:	(1)	(2)
shock^1	0.14	
	(0.016)	
shock^2		0.15
		(0.016)
Fixed-effects		
District by Year	Yes	Yes
2-digit Industry by Year	Yes	Yes
Fit statistics		
Observations	116,727	84,454
F-statistic	188.1	190.8
R^2	0.08	0.08
within-R^2	0.0017	0.0025

Table A.2: Regression results for the first stage of the 2SLS estimation of equation (3). Standard errors are clustered at the year and 4-digit industry level. First stage results for the factor demand elasticity estimation.

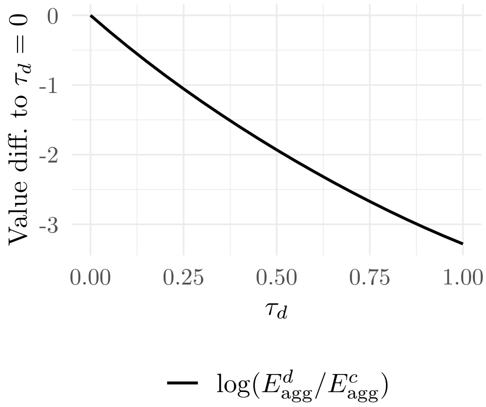
C Additional Policy Experiments

C.1 Tax on dirty energy use

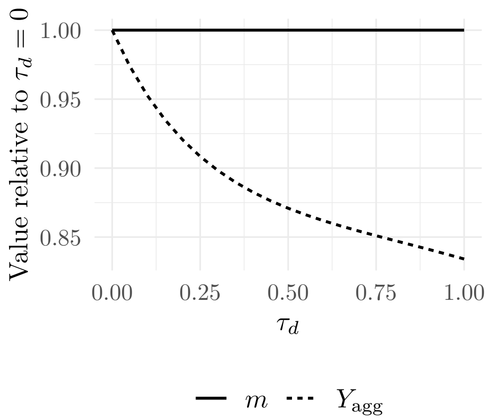
Figure A.1 shows the results of the tax on dirty energy use policy experiment keeping the distribution fixed.



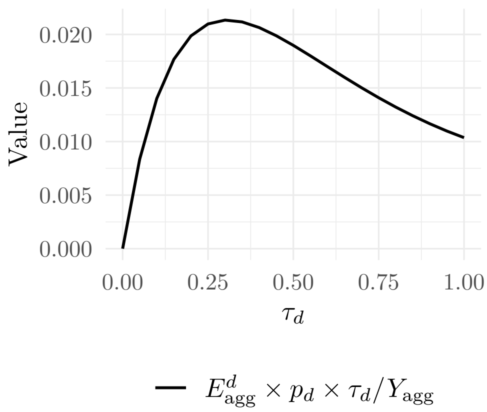
(a) Use of clean and dirty energy.



(b) Energy mix.



(c) Total output and mass of entering plants.



(d) Tax revenue as fraction of output.

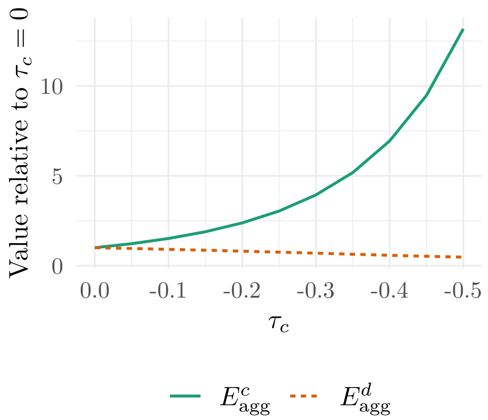
Figure A.1: Results of tax on dirty energy use policy experiment, keeping the distribution fixed at baseline.

C.2 Subsidy on clean energy use

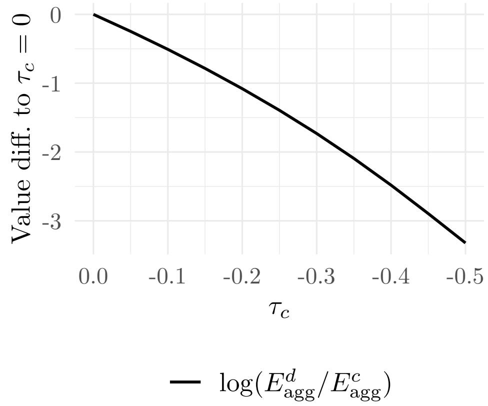
Consider a subsidy on clean energy use, such that the final price of clean energy is

$$p_c^{\text{sub}} = p_c(1 + \tau_c). \tag{27}$$

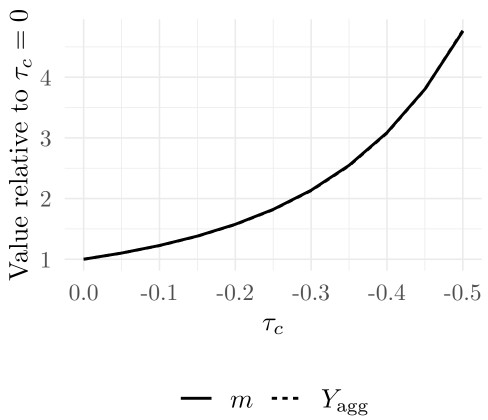
The subsidy is permanent and will be received by all plants. I consider values $\tau_c \in [-0.5, 0]$, up to a 50% subsidy on the price of clean energy. As in section 6.3, I solve the model under the new price and compute aggregates given the new equilibrium distribution of plants. Figure A.2 shows the results of the subsidy on clean energy use policy experiment.



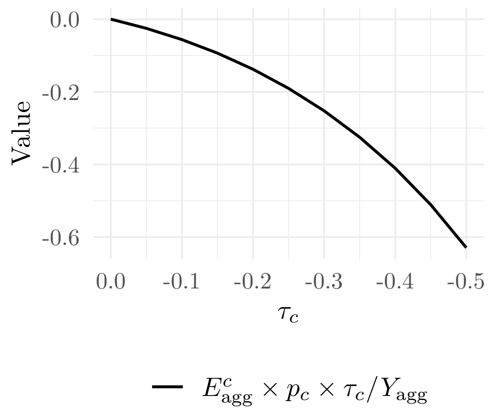
(a) Use of clean and dirty energy.



(b) Energy mix.



(c) Total output and mass of entering plants.



(d) Subsidy spending as fraction of output.

Figure A.2: Results of subsidy on clean energy use policy experiment, allowing for compositional changes.