# Growth, Firm Scale, and the Energy Intensity of Production*

Kathryn McDonald
Columbia University

Noémie Pinardon-Touati
Columbia University

Conor Walsh
Columbia University and NBER

June 2026

**Abstract**

We uncover a new mechanism linking economic growth and a decline in the energy intensity of production, observed globally since 1990. Using firm-level data from India and a causal research design, we show that the energy expenditure share, and thereby energy intensity, decline steeply with firm scale. Because average firm size increases with development, growth reduces the aggregate energy intensity of production. We develop a general equilibrium model of this mechanism, and find that it explains most of the recent decline in India's energy intensity. Projecting forward, energy use in 2050 is about 50% lower than it would be without scale-dependent energy demand.

*Keywords: Growth, Energy Cost Shares, Firm Size, Energy Efficiency*

*JEL Codes, Q43, D24, L25*

*We are very grateful to Francine Montecinos for excellent research assistance. We thank Ashley Langer for an insightful discussion, and Severin Borenstein, Danial Lashkari, Kevin Kuruc, Ezra Oberfield and Joseph Shapiro for helpful comments. We thank conference participants at the NYU LAEF, OzMacro, and Utah Winter Business Economics conferences, as well as seminar participants at Berkeley, Columbia, Georgetown, INSEAD, Middlebury, NY Fed, Richmond Fed, SF Fed, and Stanford.

# 1. INTRODUCTION

As long as production requires energy, and energy production emits CO2, there exists a fundamental trade-off between economic growth and reducing carbon emissions. A key determinant of this trade-off is the energy intensity of production, or the amount of energy required to produce a unit of output. Figure 1(A) shows that global energy intensity has declined dramatically since the 1990s. This almost entirely explains the contemporaneous decline in CO2 intensity, as the CO2-to-energy ratio remained relatively unchanged.[1] Moreover, Figure 1(B) shows that energy intensity declined most in the countries that experienced the fastest growth. Understanding the determinants of this decline in energy intensity, and of its relationship with economic growth, is central to assessing the sustainability of economic development.

In this paper, we uncover a new mechanism that endogenously links economic growth to a decline in energy intensity. It is well-known that average firm size increases as economies grow.[2] We document that the energy expenditure share, and thereby energy intensity, falls steeply as firms grow. Together, these two facts imply that economic growth endogenously lowers the energy intensity of production.

We study this mechanism in the context of India, a rapidly growing economy whose patterns of energy use are broadly representative of low- and middle-income countries. Our analysis proceeds in two steps. First, using firm-level data and a causal research design, we show that the energy expenditure share and physical energy intensity decline with firm scale. These patterns are best rationalized by a non-homotheticity in firms' production function. We then develop a model of heterogeneous firms with non-homothetic energy demand, where scale dependence arises from two forces: physical scaling laws that reduce energy losses at larger scale, and fixed-cost technology investments. Disciplining the model with our micro estimates, we quantify how firm scaling shapes aggregate energy intensity along the growth path.

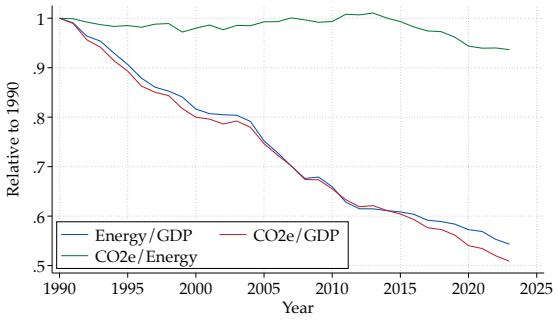
Our empirical analysis uses the Indian Annual Survey of Industries, combining a firm-level panel from 1998 to 2017 with aggregate data from 1990 to 2023. We

[1] Supplementary Material Figure SM.6 shows that this decline in energy intensity is *not* driven by reallocation across countries or sectors, which accounts for less than 10% of the total decline.

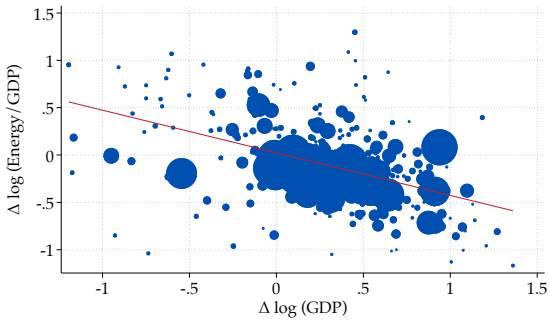
[2] See, e.g., Tybout (2000); Hsieh and Klenow (2014); Bento and Restuccia (2017); Poschke (2018); Bento and Restuccia (2021); Klenow and Li (2025). Throughout the paper, we will mainly use average firm size to refer to output per firm. Employment per firm is another measure of size used in the literature, and this also appears to rise with development.

FIGURE 1: The Global Decline in Energy Intensity

(A) In the Time Series




(B) In the Cross-Section of Countries




*Notes:* GDP is measured in constant 2021 USD PPP. CO2e is energy-related greenhouse gas emissions in tonnes of CO2 equivalent. Energy is energy use in joules. In Panel A, world values are computed by summing values across countries. Panel B plots the 10-year log change in Energy/GDP (y-axis) against 10-year GDP growth (x-axis) across countries×decades between 1990 and 2020. The dots are proportional to initial GDP. The sample is a balanced panel of 127 countries which jointly account for 98% of the world's output.

estimate the effect of firm growth (measured by sales) on the expenditure share of energy relative to other inputs, instrumenting for firm growth with firm-specific demand shocks. The instrument interacts firms' ex-ante product shares with the evolution of nationwide product sales. We estimate a scale elasticity of the relative energy expenditure share of approximately $-0.4$. This relationship is not driven by changes in relative input prices as firms grow—it is a scale effect, not a substitution effect. This decline in the energy expenditure share translates directly into lower energy intensity as firms grow, with a similar scale elasticity.

We then turn to the mechanisms that can explain this result. First, we explore *how* firms reduce their energy expenditure shares when scaling up. The reduction is fully attributable to a decline in the use of energy relative to other material inputs. That is, as scale increases, firms are using less energy to process a given amount of material input.[3] The relationship between energy expenditure shares and scale is not driven by shifting product composition, either on the input or output side.

Second, we explore the reasons *why* firms disproportionately reduce their energy usage when scaling up. We argue that energy as an input is characterized by large losses in production, and these energy losses scale less than proportionately with firm output. We attribute this to two distinct forces: (i) for a given technology, operating at a higher scale is more energy efficient because of physical scaling laws, and (ii) as they grow, firms are able to undertake fixed-cost investments that

[3]One potential concern is that the growing material share reflects higher indirect energy use as firms grow. We estimate total energy use—the sum of energy used by the firm and of the energy embedded in its material inputs—and show that total energy intensity also declines as firms grow.

improve energy efficiency. We estimate that the first force explains approximately 25% of our effect, while the latter accounts for the remaining 75%.

How does the scale dependence of energy demand affect aggregate energy intensity along the growth path? To answer this, we develop a model of heterogeneous firms with non-homothetic energy demand. We consider a non-homothetic CES (nhCES) firm-level production function that combines energy and non-energy inputs (Sato 1974, 1977; Comin, Lashkari, and Mestieri 2021; Lashkari, Bauer, and Boussard 2024). This specification assumes that the output elasticity of energy expenditure relative to other factors is a constant structural parameter. This parameter is identified by our reduced-form evidence on the causal relationship between firm scale and the energy expenditure share. Fixed costs paid in units of labor determine the number of firms and average firm size.

We find that the scale dependence of energy demand has economically meaningful effects on the aggregate energy expenditure share. Calibrating productivity growth to match observed manufacturing output growth in India from 1990 to 2023, our model generates an increase in average firm size similar to that observed in the data, and reduction in the energy expenditure share of about 45%. This decline is entirely due to the non-homotheticity in the production function; a model with a CES production function predicts virtually no change. The decline in the energy expenditure share implies a 50% reduction in energy intensity.

With the calibrated model, we perform two further exercises. First, we use the model to project the path of aggregate energy intensity in India in future decades. Compared to a benchmark without scale dependence of energy demand, our analysis implies a 49% reduction in total energy demanded in 2050. Second, we use our model to examine the paths of CO2e/Energy that would be required to meet India's Paris Climate Agreement emission goals. We find that the CO2e/Energy ratio would be required to fall significantly more from current levels, and thus require significantly more policy-led decarbonization, without accounting for the mechanism we uncover.

Our study has three key policy implications. First, it suggests that the trade-off between economic outcomes and environmental degradation endogenously improves as LMICs grow. Second, removing distortions that favor small firms will also reduce the aggregate energy intensity of production, with resulting environmental benefits. Finally, our findings encourage the developers of Integrated Assessment Models (IAMs) to consider incorporating energy scaling laws with development. Doing so could be a fruitful avenue for future research.

**Related Literature.** Our work most closely relates to a growing literature investigating the consequences of economic growth for energy use and CO2 emissions. A large body of work has focused on the endogenous relationship between growth and the adoption of renewable energy sources (Acemoglu, Aghion, Bursztyn, and Hémous 2012; Acemoglu, Aghion, and Hémous 2014; Arkolakis and Walsh 2023). Meanwhile, we focus on the endogenous relationship between growth and the energy intensity of production. Previous work has documented the importance of directed technical change for the energy transition (Shanker and Stern 2018; Hassler, Krusell, and Olovsson 2021; Acemoglu, Aghion, Barrage, and Hémous 2023; Casey 2024). We highlight energy intensity improvements in response to scale, as opposed to the relative price of energy. Gertler, Shelef, Wolfram, and Fuchs (2016), Caron and Fally (2022) and Aghion, Boppart, Peters, Schwartzman, and Zilibotti (2024) explore the role of non-homothetic consumption for the relationship between growth and the energy/CO2 intensity of consumption. We, on the other hand, document a meaningful non-homotheticity in production.

Second, our paper also relates to the literature that decomposes the time series of CO2 or pollution intensity into the composition of industries, the composition of firms within industries, and within-firm changes (Levinson 2015; Shapiro and Walker 2018; Martin 2011; Barrows and Ollivier 2018; Copeland, Shapiro, and Taylor 2022). Compared to this literature, our work provides an endogenous mechanism that links growth and energy intensity. In their work, Barrows and Ollivier (2018) document that CO2 intensity is falling in the cross-section of firm size in their sample of 2,500 Indian firms. While their sample only covers the extreme right tail of the firm size distribution, we show that this pattern holds in a sample approximately 80 times larger. In addition, we go beyond correlations to document a causal effect of firm scale on relative energy demand.[4]

Third, we contribute to the literature examining the nexus between economic development, energy usage, and CO2 emissions in the context of LMICs. Closest to the spirit of our exercise, several papers (e.g., Bruckner, Hubacek, Shan, Zhong, and Feng 2022; Wollburg, Hallegatte, and Mahler 2023) quantify the effect of growth in low-income countries on global CO2 emissions. However, these papers build projections using historical levels of CO2 intensity. Our work shows that energy intensity endogenously evolves as these countries grow. Several papers investigate the role of the electricity sector in the process of growth (Abeberese

[4]The literature on firms and CO2 intensity has also shown that CO2 intensity is negatively correlated with TFPR (Klenow, Pastén, and Ruane 2024; Gao 2025; Kim 2025; Kabir and Tan 2025) with implications for carbon taxation and the effects of green financing on aggregate production.

2017; Lee, Miguel, and Wolfram 2020; Singer 2024; Colmer, Lagakos, and Shu 2024). Abeberese (2017) and Singer (2024) investigate how shocks to the relative price of electricity affect firm growth and electricity intensity in India. We study how growth affects relative energy demand.

Finally, a distinct literature investigates the effects of trade on CO2 and energy intensity. This literature has documented that exporting firms—which tend to be larger—are cleaner (Batrakova and Davies 2012; Forslid, Okubo, and Ulltveit-Moe 2018; Barrows and Ollivier 2021). Barrows and Ollivier (2021) show that, among Indian exporters, exposure to foreign demand shocks leads to reductions in the CO2 intensity of output. Also in India, Martin (2011) documents improvements in firms' energy efficiency following the 1991 trade liberalization. It is critical to evaluate the relationship between scale and energy intensity beyond the sample of exporters; exporting firms could for instance be cleaner because they are more exposed to foreign demand for clean products or foreign climate policy.

The remainder of the paper is structured as follows. Section 2 describes the data sources and presents aggregate facts on energy intensity. Section 3 presents our main results on firm scale and energy intensity. Section 4 proposes a theory of aggregate energy intensity. Section 5 quantifies the importance of this mechanism to explain the historical trends in India, and Section 6 explores the implications for Indian energy use and CO2 emissions going forward. Section 7 concludes.

# 2. Data and empirical context

## 2.1 Data

**Annual Survey of Industries.** Our main data source is the Annual Survey of Industries (ASI), India's annual establishment-level survey of the formal manufacturing sector. The ASI samples all firms with more than 100 employees, and a fifth of all firms with more than 20 employees (or more than 10 if they use power).[5] The firms in the ASI represent approximately 80% of total manufacturing output.

We use an annual panel from 1998 to 2017. Our cleaned sample consists of 690,775 firm$\times$year observations from 199,709 firms. A detailed description of the data and variables we use, as well as the key cleaning steps, is contained in Supplementary Material SM.1.1. Table SM.6 provides summary statistics for the key variables

[5]The ASI surveys a firm's combined establishments in a given industry-state; multi-establishment firms may appear as separate observations only if they belong to different industry-state cells. Boehm, Dhingra, and Morrow (2022) estimate that less than 7.5% of all establishments fall in this case. With that caveat in mind, we call the units of observation in our data "firms."

used in the paper. Throughout, we deflate all monetary values to the base year of 2011 using the manufacturing output deflator from the national accounts. All statistics are weighted by the sampling weight. In addition to the annual firm-level panel, we use the ASI aggregated time series, available from 1990 to 2023.

We construct an industry classification consisting of 78 manufacturing industries which are consistently defined throughout our sample, approximately corresponding to the 3-digit level of the official sectoral classification (NIC). The main variables we use are gross output, labor costs, capital, material inputs purchase value, and energy inputs purchase value. To construct the cost share of capital, we construct a user cost of capital equal to the sum of the rental rate and depreciation rate. The rental rate is measured as the average interest rate on loans, and the depreciation rate is computed from reported depreciation, both averaged at the industry$\times$year level.[6] We note three caveats with the measurement of the cost share of capital: first, we only observe the cost of loans, while the cost of capital should also reflect the cost of equity; second, our measure of the user cost of capital is not firm-specific; third, ASI only reports tangible capital, not intangibles.

A key advantage of the ASI is that it contains detailed data on energy usage by manufacturing firms: we know total expenditures on energy inputs, broken down into electricity, coal, oil, and (since 2008) gas. For each energy source, expenditures are broken down into price and quantity.[7] This allows us to construct the physical quantity of energy used by each firm.

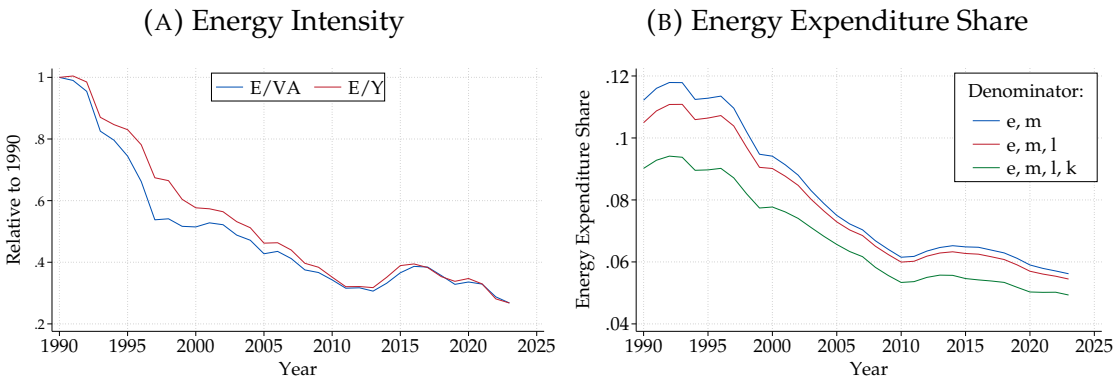
At several points, we exploit the fact that, for both the products that manufacturing establishments produce and the inputs they buy, we observe information on sales, quantities, and unit values, at the product code level. This allows us to construct firm-level changes in prices and quantities, for both outputs and material inputs.

**Other data sources.** We exploit several additional data sources, including: (i) International Energy Agency data, (ii) Indian national accounts, (iii) World Bank Enterprise Survey, (iv) United Nations Industrial Development Organization (UNIDO) Energy Efficiency and Renewable Energy Technology Compendiums, and (v) India's National Sample Survey of Unorganised Manufacturing Enterprises (NSS). These data are described in the Supplementary Material SM.1.2.

[6] We average at the industry$\times$year level because firm-level depreciation rates largely reflect accounting conventions, while interest rates are observed only for firms with positive loan balances.

[7] Oil expenditures are not broken down into price and quantity. Supplementary Material SM.1.1 details our construction of the price of oil.

FIGURE 2: Energy Intensity and Expenditure Share in Indian Manufacturing




*Notes:* Panel A plots the aggregate ratio of energy quantity E to total value added VA and total revenue Y in the ASI data. Panel B plots the aggregate cost share of energy inputs in the ASI data, relative to three denominators: energy and materials costs (e, m), energy, materials and labor costs (e, m, l), energy, materials, labor and capital costs (e, m, l, k).

## 2.2 Energy Intensity and Cost Share in Indian Manufacturing

This paper studies energy use in the Indian manufacturing sector. Similar to the global trends presented in Figure 1, overall energy intensity has declined markedly in India since 1990. As at the global level, reallocation across sectors accounts for a negligible share of this decline.[8] In this paper, we focus on the use of energy in manufacturing, which accounts for roughly 70% of energy used in the production of goods and services in India.

Figure 2(A) shows the decline in energy intensity in formal manufacturing in the ASI data. Figure 2(B) shows the cost share of energy relative to the cost share of other inputs over time. It reveals that the cost share of energy has declined *relative to that* of other inputs. Using total costs as a denominator, the energy expenditure share has declined by 45% from 1990 to 2023.[9]

# 3. FIRM SCALE AND ENERGY DEMAND

To think about the relationship between energy intensity and scale, consider a firm $i$ with four basic inputs $\mathcal{X} \subseteq \{e, m, l, k\}$ where $e$ is energy, $m$ is material inputs, $l$ is labor and $k$ is capital. They have a cost function $C_i(y_i, \mathbf{w}_i)$ for output that depends on the vector of input prices $\mathbf{w}_i$ (possibly firm-specific) and the choice of output

[8] See Figure SM.5 in the Supplementary Material.

[9] Figure SM.7 plots the energy cost share in ASI purged of reallocation across 3-digit industries, and shows that reallocation across industries plays no role in the decline in the manufacturing energy share.

$y_i$. Energy intensity, defined as energy in joules per dollar value of output, is

$$\Theta_i \equiv \frac{e_i}{p_i y_i},$$

where $p_i$ is the unit output price. Now, define the expenditure share of energy as

$$\Omega_i^e \equiv \frac{w_i^e e_i}{\sum_{x \in \mathcal{X}} w_i^x x_i} = \frac{w_i^e e_i}{C_i(y_i, \mathbf{w}_i)} \tag{1}$$

where $w_i^x$ is the price of input $x$ and $x_i$ is the input choice. Throughout the paper, we use the notation $\Omega_i^e$ when $\mathcal{X} = \{e, m, l, k\}$ (the expenditure share of total costs), and $\Omega_i^{e,\mathcal{S}}$ when we specify another denominator $\mathcal{S} \subset \mathcal{X}$. We can then write:

$$\Theta_i = \Omega_i^e (1 - \pi_i)(w_i^e)^{-1}, \tag{2}$$

where $\pi_i \equiv (p_i y_i - C_i)/(p_i y_i)$ is the profit rate of the firm. Energy intensity can vary with scale for three reasons: either the energy expenditure share, or the profit rate, or the price of energy may change with firm scale. In this paper, we show that non-homothetic energy demand driving falling energy cost shares as firms grow is the key determinant of the decline of energy intensity with scale.

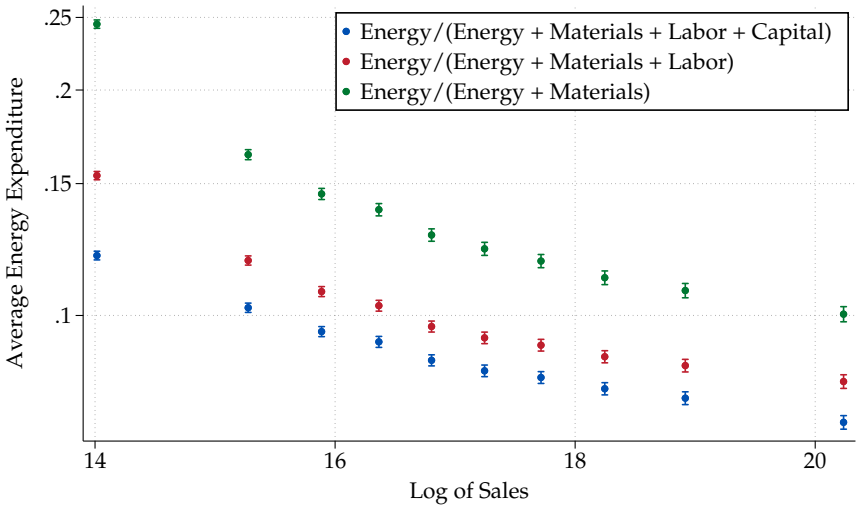
## 3.1 Firm Scale and the Energy Expenditure Share

### 3.1.1 Energy Expenditure Shares in the Cross-Section of Firms

Figure 3 shows the relationship between the expenditure share on energy relative to other inputs and firm size in the cross-section of firms. It plots the average energy expenditure share by deciles of firm size (defined within each industry×year). The denominator of the energy expenditure share is defined as either the cost of energy and materials ($\Omega^{e,em}$), energy, materials and labor ($\Omega^{e,eml}$), or energy, materials, labor and capital ($\Omega^{e,emlk} = \Omega^e$). It shows a strongly negative relationship between firm size and energy expenditure share. Consider the expenditure share of energy relative to energy, materials and labor $\Omega^{e,eml}$—which we take as our baseline to avoid uncertainties related to the measurement of the cost of capital: firms in the first decile have an average expenditure share of 15%, while firms in the last decile have an average expenditure share of 8%, that is, their expenditure share on energy is only about *half* of that of firms in the first decile.

A key threat to the interpretation of this cross-sectional relationship as a scale elasticity of energy expenditure shares to firm size is if firm size correlates with

FIGURE 3: Energy Expenditure Shares and Firm Size in the Cross-Section




*Notes:* This figure shows average energy expenditure shares by deciles of firm sales (defined within each 3-digit industry×year) in the ASI data. The dots are point estimates and bars are 95% confidence intervals.

other variables that directly affect the energy share. For instance, smaller and larger firms may face different energy prices, or differ in the probability that their energy-using machinery breaks down. Another potential concern is that size correlates with age, and firms in different cohorts use capital of different vintages, which may be more or less energy-efficient. Finally, it may be the case that size is not only correlated with factor-neutral productivity, but also factor-biased productivity, e.g., large firms are large because they are particularly productive in the use of energy. Appendix A.1.1 formalizes these concerns for the identification of the scale elasticity of energy demand through the lens of the production function used in Section 5. To address these identification concerns, we turn to an instrumental variable strategy.

### 3.1.2 Causal Effects of Firm Growth on Relative Energy Demand

To examine the causal effect of firm scale on the intensity of energy demand, we construct an instrument for firm-level output growth and estimate the effect of a change in output on the change in the energy expenditure share across $h$ years. We estimate regressions with the following general form:

$$\Delta^h \log\left(\frac{\Omega^e_{it}}{1-\Omega^e_{it}}\right) = \alpha_{sct} + \eta^h \, \Delta^h \log(\text{sales}_{it}) + u_{it} \tag{3}$$

$\Delta^h$ is the average yearly difference across $h$ years: $\Delta^h x_{it} = (x_{it+h} - x_{it})/h$.[10] $\log(\text{sales}_{it})$ is the log of firm sales (in constant 2011 INR). The outcome variable is the log of energy expenditures relative to expenditures on other inputs. We use the relative share as an outcome variable as it directly maps to the structural parameter of the model presented in Section 5. $\alpha_{sct}$ are industry×cohort×time fixed effects. Industries are time-invariant 3-digit NIC codes. Cohorts are defined by grouping firms into four-year bins based on the year of initial production (from 1974), with two additional bins for the earlier years. For horizon $h$, $\eta^h$ is the elasticity of the relative energy expenditure share to firm scale.[11]

In order to address the potential endogeneity associated with estimating equation (3) via OLS, we use an instrumental variable strategy. We construct a shift-share instrument for firm-specific demand shocks. We interact ex-ante firm-level product shares with the time-series evolution of aggregate sales at the product level. The logic is that a firm will grow if it is specialized in products that are subject to positive aggregate demand increases. More precisely, defining $\mathcal{V}$ as the set of all products, we let:

$$\mathcal{D}_{it}^h = \sum_{j \in \mathcal{V}} \omega_{ijt} \, \Delta^h \log \text{sales}_{jt} \tag{4}$$

be a demand shifter, where $\Delta^h \log \text{sales}_{jt}$ is the aggregate growth rate of sales of product $j$ from $t$ to $t+h$ and $\omega_{ijt}$ is the share of product $j$ in the sales of firm $i$ at time $t$. We use $\mathcal{D}_{it}^h$ as an instrument for $\Delta^h \log \text{sales}_{it}$.

The instrument is relevant: the Appendix first-stage results in Table OA.2 show that the F-stat varies between 212 and 505 across horizons and specifications. The instrument will be valid as long as firm exposure to aggregate product-level growth is orthogonal to changes in the unobserved determinants of the energy share $u_{it}$ in (3), in particular any firm-specific energy supply or productivity shock. That is, firms with larger energy supply shocks or greater energy-specific technical change must not have systematically different shares in high growth products (Borusyak, Hull, and Jaravel 2022). Figure OA.1 shows that the instrument is uncorrelated with firm characteristics that are likely correlates of energy supply

[10] Because smaller establishments in the ASI are not surveyed every year, requiring observations at both $t$ and $t+h$ leads to attrition. The best populated difference is $h = 1$. To preserve statistical power, we define $\Delta^h x_{it}$ using $(x_{i,t+h} - x_{it})/h$ when available and fill missing values with $(x_{i,t+h+1} - x_{it})/(h+1)$, so horizon $h$ should be understood as $h$ or $h+1$.

[11] One potential concern with our estimation of $\eta^h$ is that equation (3) restricts the sample to firms present in two periods (i.e., continuing firms). Figure SM.8 shows that changes in the energy share of continuing firms are representative of the dynamics of the energy share in the full sample.

TABLE 1: Energy Expenditure Shares and Firm Growth

| | $\Delta^5 \log(\Omega/(1-\Omega))$ | | | | | |
|---|---|---|---|---|---|---|
| | OLS | | | IV | | |
| | $\Omega^{e,em}$ | $\Omega^{e,eml}$ | $\Omega^{e,emlk}$ | $\Omega^{e,em}$ | $\Omega^{e,eml}$ | $\Omega^{e,emlk}$ |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| $\Delta^5$ log(sales) | -0.382*** | -0.304*** | -0.216*** | -0.487*** | -0.426*** | -0.355*** |
| | (0.009) | (0.009) | (0.009) | (0.042) | (0.040) | (0.040) |
| Year × Ind. × Cohort FE | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| First-stage F-Stat | | | | 293.8 | 293.9 | 293.6 |
| Observations | 184,986 | 185,021 | 184,749 | 183,964 | 183,999 | 183,727 |

*Notes:* This table presents the results of estimating equation (3). Columns 1-3 present OLS results. Columns 4-6 present IV results where output growth $\Delta^h \log \text{sales}_{it}$ is instrumented with the firm-level demand shock defined in equation (4). Regressions are weighted by firm total cost (adjusted for the sampling weight). Standard errors are clustered at the firm and main product×year level. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

or productivity, in particular the levels and past trends in factor shares. We also provide evidence that our instrument captures a demand shift by showing that it is uncorrelated with changes in TFPQ.[12] Appendix A.1.1 provides further details on identification following the "shifters-based" approach in Borusyak et al. (2022).

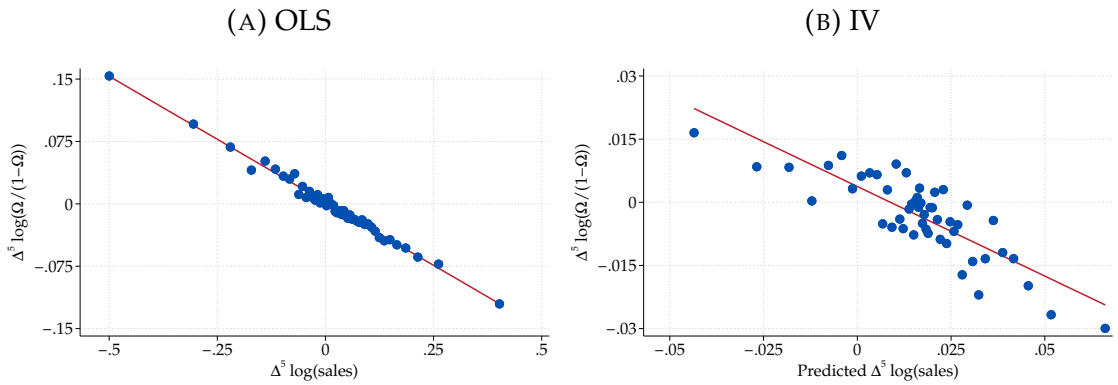
Our baseline results use $h = 5$. Regressions are weighted by firm total cost (adjusted for the sampling weight). Standard errors are double-clustered at the firm level and at the main product×year level (to account for the correlation of residuals across firms with similar product exposures, an issue raised by Adão, Kolesár, and Morales (2019) and Borusyak et al. (2022)).

**Main Results.** In Table 1 we report the results of this estimation strategy. Columns (1) to (3) report OLS results. We take as our baseline result specification (2) where the energy share is defined as $\Omega^{e,eml}$, resulting in an elasticity around $-0.3$. Columns (4) to (6) report IV estimates. The IV elasticities are slightly larger than the corresponding OLS ones, with a baseline elasticity equal to approximately $-0.4$ in column (5). Figure 4 visualizes the relationship using binned scatter plots (of variables residualized on industry×cohort×year fixed effects). The relationship is approximately log-linear, validating the choice of functional forms.

How large are the differences in the energy share across firms implied by this scale elasticity? Consider a firm that grows from the 10th to the 90th percentile of the (within-industry) size distribution. This corresponds to a 94-fold increase in sales. Using our IV estimate, the log relative energy expenditure share will fall by

[12] Scale being shifted by factor-neutral productivity is not a concern for identification. However, it is plausible that if our instrument were correlated with energy-specific productivity, it would also be correlated with measured TFPQ, so this test is informative.

FIGURE 4: Energy Expenditure Share and Firm Size: Binned Scatterplots




*Notes:* This figure presents binned scatter plots corresponding to equation (3). The energy share is defined as $\Omega^{e,eml}$ and $h = 5$. Panel A presents OLS results. Panel B presents results where output growth $\Delta^h \log \text{sales}_{it}$ is instrumented with the firm-level demand shock defined in equation (4).

$-0.4 \times \log(94) \approx -1.8$; that is, the relative energy share $\Omega^{e,eml}/(1-\Omega^{e,eml})$ falls by a factor of 6.2. This corresponds to 41% of the p90-to-p10 ratio in the relative energy share (within-industry) distribution.[13]

**Robustness.** We provide a number of robustness exercises and complementary tests for these findings in the Supplementary Material, and summarize them here.

*Alternative specifications.* Figure SM.10 shows variants of our baseline specification. We show results for $h = 1, 3, 5, 7$. Coefficients are slightly more negative at $h = 1$, and then very similar at $h = 3, 5, 7$, suggesting that our baseline $h = 5$ appropriately captures long-run elasticities. We obtain very similar coefficients for alternative fixed effects structures: using a finer definition of industries, further interacting industry×time fixed effects with state×time fixed effects or with initial characteristics (size or capital intensity). Unweighted, cost-weighted, and sales-weighted regressions yield highly similar results. Figure SM.11 shows that our results are robust to alternative definitions of our shift-share instrument excluding concentrated products or large firms from the shifters. Table SM.7 shows that the results are robust to using as outcomes the change in the expenditure shares in levels $\Delta^5 \Omega^e$, or in logs (but not in relative shares) $\Delta^5 \log(\Omega^e)$.

*Scale elasticity along the firm size distribution.* Table SM.8 shows the results from estimating equation (3) by bins of ex-ante firm size. The scale elasticity is significantly negative for both ex-ante smaller and larger firms, and does not systematically vary across bins.

*Capacity utilization.* One hypothesis is that changing cost shares merely reflect

[13] For reference, expressed in levels, the within-industry p10 and p90 values are 0.013 and 0.197, respectively (see Table SM.6).

changes in the utilization of fixed inputs. Our main results mitigate this concern by focusing on relatively long horizons, where firms are unlikely to adjust solely through capacity utilization. In addition, Table SM.9 uses the share of days worked as a proxy for capacity utilization and shows similar results for firms with and without utilization changes.

*Exposure to size-dependent regulation.* Another possibility is that larger firms face more stringent energy-related regulation, causing energy expenditure shares to fall with size. Then, our findings would reflect size-dependent regulatory exposure rather than a feature of production technologies. We compile a comprehensive list of energy-related policies in India and show in Figure SM.12 that our results are robust to excluding exposed firms.

*Informal sector.* We assess robustness to including the informal sector by combining the ASI with the National Sample Survey of Unorganised Manufacturing Enterprises (NSS). Since the NSS lacks a panel dimension, we report cross-sectional results analogous to Figure 3. Table SM.12 shows similarly negative elasticities in the NSS alone and in the pooled sample.

### 3.1.3 Scale or Substitution: Firm Scale and Relative Input Prices

In this paper, we interpret the decline in the relative energy cost share as firms grow as a direct result of firm growth—i.e., a scale elasticity of relative cost shares. An alternative explanation is that the price of energy relative to other inputs evolves endogenously with firm scale, and that cost shares evolve as firms grow in response to these relative price effects. For instance, suppose that firms face firm-specific supply curves in input markets, and that the supply curve for energy is more elastic than for other inputs. Then, one causal effect of firm growth is that the relative price of energy faced by the firm declines. If energy and other inputs are complements, this leads to a decline in the energy expenditure share.

This suggests a horse race: one can regress the change in the energy expenditure share on the log change in output and the log change in the relative price of energy. We perform this analysis in Table 2. In columns (1)-(3), we show what happens to input prices at the firm level as firms grow, with firm growth instrumented for using the strategy above. We do not find evidence that energy or material prices change with growth.[14] In columns (4)-(6), we show what happens to relative

[14]We do find that firm-level wages rise with firm growth, consistent with evidence from Mertens and Schoefer (2024); however, given the substitution elasticities we estimate in Section 5, this effect can only contribute less than 0.005 to the estimated scale elasticity of energy cost shares.

TABLE 2: Relative Input Price Changes and Firm Growth

| | $\Delta^5 \log(w)$ | | | $\Delta^5 \log(\Omega/(1-\Omega))$ | | |
|---|---|---|---|---|---|---|
| | $w^e$ | $w^m$ | $w^l$ | $\Omega^{e,em}$ | $\Omega^{e,eml}$ | $\Omega^{e,emlk}$ |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| $\Delta^5$ log(sales) | 0.010 | 0.037 | 0.048** | -0.454*** | -0.392*** | -0.327*** |
| | (0.034) | (0.121) | (0.021) | (0.048) | (0.045) | (0.044) |
| Year × Ind. × Cohort FE | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| $\Delta^5 \log(w)$ controls | | | | ✓ | ✓ | ✓ |
| F-Stat | 294.5 | 245.0 | 297.6 | 252.0 | 252.0 | 251.6 |
| Observations | 182,597 | 148,691 | 184,094 | 147,107 | 147,107 | 146,902 |

*Notes:* This table presents IV results for specification (3), with the instrument defined in equation (4). In columns (1)-(3), the outcome variable is the log change in firm-level input prices. Energy prices are the weighted average of unit prices for electricity, coal, and oil. Material prices are the weighted average of material unit prices. The wage is total wage bill divided by number of employees. In columns (4)-(6), we estimate (3) including firm-level input price changes as controls. Regressions are weighted by firm total cost (adjusted for the sampling weight). Standard errors are clustered at the firm and main product×year level. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

energy expenditures with firm growth, but including as controls the change in firm-level input prices. This leaves our results essentially unchanged, suggesting that the results above are not driven by changes in relative prices as firms grow.[15]

## 3.2 Firm Scale and the Energy Intensity

The decreasing relationship between firm size and the relative expenditure share on energy inputs translates directly into a decreasing relationship between firm size and the energy intensity of production. To show this, we employ the same specification (3), but the outcome variable is now $\Delta^h \log(e/x)$. $e$ is the firm's total energy usage measured in megajoules, and $x$ is revenue $py$, value-added $va$ or physical output $y$, respectively. Table 3 presents the results. In columns (1) to (3), the explanatory variable is sales growth (as above), and in columns (4) to (6) we present the same results where the explanatory variable is physical output growth. All coefficients are negative and statistically significant. A 1% demand-induced increase in firm sales reduces $e/py$ by 0.5% and $e/y$ by 0.3%. Finally, the elasticity

[15]Our results are also unlikely to be driven by size-dependent unpriced input wedges (we can rule out size-dependent priced input wedges from Table 2). To explain our results, such wedges would need to have a different scale elasticity for each input, and in particular for energy vs. materials. For intermediate inputs, unpriced wedges are likely primarily related to physical frictions in input access. Our results hold conditioning on state×year fixed effects, where firms arguably share the same shadow cost of accessing inputs.

TABLE 3: Energy Intensity and Firm Size

| | $\Delta^5$ log(outcome) | | | | | |
|---|---|---|---|---|---|---|
| | $e/py$ | $e/va$ | $e/y$ | $e/py$ | $e/va$ | $e/y$ |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| $\Delta^5 \log(py)$ | -0.517*** | -0.510*** | -0.340*** | | | |
| | (0.039) | (0.065) | (0.112) | | | |
| $\Delta^5 \log(y)$ | | | | -0.643*** | -0.804*** | -0.425*** |
| | | | | (0.103) | (0.206) | (0.088) |
| Year × Ind. × Cohort FE | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| F-Stat | 294.5 | 258.3 | 331.5 | 41.5 | 18.9 | 41.5 |
| Observations | 182,597 | 157,782 | 136,889 | 136,889 | 118,293 | 136,889 |

*Notes:* This table presents IV results for specification (3), with the instrument defined in equation (4). In columns (1)-(3), the endogenous variable is the log change in sales value. In columns (4)-(6), the endogenous variable is the log change in quantity produced. Regressions are weighted by firm total cost (adjusted for the sampling weight). Standard errors are clustered at the firm and main product×year level. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

TABLE 4: Energy Intensity Decomposition

| | Scale elasticity | | | |
|---|---|---|---|---|
| | $\frac{e}{py}$ | $\Omega^e$ | $w^e$ | $1-\pi$ |
| | (1) | (2) | (3) | (4) |
| Total costs excludes capital | –0.517*** | –0.397*** | 0.010 | –0.132*** |
| | (0.039) | (0.037) | (0.034) | (0.017) |
| Total costs include capital | –0.517*** | –0.334*** | 0.010 | –0.195*** |
| | (0.039) | (0.038) | (0.034) | (0.019) |

*Notes:* This table decomposes the elasticity of energy intensity to firm sales (column (1) of Table 3) into the three margins of equation (2). It shows IV results for specification (3), with the instrument defined in equation (4). The outcome variables are: (2) the energy expenditure share, (3) the firm-level energy input price, and (4) one minus the profit rate (total costs divided by sales).

of physical quantities of $e/y$ with respect to $y$ is $-0.4$.[16]

In Table 4, we decompose the energy intensity coefficient $e/py$ into the three components of equation (2). The dominant effect is the falling energy expenditure share. As shown above, firm-level energy prices do not vary with firm size. We find that the profit rate does rise with firm size, contributing to the fall in energy intensity. As we show in the model below, the same scale dependence in energy demand that lowers the energy share also generates higher profitability for larger firms (through lower returns to scale). Thus, our theory accounts not only for the

[16]The elasticities with respect to physical quantity $y$ are larger in absolute value than the elasticities with respect to sales $py$. This is because the demand shifter positively affects both $y$ and $p$, so that the first stage coefficient is larger when we instrument $py$ as opposed to $y$. Because the reduced form of both IV regressions is the same, the IV coefficient is then lower in that case.

TABLE 5: Size and Relative Input Use

| | $\Delta^5$ log(outcome) | | | |
|---|---|---|---|---|
| | $\frac{\Omega^e}{\Omega^m}$ | $\frac{\Omega^e}{\Omega^l}$ | $\frac{\Omega^e}{\Omega^l+\Omega^k}$ | $\frac{w^e e}{va}$ |
| | (1) | (2) | (3) | (4) |
| $\Delta^5$ log(sales) | -0.487*** | -0.056 | 0.053 | -0.534*** |
| | (0.042) | (0.040) | (0.041) | (0.066) |
| Year × Ind. × Cohort FE | ✓ | ✓ | ✓ | ✓ |
| F-Stat | 293.8 | 293.8 | 293.6 | 259.9 |
| Observations | 183,964 | 183,873 | 183,726 | 159,001 |

*Notes:* This table presents IV results for specification (3), with the instrument defined in equation (4). The dependent variables are log differences of (1) the energy-to-materials ratio, (2) the energy-to-labor ratio, (3) the energy-to-labor and capital ratio, and (4) total energy costs over value added. Regressions are weighted by firm total cost (adjusted for the sampling weight). Standard errors are clustered at the firm and main product×year level. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

decline in the energy expenditure share, but also—at least qualitatively—for the contribution of rising profitability to lower energy intensity.

## 3.3 Mechanisms

We now dive deeper into the underlying mechanisms behind the results above. First, we present descriptive results that provide further detail about how relative energy demand is falling with size. In particular, we discuss the role of changing intermediate input shares. Second, we present evidence on the underlying physical and economic mechanisms.

### 3.3.1 How Does Relative Energy Demand Fall with Size?

**Relative Input Use.** We investigate which inputs firms use more of when their relative energy demand declines. Table 5 shows the results of estimating equation (3) with alternative outcome variables. We find that the energy-to-materials expenditures ratio is strongly decreasing with firm growth. Consider a brick manufacturer that combines fuel and clay to produce bricks. This result implies that, as production scales up, the manufacturer requires less fuel to process a given quantity of clay. By contrast, we find that the ratio of energy expenditures to wage expenditures declines only moderately as firms grow (and is not statistically significant). Finally, the ratio of energy expenditures to value added (labor, capital, and profits) falls as firms grow.

**Scale, Firm Scope, and Indirect Energy Use.** Our interpretation of the declining relative energy demand with scale is that, for a given set of production steps, firms producing at larger scales require less energy to process a given amount of material inputs. However, an alternative interpretation is that, as firms grow, they change the set of production steps they perform. For instance, growing firms may outsource the energy-intensive parts of their production process, reducing the energy share while increasing the material share. In this case, the estimated coefficients cannot be interpreted as production function parameters mapping a given set of inputs into a given set of outputs.

First, firms could reallocate their sales or input purchases towards more or less energy-intensive products. On the output side, our findings could be explained by firms changing the composition of their product portfolio as they grow, tilting it towards less energy-intensive products. Table SM.13 investigates this point. We define the change in the energy expenditure share that can be attributed to a changing composition of products, keeping product-level energy shares fixed. Demand-induced shifts in firm scale do not affect this metric, showing that firms do not systematically tilt their sales shares towards more or less energy-intensive products. Similarly, on the input side, firms could be using different input bundles as they grow, perhaps outsourcing energy-intensive steps of the production process. Using a similar strategy, we find no meaningful role for reallocation of input purchases as firms grow.[17]

Second, even if firms do not systematically reallocate their sales or purchases towards more or less energy-intensive products, a change in the set of inputs that firms use or in the set of products they produce could reflect a change in the set of production steps performed by the firms. In Figure OA.2, we present results for the subset of firms that use the exact same disaggregated material inputs at $t$ and $t+5$, or produce the exact same disaggregated products at $t$ and $t+5$. The elasticities for this subset of firms are highly similar to those for firms with changing input or product bundles, and to our baseline results. In addition, we consider heterogeneity by change in the vertical span of firms, as defined by Boehm and Oberfield (2023). We find similarly negative effects for firms with small vs. large or positive vs. negative changes in vertical span. As such, our results are not driven by an increase in fragmentation, which could otherwise lead to a decline in the non-material (and hence energy) share.

Finally, when studying the aggregate evolution of energy intensity, it is important

[17] We also find that as firms grow, their energy *mix* from various sources (coal, gas, oil or electricity) does not change (Table SM.14).

to consider indirect energy use through firms' input purchases. Keeping the energy intensity of materials constant, the firm-level increase in the relative material share implies an increase in energy use embodied in materials relative to direct energy use. In general equilibrium, this indirect energy use may partly undo the declines in direct energy use. We address this question in two ways. First, we construct the total energy expenditure of a firm as the sum of the direct and indirect energy purchases embodied in the input purchases of that firm. Table SM.15 shows that total energy expenditure shares also decline with scale.[18] Second, in the model below, we quantify the general equilibrium implications of the increase in energy embodied in intermediate inputs for the aggregate path of energy intensity.

### 3.3.2 Why Does Relative Energy Demand Fall with Size?

In this section, we present two mechanisms that can jointly explain the negative scale elasticity of energy demand. First, physical scaling laws imply that energy losses rise less than proportionally with scale. Second, many energy-saving technologies require substantial fixed costs and are therefore more likely to be adopted by larger firms. We provide empirical evidence for both mechanisms and show in Section 4 that, together, they can quantitatively account for the scale elasticity documented above.

First, it is useful to understand how energy is used in manufacturing. Table 6 shows the share of energy by end-uses in manufacturing.[19] The majority of energy (59%) is used to generate heat, mostly in fired systems like furnaces or kilns (50%). The remainder (9%) represents steam systems (e.g., a steam dryer). The second key use is mechanical work, representing 18%, such as a pump or compressor. Second, a distinctive feature of energy use is that a large fraction is lost before performing useful work. In 2018, only 52% of the energy entering a U.S. industrial plant was applied to its intended use (U.S. Department of Energy 2022). These losses occur in on-site energy generation (e.g., on-site electricity generators), distribution (pipes, transmission lines, etc.), conversion (e.g., electrical to mechanical, heat exchangers), and in final use (wasted heat, by-products, etc.). We think of the share of energy lost in the United States as a lower bound for a country like India.

[18] As clarified in Appendix SM.3.5, these specifications suffer from a missing intercept problem, because the energy intensity of a given input enters as a time constant. We develop a strategy there to address this problem.

[19] See Table SM.5 in the Supplementary Material for more detail across industries.

TABLE 6: Modeled Scale Elasticity of Energy Intensity Within Technology

| End-Use Category | Examples of Technology | Share (%) | Scale Elasticity |
|---|---|---|---|
| Process heating - fired | Furnace, kiln | 50 | $-0.15$ |
| Process heating - steam | Steam dryer, evaporator | 9 | $-0.05$ |
| Process mechanical work | Pump, compressor | 18 | $-0.05$ |
| Other process | Electrochemical, process refrigeration | 12 | 0 |
| Non-process | Building heat, on-site transport | 11 | $-0.10$ |
| Total | | 100 | $-0.10$ |

*Notes*: This table provides estimated scaling laws for energy intensity by end-use in manufacturing. Computations are detailed in Appendix C. End-use shares are computed by combining industry-specific end-use shares from the U.S. Manufacturing Energy Consumption Survey (2018) and industry weights from the ASI.

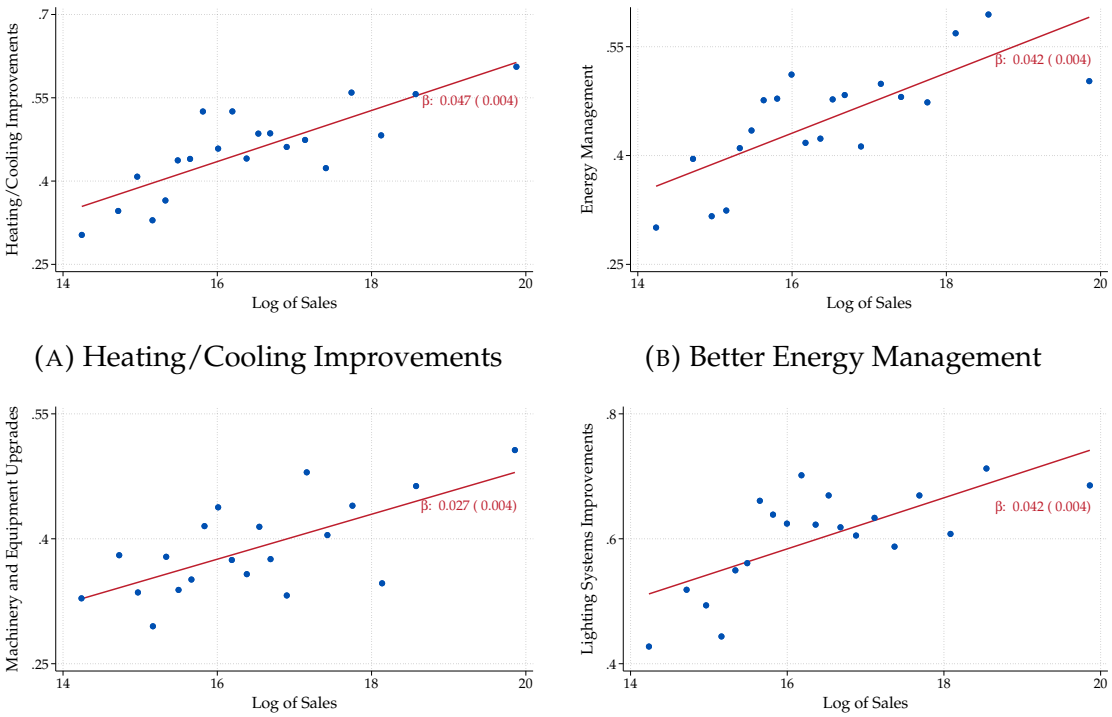
**Physical Scaling Laws.** Larger production units can achieve lower energy use per unit of output simply by virtue of physical scaling laws. As firms expand capacity, they benefit from geometric relationships that favor reduced losses at higher scales. These effects arise when scaling up a given technology, for instance, by increasing the capacity of a furnace while keeping its design unchanged. We use engineering models of key technologies within each energy end-use class to develop estimates of the implied scale elasticities, which we report in Table 6. These are read as the percent decline in energy required per unit of output for a given percent increase in output. We provide some intuition on the sources of scale dependence in energy intensity here, and leave derivations to Appendix C.

*Process heating.* One key source of energy losses in heat generation is heat leaking from hot surfaces to their surroundings, through direct contact with nearby materials, movement of air, and radiation. These losses scale with the surface area of the heated object, while useful heat output scales with the volume of the heated object. Because of this square-cube relationship—where volume grows as the cube of linear dimensions while surface area grows only as the square—surface losses per unit of output will decline as output increases with an elasticity equal to $-1/3$. Driven by this force, we estimate a scale elasticity of energy losses around $-0.15$ for fired systems like furnaces (where surface losses are most important) and $-0.05$ for steam systems (where they are much lower).[20]

*Mechanical work.* In motors, magnetic core, friction and windage losses scale less than proportionally with motor power, so that motor efficiency increases at larger scales. More generally, in systems delivering mechanical work, useful mechanical work tends to grow faster with scale than parasitic losses, improving the ratio of

[20] Bosker and Buringh (2020) documents this mechanism in ice shipping, with the physics of the melt process implying economies of scale.

FIGURE 5: Energy Efficiency Investments by Firm Size




(A) Heating/Cooling Improvements

(B) Better Energy Management

(C) Machinery Improvements

(D) Lighting Improvements

*Notes:* Data from the World Bank Enterprise Survey module on energy efficiency for India in 2022. Panel A: indicates adoption of heating/cooling improvements. Panel B: indicates whether establishment is doing active energy management. Panel C: indicates whether establishment made improvements to machinery. Panel D: indicates adoption of lighting improvements. Plots show binned means, controlling for 2-digit industry fixed effects and applying sampling weights.

output to input energy. We model a combined motor-pump system, and find an elasticity of $-0.05$.

We leave the description of the (smaller) "Other process" and "Non-process" categories to Appendix C.

Averaging across end-uses, we find that these physical forces can rationalize an average "within-technology" scale elasticity around $-0.10$. This corresponds to 24% of the $-0.425$ physical elasticity we document in Table 3. The largest elasticity is for heat generation in fired systems. To confirm the relevance of this mechanism, in Table SM.16 we split our sample by the share of energy used for heat generation in fired systems (defined at the industry-level), and find substantially more negative scale elasticities for fired-systems-intensive industries.

**Technology Investment.** Many energy-saving technologies, such as waste heat recovery systems, improved insulation, or combined heat and power units, require substantial upfront investment. Because these costs are largely fixed, they are

more easily justified at large scales where savings can be spread over higher output. In addition, smaller firms often face financial constraints, making even profitable efficiency investments difficult to undertake.[21]

We document this fact using data from the World Bank Enterprise Survey (WBES) module on energy efficiency for India in 2022.[22] The WBES is a firm-level survey representative of the formal private sector. In 2022, the WBES asks sampled firms about adoption of four major types of energy efficiency investments over the past 3 years: heating and cooling improvements, energy management systems, machinery and equipment upgrades, and lighting systems upgrades. We obtain the answers to these yes/no questions, as well as the value of sales, and a two-digit industry, for 9,376 firms. Figure 5 shows that the probability of undertaking these investments is strongly increasing in firm size (measured by sales).

To understand why this pattern arises, in Figure 6 we present evidence that manufacturing firms face a menu of fixed cost investments, in which energy savings scale in the size of the upfront investment. We gather data from the United Nations Industrial Development Organization (UNIDO) Energy Efficiency and Renewable Energy Technology Compendiums for 2022. These compendiums document, for a set of key manufacturing industries in India, the existing technologies to improve energy efficiency at the plant level, and include data on annual energy savings and upfront costs for a typical plant (see details in Supplementary Material SM.1.2).

Figure 6 shows how energy savings scale with the size of the upfront investment for four industries: foundries, hand-tools, ceramics, and dairies. Take, for example, the case of foundries in Panel A. Foundries heat and melt metals to produce castings, both for use in the end-products manufactured by the firm, and made to order and sold on the market. The majority of energy wastage is in heat loss, and the compendium details several technologies to mitigate heat loss. At the smaller end, one technology is to place a lid mechanism over the mouth of the furnace, which is generally open, thereby reducing surface losses. This is costed at INR 0.35 million, and almost halves losses per heating cycle, saving 293 gigajoules per year for a typical plant. At the larger end, a cupola furnace can be replaced with an electric induction furnace. This bypasses the fuel combustion step where a large fraction of energy is lost, allows for finer temperature control, and reduces the manufacturing rejection rate. This is costed at INR 4 million, and saves 4,229

[21] Jordà, Nechio, Phan, and Schwartzman (2026) examine the link between financial conditions and energy-efficient investment, showing that lower discount rates tilt the capital stock towards greater energy efficiency.

[22] See Section SM.1.2 for further details on this data source.

gigajoules annually. A similar range of technologies exists across industries, with higher cost technologies generally giving higher energy savings.

When investment costs rise less than proportionally with scale (i.e., the investment has a fixed cost component), larger firms will optimally choose to adopt higher fixed cost and more energy-efficient technologies, contributing to a negative scale elasticity of energy demand. We formalize this intuition in the model presented in the next section, and show that, if firms act optimally, this margin can explain roughly 75% of our estimated effect.[23]

**Mechanisms and External Validity.** Because the mechanisms described are not specific to India, we suspect that the negative relationship between firm scale and energy use is a general feature of production. Figure SM.14 provides suggestive evidence for the generality of our finding using data from the U.S. Manufacturing Energy Consumption Survey. It plots energy intensity by bins of firm size (residualized on industry$\times$time fixed effects) for the cross-section of firms. While we cannot employ our causal research design here, we find that energy intensity falls with scale in a similar manner in the United States.[24]

# 4. A Theory of Aggregate Energy Intensity

We develop a model of manufacturing firms that turn raw materials into finished products. Following the preceding discussion, the energy intensity of production may display physical returns to scale within a given technology, and firms choose among technologies where higher fixed costs improve energy efficiency. We first characterize firms' choices, and then analyze how they shape aggregate energy intensity as economy-wide productivity improves. For analytical tractability, we make several simplifying assumptions, which we relax when taking the model to the data in Section 5. The proofs for all key results are in Appendix B.

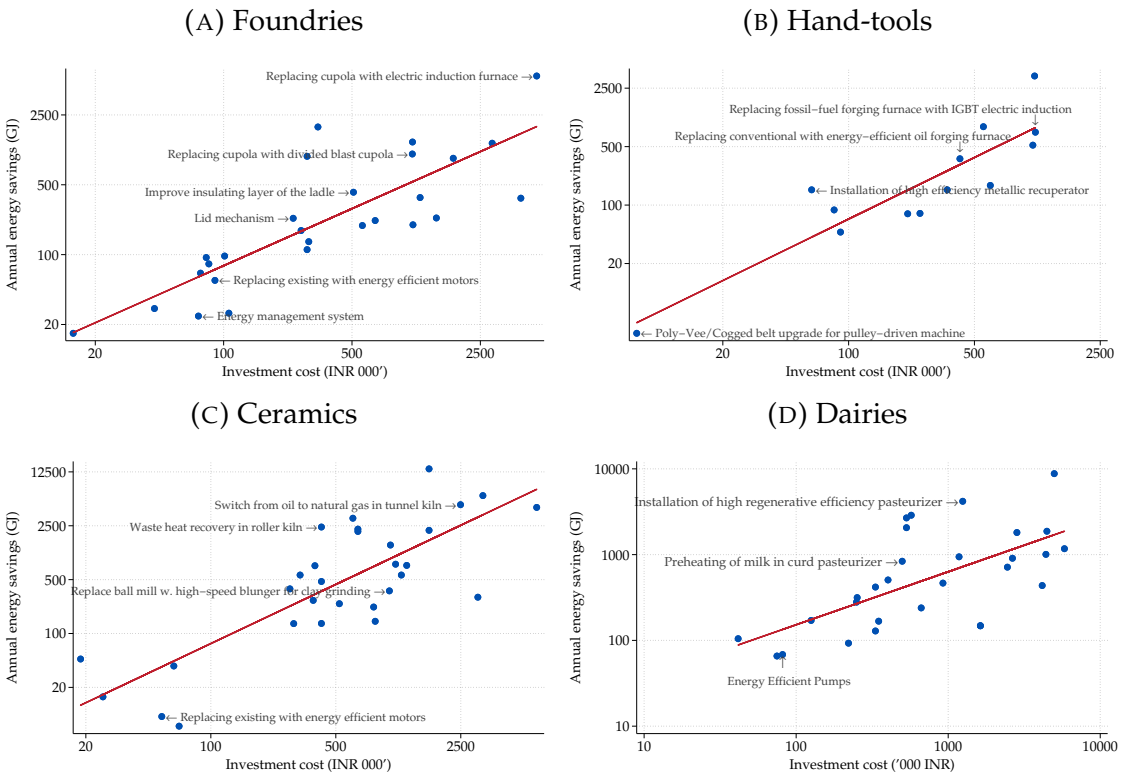
## 4.1 Setting

Time is discrete, and indexed by $t$. There is a continuum of identical households of total mass $L_t$. These households have static preferences over a final production

[23]This channel is powerful even if many of these energy-efficiency improvements require only modest upfront investments. Figure SM.13 shows that most technologies have payback periods below two years (defined as upfront cost divided by annual monetary savings). Figure OA.4 reports the implied investment rates, falling mostly below 2%. These technological improvements can be implemented in a standalone way or several may be combined.

[24]For the fixed cost investment channel to operate in the U.S., it must be that firms operate at a scale that is insufficient to rationalize some energy-efficiency investments. Figures SM.15 and SM.16 provide suggestive evidence that that is indeed the case.

FIGURE 6: Energy Technology Menus for Key Industries




*Notes:* Data from United Nations Industrial Development Organization (UNIDO) Energy Efficiency and Renewable Energy Technology Compendiums. When for a given industry×technology (e.g., "Waste heat recovery in tunnel kiln" in ceramics) we have several estimates (different case studies in the same industrial cluster, or estimates from different clusters), we take the average value.

good, whose output is given by

$$Y_t = \left[ \int_0^{N_t} a_{it} y_{it}^{\frac{\lambda-1}{\lambda}} di \right]^{\frac{\lambda}{\lambda-1}}, \tag{5}$$

where $N_t$ is the total mass of varieties operating in the economy. $a_{it}$ is a demand shifter that induces firms to produce at different scales. Households own all firms, and supply their labor inelastically to firms for a wage $w_t^l$, consuming $C_t$ of the aggregate production. We take the aggregate price index as our numeraire.

## 4.2 Firms

Firms produce output $y_{it}$ of a single variety $i$. For a given technology, the output of the firm is given implicitly by

$$y_{it} = A_{Yt} \min\{m_{it}, \phi_{it} \frac{e_{it}}{y_{it}^{\epsilon}}\}, \tag{6}$$

where $m_{it}$ is the amount of materials used, and $e_{it}$ is the amount of energy used, available at competitive prices $w_t^m$ and $w_t^e$ respectively. $A_{Yt}$ denotes an aggregate productivity shifter. We let $\epsilon < 0$ govern a physical scaling law such that producing more output can be done at a lower unit requirement of energy.[25]

The technology of the firm is indexed by $\phi$, which denotes an energy efficiency shifter, and a total capacity $\bar{y}$. A higher value for $\phi$ lowers the unit requirement of energy to produce $y$, and a higher value of $\bar{y}$ raises the total output that the firm can produce, since we suppose that $y \leq \bar{y}$. Before producing, the firm chooses investment in a capital and a labor bundle (e.g., a kiln with attendant staffing levels) that corresponds to a pair $(\phi, \bar{y})$. We suppose first that the cost of this technology bundle is given by

$$\tilde{C}_t(\phi, \bar{y}) = \frac{1}{\gamma A_{Yt}} \phi^{\gamma} \bar{y}^{\delta} (w_t^l)^{\alpha} (w_t^k)^{1-\alpha}, \tag{7}$$

where $w_t^l$ is the wage per unit of labor, and $w_t^k$ is the return on a unit of the capital good. In this way, firms can invest in a larger capital-labor bundle as a fixed cost, in order to both lower their marginal costs and increase the capacity of their plant.

Firms can be created by hiring $\nu$ units of labor. Once a firm is created, it draws its efficiency $a_{it}$ from a distribution $\Gamma(a)$. Firms live for a single period.

Energy can be created at rate $\Xi_t^e$ using the final good, and similarly for materials and capital at rates $\Xi_t^m$ and $\Xi_t^k$. As such, $w_t^e = 1/\Xi_t^e$, $w_t^m = 1/\Xi_t^m$, and $w_t^k = 1/\Xi_t^k$. In this Section, we assume $\Xi_t^m = \Xi_t^e = \Xi_t^k = 1$, and relax this in Section 5.

## 4.3 Optimality

In what follows, we suppress indexing by $i$. One can show that the energy efficiency of the technology chosen by the firm is given by

$$\phi = \left( \frac{w_t^e}{(w_t^l)^{\alpha} (w_t^k)^{1-\alpha}} \right)^{\frac{1}{1+\gamma}} y^{\frac{1+\epsilon-\delta}{1+\gamma}}. \tag{8}$$

As such, as long as the physical scale efficiencies are not too strong, and the investment cost doesn't rise too quickly with capacity so that $\delta < 1 + \epsilon$, the firm will choose to invest in a more efficient technology as it increases output.

[25] Appendix equation (OA.11) shows an example of a physical process for which the scaling law implies this functional form.

Firm cost minimization leads to a cost function over output $y$ given by

$$\mathcal{C}_t(y) = \frac{1}{A_{Yt}}\left(\frac{\gamma+1}{\gamma}((w_t^l)^{\alpha}(w_t^k)^{1-\alpha})^{\frac{1}{1+\gamma}}(w_t^e)^{\frac{\gamma}{1+\gamma}}y^{\tilde{\epsilon}+1} + w_t^m y\right) \tag{9}$$

where $\tilde{\epsilon} \equiv \frac{\epsilon\gamma+\delta-1}{1+\gamma}$. Importantly, if we define the non-material share of total cost at the firm level as

$$\Omega_t^{NM}(y) \equiv \frac{w_t^e e + w_t^l l + w_t^k k}{\mathcal{C}_t(y)},$$

then we can derive a useful expression relating this share to firm scale, as in

$$\frac{\Omega_t^{NM}(y)}{1-\Omega_t^{NM}(y)} = \frac{\gamma+1}{\gamma}((w_t^l)^{\alpha}(w_t^k)^{1-\alpha})^{\frac{1}{1+\gamma}}(w_t^e)^{\frac{\gamma}{1+\gamma}}y^{\tilde{\epsilon}}/w_t^m.$$

The non-material share of output is falling with firm size when $\tilde{\epsilon} < 0$. This can happen for two reasons. Even if the cost of the technology investment scales perfectly with capacity ($\delta = 1$), the existence of physical scaling laws $\epsilon < 0$ suffices for $\tilde{\epsilon} < 0$. Alternatively, even in the absence of physical scaling laws ($\epsilon = 0$), the technology investment scaling less than proportionally with capacity $\delta < 1$ also suffices for $\tilde{\epsilon} < 0$. The energy cost share is given by $\Omega_t^e(y) = \frac{\gamma}{1+\gamma}\Omega_t^{NM}(y)$, so that it inherits movements in the non-material cost share. Note that this implies that, as in the data, the energy share of the non-material bundle, $\Omega_t^{e,elk}(y)$, does not vary with firm scale.

A feature of non-homothetic production functions is that returns to scale vary across firms. To see this, define the cost elasticity (the inverse of returns to scale):

$$\mathcal{E}_t(y) \equiv \frac{d\log \mathcal{C}_t}{d\log y} = \tilde{\epsilon}\Omega_t^{NM}(y) + 1. \tag{10}$$

When $\tilde{\epsilon} < 0$, larger firms with lower non-material shares have higher cost elasticities, that is, lower returns to scale.

The standard CES demand function gives us an expression for inverse demand as $p(a,y) = ay^{-\frac{1}{\lambda}}Y_t^{\frac{1}{\lambda}}$. We can then write revenue as: $R(y) = ayy^{-\frac{1}{\lambda}}Y_t^{\frac{1}{\lambda}} \equiv ay^{\zeta}Y_t^{\frac{1}{\lambda}}$, where for convenience $\zeta \equiv \frac{\lambda-1}{\lambda}$. Then the problem of the firm is

$$V_t(a) = \max_y \quad ay^{\zeta}Y_t^{\frac{1}{\lambda}} - \mathcal{C}_t(y). \tag{11}$$

For a given distribution $\Gamma(a)$, the free entry condition then requires

$$w_t^l \nu = \int_a V_t(a) d\Gamma(a). \tag{12}$$

**Accounting for the Estimated Micro Scale Elasticity.** In the model, the firm-level scale elasticity in energy demand per unit of output arises from two channels: within technology, energy intensity changes with scale with elasticity $\epsilon$, and optimal technology choice $\phi$ changes with scale with elasticity $\frac{\epsilon+1-\delta}{1+\gamma}$.

$$\left.\frac{d \log\left(\frac{e}{y}\right)}{d \log y}\right|_{\mathbf{w}_t, A_{Yt}} = \underbrace{\tilde{\epsilon}}_{\text{Overall}} = \underbrace{\epsilon}_{\text{Within-tech.}} - \underbrace{\frac{\epsilon+1-\delta}{1+\gamma}}_{\text{Tech. improvement}}.$$

In the data, this elasticity is equal to $-0.425$ (Table 3, column 6). In the previous section, we showed that a reasonable estimate of $\epsilon$ is $-0.10$ (Table 6). That is, the within technology channel can rationalize 24% of the observed effect. The UNIDO data on technology menus faced by Indian manufacturing firms allows us to estimate the cost elasticity to energy savings $\gamma$. We estimate $\gamma = 0.20$. Finally, we use data on equipment prices by capacity to estimate $\delta = 0.52$. Details on the estimation of $\gamma$ and $\delta$ are provided in Appendix D. This implies that the technological improvement channel can account for an elasticity of $-0.32$, approximately 75% of our estimated effect. That is, assuming firms make the optimal technology decision, the two mechanisms quantitatively rationalize the scale elasticity estimated in the data.[26]

## 4.4 Equilibrium

**Definition 1.** *An equilibrium is a price vector* $\{w_t^l, w_t^e, w_t^m, w_t^k\}$*, and an allocation* $\{Y_t, N_t, E_t, M_t, K_t, C_t\}$*, such that*

*1. Given prices and aggregate output, firms solve* (11)*, with resulting input choices* $\{l_t(a), e_t(a), m_t(a), k_t(a)\}$ *and output choice* $y_t(a)$*,*

*2. The free entry condition* (12) *holds,*

[26] Table OA.8 shows that this conclusion is robust to realistic deviations of parameter values from our baseline.

*3. The labor market, the energy market, the materials market, and the capital market clear,*

$$L_t = N_t \int_a l_t(a) d\Gamma(a) + \nu N_t. \qquad E_t = N_t \int_a e_t(a) d\Gamma(a).$$
$$M_t = N_t \int_a m_t(a) d\Gamma(a). \qquad K_t = N_t \int_a k_t(a) d\Gamma(a).$$

*4. The goods market clears,*

$$Y_t = C_t + K_t/\Xi_t^k + M_t/\Xi_t^m + E_t/\Xi_t^e.$$

## 4.5 Growth and the Aggregate Energy Intensity of Production

We now discuss how economic growth affects aggregate energy intensity. We consider general growth in Hicks-neutral TFP, $A_Y$, as exogenous technical progress. Due to the lack of firm dynamics, we suppress indexing on $t$.

To do so, it will become apparent that wage growth, pushing up energy intensity through substitution effects, plays a crucial role in the overall movement of energy intensity. To shed light on the role of this force, we make one generalization of the above and rewrite the cost function in (9) as:

$$\mathcal{C}_t(y) = \frac{1}{A_{Yt}} \left( \left( \left(w_t^e y^{1+\tilde{\epsilon}}\right)^{1-\sigma_e} + \left( (w_t^l)^{\alpha} (w_t^k)^{1-\alpha} y^{1+\tilde{\epsilon}} \right)^{1-\sigma_e} \right)^{\frac{1}{1-\sigma_e}} + w_t^m y \right). \tag{13}$$

The scale elasticity $\tilde{\epsilon} \in (-1, 0)$ is as before and applies to both energy and the capital-labor bundle, but we allow for energy to enter as a CES nest with capital and labor, with an elasticity of substitution $\sigma_e$. As above, this cost function implies that the energy share of the non-material bundle, $\Omega_t^{e,elk}(y)$, does not vary with scale and hence is constant across firms.

We begin by characterizing the ratio of energy to gross output, and then show how this translates into movements in the ratio of energy to value added (GDP). The energy-to-gross-output ratio, which we call *gross energy intensity*, is given by

$$\bar{\Theta} \equiv \frac{N \int_0^\infty e(a) d\Gamma(a)}{Y} = \int_0^\infty \underbrace{\Theta(a)}_{\text{Micro energy intensity}} \cdot \underbrace{\frac{p(a)y(a)}{Y/N}}_{\text{Revenue weights}} d\Gamma(a), \tag{14}$$

where the micro energy intensity $\Theta(a) = \frac{e(a)}{p(a)y(a)}$ is the ratio of energy in joules to

dollars of revenue. As such, for any change due to fundamentals, we have

$$d\log\bar{\Theta} = \underbrace{\mathbb{E}_R[d\log\Theta(a)]}_{\text{Micro intensity changes}} + \underbrace{Cov_R\left[\frac{\Theta(a)}{\bar{\Theta}}, d\log\frac{p(y)y(a)}{Y/N} + d\log\Theta(a)\right]}_{\text{Reallocation}}, \tag{15}$$

where expectations $\mathbb{E}_R$ are taken over revenue weights.

Given the structure of production above, we can write the micro-level change in energy intensity as:

$$d\log\Theta(a) = \tilde{\epsilon}\, d\log y(a) - d\log(A_Y) + \sigma_e\left(1-\Omega^{e,elk}\right)\alpha\, d\log\left(w^l\right) - d\log p(a), \tag{16}$$

As such, there are four forces that change the micro energy intensity when Hicks-neutral productivity $A_Y$ increases. First, firm size increasing reduces $\Theta(a)$, moderated through $\tilde{\epsilon}$, the micro scale elasticity of physical intensity $e/y$. Second, increasing Hicks-neutral productivity $A_Y$ increases output $y$ holding inputs constant, and causes a fall in energy intensity directly across all firms. Third, the wage $w^l$ changes, inducing factor substitution with energy, which is governed by $\sigma_e$. Finally, energy intensity relative to revenue is affected by the change in the firm's price $p(a)$ in response to aggregate growth. We separate this fourth term: since $\mathbb{E}_R[d\log p(a)] = 0$, this term does not affect aggregate energy intensity.

Combining (15) and (16), we characterize how the aggregate energy intensity of gross output changes as aggregate productivity $A_Y$ changes.

**Proposition 1.** *The change in the aggregate gross energy intensity $\bar{\Theta}$ with a rise in factor-neutral productivity $A_Y$ is*

$$d\log(\bar{\Theta}) = \underbrace{\tilde{\epsilon}\mathbb{E}_R\left[\,d\log y(a)\right]}_{\textit{Micro Scale Up}} \underbrace{+\alpha\,\sigma_e\left(1-\Omega^{e,elk}\right)\,d\log\left(w^l\right)}_{\textit{Factor Price Change}}$$

$$\underbrace{-d\log(A_Y) - \left(d\log(Y/N) - \mathbb{E}_R[d\log y(a)]\right)}_{\textit{Aggregate TFP and Returns to Scale}} + \underbrace{Cov_R\left[\frac{\Theta(a)}{\bar{\Theta}}, (\tilde{\epsilon}+1)d\log y(a)\right]}_{\textit{Reallocation}}, \tag{17}$$

The first term corresponds to the direct contribution of the firm scaling up channel: the product of the micro elasticity $\tilde{\epsilon}$ and the average increase in firm size.[27] This expression clarifies why the micro energy intensity elasticity, $\tilde{\epsilon}$, may differ from the macro energy intensity elasticity to aggregate output; or equivalently, why

[27] Note that this term is mediated by the micro scale elasticity of physical intensity $e/y$, as opposed to revenue intensity $e/py$. Indeed, the micro elasticity of revenue intensity accounts for changes in relative prices in response to relative quantity growth, yielding an elasticity $\tilde{\epsilon} + \frac{1}{\lambda}$. This relative price response does not enter the aggregate elasticity of energy revenue intensity.

the aggregate reduction in energy intensity in response to an aggregate increase in scale is not simply the first term. The second and third terms are as in (16), but where the increase in TFP is augmented by aggregate returns to scale coming from the CES aggregator for the final good, if the number of firms increases. The final term—reallocation—captures the fact that smaller firms, with higher energy cost shares, have greater returns to scale, and so grow more in response to an increase in TFP. This acts to push up energy intensity.

While the first and third terms are unambiguously negative, the second and fourth terms are positive. Aggregate growth can in fact increase or decrease energy intensity in this simple model, depending on the strength of the substitution effect induced by rising wages (the Factor Price Change term in (17)). This in turn depends on the value of the elasticity of substitution between electricity and the capital and labor bundle, $\sigma_e$. To see this, we can shut down firm heterogeneity, letting $\Gamma(a)$ be degenerate. We then have the following result.

**Proposition 2.** *There exists a* $\sigma_e^* < 1$ *such that, at any regular, symmetric equilibrium in which* $d \log y / d \log A_Y > 0$, *if* $\sigma_e < \sigma_e^*$, $d \log \bar{\Theta} / d \log A_Y < 0$

In words, if energy is sufficiently complementary with other inputs, then a rising wage does not cause much of a reallocation towards energy as aggregate growth pushes up wages, and aggregate energy intensity falls with TFP. The dominant force is then the micro-scaling up channel of Proposition 1, with firms decreasing their energy intensities at higher scale.

This characterizes direct energy use by firms relative to gross output. However, producing a unit of output also requires energy indirectly through the energy embodied in intermediate inputs. Let us define the total energy content of production as the sum of direct energy use and indirect energy use embodied in material and energy inputs used along the entire input-output chain, normalized by gross output. Let $\Psi_t^x \equiv w_t^x x_t / Y_t, x \in \{M, E, L, K\}$ be the cost shares of inputs in total output. Then, the total energy content of production is given by:

$$\frac{E^{tot}}{Y} = \frac{\bar{\Theta}}{1 - \Psi^e - \Psi^m} = \frac{E}{VA} \tag{18}$$

That is, the total energy content of production is equal to the ratio of energy use to value added, or *net energy intensity*.[28] This quantity also falls with growth

[28] See the proof in Supplementary Material SM.3.5. Here, we define value added including the cost of capital, even if capital is a produced input in our model, to match the definition of GDP in the data.

when the condition of Proposition 2 is satisfied. This is despite the fact that the denominator $1 - \Psi^e - \Psi^m = 1 - w^e\bar{\Theta} - w^m M/Y$ may increase with aggregate output. This may not be true in richer models of production in which the ratio of intermediates to production $M/Y$ rises with aggregate scale. In such models, it is at least theoretically possible that joules per dollar of GDP could rise. However, this is not the case in the Indian data, nor for the world as a whole (see Figure 1).

# 5. Quantifying the Theory

We now add several extensions to take the model to the data. In particular, we allow a more flexible cost function that nests our model above as a special case.

We begin with equation (13), and make three strict generalizations. First, we allow materials to have a non-zero elasticity of substitution with other factors of production, denoted by $\sigma_m$. Second, we allow for general scale elasticities that may be different for each factor, denoted $\epsilon_e$, $\epsilon_l$ and $\epsilon_m$. These parameters regulate how the shares of energy, capital/labor and materials evolve with firm scale. In conjunction, they also control the overall returns to scale of the production function. Third, we allow for the time-varying aggregate shifters $A_{Yt}$, $A_{et}$ and $A_{mt}$, as exogenous forms of technical progress. $A_{Yt}$ is a factor-neutral shifter, while $A_{et}$ and $A_{mt}$ are factor-augmenting technical change for energy and materials, respectively.

This leaves us with a generalized non-homothetic CES cost function[29] given by

$$C_t(y) = A_{Yt}^{-1}\left((A_{mt}^{-1}w_t^m y^{\epsilon_m})^{1-\sigma_m} + \left((A_{et}^{-1}y^{\epsilon_e}w_t^e)^{1-\sigma_e} + (y^{\epsilon_l}(w_t^l)^{\alpha}(w_t^k)^{1-\alpha})^{1-\sigma_e}\right)^{\frac{1-\sigma_m}{1-\sigma_e}}\right)^{\frac{1}{1-\sigma_m}}. \tag{19}$$

## 5.1 Estimation Strategy

Our estimation strategy proceeds in three stages. First, we estimate all production function parameters from (19)—the two substitution elasticities $\sigma_m$ and $\sigma_e$, and the three scale elasticities $\epsilon_e$, $\epsilon_m$ and $\epsilon_l$—using causal empirical estimates from the micro data. Identification follows from analytical mappings from the estimated causal elasticities in the micro data to the production function parameters. These parameters form the core of our quantitative exercise, and the ability to estimate them directly with causal estimates is a considerable strength of the approach. We set $\alpha$ equal to labor's share of the labor-capital bundle in the base year $\Omega_{1990}^{l,lk}$. Second, we calibrate the demand elasticity externally ($\lambda = 4$), following the firm

[29] See Lashkari et al. (2024), Eckert, Ganapati, and Walsh (2025) and Trottner (2019) for other recent examples of production functions of this form in aggregate analysis.

dynamics literature (see e.g. Garcia-Macia, Hsieh, and Klenow (2019); Peters and Walsh (2026)). Third, we use Simulated Method of Moments (SMM) to estimate the remaining model parameters (the productivity shifters $A_e, A_m$, the entry cost $\nu$, the shape parameter of the Pareto distribution $\theta$).

**Elasticities of Substitution ($\sigma_e, \sigma_m$).** We estimate the elasticity of substitution between energy and the capital-labor bundle, $\sigma_e$, and between materials and other factors, $\sigma_m$. To do so, we estimate the effect of changes in the price of energy on the energy-to-labor and energy-to-material expenditure ratios. We instrument for the change in the price of energy: we construct a shift-share instrument which interacts changes in the aggregate prices of different fuels (coal, electricity, oil) and firms' ex-ante exposure to these different fuel types. We obtain $\sigma_e = 0.48$ and $\sigma_m = 0.30$. That is, energy is strongly complementary with the capital-labor bundle, and materials are strongly complementary with other inputs.[30]

**Scale Elasticity Parameters ($\epsilon_l, \epsilon_e, \epsilon_m$).** The parameters governing the scale elasticities $\epsilon_e$, $\epsilon_l$, and $\epsilon_m$ are chosen to target three moments: (i) how the *energy share relative to materials and labor* changes with firm size (the main scale elasticity of our empirical analysis), (ii) how the *energy share relative to labor* changes with firm size, and (iii) the average cost elasticity, or the inverse of the *degree of returns to scale* at the firm level. These scale elasticities are defined with respect to firm physical output (as opposed to revenues):

$$\eta_{e/ml} \equiv \frac{d\log\left(\frac{\Omega^e}{\Omega^m+\Omega^l}\right)}{d\log y}, \quad \eta_{e/l} \equiv \frac{d\log\left(\frac{\Omega^e}{\Omega^l}\right)}{d\log y}, \quad \mathcal{E} \equiv \frac{d\log \mathcal{C}}{d\log y}$$

We derive analytical expressions for these elasticities in terms of $\epsilon_e$, $\epsilon_l$ and $\epsilon_m$ as:

$$\eta_{e/ml} = (1-\sigma_e)(\epsilon_e-\epsilon_l)\left[1-\Omega^{m,ml}\Omega^{e,elk}\right] + (1-\sigma_m)\left[\epsilon_e\Omega^{e,elk}+\epsilon_l(1-\Omega^{e,elk})-\epsilon_m\right]\Omega^{m,ml}, \tag{20}$$

$$\eta_{e/l} = (1-\sigma_e)(\epsilon_e-\epsilon_l), \tag{21}$$

$$\mathcal{E} = \Omega^m\epsilon_m + (1-\Omega^m)(\Omega^{e,elk}\epsilon_e + (1-\Omega^{e,elk})\epsilon_l). \tag{22}$$

where $\Omega^{e,elk} \equiv \Omega^e/(\Omega^e+\Omega^l+\Omega^k)$ and $\Omega^{m,ml} \equiv \Omega^m/(\Omega^m+\Omega^l)$.[31]

We estimate the cost-weighted elasticities on the left-hand side of (20) and (21) by using specification (3) with our demand shifter as an instrument for the change

[30] We describe this analysis in detail in the Supplementary Material SM.4.2.
[31] See details in Supplementary Material SM.4.

TABLE 7: Model Estimation

| | Model Parameters | Value | Targeted Moments | Data | Model |
|---|---|---|---|---|---|
| *Production Function Parameters* | | | | | |
| $\epsilon_m$ | Scale elasticity of materials | 1.19 | Avg. cost elasticity | 0.96 | 0.96 |
| $\epsilon_l$ | Scale elasticity of L+K bundle | 0.40 | Output elasticity $\eta_{e/l}$ | 0.0 | 0.0 |
| $\epsilon_e$ | Scale elasticity of energy | 0.40 | Output elasticity $\eta_{e/ml}$ | -0.51 | -0.51 |
| $\sigma_e$ | E.O.S. energy and labor | 0.48 | Price elasticity $\frac{d \log \Omega^{e,l}}{d \log w^e}$ | 0.53 | 0.53 |
| $\sigma_m$ | E.O.S. materials and other | 0.30 | Price elasticity $\frac{d \log \Omega^{e,m}}{d \log w^e}$ | 0.58 | 0.58 |
| $\alpha$ | Labor share in inner-most nest | 0.28 | Directly calculated, $\Omega^{l,lk}_{1990}$ | | |
| *Other Parameters* | | | | | |
| $A_{e,1990}$ | Energy shifter | 5.75 | Energy share $\Omega^e_{1990}$ | 0.09 | 0.09 |
| $A_{m,1990}$ | Materials shifter | 18.96 | Materials share $\Omega^m_{1990}$ | 0.71 | 0.71 |
| $\nu$ | Entry cost | 62.71 | Avg. employees/firm | 75 | 75 |
| $\theta$ | Pareto shape | 5.56 | Size dist. (p75-p25 $\log(\Omega^m)$) | 0.15 | 0.15 |
| $\lambda$ | Demand elasticity | 4 | Set externally | | |

*Notes*: This table shows the estimated parameters for the quantitative model of Section 5.

in firms' physical output, with the results reported in Table OA.3. We estimate the cost-weighted average total cost elasticity (the left-hand side of equation (22)) by using our demand shifter $\mathcal{D}^h_{it}$ to trace out how cost moves with increases in scale. We estimate average returns to scale of approximately 1.[32] Evaluating the cost-shares on the right-hand side of (20)-(22) at their cost-weighted averages ($\Omega^x = \mathbb{E}_C[\Omega^x_i]$), we solve these three equations to obtain values $\epsilon_m = 1.19$ and $\epsilon_l = \epsilon_e = 0.40$.[33]

**Other Parameters.** $\alpha$ is the expenditure share of labor in the inner-most nest in the base year, $\Omega^{l,lk}_{1990} = 0.28$. All the other parameters are jointly estimated by SMM. We estimate the exogenous shifters $A_{m,1990}$, $A_{e,1990}$ in the base period to match the expenditure share of materials $\Omega^m_{1990} = 0.71$ and energy $\Omega^e_{1990} = 0.09$, respectively, in the year 1990. We estimate the entry cost, $\nu$, measured in units of labor, to match the average firm size, which is 75 employees in 1990.

Finally, we assume that the distribution of the idiosyncratic firm demand shifters, $a_i$, follows a Pareto distribution with shape parameter $\theta$. Firm heterogeneity matters to the extent that it shapes the reallocation term highlighted in equation (17). This term depends on the dispersion in firm-level growth rates in response to changes in aggregate TFP $A_{Yt}$, which is fully determined by the dispersion in

[32] We describe this analysis in detail in the Supplementary Material SM.4.3.

[33] Note that we get almost identical quantitative results when calibrating to alternative scale elasticity targets that include capital in the denominator $\eta_{e/mlk}, \eta_{e,lk}$.

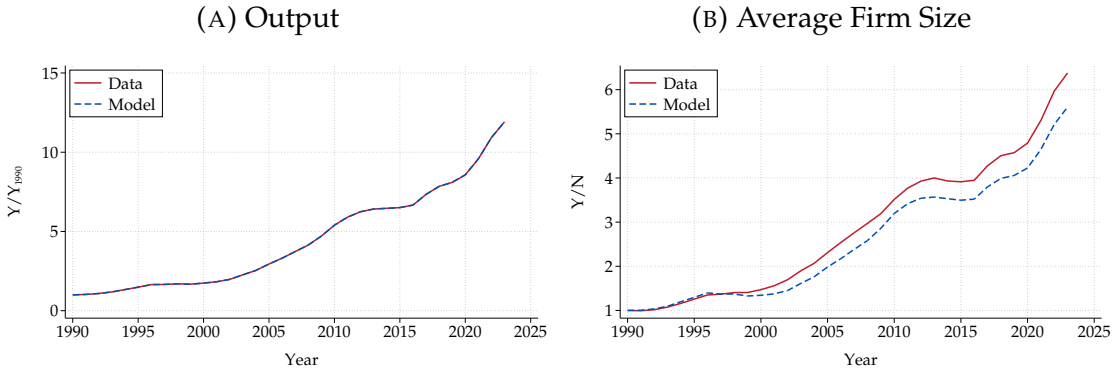
returns to scale. Firm-level returns to scale are themselves fully determined by the firm material share.[34] We thus estimate $\theta$ as follows: we project the log material share on firm size (since in our model input shares solely depend on firm size), and then target the interquartile range (p75–p25) of these fitted material shares in the data. Our estimation yields $\theta = 5.56$.

**Model Fit.** By construction, the model matches the firm-level relationship between firm size and the shares of energy relative to materials and labor $\Omega^e/(\Omega^m + \Omega^l)$, and of energy relative to labor $\Omega^e/\Omega^l$ (Figure OA.3(A) and (B), respectively).

Turning to untargeted moments, the model closely matches the positive scale elasticity of the relative materials expenditure share $\Omega^m/(1-\Omega^m)$, as can be seen from Figure OA.3(C). It also closely matches the scale elasticity of the relative labor share $\Omega^l/(1-\Omega^l)$ and relative energy share $\Omega^e/(1-\Omega^e)$ (Figure OA.3(D,E)).[35] The model also reproduces the fact that the energy-output-share elasticity ($w^e e/py$) is more negative than the energy cost-share elasticity (Figure OA.3(F)). This occurs because profit rates rise with firm size, both in the model and in the data. In the model, this is driven by the fact that larger firms have lower returns to scale, and hence higher profit rates. Quantitatively, the model predicts an energy-output-share elasticity of $-0.50$, smaller than the empirical estimate of $-0.66$. This is because, while the model qualitatively captures the positive relationship between size and profitability, it understates its magnitude.[36] We also closely match the firm-level scale elasticity of physical energy intensity, defined as energy quantity over output quantity ($e/y$, Figure OA.3(G)).

To assess the ability of our model to capture the dispersion in returns to scale—which is key for the reallocation effect—we compare the cost-weighted and unweighted average cost elasticity (the inverse of returns to scale) in the model and in the data. We explicitly target the cost-weighted estimate of 0.96 in the data. The unweighted cost elasticity, which we do not explicitly target, is 0.91 in the model and 0.88 in the data (Figure OA.3(H)). The model and data estimate agree on returns to scale being higher for small compared to large firms, and on the quantitative difference between the unweighted and cost-weighted elasticity.

FIGURE 7: Aggregate Output and Average Firm Size: Model vs. Data




*Notes:* This figure plots the model implied path of output and average firm size in response to an increase in output driven by scaling the aggregate productivity shifter $A_Y$ in the baseline model. The dashed blue line follows the model calibration assuming a non-homothetic CES production function described in Section 5.1. The red line represents the aggregate ASI time series.

## 5.2 Results

We model growth as driven by growth of the labor force $L_t$ and productivity growth $A_{Yt}$. Specifically, we feed in the growth rate of the labor force $L_t$ (taken to be the population in India aged 15-64, from UN data) and estimate $A_{Yt}$ to exactly match the aggregate growth rate of output in the ASI data over 1990-2023. Figure 7(A) shows that we (by construction) perfectly match this series. We analyze how the aggregate expenditure share of energy and energy intensity change in our model in response to output growth. In doing so, we solve the full general equilibrium solution, including endogenous wages and firm choices. In our baseline calibration, we hold fixed the technical efficiencies $\Xi_t^k$, $\Xi_t^m$ and $\Xi_t^e$, which govern the relative prices of capital, materials, and energy, respectively. We return to these momentarily.

As shown in Figure 7(B), the model closely mirrors the increase in average firm size in the data. In the model, this emerges because entry costs are paid in units of labor, and hence rise with TFP growth (Klenow and Li 2025). This parsimonious assumption fits the Indian time series remarkably well. Fitting the increase in average size is important for our purpose, since it is the key forcing variable of our mechanism.

Panel A of Figure 8 presents our key result: our estimated model generates a 45% decline in the aggregate energy expenditure share from 1990 to 2023. This

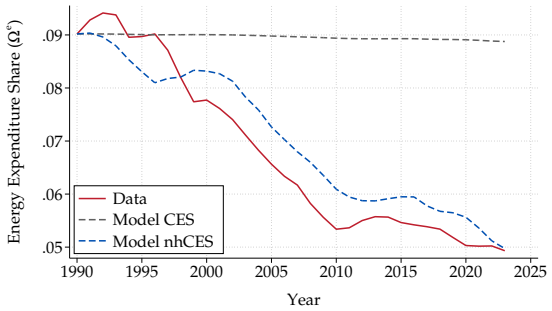
[34] See equation (22) for the cost elasticity when $\epsilon_e = \epsilon_l$.

[35] Because cost shares sum to 1, it follows that we match the scale elasticity of the capital share.

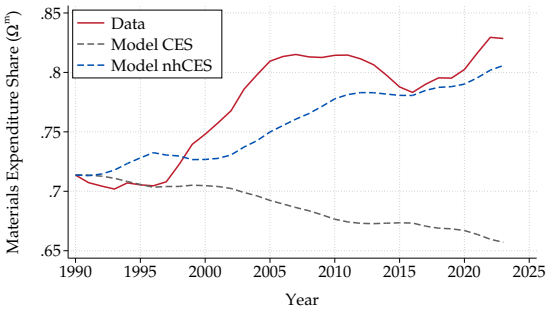
[36] A likely explanation, outside of the scope of our model, is that markups increase in firm size (Edmond, Midrigan, and Xu 2023).

FIGURE 8: Energy and Materials Cost Shares: Model vs. Data

(A) Energy Share




(B) Materials Share




*Notes:* This figure plots the model path of the aggregate energy and materials expenditure shares in response to an increase in output matching the data (dashed blue line). The dashed gray line assumes a homothetic (CES) production function. The red line is the aggregate ASI time series.

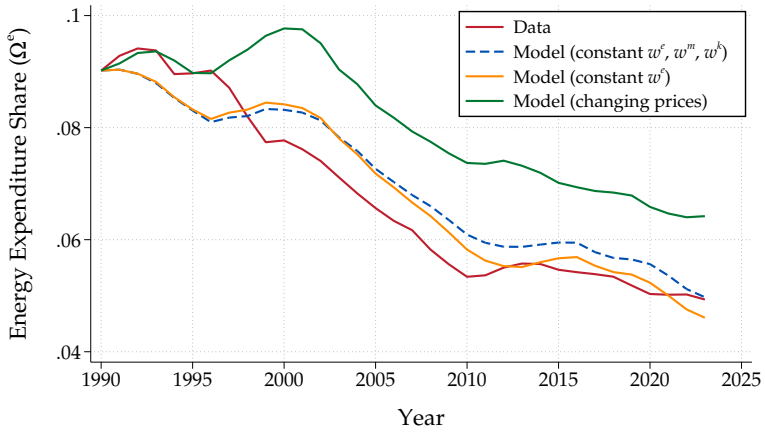
corresponds to almost all of the actual decline during this period, as illustrated by the red line. The decline in the energy share is entirely attributable to the non-homotheticity in production. To isolate this mechanism, the dashed gray line reports a counterfactual with a homothetic CES production function (obtained by setting $\epsilon_m = \epsilon_e = \epsilon_l = \mathcal{E}$ in equation (22), disciplined by the average cost elasticity estimated in the data). In this case, the energy share barely changes.[37] The non-homothetic CES model predicts a macro-elasticity of the relative energy expenditure share with respect to total output equal to $-0.39$ (Table OA.4). This is very close to the firm-level counterpart of this elasticity in the data, equal to $-0.43$ (Table OA.3).

In addition, the model matches well the trend in the share of materials expenditures (Figure 8(B)). As with energy, the pattern is entirely driven by non-homotheticity. Matching this ensures that the model captures both direct and indirect energy use, and therefore provides a consistent account of total energy demand.

We then consider a scenario in which the relative price of energy changes in our model to match the rise in the relative energy price observed in the data from 1990-2023 (Figure SM.4). Figure 9 displays three model-predicted time series. In blue, we keep the price of capital, intermediates and energy equal to that of the final good, reproducing Panel A of Figure 8. In orange, we match the time series for the relative price of capital and intermediates, but leave the relative price of energy constant. The model prediction remains very similar. In green, we show the model prediction when we additionally match the rise in the

[37] Note that the change is not exactly 0 because TFP growth raises the wage relative to other input prices (equal to the final good's price) and substitution is non-unitary.

FIGURE 9: Impact of Changing Energy Price on Modeled Energy Share




*Notes:* This figure plots the model implied decline in the aggregate energy share under three scenarios. The dashed blue line simulates the model holding exogenous technical efficiencies fixed (fixed prices). The orange line allows the price of materials and the price of capital to match the data series. The green line also lets the price of energy match the relative price series in the data. The red line represents the aggregate ASI time series.

relative energy price. This causes the overall energy share to decline less than in a scenario assuming constant prices. The reason is that energy is estimated to be complementary with other inputs in our model, so a rising price, all else equal, implies an *increasing share* of energy expenditure at the micro level, and this carries over to the aggregate. In this scenario—where model prediction and data are most comparable—the model explains about 64% of the overall decline in the energy cost share. In addition, had the relative price of energy in India stayed constant, the energy expenditure share would have fallen even more.

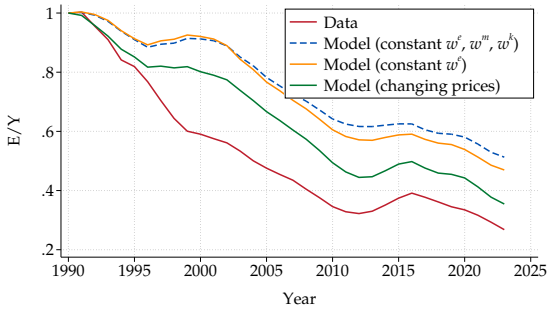
Finally, we analyze what our model predicts for the path of energy intensity from 1990-2023 compared to the data. In Panel A of Figure 10, energy intensity is defined as the physical energy in production over total output, while in Panel B, it is defined relative to total value added. Allowing the relative price of energy to match the data, our model predicts a 65% decline in energy intensity, highly similar in panels A and B, corresponding to 88% of the decline in the data.

## 5.3 Growth Decomposition

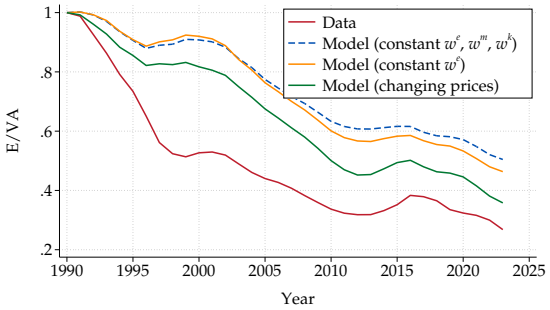
To further unpack the drivers of the decline in the aggregate energy share through the lens of our model, we perform a full growth accounting exercise. We analyze which components of the quantitative model are most important in explaining the aggregate trends of two key measures: (1) aggregate energy intensity $\bar{\Theta}_t$, and (2) the energy cost share $\Omega^e_t$.

FIGURE 10: Energy Intensity: Model vs. Data

(A) Energy / Output




(B) Energy / Value Added




*Notes:* This figure plots the model implied paths of energy intensity measured as Energy/Aggregate Output in Panel A, and Energy/Value Added in Panel B. The dashed blue line simulates the model holding exogenous technical efficiencies fixed (fixed prices). The orange line allows the price of materials and the price of capital to match the data series. The green line also lets the price of energy match the relative price series in the data. The red line is the ASI time series for energy over output and energy over value added.

Using the cost function in (19), we can derive a similar result to Proposition 1:

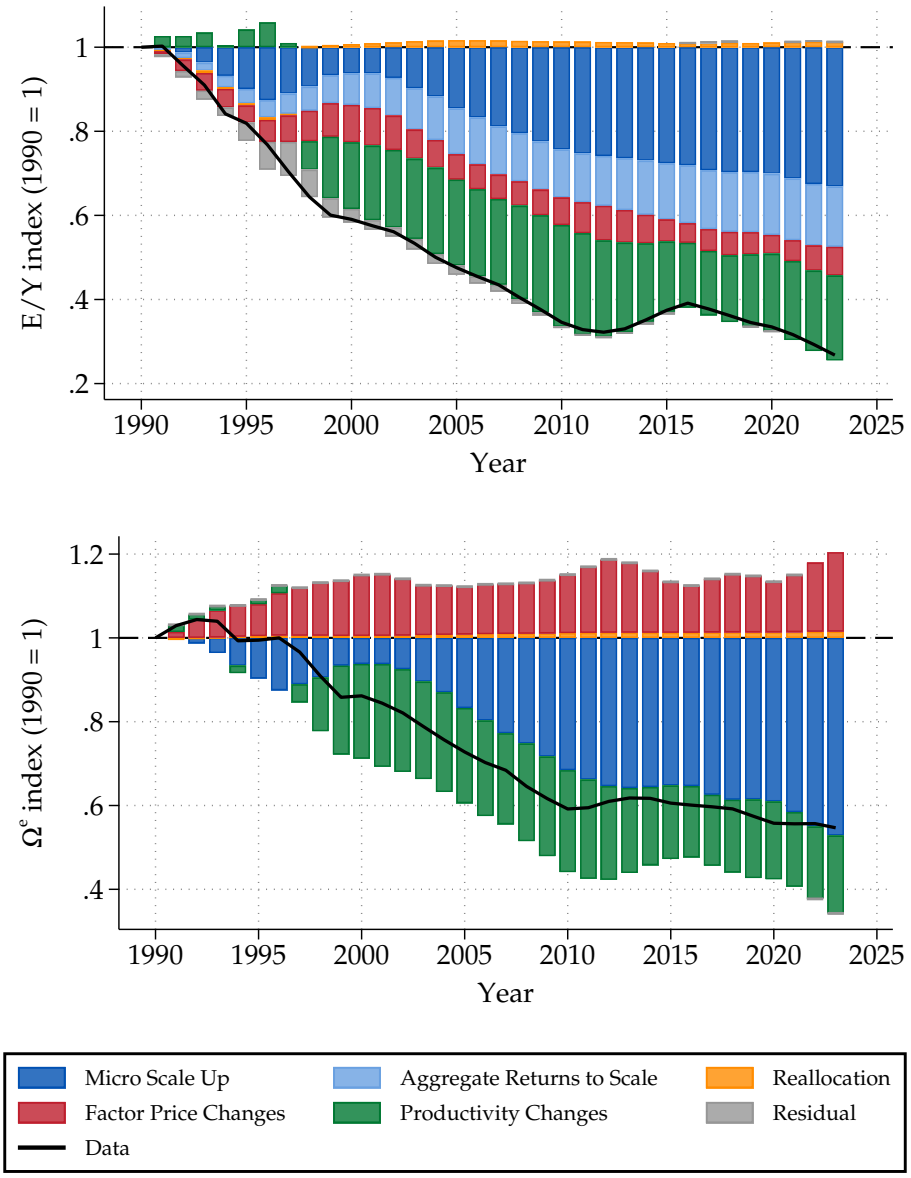
$$
\begin{aligned}
d\log\bar{\Theta} = & \underbrace{\mathbb{E}_R\left[\tilde{\epsilon}_Q(a)\, d\log y(a)\right]}_{\text{Micro Scale Up}} \\
& + \underbrace{\left(\mathbb{E}_R[d\log y(a)] - d\log(Y_t/N_t)\right)}_{\text{Aggregate Returns to Scale}} + \underbrace{\text{Cov}_R\left[\frac{\Theta(a)}{\bar{\Theta}},\, \left(\tilde{\epsilon}_Q(a)+1\right)\, d\log y(a)\right]}_{\text{Reallocation}} \\
& + \underbrace{\mathbb{E}_R^{\Theta}[\beta_m(a)]d\log w_t^m + \left(\mathbb{E}_R^{\Theta}[\beta_e(a)]-1\right) d\log w_t^e + \mathbb{E}_R^{\Theta}[\beta_l(a)]\left[\alpha d\log w_t^l + (1-\alpha)d\log w_t^k\right]}_{\text{Factor Price Changes}} \\
& - \underbrace{\mathbb{E}_R^{\Theta}[\beta_m(a)]d\log A_{mt} - \mathbb{E}_R^{\Theta}[\beta_e(a)]d\log A_{et} - d\log A_{Yt}}_{\text{Productivity Changes}}.
\end{aligned}
$$

where $\beta_m(a)$, $\beta_e(a)$ and $\beta_l(a)$ are functions of the substitution elasticities $\sigma_e$ and $\sigma_m$, $\mathbb{E}_R^{\Theta}$ are revenue-and-intensity-weighted expectations, and $\tilde{\epsilon}_Q(a)$ is the micro scale elasticity of energy intensity in the quantitative model.[38] The main differences from Proposition 1 are two-fold. First, there are effects coming from the additional technology shifters $A_{mt}$ and $A_{et}$. Second, we allow factor prices to vary to match the data, and this can induce additional substitution.

The result is presented in Figure 11, along with a parallel decomposition of the aggregate energy expenditure share, $\Omega^e$. The firm-level scale channel we have emphasized accounts for essentially half of the aggregate energy intensity decline, and all of the decline in the expenditure share. In addition, the growth in the productivity shifters $A_Y$, $A_e$ and $A_m$ and aggregate returns to scale play

[38] See derivations in Supplementary Material SM.4.4.

FIGURE 11: Decomposition of Energy Cost Share and Energy Intensity Changes




*Notes:* The top panel decomposes aggregate energy intensity into 6 components: (1) firm scale, (2) aggregate returns to scale, (3) reallocation, (4) the change in factor prices, (5) the effect of factor-specific and factor-neutral productivity shifters $A_m$, $A_e$ and $A_Y$, and (6) a numerical residual. The residual in the E/Y decomposition also contains differences between model and data time series. The bottom panel decomposes the aggregate energy cost share into 5 components: (1) firm scale, (2) reallocation, (3) the change in factor prices, (4) the effect of factor-specific productivity shifters $A_m$, $A_e$, and (5) a numerical residual.

meaningful roles in the decline in energy intensity.[39] An increase in the price of energy also contributed to the decline in $E/Y$, but plays the opposite role for

[39]The exact contribution of aggregate returns to scale relative to $A_Y$, $A_e$ and $A_m$ depends on the calibration of $\lambda$, with lower values increasing the importance of this channel.

the expenditure share $\Omega^e$.[40] Again, this is because we estimate that energy is complementary to other inputs, and the increase in its relative price in the data acts to push up the expenditure share. Almost fully offsetting this is an increase in energy-specific technical change through $A_e$, pushing the expenditure share the other way.

**Direct and Indirect Energy Intensity.** Figure 12 decomposes the change in the *total* energy content of production (comprised of both direct and indirect energy use embodied in materials, energy, and capital) into a change in the direct energy share and a change in indirect energy use, due to changing shares of other inputs. From equation (18), this is equal to the energy-to-value added ratio.[41] Denote $\Lambda \equiv \Psi^m + \Psi^k + \Psi^e$. We can then decompose:

$$d\log\left(E^{tot}/Y\right) = \underbrace{d\log(E/Y)}_{\text{direct}} + \underbrace{\frac{\Psi^m}{1-\Lambda} d\log\Psi^m}_{\text{indirect via M}} + \underbrace{\frac{\Psi^k}{1-\Lambda} d\log\Psi^k}_{\text{indirect via K}} + \underbrace{\frac{\Psi^e}{1-\Lambda} d\log\Psi^e}_{\text{indirect via E}} \tag{23}$$

The majority of the decline in the total energy intensity of production $E^{tot}/Y$ comes from the decline in direct energy use $E/Y$, in green. Since the materials share rises, the indirect energy share via materials (in blue) contributes positively to the total energy share, a small counteracting force.

## 6. Energy and Emissions in India: 2025-2050

In this final section, we consider the implications of our findings for the future trajectory of energy use and emissions in India.

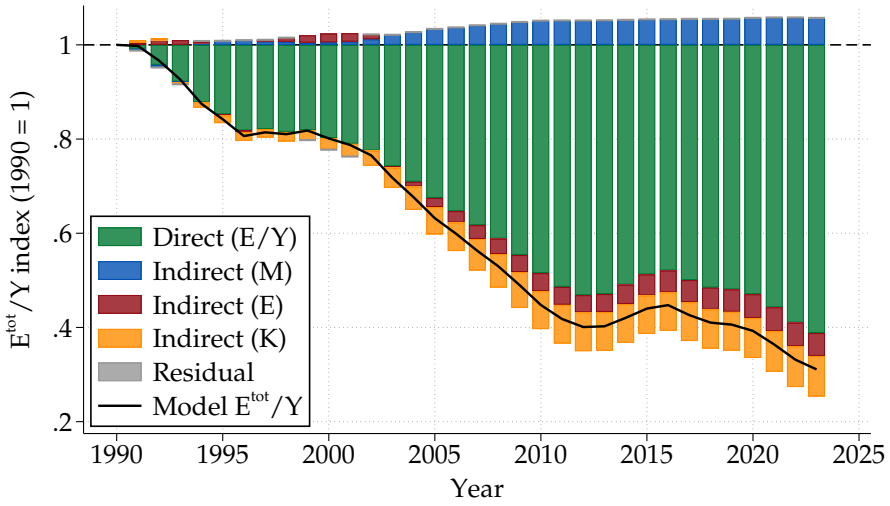
We analyze what GDP and labor supply growth projections through 2050 imply for total energy demand in the medium term. We assume the labor force grows according to the projections from the UN Population Division through 2050. For aggregate output growth, we use the IMF's GDP growth projections from 2025-2031, and assume that the annual growth rate in GDP from 2032-2050 matches the annual growth rate from 2025-2031. This implies a path for productivity growth consistent with the observed manufacturing productivity growth in India over the last 20 years, obtained from KLEMS. Lastly, this exercise assumes that our analysis of the manufacturing sector is representative of the economy as a whole.

Figure 13(A) shows the projected path of energy demanded under two scenarios.

[40] Figure SM.17 breaks out the sub-components of the productivity and price terms.

[41] For this decomposition, we treat capital as a produced input (consistent with our model), so that equation (18) becomes $E^{tot}/Y = \bar{\Theta}/(1-\Psi^e-\Psi^m-\Psi^k)$. This equals the energy-to-value added ratio, when value added excludes capital costs. This time series closely follows our baseline energy-to-value added series.

FIGURE 12: Direct and Indirect Energy Intensity




*Notes:* This figure implements the decomposition in equation (23), for the model with changing input prices. The black line is the left-hand-side; the bars are the components on the right-hand-side.

The first scenario extrapolates future energy demand while allowing for the scale dependence of energy use—our baseline model. The second scenario assumes that the energy cost share remains flat at 4.9% from 2023-2050. We find the total energy demanded in 2050 is 49% smaller in our model, compared to a scenario where the energy share stays constant.[42]

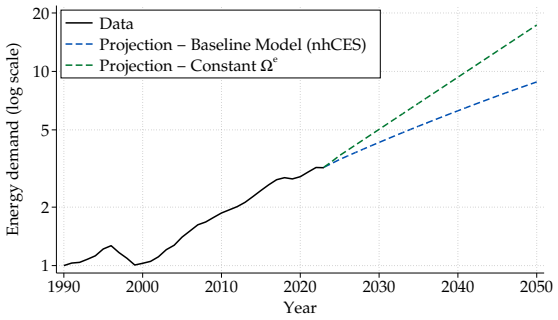
Finally, in Figure 13(B), we analyze what our model requires for the evolution of the ratio of Greenhouse Gases to Energy (CO2e/Energy) to meet the Paris Climate Agreement goals. India's goal under the UN's Nationally Determined Contributions (NDCs) is to reduce its emissions intensity by 47% by 2035 from its 2005 levels (Government of India 2026). The latest climate data indicate that the country is on target to meet these goals, with a 36% reduction in CO2e/GDP already recorded through 2023. This has occurred alongside a relatively small reduction in the CO2e/Energy ratio (the black line of Figure 13(B)).

We use our model to project what the ratio of CO2e/Energy must be to achieve India's target of a 47% decline in CO2e/GDP through 2035, as well as a 60% reduction by 2050. We find that allowing for scale effects in energy demand, the ratio of CO2e/Energy can actually expand in the decades ahead while still achieving these targets. This is because as the economy expands, and average firm size grows, firms endogenously reduce their energy intensity. By contrast,

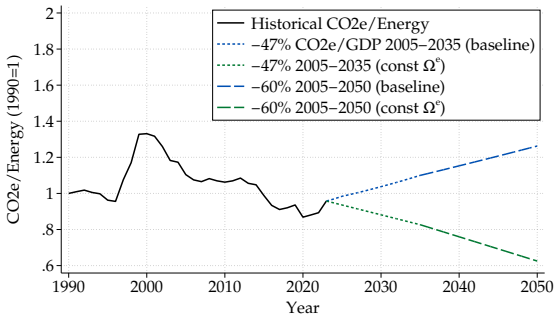
[42]The non-homothetic model predicts that energy expenditure share continues to decline, falling from 4.9% in 2023 to 2.6% in 2050. For reference, the energy cost share in U.S. manufacturing is around 2% (NBER-CES Manufacturing Industry Database).

FIGURE 13: Projections of Energy Demand and GHG Intensity Requirements

(A) Energy Demand




(B) CO2e Intensity Requirements




*Notes:* Panel A plots energy demand in the data (black line), and model predictions from 2023. The dashed blue line projects energy demand in the model. The dashed green line projects energy demand under a constant energy cost share of 4.9%. Panel B plots the ratio of greenhouse gas emissions to energy in India, where GHG are measured in CO2 equivalent (CO2e) from ClimateWatch. The black line is data from 1990-2023. The dashed blue line projects the decline in the CO2e/Energy ratio that is required to meet a 47% reduction in CO2e/GDP from 2005 levels by 2035. The long dashed blue line plots the required CO2e/Energy needed to meet a total reduction of 60% below 2005 levels by 2050. The dashed green line plots the same assuming a constant energy expenditure share.

the green line shows the path of CO2e/Energy required in the alternative scenario where we assume that the energy share remains constant. We find that the CO2e/Energy ratio would be required to fall to 25% below 2005 levels to achieve a 60% decline in the emissions intensity of GDP by 2050.

# 7. CONCLUSION

This paper uncovers a new mechanism linking economic growth to declining energy intensity: as economies expand, firms grow larger, and larger firms use less energy per unit of output. This relationship is driven by a decline in the cost share of energy relative to other inputs as firms grow, and can be rationalized by a non-homotheticity in the production function. In particular, this relationship is not driven by input price differences or compositional shifts. The scale elasticity of energy demand likely reflects technological forces: physical scaling laws that reduce energy losses at higher capacity, and fixed-cost efficiency investments. Embedding these forces into a model of heterogeneous firms, we find that scale-dependent energy demand can account for almost all of the observed fall in India's energy expenditure share since 1990. Looking forward, India's growth implies substantial additional reductions in energy intensity through 2050.

Our findings have broad implications for growing energy demand in low- and middle-income countries. They suggest that the environmental costs of develop-

ment may be lower than conventionally believed, as growth itself induces cleaner production. They also point toward the value of incorporating scale-dependent production technologies into Integrated Assessment Models and policy frameworks that evaluate the energy transition. Doing so would allow such models to capture an important and previously overlooked channel through which economic development can support, rather than hinder, long-run decarbonization.

# REFERENCES


ABEBERESE, A. B. (2017): "Electricity Cost and Firm Performance: Evidence from India," *Review of Economics and Statistics*, 99, 839–852.

ACEMOGLU, D., P. AGHION, L. BARRAGE, AND D. HÉMOUS (2023): "Climate Change, Directed Innovation, and Energy Transition: The Long-Run Consequences of the Shale Gas Revolution," Tech. rep., National Bureau of Economic Research.

ACEMOGLU, D., P. AGHION, L. BURSZTYN, AND D. HÉMOUS (2012): "The Environment and Directed Technical Change," *American Economic Review*, 102, 131–166.

ACEMOGLU, D., P. AGHION, AND D. HÉMOUS (2014): "The Environment and Directed Technical Change in a North–South Model," *Oxford Review of Economic Policy*, 30, 513–530.

ADÃO, R., M. KOLESÁR, AND E. MORALES (2019): "Shift-share designs: Theory and inference," *The Quarterly Journal of Economics*, 134, 1949–2010.

AGHION, P., T. BOPPART, M. PETERS, M. SCHWARTZMAN, AND F. ZILIBOTTI (2024): "A Theory of Endogenous Degrowth and Environmental Sustainability," .

ARKOLAKIS, C. AND C. WALSH (2023): "Clean Growth," Tech. rep., National Bureau of Economic Research.

BARROWS, G. AND H. OLLIVIER (2018): "Cleaner Firms or Cleaner Products? How Product Mix Shapes Emission Intensity from Manufacturing," *Journal of Environmental Economics and Management*, 88, 134–158.

——— (2021): "Foreign Demand, Developing Country Exports, and CO2 Emissions: Firm-Level Evidence from India," *Journal of Development Economics*, 149, 102587.

BATRAKOVA, S. AND R. B. DAVIES (2012): "Is There an Environmental Benefit to Being an Exporter? Evidence from Firm-Level Data," *Review of World Economics*, 148, 449–474.

BENTO, P. AND D. RESTUCCIA (2017): "Misallocation, Establishment Size, and Productivity," *American Economic Journal: Macroeconomics*, 9, 267–303.

——— (2021): "On Average Establishment Size across Sectors and Countries," *Journal of Monetary Economics*, 117, 220–242.

BOEHM, J., S. DHINGRA, AND J. MORROW (2022): "The Comparative Advantage of Firms," *Journal of Political Economy*, 130, 3025–3100.

BOEHM, J. AND E. OBERFIELD (2020): "Misallocation in the Market for Inputs: Enforcement and the Organization of Production," *The Quarterly Journal of Economics*, 135,

2007–2058.

——— (2023): "Growth and the Fragmentation of Production," Tech. rep.

BORUSYAK, K., P. HULL, AND X. JARAVEL (2022): "Quasi-Experimental Shift-Share Research Designs," *The Review of Economic Studies*, 89, 181–213.

BOSKER, M. AND E. BURINGH (2020): "Ice (berg) transport costs," *The Economic Journal*, 130, 1262–1287.

BRUCKNER, B., K. HUBACEK, Y. SHAN, H. ZHONG, AND K. FENG (2022): "Impacts of Poverty Alleviation on National and Global Carbon Emissions," *Nature Sustainability*, 5, 311–320.

CARON, J. AND T. FALLY (2022): "Per Capita Income, Consumption Patterns, and CO2 Emissions," *Journal of the Association of Environmental and Resource Economists*, 9, 235–271.

CASEY, G. (2024): "Energy Efficiency and Directed Technical Change: Implications for Climate Change Mitigation," *Review of Economic Studies*, 91, 192–228.

COLMER, J. M., D. LAGAKOS, AND M. SHU (2024): "Is the Electricity Sector a Weak Link in Development?" .

COMIN, D., D. LASHKARI, AND M. MESTIERI (2021): "Structural Change with Long-Run Income and Price Effects," *Econometrica*, 89, 311–374.

COPELAND, B. R., J. S. SHAPIRO, AND M. S. TAYLOR (2022): "Globalization and the Environment," in *Handbook of international economics*, Elsevier, vol. 5, 61–146.

ECKERT, F., S. GANAPATI, AND C. WALSH (2025): "Urban-Biased Growth: A Macroeconomic Analysis," Tech. rep., National Bureau of Economic Research.

EDMOND, C., V. MIDRIGAN, AND D. Y. XU (2023): "How costly are markups?" *Journal of Political Economy*, 131, 1619–1675.

FORSLID, R., T. OKUBO, AND K. H. ULLTVEIT-MOE (2018): "Why Are Firms That Export Cleaner? International Trade, Abatement and Environmental Emissions," *Journal of Environmental Economics and Management*, 91, 166–183.

GAO, Y. (2025): "Beyond Emissions: The Implications of Capital Misallocation on Optimal Carbon Tax," .

GARCIA-MACIA, D., C.-T. HSIEH, AND P. J. KLENOW (2019): "How Destructive Is Innovation?" *Econometrica*, 87, 1507–1541.

GERTLER, P. J., O. SHELEF, C. D. WOLFRAM, AND A. FUCHS (2016): "The Demand for Energy-Using Assets among the World's Rising Middle Classes," *American Economic Review*, 106, 1366–1401.

GHODBANAN, S., R. ALIZADEH, S. SHAFIEI, AND J. RAHBAR SHAHROUZI (2024): "Improving Energy Efficiency and Heat Recovery Performance in Paper Machine Drying Section," *Journal of Computational & Applied Research in Mechanical Engineering (JCARME)*, 14, 47–62.

GOVERNMENT OF INDIA (2026): "India's Nationally Determined Contribution (2031–2035)," Tech. rep., Government of India.

GÜLICH, J. F. (2020): *Centrifugal Pumps*, Cham, Switzerland: Springer Nature Switzerland, 4 ed.

HASSLER, J., P. KRUSELL, AND C. OLOVSSON (2021): "Directed Technical Change as a Response to Natural Resource Scarcity," *Journal of Political Economy*, 129, 3039–3072.

HSIEH, C.-T. AND P. J. KLENOW (2014): "The Life Cycle of Plants in India and Mexico," *The Quarterly Journal of Economics*, 129, 1035–1084.

JAMES, R. E., S. LEPTINSKY, M. TURNER, ET AL. (2022): "Quality Guidelines for Energy System Studies: Capital Cost Scaling Methodology: Revision 4a Report," Tech. Rep. DOE/NETL-2022/3340, National Energy Technology Laboratory (NETL), U.S. Department of Energy, technical report; Capital cost scaling guidance for energy system studies (Revision 4a).

JORDÀ, Ò., F. NECHIO, T. PHAN, AND F. SCHWARTZMAN (2026): "Financial Conditions and Capital Investment Choices," Federal Reserve Bank of San Francisco.

KABIR, P. AND E. TAN (2025): "Green Financing, Capital (Mis) Allocation, and Unintended Consequences," *Available at SSRN 5849842*.

KIM, S. (2025): "Optimal Carbon Taxes and Misallocation across Heterogeneous Firms," *Available at SSRN 4844030*.

KLENOW, P. J. AND H. LI (2025): "Entry Costs Rise with Growth," *Journal of Political Economy Macroeconomics*, 3, 43–74.

KLENOW, P. J., E. PASTÉN, AND C. RUANE (2024): "Carbon Taxes and Misallocation in Chile," Tech. rep., Mimeo, Stanford.

LASHKARI, D., A. BAUER, AND J. BOUSSARD (2024): "Information Technology and Returns to Scale," *American Economic Review*, 114, 1769–1815.

LEE, K., E. MIGUEL, AND C. WOLFRAM (2020): "Does Household Electrification Supercharge Economic Development?" *Journal of Economic Perspectives*, 34, 122–144.

LEVINSON, A. (2015): "A Direct Estimate of the Technique Effect: Changes in the Pollution Intensity of US Manufacturing, 1990–2008," *Journal of the Association of Environmental and Resource Economists*, 2, 43–56.

LOH, H. P., J. LYONS, AND I. WHITE, CHARLES W. (2002): "Process Equipment Cost Estimation: Final Report," Tech. Rep. DOE/NETL-2002/1169, U.S. Department of Energy, National Energy Technology Laboratory.

MARTIN, L. A. (2011): "Energy Efficiency Gains from Trade: Greenhouse Gas Emissions and India's Manufacturing Sector," *Mimeograph, Berkeley ARE*.

MERTENS, M. AND B. SCHOEFER (2024): "From Labor to Intermediates: Firm Growth, Input Substitution, and Monopsony," Tech. rep., National Bureau of Economic Research.

PETERS, M. AND C. WALSH (2026): "Population Growth and Firm Dynamics," *Journal of Political Economy Macroeconomics*.

POSCHKE, M. (2018): "The Firm Size Distribution across Countries and Skill-Biased Change in Entrepreneurial Technology," *American Economic Journal: Macroeconomics*, 10, 1–41.

SATO, R. (1974): "On the Class of Separable Non-Homothetic CES Functions," *The Economic Studies Quarterly*, 25, 42–55.

——— (1977): "Homothetic and Non-Homothetic CES Production Functions," *The American Economic Review*, 67, 559–569.

SCHLICHTING, H. AND K. GERSTEN (2017): *Boundary-Layer Theory*, Berlin, Germany: Springer, 9 ed.

SHANKER, A. AND D. I. STERN (2018): "Energy Intensity, Growth and Technical Change," *CAMA Working Paper*.

SHAPIRO, J. S. AND R. WALKER (2018): "Why Is Pollution from US Manufacturing Declining? The Roles of Environmental Regulation, Productivity, and Trade," *American Economic Review*, 108, 3814–3854.

SINGER, G. (2024): "Complementary Inputs and Industrial Development: Can Lower Electricity Prices Improve Energy Efficiency?" Tech. rep.

TRIBE, M. AND R. ALPINE (1986): "Scale Economies and the "0.6 Rule"," *Engineering Costs and Production Economics*, 10, 271–278.

TROTTNER, F. (2019): "Who Gains from Scale?" Tech. rep., Working Paper.

TYBOUT, J. R. (2000): "Manufacturing Firms in Developing Countries: How Well Do They Do, and Why?" *Journal of Economic Literature*, 38, 11–44.

US DEPARTMENT OF ENERGY (1997): "Determining Electric Motor Load and Efficiency," *Factsheet*.

U.S. DEPARTMENT OF ENERGY (2022): "2018 Manufacturing Static Energy Sankey Diagrams," Tech. rep., U.S. Department of Energy, page updated June 2022; accessed 2026-02-02.

WOLLBURG, P., S. HALLEGATTE, AND D. G. MAHLER (2023): "Ending Extreme Poverty Has a Negligible Impact on Global Greenhouse Gas Emissions," *Nature*, 623, 982–986.

# A. Additional Empirical Results

## A.1 Firm-Level Evidence on Firm Scale and Energy Demand

### A.1.1 Instrument Validity

To clarify the discussion of identification, we relate our estimating equation to a simplified version of the production function used in the quantitative analysis of Section 5, where we collapse the inner nest to include only energy. This version suffices to highlight the concerns for the identification of the scale elasticity.

Consider a non-homothetic CES production function, where output $y_i$ is implicitly defined through:

$$\left(\frac{e^{z_i} m_i}{y_i^{\gamma}}\right)^{\frac{\sigma-1}{\sigma}} + \left(\frac{e^{z_i+z_i^e} e_i}{y_i^{\gamma+\epsilon}}\right)^{\frac{\sigma-1}{\sigma}} = 1,$$

where $e_i$ is energy and $m_i$ is materials. $z_i$ and $z_i^e$ are factor-symmetric and energy-biased (log) productivity states, potentially heterogeneous across firms.

Solving cost minimization yields the equation for the relative energy cost share:

$$\log\left(\frac{\Omega_{it}^e}{1-\Omega_{it}^e}\right) = \underbrace{(1-\sigma)\epsilon \log y_{it}}_{\text{Scale}} + \underbrace{(1-\sigma)\log\left(\frac{w_{it}^e}{w_{it}^m}\right)}_{\text{Substitution}} + \underbrace{(\sigma-1)\, z_{it}^e}_{\substack{\text{Energy-specific} \\ \text{productivity shock}}} \tag{OA.1}$$

The first term captures the scale elasticity of relative energy demand: as long as $(1-\sigma)\epsilon \neq 0$, the energy expenditure share depends on scale. The goal of our empirical exercise is to identify $(1-\sigma)\epsilon$. The equation shows that the relative energy cost share depends on two additional forces: a substitution effect dependent on relative prices, and energy-biased productivity. Therefore, to estimate $(1-\sigma)\epsilon$, we require variation in firm scale that is orthogonal to: (i) the relative price of energy, (ii) energy-specific productivity shocks.

*Empirical specification and instrument.* We estimate:

$$\Delta^h \log\left(\frac{\Omega_{it}^e}{1-\Omega_{it}^e}\right) = \alpha_{sct} + \eta^h\, \Delta^h \log \text{output}_{it} + u_{it} \tag{OA.2}$$

where $\Delta^h \log \text{output}_{it}$ is defined as the log change in sales or quantity produced. $\Delta^h \log \text{output}_{it}$ is instrumented with:

$$\mathcal{D}_{it}^h = \sum_{j \in \mathcal{V}} \omega_{ijt}\, \Delta^h \log \text{sales}_{jt} \tag{OA.3}$$

The coefficient $\eta^h$ provides an estimate of the structural scale elasticity of energy demand $(1-\sigma)\epsilon$ as long as the following orthogonality condition is satisfied:[43]

$$\mathbb{E}\left[\mathcal{D}_{it}^{h} u_{it} \mid s(i), t\right] = 0. \tag{OA.4}$$

The error term $u_{it}$ captures the firm-specific determinants of the energy share, in particular the relative price of energy and energy-specific productivity shocks. In what follows, we refer to both as "energy supply shocks".

**Instrument Validity.** It is useful to distinguish two types of potential issues. A first concern is that we do not observe the true product-level demand shocks, but instead estimate them as $\Delta^h \log \text{sales}_{jt}$, which may induce a mechanical correlation between $\mathcal{D}_{it}^{h}$ and $u_{it}$. Second, even if we observed true product-level demand shocks, assumption (OA.4) requires that firms exposed to products with large demand shocks are not subject to systematically different energy price or productivity shocks. We discuss both concerns in turn.

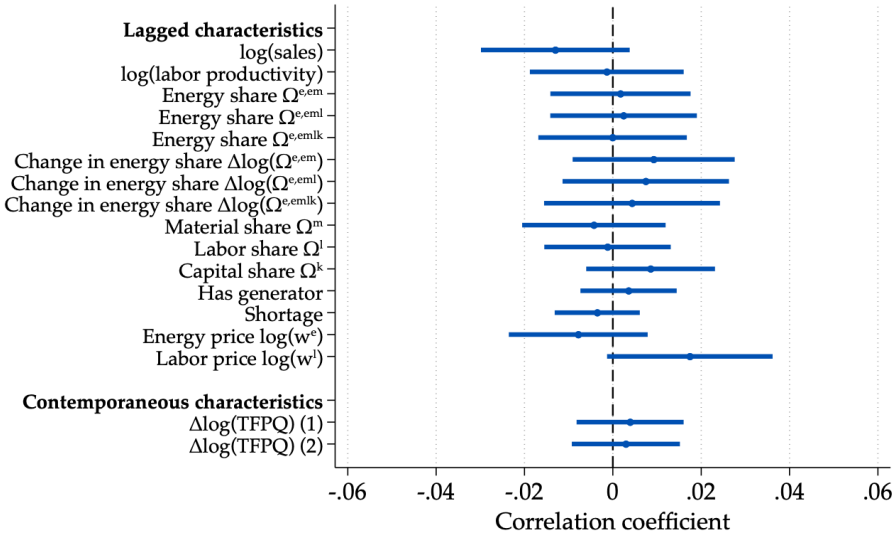
*Measurement of the shifters.* The shifters $\Delta^h \log \text{sales}_{jt}$ are not exogenous shocks but product-level sales growth, an equilibrium outcome. In particular, $\Delta^h \log \text{sales}_{jt}$ aggregates firm-specific shocks of firms producing product $j$ that may directly enter the residual $u_{it}$. As shown in Borusyak et al. (2022), this mechanical bias is most severe when only few firms contribute to each shifter estimate. We alleviate this concern in two ways. First, we drop products $j$ with fewer than $N$ firms in the product$\times$year cell for $N \in \{5, 10, 20\}$. Second, we construct $\Delta^h \log \text{sales}_{jt}$ excluding any firms with market share above $x$% in the production of product $j$ for $x \in \{5, 10, 20\}$. All variants yield highly similar results (Figure SM.11), suggesting that our strategy is relatively immune to this concern.

*Shift-share identification.* Assumption (OA.4) requires that firm exposure to product-level growth is orthogonal to the unobserved determinants of energy shares $u_{it}$, in particular any firm-specific energy supply shock. That is, firms must not sort across products such that firms with high (low) energy supply shocks systematically have high shares in high (low) growth products (Borusyak et al. 2022). In our context, the main threat to identification is that product-level sales growth is correlated with product-level energy-specific supply shocks. For instance, it may be that product $j$ is growing fast because of an energy-specific productivity improvement in the production process for this product. Then, firms specialized in $j$ would benefit from this energy-specific productivity shock, and $\mathcal{D}_{it}^{h}$ would be

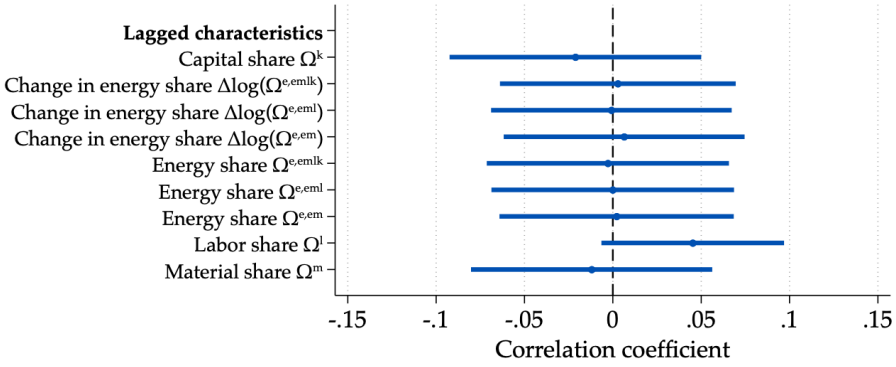
[43] When the explanatory variable is the change in quantity produced, when it is the change in sales, the estimated parameter is the scale elasticity modified by the demand elasticity.

FIGURE OA.1: Instrument Balance Tests

(A) Firm-level




(B) Product-level




*Notes:* Panel A shows the coefficients of firm-level regressions of the firm-level demand shifter (defined in (4)) on firm characteristics. The first block reports correlations with lagged characteristics. The second block reports correlations with the contemporaneous change in productivity (TFPQ), defined in Supplementary Material SM.1.4. Regressions include industry×year×cohort fixed effects and are weighted by firm total cost. Standard errors are clustered at the firm and main product×year level. Panel B shows the coefficients of product-level regressions of the product-level shifter $\Delta^{h} \log \text{sales}_{jt}$ on lagged product-level characteristics. Regressions include year fixed effects and are weighted by product-level total sales. Standard errors are clustered at the product level. The dots are point estimates and bars are 95% confidence intervals. All variables are standardized so that the coefficients can be interpreted as correlation coefficients. All regressions are for horizon $h = 5$.

correlated with firm-level energy-specific productivity shocks.

The most direct test supporting assumption (OA.4) is firm-level balance on observables. Figure OA.1(A) shows that firms with high and low $\mathcal{D}_{it}^{h}$ are similar on

variables that are likely correlates of energy supply or productivity, e.g., factor shares, factor prices, having an electricity generator, or facing electricity shortages. Figure OA.1(A) also reports correlations between the instrument and the lagged changes in the energy share, showing the absence of pre-trends. Balanced firm-level characteristics make it less likely that high $\mathcal{D}_{it}^{h}$ firms are systematically subject to different energy-specific supply shocks.

Following the argument in Borusyak et al. (2022), the orthogonality condition (OA.4) can be expressed at the shifter level: the necessary requirement for identification is that product-level shifters are uncorrelated with the average firm-level determinants of the energy share for the firms most exposed to each product. A sufficient condition is that the product-level shifters are "as good as random". Figure OA.1(B) shows that the change in product-level sales is not correlated with lagged product-level factor shares, or lagged changes in these shares.

*Demand vs. productivity shifters.* We further show that our instrument is a demand shift, not a productivity shock. It is *not* a problem for identification if scale is shifted by a factor-neutral productivity shock—$z_i$ does not appear in the residual of equation (OA.1). However, it is plausible that if our instrument were correlated to energy-specific productivity, then it would be correlated to overall productivity too.[44] It is therefore useful to validate that our instrument affects scale via a demand shift. Using our estimates of the change in physical productivity (TFPQ, see details in Section SM.1.4), the orthogonality tests in the bottom block of Figure OA.1(A) show that it is uncorrelated with our demand instrument.

*Consistency.* Exposure to common product-level shocks induces dependencies across firms with similar exposure shares, so that the setting is not *iid*. Borusyak et al. (2022) show that the conditions for consistency are that (i) there is a sufficiently large number of shocks with sufficient shock-level variation, and (ii) that shock exposure is not too concentrated. Panel A of Table OA.1 documents a large dispersion in $\Delta^h \log \text{sales}_{jt}$, which persists when residualizing on fixed effects. Define product-level weights as $s_{jt} = \sum_i \text{sales}_{it}\, \omega_{ijt}$. Panel B shows that the largest weight is small (0.44%) and the inverse Herfindahl index is large (1,733). We report the same statistics when exposure weights are aggregated at the product-level, and there is sufficient product-level dispersion even when shocks are allowed to be serially correlated.[45]

[44] Equation (SM.4) shows that measured changes in firm productivity conflate factor-neutral productivity changes with a weighted average of factor-specific productivity changes.

[45] As a benchmark, Borusyak et al. (2022) show that their methodology is relevant in the canonical "China shock" setting where the inverse Herfindahl is 58.4 and the largest share is 6.5%.

TABLE OA.1: Shock-Level Summary Statistics

**Panel A:** Summary statistics on product-level shocks

| | count | mean | sd | p25 | p75 |
|---|---|---|---|---|---|
| Product-level $\Delta^5$ log(sales) | 10,314 | -0.024 | 0.672 | -0.282 | 0.340 |
| Residualized on time FE | 10,314 | 0.000 | 0.653 | -0.265 | 0.370 |
| Residualized on industry × time FE | 10,247 | 0.000 | 0.563 | -0.184 | 0.267 |
| Residualized on product FE | 10,251 | -0.000 | 0.560 | -0.236 | 0.263 |

**Panel B:** Summary statistics on exposure shares

| | Across products and dates | Across products |
|---|---|---|
| Inverse HHI | 1732.721 | 131.864 |
| Largest weight | .44% | 3.07% |

*Notes:* This table presents descriptive statistics relevant for the shift-share design. Panel A presents summary statistics of the product-level shocks. Panel B presents summary statistics of product-level weights $s_{jt} = \sum_i \text{sales}_{it}\ \omega_{ijt}$, normalized to sum to 1.

TABLE OA.2: First Stage

| | $\Delta^h$ log(sales) | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | $h=1$ | | $h=3$ | | $h=5$ | | $h=7$ | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| $\Delta^h \mathcal{D}$ | 0.286*** | 0.275*** | 0.290*** | 0.278*** | 0.273*** | 0.256*** | 0.262*** | 0.240*** |
| | (0.013) | (0.012) | (0.014) | (0.014) | (0.016) | (0.015) | (0.018) | (0.016) |
| Year × Ind. FE | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Year × Ind. × Cohort FE | | ✓ | | ✓ | | ✓ | | ✓ |
| F-stat | 504.8 | 500.6 | 414.5 | 410.5 | 293.6 | 293.7 | 214.8 | 212.4 |
| R-squared | 0.063 | 0.18 | 0.087 | 0.22 | 0.10 | 0.25 | 0.11 | 0.27 |
| Observations | 371,703 | 367,485 | 263,478 | 259,789 | 188,128 | 184,760 | 135,729 | 132,736 |

*Notes:* This table shows the first stage of the IV regression in (3). Regressions are weighted by firm total cost (adjusted for the sampling weight). Standard errors are clustered at the firm and main product×year level. ***, **, and * denote significance at the 1%, 5% and 10% levels, respectively.

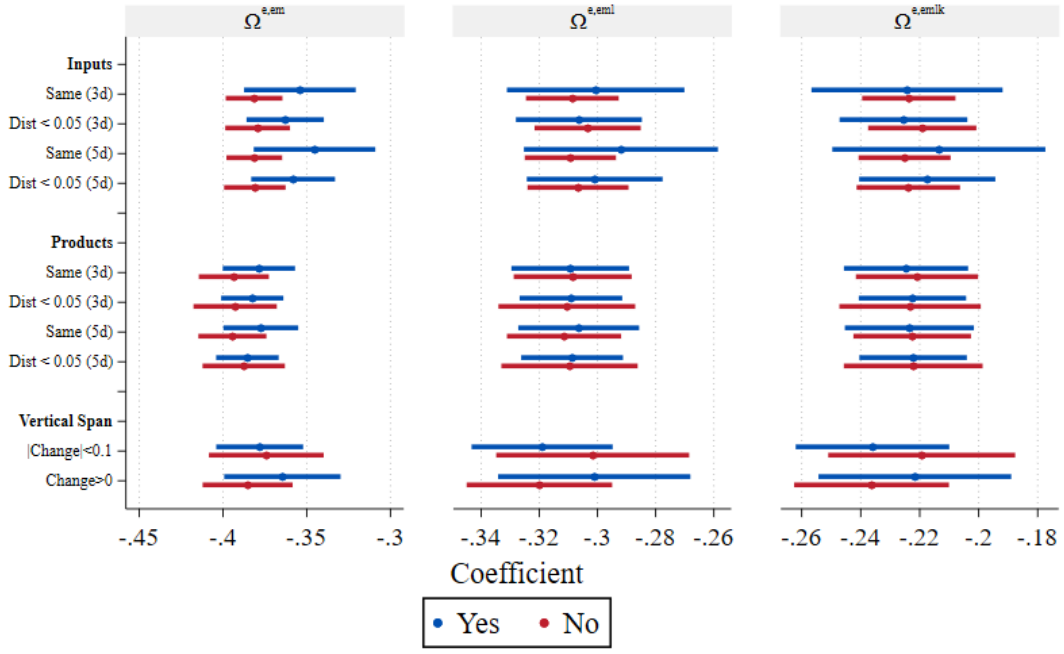
**First Stage.** Table OA.2 reports the first stage results.

### A.1.2 Additional Firm-Level Results

**Keeping Products, Inputs, and Vertical Scope Fixed.** Figure OA.2 investigates whether the decline in energy intensity with scale is concentrated in firms that at the same time experience changes in their product portfolio, input portfolio, or vertical span. The measure of the firm-level vertical span of production comes from Boehm and Oberfield (2020).[46]

[46] Denote $\delta_{\omega\omega'}$ the vertical distance between input $\omega$ and $\omega'$, defined as the average number of production steps between an output $\omega$ and an input $\omega'$, weighted by the overall expenditure of each product chain in industry $\omega$'s mix of materials, and excluding any circular product chains

FIGURE OA.2: Results by Inputs or Products Adjustments




*Notes:* This figure presents OLS results for specification (3), by subsamples. The blue (red) dots show results for the subsample indicated in each line (other firms). The blocks "Inputs" and "Products" consider the subset of firms that use similar material inputs or produce similar products at $t$ and $t+5$. "Same (3d)" and "Same (5d)" are the cases where firms use/produce the exact same set of product codes at $t$ and $t+5$, where product codes are defined at the 3-digit and 5-digit level, respectively. "Dist $<$ 0.05 (3d)" and "Dist $<$ 0.05 (5d)" are the cases where the cosine distance between product code shares at $t$ and $t+5$ is lower than 0.05, defined at the 3-digit and 5-digit level, respectively. The block "Vertical Span" considers subsamples defined by the $t$ to $t+5$ change in their vertical span, following the definition in Boehm and Oberfield (2020). $|\text{Change}| < 0.1$ is where the absolute change in vertical span is lower than 0.1. Change $> 0$ is the case where firms increased their vertical span. The dependent variable is the relative energy expenditure share, $\Delta^5 \log(\Omega^e/(1-\Omega^e))$, where the definition of $\Omega^e$ is indicated in the column heading. Regressions are weighted by firm total cost (adjusted for the sampling weight). Standard errors are clustered at the firm and main product$\times$year level. The bars are 95% confidence intervals.

**Quantity Regressions.** Table OA.3 reports IV estimates of equation (3), where the main regressor is $\Delta^5 \log(\text{quantity})$.

## A.2 Additional Model Results

**Model Fit.** In Figure OA.3, we present the key firm-level elasticities with respect to firm output $y$ in our model and compare them to those estimated in the data.

(see Boehm and Oberfield (2020) Appendix B for details). Then, the vertical span of firm $i$ having as main product $\omega$ is the expenditure-weighted average of the distances $\delta_{\omega\omega'}$ over its input bundle.With this measure, we replicate the key fact in Boehm and Oberfield (2023) that plant growth correlates negatively with vertical span.

TABLE OA.3: Quantity Regressions

| | $\Delta^5$ log(outcome) | | | | | | |
|---|---|---|---|---|---|---|---|
| | Targeted | | Untargeted | | | | |
| | $\frac{\Omega^e}{\Omega^m+\Omega^l}$ | $\frac{\Omega^e}{\Omega^l}$ | $\frac{\Omega^m}{1-\Omega^m}$ | $\frac{\Omega^l}{1-\Omega^l}$ | $\frac{\Omega^e}{1-\Omega^e}$ | $\frac{w^e e}{py}$ | $e/y$ |
| $\Delta^5$ log(quantity) | -0.511*** | -0.044 | 0.639*** | -0.419*** | -0.430*** | -0.660*** | -0.425*** |
| | (0.089) | (0.047) | (0.099) | (0.072) | (0.080) | (0.105) | (0.088) |
| Year × Ind. × Cohort FE | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| F-Stat | 41.7 | 41.6 | 41.1 | 41.1 | 41.6 | 41.7 | 41.5 |
| Observations | 138,011.0 | 137,913.0 | 138,109.0 | 138,032.0 | 137,776.0 | 138,021.0 | 136,889.0 |

*Notes*: This table reports IV estimates of equation (3), where the main regressor is $\Delta^5 \log(\text{quantity})$, instrumented with the firm-level demand shock defined in equation (4). The dependent variable is indicated in column headings, with targeted coefficients in columns (1)–(2) and untargeted ones in columns (3)–(7). Regressions are weighted by firm total cost (adjusted for the sampling weight). Standard errors are clustered at the firm and main product×year level. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

## A.3 Macro (Time-Series) Elasticities

We estimate the macro elasticity of the energy share using the aggregate time series computed from our model, described in Section 5.2. To do this we regress log changes in input shares on log changes in output, running the following regression in the time series from 1990-2023:

$$\Delta^5 \log(\Omega_t) = \alpha + \eta^{macro} \Delta^5 \log(\text{output}_t) + \varepsilon_t \tag{OA.5}$$

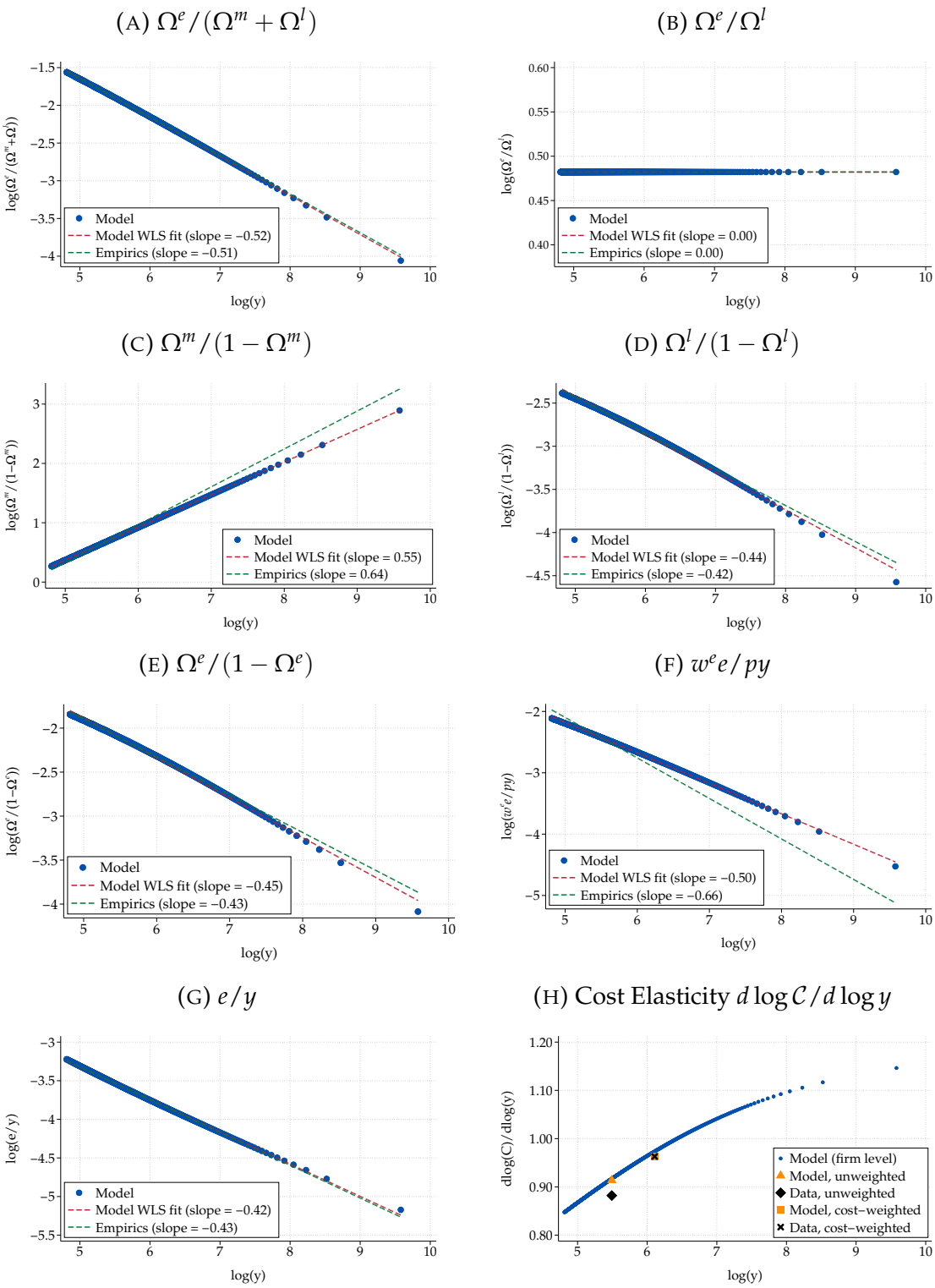
where $\Omega_t$ are the outcomes displayed in Table OA.4. We use the version of the model where energy, materials and capital prices are held constant.

TABLE OA.4: Model Fit: Macro (Time-Series) Elasticities

| | $\Delta^5$ log(outcome) | | | | | | |
|---|---|---|---|---|---|---|---|
| | $\frac{\Omega^e}{\Omega^m+\Omega^l}$ | $\frac{\Omega^e}{\Omega^l}$ | $\frac{\Omega^m}{1-\Omega^m}$ | $\frac{\Omega^l}{1-\Omega^l}$ | $\frac{\Omega^e}{1-\Omega^e}$ | $\frac{w^e e}{py}$ | $e/y$ |
| $\Delta^5$ log(quantity) | -0.431 | -0.165 | 0.334 | -0.210 | -0.394 | -0.415 | -0.415 |
| | (0.004) | (0.001) | (0.003) | (0.006) | (0.004) | (0.004) | (0.004) |
| Observations | 29 | 29 | 29 | 29 | 29 | 29 | 29 |

*Notes:* This table displays the model-implied macro elasticities estimated from time-series regressions of changes in log shares on changes in log output in the model, using equation (OA.5). Newey West standard errors are in parentheses.

FIGURE OA.3: Firm-level Elasticities: Model vs. Data




*Notes:* Panels A-G plot the empirical slope estimates (green dashed line) from the regressions in Table OA.3 against the model predictions. Each blue dot is a grid point in the state vector of preference shifters $a_i$ in the quantitative model, and the model slope (red dashed line) is estimated by weighted least squares, weighting each node by its total cost share. Panel H plots firm-level cost elasticities in the model (blue dots); the yellow triangle and square mark their unweighted and cost-weighted model averages, respectively, and the black diamond and cross are the corresponding averages in the data (Table SM.18).

# B. Proofs and Derivations

## B.1 Derivation of Equation (8) for Optimal Technology Choice

The problem of the firm is to choose technology to minimize the cost of output $y$, subject to producing at capacity. In other words, they solve

$$\min_{\phi} w^e e + \frac{1}{\gamma A_Y} \phi^\gamma \bar{y}^\delta (w^l)^\alpha (w^k)^{1-\alpha}$$

subject to $y = A_Y \phi e / y^\epsilon$ and $y = \bar{y}$ for a given $\bar{y}$. This has the first order condition

$$(w^l)^\alpha (w^k)^{1-\alpha} \phi^{\gamma-1} y^\delta = \frac{w^e y^{1+\epsilon}}{\phi^2},$$

which we can solve for optimal $\phi$ as

$$\phi = \left( \frac{w^e y^{1+\epsilon-\delta}}{(w^l)^\alpha (w^k)^{1-\alpha}} \right)^{\frac{1}{1+\gamma}}.$$

Energy demand for a given level of output $y$ is then

$$e = \frac{y^{1+\epsilon}}{\phi A_Y} = \frac{1}{A_Y} y^{\frac{(1+\epsilon)\gamma+\delta}{1+\gamma}} \left( \frac{(w^l)^\alpha (w^k)^{1-\alpha}}{w^e} \right)^{\frac{1}{1+\gamma}}$$

Then energy intensity, measured in joules per dollar of output, is

$$\Theta(a) = \frac{e}{p(y)y} = \frac{1}{a A_Y Y^{\frac{1}{\lambda}}} y^{\frac{(1+\epsilon)\gamma+\delta}{1+\gamma} - \zeta} \left( \frac{(w^l)^\alpha (w^k)^{1-\alpha}}{w^e} \right)^{\frac{1}{1+\gamma}}$$

## B.2 Proof of Proposition 1

We note that, by Shephard's Lemma, the energy demand of the firm is given from the cost function in equation (13) by

$$e = \frac{\partial \mathcal{C}(y)}{\partial w^e} = \frac{y^{1+\tilde{\epsilon}}}{A_Y} \left( \Omega^{e,elk} \right)^{-\frac{\sigma_e}{1-\sigma_e}},$$

where

$$\Omega^{e,elk} = \frac{(w^e)^{1-\sigma_e}}{(w^e)^{1-\sigma_e} + \left( (w^l)^\alpha (w^k)^{1-\alpha} \right)^{1-\sigma_e}}$$

is the share of energy in the cost of the inner energy–labor–capital nest: since both inputs of the nest carry the common scale exponent $1+\tilde{\epsilon}$, the factor $y^{(1+\tilde{\epsilon})(1-\sigma_e)}$ cancels and $\Omega^{e,elk}$ is independent of firm output, and hence common across firms. Using the expression for revenue $p(a)y(a) = ay^{\zeta}Y^{\frac{1}{\lambda}}$, we have

$$\Theta(a) \;=\; \frac{1}{aA_Y Y^{\frac{1}{\lambda}}}\, y(a)^{1+\tilde{\epsilon}-\zeta}\left(\Omega^{e,elk}\right)^{-\frac{\sigma_e}{1-\sigma_e}},$$

Differentiating the expression for $\Omega^{e,elk}$ above,

$$d\log\Omega^{e,elk} \;=\; -(1-\sigma_e)\left(1-\Omega^{e,elk}\right)\, d\log\left(\frac{(w^l)^{\alpha}(w^k)^{1-\alpha}}{w^e}\right),$$

and so, given a general equilibrium movement in productivity $A_Y$, and using the fact that $d\log(w^k) = d\log(w^e) = 0$ by the linearity of their production in the final good, we have

$$d\log\Theta(a) \;=\; (\tilde{\epsilon}+1-\zeta)\, d\log y(a) \;-\; d\log\left(A_Y Y^{\frac{1}{\lambda}}\right) \;+\; \alpha\,\sigma_e\left(1-\Omega^{e,elk}\right) d\log(w^l),$$

where $\tilde{\epsilon}$ is the scale elasticity, common across firms. Then use equation (15) to write

$$\begin{aligned}
d\log\bar{\Theta} &= \mathbb{E}_R[(\tilde{\epsilon}+1-\zeta)\, d\log y(a)] - d\log\left(A_Y Y^{\frac{1}{\lambda}}\right) + \alpha\,\mathbb{E}_R\left[\sigma_e\left(1-\Omega^{e,elk}\right)\right] d\log(w^l) \\
&\quad + Cov_R\left[\frac{\Theta(a)}{\bar{\Theta}}, d\log\frac{p(a)y(a)}{Y/N} + d\log\Theta(a)\right], \\
&= \mathbb{E}_R[(\tilde{\epsilon}+1-\zeta)\, d\log y(a)] - d\log\left(A_Y Y^{\frac{1}{\lambda}}\right) + \alpha\,\mathbb{E}_R\left[\sigma_e\left(1-\Omega^{e,elk}\right)\right] d\log(w^l) \\
&\quad + Cov_R\left[\frac{\Theta(a)}{\bar{\Theta}}, (\tilde{\epsilon}+1)\, d\log y(a)\right], \\
&= \mathbb{E}_R[(\tilde{\epsilon}+1-\zeta)\, d\log y(a)] - d\log\left(A_Y Y^{\frac{1}{\lambda}}\right) + \alpha\,\sigma_e\left(1-\Omega^{e,elk}\right) d\log(w^l) \\
&\quad + Cov_R\left[\frac{\Theta(a)}{\bar{\Theta}}, (\tilde{\epsilon}+1)\, d\log y(a)\right] + \left(\mathbb{E}_R[d\log p(a)y(a)] - d\log(Y/N)\right),
\end{aligned}$$

where the second equality substitutes $d\log p(a)y(a) \;=\; \zeta\, d\log y(a) + \frac{1}{\lambda}\, d\log Y$ and drops from the covariance the terms that are common across firms (including the wage term, and the third equality adds a term that is identically zero, since the revenue-weighted change in revenue weights must be zero. Rearranging, we get

$$\begin{aligned}
d\log(\bar{\Theta}) &= \tilde{\epsilon}\mathbb{E}_R[\, d\log y(a)] - d\log(A_Y) + \alpha\,\sigma_e\left(1-\Omega^{e,elk}\right) d\log(w^l) \\
&\quad + \left(\mathbb{E}_R[d\log y(a)] - d\log(Y/N)\right) \quad + Cov_R\left[\frac{\Theta(a)}{\bar{\Theta}}, (\tilde{\epsilon}+1)\, d\log y(a)\right].
\end{aligned}$$

## B.3 Proof of Proposition 2

**Step 1: A useful decomposition of $d\log\Theta$** The cost elasticity can be shown to be

$$\mathcal{E} \;=\; \frac{d\log\mathcal{C}}{d\log y} \;=\; 1+\tilde{\epsilon}\,\Omega^{NM}$$

By Shephard's Lemma, energy and production labor payments are

$$w^e e \;=\; \Omega^{e,elk}\,\Omega^{NM}\,\mathcal{C}(y), \qquad w^l l \;=\; \alpha\left(1-\Omega^{e,elk}\right)\Omega^{NM}\,\mathcal{C}(y).$$

Equate marginal revenue to marginal cost, $\zeta p = \mathcal{C}'(y)$, so that revenue satisfies $py = (\mathcal{E}/\zeta)\,\mathcal{C}(y)$. Hence

$$\bar{\Theta} \;=\; \frac{Ne}{Y} \;=\; \frac{e}{py} \;=\; \zeta\,\frac{\Omega^{e,elk}\,\Omega^{NM}}{\mathcal{E}}, \qquad d\log\bar{\Theta} \;=\; d\log\Omega^{e,elk}+d\log\Omega^{NM}-d\log\mathcal{E},$$

which is the decomposition in the main text. Differentiating the two shares and the cost elasticity,

$$\begin{aligned} d\log\Omega^{e,elk} &= -(1-\sigma_e)\left(1-\Omega^{e,elk}\right)\alpha\,d\log w^l,\\ d\log\Omega^{NM} &= \left(1-\Omega^{NM}\right)\left[\tilde{\epsilon}\,d\log y-\frac{d\log\Omega^{e,elk}}{1-\sigma_e}\right],\\ d\mathcal{E} &= \tilde{\epsilon}\,d\Omega^{NM}. \end{aligned}$$

Combining these with the identity $\mathcal{E}=1+\tilde{\epsilon}\,\Omega^{NM}$ yields, after rearranging,

(OA.6)

$$\mathcal{E}\,d\log\bar{\Theta} \;=\; \tilde{\epsilon}\left(1-\Omega^{NM}\right)d\log y \;+\; \left[(1+\tilde{\epsilon})\,\Omega^{NM}-\sigma_e\left(1+\tilde{\epsilon}\,\Omega^{NM}\right)\right]\frac{d\log\Omega^{e,elk}}{1-\sigma_e}.$$

The first term is strictly negative, since $\tilde{\epsilon}<0$ and $d\log y>0$. Step 2 shows that free entry requires $d\log\Omega^{e,elk}<0$.

**Step 2: Free entry requires the energy share of the nest to fall.** Combining free entry, equation (12), with $py = (\mathcal{E}/\zeta)\,\mathcal{C}(y)$ gives $w^l v = py-\mathcal{C}(y) = ((\mathcal{E}-\zeta)/\zeta)\,\mathcal{C}(y)$. Since $w^l, v, \mathcal{C} > 0$, any equilibrium has $\mathcal{E}>\zeta$. Dividing the expression for $w^l l$ above by $w^l v$,

$$\frac{l}{v} \;=\; \frac{\alpha\,\zeta\left(1-\Omega^{e,elk}\right)\Omega^{NM}}{\mathcal{E}-\zeta}.$$

With symmetric firms the aggregate price index is $N^{\frac{1}{1-\lambda}} p$, so the numeraire implies $p = N^{\frac{1}{\lambda-1}}$. Finally, writing $\zeta p = \mathcal{C}'(y)$ per unit of output,

$$\zeta A_Y \, p \;=\; (1+\tilde{\epsilon}) \, \frac{\mathcal{P}(y)}{y} + w^m. \tag{OA.7}$$

where

$$\mathcal{P}(y) \;\equiv\; y^{1+\tilde{\epsilon}} \Big( (w^e)^{1-\sigma_e} + \big( (w^l)^{\alpha} (w^k)^{1-\alpha} \big)^{1-\sigma_e} \Big)^{\frac{1}{1-\sigma_e}}$$

Suppose, toward a contradiction, that $d \log \Omega^{e,elk} \geq 0$. Since we assume $\sigma_e < 1$, this is equivalent to $d \log w^l \leq 0$. Then:

1. $\Omega^{NM}$ falls strictly: in the expression for $d \log \Omega^{NM}$ above, the first term in the bracket is strictly negative and the second is non-positive.
2. $\mathcal{E}$ rises strictly: $d\mathcal{E} = \tilde{\epsilon} \, d\Omega^{NM} > 0$, since $\tilde{\epsilon} < 0$ and $d\Omega^{NM} < 0$.
3. $l/\nu$ falls strictly: $\left(1 - \Omega^{e,elk}\right)$ weakly falls, $\Omega^{NM}$ strictly falls, and $\mathcal{E} - \zeta$ strictly rises.
4. $N$ and $p$ rise strictly: since $\nu$ is a parameter, $l$ falls strictly, so $N = L/(l+\nu)$ rises strictly, and $p = N^{\frac{1}{\lambda-1}}$ rises strictly.
5. Equation (OA.7) is violated: its left-hand side strictly rises, since $d \log A_Y > 0$ and $d \log p > 0$. On the right-hand side, $w^m$ is constant, while, using $1 + \tilde{\epsilon} > 0$ and $\mathcal{P}(y)/y = y^{\tilde{\epsilon}} \, \mathcal{P}(1)$,

$$d \log \left( (1+\tilde{\epsilon}) \, \frac{\mathcal{P}}{y} \right) = \tilde{\epsilon} \, d \log y + \left(1 - \Omega^{e,elk}\right) \alpha \, d \log w^l \;<\; 0,$$

since $\tilde{\epsilon} < 0$, $d \log y > 0$, and $d \log w^l \leq 0$ under the supposition. So the right-hand side strictly falls.

This contradiction establishes $d \log \Omega^{e,elk} < 0$: free entry requires the wage to strictly rise, $d \log w^l > 0$.

**Step 3: The threshold $\sigma_e^*$.** By Step 2, $d \log \Omega^{e,elk} < 0$, so the first term on the right-hand side of (OA.6) is strictly negative, and the factor $d \log \Omega^{e,elk} / (1 - \sigma_e)$ multiplying the bracketed coefficient is strictly negative as well. Whenever $\sigma_e \leq (1+\tilde{\epsilon})\Omega^{NM} / \left(1 + \tilde{\epsilon}\Omega^{NM}\right)$, the bracketed coefficient in (OA.6) is non-negative and $d \log \bar{\Theta} < 0$ follows immediately. More generally, dividing (OA.6) by the positive quantity $(1 + \tilde{\epsilon}\Omega^{NM}) \cdot (-\, d \log \Omega^{e,elk} / (1 - \sigma_e))$ and rearranging, $d \log \bar{\Theta} < 0$ if

and only if

$$\sigma_e \;<\; \sigma_e^* \;\equiv\; \frac{(1+\tilde{\epsilon})\,\Omega^{NM}}{1+\tilde{\epsilon}\,\Omega^{NM}} \;+\; \frac{(-\tilde{\epsilon})\,\left(1-\Omega^{NM}\right)\,d\log y}{(1+\tilde{\epsilon}\,\Omega^{NM})\,\left(-\,d\log\Omega^{e,elk}/(1-\sigma_e)\right)},$$

where all objects are evaluated at the equilibrium under consideration, and $d\log\Omega^{e,elk}/(1-\sigma_e) = -\left(1-\Omega^{e,elk}\right)\alpha\, d\log w^l$. The second term in $\sigma_e^*$ is strictly positive, so

$$\sigma_e^* \;>\; \frac{(1+\tilde{\epsilon})\,\Omega^{NM}}{1+\tilde{\epsilon}\,\Omega^{NM}} \;>\; 0.$$

Hence, at any equilibrium satisfying the hypotheses of the proposition, $\sigma_e < \sigma_e^*$ implies $d\log\bar{\Theta} < 0$; the equilibrium statistic $\sigma_e^*$ is the threshold whose existence the proposition asserts. Finally, note that $\sigma_e^* < 1$ if and only if $d\log\Omega^{NM} > 0$: the threshold binds below one precisely when the general equilibrium increase in the non-material share emphasized in the main text is operative; otherwise $\sigma_e^* \geq 1 > \sigma_e$, and energy intensity necessarily falls at such an equilibrium.

# C. PHYSICAL MECHANISMS

In this Appendix, we derive the scale elasticities for typical manufacturing technologies corresponding to the main energy end-uses. This section uses the notation $\dot{x}$ for the derivative of variable $x$ with respect to time.

## C.1 Process Heating (Fired Systems)

This model estimates the scale elasticity of energy intensity for a crucible furnace used for melting and holding molten metal. For the numerical application, we consider the case of aluminum melting. This is a simple system for which we can derive a straightforward characterization. Below, we also provide estimates of the scale elasticity for more complex systems.

**Step 1: Combustion.** We estimate the energy input $Q_{\text{total}}$ required to deliver $Q_{\text{furnace}}$ to the furnace. The main combustion loss is flue-gas loss, which is proportional to fuel input, so that

$$Q_{\text{total}} = \frac{Q_{\text{furnace}}}{\eta_{\text{comb}}},$$

where $\eta_{\text{comb}}$ is the combustion efficiency. To obtain this, start from $Q_{\text{total}} = Q_{\text{furnace}} + L_{\text{flue}}$. The flue-gas loss per unit of time $t$ depends on the gas–environment

temperature difference and the flue-gas mass flow rate $\dot{m}_{\text{flue}}$, so that

$$L_{\text{flue}} = \dot{m}_{\text{flue}}\, c_p \left(T_{\text{flue}} - T_{\text{ref}}\right) t,$$

with $c_p$ the mean specific heat of the flue gas. The flue-gas flow is proportional to the fuel flow, $\dot{m}_{\text{flue}} \approx (1+\lambda)\, A_{\text{st}}\, \dot{m}_{\text{fuel}}$ ($\lambda$ the excess-air ratio, $A_{\text{st}}$ the stoichiometric air-to-fuel ratio). Since the fuel energy input is $Q_{\text{total}} = \dot{m}_{\text{fuel}}\, \text{LHV}\, t$ where LHV is the lower heating value of the fuel, the flue-gas loss is a scale-independent fraction of input:

$$\frac{L_{\text{flue}}}{Q_{\text{total}}} = \frac{(1+\lambda)\, A_{\text{st}}\, c_p \left(T_{\text{flue}} - T_{\text{ref}}\right)}{\text{LHV}} \equiv 1 - \eta_{\text{comb}}.$$

This analysis only considers heat losses in dry flue gas. There are also (smaller) evaporation losses from hydrogen and moisture in the fuel. These losses are also proportional to $\dot{m}_{\text{fuel}}$. Hence, they lower $\eta_{\text{comb}}$ but preserve proportionality.

**Step 2: Energy required by the furnace.** We now estimate the energy required by the furnace to melt a mass $M$ of metal.

*Energy to heat and melt the metal:* The useful energy raises the metal mass $M$ from the ambient temperature $T_{\text{init}}$ to the operating temperature $\overline{T}$ (sensible heat) and melts it (latent heat of fusion):

$$Q_{\text{useful}} = c\,(\overline{T} - T_{\text{init}})\, M + L_f\, M \equiv a_{\text{useful}}\, M,$$

with $c$ the specific heat capacity and $L_f$ the latent heat of fusion of the metal. Useful energy scales proportionally with $M$.

*Energy required to maintain molten state (holding losses):* We now quantify the surface losses. They are of three types:

Radiative loss (electromagnetic heat loss from radiation): $\dot{Q}_{\text{rad}} = \varepsilon\sigma A_s \left(\overline{T}^4 - T_{\text{init}}^4\right)$ where $\varepsilon$ is the emissivity for molten metal, $\sigma$ is the Stefan-Boltzmann constant, and $A_s$ is the surface area of molten metal in $\text{m}^2$.

Convective loss (heat transfer through metal-air contact): $\dot{Q}_{\text{conv}} = hA_s(\overline{T} - T_{\text{init}})$ where $h$ is the natural convective heat transfer coefficient.

Conductive loss through furnace walls (heat transfer through contact between materials): $\dot{Q}_{\text{cond}} = \frac{k \cdot A_c}{d} \cdot (\overline{T} - T_{\text{init}})$ where $k$ is the effective thermal conductivity of the furnace insulation, and $d$ is the wall thickness of the insulation layer.

Hence, the energy required to maintain the molten state, which equals the holding

losses, is given by: $L_{\text{surface}} = Q_{\text{holding}} = \left(\dot{Q}_{\text{rad}} + \dot{Q}_{\text{conv}} + \dot{Q}_{\text{cond}}\right) \cdot t$.

We assume a cubic furnace for simplicity (results would be similar for a cylinder). The edge length $\ell$ for a given $M$ is given by $M = \rho V = \rho \ell^3$ where $\rho$ is the metal density. Then, the open surface area (top) is $A_s = \ell^2$ and the conductive surface area (sides + bottom) is $A_c = 5\ell^2$. Therefore,

$$L_{\text{surface}} = \left( \left( \varepsilon \cdot \sigma \cdot \left( \overline{T}^4 - T_{\text{init}}^4 \right) + h \left( \overline{T} - T_{\text{init}} \right) \right) \cdot t + \frac{k \cdot 5}{d} \cdot \left( \overline{T} - T_{\text{init}} \right) \cdot t \right) \left( \frac{M}{\rho} \right)^{\frac{2}{3}} = a_{\text{surface}} M^{\frac{2}{3}}.$$

The surface losses scale less than proportionally with mass $M$. Finally, the energy requirement for the furnace (after combustion) is $Q_{\text{furnace}} = Q_{\text{useful}} + L_{\text{surface}}$.

**Step 3: Scale elasticity of energy intensity** The total energy per unit mass is:

$$\frac{Q_{\text{total}}}{M} = \frac{1}{\eta_{\text{comb}}} \left( \frac{Q_{\text{useful}}}{M} + \frac{L_{\text{surface}}}{M} \right) = \frac{1}{\eta_{\text{comb}}} \left( a_{\text{useful}} + a_{\text{surface}} \, M^{-\frac{1}{3}} \right). \tag{OA.8}$$

Then,

$$\frac{d \log \frac{Q_{\text{total}}}{M}}{d \log M} = -\frac{L_{\text{surface}}}{Q_{\text{furnace}}} \frac{1}{3}. \tag{OA.9}$$

**Numerical implementation for the aluminum melting furnace.** We use physical constants and typical values from the engineering literature:

- Aluminum: $c = 900\,\text{J}\,\text{kg}^{-1}\,\text{K}^{-1}$, $L_f = 397{,}000\,\text{J}\,\text{kg}^{-1}$, $\rho = 2700\,\text{kg}\,\text{m}^{-3}$, emissivity $\varepsilon = 0.2$.
- Temperatures: operating $\overline{T} = 973\,\text{K}$, initial $T_{\text{init}} = 293\,\text{K}$.
- Furnace: convective coefficient $h = 20\,\text{W}\,\text{m}^{-2}\,\text{K}^{-1}$, insulation conductivity $k = 2\,\text{W}\,\text{m}^{-1}\,\text{K}^{-1}$, wall thickness $d = 0.25\,\text{m}$.
- Holding time $t = 10\,\text{h} = 36{,}000\,\text{s}$—a typical work shift ranging from 4 h (advanced methods such as CRIMSON) to 16 h (traditional casting); melt mass $M = 250\,\text{kg}$.

We find that $\frac{L_{\text{surface}}}{Q_{\text{furnace}}} = 0.6$. Foundries have a large surface-loss share because the furnace holds metal molten for long periods at high temperatures. Therefore,

$$\frac{d \log \frac{Q_{\text{total}}}{M}}{d \log M} = -0.2. \tag{OA.10}$$

**Alternative calibrations.** *Billet furnace.* The Indian Bureau of Energy Efficiency (link) details the energy balance for an oil-fired billet reheating furnace for rolling mills, processing 6000 kg/h, yielding $\frac{L_{\text{surface}}}{Q_{\text{furnace}}} \approx 0.26$. The billet furnace has lower surface losses: it does not hold metal molten for hours, and measured outer-surface temperatures are around 100 °C. This implies a scale elasticity equal to $-0.09$. This magnitude, derived from first principles, is broadly in line with the one that can be computed from the furnace efficiency by capacity tables provided by the Indian Bureau of Energy Efficiency (link) for a pusher-type billet reheating furnace, equal to $-0.16$.

*Forging furnace.* The same document provides furnace efficiency by capacity tables for a pusher forging furnace, from which we estimate a scale elasticity of $-0.23$.

**Taking stock.** From this analysis, we take $-0.15$ to be a reasonable scale elasticity for fired systems, and use this value in our summary table.

**Functional form for the model.** In the model, we write $y = A_Y \min\{m, \phi \frac{e}{y^\epsilon}\}$ where $e$ is energy, $m$ is materials, $\phi$ is a technology quality parameter, and $A_Y$ is aggregate productivity. At the optimum, $m = \phi \frac{e}{y^\epsilon}$ so that output writes $y = A_Y \phi \frac{e}{y^\epsilon}$. To show how this functional form can emerge from the physical problem, we consider the case where $a_{\text{useful}}/a_{\text{surface}} \approx 0$. This will for instance occur for processes where the holding time is very large, so that most energy is used to compensate for surface losses. Concrete examples include hot-dip galvanizing baths where molten zinc is held at ∼450 °C continuously, or heated storage tanks for asphalt, molten sulphur, or molten polymers. Then, taking equation (OA.8) and replacing $y = M$ and $e = Q_{\text{total}}$, we can write:

$$y = \frac{\eta_{\text{comb}}}{a_{\text{surface}}} \frac{e}{y^{-\frac{1}{3}}}. \tag{OA.11}$$

Here, the within technology physical elasticity is $\epsilon = -\frac{1}{3}$, and technology quality $\phi = \frac{\eta_{\text{comb}}}{a_{\text{surface}}}$ captures both combustion efficiency and the parameters determining surface losses (wall thickness, insulation conductivity, etc.).

## C.2 Process Heating (Steam System)

**Step 1. Steam production in a boiler.** We estimate the energy input $Q_{\text{total}}$ required to deliver $Q_{\text{steam}}$ to the process. In a boiler, most of the energy loss comes from the combustion step (due, as explained above, to heat losses in dry flue gas, and evaporation of the water due to the hydrogen or moisture in the fuel). Surface heat losses occur at the shell of the boiler, but they are quantitatively negligible

(of the order of magnitude of 0.25% of the energy input). Hence, we ignore them. Therefore, as in step 1 in the section above, we can write:

$$Q_{\text{total}} = \frac{Q_{\text{steam}}}{\eta_{\text{boiler}}} \tag{OA.12}$$

where $\eta_{\text{boiler}}$ is boiler efficiency and does not depend on scale. The energy out of the boiler and entering the process can be written as $Q_{\text{steam}} = h\, m_s$ where $m_s$ is the mass of steam and $h$ is its specific enthalpy (measured relative to the feedwater). Of this, the part transferred to the dryer is the latent heat released as the steam condenses, $m_s\, h_{fg}$ where $h_{fg}$ is the specific latent enthalpy of the steam at the supply pressure.

**Step 2. Steam-based process heating.** We consider a drying task, akin to the use of steam in the paper products industry. The task is to remove $r$ kg water per kg product (drying duty) for a mass of dry product $M$ (kg).

The energy supplied by the steam is split between energy delivered to the dryer heat-exchange interface $Q_{\text{in}}$, flow enthalpy losses $L_{\text{flow}}$, and surface losses $L_{\text{surf}}$: $m_s h_{fg} = Q_{\text{in}} + L_{\text{flow}} + L_{\text{surf}}$.

To obtain $Q_{\text{in}}$, consider the energy required to vaporize the mass $r\,M$ of water, given by $r\,M\,h_{\text{vap}}$ where $h_{\text{vap}}$ is latent heat of vaporization of the water in the product. Given a fraction $\eta_{ht}$ of the latent heat entering the dryer heat-exchange interface that reaches the product (exchanger/drum effectiveness), we obtain:

$$Q_{\text{in}} = \frac{1}{\eta_{ht}} r\, M\, h_{\text{vap}}. \tag{OA.13}$$

Denote $\psi$ the fractional enthalpy loss per kg steam across the loop (dimensionless). This captures condensate losses, flash steam venting, steam traps and valve leakages, all of which are primarily proportional to the mass of steam. Then,

$$L_{\text{flow}} = \psi\, m_s\, h_{fg}. \tag{OA.14}$$

Surface losses are proportional to the surface area of the dryer. Considering that the volume of the dryer scales proportionally with the mass to dry, we obtain:

$$L_{\text{surf}} = c_{\text{surf}}\, M^{2/3}. \tag{OA.15}$$

**Step 3. Total energy per kg product and scale elasticity.** Solving for $m_s$ and substituting,

$$\frac{Q_{\text{total}}}{M} = \frac{1}{\eta_{\text{boiler}}} \frac{h}{h_{fg}} \frac{1}{1-\psi} \left( \frac{r\, h_{\text{vap}}}{\eta_{ht}} + c_{\text{surf}}\, M^{-1/3} \right). \tag{OA.16}$$

$$\frac{d \log \frac{Q_{\text{total}}}{M}}{d \log M} = -\frac{L_{\text{surf}}}{Q_{\text{in}} + L_{\text{surf}}} \frac{1}{3}. \tag{OA.17}$$

From the detailed drying-section energy balance of an industrial fluting-paper machine (Ghodbanan, Alizadeh, Shafiei, and Rahbar Shahrouzi 2024), the surface loss is approximately 15% of the combined dryer input plus surface loss. This yields a scale elasticity around $-0.05$ for the combined boiler-dryer system.

## C.3 Mechanical Work

Consider the case of a pump driven by an electric motor, a common system in industrial plants. Electrical power $P_{\text{elec}}$ is converted by the motor into shaft power $P_{\text{shaft}}$; the pump converts the shaft power into useful hydraulic power $P_{\text{hyd}}$. The useful output is the power to move fluid against a head $H$: for fluid of density $\rho$ at volumetric flow rate $\dot{V}$, $P_{\text{hyd}} = \rho\, g\, H\, \dot{V}$, with $g$ gravitational acceleration. We take the flow rate $\dot{V}$ as the measure of scale.

**Step 1: the motor.** $\eta_{\text{motor}}(\ell, P_r)$ is motor efficiency defined as the ratio of the mechanical power delivered at the rotating shaft to the electrical power input at its terminals. $P_r$ is the nameplate-rated power of the motor (often expressed in horsepower, or watts). $\ell$ is the load, defined by dividing the actual power by $P_r$. Fixing the pump efficiency, we can write energy intensity as a function of scale as

$$\frac{d \log \frac{P_{\text{elec}}}{\dot{V}}}{d \log \dot{V}} = -\frac{d \log \eta_{\text{motor}}(\ell, P_r)}{d \log \dot{V}}.$$

Motor efficiency depends on the magnitude of two types of losses. First, magnetic core losses and friction and windage losses are essentially fixed for a given motor design ($P_r$), and do not scale with load $\ell$. They also scale less than linearly with $P_r$. Second, resistance losses are proportional to the square of the current, and hence vary with the load. Therefore, the scale elasticity of energy intensity depends on whether the increase in scale is accommodated by an increase in $P_r$ at constant load, or an increase in load $\ell$ at constant nameplate horsepower. We assume that, in the long run, scale increases have to be accommodated by higher nameplate

horsepower motors.

To estimate the elasticity of efficiency with respect to nameplate horsepower, we use the reference tables produced in US Department of Energy (1997). We obtain:

$$\frac{d \log \eta_{\text{motor}}(\ell, P_r)}{d \log P_r} = 0.03$$

**Step 2: the pump.** The pump's shaft power $P_{\text{shaft}}$ is the useful hydraulic power $P_{\text{hyd}}$ plus internal losses: $P_{\text{shaft}} = \rho g H \dot{V} + L_{\text{hydraulic}} + L_{\text{disk}} + L_{\text{leak}} + L_{\text{rub}}$.

To estimate the scale elasticity, we compare geometrically identical pumps, indexed by impeller diameter $D$, each raising the fluid by the same head $H$. The Euler relation makes head proportional to the square of the impeller tip speed $u$, $H \propto u^2/g$, so a fixed head demands a fixed tip speed. With the tip speed fixed, every flow velocity inside the pump is fixed too. The flow rate is then (flow area) $\times$ (velocity) $\propto D^2$, so $\dot{V} \propto D^2$ and $D \propto \dot{V}^{1/2}$. The Reynolds number—which gauges inertia against viscous friction—scales as $Re \propto (\text{velocity}) \cdot D/\nu \propto \dot{V}^{1/2}$, with $\nu$ the kinematic viscosity.

*Hydraulic losses* are due to friction and turbulence as fluid is forced through the impeller and volute passages. Part is skin friction along the passage walls, whose coefficient $C_f$ falls with Reynolds number, $C_f \propto Re^{-0.2}$ (Schlichting and Gersten 2017): a larger pump carries thinner boundary layers and loses slightly less. The rest is "form" loss—incidence at the blade, diffusion, recirculation, wake mixing—which grows in step with the flow. The hydraulic-efficiency correlation of Gülich (2020) bundles the two: across pump sizes at a fixed specific speed $n_q \approx 30$ (a dimensionless shape index), the hydraulic loss writes $L_{\text{hydraulic}} = a_{\text{hydraulic}} \dot{V}^{0.86}$.

*Disk friction*. Shearing the fluid in the narrow gaps between the impeller shrouds and the casing walls costs power $P_{RR} = k_{RR}\, \rho\, \omega^3 R^5$, with $\omega$ rotational speed, $R$ impeller radius, and $k_{RR}$ disk-friction coefficient. At fixed tip speed $\omega^3 R^5 \propto D^2 \propto \dot{V}$ ($\omega \propto D^{-1}$, $R \propto D$), while $k_{RR}$ declines as $Re^{-0.2}$. Together, the disk-friction loss scales as $L_{\text{disk}} = a_{\text{disk}} \dot{V}^{0.90}$ (Gülich 2020).

*Leakage*. Driven by the impeller's pressure rise, some fluid escapes back through the wear-ring clearance gap, from the high-pressure discharge to the low-pressure inlet, only to be pumped again. The leak rate is (gap area) $\times$ (gap velocity). The gap is set by manufacturing tolerance and shaft deflection. The empirical correlations of Gülich (2020) show that the wear-ring gap area grows only sublinearly with size while the throttling velocity declines, so that $L_{\text{leak}} = a_{\text{leak}} \dot{V}^{0.64}$.

TABLE OA.5: Non-Process Energy Use by End-Use

| End-Use Category | Share (%) | Scale Elasticity |
|---|---|---|
| Facility HVAC | 49.5 | $\approx -0.65 \times 0.95 \times \frac{1}{3}$ |
| Facility Lighting | 34.7 | $\approx 0$ |
| Other Facility Support | 10.6 | $\approx 0$ |
| Onsite Transportation | 2.9 | $\approx 0$ |
| Other Nonprocess Use | 2.5 | $\approx 0$ |
| Total Nonprocess | 100 | $\approx -0.10$ |

*Notes*: Shares are from the U.S. Manufacturing Energy Consumption Survey (2018), scale elasticities are authors' own calculations.

*Bearing and seal friction.* Mechanical rubbing in bearings and shaft seals is a near-fixed overhead, growing weakly with size; Gülich (2020) yields $L_{\text{rub}} = a_{\text{rub}} \dot{V}^{0.75}$.

Denoting each loss $L_i$'s exponent as $\beta_i$, we obtain:

$$\frac{d \log \frac{P_{\text{shaft}}}{\dot{V}}}{d \log \dot{V}} = \sum_i \frac{L_i}{P_{\text{shaft}}} (\beta_i - 1)$$

For a mid-specific-speed pump ($n_q \approx 30$), the component correlations of Gülich (2020) give loss shares (of shaft power) for hydraulic 7.1%, disk friction 3.2%, leakage 1.4%, bearings/seals 1.0%. This yields a pump scale elasticity of $-0.020$.

**Step 3: Combined motor-pump system.** Adding the motor and pump channels gives a combined scale elasticity of the system of $-(0.03 + 0.02) = -0.05$.

## C.4 Other Process

This includes electro-chemical processes (that convert electrical energy directly into chemical change at an electrode, 19%), process refrigeration (28%), and process uses not classified elsewhere (53%). We assume an elasticity of 0 for this residual category. Electro-chemical uses are a true zero: by Faraday's law the energy required is proportional to the mass transformed, so energy per unit of output is fixed by the cell voltage and independent of scale. Process refrigeration is likely mildly negative, since cooling loads carry a surface-area holding-loss component analogous to heating; but the driving temperature difference is an order of magnitude smaller than in typical process heat, and these losses are a small share of the load, so that the scale economy is very small. The remaining uses not classified elsewhere are a heterogeneous residual; we therefore assign it a neutral elasticity of 0.

### C.5 Non-process

Non-process covers a range of diverse energy uses, summarized in Table OA.5. Among these end-uses, only facility heating, ventilation, and air-conditioning (HVAC) exhibits a significant scale elasticity. Indeed, in facility heating, the same negative scale elasticity due to surface losses as the one derived in (OA.9) applies, and the share of holding losses in total energy is close to 1 since heating is often continuous (i.e. $t$ tends to infinity). Using a share of holding losses equal to 0.95, we get a negative scale elasticity for facility heating equal to $-0.32$. This effect is less important for ventilation and air-conditioning, which have a lower fraction of energy lost via the surface (compared to other sources of loss), so we assume a scale elasticity of 0 for ventilation and air-conditioning. To obtain the fraction of heating in HVAC (not reported in the Manufacturing Energy Consumption Survey), we use the 2018 Commercial Buildings Energy Consumption Survey and obtain a value equal to 65%.

## D. Identification of technology cost function

We suppose that the cost of a given technology bundle is given by:

$$\tilde{C}_t(\phi, \bar{y}) = \frac{1}{\gamma A_{Yt}} \phi^{\gamma} \bar{y}^{\delta} (w_t^l)^{\alpha} (w_t^k)^{1-\alpha},$$

Note that we also have $\tilde{C}_{it} = w_t^l l_{it} + w_t^k k_{it}$ by assumption.

**Identification of $\gamma$.** At the optimum, $\phi = \frac{y^{1+\epsilon}}{A_{Ye}}$ and $\bar{y} = y$. Therefore, we have:

$$\log \tilde{C}_{it} = -\gamma \log e_{it} + (\delta + \gamma + \gamma\epsilon) \log y_{it} + \log \left( (w_t^l)^{\alpha} (w_t^k)^{1-\alpha} \right) - (1+\gamma) \log(A_{Yt})$$

$-\gamma$ is the elasticity of the cost function with respect to energy used, keeping production $y$ constant. This elasticity can be measured using the data on required investment and associated annual energy savings obtained from the UNIDO technology compendia data. Note that the technology compendia report these amounts for a given firm assumed to produce the same quantity before and after the technology investment, hence they keep production $y$ constant.

This data reports the required investment $\Delta k_{\tau is}$ in INR and associated annual energy savings $\Delta e_{\tau is}$ in gigajoules for a firm $i$ in industry $s$ adopting technology $\tau$. The coefficient $\gamma$ cannot be directly read from Figure 6, because both axes are

TABLE OA.6: Estimation of $\gamma$

**Panel A:** Baseline $\Delta \log(\tilde{\mathcal{C}}_{\tau s})$

| | Investment cost (%) | | | | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Energy savings (%) | 0.113*** | 0.095*** | 0.114*** | 0.112*** | 0.099*** | 0.098*** |
| | (0.020) | (0.026) | (0.008) | (0.008) | (0.020) | (0.017) |
| Industry | Foundry | Foundry | Hand-tool | Hand-tool | Ceramic | Ceramic |
| Sample | Baseline | Expanded | Baseline | Expanded | Baseline | Expanded |
| Observations | 18 | 27 | 9 | 14 | 20 | 30 |
| R-squared | 0.68 | 0.35 | 0.96 | 0.94 | 0.59 | 0.55 |

**Panel B:** Alternative $\Delta \log(\tilde{\mathcal{C}}_{\tau s})$

| | Investment cost (%) | | | | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Energy savings (%) | 0.259*** | 0.218*** | 0.498*** | 0.488*** | 0.174*** | 0.172*** |
| | (0.045) | (0.060) | (0.037) | (0.037) | (0.035) | (0.029) |
| Industry | Foundry | Foundry | Hand-tool | Hand-tool | Ceramic | Ceramic |
| Sample | Baseline | Expanded | Baseline | Expanded | Baseline | Expanded |
| Observations | 18 | 27 | 9 | 14 | 20 | 30 |
| R-squared | 0.68 | 0.35 | 0.96 | 0.94 | 0.59 | 0.55 |

*Notes*: This table reports the results of estimating equation (OA.18) separately for each industry $s$. The two methods for the construction of $\Delta \log(\tilde{\mathcal{C}}_{\tau s})$ are detailed in the text. "Sample" refers to the construction of the sample, as detailed in the text. ***, **, and * denote significance at the 1%, 5% and 10% levels, respectively.

expressed in levels as opposed to log changes. To estimate $\gamma$, we now restrict our attention to the observations for which we observe pre-existing energy consumption, allowing us to construct $\Delta \log(e_{\tau is})$, the log change in energy consumption upon the adoption of technology $\tau$. To construct the log change in fixed cost, we proceed as follows. We construct the cost of the pre-existing capital-labor bundle $\tilde{\mathcal{C}}^0_{\tau is} = (w^k k^0_{\tau is} + w^l l^0_{\tau is})$ by multiplying the production level $y_{\tau is}$ reported in the UNIDO data by the median $(w^k k + w^l l)/y$ ratio for this industry found in the ASI data.[47] Then, $\Delta \log(\tilde{\mathcal{C}}_{\tau is})$, the change in the fixed cost of the technology, is defined as $\Delta \log(\tilde{\mathcal{C}}_{\tau is}) = w^k \Delta k_{\tau is} / (w^k k^0_{\tau is} + w^l l^0_{\tau is})$. This is the user cost of increasing the capital stock by $\Delta k_{\tau is}$, relative to the pre-existing total fixed cost. We use the same $w^k$ as the one used to construct the cost of capital. This measure implicitly assumes that investing in the more energy-efficient capital does not require new staff or training cost. The fixed cost increase is only the capital amount. We also

[47] For instance, if for industry $s$, UNIDO reports production in tons, we compute fixed costs in INR per ton in the ASI data, and multiply this value by production in the UNIDO data.

propose an alternative measure $\Delta \log(\tilde{\mathcal{C}}_{\tau is}) = \Delta k_{\tau is}/k^0_{\tau is}$. Here, by contrast, we assume that investing in the more energy-efficient capital comes with an unobserved labor cost proportional to the increase in the capital stock. When several observations exist for a given industry×technology pair (typically coming from different clusters), we take the average of all variables. We then estimate the following specification separately for each sector $s$:

$$\Delta \log \tilde{\mathcal{C}}_{\tau s} = \alpha + \beta(-\Delta \log e_{\tau s}) + \varepsilon_{\tau s} \tag{OA.18}$$

where $(-\Delta \log e_{\tau s})$ is energy savings relative to the baseline and the regression coefficient identifies $\gamma$.[48]

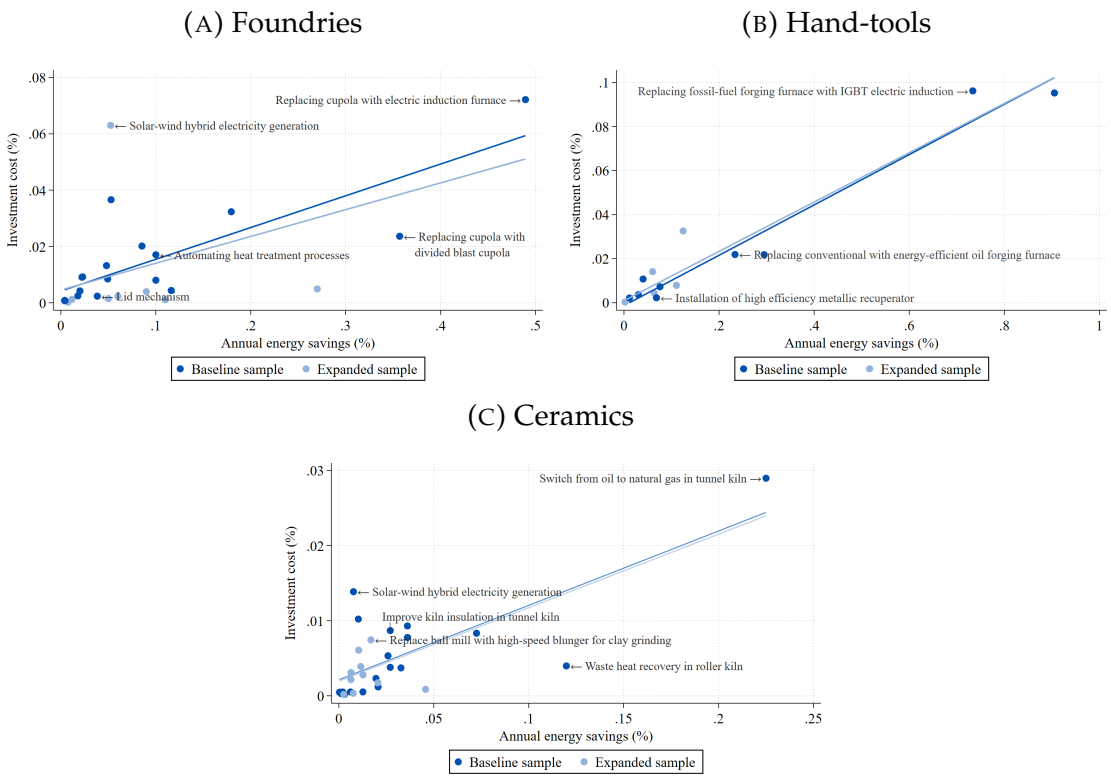
The results are presented in Table OA.6. Panel A reports results for the baseline measure of $\Delta \log(\tilde{\mathcal{C}}_{\tau s})$, while Panel B presents results for the alternative version. Odd columns use only observations for which we observe pre-existing energy consumption and production quantity. In even columns, we expand the sample by imputing these variables with the industry-specific average when missing. Looking at Panel A, we obtain $\gamma \approx 0.1$ with coefficients very stable across industries and sample choice. The R-squared are high, in particular in odd columns where we use only the highest quality data points, suggesting that the functional form is appropriate. The results in Panel B yield higher slopes by construction: averaging across the three industries, this second set of results implies $\gamma \approx 0.30$.

**Identification of $\delta$.** $\delta$ is the elasticity of the fixed cost with respect to capacity, for a given technology quality $\phi$. A lower value of $\delta$ implies that the capital-labor bundle behaves more like a fixed cost, providing more room for a scale elasticity of technological upgrading.

We estimate $\delta$ as follows. Our main data source is data on equipment costs by capacity compiled by the U.S. Department of Energy (Loh, Lyons, and White 2002). The goal of this dataset is to allow engineers to rapidly perform cost estimates when evaluating a new system. It covers a relatively comprehensive set of 20 conventional process equipment types. While the report has a chemical engineering focus, it covers most of the equipment found in the UNIDO compendia for foundries, hand-tools, ceramics, and dairies. An observation is characterized by an equipment type $e$ (e.g., dryer), a model $j$ (e.g., direct contact rotary), and a capacity $i$ (e.g., 100 sq. ft.). The capacity unit is specific to each type of equipment.

[48] This analysis cannot be conducted for dairies because for this industry the UNIDO data does not report output in a standardized unit.

FIGURE OA.4: Estimation of $\gamma$




*Notes:* This figure shows the scatter plots corresponding to equation (OA.18), for each industry $s$. We use the baseline version of the outcome variable $\Delta \log\left(\tilde{\mathcal{C}}_{\tau s}\right)$. Dark (light) blue dots correspond to the baseline (expanded) sample, as described in the text.

For each equipment type $e$, we then estimate the following model:

$$\log\left(\text{Price}_{eij}\right) = \alpha_{ej} + \beta \log\left(\text{Capacity}_{eij}\right) + \varepsilon_{eij} \tag{OA.19}$$

$\alpha_{ej}$ are equipment$\times$model fixed effects, so that $\beta$ is estimated using variation in prices across capacities for a given equipment$\times$model pair (hence keeping model quality $\phi$ constant). The results are reported in Table OA.7. Across all equipment types, we find an average scale elasticity equal to 0.52, with most estimates in the 0.4-0.6 range. The R-squared values are high, providing support for the log-log functional form.

We complement this analysis in several ways. First, we collect data on price by capacity for additional equipments mentioned in the UNIDO compendia but that do not appear in the U.S. Department of Energy data (chiller, bulk milk cooler, pasteurizer, variable frequency drive, motor, and biomass boiler). Across these equipment categories, we find an average elasticity equal to 0.54, very close to our baseline estimate. Second, we note that these elasticities all relate to process

TABLE OA.7: Estimation of $\delta$

| Equipment type | Coefficient | Std. Error | R-squared | Observations |
|---|---|---|---|---|
| Vessel | 0.314*** | (0.023) | 0.75 | 44 |
| Storage Tank | 0.544*** | (0.039) | 0.95 | 18 |
| Column | 0.330*** | (0.014) | 0.99 | 230 |
| Heat Exchanger | 0.521*** | (0.035) | 0.94 | 40 |
| Air Cooler | 0.404*** | (0.045) | 0.91 | 15 |
| Furnace | 0.713*** | (0.018) | 0.99 | 9 |
| Cooling Tower | 0.345*** | (0.042) | 0.91 | 9 |
| Package Steam Boiler | 0.434*** | (0.033) | 0.97 | 8 |
| Evaporator | 0.501*** | (0.005) | 1.00 | 20 |
| Crusher | 0.951*** | (0.116) | 0.97 | 19 |
| Mill | 0.574*** | (0.021) | 0.99 | 16 |
| Dryer | 0.664*** | (0.080) | 0.98 | 19 |
| Centrifuge | 0.792*** | (0.065) | 0.99 | 24 |
| Filter | 0.379*** | (0.044) | 0.99 | 26 |
| Agitator | 0.495*** | (0.057) | 0.93 | 6 |
| Pump | 0.360*** | (0.025) | 0.95 | 81 |
| Compressor | 0.509*** | (0.031) | 0.92 | 33 |
| Centrifugal Fan | 0.568*** | (0.064) | 0.92 | 9 |
| Rotary Blower | 0.415*** | (0.034) | 0.97 | 6 |
| Turbine | 0.650*** | (0.055) | 0.98 | 24 |
| Average | 0.523 | | | |

*Notes*: This table provides the results of estimating equation (OA.19) by equipment type, as well as the average coefficient across equipments.

equipments. Scale elasticities for items related to instrumentation and control, site improvements, and facilities are generally much lower—estimated around 0.15 in James, Leptinsky, Turner et al. (2022)—so that our estimate of $\delta$ is likely to be on the conservative side. Finally, our baseline estimate is very close to 0.6, a benchmark in the scale elasticity literature (see, e.g., Tribe and Alpine 1986).

**Taking Stock: Implied Energy Intensity Scale Elasticity.** Putting together our estimates of $\epsilon$, $\gamma$, and $\delta$, we can estimate the scale elasticity of energy intensity implied by our model:

$$\frac{d\log\left(\frac{e}{y}\right)}{d\log y} = \underbrace{\tilde{\epsilon}}_{\text{Overall}} = \underbrace{\epsilon}_{\text{Within-tech.}} - \underbrace{\frac{\epsilon+1-\delta}{1+\gamma}}_{\text{Tech. improvement}}$$

Table OA.8 shows the obtained scale elasticity. Our baseline calibration uses $\epsilon = -0.10$, $\gamma = 0.20$ (we use the mid-point of the $\gamma$ estimate obtained with the two alternative definitions of the fixed cost), and $\delta = 0.52$. We obtain a scale elasticity equal to $-0.42$, virtually equal to the one we measure in the data ($-0.425$). The table also shows the robustness of this estimate to the choices of

TABLE OA.8: Scale Elasticity of Energy Intensity

| $\epsilon$ | $\delta$ | $\gamma$ | Within tech. | Tech. improv. | Overall |
|---|---|---|---|---|---|
| *Baseline* | | | | | |
| $-0.10$ | 0.52 | 0.20 | $-0.10$ | $-0.32$ | $-0.42$ |
| *Robustness* | | | | | |
| $-0.10$ | 0.52 | 0.10 | $-0.10$ | $-0.35$ | $-0.45$ |
| $-0.10$ | 0.52 | 0.30 | $-0.10$ | $-0.29$ | $-0.39$ |
| $-0.10$ | 0.60 | 0.20 | $-0.10$ | $-0.25$ | $-0.35$ |
| $-0.10$ | 0.52 | 0.20 | $-0.10$ | $-0.32$ | $-0.42$ |
| $-0.05$ | 0.52 | 0.20 | $-0.05$ | $-0.36$ | $-0.41$ |
| $-0.15$ | 0.52 | 0.20 | $-0.15$ | $-0.27$ | $-0.42$ |

*Notes:* This table reports the scale elasticity of energy intensity as a function of model parameters.

underlying parameters: for $\gamma \in \{0.10, 0.30\}$ (the estimates obtained with the two alternative definitions of the fixed cost), a more conservative value of $\delta = 0.60$, and $\epsilon \in \{-0.05, -0.15\}$.

SUPPLEMENTARY MATERIAL

# GROWTH, FIRM SCALE, AND THE ENERGY INTENSITY OF PRODUCTION

BY KATHRYN MCDONALD, NOÉMIE PINARDON-TOUATI, AND CONOR WALSH

August 2026

# SM.1. DATA SOURCES AND VARIABLES CONSTRUCTION

## SM.1.1 Annual Survey of Industries: description and variable definitions

The ASI is a dataset put together by India's Ministry of Statistics and Programme Implementation (MOSPI). The reference period for each survey is the accounting year, which in India begins on the 1st of April and ends on the 31st of March the following year. Throughout the paper we reference the surveys by the earlier of the two years covered.

**Coverage and sampling methodology.** The ASI contains information on a representative sample of manufacturing establishments, conditional on them taking part in the organized sector, and either employing more than 20 employees, or employing more than 10 employees and using electricity. We refer to establishments as firms. We call the subpopulation of firms satisfying this criterion the ASI population. Within the ASI population, ASI defines a Census sector which is sampled exhaustively and a Sample sector for which the microdata contains only a representative sample. Details of the sampling methodology for the ASI are shown in Table SM.1. ASI provides sampling weights, which we use to weight all data moments.

**Sample selection.** We start with 1,068,114 firm×year observations. We subsequently employ multiple sample selection rules. First, we restrict the sample to firm×year observations with either positive reported gross sales, or positive reported sales at the factory gate. This drops one third of all observations (∼360,000). Next, we disregard all observations that, even when reporting positive sales, are marked as closed and have their sales value exactly copied from the previous year, suspecting these firms to be actually closed. This drops $\sim 1,000$ additional observations. Third, we drop all firms that report no days worked and no persons employed, dropping a supplementary $\sim 200$ observations. These cleaning steps follow Martin, Nataraj, and Harrison (2017).We drop observations in non-manufacturing industries (NIC-98 codes above 36). In particular, ASI contains a few observations belonging to the "Electricity, gas, and water supply sector" (NIC-98 industry 40), which we drop, so that our sample includes only final users of energy. For the same reason, we drop NIC-98 industry 23 "Manufacture of coke, refined petroleum products, and nuclear fuel."

**Construction of key variables.** We detail below the construction of the main variables for our analysis.

*Firm-level variables.*

- Sales: We construct sales as the sum of the gross value of products sold (including distribution expenses, as well as taxes and subsidies). We include the value of electricity sold by the establishment (consistent with how we define fuels), if any.

TABLE SM.1: Sampling Methodology for Indian ASI

| Period | Census Sector | Sample Sector |
|---|---|---|
| 1998 | Complete enumeration states, establishments with $>$ 200 workers, all joint returns | Stratified within state $\times$ 4-digit industry (NIC-98), minimum of 8 establishments per stratum |
| 1999-2003 | Complete enumeration states, establishments with $\geq$ 100 workers, all joint returns | Stratified within state $\times$ 4-digit industry (NIC-98), 12% sampling fraction (20% in 2002), minimum of 8 establishments per stratum |
| 2004-2006 | 6 less industrially developed states, 100 or more workers, all joint returns, all establishments within state $\times$ 4-digit industry with $\leq$ 4 units | Stratified within state $\times$ 4-digit industry, minimum 4 establishments, 20% sampling fraction |
| 2007 | 5 less industrially developed states, 100 or more workers, all joint returns, all establishments within state $\times$ 4-digit industry with $\leq$ 6 units | Stratified within state $\times$ 4-digit industry, minimum 6 establishments, 12% sampling fraction |
| 2008-2011 | 6 less industrially developed states, 100 or more workers, all joint returns, all establishments within state $\times$ 4-digit industry with $\leq$ 4 units | Stratified within state $\times$ 4-digit industry, minimum 4 establishments, 20% sampling fraction (19% in 2009, 18% in 2010) |
| 2012-2013 | 6 less industrially developed states, 100 or more workers, all joint returns, all establishments within district $\times$ 4-digit industry with $\leq$ 4 units | Stratified within district $\times$ 4-digit industry, minimum 4 establishments, 20% sampling fraction |
| 2014 | 7 less industrially developed states, 100 or more employees, all joint returns, all establishments within district $\times$ 4-digit industry with $\leq$ 4 units | Stratified within district $\times$ 4-digit industry, minimum 4 establishments, 17% sampling fraction |
| 2015-2017 | 7 less industrially developed states, 100 or more employees (75 or more in six states, 50 or more in three states), all joint returns, all establishments within district $\times$ 3-digit industry with $\leq$ 4 units | Stratified within district $\times$ 3-digit industry, minimum 4 establishments, 16-20% sampling fraction |

*Notes:* Baseline sampling fractions are shown, not accounting for state-specific exceptions. Reported fractions refer to the sample sector as a whole. Sourced from Bils, Klenow, and Ruane (2021) and ASI instruction manuals.

- Capital: Capital is the closing book value of fixed assets (net of depreciation). It excludes intangible assets and current assets.
- Cost of capital: Product of the capital stock and of the user cost of capital, plus cost of rented capital. The user cost of capital is the depreciation rate plus the interest rate on loans, aggregated at the industry$\times$year level.
- Cost of materials: We construct the cost of materials as the sum of the value of non-fuel materials consumed, operating expenses (freight and transportation charges, taxes paid) and insurance charges.
- Cost of fuels: Sum of expenses on electricity, oil, coal, gas, and other fuels. We do not

include the purchase value of electricity generated within the firm, since we account for the fuels used to generate that electricity.

- Labor: Total days worked.
- Cost of labor: We construct labor costs as total payments to labor over the course of the year. These payments include wages and salaries, bonuses, contributions to old-age pension funds (and other funds), and all welfare expenses. We also include the costs of contract and commission work.

*Firm×product-level variables and firm-level price and quantity indices.* A key advantage of the ASI is that both for the products that manufacturing establishments produce and the inputs they buy, we observe information on sales, quantities, and unit values, with products classified at the 5-digit NPC level (around 1,200 distinct product codes for our sample of manufacturing firms). Some sections of our analysis exploit this data. We detail the construction of firm×product- and firm-level price and quantity variables here, and provide more details on the product classification below.

We construct a panel of firm×product prices and quantities. We denote $\Delta \log p_{ijt}$ and $\Delta \log y_{ijt}$ as the changes in the log of price $p$ and log of quantity $y$ of product $j$ sold by firm $i$ at time $t$, respectively. The key cleaning steps are: (i) harmonizing product codes within firms, as sometimes firms report different codes for the same product in consecutive years; (ii) correcting observations for which the product of unit values and quantity sold differ significantly from the reported sales value; (iii) correcting unit mistakes: as noted by Boehm et al. (2022), the data contains incorrect entries due to misplaced commas, which we address by rescaling values up or down when the price was multiplied by $\sim 10^n$ and the quantity by $\sim 10^{-n}$ with respect to the previous year.

Even working with narrowly-defined product categories, unobserved heterogeneity could prevent a meaningful comparison of prices across firms. We always work with within firm×product changes in the price (or quantity), largely alleviating this concern. The set of products for which we observe a valid price (quantity) change for two consecutive years, which we denote $\mathcal{J}_i$, accounts for, on average, 75% of firm-level total sales.

We define the firm-level price index as the Törnqvist-weighted change in the observed firm×product-level price changes: $\underline{\Delta} \log p_{it} = \sum_{j \in \mathcal{J}_i} \bar{s}_{ijt} \Delta \log p_{ijt}$. We use the convention of placing a bar on top of the share to denote that these shares are the mid-point of the shares in $t-1$ and $t$, and the bar under the $\Delta$ sign indicates that we take the average price change over the set of *observed* products. We construct the firm-level change in quantities as $\Delta \log y_{it} = \Delta \log \text{sales}_{it} - \underline{\Delta} \log p_{it}$.[1]

Similarly, we observe purchase value, unit price, and quantity purchased for materials classified in the same 1,200 products. We denote by $\Delta \log w_{ikt}$ and $\Delta \log x_{ikt}$ the log

[1] Because we do not observe the price and quantity change for all total sales, in general $\underline{\Delta} \log p_{it} + \underline{\Delta} \log y_{it} \neq \Delta \log \text{sales}_{it}$. We assume that the price change of observed products is on average equal to the price change for all products.

change in prices and quantities of input $k$ used by firm $i$. We perform the same cleaning steps as described for products. The inputs for which we observe a valid price (quantity) change across two consecutive years, which we denote $\mathcal{K}_i$, account for on average 57% of firm-level total input purchases. We define the firm-level input quantity change as the Törnqvist-weighted change in the observed firm×input-level quantity changes: $\underline{\Delta} \log x_{it} = \sum_{k \in \mathcal{K}_i} \bar{s}_{ikt} \Delta \log x_{ikt}$. We construct the firm-level intermediate input price index as $\Delta \log w^x_{it} = \Delta \log C^x_{it} - \underline{\Delta} \log x_{it}$.[2]

*Firm-level energy prices and quantities.* Energy purchases are broken down into electricity, oil, coal, gas (since 2008), and other fuels (including gas prior to 2008). For electricity, coal, and gas (since 2008), firms separately report unit values and quantities. For oil, ASI does not report the unit value/quantity breakdown, hence we proceed as follows. We construct an oil price series using India's Wholesale Price Index (WPI). The two main oil products used in manufacturing for which WPI series are available are furnace oil and high-speed diesel. Using the Prowess CMIE dataset, which reports detailed fuel expenditures, we estimate the expenditure shares of these two fuels in manufacturing and construct an oil price index as their weighted average. To recover a price level, we normalize the index using the average price per cubic meter reported in Prowess in 1990. In ASI, we then recover the quantity of oil purchased (in cubic meters) by dividing expenditures by unit value. For the other fuels residual, which accounts for around 13% of fuel expenditures before 2008 (including gas) and 7% of fuel expenditures after 2008 (excluding gas), we do not have a price/quantity breakdown. We define firm-level energy price and quantityindices by using Törnqvist-weighted changes over electricity (on average 57% of expenditures), coal (on average 11% of expenditures), and oil (on average 19% of expenditures), for which price/quantity breakdowns are observed over the whole sample. To construct firm-level energy quantity used in levels, we convert the quantity of oil, electricity, coal, and gas from their respective units to megajoules, and take the sum.

**Industry classification.** Our data relies on three distinct industry classification systems: NIC-98 (1998–2003), NIC-04 (2004–2007), and NIC-08 (2008 and beyond). We first address issues with the 5-digit industry codes in NIC-98, where codes are sometimes masked with zeroes or absent from the official documentation, by replacing them with the most frequent 5-digit code within each 4-digit grouping following the approach of Martin et al. (2017). Next, we apply concordances from NIC-08 to NIC-04 using the mapping provided by Rijesh (2023), and from NIC-04 to NIC-98 using the mapping provided by Martin et al. (2017). We manually supplement the mappings for industries not covered by these concordances. In case of 1:m mappings, we select the appropriate industry based on transition matrices in the microdata. We obtain 78 time-invariant industries, mostly

[2]The assumption is that $\underline{\Delta} \log x_{it}$ (the average increase in input quantity for the inputs $\underline{\mathcal{K}}_i$ for which we observe input-level data) is equal to the average change in input quantity for all inputs $\Delta \log x_{it}$. This assumption is the most natural when different material inputs are strong complements (and it is exactly true if production is Leontief).

corresponding to the 3-digit level of the NIC-98 classification.[3]

**Product classification.** Our analysis standardizes product classifications across four distinct classifications used in our sample: NPCMS 2015 (2016-2017), NPCMS 2011 (2010-2015), ASICC 2009 (2008-2009), and ASICC 2008 (pre-2008). We harmonize all product codes to NPCMS 2011, as it provides a well-defined five-digit structure that balances granularity and coverage. Given the absence of an official concordance between NPCMS 2015 and NPCMS 2011, we construct a mapping using fuzzy matching (based on product codes, descriptions, and units) and semantic embeddings (OpenAI's AA2 model). For ASICC 2009, we utilize the official concordance to NPCMS 2011 but address its limitations (such as missing or m:m mappings) by constructing a transition matrix using the ASI micro data and a mapping obtained from semantic embeddings proximity. The harmonization of ASICC 2008 to ASICC 2009 follows the concordance from Boehm et al. (2022). Table SM.2 shows an excerpt of the product classification.

TABLE SM.2: Example of NPCMS 2011 5-Digit Classification

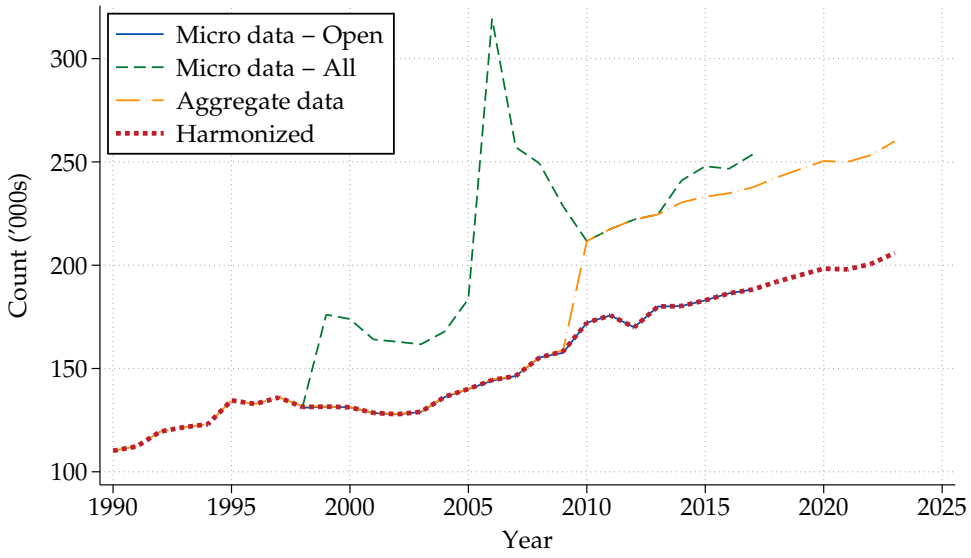
| Code | Description |
|---|---|
| 35 | Other chemical products; man-made fibres |
| 351 | Paints and varnishes and related products; artists' colours; ink |
| 35110 | Paints and varnishes and related products |
| 35120 | Artists', students' or signboard painters' colours, modifying tints, amusement colours and the like |
| 35130 | Printing ink |
| 35140 | Writing or drawing ink and other inks |
| 352 | Pharmaceutical products |
| 353 | Soap, cleaning preparations, perfumes and toilet preparations |
| 354 | Chemical products n.e.c. |
| 355 | Man-made fibres |
| 36 | Rubber and plastics products |
| 361 | Rubber tyres and tubes |
| 36111 | New pneumatic tyres, of rubber, of a kind used on motor cars |
| 36112 | New pneumatic tyres, of rubber, of a kind used on motorcycles or bicycles |
| 36113 | Other new pneumatic tyres, of rubber |
| 36114 | Inner tubes, solid or cushion tyres, interchangeable tyre treads and tyre flaps, of rubber |
| 36115 | Camel back strips for retreading rubber tyres |
| 36120 | Retreaded pneumatic tyres, of rubber |
| 362 | Other rubber products |
| 36210 | Reclaimed rubber |
| 36220 | Unvulcanized compounded rubber, in primary forms or in plates, sheets or strip; unvulcanized rubber in forms other than primary forms or plates, sheets or strip |
| 36230 | Tubes, pipes and hoses of vulcanized rubber other than hard rubber |
| 36240 | Conveyor or transmission belts or belting, of vulcanized rubber |
| 36250 | Rubberized textile fabrics, except tyre cord fabric |
| 36260 | Articles of apparel and clothing accessories (including gloves) of vulcanized rubber other than hard rubber |
| 36270 | Articles of vulcanized rubber n.e.c.; hard rubber; articles of hard rubber |
| 363 | Semi-manufactures of plastics |
| 364 | Packaging products of plastics |
| 369 | Other plastics products |

*Notes:* For codes other than 351, 361 & 362, the 5-digit classifications are not shown.

**Number of factories in ASI.** We construct the number of factories as follows. From 1990 to 2009, we use the number of factories reported in the ASI aggregated annual series. This perfectly corresponds to the number of factories declared as open in the micro data. In 2010, the aggregate series switches to counting all factories, including those declared as closed. To obtain a consistent time series, for years after 2009, we use the number of factories declared as open in the micro data. This series is available until 2017. For 2018-2023, we extend this series by using the growth rate in the number of factories found in the aggregate series. Our harmonized series is plotted in dotted red on Figure SM.1.

[3]For well-populated 3-digit industries with very heterogeneous energy shares, we keep the 4-digit level, while for a few industries with insufficient observations, we use the 2-digit level.

FIGURE SM.1: Number of Factories in ASI

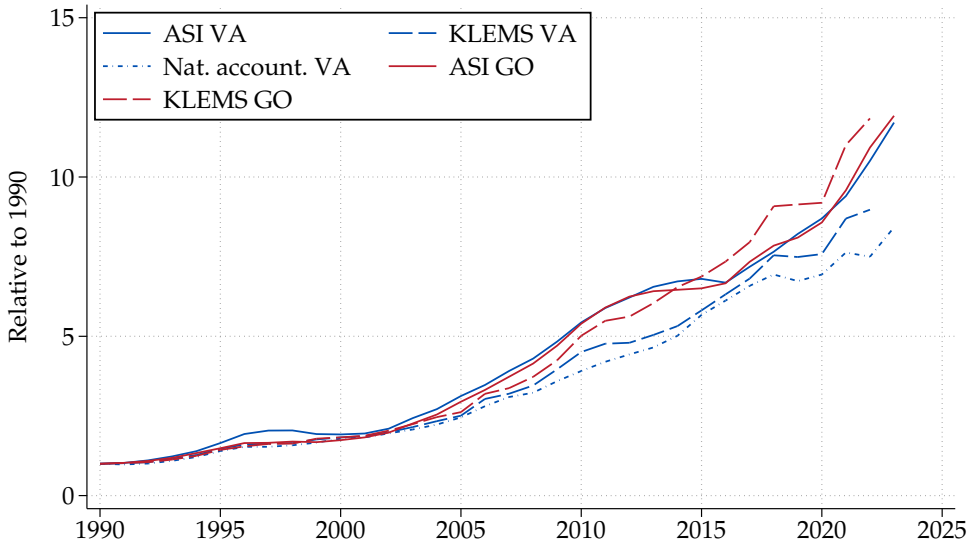



*Notes:* The figure shows the number of factories in ASI from 1990 to 2023, computed using three underlying series. "Micro data – Open" and "Micro data – All" plot, respectively, the number of factories declared open and the total number of factories in the ASI micro data. "Aggregate data" plots the number of factories reported in the ASI aggregated annual series, which counts open factories up to 2009 and all factories thereafter. The dotted red "Harmonized" series is our preferred measure: it uses the aggregate series (which coincides with open factories) up to 2009, the number of open factories in the micro data from 2010–2017, and then extrapolates to 2023 using the growth rate of the aggregate series.

**Manufacturing in ASI and in the national accounts.** ASI only covers the organized sector. Compared to the national accounts (which include the informal sector), ASI covers 64% of manufacturing value added and 83% of manufacturing output on average across years. Figure SM.2 plots gross output and value added growth in ASI, in KLEMS, and in the national accounts, indexed to 1 in 1990. Gross output trends are highly similar in ASI and in KLEMS. Value added growth is highly similar to gross output growth in ASI, but somewhat slower than gross output growth in KLEMS and the national accounts. Figure SM.3 reports energy intensity in ASI, and our estimate of energy intensity in the overall manufacturing sector, where the numerator is manufacturing energy use from the International Energy Agency (IEA) data and the denominator is gross output or value added from KLEMS.[4] The left panel reports energy intensity over gross output. The trend in ASI is broadly consistent with the IEA/KLEMS series. The ASI series shows a slightly stronger decline, suggesting stronger energy intensity improvements in the formal sector, compared to the informal sector. The right panel reports energy intensity over value added. There, the difference is more marked. As explained above, this is because while in ASI gross output and value added grew at the same rate over our sample, in KLEMS value added growth appears slower than gross output growth.

[4]Note that manufacturing energy use from the IEA data includes an item "Industry not elsewhere specified" that includes some manufacturing industry but also some components of non-manufacturing industrial use (mining, construction) when a clean breakdown is not possible.

FIGURE SM.2: Value Added and Output Growth in ASI




*Notes:* The figure plots gross output (GO) and value added (VA) indices for Indian manufacturing constructed from ASI, India KLEMS, and the national accounts. All series are expressed in 2011 constant prices and normalized to 1 in 1990. The national accounts do not report gross output.

### SM.1.2 Other data sources

**International Energy Agency.** We exploit the World Energy Balance dataset produced by the International Energy Agency (link) to obtain energy use by broad sectors for all countries.

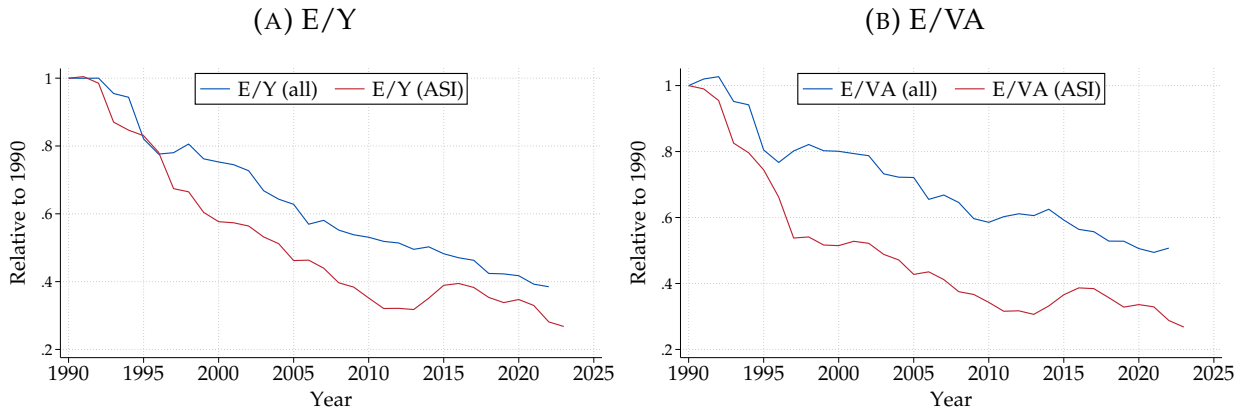
**Indian national accounts and KLEMS.** For macroeconomic series, we use the Indian national accounts. We use the detailed input-output matrix (130×130 distinct commodities) for the year 2008. For some manufacturing-specific series, we use the KLEMS database produced by the Reserve Bank of India (link).

**World Bank Enterprise Survey.** We exploit the "Green Economy" module of the World Bank Enterprise Survey administered in India in 2022 (link). The World Bank Enterprise Survey is a firm-level survey of a representative sample of an economy's private sector. The Enterprise Surveys are conducted across all geographic regions and cover small, medium, and large companies, and are representative of firms in the non-agricultural formal private economy.

In 2022, as part of the "Green Economy" module, the WBES asks sampled firms about adoption of four major types of energy efficiency investments over the past 3 years: heating and cooling improvements, energy management systems, machinery and equipment upgrades, and lighting systems upgrades. We obtain these variables, as well as the value of sales, and a two-digit industry, for 9,376 firms.

**UNIDO Technology Compendia.** The United Nations Industrial Development Organization (UNIDO), in collaboration with the Indian Bureau of Energy Efficiency (BEE),

FIGURE SM.3: Energy Intensity: ASI (Formal Manufacturing) vs. Alternative Sources (All Manufacturing)




*Notes:* The figure compares energy intensity in ASI (covering the formal manufacturing sector) and in other sources (covering the whole manufacturing sector). Panel A plots energy use relative to gross output, and Panel B plots energy use relative to value added. All series are normalized to 1 in 1990. The aggregate manufacturing series combines IEA manufacturing energy use with output and value-added measures from India KLEMS.

undertook a project titled "Promoting energy efficiency and renewable energy in selected MSME clusters in India" over 2020-2022 (link). The project aimed to promote energy efficiency and renewable energy technologies in process applications in energy-intensive industrial clusters, comprising micro, small and medium enterprises. In this context, technology compendia were produced to provide a list of available energy-efficient and renewable energy technologies. There is a compendium for each industry×cluster where the project was implemented (e.g., ceramics in Morbi). Each compendium details a menu of technologies available to firms in this industry×cluster. For each technology, the document details the baseline scenario, the energy-efficient alternative, and provides a cost-benefit analysis. Some cost-benefit analyses are estimates for the typical firm in the industry×cluster, while others rely on data from actual implementations.

We systematically collect all the data on technology menus for a number of industries for which sufficiently many clusters are available:

- Foundries: Faridabad, Belgaum, Howrah, Agra, Ahmedabad, Rajkot
- Hand-tools: Jalandhar, Nagaur, Ludhiana
- Ceramics: Khurja, Morbi, Thangadh
- Dairies: Andhra Pradesh and Telangana, Gujarat, Haryana, Kerala, Madhya Pradesh, Maharashtra, Odisha, Punjab, Rajasthan, Tamil Nadu

We collapse this data at the industry×technology level, averaging values when several firm observations are provided for a given industry×technology. We obtain 27 distinct technology data points for foundries, 14 for hand-tools, 30 for ceramics, and 29 for dairies.

From this data, we obtain the required investment $\Delta k_{\tau s}$ in INR, and associated annual energy savings $\Delta e_{\tau s}$ in gigajoules, for a number of alternative technologies $\tau$ available to a given industry $s$.

**U.S. Manufacturing Energy Consumption Survey.** We use data from the Manufacturing Energy Consumption Survey (MECS), a nationally representative survey conducted every four years by the U.S. Energy Information Administration (EIA) covering energy use in U.S. manufacturing establishments (link). We use the waves from 2002 to 2018. The survey provides information on energy consumption by industry and bins of firm size. It also records industry-level breakdowns of energy by end-uses.

**Wholesale Price Index (WPI).** The Wholesale Price Index (link), produced by India's Office of the Economic Adviser, provides monthly producer price indices for a broad set of commodities. We use these data to construct input price series.

**Centre for the Monitoring of the Indian Economy (CMIE) Prowess.** To construct the price of oil purchased by manufacturing firms, we use the Prowess database (link) which compiles information from the income statements and balance sheets of all publicly traded firms as well as a large number of private firms.

**India National Sample Survey of Unorganised Manufacturing Enterprises.** The NSS is a nationally representative survey conducted by the National Sample Survey Office (NSSO) covering unincorporated manufacturing enterprises in India's informal sector. The survey collects information on output, inputs, employment, assets, and other operational characteristics of manufacturing establishments not covered by the ASI. We use the waves 2000, 2005, 2010, and 2015.

### SM.1.3 Aggregate output, materials, capital, and energy price indices

The price index for manufacturing output is obtained from the national accounts.

We construct a materials price index by combining: (i) WPI product-level price indices at the level of the WPI sub-headings (30 indices), and (ii) the share of individual products within total material inputs, obtained from the ASI data over 1998-2017 and extrapolated for other years.

We construct a capital price index by combining: (i) Penn World Tables capital deflators for India for four capital types (structures, machinery, transport equipment, and a residual other asset group), and (ii) the share of each type of capital within the total capital stock, obtained from the ASI data over 1998-2017 and extrapolated for other years.

We construct the aggregate energy price index as a weighted average of the price of electricity, oil, coal, gas, and other fuels, using the fuel shares observed in ASI over 1998-2017 and extrapolated for other years.[5] For each fuel, we use the weighted median price observed in the ASI micro data for the years when this price is available, and extend these price series using corresponding WPI series pre-1997 and post-2018.[6] We assume that the

[5] For gas, reported only since 2008, we compute the share of gas in the sum of gas and other fuels after 2008, and apply this same share before 2008.

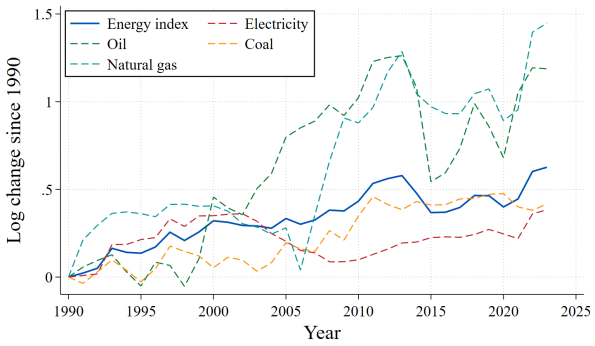
[6] The ASI micro data-based fuel price series, the WPI series, and micro data-based fuel price series constructed from the CMIE Prowess dataset closely coincide over overlapping years. For oil, we do not observe a price in the ASI micro data and use the aggregate oil price index described in SM.1.1.

relative price of the "other fuels" residual has remained constant over time.

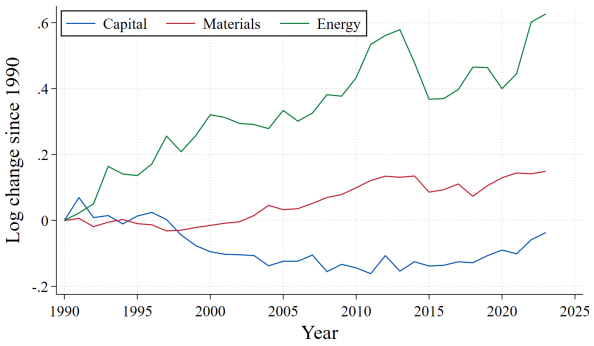
Figure SM.4(A) plots the price of individual energy inputs as well as the energy price index, all deflated by the output deflator and plotted in log change since 1990. Figure SM.4(B) plots the energy, material, and capital price index, relative to the output price index and measured in log change since 1990.

FIGURE SM.4: Relative Prices of Inputs

(A) Energy Inputs




(B) Capital, Materials, and Energy Inputs




*Notes:* The figure plots price indices for inputs used by Indian manufacturing establishments, expressed relative to the manufacturing output price index. Panel A plots indices for specific energy inputs (electricity, oil, coal, and natural gas) together with a composite energy price index. Panel B plots the capital, materials, and energy price indices. The construction of these indices is detailed in the text.

### SM.1.4 Construction of Firm-Level Productivity Growth in ASI

We measure the change in physical productivity as follows. Consider firm $i$ at time $t$ producing a single physical output:

$$y_{it} = z_{it}\, F\left(\phi_{1,it} x_{1,it},\ \ldots,\ \phi_{J,it} x_{J,it}\right), \tag{SM.1}$$

where $z_{it}$ is Hicks-neutral physical productivity, $x_{j,it}$ is the physical quantity of input $j$, $\phi_{j,it}$ is input-specific (input-augmenting) productivity, and $F(\cdot)$ is a differentiable production function. $F$ can have arbitrary curvature (Cobb–Douglas, CES, nested CES, translog, etc.), arbitrary returns to scale, and arbitrary (non-)homotheticity. Let $p_{it}$ denote the output price faced by the firm and $w_{j,it}$ the firm-specific price of input $j$. Firms may face arbitrary firm-specific input prices.

Define the output elasticity of physical input $x_{j,it}$: $\theta_{j,it} \equiv \frac{\partial \log y_{it}}{\partial \log x_{j,it}}$. Differentiating (SM.1):

$$\Delta \log y_{it} = \Delta \log z_{it} + \sum_{j=1}^{J} \theta_{j,it} \Delta \log x_{j,it} + \sum_{j=1}^{J} \theta_{j,it} \Delta \log \phi_{j,it} \tag{SM.2}$$

Define measured physical TFPQ growth as:

$$\Delta \widehat{\log TFPQ_{it}} \equiv \Delta \log y_{it} - \sum_{j=1}^{J} \widehat{\theta}_{j,it} \Delta \log x_{j,it}. \tag{SM.3}$$

with $\widehat{\theta}_{j,it}$ an estimate of the output elasticity. If $\widehat{\theta}_{j,it} = \theta_{j,it}$, then:

$$\Delta \widehat{\log TFPQ}_{it} = \Delta \log z_{it} + \sum_j \theta_{j,it} \Delta \log \phi_{j,it}, \tag{SM.4}$$

so measured TFPQ growth captures both Hicks-neutral and input-augmenting productivity growth. Note that physical productivity measurement is made possible by the availability of price and quantity data separately, for both outputs and inputs.

*Measurement of output elasticities.* Under the assumption that firms are price-takers on the input and output markets, and for any flexible input $j$ chosen at an interior optimum of the static cost minimization problem, we can show that:

$$\theta_{j,it} = \frac{w_{j,it} x_{j,it}}{p_{it} y_{it}} \equiv s_{j,it} \tag{SM.5}$$

That is, output elasticities are equal to output cost shares. In the presence of markups $\mu_{it}$ and non-priced input wedges $\tau_{j,it}$, we instead have

$$\theta_{j,it} = \mu_{it}(1 + \tau_{j,it})\, s_{j,it} \tag{SM.6}$$

We consider four inputs: raw materials, energy, capital, and labor. We measure output elasticities in two different ways. Option 1 uses average revenue shares by industry×year×size decile, and option 2 uses firm-level revenue shares. Firm-level shares are conceptually preferable under heterogeneity (e.g., if the production function is non-homothetic) but may reflect firm-specific wedges (markups, adjustment costs). In addition, firm-level shares are unreliable for quasi-fixed inputs such as capital, so we always use industry-level shares for capital. On the other hand, industry×year×size decile-level shares average out input wedges and statistical noise, but may not capture the true output elasticity for any individual firm.

For both firm-level and industry-level shares, we implement the following correction for the markup-wedge multiplicative factor. We exploit the fact that $\sum_j \theta_{j,it} = \text{RTS}_{it}$, the returns to scale for firm $i$. In Section 5, where we calibrate our quantitative model, we estimate average returns to scale $\widehat{\text{RTS}}$ by estimating the elasticity of the cost function with respect to output (Table SM.18). Assuming a constant combined distortion $\mu(1+\tau)$, we can back out $\mu(1+\tau)$ from the relationship $\mu(1+\tau) = \widehat{\text{RTS}}/\mathbb{E}[\sum_j s_{j,it}]$.

*Validation of measured productivity growth.* In Table SM.3, we show that, as predicted by theory, our measure of physical productivity growth predicts a decline in firm-level prices, an increase in firm-level quantities, and an increase in firm-level revenues.

TABLE SM.3: Validation of Measured Productivity Growth

| | $\Delta^5 \log(p)$ | | $\Delta^5 \log(y)$ | | $\Delta^5 \log(py)$ | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| $\Delta^5$ log(tfpq) | -0.814*** | -0.847*** | 0.847*** | 0.880*** | 0.045*** | 0.046*** |
| | (0.014) | (0.014) | (0.014) | (0.013) | (0.004) | (0.004) |
| Year × Ind. × Cohort FE | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Def. TFPQ | (1) | (2) | (1) | (2) | (1) | (2) |
| R-squared | 0.678 | 0.695 | 0.681 | 0.696 | 0.288 | 0.288 |
| Observations | 114,081 | 114,081 | 114,081 | 114,081 | 114,081 | 114,081 |

*Notes:* This table presents the results of regressions of firm-level variables on firm-level physical productivity growth, defined in (SM.3). The outcome variable is the change in the firm-level price index (columns 1-2), the change in the firm-level quantity index (columns 3-4), and the change in firm-level revenues (columns 5-6). "Def. TFPQ" indicates the chosen measure for output elasticities, as detailed in the text. Regressions are weighted by firm total cost (adjusted for the sampling weight). Standard errors are clustered at the firm and main product×year level. ***, **, and * indicate significance at the 1%, 5% and 10% levels, respectively.

# SM.2. Descriptive Statistics

## SM.2.1 Descriptive Statistics: Aggregate Data

### SM.2.1.1 Energy Intensity

To analyze the sources of change in aggregate energy intensity, we apply a standard within-across decomposition approach. Denote $E_{jt}$ energy use in sector $j$ at time $t$, $Y_{jt}$ sectoral output, and $s_{jt} = Y_{jt}/Y_t$ is the sectoral output share. The change in aggregate energy intensity between two periods can be decomposed as follows:

$$\Delta \frac{E_t}{Y_t} = \underbrace{\sum_{j \in J} s_{j,t-1} \Delta \left( \frac{E_{jt}}{Y_{jt}} \right) + \Delta s_{jt} \Delta \left( \frac{E_{jt}}{Y_{jt}} \right)}_{\text{Within + Cross}} + \underbrace{\sum_{j \in J} \Delta s_{jt} \frac{E_{j,t-1}}{Y_{j,t-1}}}_{\text{Reallocation}} \tag{SM.7}$$

The reallocation term reflects changes in aggregate energy intensity attributable to shifts in the composition of output across sectors, keeping sectoral energy intensities fixed. This term is orthogonal to our theory. The within-sector term captures changes in sectoral energy intensity, keeping sectoral shares constant. The cross term captures the fact that faster-growing industries may experience a larger decline in sectoral energy intensity. Our theory is consistent with both of these terms.

Estimating this decomposition requires matching sectoral energy use and sectoral output. We proceed as follows. Sectoral energy use comes from the IEA. Sectoral output shares come from the UN. We match the sectors in these two datasets as described in Table SM.4.

The sectors are: Agriculture and Fishing, Industry (mining, manufacturing, construction), Transport and Communications, and Other Services. Transport energy use in the IEA data corresponds to both the household sector (energy use that cannot be linked to value added in the transportation sector, like personal car use) and the production sector. We estimate energy used by the transportation sector as follows. For road transport, we exclude fuel use by passenger cars, two-wheelers and three-wheelers. To obtain their share of total road fuel use, we rely on the Petroleum Planning & Analysis Cell (PPAC) (2014) report for India and on the world aggregates in the IEA (2025) report for other countries. All non-road uses (rail, pipeline, domestic aviation, world aviation bunkers, world marine bunkers, domestic navigation, n.e.c.) are treated as production of goods and services.

**India.** The results for India are presented in Figure SM.5. Reallocation accounts for little of the decline. The manufacturing share remains roughly constant, and the decline in moderately energy-intensive agriculture is offset by a rise in both low-energy services and highly energy-intensive transport.

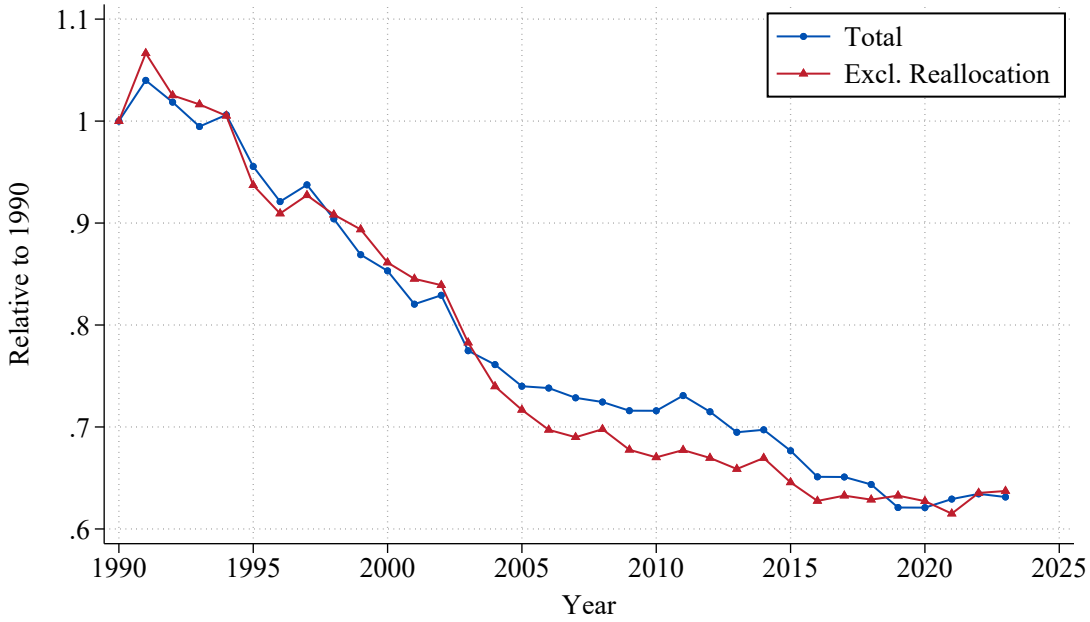
**Global.** We document the same decomposition at the global level in Figure SM.6. The within cell $j$ is defined as a country in Panel A, a sector in Panel B, and a country$\times$sector pair in Panel C. That is, in this last panel, the red line excludes reallocation across both

TABLE SM.4: Sectors for the Within-Across Decomposition

| Sector | IEA | UN |
|---|---|---|
| Agriculture, forestry and fishing | Agriculture, forestry (01–02) | Agriculture, forestry and fishing (01–03) |
| Industry (mining, manufacturing, construction) | Mining (07–08–099), Manufacturing (10–32, except 19), Construction (41–43) | Mining (05–09), Manufacturing (10–33), Construction (41–43), Utilities (35) |
| Transport and communications | Transport* (49–51) | Transport, storage and communications (49–53) |
| Services | Commercial and Public Services (33, 36–39, 45–47, 52–53, 55–56, 58–66, 68–75, 77–82, 84–88, 90–96, 99) | Wholesale and Retail Trade, Restaurants and Hotels (45–47), Other Activities (58–66, 68–75, 84–88, 90–96, 97–98). |

*Notes:* This table shows the sectors used in the within-across sectors decomposition of energy intensity. Sector codes correspond to ISIC Rev. 4. For transport energy use, please refer to explanations in the main text. Utilities energy use is excluded because the utility sector is not an end-user of energy.

FIGURE SM.5: Energy Intensity in India




*Notes:* This figure displays the change in energy intensity in India. It is defined as energy used in the production of goods and services divided by GDP (measured in constant 2011 INR), and is normalized to 1 in 1990. The red line excludes the reallocation component (see equation (SM.7)).

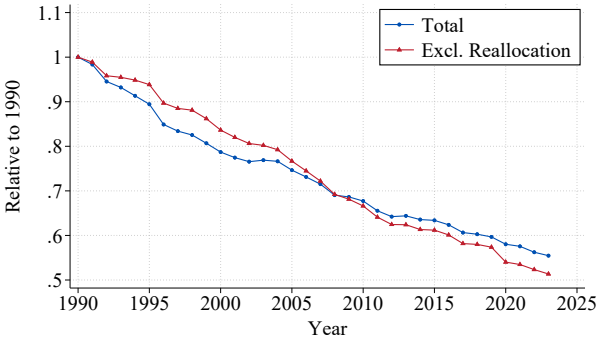
sectors and countries. In terms of sectoral shifts, the small decline in highly energy-intensive manufacturing and large decline in moderately energy-intensive agriculture are offset by a rise in both low-energy services and highly energy-intensive transport. In addition, reallocation of economic activity across countries tends to push up the aggregate energy intensity.

### SM.2.1.2 Other Descriptive Statistics on Aggregate Data

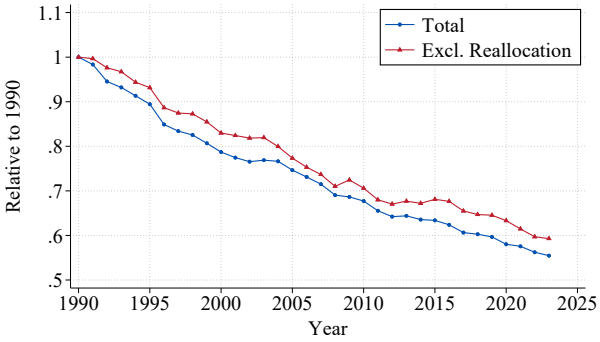
**Energy by End Use.** Table SM.5 uses data from US 2018 Manufacturing Energy Consumption Survey energy flowcharts to calculate the fraction of energy by end use across the manufacturing industries present in the ASI.

FIGURE SM.6: Energy Intensity in the World

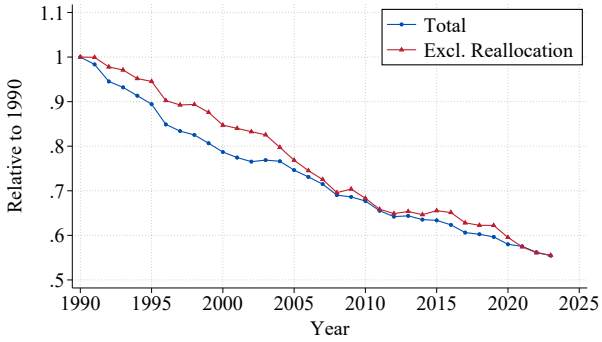
(A) Country Decomposition




(B) Sector Decomposition




(C) Country and Sector Decomposition




*Notes:* This figure displays the change in energy intensity at the global level. It is defined as energy used in the production of goods and services divided by GDP (measured in constant 2021 USD PPP), and is normalized to 1 in 1990. The red line excludes the reallocation component (see equation (SM.7)). The within cell $j$ is defined as a country in Panel A, a sector in Panel B, and a country×sector pair in Panel C.

TABLE SM.5: Energy by End Use

| | Process | | | | Non-Process |
|---|---|---|---|---|---|
| | Heat: fired | Heat: steam | Mechanical | Other | |
| All Manufacturing | 0.50 | 0.09 | 0.18 | 0.12 | 0.11 |
| Alumina and Aluminum | 0.46 | 0.01 | 0.11 | 0.31 | 0.10 |
| Automobile and Light Duty Motor Vehicle | 0.36 | 0.03 | 0.21 | 0.10 | 0.30 |
| Cement | 0.75 | 0.00 | 0.21 | 0.03 | 0.01 |
| Chemicals | 0.31 | 0.23 | 0.19 | 0.17 | 0.09 |
| Computers, Electronics, and Electrical Equipment | 0.20 | 0.01 | 0.17 | 0.17 | 0.45 |
| Fabricated Metals | 0.38 | 0.01 | 0.22 | 0.07 | 0.32 |
| Food and Beverage | 0.19 | 0.24 | 0.18 | 0.17 | 0.21 |
| Forest Products | 0.13 | 0.44 | 0.24 | 0.06 | 0.13 |
| Foundries | 0.62 | 0.00 | 0.15 | 0.05 | 0.18 |
| Glass and Glass Products | 0.79 | 0.00 | 0.11 | 0.03 | 0.08 |
| Iron and Steel | 0.69 | 0.05 | 0.09 | 0.09 | 0.08 |
| Machinery | 0.18 | 0.01 | 0.34 | 0.06 | 0.42 |
| Plastic Material and Resins | 0.19 | 0.29 | 0.20 | 0.24 | 0.08 |
| Plastics and Rubber Products | 0.21 | 0.05 | 0.36 | 0.10 | 0.27 |
| Textiles | 0.23 | 0.11 | 0.38 | 0.05 | 0.22 |
| Transportation Equipment | 0.24 | 0.02 | 0.24 | 0.09 | 0.41 |

*Notes:* The table reports end-use energy shares by industry, obtained from the US 2018 Manufacturing Energy Consumption Survey energy flowcharts. The "All Manufacturing" row aggregates industries using Indian ASI industry-level energy use as weights. Industries without an ASI match are excluded. Row shares sum to one by construction.

### SM.2.2 Descriptive Statistics: Micro Data

**Summary Statistics of Key Variables.** Table SM.6 produces summary statistics of the key variables used in the analysis. This table reveals a large dispersion in energy expenditure shares, even when residualized on industry×year fixed effects. For instance, for our baseline outcome variable $\log\left(\Omega^{e,eml}/(1-\Omega^{e,eml})\right)$, the 10th and 90th percentiles of the within-industry distribution are $-4.33$ and $-1.62$, which expressed in levels correspond to 0.013 and 0.197, respectively.

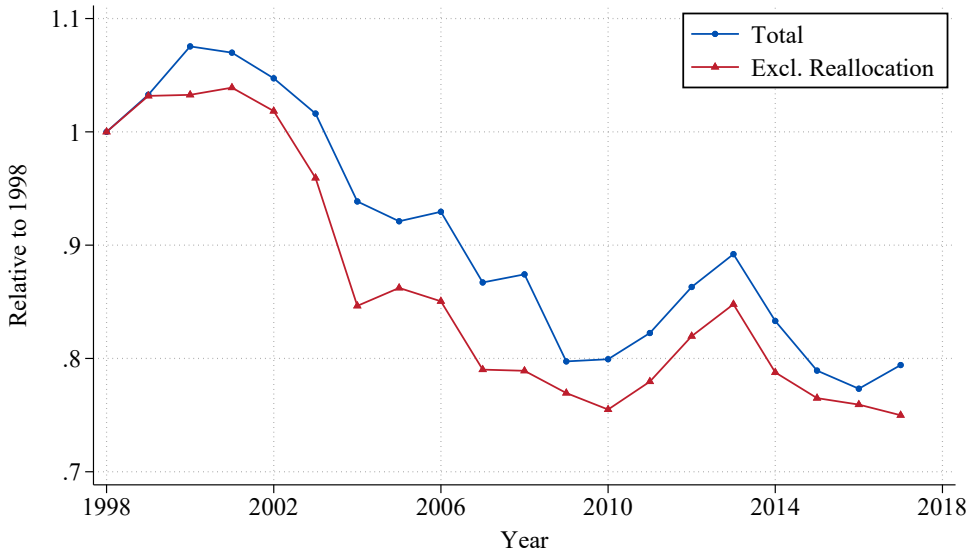
TABLE SM.6: Summary Statistics of Key Variables

| | Raw | | | | | Residualized on industry × year fixed effects | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | mean | sd | p10 | p50 | p90 | mean | sd | p10 | p50 | p90 |
| log(sales) | 17.1 | 2.03 | 14.5 | 16.9 | 19.7 | 17.1 | 1.80 | 14.8 | 17.0 | 19.4 |
| $\Omega^e$ | 0.085 | 0.11 | 0.0086 | 0.041 | 0.22 | 0.085 | 0.079 | 0.021 | 0.071 | 0.17 |
| $\Omega^m$ | 0.67 | 0.23 | 0.32 | 0.73 | 0.92 | 0.67 | 0.18 | 0.43 | 0.71 | 0.86 |
| $\Omega^k$ | 0.092 | 0.11 | 0.012 | 0.056 | 0.22 | 0.092 | 0.10 | 0.013 | 0.061 | 0.21 |
| $\Omega^l$ | 0.15 | 0.14 | 0.021 | 0.11 | 0.35 | 0.15 | 0.11 | 0.040 | 0.13 | 0.29 |
| $\log(\Omega^{e,em}/(1-\Omega^{e,em}))$ | -2.66 | 1.60 | -4.52 | -2.82 | -0.63 | -2.66 | 1.22 | -4.14 | -2.66 | -1.23 |
| $\log(\Omega^{e,eml}/(1-\Omega^{e,eml}))$ | -2.95 | 1.41 | -4.67 | -3.02 | -1.05 | -2.95 | 1.10 | -4.33 | -2.91 | -1.62 |
| $\log(\Omega^{e,emlk}/(1-\Omega^{e,emlk}))$ | -3.07 | 1.35 | -4.73 | -3.14 | -1.27 | -3.07 | 1.06 | -4.41 | -3.02 | -1.81 |
| Observations | 690775 | | | | | 690771 | | | | |

*Notes:* This table produces summary statistics of the key variables used in the analysis. The left panel uses raw variables, while the right panel uses variables residualized on industry×year fixed effects, normalized to the same mean. All statistics are weighted by sampling weights.

**Manufacturing Energy Share in ASI: Excluding Reallocation Across Sectors.** Figure SM.7 shows the change in the energy expenditure share in ASI excluding the reallocation component. Reallocation is defined as shifting sectoral shares across 3-digit industries. This decomposition requires the micro data, hence we report results for 1998-2017.

FIGURE SM.7: Manufacturing Energy Share: Excluding Reallocation Across Industries




*Notes:* This figure displays the change in the energy expenditure share in ASI. It is normalized to 1 in 1998. The red line excludes the reallocation component (see equation (SM.7)). The sectors are 3-digit industries.

**Analysis of Continuing Firms.** One potential concern with our estimation of $\eta$ is that our identification strategy exploiting within-firm changes restricts the sample to firms present in two consecutive periods (i.e., continuing firms). Here, we show that changes in the energy share of continuing firms are representative of the dynamics of the energy share in the whole sample.

Let $CN$ denote continuing firms, $EN$ denote entrants, and $EX$ denote exiting firms. Consider the energy expenditure share with a given denominator; here we use total costs as a denominator. The aggregate energy share $\Omega_t^e$ can be written as:

$$\Omega_t^e = \sum_i \omega_{it} \Omega_{it}^e \tag{SM.8}$$

where $\omega_{it} = \frac{X_{it}}{X_t}$ are cost weights.

First, we show that in levels continuing firms are representative of the broader sample. Write:

$$\Omega_t^e = \omega_t^{CN} \Omega_t^{e,CN} + \omega_t^{EN} \Omega_t^{e,EN} \tag{SM.9}$$

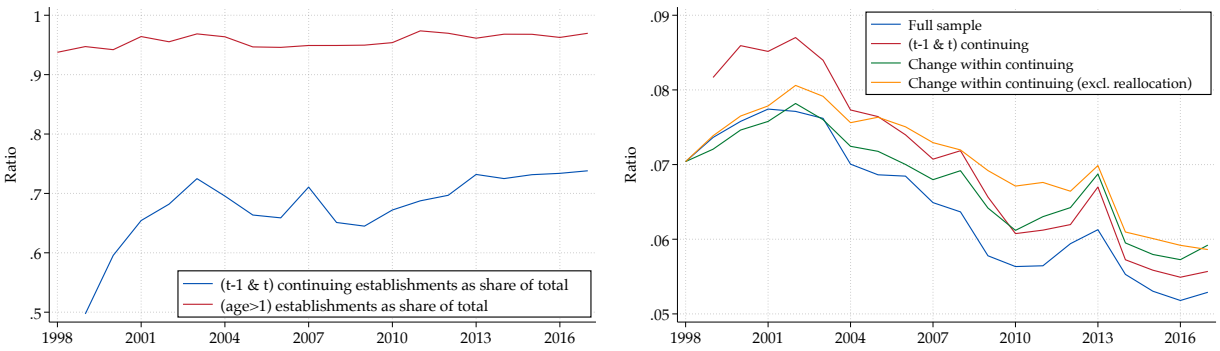
where $\Omega_t^{e,CN} = \sum_{i \in CN} \frac{\omega_{it}}{\omega_t^{CN}} \Omega_{it}^e$ is the energy share within continuing firms, and we use similar notations for other groups of firms.

The left panel of Figure SM.8 plots the cost share of continuing firms $\omega_t^{CN}$ in each time period. Continuing firms account for around 50% of the sample in 1999 and reach approximately 75% by the end of the sample. This ratio is considerably below 1 mostly because of entry *into the sample* as opposed to new establishment creations. To show this, we exploit the fact that the ASI reports the year of creation of each establishment and

FIGURE SM.8: Energy Share: Representativeness of Continuing Firms

(A) Share of continuing establishments

(B) Energy share and role of continuing establishments




*Notes:* Panel A reports the ratio of total costs by continuing establishments (defined as establishments present in the sample in year $t-1$ and $t$) over total costs for all establishments (blue line) and the ratio of total costs by establishments created before year $t-1$ over total costs for all establishments (red line). Panel B reports three versions of the energy expenditure share, defined in the text. The blue line is $\Omega_t^e$. The red line is $\Omega_t^{e,CN}$. The green line is $\tilde{\Omega}_t^e$.

reports the share of establishments created before year $t-1$. This share is significantly higher, averaging 96%. This also suggests that the increase in $\omega_t^{CN}$ over time is mostly attributable to an increase in the sampling rate, as opposed to an increase in the share of incumbents in total costs.

The right panel of Figure SM.8 plots $\Omega_t^e$ (blue line) and $\Omega_t^{e,CN}$ (red line) over time. The figure shows that energy intensity among continuing firms closely tracks aggregate energy intensity. Note that here the definition of incumbents changes in each period, hence two consecutive points do not include the same set of firms.

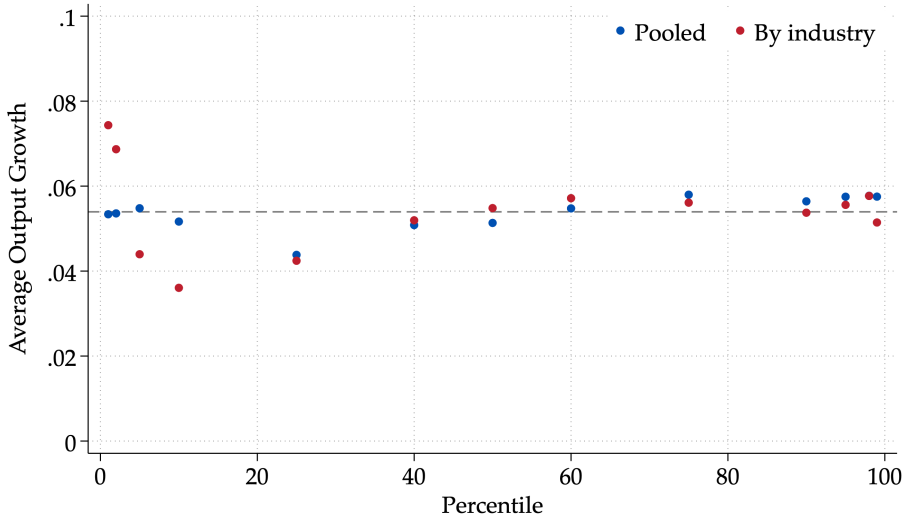
Second, we show that changes in energy intensity among continuing firms account for most of the decline in the aggregate energy intensity. Note that,

$$\begin{aligned}\Omega_t^e - \Omega_{t-1}^e = & \underbrace{\Omega_t^{e,CN} - \Omega_{t-1}^{e,CN}}_{\text{Change among incumbents}} + \underbrace{\omega_t^{EN}(\Omega_t^{e,EN} - \Omega_t^{e,CN}) - \omega_{t-1}^{EX}(\Omega_{t-1}^{e,EX} - \Omega_{t-1}^{e,CN})}_{\text{Contribution of entry and exit}} \\ = & \underbrace{\underbrace{\sum_{i\in CN} \tilde{\omega}_{it-1}^{CN}\Delta\Omega_{it}^e + \sum_{i\in CN}\Delta\tilde{\omega}_{it}^{CN}\Delta\Omega_{it}^e}_{\text{Within + cross}} + \underbrace{\sum_{i\in CN}\Delta\tilde{\omega}_{it}^{CN}\Omega_{it-1}^e}_{\text{Reallocation}}}_{\text{Change among incumbents}} \\ & + \underbrace{\omega_t^{EN}(\Omega_t^{e,EN} - \Omega_t^{e,CN}) - \omega_{t-1}^{EX}(\Omega_{t-1}^{e,EX} - \Omega_{t-1}^{e,CN})}_{\text{Contribution of entry and exit}}\end{aligned}$$

where the weights $\tilde{\omega}_{it}^{CN}$ correspond to the share of firm $i$ in the total costs of continuing firms. To assess the contribution of incumbents to the overall dynamics of the energy share, we define:

$$\tilde{\Omega}_t^e = \tilde{\Omega}_{t-1}^e + \Omega_t^{e,CN} - \Omega_{t-1}^{e,CN} \text{ for } t > 1998,\ \tilde{\Omega}_{1998}^e = \Omega_{1998}^e$$

FIGURE SM.9: Output Growth Along the Firm Size Distribution




*Notes:* This figure displays output growth along percentiles of the firm size distribution. The construction of the figure is detailed in the text.

$\tilde{\Omega}_t^e$ is plotted in the green line in Panel B of Figure SM.8. $\tilde{\Omega}_t^e$ closely tracks $\Omega_t^e$. That is, the dynamics of the energy share among continuing firms for each pair of consecutive dates $(t-1, t)$ closely tracks the change in the aggregate energy share. This suggests that the contribution of entry and exit (the residual) is minimal. The yellow line in Panel B of Figure SM.8 additionally shows that the change within continuing firms is not driven by reallocation of economy activity across firms with ex-ante higher or smaller energy cost shares.

**Output Growth Along the Firm Size Distribution.** To investigate how output growth varies along the firm size distribution, we proceed as follows. In a first exercise, in each year $t$ and for each percentile $p$, we compute $Sales(p)_t$ as the $p$-th percentile of the distribution of real output (sales value) in year $t$. We define percentile $p$ growth in year $t$ as $\log(Sales(p)_t) - \log(Sales(p)_{t-1})$, and average these values over time for each percentile $p$. In a second exercise, we perform this procedure by 3-digit industries and aggregate using fixed industry sales weights to remove any effect driven by changes in the composition of industries.

Figure SM.9 shows that growth in the percentiles of the distribution is roughly even, except for some slight deviation at the bottom of the distribution. In the upper half, annual sales growth for the very largest firms is approximately equal to growth at the median.

TABLE SM.7: Alternative Outcome Variables

**Panel A:** Logs

| | $\Delta^5$ log(outcome) | | | | |
|---|---|---|---|---|---|
| | $\Omega^{e,em}$ | $\Omega^{e,eml}$ | $\Omega^{e,emlk}$ | $\frac{w^e e}{py}$ | $\frac{w^e e}{va}$ |
| | (1) | (2) | (3) | (4) | (5) |
| $\Delta^5$ log(sales) | -0.443*** | -0.397*** | -0.334*** | -0.532*** | -0.534*** |
| | (0.038) | (0.037) | (0.038) | (0.039) | (0.066) |
| Year × Ind. × Cohort FE | ✓ | ✓ | ✓ | ✓ | ✓ |
| Observations | 184,011 | 184,011 | 183,727 | 184,011 | 159,001 |

**Panel B:** Levels

| | $\Delta^5$ (outcome) | | | | |
|---|---|---|---|---|---|
| | $\Omega^{e,em}$ | $\Omega^{e,eml}$ | $\Omega^{e,emlk}$ | $\frac{w^e e}{py}$ | $\frac{w^e e}{va}$ |
| | (1) | (2) | (3) | (4) | (5) |
| $\Delta^5$ log(sales) | -0.029*** | -0.019*** | -0.013*** | -0.033*** | -0.153*** |
| | (0.004) | (0.003) | (0.003) | (0.005) | (0.040) |
| Year × Ind. × Cohort FE | ✓ | ✓ | ✓ | ✓ | ✓ |
| Observations | 184,496 | 184,498 | 184,206 | 184,498 | 159,399 |

*Notes:* This table presents IV results for specification (3), with the instrument defined in equation (4). The dependent variable is indicated in the column heading. Regressions are weighted by firm total cost (adjusted for the sampling weight). Standard errors are clustered at the firm and main product×year level. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

# SM.3. FIRM SIZE AND ENERGY DEMAND: ADDITIONAL RESULTS

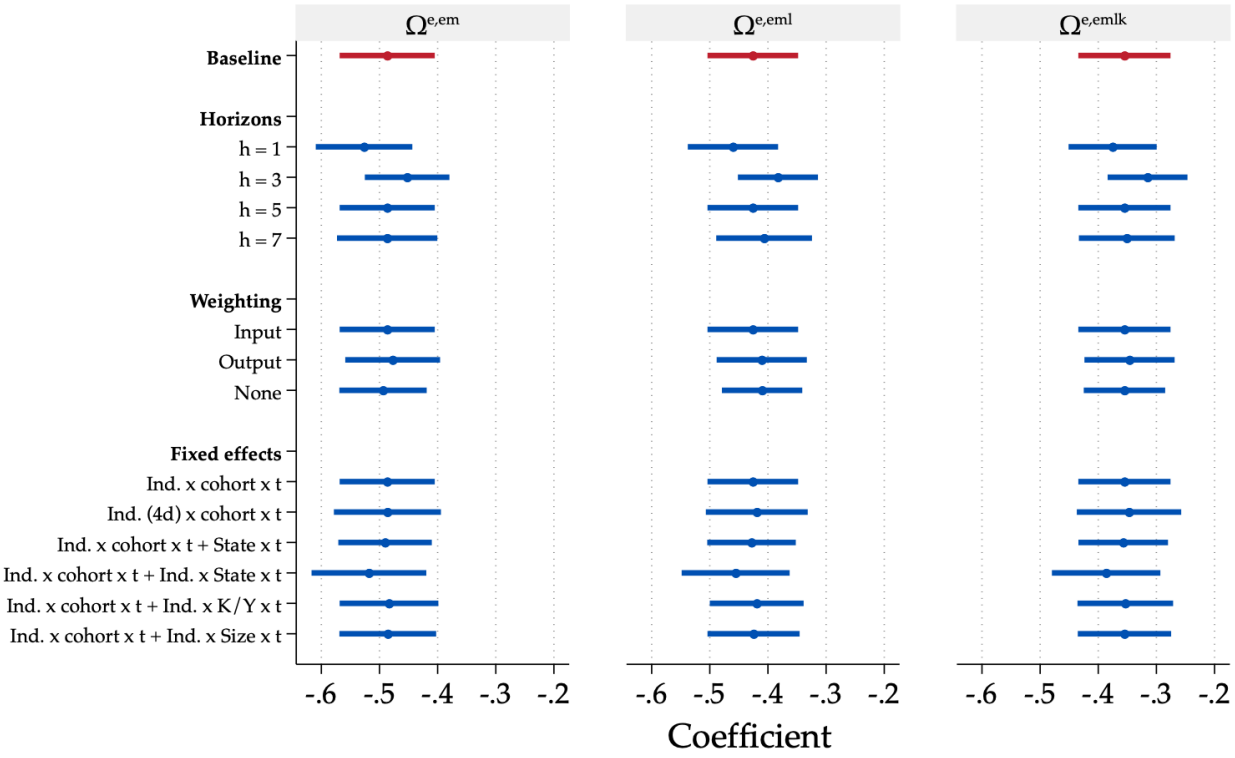
## SM.3.1 Robustness to Choice of Specification

Table SM.7 shows our main specification results for different outcome variables. Figure SM.10 shows the results of our baseline specification at different horizons, for different observation weights, and with different fixed effect structures.

Figure SM.11 shows the results of our baseline specification for different definitions of our shift-share instrument, defined in equation (4). To alleviate concerns related to a mechanical bias when only a few firms contribute to each shifter $\Delta^h \log \text{sales}_{jt}$ (Borusyak et al. 2022), we drop products $j$ with fewer than $N$ firms in the product×year cell for $N \in \{5, 10, 20\}$, or construct $\Delta^h \log \text{sales}_{jt}$ excluding any firms with market share above $x$% in the production of product $j$ for $x \in \{5, 10, 20\}$. All variants yield highly similar results, suggesting that our strategy is relatively immune to this concern.

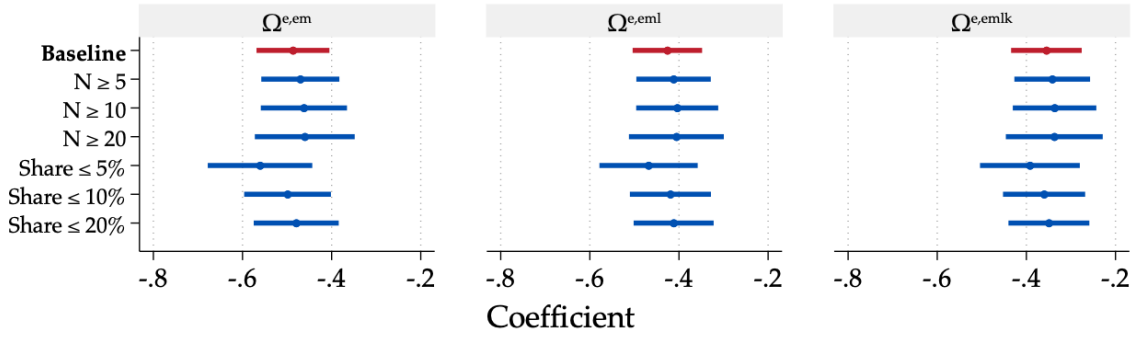
Table SM.8 shows the results of our baseline specification by bins of ex-ante size (as defined by sales). Table SM.9 shows the results of our baseline specification by changes in capacity utilization (defined as the share of days worked during the year) and by positive vs. negative sales growth.

FIGURE SM.10: Robustness to Choice of Specification




*Notes:* This figure presents IV results for specification (3), with the instrument defined in equation (4). The dependent variable is the relative energy expenditure share, $\Delta^h \log(\Omega^e/(1-\Omega^e))$, where the definition of $\Omega^e$ is indicated in the panel headings. Unless otherwise specified, "Ind." fixed effects refer to 3-digit industries. "$\times$ K/Y" indicates that fixed effects are interacted with within-industry$\times$time quintiles of the time-$t$ capital-to-output ratio. "$\times$ Size" indicates that fixed effects are interacted with within-industry$\times$time quintiles of time-$t$ size. Regressions are weighted by firm total cost (adjusted for the sampling weight). Standard errors are clustered at the firm and main product$\times$year level. The dots are point estimates and bars are 95% confidence intervals.

FIGURE SM.11: Robustness to Instrument Definition




*Notes:* This figure presents IV results for specification (3), with the baseline instrument defined in equation (4). Each line refers to a different definition of the instrument (4). In lines $N \geq x$, we restrict the set of shifters to product-year cells with at least $x \in \{5, 10, 20\}$ observations. In lines Share $\leq x$%, the construction of the shifter excludes any firms with product-level market share above $x$% for $x \in \{5, 10, 20\}$. The dependent variable is the relative energy expenditure share, $\Delta^5 \log(\Omega^e/(1-\Omega^e))$, where the definition of $\Omega^e$ is indicated in the panel headings. Regressions are weighted by firm total cost (adjusted for the sampling weight). Standard errors are clustered at the firm and main product×year level. The dots are point estimates and bars are 95% confidence intervals.

TABLE SM.8: Heterogeneity: By Ex-Ante Sales

| | $\Delta^5 \log(\Omega/(1-\Omega))$ | | | | | |
|---|---|---|---|---|---|---|
| | By time-$t$ sales | | | | | |
| | $\Omega^{e,em}$ | | $\Omega^{e,eml}$ | | $\Omega^{e,emlk}$ | |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| $\Delta^5$ log(sales) × Small | -0.406*** | -0.537*** | -0.308*** | -0.426** | -0.278*** | -0.434** |
| | (0.111) | (0.183) | (0.101) | (0.171) | (0.106) | (0.182) |
| $\Delta^5$ log(sales) × Medium | -0.565*** | -0.584*** | -0.482*** | -0.472*** | -0.426*** | -0.404*** |
| | (0.064) | (0.104) | (0.061) | (0.098) | (0.061) | (0.098) |
| $\Delta^5$ log(sales) × Large | -0.463*** | -0.469*** | -0.409*** | -0.419*** | -0.355*** | -0.343*** |
| | (0.036) | (0.051) | (0.034) | (0.048) | (0.033) | (0.049) |
| Year × Ind. × Cohort × Bin FE | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Weighted | | ✓ | | ✓ | | ✓ |
| Observations | 179,915 | 179,260 | 179,950 | 179,295 | 179,015 | 179,015 |

*Notes:* This table presents IV results for specification (3), with the instrument defined in equation (4), by bins of time-$t$ sales. Bins are within-industry×year terciles of firm sales. The dependent variable is the relative energy expenditure share, $\Delta^5 \log(\Omega^e/(1-\Omega^e))$, where the definition of $\Omega^e$ is is indicated in the column heading. All columns use the sampling weight and the line "Weighted" indicates that regressions are additionally weighted by firm total cost. Standard errors are clustered at the firm and main product×year level. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

TABLE SM.9: Heterogeneity: By Change in Days Worked and Output Sign

| | $\Delta^5 \log(\Omega/(1-\Omega))$ | | | | | |
|---|---|---|---|---|---|---|
| | By change in % of days worked | | | By $\Delta^5 \log(\text{sales})$ sign | | |
| | $\Omega^{e,em}$ | $\Omega^{e,eml}$ | $\Omega^{e,emlk}$ | $\Omega^{e,em}$ | $\Omega^{e,eml}$ | $\Omega^{e,emlk}$ |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| $\Delta^5$ log(sales) × Negative | -0.366*** | -0.285*** | -0.179*** | | | |
| | (0.015) | (0.014) | (0.015) | | | |
| $\Delta^5$ log(sales) × Close-zero | -0.410*** | -0.339*** | -0.275*** | | | |
| | (0.011) | (0.010) | (0.011) | | | |
| $\Delta^5$ log(sales) × Positive | -0.337*** | -0.268*** | -0.159*** | | | |
| | (0.017) | (0.017) | (0.017) | | | |
| $\Delta^5$ log(sales) × Negative | | | | -0.403*** | -0.287*** | -0.163*** |
| | | | | (0.016) | (0.015) | (0.016) |
| $\Delta^5$ log(sales) × Non-negative | | | | -0.323*** | -0.273*** | -0.194*** |
| | | | | (0.015) | (0.014) | (0.015) |
| Year × Ind. × Cohort × Group FE | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Observations | 180,279 | 180,314 | 180,036 | 182,955 | 182,991 | 182,722 |

*Notes:* This table presents OLS results for specification (3), by subsamples. Columns (1)–(3) show results when firms are sorted by the change in the share of days worked. Columns (4)–(6) show results by sales growth sign. Since the variable used to define subsamples is a $t$ to $t+5$ change, we present only OLS results. The dependent variable is the relative energy expenditure share, $\Delta^5 \log(\Omega^e/(1-\Omega^e))$, where the definition of $\Omega^e$ is indicated in the column heading. Regressions are weighted by firm total cost (adjusted for the sampling weight). Standard errors are clustered at the firm and main product×year level. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

### SM.3.2 Scale Elasticity of Energy Demand and India's Energy Policies

One potential explanation for our results is that firms become subject to more stringent energy-related policies as they grow, which could depress their energy expenditure share. In that case, our findings would reflect size-dependent regulation rather than a feature of production technologies. To address this concern, we compile a comprehensive list of energy-related policies in India and show that our results are robust to excluding firms exposed to these policies.

The background for India's energy policy is the Energy Conservation Act enacted in 2001 with the goal of reducing the energy intensity of the Indian economy. The Bureau of Energy Efficiency (BEE) was set up as the statutory body to facilitate the implementation of the Energy Conservation Act. The Act provides a regulatory mandate for standards and labeling of equipment and appliances, energy conservation codes for commercial buildings, and energy consumption goals for energy-intensive industries (IEA 2023). Several policies affecting manufacturing firms were implemented as part of this mandate.

*Perform, Achieve & Trade (PAT).* India's PAT scheme is a regulatory instrument designed to reduce specific energy consumption (energy use per unit of production) in energy-intensive industries (Bureau of Energy Efficiency 2026). The policy applies only to the largest establishments (called Designated Consumers) in a number of selected energy-intensive industries. Each designated consumer's energy consumption is assessed and an establishment-specific reduction target is set. The excess energy savings made by the designated consumers are converted into tradable certificates, to be bought by designated consumers not meeting their targets. Within India's energy policy framework, the PAT scheme is the central policy targeting the industrial sector. PAT is implemented in three-year cycles, each targeting specific sectors and designated consumers. The first three cycles overlap with our micro data sample, and all six overlap with our aggregate sample. Table SM.10 provides an overview of each PAT cycle. One notable fact is that the policy targets very few establishments.

To exclude observations exposed to the PAT scheme from our sample, we proceed in two ways. First, we exclude all firms likely to be considered Designated Consumers (option A). Designated Consumers are firms above a certain energy consumption threshold, specific to each sector. We use this threshold and the energy consumption data in ASI (both expressed in physical energy units) to construct an indicator for likely Designated Consumer. We obtain 1000 such unique firms, corresponding to 0.5% of the observations in our sample. Second, we exclude all industry$\times$year pairs covered by the PAT scheme, around 16,000 unique firms corresponding to 6% of our sample (option B).

*BEE MSME policy.* The Bureau of Energy Efficiency (BEE) implements energy-efficiency policies targeted at the Micro-, Small- and Medium-sized Enterprises sector (MSMEs). These policies are structured around industrial clusters, which are groups of establishments operating in the same industry in a given area. The first phase of this policy

TABLE SM.10: Coverage of the Perform, Achieve and Trade (PAT) Scheme

| Cycle | Duration | # Sectors | # Designated Consumers | Sectors Covered |
|---|---|---|---|---|
| I | 2012–15 | 8 | 478 | aluminum, cement, chlor-alkali, fertilizer, iron and steel, pulp and paper, textile, and thermal power |
| II | 2016–19 | 11 | 621 | aluminum, cement, chlor-alkali, fertilizer, iron and steel, pulp and paper, textile, thermal power, railways, refineries, and electricity distribution |
| III | 2017–20 | 6 | 116 | aluminum, cement, iron and steel, pulp and paper, textile, and thermal power |
| IV | 2018–22 | 8 | 109 | aluminum, cement, iron and steel, pulp and paper, textile, thermal power, petrochemicals, and commercial buildings (hotels) |
| V | 2019–22 | 8 | 110 | aluminum, cement, chlor-alkali, commercial buildings (hotels), iron and steel, pulp and paper, textile, and thermal power |
| VI | 2020–23 | 6 | 135 | cement, commercial buildings (hotels), iron and steel, petroleum refinery, pulp and paper, and textiles |

*Source:* Ministry of Power, Government of India (2022)

(Eleventh Plan, 2007-2012) primarily consisted of energy audits and technology gap assessments in 25 MSME clusters. A key output was the development of cluster-specific energy-efficiency manuals. In the second phase (Twelfth Plan, 2012–2017), 100 pilot projects were implemented in five MSME sectors, with a goal to encourage large-scale replication. Subsequent interventions were then implemented in partnership with the World Bank (World Bank-GEF-BEE program, 2010-2014; 2014-2016; 2017-2019) and the UNIDO (UNIDO-BEE program, 2011-2022). These interventions targeted new clusters and followed a similar approach, combining energy audits, technology gap assessments, and technical support for the adoption of energy-efficient technologies. Table SM.11 summarizes the coverage of these interventions.

To exclude observations exposed to these schemes from our sample, we proceed as follows. We collect the list of all industrial clusters covered by each phase of the plans and map each cluster to a location (district, and in a few cases state-wide) and an industry indicator. Observations in these industry$\times$location$\times$year cells are then excluded from our sample. We either exclude the BEE interventions from the Eleventh and Twelfth Plans (option A) or additionally exclude the World Bank-GEF-BEE and UNIDO-BEE programs (option B).

*Renewable Energy Certificates.* Starting in 2010, producers of renewable electricity are awarded Renewable Energy Certificates for each 1 MWh of renewable electricity sold. This mostly targets electricity utilities, but could apply to manufacturing establishments with large on-site generation capacities that sell surplus electricity. These Renewable Energy Certificates must be purchased by buyers subject to Renewable Purchase Obligations, mostly electricity distribution companies, but also large industrial consumers.

We identify firms plausibly exposed to the Renewable Energy Certificate system from 2010 onwards. On the seller side, a firm$\times$year is flagged if it generates electricity above the

TABLE SM.11: Coverage of the BEE MSME Interventions

| Intervention | Years | # Clusters | Industries |
|---|---|---|---|
| Eleventh Plan | 2007 - 2012 | 25 | chemicals and dyes, brass, ceramics, textiles, rice milling, oil milling, machine tools, foundry, ice making, refractories, dairy, galvanizing, brass and aluminum, limestone, tea, seafood processing, paper, sponge iron, brick, glass, coir, tiles, khandsari, and food processing |
| Twelfth Plan | 2012 - 2017 | 5 | food processing, forging, brick, and textile |
| WB - GEF - BEE (Phase I) | 2010 - 2014 | 5 | forging, chemicals, limekiln, foundry, and mixed industries |
| WB - GEF - BEE (Phase II) | 2014 - 2016 | 10 | forging, chemicals, limekiln, foundry, textiles, automobile, and mixed industries |
| WB - GEF - BEE (Phase III) | 2017 - 2019 | 26 | forging, chemicals, limekiln, foundry, textiles, automobile, and mixed industries |
| UNIDO - BEE (MSME) | 2011 - 2022 | 12 | brass, ceramics, dairy, foundry, and hand-tools |

*Notes:* This table lists the energy-efficiency interventions implemented by the Bureau of Energy Efficiency (BEE) in MSME industrial clusters. A cluster is a group of establishments operating in the same industry in a given location, defined at the district level and in a few cases state-wide. The targets of the Eleventh Plan, Twelfth Plan, World Bank-GEF-BEE program, and UNIDO-BEE interventions are taken from Bureau of Energy Efficiency (2017), Shakti Sustainable Energy Foundation (2017), World Bank (2019), and UNIDO (2022), respectively. Clusters labeled "mixed industries" target an undefined set of industries in the location.

250 kW threshold (proxied as 400,000 kWh/year), and reports positive electricity sales. On the buyer side, a firm×year is flagged if it is a large captive generator (generation above 400,000 kWh/year and not a seller) or a large consumer (defined as top 5% of electricity quantity purchased).

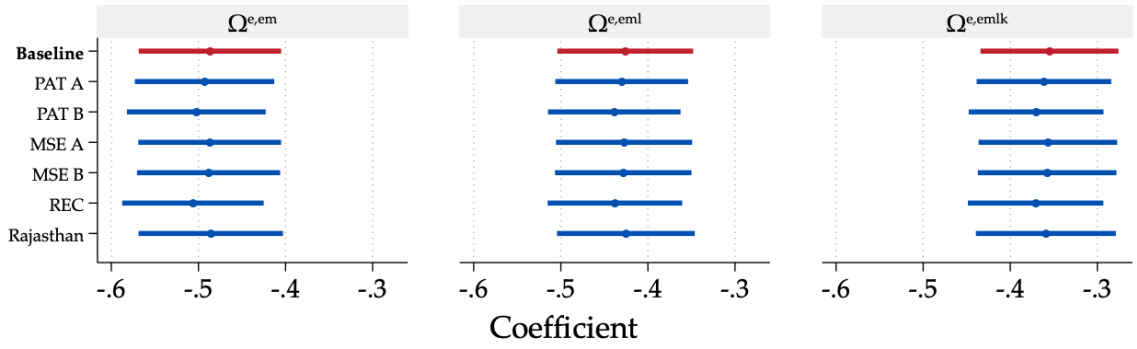
*Rajasthan Solar Policy.* The Rajasthan Solar Policy (2011–2014) promoted the adoption of solar technologies among specific building categories, including industrial facilities requiring hot water. We consider that a firm×year observation is exposed to this policy if the firm is located in Rajasthan and the year is between 2011 and 2014.

Figure SM.12 reports results excluding observations potentially exposed to each of these policies. The results are highly similar to our baseline.

*Other policies.* India's energy policy has three other components that target energy efficiency in the industrial sector. First, a standards and labeling program was initiated in 2006. It sets energy-consumption ratings visible to potential consumers for a number of common equipment types and appliances. This may have affected the energy use of users of these equipment. However, this information provision policy has no reason to interact with firms' scaling up in a meaningful way. Second, in 2010, the Energy Efficiency Financing Platform was set up as a coordination platform designed to connect banks, project developers, energy service companies, and firms to scale up energy-efficiency investment. This platform has two main activities: the organization of "Investment Bazaars" (34 to date) in which energy efficiency projects can be showcased in the presence of potential financiers and business partners, and a training program for financial institutions. Third, the Framework for Energy Efficient Economic Development is a set of financial instru-

ments to mobilize private capital for energy-efficiency. The most important is the Partial Risk Sharing Facility for Energy Efficiency, initiated in 2015, which provides partial credit guarantees to financial institutions financing energy-efficiency investments. As of 2026, 79 energy efficiency projects have been supported. These last two programs are very limited in scale and are hence unlikely to affect our results.

FIGURE SM.12: Robustness to Exposure to Energy-Efficiency Policies




*Notes:* This figure presents IV results for specification (3), with the instrument defined in equation (4). The dependent variable is the relative energy expenditure share, $\Delta^5 \log(\Omega^e/(1-\Omega^e))$, where the definition of $\Omega^e$ is indicated in the panel headings. Each line refers to a specific sample selection rule, as detailed in the text. Regressions are weighted by firm total cost (adjusted for the sampling weight). Standard errors are clustered at the firm and main product×year level. The dots are point estimates and bars are 95% confidence intervals.

## SM.3.3 Results in the Combined ASI-NSS Sample

TABLE SM.12: Energy Expenditure Shares and Firm Size: ASI vs. NSS

| | $\log(\Omega^{e,em})$ | | | $\log(\Omega^{e,eml})$ | | | $\log\left(\frac{w^e e}{py}\right)$ | | |
|---|---|---|---|---|---|---|---|---|---|
| | ASI (1) | NSS (2) | Pooled (3) | ASI (4) | NSS (5) | Pooled (6) | ASI (7) | NSS (8) | Pooled (9) |
| log(sales) | -0.202*** | -0.269*** | -0.263*** | -0.160*** | -0.278*** | -0.270*** | -0.232*** | -0.168*** | -0.165*** |
| | (0.002) | (0.006) | (0.006) | (0.002) | (0.006) | (0.006) | (0.002) | (0.005) | (0.005) |
| Ind. × Year FE | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| R-squared | 0.428 | 0.396 | 0.395 | 0.405 | 0.403 | 0.403 | 0.402 | 0.310 | 0.312 |
| Observations | 139,761 | 205,187 | 344,949 | 139,761 | 205,187 | 344,949 | 139,761 | 205,187 | 344,949 |

*Notes:* This table reports the results of OLS regressions of firm-level energy shares (in logs) on sales (in logs). The dependent variable is indicated in column headings. Columns labeled "ASI" and "NSS" are estimated separately by source, while "Pooled" stacks both datasets. We keep only years common to the NSS and ASI sample. All regressions are weighted by survey sampling weights. Standard errors are clustered at the firm level. *** , **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

## SM.3.4 The Role of Reallocation Across Products and Inputs

**Changes in Energy Shares Due to Changing Product or Input Mix.** First, we examine whether growing firms systematically reallocate their production to more or less energy-intensive products. To do so, we define product-level energy expenditure shares $\Omega^e_j$. We start from product×year energy expenditure shares $\Omega^e_{jt}$, defined as the average

expenditure share of single-product firms $i$ producing product $j$, regress $\Omega^e_{jt}$ on product and year fixed effects and use the product fixed effect as our measure of product-level energy expenditure shares $\Omega^e_j$. Then, we define firm-level expenditure shares capturing only the firm's product portfolio as:

$$\ddot{\Omega}^e_{it} = \sum_{j \in \mathcal{V}} \omega_{ijt} \Omega^e_j \tag{SM.10}$$

where $\omega_{ijt} = \frac{\text{sales}_{ijt}}{\text{sales}_{it}}$ and $\mathcal{V}$ is the set of all product varieties. Therefore, the change in $\ddot{\Omega}^e_{it}$, $\Delta^h \ddot{\Omega}^e_{it}$, captures the change in the firm's energy expenditure share that can be attributed to the reallocation of sales shares to more or less energy-intensive products. The results in columns (1)-(6) of Table SM.13 show that demand-driven expansion in firm size does not cause a reallocation of production towards more or less energy-intensive products.

Second, we conduct a similar analysis for inputs: are increases in firm scale accompanied by shifts of the input mix towards more or less energy-intensive inputs? Starting from the input code×year-level total energy content of inputs $\Psi^{e,tot}_{jt}$ (defined in equation (SM.22)), we regress this variable on input code and year fixed effects and use the input fixed effect as our measure of input-level energy content $\Psi^{e,tot}_j$. Define the average energy content of inputs as:

$$\Phi^e_{it} = \sum_{j \in \mathcal{V}_m} \omega^m_{ijt} \Psi^{e,tot}_{jt} \tag{SM.11}$$

where $\omega^m_{ijt} = \frac{\text{material expenditures}_{ijt}}{\text{material expenditures}_{it}}$ and $\mathcal{V}_m$ is the set of material input varieties. Then, we define firm-level energy content of inputs capturing only the firm's input mix as:

$$\ddot{\Phi}^e_{it} = \sum_{j \in \mathcal{V}_m} \omega^m_{ijt} \Psi^{e,tot}_{j} \tag{SM.12}$$

Hence, the change in $\ddot{\Phi}^e_{it}$, $\Delta^h \ddot{\Phi}^e_{it}$, captures the change in the firm's energy content of inputs that can be attributed to the reallocation of input purchases to more or less energy-intensive inputs. Columns (7)-(8) of Table SM.13 show the effect of a demand-driven expansion in firm size on such reallocation. The coefficients are positive, but only mildly significant, and the economic magnitudes are very small. To interpret this magnitude, we rely on the framework introduced in the next subsection. Define the firm-level share of (direct) energy expenditures relative to output $\Psi^e_{it} = w^e_{it} e_{it} / p_{it} y_{it}$, and similarly for other inputs. The total share of energy expenditures relative to output—including indirect energy expenditures embodied in material inputs (but not capital and energy)—is: $\Psi^{e,tot(m)}_{it} = \Psi^e_{it} + \Psi^m_{it} \Phi^e_{it}$.

Then, we can write:

$$\underbrace{d\log\Psi_{it}^{e,tot(m)}}_{\text{Change in total energy share}} \approx \underbrace{(1-\Psi_{it}^{m})d\log\Psi_{it}^{e}}_{\text{Contribution of the direct effect}} + \underbrace{\underbrace{\Psi_{it}^{m}d\log\Psi_{it}^{m}}_{\text{Change in the material share}} + \underbrace{\Psi_{it}^{m}d\log\Phi_{it}^{e}}_{\text{Change in the average energy content of materials}}}_{\text{Contribution of the indirect effect}}$$

where we use $\frac{\Psi_{it}^{e}}{\Psi_{it}^{e,tot(m)}} \approx 1-\Psi_{it}^{m}$.[7] Finally, we can always decompose:

$$d\log\Phi_{it}^{e} = \underbrace{d\log\ddot{\Phi}_{it}^{e}}_{\text{Reallocation}} + \underbrace{d\log\tilde{\Phi}_{it}^{e}}_{\text{Residual}}$$

where the residual captures the contribution of within-product changes in energy content. Therefore, the estimated elasticity of $d\log\ddot{\Phi}_{it}^{e} = 0.012$ affects the elasticity of the total energy content by 0.008 (0.012 multiplied by the average materials output share equal to 0.68). This is negligible compared to the contribution of the direct effect (elasticity of $-0.53$ times one minus the material share yielding $-0.171$). In addition, it is worth noting that the total effect on the average energy intensity of the input mix $d\log\Phi_{it}^{e}$ is not statistically different from 0.

TABLE SM.13: The Role of Reallocation Across Products and Inputs

| | Products | | | | | | Inputs | |
|---|---|---|---|---|---|---|---|---|
| | $\Delta^5\ddot{\Omega}^e$ | | | $\Delta^5\log(\ddot{\Omega}^e)$ | | | $\Delta^5\ddot{\Phi}^e$ | $\Delta^5\log(\ddot{\Phi}^e)$ |
| | $\Omega^{e,em}$ | $\Omega^{e,eml}$ | $\Omega^{e,emlk}$ | $\Omega^{e,em}$ | $\Omega^{e,eml}$ | $\Omega^{e,emlk}$ | | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| $\Delta^5$ log(sales) | -0.000 | 0.000 | 0.002 | -0.000 | 0.001 | 0.017 | 0.003* | 0.012* |
| | (0.002) | (0.002) | (0.002) | (0.031) | (0.031) | (0.030) | (0.002) | (0.007) |
| Year × Ind. × Cohort FE | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Observations | 184,429 | 184,429 | 184,429 | 184,389 | 184,389 | 184,412 | 182,136 | 182,123 |

*Notes:* This table presents IV results for specification (3), with the instrument defined in equation (4). The dependent variables are the change in average product (column 1–6) and input (column 7–8) energy intensity, as defined in equations (SM.10) and (SM.12), respectively. We consider both level and log changes. Regressions are weighted by firm total cost (adjusted for the sampling weight). Standard errors are clustered at the firm and main product×year level. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

**Changes in the Fuel Mix.** Table SM.14 shows that growing firms do not systematically change the mix of energy inputs they use.

[7] This is exactly true in the case where material varieties are undifferentiated, see equation SM.18, and approximately true otherwise.

TABLE SM.14: Firm Growth and Energy Sources

| | Energy sources | | | | On-site elec. generation |
|---|---|---|---|---|---|
| | Coal | Oil | Elec. | Gas | |
| | (1) | (2) | (3) | (4) | (5) |
| $\Delta^5$ log(sales) | 0.005 | -0.001 | -0.008 | 0.001 | 0.018 |
| | (0.007) | (0.014) | (0.015) | (0.009) | (0.015) |
| Year × Ind. × Cohort FE | ✓ | ✓ | ✓ | ✓ | ✓ |
| Observations | 184,011 | 184,011 | 184,011 | 75,735 | 184,760 |

*Notes:* This table presents IV results for specification (3), with the instrument defined in equation (4). In columns (1)-(4), the dependent variable is the change in the cost share of energy source $f$ in total energy expenditures, where the energy source $f$ is indicated in the column heading. In column (5), the dependent variable is the change in the share of electricity generated on site. We use changes in shares (as opposed to log shares) to accommodate the many zeros in energy source-specific shares. Data on expenditures on natural gas is available only since 2008. Regressions are weighted by firm total cost (adjusted for the sampling weight). Standard errors are clustered at the firm and main product×year level. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

### SM.3.5 Indirect Energy Consumption

In this section, we discuss the determination of the total energy content of production, including that which is indirectly used in material inputs, capital and energy itself.

We first consider the case where materials, capital and energy are produced from the final good, as is the case in the model. Namely, energy, material and capital are produced by the final good at rates $\Xi_t^e$, $\Xi_t^m$ and $\Xi_t^k$, respectively. This most clearly illustrates the link between the direct and indirect energy use and the identification strategy for the effects on indirect energy consumption. We then consider the case with differentiated intermediate inputs, where the formulas are more complex but the key insights are unchanged.

Take an economy producing output $Y_t$. Let $\{y_{it}\}$ be the cost-minimizing allocation that produces $Y_t$. The factor prices are denoted by $\mathbf{w}_t$. For each firm $i$, the cost-minimization problem yields input demand functions that we denote: $m_{it}(y, \mathbf{w})$, $k_{it}(y, \mathbf{w})$, $e_{it}(y, \mathbf{w})$ for materials, capital, and energy, respectively. Let us denote the expenditures on input $x$ by $\mathcal{C}_{it}^x(y, \mathbf{w}) = w_t^x x_{it}(y, \mathbf{w})$. Finally, denote cost shares of output as $\Psi_{it}^x(y, \mathbf{w}) = \frac{\mathcal{C}_{it}^x(y, \mathbf{w})}{p_{it} y_{it}}$. Then, the aggregate quantities of energy, materials and capital required in the production of $Y_t$ are given by: $E_t(Y_t, \mathbf{w}_t) = \int_i e_{it}(y_{it}, \mathbf{w}_t) di$, $M_t(Y_t, \mathbf{w}_t) = \int_i m_{it}(y_{it}, \mathbf{w}_t) di$, $K_t(Y_t, \mathbf{w}_t) = \int_i k_{it}(y_{it}, \mathbf{w}_t) di$. In what follows, we drop the arguments of the input demand and cost functions to ease notation.

**Final good requirements of producing inputs.** At the aggregate level, the final goods requirement of producing the inputs used in the production of $Y_t$ is given by:

$$\begin{aligned}
&(\Xi_t^m)^{-1}M_t + (\Xi_t^k)^{-1}K_t + (\Xi_t^e)^{-1}E_t \\
&= w_t^m M_t + w_t^k K_t + w_t^e E_t \\
&= Y_t \int \frac{w_t^m m_{it}}{p_{it}y_{it}} \frac{p_{it}y_{it}}{Y_t} di + Y_t \int \frac{w_t^k k_{it}}{p_{it}y_{it}} \frac{p_{it}y_{it}}{Y_t} di + Y_t \int \frac{w_t^e e_{it}}{p_{it}y_{it}} \frac{p_{it}y_{it}}{Y_t} di \\
&= (\Psi_t^m + \Psi_t^k + \Psi_t^e) Y_t \\
&\equiv \Lambda_t Y_t
\end{aligned}$$

The second line uses the fact that the price of input $x$ relative to the final good is $w_t^x = (\Xi_t^x)^{-1}$. The last line defines $\Lambda_t = \Psi_t^m + \Psi_t^k + \Psi_t^e$ as the average final goods requirement of producing the inputs used in the production of $Y_t$. Similarly, for every firm $i$, the average per-unit final goods requirement of producing the inputs used in the production of $y_{it}$ is given by $\Lambda_{it} \equiv \Psi_{it}^m + \Psi_{it}^k + \Psi_{it}^e$.

**Total energy content of production.** At the aggregate level,

$$\mathcal{C}_t^{e,tot} = \underbrace{\mathcal{C}_t^e}_{\substack{\text{Direct energy use}}} + Y_t \times \underbrace{\Lambda_t}_{\substack{\text{Indirect final} \\ \text{good usage} \\ \text{Round (1)}}} \times \left( \underbrace{\Psi_t^e}_{\substack{\text{Direct energy} \\ \text{use of indirect} \\ \text{final good (1)}}} + \underbrace{\Lambda_t}_{\substack{\text{Indirect final} \\ \text{good usage} \\ \text{Round (2)}}} \left( \underbrace{\Psi_t^e}_{\substack{\text{Direct energy} \\ \text{use of indirect} \\ \text{final good (2)}}} + \underbrace{\Lambda_t}_{\substack{\text{Indirect final} \\ \text{good usage} \\ \text{Round (3)}}} \times ... \right) \right)$$

Then, the total energy expenditure share of output is given by:

$$\Psi_t^{e,tot} = \frac{\Psi_t^e}{1 - \Lambda_t} \tag{SM.13}$$

Therefore, total energy expenditures are equal to direct expenditures times a multiplier $1/(1-\Lambda_t)$. We can write $\Psi_t^{e,tot}$ as:

$$\Psi_t^{e,tot} = \frac{w_t^e E_t}{w_t^l L_t^P + \text{Profits}_t} \tag{SM.14}$$

where $L_t^P$ is labor used in production. That is, the total energy cost share of gross output is simply energy expenditures divided by value added.[8] Finally, this expression can be

[8] Value added takes this definition because capital is a produced factor here. If we treated capital as a primary factor, and excluded indirect energy use through capital, we would instead obtain that the total energy cost share of gross output is $\frac{w_t^e E_t}{w_t^l L_t^P + w_t^k K_t + \text{Profits}_t}$.

rewritten as:

$$\underbrace{\Psi_t^{e,tot}}_{\substack{\text{Total energy}\\ \text{share of}\\ \text{final good}}} = \underbrace{\Psi_t^{e}}_{\substack{\text{Direct energy share}\\ \text{of final good}}} + \underbrace{\underbrace{\Lambda_t}_{\substack{\text{Indirect final}\\ \text{good usage of}\\ \text{final good}}} \times \underbrace{\Psi_t^{e,tot}}_{\substack{\text{Total energy}\\ \text{share of}\\ \text{final good}}}}_{\substack{\text{Indirect energy share}\\ \text{of final good}}} \tag{SM.15}$$

Following the same logic, for each firm $i$, the total energy expenditure share is given by:

$$\underbrace{\Psi_{it}^{e,tot}}_{\substack{\text{Total energy}\\ \text{share of}\\ \text{firm } i}} = \underbrace{\Psi_{it}^{e}}_{\substack{\text{Direct energy share}\\ \text{of firm } i}} + \underbrace{\underbrace{\Lambda_{it}}_{\substack{\text{Indirect final}\\ \text{good usage of}\\ \text{firm } i}} \times \underbrace{\Psi_t^{e,tot}}_{\substack{\text{Total energy}\\ \text{share of}\\ \text{final good}}}}_{\substack{\text{Indirect energy share}\\ \text{of firm } i}} \tag{SM.16}$$

Note that at the firm level, the total energy expenditure share depends on the firm's direct energy expenditure share $\Psi_{it}^e$ and the firm's final good usage $\Lambda_{it}$ multiplied by the *economy-wide* total energy expenditure share $\Psi_t^{e,tot}$. This last component is not firm-specific: the firm purchases inputs produced using the final good, with energy content corresponding to the economy-wide average.

Note that we can apply the same logic to construct the total energy share of any cost component. For instance, the total energy share of total costs is given by:

$$\Omega_t^{e,tot} = \frac{\Omega_t^e}{1 - \Lambda_t} \tag{SM.17}$$

Finally, we can consider the indirect energy content of all inputs, or of some inputs only. For instance, the total output share of energy including only the indirect energy content of materials is given by:

$$\Psi_t^{e,tot(m)} = \frac{\Psi_t^e}{1 - \Psi_t^m} \tag{SM.18}$$

**Effect of firm scale on the total energy content of output.** How does the scaling up channel affect the total energy expenditure share, and what can we learn from cross-sectional regressions?

*Macro and micro elasticities.* Consider a change in aggregate TFP raising aggregate output by 1%. What is the effect on the total energy expenditure share? Differentiating equation

(SM.15) with respect to $\log Y_t$ gives:

$$\frac{d\log \Psi_t^{e,tot}}{d\log Y_t} = (1-\Lambda_t)\frac{d\log \Psi_t^{e}}{d\log Y_t} + \Lambda_t\left(\frac{d\log \Lambda_t}{d\log Y_t} + \frac{d\log \frac{\Psi_t^e}{1-\Lambda_t}}{d\log Y_t}\right) \tag{SM.19}$$
$$= \frac{d\log \Psi_t^{e}}{d\log Y_t} - \frac{d\log(1-\Lambda_t)}{d\log Y_t}$$

where the second line uses the equivalence between equations (SM.13) and (SM.15). Note that fraction of the direct share in the total energy expenditure share is given by $\frac{\Psi_t^e}{\Psi_t^{e,tot}} = 1-\Lambda_t$. The three components of the elasticity are the scale elasticity of the direct energy expenditure share, the scale elasticity of the final good usage, and the scale elasticity of the total energy content.

Now, consider what a firm-level regression of changes in total energy content on exogenous demand-driven changes in output identifies. Suppose that we can give random demand shocks to a subset of firms, that change their values of $a_{it}$. These shocks yield proportional changes $\hat{y}_{it}$ such that $Cov[\hat{y}_{it}, a_{it-1}] = 0$. For any given firm, we can decompose the change in total energy expenditure shares as:

$$\frac{d\log \Psi_{it}^{e,tot}}{d\log y_{it}}\hat{y}_{it} = (1-\Lambda_{it})\frac{d\log \Psi_{it}^{e}}{d\log y_{it}}\hat{y}_{it} + \Lambda_{it}\left(\frac{d\log \Lambda_{it}}{d\log y_{it}}\hat{y}_{it} + \frac{d\log \frac{\Psi_t^e}{1-\Lambda_t}}{d\log y_{it}}\hat{y}_{it}\right)$$
$$= (1-\Lambda_{it})\frac{d\log \Psi_{it}^{e}}{d\log y_{it}}\hat{y}_{it} + \Lambda_{it}\frac{d\log \Lambda_{it}}{d\log y_{it}}\hat{y}_{it}$$

where the final term is zero because individual firm output does not affect aggregates. Take revenue-weighted expectations of both sides, and we have

$$\mathbb{E}_R[\frac{d\log \Psi_{it}^{e,tot}}{d\log y_{it}}]\mathbb{E}_{\mathbb{R}}[\hat{y}_{it}] = \mathbb{E}_{\mathbb{R}}[(1-\Lambda_{it})\frac{d\log \Psi_{it}^{e}}{d\log y_{it}}]\mathbb{E}_{\mathbb{R}}[\hat{y}_{it}] + \mathbb{E}_R[\Lambda_{it}\frac{d\log \Lambda_{it}}{d\log y_{it}}]\mathbb{E}_{\mathbb{R}}[\hat{y}_{it}]$$

where the expectations factor by assumption of the random allocations of demand shocks. Write this as:

$$\mathbb{E}_R[\frac{d\log \Psi_{it}^{e,tot}}{d\log y_{it}}] = \mathbb{E}_{\mathbb{R}}[(1-\Lambda_{it})\frac{d\log \Psi_{it}^{e}}{d\log y_{it}}] + \mathbb{E}_R[\Lambda_{it}\frac{d\log \Lambda_{it}}{d\log y_{it}}] \tag{SM.20}$$

*The missing intercept problem.* Comparing equation (SM.20) with the first line of equation (SM.19), we can see that the mean causal elasticity of total energy expenditure shares at the firm level—the coefficient identified by our cross-sectional regression—will differ from the aggregate elasticity. First, even given random allocation of the demand shocks, it is not the case in general that $(1-\Lambda_t)\frac{d\log \Psi_t^e}{d\log Y_t}$ is equal to $\mathbb{E}_{\mathbb{R}}[(1-\Lambda_{it})\frac{d\log \Psi_{it}^e}{d\log y_{it}}]$, and that $\Lambda_t\frac{d\log \Lambda_t}{d\log Y_t}$ is equal to $\mathbb{E}_R[\Lambda_{it}\frac{d\log \Lambda_{it}}{d\log y_{it}}]$, since the aggregate terms will feature general equilibrium reallocation with an increase in aggregate productivity (see the discussion of equation (15)

in the main text). However, even if these differences can be ignored, the term $\Lambda_t \frac{d \log \frac{\Psi^e_t}{1-\Lambda_t}}{d \log Y_t}$ is entirely missing from the cross-sectional regressions. The cross-sectional coefficient is a weighted average of the firm-level scale elasticities of $\Psi^e_{it}$ and $\Lambda_{it} = \Psi^m_{it} + \Psi^e_{it} + \Psi^k_{it}$, and it does *not* depend on the energy intensity of inputs $\frac{\Psi^e_t}{1-\Lambda_t}$.

**Identification from firm-level elasticities.** All the terms in (SM.19) correspond to the aggregate elasticities of firm input shares with respect to aggregate scale. Let us now assume that the size-weighted micro-elasticities are approximately equal to aggregate elasticities. As highlighted above, this is not true in general; however, it is useful as a back-of-the-envelope exercise relying on reduced-form evidence only, and in our quantitative model, we test and find support for this assumption.[9] Then, we can write:

$$\frac{d \log \Psi^{e,tot}_t}{d \log Y_t} = \frac{d \log \Psi^e_t}{d \log Y_t} - \frac{d \log(1-\Lambda_t)}{d \log Y_t} \approx \underbrace{\mathbb{E}_{\mathbb{R}}[\frac{d \log \Psi^e_{it}}{d \log y_{it}}]}_{\equiv \eta^{\Psi^e}} - \underbrace{\mathbb{E}_R[\frac{d \log(1-\Lambda_{it})}{d \log y_{it}}]}_{\equiv \eta^{(1-\Lambda)}}$$

where $\eta^{\Psi^e}$ and $\eta^{(1-\Lambda)}$ are the coefficients of the firm-level IV specification (3) with dependent variables $\Delta \log \Psi^e_{it}$ and $\Delta \log(1-\Lambda_{it})$, respectively.

That is, to estimate how the total energy expenditure share, equal to energy expenditures over value added, responds to an aggregate increase in scale, we estimate how firm-level energy expenditures over value added respond to a firm-level increase in scale. The intuition is that if, on average for all firms, energy expenditures relative to value added go down with scale, then it cannot be that in general equilibrium indirect energy use undoes the negative effect of scale on direct energy use.

This identification argument similarly holds for the other versions of the total cost or energy expenditure share. For instance, we will estimate the aggregate elasticity of the total energy share considering only indirect energy use through materials $\frac{d \log \Psi^{e,tot(m)}_t}{d \log Y_t}$ as $\eta^{\Psi^e} - \eta^{(1-\Psi^m)}$.

Panel A of Table SM.15 presents the results of this analysis. For each choice of denominator of the energy share, it reports three regression coefficients: the scale elasticity of the direct energy expenditure share (line labeled "Direct"), the firm-level scale elasticity of the total share (line labeled "Direct + Indirect: Firm-level", corresponding to equation SM.20), and the implied aggregate scale elasticity of the total energy expenditure share (line labeled "Direct + Indirect: Aggregate", corresponding to equation SM.19). In odd columns, we are considering the energy content associated to the final good usage of all inputs, while in even columns we are considering the energy content associated to the final good usage of materials only. The coefficients for the firm-level effect on the total share (line "Direct + Indirect: Firm-level") tend to be negative but meaningfully smaller than the direct effect (line "Direct"). This is because these coefficients are weighted averages of the scale

[9] Using our estimated model, we can estimate the macro-elasticities (Table OA.4) and find that they are highly similar to their micro counterparts (in Table OA.3).

elasticity of the direct energy share, which is negative, and of the scale elasticity of the total material, energy, and capital share (in odd columns), which is negative, but less so than that of the direct share. In even columns, the second term is the scale elasticity of the material share, which is positive. By contrast, our estimates of the aggregate elasticity of the total energy share (line "Direct + Indirect: Aggregate") are relatively similar (or larger than) the elasticity of the direct share. Including only the indirect energy content of materials slightly dampens the negative scale effect, because of the increase in the material share in total costs, while accounting for the indirect energy content of all inputs amplifies the negative effect, because of the fall in the indirect energy use due to capital and energy inputs.

**Introducing differentiated inputs.** In the derivations above, the material, energy, and capital good are produced from the final good. We now introduce differentiated inputs. Assume that there exists $\mathcal{N}$ distinct product varieties. We index all varieties by $\{1, \ldots, \mathcal{N}\}$, partitioned into the disjoint sets $\mathcal{V}_m, \mathcal{V}_k, \mathcal{V}_e, \mathcal{V}_c$ of material, capital, energy, and consumption varieties, and denote $|\mathcal{V}_x| = \mathcal{N}_x$ for $x \in \{m, k, e, c\}$. The total energy expenditure share of firm $i$ is given by:

$$\Psi_{it}^{e,tot} = \Psi_{it}^{e} + \Psi_{it}^{m} \sum_{j' \in \mathcal{V}_m} \omega_{ij't}^{m} \Psi_{j't}^{e,tot} + \Psi_{it}^{k} \sum_{j' \in \mathcal{V}_k} \omega_{ij't}^{k} \Psi_{j't}^{e,tot} + \Psi_{it}^{e} \sum_{j' \in \mathcal{V}_e} \omega_{ij't}^{e} \Psi_{j't}^{e,tot} \tag{SM.21}$$

where $\omega_{ij't}^{m}$ is the share of firm $i$ material spending that flows to material variety $j'$, with $\sum_{j' \in \mathcal{V}_m} \omega_{ij't}^{m} = 1$ (and analogously for capital and energy, each summing to one over its own input category's varieties). Aggregating across firms producing the same variety $j$ and writing in matrix form, we obtain:

$$\begin{aligned} &\boldsymbol{\Psi}_t^{e,tot} = \boldsymbol{\Psi}_t^{e} + \mathbf{B}_t \, \boldsymbol{\Psi}_t^{e,tot}, \quad \mathbf{B}_t \equiv \operatorname{diag}(\boldsymbol{\Psi}_t^{m}) \, \boldsymbol{\Omega}^{m,m} + \operatorname{diag}(\boldsymbol{\Psi}_t^{k}) \, \boldsymbol{\Omega}^{k,k} + \operatorname{diag}(\boldsymbol{\Psi}_t^{e}) \, \boldsymbol{\Omega}^{e,e} \\ &\boldsymbol{\Psi}_t^{e,tot} = (\mathbf{I} - \mathbf{B}_t)^{-1} \, \boldsymbol{\Psi}_t^{e}. \end{aligned} \tag{SM.22}$$

where $\boldsymbol{\Psi}_t^{e,tot}, \boldsymbol{\Psi}_t^{e}, \boldsymbol{\Psi}_t^{m}, \boldsymbol{\Psi}_t^{k}$ are $\mathcal{N} \times 1$ vectors of variety-level shares; $\boldsymbol{\Omega}^{m,m}, \boldsymbol{\Omega}^{k,k}, \boldsymbol{\Omega}^{e,e}$ are $\mathcal{N} \times \mathcal{N}$ input-share matrices with $(\boldsymbol{\Omega}^{m,m})_{jj'} = \omega_{jj't}^{m}$ the share of product $j$'s material spending on variety $j'$ (nonzero only for $(j, j')$ in the materials bundle), and similarly for energy and capital.[10]

As before, we can also define total cost shares (as opposed to output shares) as $\boldsymbol{\Omega}_t^{e,tot} = (\mathbf{I} - \mathbf{B}_t)^{-1} \, \boldsymbol{\Omega}_t^{e}$, and focus on the indirect energy use due to one input only, which in the case of materials yields $\boldsymbol{\Psi}_t^{e,tot(m)} = (\mathbf{I} - \operatorname{diag}(\boldsymbol{\Psi}_t^{m}) \, \boldsymbol{\Omega}^{m,m})^{-1} \, \boldsymbol{\Psi}_t^{e}$. Solving for the total energy shares of all varieties, we can then construct the firm-level total energy expenditure share using equation (SM.21).

How does a change in aggregate scale affect total energy shares in this case? Define aggregate total energy share as $\Psi_t^{e,tot} = \mathbf{s}_t{}^{\top} \boldsymbol{\Psi}_t^{e,tot}$ where $\mathbf{s}_t$ is the $\mathcal{N} \times 1$ vector of product output

[10] These matrices are defined at the product (and not firm) level and do not have a $t$ index because detailed input-output tables are available only at the product level (we do not observe the firm×firm network) and for one date during our sample period.

(sales) shares, $\sum_k s_{kt} = 1$ (distinct from the input-mix weights $\omega^x$). Then, differentiating $\Psi_t^{e,tot}$ with respect to $Y_t$, holding product shares $\mathbf{s}_t$ fixed, gives:

$$\frac{d \log \Psi_t^{e,tot}}{d \log Y_t} = \frac{d \log \Psi_t^{e}}{d \log Y_t} + \frac{d \log \Psi_t^{m}}{d \log Y_t} \bar{r}_t^m + \frac{d \log \Psi_t^{k}}{d \log Y_t} \bar{r}_t^k + \frac{d \log \Psi_t^{e}}{d \log Y_t} \bar{r}_t^e$$

where $r_t^x \equiv \mathbf{s_t}^\top (\mathbf{I} - \mathbf{B}_t)^{-1} \operatorname{diag}(\mathbf{\Psi_t^x}) \mathbf{\Omega^x} \mathbf{\Psi_t^{e,tot}} (\Psi_t^{e,tot})^{-1}$. The formula is highly similar to the homogeneous final good case, except that each input carries its own aggregate elasticity times a channel-specific indirect-to-total ratio $\bar{r}_t^x$. $\bar{r}_t^x$ captures the total energy content of the input bundle $x$, that can differ from the total energy content of the representative good $\Psi_t^{e,tot}$.

The firm-level counterpart of this relationship suffers from a missing intercept problem, because the total energy share terms in equation (SM.21) are not firm-specific and will be soaked up by the intercept of the regression. However, again assuming that size-weighted regressions identify the aggregate elasticities, we can write

$$\frac{d \log \Psi_t^{e,tot}}{d \log Y_t} \approx \eta^{\Psi^e} + \eta^{\Psi^m} \bar{r}_t^m + \eta^{\Psi^k} \bar{r}_t^k + \eta^{\Psi^e} \bar{r}_t^e$$

where the ratios $\bar{r}_t^x$ can be computed from the data.

Panel B of Table SM.15 presents the results of this analysis. The aggregate elasticities are noticeably more negative than in Panel A. This is because the inputs used by manufacturing firms tend to have a lower total energy content than the final good produced by the manufacturing sector. In particular in the case of materials, the (positive) material share elasticity is amplified by a smaller indirect multiplier under differentiation ($\bar{r}_t^m$ below its scalar value), so the aggregate elasticity moves toward the more negative direct elasticity.[11]

**Linear specification.** In Panel C, we additionally present results for a linear specification. The linear specification also suffers from the missing intercept problem, which can be seen by comparing:

$$\frac{d\Psi_t^{e,tot}}{d \log Y_t} = \frac{d\Psi_t^{e}}{d \log Y_t} + \frac{d\Lambda_t}{d \log Y_t} \times \frac{\Psi_t^e}{1-\Lambda_t} + \Lambda_t \times \frac{d \frac{\Psi_t^e}{1-\Lambda_t}}{d \log Y_t}$$
$$\frac{d\Psi_{it}^{e,tot}}{d \log y_{it}} = \frac{d\Psi_{it}^{e}}{d \log y_{it}} + \frac{d\Lambda_{it}}{d \log y_{it}} \times \frac{\Psi_t^e}{1-\Lambda_t}$$

One advantage of the linear specification is that the partial equilibrium indirect effect $\frac{d\Lambda_{it}}{d \log y_{it}} \times \frac{\Psi_t^e}{1-\Lambda_t}$ enters additively with the direct effect $\frac{d\Psi_{it}^e}{d \log y_{it}}$, so that comparing the two coefficients allows us to gauge the magnitude of the (partial equilibrium) amplification or dampening due to indirect input usage. Taking our baseline outcome $\Omega^{e,eml}$, the effect

[11] Because our model assumes that the input has the same energy intensity as the final good, the aggregate elasticities implied by the model will therefore tend to be conservative.

of scale on indirect usage via all inputs tends to amplify the direct effect by 50%, while indirect usage via materials only dampens the direct effect by 50%. We also produce an estimate of aggregate elasticities for the linear specification by estimating the effect of scale on the firm level value of $\frac{\Psi^e_{it}}{1-\Lambda_{it}}$ (in levels). In this case, this assumes that the weighted regression identifies the aggregate level coefficient, which is a different assumption from the log specification.

TABLE SM.15: Direct and Indirect Energy Content

**Panel A**: Log Specification – Inputs are the final good

| | $\Delta^5 \log(\Omega^{e,em})$ | | $\Delta^5 \log(\Omega^{e,eml})$ | | $\Delta^5 \log(\Omega^{e,emlk})$ | | $\Delta^5 \log(\Psi^{e})$ | |
|---|---|---|---|---|---|---|---|---|
| | All | Only M | All | Only M | All | Only M | All | Only M |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Direct | -0.443*** | -0.443*** | -0.397*** | -0.397*** | -0.334*** | -0.334*** | -0.532*** | -0.532*** |
| | (0.038) | (0.038) | (0.037) | (0.037) | (0.038) | (0.038) | (0.039) | (0.039) |
| Direct + Indirect: Firm-level | -0.118*** | -0.086*** | -0.066*** | -0.035*** | -0.009 | 0.025** | -0.204*** | -0.168*** |
| | (0.010) | (0.010) | (0.010) | (0.010) | (0.008) | (0.012) | (0.020) | (0.020) |
| Direct + Indirect: Aggregate | -0.606*** | -0.429*** | -0.553*** | -0.382*** | -0.525*** | -0.312*** | -0.601*** | -0.449*** |
| | (0.082) | (0.046) | (0.082) | (0.045) | (0.081) | (0.047) | (0.090) | (0.052) |

**Panel B**: Log Specification – Differentiated inputs

| | $\Delta^5 \log(\Omega^{e,em})$ | | $\Delta^5 \log(\Omega^{e,eml})$ | | $\Delta^5 \log(\Omega^{e,emlk})$ | | $\Delta^5 \log(\Psi^{e})$ | |
|---|---|---|---|---|---|---|---|---|
| | All | Only M | All | Only M | All | Only M | All | Only M |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Direct | -0.443*** | -0.443*** | -0.397*** | -0.397*** | -0.334*** | -0.334*** | -0.532*** | -0.532*** |
| | (0.038) | (0.038) | (0.037) | (0.037) | (0.038) | (0.038) | (0.039) | (0.039) |
| Direct + Indirect: Firm-level | -0.120*** | -0.092*** | -0.069*** | -0.041*** | -0.009 | 0.020 | -0.203*** | -0.173*** |
| | (0.013) | (0.014) | (0.013) | (0.014) | (0.012) | (0.016) | (0.021) | (0.022) |
| Direct + Indirect: Aggregate | -0.797*** | -0.505*** | -0.752*** | -0.460*** | -0.687*** | -0.396*** | -0.882*** | -0.587*** |
| | (0.052) | (0.038) | (0.052) | (0.038) | (0.053) | (0.040) | (0.062) | (0.047) |

**Panel C**: Linear Specification – Inputs are the final good

| | $\Delta^5 \Omega^{e,em}$ | | $\Delta^5 \Omega^{e,eml}$ | | $\Delta^5 \Omega^{e,emlk}$ | | $\Delta^5 \Psi^{e}$ | |
|---|---|---|---|---|---|---|---|---|
| | All | Only M | All | Only M | All | Only M | All | Only M |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Direct | -0.029*** | -0.029*** | -0.019*** | -0.019*** | -0.013*** | -0.013*** | -0.033*** | -0.033*** |
| | (0.004) | (0.004) | (0.003) | (0.003) | (0.003) | (0.003) | (0.005) | (0.005) |
| Direct + Indirect: Firm-level | -0.064*** | -0.024*** | -0.032*** | -0.010*** | -0.006** | 0.001 | -0.112*** | -0.045*** |
| | (0.006) | (0.003) | (0.005) | (0.003) | (0.003) | (0.002) | (0.015) | (0.006) |
| Direct + Indirect: Aggregate | -0.293*** | -0.081*** | -0.234** | -0.062*** | -0.207*** | -0.040*** | -0.226*** | -0.069*** |
| | (0.104) | (0.014) | (0.092) | (0.013) | (0.079) | (0.011) | (0.085) | (0.016) |

*Notes:* This table presents the analysis of the effect of firm scale on direct and indirect energy use. All cells correspond to IV results for specification (3), with the instrument defined in equation (4), with different outcome variables. The first column heading defines the outcome variable. For each outcome variable, we report six coefficients. The lines "Direct" report the effect on the direct energy share. The lines "Direct + Indirect: Firm-level" report the effect on the firm-level total energy expenditure share as defined in (SM.16). The lines "Direct + Indirect: Aggregate" report the effect on the firm-level value of the aggregate total energy share variable defined in (SM.13). For the lines "Direct + Indirect", we consider either indirect energy use associated to all inputs (columns labeled "All") or indirect energy use associated to materials only (columns labeled "Only M"). Regressions are weighted by firm total cost (adjusted for the sampling weight). Standard errors are clustered at the firm and main product×year level. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

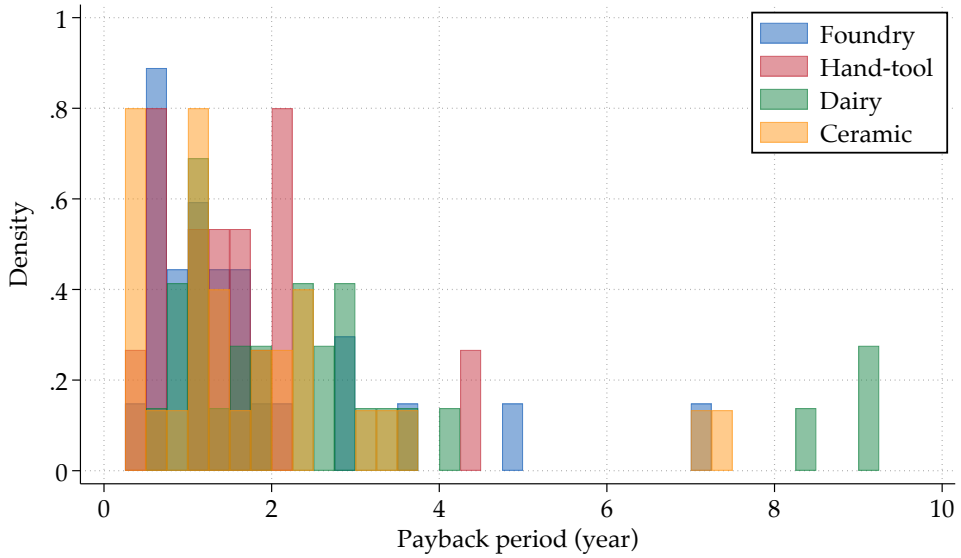
## SM.3.6 Additional results

TABLE SM.16: Heterogeneity: By Fraction of Energy Used in Fired Systems

| | $\Delta^5 \log(\Omega/(1-\Omega))$ | | | | | | $\Delta^5 \log\left(\frac{e}{y}\right)$ | |
|---|---|---|---|---|---|---|---|---|
| | $\Omega^{e,em}$ | | $\Omega^{e,eml}$ | | $\Omega^{e,emlk}$ | | | |
| | Low | High | Low | High | Low | High | Low | High |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| $\Delta^5 \log(py)$ | -0.434*** | -0.671*** | -0.372*** | -0.620*** | -0.303*** | -0.537*** | | |
| | (0.042) | (0.134) | (0.040) | (0.129) | (0.041) | (0.131) | | |
| $\Delta^5 \log(y)$ | | | | | | | -0.361*** | -0.529*** |
| | | | | | | | (0.115) | (0.182) |
| Year × Ind. × Cohort FE | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Observations | 133,730 | 38,319 | 133,740 | 38,338 | 133,521 | 38,298 | 99,622 | 28,274 |

*Notes:* This table presents IV results for specification (3), where sales growth $\Delta^5 \log(py)$ and quantity growth $\Delta^5 \log(y)$ are instrumented with the firm-level demand shock defined in equation (4). Results are reported separately for firms in the top 6 industries in terms of their fraction of total energy used in fired systems (see Table SM.5, these industries correspond to approximately 25% of ASI firms). Regressions are weighted by firm total cost (adjusted for the sampling weight). Standard errors are clustered at the firm and main product×year level. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

FIGURE SM.13: Technology Menus: Distribution of Payback Periods




*Notes:* This figure shows the distribution of payback periods in the technology menu data from United Nations Industrial Development Organization (UNIDO) Energy Efficiency and Renewable Energy Technology Compendiums. The payback period is defined as the yearly monetary gain from energy savings, divided by the upfront investment cost.

### SM.3.7 External Validity: Evidence from the United States

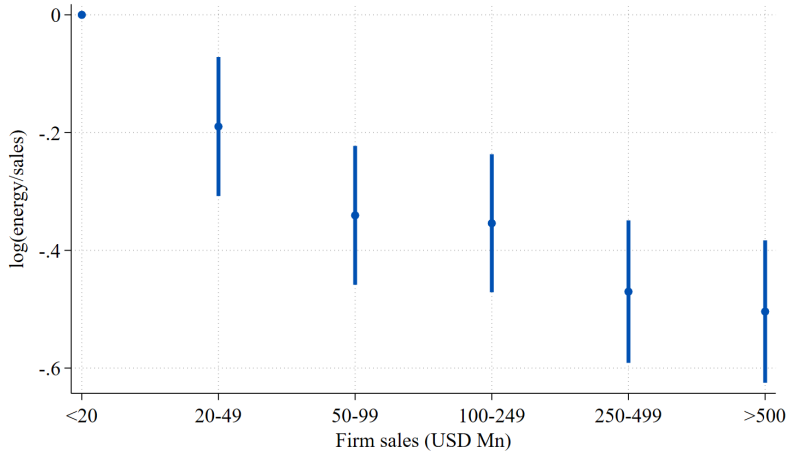
In this section we provide suggestive evidence that our results generalize beyond LMICs like India to high-income countries.

**Energy Intensity by Firm Size.** The Manufacturing Energy Consumption Survey for the United States reports energy intensity (energy in physical unit per value of shipments in USD millions) by bins of firm size (6 bins defined by value of shipments in USD millions). The data is provided for a selected number of industries: in the modal wave we have forty-five 6-digit industries, one 5-digit industry, seven 4-digit industries, and eleven 3-digit industries. We estimate the following specification:

$$\log(e/py)_{qst} = \alpha_{st} + \sum_{x=1}^{6} \beta_x [Bin = x]_{qst} + \varepsilon_{qst} \tag{SM.23}$$

where $q$ indexes the six size bins, and $\alpha_{st}$ are industry×time fixed effects.

FIGURE SM.14: Energy Intensity by Firm Size in the United States



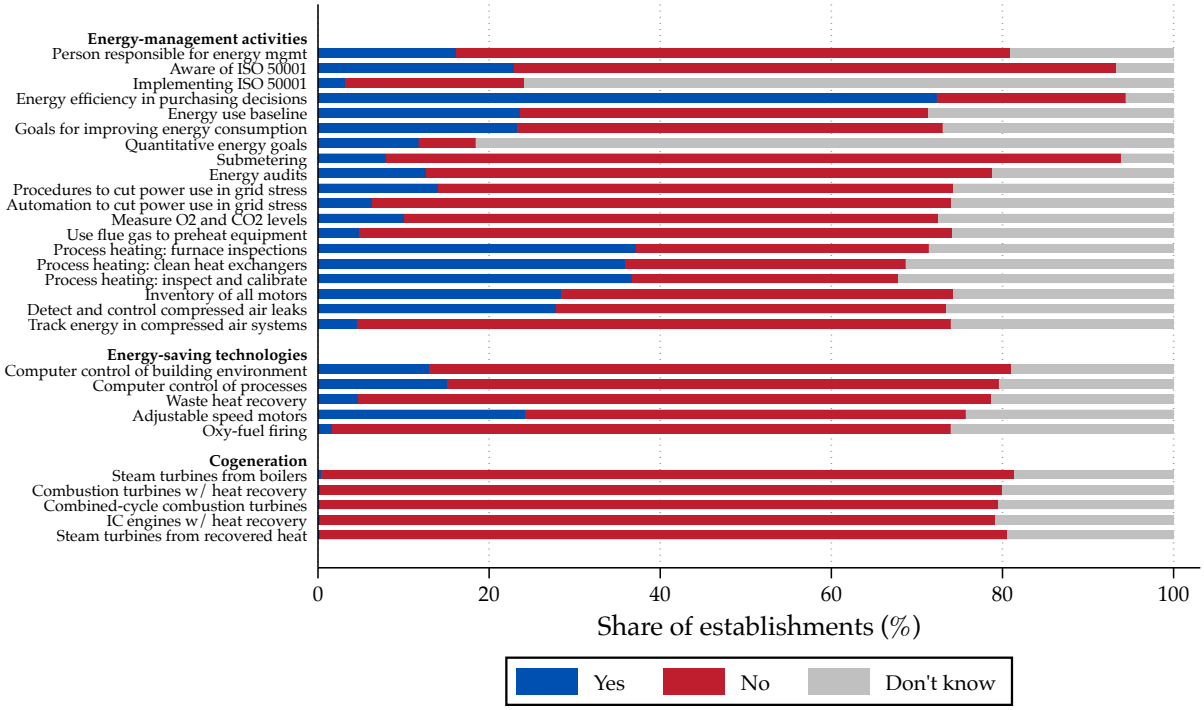

*Notes:* The figure shows the results of estimating equation (SM.23). The dots are point estimates and bars are 95% confidence intervals.

The results are presented in Figure SM.14. The figure shows that energy intensity declines in firm size (in the cross-section) in the United States too.

**Adoption of Energy-Efficiency Measures.** For India, we propose that one of the two mechanisms behind the scale dependence in energy demand is energy-efficiency investments with a fixed cost component, that are more likely to be undertaken by larger firms. For this mechanism to apply to the United States too, it must be that some U.S. firms operate at a scale that is insufficient to rationalize the adoption of some energy-efficiency investments. While U.S. firms are on average much larger than Indian firms, we provide two suggestive pieces of evidence that this mechanism is active in the U.S. too. First, the U.S. Manufacturing Energy Consumption Survey reports adoption rates for a large number of energy efficiency measures. Figure SM.15 shows that for most of this measures, there is still significant under-adoption. In addition, for one key energy-efficiency invest-

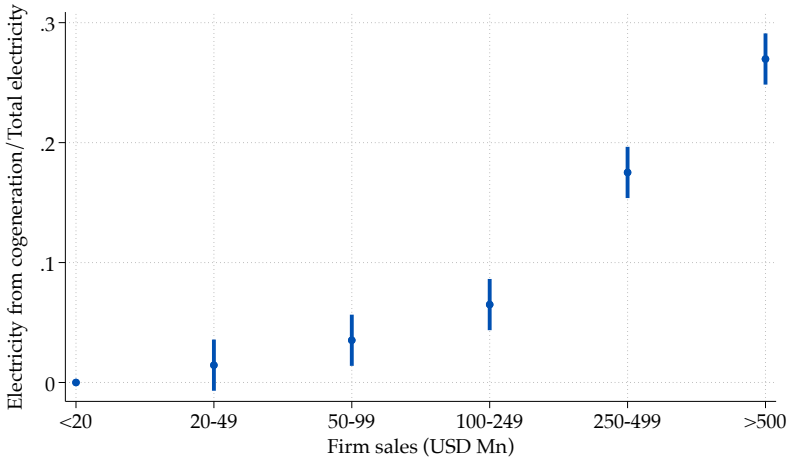
ment, cogeneration, the survey reports adoption by firm size. Figure SM.16 shows that the share of electricity obtained from cogeneration is steeply increasing in firm size.

FIGURE SM.15: Adoption of Energy-Efficiency Measures by U.S. Establishments




*Notes:* This figure reports the shares of firms declaring Yes/No/Don't know to questions about the adoption of various energy efficiency measures. The data is from the 2018 wave of the U.S. Manufacturing Energy Consumption Survey.

FIGURE SM.16: Cogeneration by Firm Size




*Notes:* The figure shows the share of electricity obtained from cogeneration by bins of firm size. The dots are point estimates and bars are 95% confidence intervals.

# SM.4. QUANTITATIVE MODEL RESULTS

Consider the cost function in the quantitative model:

$$\mathcal{C}(y) = A_{Yt}^{-1}\left((A_{mt}^{-1}w_t^m y_t^{\epsilon_m})^{1-\sigma_m} + \left((A_{et}^{-1}\, y_t^{\epsilon_e} w_t^e)^{1-\sigma_e} + (y_t^{\epsilon_l}(w_t^l)^{\alpha}(w_t^k)^{1-\alpha})^{1-\sigma_e}\right)^{\frac{1-\sigma_m}{1-\sigma_e}}\right)^{\frac{1}{1-\sigma_m}}$$

To pin down the scale elasticities $\epsilon_m, \epsilon_e$ and $\epsilon_l$ we derive analytical expressions for three moments which we compute in the micro data: (1) the elasticity of the energy share relative to the material and labor bundle with respect to firm size ($\eta_{e/ml} \equiv \frac{d\log\left(\Omega^e/(\Omega^m+\Omega^l)\right)}{d\log y}$), (2) the elasticity of energy over labor cost share with respect to firm size ($\eta_{e/l} \equiv \frac{d\log\left(\Omega^e/\Omega^l\right)}{d\log y}$), and (3) the cost elasticity $\mathcal{E} \equiv \frac{d\log \mathcal{C}}{d\log y}$. We solve for them in closed form below, dropping time subscripts.

## SM.4.1 Closed Form Expression for $\epsilon_m, \epsilon_l, \epsilon_e$

**Energy to Materials and Labor: $\eta_{e/ml}$.**

$$\frac{d\log\left(\frac{\Omega^e}{\Omega^l+\Omega^m}\right)}{d\log y} = \frac{d\log\Omega^e}{d\log y} - \frac{\Omega^m}{\Omega^m+\Omega^l}\frac{d\log\Omega^m}{d\log y} - \frac{\Omega^l}{\Omega^m+\Omega^l}\frac{d\log\Omega^l}{d\log y}.$$

First, we solve for $\frac{d\log\Omega^e}{d\log y}$ as:

$$\frac{d\log\Omega^e}{d\log y} = (1-\sigma_e)(1-\Omega^{e,elk})(\epsilon_e-\epsilon_l) + \Omega^m(1-\sigma_m)\left[\epsilon_e\Omega^{e,elk} + \epsilon_l(1-\Omega^{e,elk}) - \epsilon_m\right].$$

Second, we solve for $\frac{d\log\Omega^m}{d\log y}$ as:

$$\frac{d\log\Omega^m}{d\log y} = -(1-\Omega^m)(1-\sigma_m)\left[\epsilon_e\Omega^{e,elk} + \epsilon_l(1-\Omega^{e,elk}) - \epsilon_m\right].$$

Third, we solve for $\frac{d\log\Omega^l}{d\log y}$ as:

$$\frac{d\log\Omega^l}{d\log y} = \Omega^m(1-\sigma_m)\left[\epsilon_e\Omega^{e,elk} + \epsilon_l(1-\Omega^{e,elk}) - \epsilon_m\right] - \Omega^{e,elk}(1-\sigma_e)(\epsilon_e-\epsilon_l).$$

Substituting in the previous expressions yields

$$\begin{aligned}\frac{d\log\left(\frac{\Omega^e}{\Omega^l+\Omega^m}\right)}{d\log y} &= (1-\sigma_e)(1-\Omega^{e,elk})(\epsilon_e-\epsilon_l)+\Omega^m(1-\sigma_m)\left[\epsilon_e\Omega^{e,elk}+\epsilon_l(1-\Omega^{e,elk})-\epsilon_m\right]\\ &+\frac{\Omega^m}{\Omega^m+\Omega^l}\left[(1-\Omega^m)(1-\sigma_m)\left[\epsilon_e\Omega^{e,elk}+\epsilon_l(1-\Omega^{e,elk})-\epsilon_m\right]\right]\\ &-\frac{\Omega^l}{\Omega^m+\Omega^l}\left[\Omega^m(1-\sigma_m)\left[\epsilon_e\Omega^{e,elk}+\epsilon_l(1-\Omega^{e,elk})-\epsilon_m\right]-\Omega^{e,elk}(1-\sigma_e)(\epsilon_e-\epsilon_l)\right].\end{aligned}$$

Simplifying, we obtain:

$$\eta_{e/ml}\equiv\frac{d\log\left(\frac{\Omega^e}{\Omega^l+\Omega^m}\right)}{d\log y}=A\left[1-\frac{\Omega^m\Omega^{e,elk}}{\Omega^m+\Omega^l}\right]+B\left[\frac{\Omega^m}{\Omega^m+\Omega^l}\right],$$

where

$$A=(1-\sigma_e)(\epsilon_e-\epsilon_l),\qquad B=(1-\sigma_m)\left[\epsilon_e\Omega^{e,elk}+\epsilon_l(1-\Omega^{e,elk})-\epsilon_m\right].$$

**Energy to Labor: $\eta_{e/l}$.** Define

$$\eta_{e/l}\equiv\frac{d\log\left(\frac{\Omega^e}{\Omega^l}\right)}{d\log y}=\frac{d\log(\Omega^e)}{d\log y}-\frac{d\log(\Omega^l)}{d\log y}.$$

Plugging in our solved expressions, we have

$$\eta_{e/l}\;=\;(1-\sigma_e)\,(\epsilon_e-\epsilon_l)$$

Note that the analytical expression for $\eta_{e/lk}$ is identical to that of $\eta_{e/l}$. In our empirical estimates we find that both scale elasticities are insignificant and thus we set $\eta_{e/l}=0$ in our calibration which implies that $\epsilon_e=\epsilon_l$.

**Cost Elasticity: $\mathcal{E}$.** Taking logs of the cost function and differentiating, holding input prices fixed, we have

$$\frac{d\log\mathcal{C}}{d\log y}=\Omega^m\epsilon_m+(1-\Omega^m)(\Omega^{e,elk}\epsilon_e+(1-\Omega^{e,elk})\epsilon_l)$$

**Analytical Calibration of $\epsilon_m, \epsilon_e, \epsilon_l$** Given these formulas for our scale elasticities, we can solve for $\epsilon_e$, and $\epsilon_l$ in closed form:

$$\epsilon_m = \mathcal{E} - (1-\Omega^m)\frac{\eta_{e/ml} - \eta_{e/l} + \Omega^{m,ml}\Omega^{e,elk}\eta_{e/l}}{(1-\sigma_m)\Omega^{m,ml}}, \tag{SM.24}$$

$$\epsilon_l = \mathcal{E} + \Omega^m\frac{\eta_{e/ml} - \eta_{e/l} + \Omega^{m,ml}\Omega^{e,elk}\eta_{e/l}}{(1-\sigma_m)\Omega^{m,ml}} - \Omega^{e,elk}\frac{\eta_{e/l}}{1-\sigma_e}, \tag{SM.25}$$

$$\epsilon_e = \mathcal{E} + \Omega^m\frac{\eta_{e/ml} - \eta_{e/l} + \Omega^{m,ml}\Omega^{e,elk}\eta_{e/l}}{(1-\sigma_m)\Omega^{m,ml}} + (1-\Omega^{e,elk})\frac{\eta_{e/l}}{1-\sigma_e}. \tag{SM.26}$$

Since $\eta_{e/l} = 0$ we set $\epsilon_l = \epsilon_e$. Given our empirical estimates of $\eta_{e/ml} = -0.51$, $\mathcal{E} = 0.96$, $\sigma_m = 0.3$, and cost shares of $\Omega^{e,elk}_{1990} = 0.315$, $\Omega^m_{1990} = 0.714, \Omega^l_{1990} = 0.06$ as measured in the data, we find that $\epsilon_m = 1.19$ and $\epsilon_l = \epsilon_e = 0.40$.

### SM.4.2 Elasticity of Substitution Between Energy and Other Inputs

To estimate the elasticity of substitution parameters $\sigma_m$ and $\sigma_e$ for our quantitative model we use the ASI data to estimate in the elasticity of relative energy demand in response to changes in the relative price of energy. We then back out the production function parameters from analytical expressions for these elasticities derived below.

We estimate two elasticities: (1) the elasticity of the energy-to-labor ratio to changes in energy prices, and (2) the elasticity of the energy-to-materials ratio to changes in energy prices. To understand how these estimates map to the model parameters $\sigma_m$ and $\sigma_e$, recall:

$$\Omega^e/\Omega^l \equiv \frac{w^e e}{w^l l} = \frac{A_e^{\sigma_e - 1}(w^e)^{1-\sigma_e} y^{(1-\sigma_e)\epsilon_e}}{\alpha (w^l)^{\alpha(1-\sigma_e)}(w^k)^{(1-\alpha)(1-\sigma_e)} y^{(1-\sigma_e)\epsilon_l}}$$

$$\Omega^e/\Omega^m \equiv \frac{w^e e}{w^m m} = \frac{A_e^{\sigma_e - 1} x^{\frac{\sigma_e - \sigma_m}{1-\sigma_e}} (w^e)^{1-\sigma_e} y^{(1-\sigma_e)\epsilon_e}}{A_m^{\sigma_m - 1}(w^m)^{1-\sigma_m} y^{(1-\sigma_m)\epsilon_m}}$$

where $x \equiv (A_e^{-1} w^e y^{\epsilon_e})^{1-\sigma_e} + ((w^l)^\alpha (w^k)^{1-\alpha} y^{\epsilon_l})^{1-\sigma_e}$. This implies

$$d\log\left(\Omega^e/\Omega^l\right) = (1-\sigma_e) d\log w^e \tag{SM.27}$$

$$d\log(\Omega^e/\Omega^m) = [(1-\sigma_e) + (\sigma_e - \sigma_m)\Omega^{e,elk}] d\log w^e \tag{SM.28}$$

holding fixed any changes in firm scale or in the prices of other inputs. We therefore estimate the following expressions

$$\Delta^h \log\left(\frac{\Omega^e_{it}}{\Omega^l_{it}}\right) = \alpha_{sct} + \beta^l \Delta^h \log w^e_{it} + u_{it} \tag{SM.29}$$

$$\Delta^h \log\left(\frac{\Omega^e_{it}}{\Omega^m_{it}}\right) = \alpha_{sct} + \beta^m \Delta^h \log w^e_{it} + u_{it} \tag{SM.30}$$

$\Delta^h \log w^e_{it}$ is the log change in the price of the firm-specific energy bundle, constructed as

the Törnqvist-weighted change in the observed firm×fuel-level price changes.

The key threat to identification is that firm-specific energy prices themselves respond to firm-level energy demand. To identify $\sigma_e$, we construct a supply shifter affecting $\Delta^h \log w^e_{it}$, but plausibly orthogonal to other shifts in firm-specific energy demand. We construct a shift-share instrument exploiting changes in the aggregate prices of different fuels $f \in \mathcal{F} = \{\text{coal}, \text{electricity}, \text{oil}\}$, as well as firm-level exposure to different fuels:

$$\mathcal{S}^{E,h}_{it} = \sum_{f \in \mathcal{F}} \omega^e_{fit} \, \Delta^h \log w_{ft}, \tag{SM.31}$$

where $\Delta^h \log w_{ft} = \log w_{ft+h} - \log w_{ft}$, and $\omega^e_{fit}$ is the share of fuel type $f$ in total energy expenses at time $t$ used by firm $i$. For oil, we construct the aggregate oil price index (see details in Supplementary Material SM.1). For electricity, given that electricity prices are set by state-level utilities, we use the change in the average state-level electricity price. For coal, we account for the fact that coal markets are highly local and use the district-level change in the price of coal. The identifying assumption is that firms most exposed to changes in aggregate energy prices due to their fuel mix are not systematically subject to other energy-specific demand or supply shocks. We check that this instrument is uncorrelated with the change in other input prices (price of the material bundle, wage) and with the demand shifter defined in (4).[12]

The results are summarized in Table SM.17. For our baseline model, we use the specification where we control for changes in firm scale instrumented by our demand shifter and changes in other input prices (columns (4) and (8)). Using these estimates for $\beta^l$ and $\beta^m$ and equations SM.27 and SM.28, we find $\sigma_e = 0.48, \sigma_m = 0.30$. Our quantitative model results are robust to using any of the specifications in Table SM.17.

### SM.4.3 Elasticity of the Cost Function With Respect to Output

By definition,

$$\text{RTS}_i^{-1} = \mathcal{E}_i = \frac{\partial \log \mathcal{C}_i(y_i, \mathbf{w}_i)}{\partial \log y_i} \tag{SM.32}$$

where the $i$ index on the cost function makes explicit that the mapping from $(y_i, \mathbf{w}_i)$ to costs may be firm-specific, and in particular depend on firm-level (physical) productivity.

We obtain the average cost elasticity in our sample by estimating the following regression model:

$$\Delta^h \log \mathcal{C}_{it} = \alpha_{sct} + \beta \, \Delta^h \log y_{it} + u_{it} \tag{SM.33}$$

where $\Delta^h \log y_{it}$ is instrumented by the demand shifter $\mathcal{D}^h_{it}$. The identifying assumption

[12] There is a small correlation of the energy price instrument with the wage, but as shown in Table SM.17, controlling for input price changes leaves the coefficients essentially unchanged.

TABLE SM.17: Estimation of the Elasticity of Substitution

| | $\Delta^5 \log(\Omega^e/\Omega^l)$ | | | | $\Delta^5 \log(\Omega^e/\Omega^m)$ | | | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| $\Delta^5 \log(w^e)$ | 0.496*** | 0.519*** | 0.506*** | 0.525*** | 0.508*** | 0.554*** | 0.558*** | 0.581*** |
| | (0.072) | (0.071) | (0.084) | (0.080) | (0.070) | (0.074) | (0.081) | (0.083) |
| Year × Ind FE | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| $\Delta^5 \log(w)$ controls | | ✓ | | ✓ | | ✓ | | ✓ |
| $\Delta^5 \log(y)$ controls | | | ✓ | ✓ | | | ✓ | ✓ |
| Observations | 204,202 | 203,858 | 186,451 | 186,133 | 204,279 | 157,718 | 186,516 | 150,548 |

*Notes:* This table presents the results of estimating the elasticity of substitution using different specifications of the relative energy share: (i) energy costs relative to labor costs or (ii) energy costs relative to materials costs. $\Delta^h \log(w^e)$ is instrumented using a shift-share instrument which exploits changes in the aggregate prices of different fuels. $\Delta^5 \log(w)$ controls indicates that the regression controls for the firm-level changes in input prices (the price of labor in $\Omega^e/\Omega^l$ regressions, and the price of labor and materials in the $\Omega^e/\Omega^m$ regressions). $\Delta^5 \log(y)$ controls indicates that the regression controls for the firm-level changes in output, instrumented with the firm-level demand shock defined in equation (4). Regressions are weighted by firm total cost (adjusted for the sampling weight). Standard errors are clustered at the firm and state level. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

is that $\mathcal{D}^h_{it}$ is orthogonal to unobserved determinants of the change in total costs. In particular, $\mathcal{D}^h_{it}$ must be orthogonal to: (i) the change in firm-level physical productivity, and (ii) the change in firm-level input prices. Note that this is a more stringent identifying assumption than that underlying the identification of the scale elasticity of energy demand, which only required that $\mathcal{D}^h_{it}$ be orthogonal to energy-specific productivity shocks. To alleviate these identification concerns, we include the change in firm-specific input prices and the change in physical productivity as controls.

Table SM.18 shows the results. We find an average cost-weighted elasticity close to 1. Table SM.19 investigates heterogeneity in the cost elasticity across firms. Our model predicts that smaller firms, with smaller material shares, have lower cost elasticities (higher returns to scale). To investigate this point, we compare the unweighted, cost-weighted and output-weighted average cost elasticity (Panel A) and estimate the cost elasticity splitting the sample by firm size or material share (Panel B). The evidence is in line with the model prediction.

TABLE SM.18: Estimation of the Cost Elasticity

| | $\Delta^5 \log(C)$ | | | | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| $\Delta^5 \log(y)$ | 0.943*** | 0.924*** | 1.007*** | 0.963*** | 0.987*** | 0.939*** |
| | (0.132) | (0.150) | (0.103) | (0.053) | (0.098) | (0.052) |
| Year × Ind. × Cohort FE | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| $\Delta^5 \log(w)$ controls | | ✓ | | ✓ | | ✓ |
| $\Delta^5 \log(\text{tfpq})$ controls | | | ✓ (1) | ✓ (1) | ✓ (2) | ✓ (2) |
| Observations | 138,147 | 114,185 | 114,081 | 114,075 | 114,081 | 114,075 |
| Returns to scale | 1.061 | 1.083 | 0.993 | 1.038 | 1.013 | 1.065 |

*Notes:* This table presents the results of estimating equation (SM.33). $\Delta^h \log(y)$ is instrumented with the firm-level demand shock defined in equation (4). Returns to scale are defined as the inverse of the estimated coefficient. $\Delta^5 \log(w)$ controls indicates that the regression controls for the firm-level changes in input prices (wage, energy price index, raw materials price index). $\Delta^5 \log(\text{tfpq})$ controls for the change in firm physical productivity, using the two different variants described in Appendix SM.1.4. Regressions are weighted by firm total cost (adjusted for the sampling weight). Standard errors are clustered at the firm and main product×year level. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

TABLE SM.19: Cost Elasticity: Heterogeneity Analysis

**Panel A:** Unweighted vs. Weighted Estimation

| | $\Delta^5 \log(\mathcal{C})$ | | | | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| $\Delta^5 \log(y)$ | 0.882*** | 0.887*** | 0.963*** | 0.939*** | 0.974*** | 0.969*** |
| | (0.042) | (0.045) | (0.053) | (0.052) | (0.054) | (0.055) |
| Year × Ind. × Cohort FE | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Weight | | | Cost | Cost | Output | Output |
| $\Delta^5 \log(w)$ controls | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| $\Delta^5 \log(\text{tfpq})$ controls | ✓ (1) | ✓ (2) | ✓ (1) | ✓ (2) | ✓ (1) | ✓ (2) |
| Observations | 114,075 | 114,075 | 114,075 | 114,075 | 114,075 | 114,075 |
| Returns to scale | 1.133 | 1.128 | 1.038 | 1.065 | 1.027 | 1.032 |

**Panel B:** Heterogeneity Across Samples

| | $\Delta^5 \log(\mathcal{C})$ | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | By material share | | | | By output | | | |
| | Low | | High | | Low | | High | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| $\Delta^5 \log(y)$ | 0.843*** | 0.850*** | 1.019*** | 0.970*** | 0.842*** | 0.785*** | 0.961*** | 0.935*** |
| | (0.093) | (0.096) | (0.069) | (0.063) | (0.101) | (0.091) | (0.059) | (0.057) |
| Year × Ind. × Cohort FE | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| $\Delta^5 \log(w)$ controls | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| $\Delta^5 \log(\text{tfpq})$ controls | ✓ (1) | ✓ (2) | ✓ (1) | ✓ (2) | ✓ (1) | ✓ (2) | ✓ (1) | ✓ (2) |
| Observations | 55,455 | 55,455 | 56,000 | 56,000 | 53,697 | 53,697 | 57,728 | 57,728 |
| Returns to scale | 1.186 | 1.176 | 0.982 | 1.031 | 1.188 | 1.274 | 1.040 | 1.070 |

*Notes:* This table presents the results of estimating equation (SM.33). $\Delta^h \log(y)$ is instrumented with the firm-level demand shock defined in equation (4). Returns to scale are defined as the inverse of the estimated coefficient. In Panel A, we estimate this specification unweighted (only sampling weights), weighted by firm total cost, and weighted by firm sales. In Panel B, we estimate this specification separately for firms below and above the within-industry median material share (sales, respectively). $\Delta^5 \log(w)$ controls indicates that the regression controls for the firm-level changes in input prices (wage, energy price index, raw materials price index). $\Delta^5 \log(\text{tfpq})$ controls for the change in firm physical productivity, using the two different variants described in Appendix SM.1.4. Standard errors are clustered at the firm and main product×year level. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

### SM.4.4 Energy Intensity and Expenditure Share Decompositions

Each firm $a$ is characterized by its physical (revenue) energy intensity and its energy cost share,

$$\Theta_t(a) \equiv \frac{e_t(a)}{p_t(a)\, y_t(a)}, \qquad \Omega^e_t(a) \equiv \frac{w^e_t\, e_t(a)}{C_t(a)},$$

with aggregate counterparts $\bar{\Theta}_t \equiv E_t / Y_t$ and $\Omega^e_t \equiv \left(\int w^e_t e_t(a)\, d\Gamma\right) / \left(\int C_t(a)\, d\Gamma\right)$. All cost objects are taken from the quantitative non-homothetic CES cost function (19). We write $\Omega^m(a)$, $\Omega^e(a)$, $\Omega^l(a)$, and $\Omega^k(a)$ for its factor cost shares and $\Omega^{e,elk}(a) \equiv \Omega^e(a)/(\Omega^e(a) + \Omega^l(a) + \Omega^k(a))$ for the energy share within the energy–labor–capital bundle. Writing $r(a) \equiv p(a)y(a)/(Y_t/N_t)$ for the revenue weight, the revenue-weighted and intensity-weighted expectations of a firm-level object $X(a)$ are

$$\mathbb{E}_R[X] \equiv \int X(a)\, r(a)\, d\Gamma(a), \qquad \mathbb{E}^{\Theta}_R[X] \equiv \mathbb{E}_R\left[\frac{\Theta(a)}{\bar{\Theta}}\, X(a)\right]$$

Here $\Gamma$ is the unit-mass productivity distribution from which firms draw and $N_t$ is the mass of firms, so that $Y_t = N_t \int p(a)y(a)\, d\Gamma(a)$; hence $Y_t/N_t$ is revenue per firm and the revenue weight integrates to one, $\int r(a)\, d\Gamma(a) = 1$. Since $\mathbb{E}_R[\Theta/\bar{\Theta}] = 1$, the intensity weight has unit mean, and the gap between the two expectations is a covariance,

$$\mathbb{E}^{\Theta}_R[X] = \mathbb{E}_R[X] + \mathrm{Cov}_R\left(\frac{\Theta(a)}{\bar{\Theta}}, X\right), \tag{SM.34}$$

an identity that will isolate the reallocation channel below.

Because $\Theta(a) = \Omega^e(a)\, C(a) / \left(w^e_t\, p(a)y(a)\right)$, the firm-level intensity moves according to

$$d\log\Theta(a) = d\log\Omega^e(a) + d\log C_t(a) - d\log w^e_t - d\log\left(p(a)y(a)\right) \tag{SM.35}$$

Differentiating the cost function (19) and applying Shephard's lemma gives the two required micro differentials,

$$\begin{aligned} d\log C_t(a) = &- d\log A_{Yt} + \mathcal{E}(a)\, d\log y(a) \\ &+ \sum_{x\in\{m,e,l,k\}} \Omega^x(a)\, d\log w^x_t - \Omega^e(a)\, d\log A_{et} - \Omega^m(a)\, d\log A_{mt} \end{aligned} \tag{SM.36}$$

$$\begin{aligned} d\log \Omega^e_t(a) = &\ \phi_y(a)\, d\log y(a) \\ &+ \sum_{x\in\{e,l,k,m\}} \phi_x(a)\, d\log w^x_t + \phi_{A_e}(a)\, d\log A_{et} + \phi_{A_m}(a)\, d\log A_{mt} \end{aligned} \tag{SM.37}$$

where $\mathcal{E}(a) \equiv \Omega^m(a)\epsilon_m + \left(1-\Omega^m(a)\right)\left[\Omega^{e,elk}(a)\epsilon_e + \left(1-\Omega^{e,elk}(a)\right)\epsilon_l\right]$ is the firm cost elasticity and the coefficients $\phi_x(a)$ are defined in SM.4.4.3 below.

### SM.4.4.1 Energy Intensity Decomposition

Substituting (SM.36)–(SM.37) into (SM.35) and collecting the coefficients on each factor price into the firm-level share–substitution coefficients $\beta_m(a), \beta_e(a), \beta_l(a)$ yields

$$\begin{aligned} d\log\Theta(a) = &\left[\beta_m(a)\epsilon_m + \beta_e(a)\epsilon_e + \beta_l(a)\epsilon_l - 1 + \tfrac{1}{\lambda}\right] d\log y(a) - d\log\left(A_{Yt}Y_t^{1/\lambda}\right) \\ &+ \beta_m(a)\, d\log w_t^m + (\beta_e(a)-1)\, d\log w_t^e + \beta_l(a)\left[\alpha\, d\log w_t^l + (1-\alpha)\, d\log w_t^k\right] \\ &- \beta_m(a)\, d\log A_{mt} - \beta_e(a)\, d\log A_{et} \end{aligned} \tag{SM.38}$$

where each coefficient is a cost-share-weighted combination of the substitution elasticities,

$$\begin{aligned} \beta_m(a) &= \Omega^m(a)\sigma_m, \\ \beta_e(a) &= \Omega^{e,elk}(a)\left(1-\Omega^m(a)\sigma_m\right) + \left(1-\Omega^{e,elk}(a)\right)(1-\sigma_e), \\ \beta_l(a) &= \left(1-\Omega^{e,elk}(a)\right)\left(\sigma_e - \Omega^m(a)\sigma_m\right), \end{aligned}$$

which satisfy $\beta_m(a)+\beta_e(a)+\beta_l(a)=1$. Writing the scale coefficient in terms of the micro non-homothetic scale elasticity of energy intensity, $\tilde{\epsilon}_Q(a) \equiv \beta_m(a)(\epsilon_m - 1) + \beta_e(a)(\epsilon_e - 1) + \beta_l(a)(\epsilon_l - 1)$, the coefficient on $d\log y(a)$ in (SM.38) is $\tilde{\epsilon}_Q(a) + \frac{1}{\lambda}$.

To aggregate, inverse demand implies $d\log(p(a)y(a)) = (1-\frac{1}{\lambda})d\log y(a) + \frac{1}{\lambda}\, d\log Y_t$, so the revenue weight evolves as $d\log r(a) = (1-\frac{1}{\lambda})d\log y(a) + \frac{1}{\lambda}\, d\log Y_t - d\log(Y_t/N_t)$. As $\bar{\Theta}$ is the revenue-weighted average of $\Theta(a)$, its change equals the intensity-weighted average of the within-firm change and the change in weights,

$$d\log\bar{\Theta} = \mathbb{E}_R^{\Theta}\left[d\log\Theta(a) + d\log r(a)\right]$$

Using the demand curve and rearranging,

$$\begin{aligned} d\log\bar{\Theta} = &\,\mathbb{E}_R^{\Theta}\left[\tilde{\epsilon}_Q(a)\, d\log y(a)\right] + \left(\mathbb{E}_R^{\Theta}\left[d\log y(a)\right] - d\log(Y_t/N_t)\right) - d\log A_{Yt} \\ &+ \mathbb{E}_R^{\Theta}[\beta_m(a)]\, d\log w_t^m + \left(\mathbb{E}_R^{\Theta}[\beta_e(a)]-1\right) d\log w_t^e + \mathbb{E}_R^{\Theta}[\beta_l(a)]\left[\alpha\, d\log w_t^l + (1-\alpha)\, d\log w_t^k\right] \\ &- \mathbb{E}_R^{\Theta}[\beta_m(a)]\, d\log A_{mt} - \mathbb{E}_R^{\Theta}[\beta_e(a)]\, d\log A_{et} \end{aligned} \tag{SM.39}$$

Finally, applying identity (SM.34) to the two intensity-weighted scale terms,

$$\begin{aligned} d\log\bar{\Theta} = &\underbrace{\mathbb{E}_R\left[\tilde{\epsilon}_Q(a)\, d\log y(a)\right]}_{\text{Micro Scale Up}} \\ &+ \underbrace{\left(\mathbb{E}_R\left[d\log y(a)\right] - d\log(Y_t/N_t)\right)}_{\text{Aggregate Returns to Scale}} + \underbrace{\mathrm{Cov}_R\left[\frac{\Theta(a)}{\bar{\Theta}}, \left(\tilde{\epsilon}_Q(a)+1\right)\, d\log y(a)\right]}_{\text{Reallocation}} \\ &+ \underbrace{\mathbb{E}_R^{\Theta}[\beta_m(a)]d\log w_t^m + \left(\mathbb{E}_R^{\Theta}[\beta_e(a)]-1\right) d\log w_t^e + \mathbb{E}_R^{\Theta}[\beta_l(a)]\left[\alpha d\log w_t^l + (1-\alpha)d\log w_t^k\right]}_{\text{Factor Price Changes}} \\ &- \underbrace{\mathbb{E}_R^{\Theta}[\beta_m(a)]d\log A_{mt} - \mathbb{E}_R^{\Theta}[\beta_e(a)]d\log A_{et} - d\log A_{Yt}}_{\text{Productivity Changes}}. \end{aligned} \tag{SM.40}$$

#### SM.4.4.2 The Micro Non-homothetic Scale Elasticity

The scale elasticity $\tilde{\epsilon}_Q(a)$ is the elasticity of physical energy intensity to firm output, holding prices and productivities fixed,

$$\tilde{\epsilon}_Q(a) \equiv \left.\frac{\partial \log\left(e(a)/y(a)\right)}{\partial \log y(a)}\right|_{w,A}$$

Since $e(a)/y(a) = \Omega^e(a)C_t(a)/\left(w_t^e y(a)\right)$, this elasticity splits into an energy-share and a cost component,

$$\tilde{\epsilon}_Q(a) = \left.\frac{\partial \log \Omega^e(a)}{\partial \log y(a)}\right|_{w,A} + \left.\frac{\partial \log C_t(a)}{\partial \log y(a)}\right|_{w,A} - 1 \tag{SM.41}$$

For the quantitative cost function the cost component is the cost elasticity $\mathcal{E}(a)$ introduced above, while the energy-share component is

$$\begin{aligned}\left.\frac{\partial \log \Omega^e(a)}{\partial \log y(a)}\right|_{w,A} =& \Omega^m(a)(1-\sigma_m)\left[\Omega^{e,elk}(a)\epsilon_e + (1-\Omega^{e,elk}(a))\epsilon_l - \epsilon_m\right] \\ &+ (1-\Omega^{e,elk}(a))(1-\sigma_e)(\epsilon_e - \epsilon_l)\end{aligned}$$

Adding these and subtracting one recovers the weighted form used above,

$$\tilde{\epsilon}_Q(a) = \beta_m(a)(\epsilon_m - 1) + \beta_e(a)(\epsilon_e - 1) + \beta_l(a)(\epsilon_l - 1), \tag{SM.42}$$

with the same $\beta_m(a), \beta_e(a), \beta_l(a)$ as in (SM.38). This is the quantitative-model analogue of the single primitive $\tilde{\epsilon}$ of the simple model: it replaces $\tilde{\epsilon}$ by a share-weighted combination of the three scale elasticities, and reduces to zero when all input requirements scale proportionally with output, $\epsilon_m = \epsilon_e = \epsilon_l = 1$.

#### SM.4.4.3 Energy Expenditure Share Decomposition

**Aggregate definition.** The aggregate energy cost share is the cost-weighted average of the micro shares,

$$\Omega_t^e = \frac{\int w_t^e e_t(a)\, d\Gamma(a)}{\int C_t(a)\, d\Gamma(a)} = \int \Omega_t^e(a)\, \frac{C_t(a)}{\bar{C}_t}\, d\Gamma(a), \qquad \bar{C}_t \equiv \int C_t(a)\, d\Gamma(a) \tag{SM.43}$$

Define the cost-weighted and energy-cost-weighted expectations

$$\mathbb{E}_C[X] \equiv \int \frac{C(a)}{\bar{C}}\, X(a)\, d\Gamma(a), \qquad \mathbb{E}_C^{\Omega}[X] \equiv \frac{1}{\Omega^e}\int \Omega^e(a)\, \frac{C(a)}{\bar{C}}\, X(a)\, d\Gamma(a),$$

for which $\mathbb{E}_C[\Omega^e(a)/\Omega^e] = 1$ and hence the cost-weight analogue of (SM.34), $\mathbb{E}_C^{\Omega}[X] = \mathbb{E}_C[X] + \mathrm{Cov}_C(\Omega^e(a)/\Omega^e, X)$, holds. Differentiating (SM.43) and dividing by $\Omega_t^e$ gives

$$d\log\Omega_t^e = \mathbb{E}_C^{\Omega}\big[d\log\Omega_t^e(a)\big] + \mathbb{E}_C^{\Omega}\Big[d\log\tfrac{C_t(a)}{\bar{C}_t}\Big] \tag{SM.44}$$

**Micro energy cost-share change.** The micro differential $d\log\Omega^e(a)$ is (SM.37), with coefficients

(SM.45)

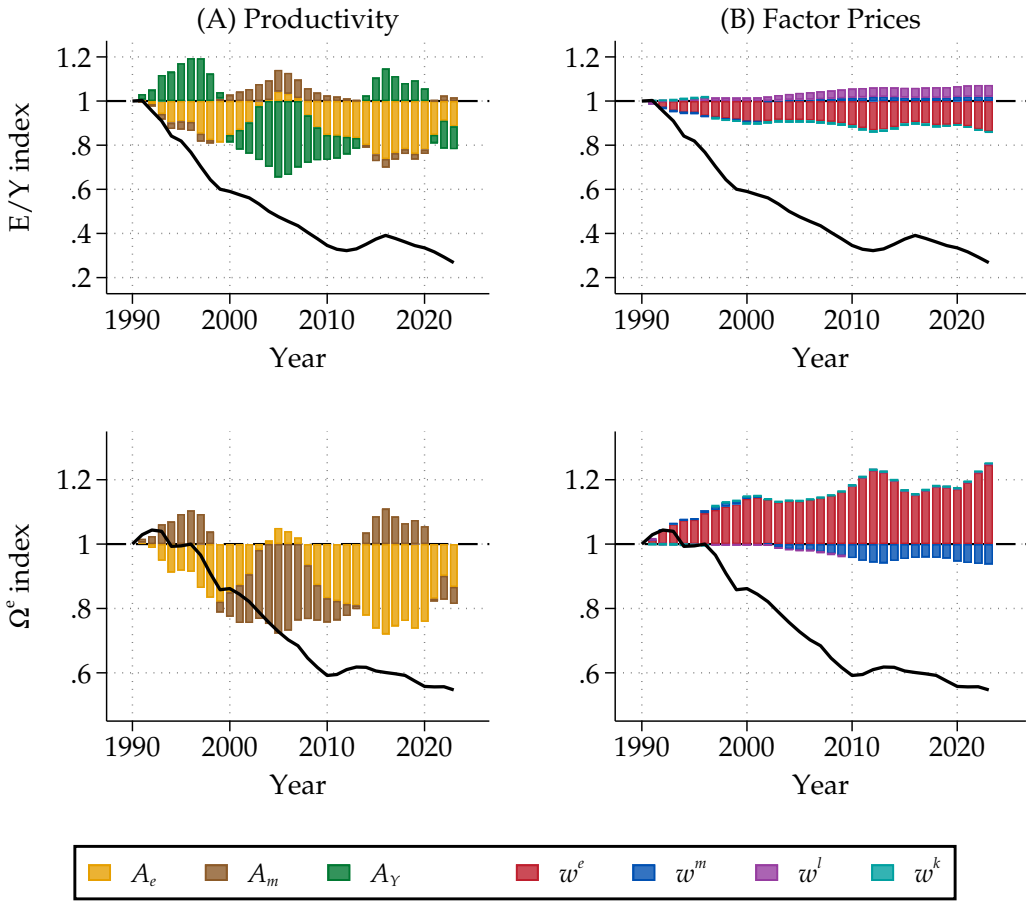
$$\begin{aligned}
\phi_y(a) &= (1-\sigma_e)(1-\Omega_t^{e,elk}(a))(\epsilon_e-\epsilon_l) + (1-\sigma_m)\Omega_t^m(a)\big[\Omega_t^{e,elk}(a)\epsilon_e + (1-\Omega_t^{e,elk}(a))\epsilon_l - \epsilon_m\big],\\
\phi_e(a) &= (1-\sigma_e)(1-\Omega_t^{e,elk}(a)) + (1-\sigma_m)\Omega_t^m(a)\Omega_t^{e,elk}(a),\\
\phi_l(a) &= \alpha\big[-(1-\sigma_e)(1-\Omega_t^{e,elk}(a)) + (1-\sigma_m)\Omega_t^m(a)(1-\Omega_t^{e,elk}(a))\big],\\
\phi_k(a) &= (1-\alpha)\big[-(1-\sigma_e)(1-\Omega_t^{e,elk}(a)) + (1-\sigma_m)\Omega_t^m(a)(1-\Omega_t^{e,elk}(a))\big],\\
\phi_m(a) &= -(1-\sigma_m)\Omega_t^m(a),\\
\phi_{A_e}(a) &= -(1-\sigma_e)(1-\Omega_t^{e,elk}(a)) - (1-\sigma_m)\Omega_t^m(a)\Omega_t^{e,elk}(a),\\
\phi_{A_m}(a) &= (1-\sigma_m)\Omega_t^m(a)
\end{aligned}$$

**Aggregate decomposition.** Substituting (SM.37) into (SM.44) and applying the cost-weight covariance identity to the scale and reallocation terms yields

$$\begin{aligned}
d\log\Omega_t^e = {} & \underbrace{\mathbb{E}_C[\phi_y(a)]\,\mathbb{E}_C[d\log y(a)] + \mathrm{Cov}_C(\phi_y(a), d\log y(a))}_{\text{micro scale-up}}\\
& + \underbrace{\mathbb{E}_C^{\Omega}\Big[d\log\tfrac{C_t(a)}{\bar{C}_t}\Big] + \mathrm{Cov}_C\Big(\tfrac{\Omega_t^e(a)}{\Omega_t^e}, \phi_y(a)\,d\log y(a)\Big)}_{\text{reallocation}}\\
& + \underbrace{\sum_{x\in\{e,l,k,m\}} \mathbb{E}_C^{\Omega}[\phi_x(a)]\,d\log w_t^x}_{\text{factor prices}} + \underbrace{\mathbb{E}_C^{\Omega}[\phi_{A_e}(a)]\,d\log A_{et} + \mathbb{E}_C^{\Omega}[\phi_{A_m}(a)]\,d\log A_{mt}}_{\text{factor-specific productivity}}.
\end{aligned} \tag{SM.46}$$

Because $\Omega_t^e$ is a ratio of costs to costs, both the aggregate-scale term and the factor-neutral productivity ($A_{Yt}$) term cancel from the share and are absent from (SM.46).

FIGURE SM.17: Contribution of Productivity and Factor Prices to the Decline in Energy Intensity and Cost Shares




*Notes:* The figure further decomposes the decline in energy intensity due to productivity shifters $A_e, A_m, A_Y$ (Panel A) and factor prices $w^e, w^m, w^l, w^k$ (Panel B). The E/Y index decomposition in row 1 is derived in Section 5.3, and the energy cost share decomposition in row 2 is derived in equation SM.46. The data series for energy intensity and energy cost shares are represented by the black lines and are indexed to 1 in 1990.

# SUPPLEMENTARY MATERIAL REFERENCES


BILS, M., P. J. KLENOW, AND C. RUANE (2021): "Misallocation or mismeasurement?" *Journal of Monetary Economics*, 124, S39–S56.

BOEHM, J., S. DHINGRA, AND J. MORROW (2022): "The Comparative Advantage of Firms," *Journal of Political Economy*, 130, 3025–3100.

BORUSYAK, K., P. HULL, AND X. JARAVEL (2022): "Quasi-Experimental Shift-Share Research Designs," *The Review of Economic Studies*, 89, 181–213.

BUREAU OF ENERGY EFFICIENCY (2017): "Energy Efficiency Interventions in MSME Clusters: Annexure 1," Tech. rep., Ministry of Power, Government of India.

——— (2026): "Perform, Achieve and Trade (PAT)," https://beeindia.gov.in, Ministry of Power, Government of India, accessed: 2026-06-23.

IEA (2023): "Energy Conservation Act," https://www.iea.org/policies/1975-energy-conservation-act, IEA Policies Database, accessed: 2026-06-23.

——— (2025): "Energy Demand in Road Transport, by Mode and Vehicle Type, in Advanced and Emerging Economies, 2023," International Energy Agency, Paris. Licence: CC BY 4.0.

MARTIN, L. A., S. NATARAJ, AND A. E. HARRISON (2017): "In with the Big, Out with the Small: Removing Small-Scale Reservations in India," *American Economic Review*, 107, 354–86.

MINISTRY OF POWER, GOVERNMENT OF INDIA (2022): "Status of Implementation of National Mission for Enhanced Energy Efficiency (NMEEE)," Press Release, accessed April 2026.

PETROLEUM PLANNING & ANALYSIS CELL (PPAC) (2014): "All India Study on Sectoral Demand of Diesel & Petrol," Tech. rep., Ministry of Petroleum and Natural Gas, Government of India, accessed 2025-09-04.

RIJESH, R. (2023): "Concordance of Manufacturing Industries Using NIC-2008 and NIC-2004: A Disaggregate Level Study," Tech. rep.

SHAKTI SUSTAINABLE ENERGY FOUNDATION (2017): "Energy Efficiency in MSMEs: Final Project Report," Tech. rep., Shakti Sustainable Energy Foundation.

UNIDO (2022): "Evaluation Report on Promoting Energy Efficiency and Renewable Energy in Selected MSME Clusters in India," Tech. rep., United Nations Industrial Development Organization.

WORLD BANK (2019): "India: Financing Energy Efficiency at MSMEs Project," Tech. Rep. ICR00004902, World Bank Group.