INTERNATIONAL MONETARY FUND

Mitigating Climate Change at the Firm Level: Mind the Laggards

Damien Capelle, Divya Kirti, Nicola Pierri, and Germán Villegas Bauer

WP/23/242

2023
NOV

WORKING PAPER

WP/23/242

IMF Working Paper
Research Department and Fiscal Affairs Department

Mitigating Climate Change at the Firm Level: Mind the Laggards

Prepared by Damien Capelle, Divya Kirti, Nicola Pierri, and Germán Villegas Bauer

Authorized for distribution by Maria Soledad Martinez Peria
November 2023

ABSTRACT: Using self-reported data on emissions for a global sample of 4,000 large, listed firms, we document large heterogeneity in environmental performance within the same industry and country. Laggards—firms with high emissions relative to the scale of their operations—are larger, operate older physical capital stocks, are less knowledge intensive and productive, and adopt worse management practices. To rationalize these findings, we build a novel general equilibrium heterogeneous-firm model in which firms choose capital vintages and R&D expenditure and hence emissions. The model matches the full empirical distribution of firm-level heterogeneity among other moments. Our counter-factual analysis shows that this heterogeneity matters for assessing the macroeconomic costs of mitigation policies, the channels through which policies act, and their distributional effects. We also quantify the gains from technology transfers to EMDEs.

JEL Classification Numbers:	Q55, H23, O44, G30
Keywords:	Climate Change; Productivity; Technology Adoption; Capital Vintages; Emissions
Author's E-Mail Address:	dcapelle@imf.org, dkirti@imf.org, npierri@imf.org, gvillegasbauer@imf.org

---

* We thank Ariadne Maridena Checo de los Santos for outstanding research assistance and Maria Soledad Martinez Peria, Benjamin Carton, Hasan Dudu, Nicolas Fernandez-Arias, Alan Finklestein Shapiro, Clara Galeazzi, Dora Iakova, Florence Jaumotte, Pedro Juarros, Yuko Kinoshita, Yueran Ma, Chris Papageorgiou, Christine Richmond, Andinet Woldemichael, and seminar participants at IMF for their insightful comments.

WORKING PAPERS

Mitigating Climate Change at the Firm Level: Mind the Laggards

Prepared by Damien Capelle, Divya Kirti, Nicola Pierri, and Germán Villegas Bauer

1 Introduction

Sharp reductions in greenhouse gas emissions are needed to mitigate climate change. Achieving these cuts will require significant improvements in efficiency, including in how firms combine energy and other inputs to produce goods and services. Indeed, policymakers around the world have been actively exploring the role that carbon pricing and other policies like subsidies and technology transfers can play in supporting these technological improvements (Georgieva, 2021; Li et al., 2022; Kammer, 2023; IMF, 2023). Careful assessment of these policies requires better understanding the investment decisions that shape emissions at the firm level.

To assess the scope for efficiency gains, we first document striking heterogeneity in environmental performance across firms. Drawing on self-reported data on greenhouse gas emissions for a global sample of about 4,200 large, listed firms, we find large heterogeneity in emissions per dollar of revenues (emission intensities), even within firms operating in the same industry and country.[1] Emission intensities for firms at the 90th percentile of the within-country and industry distribution are more than six times larger than for firms at the 10th percentile. Building on the prominent literature examining productivity differences across firms (Syverson, 2011), we show that the dispersion of emission intensities is similar in magnitude to the dispersion of labor productivity and significantly larger than the dispersion of Total Factor Productivity (TFP) in the same sample.

This heterogeneity raises several questions. Why does environmental performance vary so much across firms? How much scope is there for climate 'laggards'—firms with high emission intensities relative to industry-country peers—to improve? Which policies can support this transition and at what cost?

To shed light on these questions, we examine the drivers of firms' environmental performance both empirically and theoretically. We present new evidence across countries in different income groups on the crucial role of technological adoption and innovation in shaping firms' climate footprint. We use this empirical evidence to inform a novel heterogeneous-firm general equilibrium model that can help evaluate a wide range of climate policies. The impact of policies on climate laggards shapes both overall policy effectiveness and distributional

---

[1] We focus on a sample of firms in manufacturing, transportation, and services for which Scope 1 and 2 emissions were 5.8 Gigatons of Carbon (GtC) in 2020, or about 16 percent of global emissions.

effects and therefore matters for the relative attractiveness of policy options.

Across our global sample of firms—covering both advanced economies (AEs) and emerging and developing economies (EMDEs)—climate laggards are slower to adopt new technologies and less innovative relative to firms at the within-industry frontier of environmental performance. They operate older physical capital stocks, suggesting that older capital vintages generate more emissions. They also have lower intangible capital shares and spend less on research and development (R&D). Technology therefore appears to be a key factor that shapes performance. Climate laggards are less effective on other dimensions as well—they are less productive and profitable. Importantly, we find similar results when we use different industry classification approaches, including text-based approaches drawing on self-identified competitors (Hoberg and Phillips, 2016), or different levels of industry granularity.

Firms with better environmental performance therefore appear to be firms with more efficient production processes. Both adoption of existing frontier technologies and innovation have a role to play in reducing emissions. Again, we do not find evidence that greener firms are less profitable. Instead, a common set of drivers appears to explain environmental performance, productivity, and profitability. Additional evidence on management practices (Bloom and Van Reenen, 2007) for a subsample of US manufacturers points in the same direction. Firms that adopt better management practices also have lower emission intensities. These stylized facts can help discipline a range of models that study climate change mitigation.

We explore the causal relationship between technological adoption and innovation and environmental performance through the lens of an instrumental variable approach. We use R&D cost differentials induced by differences in tax credits across US states (Lucking, 2019) to instrument for R&D expenditure of US manufacturers. We use the pace of firm growth over the previous five years to instrument for the age of capital stock. Results from these analyses point to a causal interpretation of our empirical findings.

We propose a granular, general equilibrium heterogeneous-firm model to rationalize our empirical findings and assess policy options and tradeoffs in mitigating climate change at the firm level. Each industry consists of a distribution of firms that differ in the capital vintage they use, their knowledge intensity, and consumers' taste for their variety of goods. Each period, firms may adopt better capital vintages—newer vintages are more efficient and generate less emissions per unit of output. Updating vintages entails substantial costs due to a difficulty of mixing different vintages, compounded by a challenge of reselling legacy capital, which exhibits a high degree of firm-specificity. Firms also choose how much to invest in R&D to accumulate knowledge and benefit from positive knowledge externalities from innovation at firms in the same industry. Capital vintage and knowledge intensity shape firms' productivity levels and the scale of energy and labor uses, ultimately determining emissions. Our modeling approach adds to past work (Shapiro and Walker, 2018; Shapiro and Metcalf, 2023) by jointly considering the role of capital, labor and innovation in driving energy consumption and emissions in a tractable general equilibrium model.

This rich conceptual framework replicates our main empirical findings and formalizes the drivers of environmental and financial performance. In the model, firms that rely on older vintages of physical capital and innovate less have higher emission intensities. These climate laggards are also less productive and less profitable. The same forces—efficiency of physical capital and innovation at the firm level—drive both environmental performance and profitability.

To derive quantitative assessments of mitigation policies, we calibrate the model to aggregate, industry-level and firm-level moments. At the country-industry level, the model matches input intensities in production, emission intensities, as well as the elasticity of emissions and the elasticity of capital productivity to vintage age. At the firm-level, it matches the distribution of heterogeneity in size, knowledge intensity, and capital vintages. This granularity allows us to provide a detailed assessment of different mitigation policies.

We consider the impact of four policy instruments. Two instruments target specific inputs—subsidies for upgrading capital or for R&D. The other two target emissions directly, either via a carbon tax proportional to emissions at the firm level or through a feebate system, which combines a tax proportional to emissions and a rebate proportional to a firm's output. For comparability, we calibrate each policy to reduce total corporate emissions by 15 percent.

Four key insights emerge from our counterfactual analysis. First, mitigation policies differ in the extent to which they reduce the dispersion in environmental performance. More than any other policy option we consider, capital upgrade subsidies lead a large fraction of firms to update their vintages, significantly reducing dispersion across firms. In contrast, carbon taxes incentivize firms to act on all margins, including reducing energy consumption and conduct more research, which leads to less reduction in dispersion across firms, while research subsidies have little effect on dispersion across firms.

Second, subsidies for capital upgrades and R&D can help reduce emissions but at significantly larger macroeconomic costs and with different intertemporal tradeoffs relative to carbon taxes and feebates. Subsidies are too narrow in scope, don't give direct incentives to economize on energy, and distort the allocation of inputs. Although capital upgrade subsidies reduce dispersion across firms, when discounting at market rates we find that this reduction comes at high cost. Indeed, we find that while carbon taxes needed to cut emissions by 15 percent reduce the net present value (NPV) of global consumption by 2 percent, subsidies for capital upgrades and R&D respectively generate NPV reductions of 12.5 and 4.5 percent.[2] However, subsidies for capital upgrades generate long-run consumption gains—unlike carbon taxes and R&D subsidies—because firms that upgrade become both greener and more productive. Subsidies for capital upgrades may be attractive when the social planner's discount rate is significantly lower than the market interest rate (Campbell and Martin, 2023).

Third, the macroeconomic costs—and long-term benefits—of capital subsidies depend on the initial distribution of vintages. When many firms use vintages far from the frontier, subsidies that nudge a small fraction to upgrade—and therefore that are less costly—can generate significant emission cuts. Relatedly, differences in the initial distribution of vintages and capital intensity across countries also shape the long-term gains of capital subsidies, and therefore the relative attractiveness of policy options across countries. Our results thus highlight the value of models that take into account the heterogeneity of capital vintages across firms and that are calibrated to each specific country.

Fourth, climate mitigation policies have important distributional effects—both across firms, and between consumers and firms—with implications for political economy considerations. A key finding is that distributional effects vary across policy instruments. Consistent with our results on consumption, subsidies lead to a larger decrease in the net present value of profits than carbon taxes, due to the general equilibrium reduction in output. While carbon feebates induce emission cuts at the firm level with the same effectiveness as carbon taxes, profits decrease by less at the median firm—and they therefore may be more politically palatable. In addition, policies affect different firms differently. In particular, firms with older capital vintages are hit harder. This is intuitive with carbon taxes—these firms emit more. But

---

[2] In our model more ambitious emission reduction targets do not change this relative comparison.

firms with older capital vintages are also hit harder with capital subsidies, reflecting a combination of upgrade costs, subsidies that partially offset these costs, and general equilibrium changes in the competitive environment.

Finally, we consider the effect of technology transfers from AEs to EMDEs. We model transfers of intellectual property rights to EMDEs as large subsidies to upgrade to the best available capital vintage, funded by AEs. We find significant output and consumption gains for recipient EMDE countries. However, whether emissions decrease depends on whether these vintages are produced in AEs or EMDEs. When EMDEs import newer vintages from AEs, expansions of output in the former may offset the reduction in emission intensities and total emissions may increase.

Literature and contributions. This paper contributes to prior work on firm performance and emissions. An established literature documents large heterogeneity in productivity at the firm level, even within narrowly defined industries (Syverson, 2011). This extends to emissions: prominent work by Shapiro and Walker (2018) and Lyubich et al. (2018) finds large dispersion both in the cross section and in the time series in emissions within US manufacturing firms operating in very similar industries using US Census data. We extend the finding that there is important variation in emissions across firms within the same industry—and in our case also in the same country—to a broad set of industries and countries, across both AEs and EMDEs.

Our paper also adds to a growing literature that examines what determines and correlates with greenness at the firm level (Haller and Murphy, 2012; Greenstone et al., 2012; Goetz, 2019; De Haas and Popov, 2023). A key contribution of our paper is to provide direct evidence that productivity-enhancing investments, such as R&D and capital upgrades, help reduce the emission intensity of production. This means that, across firms, these factors drive both environmental and financial performance. In addition, we analyze how different policies impact such firm-level investments through a novel heterogeneous-firm model.

As in our work, Lanteri and Rampini (2023) study the importance of capital vintages in determining firm-level emissions. They highlight the role of financing frictions, which are likely to be much less important within our sample of listed firms, as larger firms are generally less financially constrained (Beck et al., 2005; Driver and Muñoz-Bugarin, 2019). Barahona et al. (2020) focus on vintage-specific policies within the auto industry, while we consider vintages of capital across a broad set of industries.

A key contribution of the paper is to bring a corporate finance, heterogeneous-firm perspective to the large literature that has considered the macroeconomic implications of environmental policies by examining firm-level determinants of emissions and related firm choices. Many important papers in this literature feature a representative firm in each industry (Metcalf, 1999; Annicchiarico and Di Dio, 2015; Acemoglu et al., 2016; Metcalf and Stock, 2020; Chateau et al., 2022; IMF, 2022). A notable exception is the recent and closely related paper by Shapiro and Metcalf (2023) which allows for firms to differ in productivity and technology, and for entry. As in Shapiro and Metcalf (2023), we find that accounting for technology adoption mitigates the long-term costs of policies and helps rationalize the empirical finding that carbon taxes have small GDP costs (Goulder and Hafstead, 2017; Metcalf and Stock, 2020).

Another important contribution of our paper with respect to this strand of the literature is to incorporate additional and novel sources of heterogeneity and to propose an empirically grounded technology for the production of firms' emissions, all closely motivated by our empirical investigation and disciplined by the data. We find that heterogeneity across firms plays an important role in shaping the aggregate effectiveness of policies. In addition, while the literature largely focuses on the effect of a carbon tax, we consider a broader set of mitigation policies, including research and capital subsidies and technology transfers. We find that heterogeneity in capital vintages is especially important for capital subsidies and technology transfers because these policies induce many firms to upgrade their capital stock. Finally, we analyze the distributional impact of policies across firms, an important aspect that determines whether policies are politically feasible.

Structure of the paper. The remainder of the paper is structured as follows. Section 2 describes the data sources. Section 3 documents a large degree of within-industry heterogeneity in emission intensity. Section 4 investigates the drivers of such heterogeneity. Section 5 presents the model, and Section 6 the model's calibration. Section 7 discusses policy counterfactuals. Section 8 concludes.

2 Data

Data on yearly firm-level emissions—self-reported following the Greenhouse Gas Protocol—are retrieved through ICE Data Services directly analyzed emissions dataset.[3] We focus on CO2 equivalent scope 1 (direct) and scope 2 (indirect emissions from the generation of purchased energy) emissions. We exclude observations imputed by ICE Data Services rather than reported at the firm level through regulatory or voluntary filings. Our analysis focuses on within industry heterogeneity in emissions, and this is particularly challenging to impute.[4] Financial corporation data are extracted from S&P Compustat Global. The data covers corporations that issue publicly traded securities. We obtain a combined panel of 4,233 firms, headquartered in 70 countries, for the 2010-2022 period covering a low of 418 firms in 2010 and a high of 2,590 firms in 2022.[5] 78% of the firms are headquartered in advanced economies, with meaningful concentration in the US (Figure A1). Firms for which self-reported emissions are available are larger on average than Compustat firms (Figure A2).[6]

For the main analysis in section 3 and onwards, we exclude finance, public administration, and utilities. These firms' investment decisions are often shaped by direct public interventions and ownership rather than market forces, which are at the core of our analysis.

Because of these exclusions, the counterfactual exercises in section 6 capture emission reductions conditional on a fixed level of emissions per unit of energy purchased. The counterfactual exercise does not take into account that energy producing companies may also react to carbon pricing and other policies by offering cleaner energy to their corporate customers. That is, the counterfactual exercises do capture the changes in the *quantity* of energy purchased by the sample firms' response to the policies, but do not capture utilities' potential response in terms of *the greenness of the energy mix* offered to their costumers.

We utilize four distinct industry classifications, each with multiple levels of granularity.

---

[3] The GHG Protocol Corporate Accounting and Reporting Standard consists of requirements and guidance for firms preparing GHG emissions inventories.

[4] Research has shown that using emissions estimated by data vendors can lead to biased results (Aswani et al., 2023).

[5] We also obtain data on energy consumption from DataStream. All variables are analyzed after being winsorized at top and bottom 1 percent.

[6] Transportation & Public Utilities and Manufacturing sectors produce the highest amount of emissions in our dataset. The relative share of scope 1 and scope 2 emissions varies by sector (Figure A3). The share of emissions by country in our merged dataset is fairly representative of the country shares computed using World Bank aggregate emission estimates. Our merged dataset has a higher relative representation of industrial sector emissions than estimated by the World Resources Institute (Figure A4).

Three of these, the Standard Industrial Classification (SIC), which we use in most specifications, the North American Industry Classification System (NAICS), and the Global Industry Classification Standard (GICS), assign firms to their main industry according to the production processes. The final one is a text-based industry classification developed by Hoberg and Phillips (2010, 2016) for listed US firms. This classification captures which firms compete with each other based on firms' own disclosures.[7]

For a small subset of the US manufacturing firms in the sample (23 firms), we can access data on management practices from the World Management Survey (Bloom and Van Reenen, 2007; Bloom et al., 2012). The survey captures factors that are important for efficiency for the specific goods or services offered by each firm based on consensus among consultants and industry experts (Scur et al., 2021).

Estimates of the age of capital stocks at the firm level and productivity are also important to our empirical analysis. We construct proxies of both. Our estimate of the age of the capital stock draws on the time series of past investments within firm. We begin by discounting to the present all past capital expenditures using the average depreciation rate within industry, country, and year. The denominator of the age of capital variable is the sum of the current values of past capital expenditures. The numerator is the sum of those values but first multiplying each of them times the number of years in the past in which such expenditure took place. As we do not have information on product-level prices, we estimate TFP from revenue data following the procedure detailed in Asker et al. (2014).

Series on R&D cost differentials due to US state tax credit are taken from Lucking (2019). We merge each observation within US manufacturing using the headquarter state location and the closest year.

3 Heterogeneity in Emission Intensity

This section documents the extent of heterogeneity in firm emission intensities within countries and industries. We find that the heterogeneity is large and important. Improvements in the environmental performance of laggard firms—i.e., firms with high emissions relative to

---

[7] Hoberg and Phillips (2010, 2016) provide both a distance measure between each pair of different firms in the sample and also a partition of the sample in groups of firms competing with each other—akin to the concept of industries in standard classifications. We rely on the latter as a control in our regressions.

industry-country peers—could significantly reduce total emissions.

The heterogeneity of firms' emission intensities within industries is large, comparable or larger to heterogeneity in other measures of firms' performance, such as total factor and labor productivity. We focus on the log of CO2 emissions (scope 1 plus scope 2) over revenues (megatons per million of USD) in 2019, and extract residuals after controlling for country × industry (4-digit SIC) fixed effects.[8] The distribution of residuals is plotted in Panel (a) of Figure 1, showing large heterogeneity. For instance, the standard deviation (of the levels, after controlling for country × industry fixed effects) is $4.3x10^{-4}$ megatons per million of USD, about 10 times the median emission intensity in the sample. The difference between the 90th and the 10th percentile—within the same industry and country group—is 1.84 log points, which translates into 6.3 times larger emissions. These magnitudes are substantial relative to heterogeneity in productivity: the 90-10 percentile difference of emission intensities is almost as large as for labor productivity (revenues over wage bill), which is 1.96 log points, and much larger than for revenue-based TFP, which is 0.49.[9]

These large differences in emission intensities are present with countries at different income levels. Panel (a) of Figure 2 plots the kernel densities of log emissions over revenues separating firms headquartered in advanced and in emerging markets and developing economies, after controlling for industry fixed effects common across countries (while Figure 1 controls for industry *times country* fixed effects) and the log of assets. We observe large heterogeneity in emission intensities within both country groups. Firms in AEs emit less than their EMDE counterparts, although the two distributions partially overlap.

Interestingly, firms headquartered in countries with more stringent environmental policies have lower emissions intensities than other firms in the same industry. We rely on the OECD Environmental Policy Stringency Index to measure stringency.[10] Panel (b) of Figure 2 provides suggestive evidence of a role for policy action in shaping firms' environmental performance. The impact of policies is studied more formally in section 7.

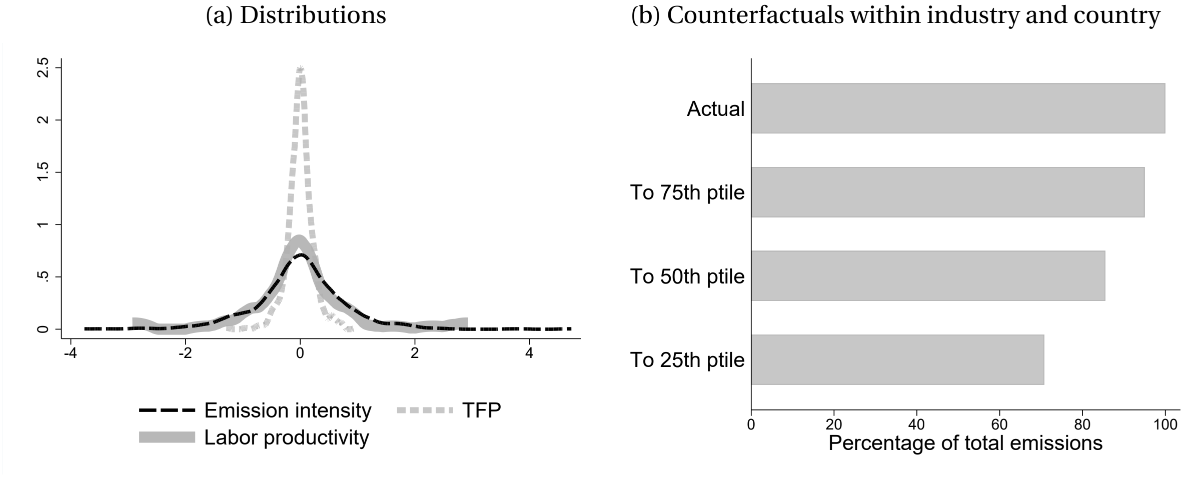
---

[8] We focus on 2019 to describe cross-sectional patterns in order to abstract from the potential effect of COVID. However, the results are robust if we focus on later years.

[9] Figure A5 presents the distribution of each of these variables through box plots.

[10] The OECD Environmental Policy Stringency Index is described by OECD website as: "*a country-specific and internationally-comparable measure of the stringency of environmental policy. Stringency is defined as the degree to which environmental policies put an explicit or implicit price on polluting or environmentally harmful behaviour. The index is based on the degree of stringency of 13 environmental policy instruments, primarily related to climate and air pollution.*"

Figure 1: Hetereogeneity in Emission Intensity - Within Industry and Country


(a) Distributions

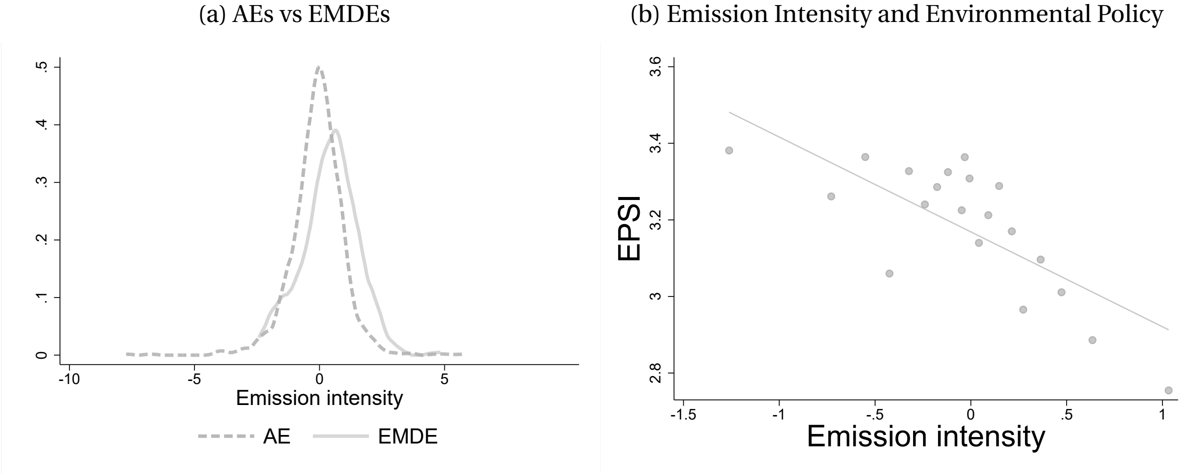
(b) Counterfactuals within industry and country

Notes: Panel (a) plots the kernel densities of the log of emissions intensity (measured as emissions over revenues), of (revenue) TFP, and of labor productivity (measured as revenues over wage bill), of all firms in the dataset, after controlling for industry × country fixed effects. Panel (b) illustrates actual emissions of firms in our sample, together with the counterfactual emissions that we would observe if every firm with emission intensity above the Xth percentile of the emission intensity distribution within the same country-industry group, saw its emission intensity reduced to that value. Only industry-country groups with at least 4 firms are included in panel (b). 4-digit SIC industry classification and 2019 data used. Finance, public administration, and utilities sectors are excluded from the calculation

A simple back-of-the-envelope exercise shows that heterogeneity in emissions is important: a counterfactual in which the worst environmentally performing firms improve would entail significantly lower total emissions. The exercise shows how total emissions would change if every firm had emission intensities at least as low as the firm at the Xth percentile of their industry-country group, holding each firm's output constant. For example, if firm's emission intensity is higher than the 25th percentile for its industry-country, we reduce its emission intensity to that percentile. Otherwise, we do not change its emission intensity. We then sum emissions across firms. Panel (b) of Figure 1 illustrates the results of this exercise performed for different percentiles of emission intensities. When using the 25th percentile of emission intensities, aggregate emissions fall by 33%. This suggests that improving the technology and production processes of laggard firms—even within technologies and processes available in the given industry and country—can play an important role in achieving emission reduction goals.

Next, we investigate hetereogeneity across countries at the industry level (Figure 3). Specifically, we explore the potential gains from achieving the performance of the best firms in the industry globally. As cross-country technology transfers may differentially benefit emerging

Figure 2: Hetereogeneity in Emission Intensity - Within Industry and Across Countries


(a) AEs vs EMDEs (b) Emission Intensity and Environmental Policy

Notes: Emission intensity is measured as the log of emissions over revenues. We residualize the variable against industry fixed effects (4-digit SIC). Finance, public administration, and utilities sectors are excluded from the calculation. 2019 data is used. Panel (a) plots the kernel density of firm emission intensity separately for firms headquartered in AEs and in EMDEs. Panel (b) plots a binned scatterplot of firm emission intensity against the Environmental Policy Stringency Index (EPSI) of OECD for the country where the firm is headquartered.

markets and developing economies (EMDEs) relative to advanced economies (AEs), we consider separate exercises within each group.[11] When using the 25th percentile of emission intensities to perform the exercise, aggregate emissions fall by 50% for AEs and 75% for EMDEs. Figure 3 suggests technology transfers to EMDE firms may be particularly beneficial.

These exercises exogenously assign emission intensities from greener to browner firms and assume nothing else changes. They do not consider whether such outcomes are feasible, which policies can achieve such gains, or at what cost. These considerations are explored later in the paper through the lens of a quantitative model.

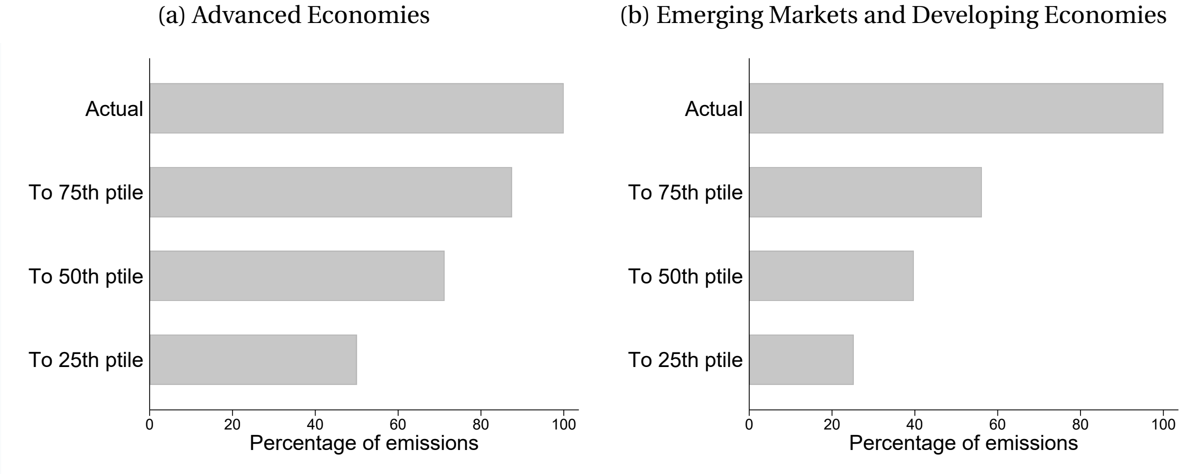
4 Emission Intensity and Firm Characteristics

Why is heterogeneity in emission intensity so large? Could differences in production processes and technological adoption drive this heterogeneity? Both academic and policy discussions have emphasized the crucial role of technology in reducing carbon emissions (e.g., Popp 2011;

---

[11] The exercise performed in Figure 3 is similar to Panel (b) of Figure 1 but computing the emission intensity thresholds only within industries, rather than within industry-country groups, when the former is smaller than the latter.

Figure 3: Emission Counterfactuals: Technology Transfers Across Countries


(a) Advanced Economies

(b) Emerging Markets and Developing Economies

Notes: This figure illustrates actual emissions of firms in our sample, together with the counterfactual emissions that we would observe if every firm was at least as clean as the firm in the Xth percentile of the emission intensity distribution of firms in the same industry, among advanced economy firms. 4-digit SIC industry classification and 2019 data used. Finance, public administration, and utilities sectors are excluded from the calculations.

Acemoglu et al. 2012; IMF 2020; Nordhaus 2021).

This section investigates the association between firm emission intensities and observable characteristics related to technology and efficiency, such as the age of physical capital stocks, the intangible share of capital, R&D expenditures, productivity, and management practices. We begin with stylized facts from the raw data. Indeed, Figure 4 shows that, in our global sample, firms with higher emission intensity operate older capital stocks, and have a lower share of intangible assets and lower TFP.

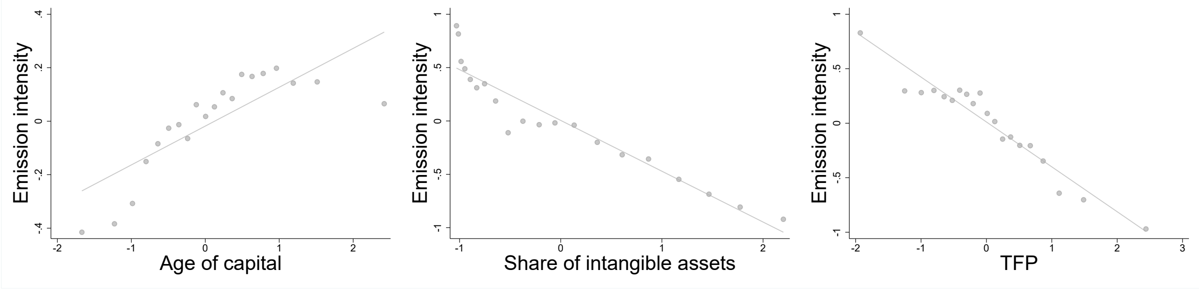
4.1 The Role of Technology Adoption and Innovation

We focus on the determinants of the heterogeneity in emission intensities *within* industries and countries. In principle, the patterns in Figure 4 could be driven by differences across industries or countries, or by firms of different size. We include a range of controls in our empirical analysis to restrict attention to the variation of interest.

Specifically, we estimate:

$$EI_{i,t} = \phi_{c(i),l(i),t} + X_{i,t}\beta + W_{i,t}\gamma + \epsilon_{i,t} \tag{1}$$

Figure 4: Emission Intensity, Age of Capital, Intangible Capital, and Productivity


(a) Age of capital

(b) Intangible assets

(c) Productivity

Notes: The figure displays binned scatter plots of emission intensities (the logarithm of emissions over revenue) against the age of capital (panel a), the share of intangible assets (panel b), and the logarithm of revenue productivity (panel c) of the firm. All variables standardized. Finance, public administration, and utilities sectors are excluded from the calculations.

where $EI_{i,t}$ is the emission intensity of firm i in year t, $\phi_{c(i),l(i),t}$ is a set of year-specific fixed effects for i's country and industry, $X_{i,t}$ are firm characteristics of interest, $W_{i,t}$ is a set of controls, and $\epsilon_{i,t}$ is an error term. In our baseline, emission intensity is calculated as the log of Scope 1 and 2 emissions scaled by revenues. All variables are standardized to have mean zero and variance of one, to make coefficients easily comparable. To avoid capturing mechanical correlations, we normalize emissions by lagged revenues while the independent variables, when calculated relative to size, are normalized by total assets. We include log assets as a control, to proxy for size and to avoid capturing economies/dis-economies of scale rather than the use of different technologies.[12]

Technological adoption and innovation seem to play an important role in driving firms' emission intensities. Results are presented in Table 1. We find that firms with older physical capital stocks emit more, relative to their size, than other firms in the same industry-country group, as reported in column (1). One standard deviation increase in the age of capital is associated with 0.03-0.06 standard deviation higher emissions per unit of revenue. This suggests that legacy machines and production processes lead to higher emissions. In column (2) we focus on the relation between knowledge and emission intensities. A one standard deviation increase in the share of intangible capital to total capital is associated with 0.14-0.18 of a standard deviation lower emissions per unit of revenues. Moreover, more productive

---

[12] Regressions only include country-industry-years with more than one observation. This requirement reduces the sample of firm-years in Table 1 from nearly 14,000 to about 6,500.

firms have lower emission intensities, as reported in column (3). Our results point to the important role that knowledge and technology can play in combating climate change.

All correlations are robust to a specification that includes all independent variables together—see column (4). This mitigates concerns that individual coefficients may be driven by omitting the other variables.

Table 1: Emission Intensity and Firm Characteristics

	(1)	(2)	(3)	(4)
	Emission Intensity (STD log emissions / revenue(t-1))			
STD age of capital	0.05***			0.03***
	(0.01)			(0.01)
STD share intangibles		-0.17***		-0.13***
		(0.01)		(0.01)
STD TFPR			-0.20***	-0.16***
			(0.01)	(0.01)
STD log(assets)	0.04***	0.08***	0.13***	0.14***
	(0.01)	(0.01)	(0.01)	(0.01)
N	6574	6574	6574	6574
R^2	0.80	0.81	0.82	0.82
Adj-R^2	0.70	0.71	0.72	0.73
Industry × country × year FE	Yes	Yes	Yes	Yes

Notes: All variables are standardized to have mean zero and standard deviation of one. Standard errors in parentheses, clustered at the country × industry × year. Finance, public administration, and utilities sectors are excluded from the calculations. By including only industry × country × year groups with more than one firm, our sample size is reduced from 13,946 to 6,574. * $p < .1$, ** $p < .05$, *** $p < .01$

These results capture correlations that are meaningful relative to total emissions. A back-of-the-envelope calculation similar to the one presented in section 3, finds that if all firms had capital in the youngest quartile, total emissions would fall by 7%, while if all firms were in the top quartile of knowledge intensity, total emissions would fall by 9%.[13]

Through which channels are the variables of interest related to emission intensity? More productive firms can economize on inputs, including energy, per unit of output. However, firms may also be able to emit less holding the quantity of inputs fixed, for instance because they use greener energy sources. To help separate between these channels, we investigate the

---

[13] We reduce firms' capital age or increase firms' knowledge intensity (share of intangible capital) to match the one of the Xth percentile of firms in their industry-country group. We rely on the coefficients reported by Table 1 to estimate the counterfactual emission intensity associated with the counterfactual age of capital or intangible share. Results using 2019 data are presented in graphical form in Figure A6.

the relationship between firm characteristics and the extent to which energy use translates into emissions, as measured by emissions normalized by energy consumption (data on the quantity of energy used is available for a subset of our sample).

Interestingly, more productive or knowledge-intensive firms do not emit less per unit of energy used—and therefore their better environmental performance is driven by lower consumption of inputs per unit of output (Table 2). In contrast, firms with newer capital stocks do emit less per unit of energy used. This suggests that newer capital is a complement to cleaner sources of energy, for instance because of more widespread use of electrical power or less use of brown (e.g., coal) in-house energy production.

Table 2: Emissions over Energy and Firm Characteristics

	(1)	(2)	(3)	(4)
	Emissions over Energy (STD log emissions / energy)			
STD age of capital	0.09***			0.10***
	(0.02)			(0.02)
STD share intangibles		0.00		-0.01
		(0.03)		(0.03)
STD TFPR			0.00	0.01
			(0.02)	(0.03)
STD log(assets)	-0.06**	-0.06**	-0.06**	-0.07**
	(0.03)	(0.03)	(0.03)	(0.03)
N	2722	2722	2722	2722
R^2	0.52	0.51	0.51	0.52
Adj-R^2	0.23	0.22	0.22	0.23
Industry × country × year FE	Yes	Yes	Yes	Yes

Notes: All variables are standardized to have mean zero and standard deviation of one. Standard errors in parentheses, clustered at the country × industry × year. Finance, public administration, and utilities sectors are excluded from the calculations. By including only industry × country × year groups with more than one firm, our sample size is reduced from 7,221 to 2,722. * $p < .1$, ** $p < .05$, *** $p < .01$

4.2 Robustness

In this section we discuss several robustness exercises to rule out alternative potential explanations to the results presented in Table 1. These additional results help address a range of concerns, including whether we compare sufficiently similar firms, and which measures best proxy for key objects of interest at the firm level. We also present evidence about management quality for a small sub-sample.

Industry classification. An important concern regarding our empirical results relates to industry classification. Our baseline empirical specification relies on the most granular industry classification available in our data: the 4-digit SIC industry classification. However, this classification remains imperfect and may mask substantial heterogeneity in the products produced by the firms we classify as belonging to the same industry and country.

To mitigate such concerns, we show that our results remain similar with a range of different approaches to classifying industries. Specifically, we re-estimate Equation 1 while changing either the classification system or the granularity of the industry considered. Results are presented in Table 3. Moving from column (1) to (3) we observe that the estimated coefficients are stable when we change the granularity of the SIC industry classification from 4 to 2 digit.[14] The same stability is observed comparing columns (4) to (5) and (6) to (7), which refer to 2 and 4 digit GICS and NAICS classification systems. Our results are both qualitatively and quantitatively similar regardless of the classification system adopted.

A further concern with traditional industry classifications—like SIC or NAICS—is that they rely on pre-determined industry descriptions. Firms' inclusion in an industry is to some extent a bureaucratic procedure with less than perfect adherence to economic boundaries of industries defined by cross-product substitutability. Moreover, boundaries between markets can shift as new products and technologies potentially change the competitive landscape. We therefore consider text-based industry classifications—derived from listed US companies' regulatory filings—that aim to capture which firms actually compete with each other (Hoberg and Phillips, 2010, 2016). Our results are robust to using this industry classification, as presented in column (8) of Table 3.[15]

Alternative measures of firm technology. An additional challenge is that empirically capturing our key objects of interest—the age of physical capital, firm technological adoption, and productivity—is inherently difficult, given the large span of industries and countries covered by our sample. To mitigate this concern, we repeat the empirical analysis with al-

---

[14] To ease the comparison across columns, we keep the sample constant across columns which differ only because of granularity of the fixed effects but rely on the same classification system.

[15] Table A3 shows that results for emissions scaled by energy are also robust to alternative industry classifications.

Table 3: Emission Intensity and Firm Characteristics: Alternative Industry Classifications

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Emission Intensity (STD log emissions / revenue(t-1))							
STD age of capital	0.03***	0.03***	0.02**	0.05***	0.06***	0.03***	0.03**	0.04*
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)
STD share intangibles	-0.13***	-0.12***	-0.16***	-0.21***	-0.22***	-0.11***	-0.17***	-0.18***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)
STD TFPR	-0.16***	-0.16***	-0.22***	-0.14***	-0.17***	-0.19***	-0.28***	-0.15***
	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)	(0.02)	(0.03)
STD log(assets)	0.14***	0.13***	0.16***	0.15***	0.17***	0.15***	0.18***	0.09***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.03)
N	6574	6574	6574	12699	12699	7865	7865	1376
R^2	0.82	0.81	0.75	0.67	0.59	0.81	0.65	0.74
Adj-R^2	0.73	0.72	0.68	0.60	0.54	0.73	0.60	0.65
Industry × country × year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry classification	SIC4	SIC3	SIC2	GICS4	GICS2	NAICS4	NAICS2	HP

Notes: Industry classification: SIC4 = 4-digit Standard Industrial Classification; GICS4 = 4-digit Global Industry Classification Standard; NAICS4 = 4-digit North American Industry Classification System; HP = Hoberg-Phillips 500. All variables are standardized to have mean zero and standard deviation of one. Standard errors in parentheses, clustered at the country × industry × year. Energy, utilities, finance, and public sectors are excluded from the calculations. * $p < .1$, ** $p < .05$, *** $p < .01$

ternative proxies for firm technological adoption and productivity. Results are similar to the baseline, as reported by Table 4, if we use the age of the firm, research and development (R&D) expenditures, and profitability as our main independent variables.[16]

Management practices. To further test whether improved production processes help lower emissions per unit of output, we analyze whether better management practices, as captured by the World Management Survey (Bloom and Van Reenen, 2007), are related to lower emission intensities. We estimate the following linear model:

$$\widetilde{EI}_{i,t} = \widetilde{MS}_i\beta + \widetilde{Assets}_{i,t}\gamma + \epsilon_{i,t} \tag{2}$$

$\widetilde{EI}_{i,t}$ are residualized log emission intensities, that is, the residuals resulting from regressing log emission intensities on a set of country × industry × year dummies. $\widetilde{MS}_i$ is the (residu-

---

[16] Table A4 shows that results for emissions scaled by energy are also robust to alternative measures of firm technology.

Table 4: Emission Intensity and Firm Characteristics: Alternative Indicators

	(1)	(2)	(3)	(4)
	Emission Intensity (STD log emissions / revenue(t-1))			
STD Age	0.04***			0.04***
	(0.01)			(0.01)
STD log(RD / assets)		-0.07***		-0.07***
		(0.02)		(0.02)
STD log(EBIT/assets)			-0.05***	-0.05***
			(0.01)	(0.01)
STD log(assets)	0.02*	0.03**	0.03***	0.01
	(0.01)	(0.01)	(0.01)	(0.01)
N	6242	6242	6242	6242
R^2	0.79	0.79	0.79	0.80
Adj-R^2	0.68	0.68	0.68	0.69
Industry × country × year FE	Yes	Yes	Yes	Yes

Notes: All variables are standardized to have mean zero and standard deviation of one. Standard errors in parentheses, clustered at the country × industry × year. Finance, public administration, and utilities sectors are excluded from the calculations. By including only industry × country × year groups with more than one firm, our sample size is reduced from 13,495 to 6,242. * $p < .1$, ** $p < .05$, *** $p < .01$

alized) overall management score of firm i according to the World Management Survey.[17] (Residualized) log assets are used to control for size. Before residualizing, all variables are standardized to have mean zero and standard deviation of one. As reported in Table 5, firms with higher management scores have lower emission intensity with respect to other firms in the same industry. These results support our hypothesis, and are in line with Bloom et al. (2010). Because the World Management Survey captures managerial practices mostly related to the efficiency of production processes, this result mitigates the concern that the positive correlation between TFP and emission intensities is solely driven by heterogeneous markups across firms, which can bias revenue-based measures of TFP.

---

[17] World Management Survey data is only available for a small subset of the US manufacturing firms in the sample (23 firms). Given the small sample size, we cannot include the fine set of fixed effects directly in the regression. Therefore, for emission intensities and total assets we first obtain the residuals of the variables against the fine set of fixed effects from a regression including all the firms in the matched Compustat-ICE data. For the management score, we use the score of the latest survey wave at our disposal, which is 2015 for most firms. We residualize this variable against country × industry dummies, from a regression including all firms in the World Management Survey data. We then include these residuals in Equation 2.

Table 5: Emission Intensity and Management Practices

	(1)	(2)
	Emission Intensity (STD log emissions / revenue(t-1))	
Management score	-0.465***	-0.532***
	(0.149)	(0.189)
STD log(assets)		0.107
		(0.088)
N	92	92
R^2	0.191	0.193
Adj-R^2	0.182	0.175

Notes: Robust standard errors in parentheses. Intensity and assets residualized against SIC3-country-year fixed effects, management scores are residualized against SIC3-country-wave fixed effects. * $p < .1$, ** $p < .05$, *** $p < .01$

Additional robustness tests. We perform a battery of additional robustness tests. Large firms may differ from small firms in important ways: they may be older, with both older workforce and physical capital stocks, and may have more inertia preventing the adoption of greener production processes. Columns (1) to (4) in Table A5 show that our results are robust to including only large firms in the regressions, addressing these concerns. Columns (5) to (8) reveal that our results are robust to the inclusion of several financial controls (lagged leverage, liquidity, and capitalization ratios, and market share). This exercise mitigates the concern that our results may be driven by financial frictions or corporate finance decisions which may impact both technology adoption and emissions.[18] Columns (9) to (12) show that the results are robust to removing 2020 and subsequent years, and hence are not driven by the COVID pandemic. Columns (13) to (16) illustrate that results are similar when we focus only on scope 1 emissions, which may be easier to measure for the reporting firms.[19] Table A7 shows that results are robust to focusing only on firms headquartered in AEs or EMDEs. Finally, Table A8 exhibits that our results are robust to computing emission intensities as emissions over total assets, rather than revenues, to mitigate the concern of a potential mechanical correlation between emission intensities and productivity estimated from data on sales.

---

[18] For this reason, the model proposed in section 5 abstracts from such frictions.

[19] Table A6 shows robustness for results for emissions scaled by energy.

4.3 Instrumental Variable Strategy

We document a robust association between firm emission intensity and technological factors. This association need not speak to causal effects. Moreover, the measurement of the main regressors of interest is challenging and could lead to a dampening of the estimated coefficients because of measurement error bias. For instance, we measure the age of capital stock using the timing of investments but we have no information on the actual equipment purchased, which could be older. In this subsection, we provide evidence based on instrumental variables suggesting that these associations are at least partially driven by a causal impact of technological factors on emissions.

Corporate investments in R&D are a major way to accumulate intangible capital and improve productivity and product quality. Policies to promote such investments, for instance by offering subsidies or tax credits, are common. A robust literature has aimed at evaluating the impact of public R&D incentives (Becker, 2015). In the United States, incentives to invest in R&D are provided both by the Federal and state governments (Chang, 2018).

We build on this literature to investigate whether policy-induced heterogeneity in R&D investments impacts firms' environmental performance. In particular, Lucking (2019) constructs a measure of differences in R&D cost for manufacturing firms by US states due to differences in tax credit policies, and shows that establishments in states with lower R&D costs grow and innovate more, and invest more in R&D. We observe where US firms are headquartered and use the state tax credit in the headquarter location to proxy for the actual state tax credit faced. This assumption is reasonable given that physical distance from headquarters is relevant for corporate governance (Giroud, 2013). We can then use the variation induced by state tax credit on R&D cost as an instrument for R&D expenditures.

Table 6 presents the results of this analysis. Column (1) presents the results of a OLS regression of firm emission intensity on R&D (normalized by revenues) while controlling for industry times year fixed effect and for firm size, for the sample of US manufacturers. It shows that firms that spend more on R&D also emit less per unit of output. Column (2) presents the first stage of the IV strategy, which is a regression of firm R&D on the cost of R&D due to state tax credit. We find that firms in low-R&D cost states invest more in R&D. The coefficient is different than zero only at the 10% confidence level, suggesting that state headquarter is a noisy proxy for the company's establishment location. Column (3) shows that firms in low-cost R&D not only spend more in R&D, but also have lower emission intensity, indicating a causal role of R&D on emissions. Column (4) presents a two-stage least square model where R&D expenditure is instrumented by cost of R&D credit. We find a negative and statistically significant coefficient for R&D, indicating that R&D is effective in lowering emission intensity.

Table 6: Emission Intensity and R&D

	(1) Emission Intensity	(2) R&D	(3) Emission Intensity	(4) Emission Intensity
R&D	-0.231***			-2.310***
	(0.0570)			(0.884)
Cost of R&D		-11.13*	24.83***	
(due to Tax-credit)		(6.090)	(8.715)	
N	1,264	1,264	1,264	1,264
R^2	0.821	0.891	0.820	-0.068
Industry × Year FE	Yes	Yes	Yes	Yes
		90% Confidence Intervals (Robust to Weak Instruments)		
Wald CI:				[-3.76,-.86]
AR CI:				[-5.19 ,-1.22]
		95% Confidence Intervals (Robust to Weak Instruments)		
Wald CI:				[-4.04,-.58]
AR CI:				[-6.83 ,-1.06]

Notes: Regressions include only US manufacturing firms. The cost of R&D due to US state tax credit is measured by Lucking (2019) and applied to the state where each US manufacturing firm is headquartered. All regressions add as a control variable the size of the firm (level of assets measured in logs). For the last regression, the first stage F-statistics is 3.396 (Montiel-Pflueger). Standard errors clustered at state level, * $p < .1$, ** $p < .05$, *** $p < .01$

Weak instrument techniques help us conclude that R&D helps improve environmental performance. Given that our first stage has an F-statistic below 4, we can obtain robust confidence intervals but not robust point estimates (Andrews et al., 2006; Isaiah et al., 2018; Pierri and Timmer, 2022). The Wald and Anderson-Rubin confidence intervals reported at the bottom of Table 6 exclude zero (they have upper bounds below zero). This means that the instrumental variable approach excludes a null impact of R&D on emission intensity at the 95% confidence level.

We also explore the potential for a causal relationship between the age of capital stock and emission intensity. Two firms of the same size may have physical capital stocks with very different ages depending on whether they achieved that size in a short or long period. Firms that grow faster are likely to have newer capital. This suggests instrumenting the age of capital stock with firms' growth rates in recent years. This instrument is valid under the assumption that given a high-growth firm and a low-growth firm of the same size and TFP, environmental performance is affected the only by the age of their capital stocks.

Under this assumption, we find that instrumental variable estimates are qualitatively similar to OLS ones (as reported in Table A9 and Table A10), indicating a causal impact of newer capital in decreasing emission intensity. Consistent with significant attenuation bias in our OLS specifications, the instrumental variable estimate is significantly larger. The F-statistic for the first stage indicates that the instrument is strong.

These instrumental variable results provide evidence supporting a causal impact of technological adoption and innovation in improving firms' environmental performance.

5 A Model of Capital Vintage and Knowledge Intensity

To rationalize the empirical findings and investigate their implications for mitigation policies, we propose a granular, general equilibrium heterogeneous-firm model of capital vintage, knowledge accumulation, and emissions, incorporating both entry and exit. This section briefly describes the model setting and the main equations. The next section uses a calibrated version of the model to analyze the impact of climate policies on firms' emissions, output, and profits as well as on households' consumption.

5.1 Setting

There is a finite number of countries indexed by j. In each country, there is a large set of firms and a representative household. We abstract from trade in goods and assets across countries: each economy is closed. Time is discrete and runs to infinity, $t = 1, 2, \dots$. When no confusion results, we omit the country and time subscripts.

Households' Preferences. Preferences of the representative household in a given country over streams of consumption bundles from 1 to infinity, $\{\mathscr{C}_t\}_{t=1}^{\infty}$, are given by $\mathscr{U} = \sum_{t=1}^{\infty} U(\mathscr{C}_t)$ where $U(.)$ is a strictly increasing and concave utility function. At each period, households consume differentiated final goods produced by firms in S different sectors $s = 1, \dots, S$. If $\mathscr{S}_s$ denotes the set of firms in sector s, each firm supplies a single variety $i \in \mathscr{S}_s$, and firms compete monopolistically. Households have Cobb-Douglas preferences over sectors and constant elasticity of substitution preferences over varieties within each sector s, $\mathscr{C}_t = \prod_{s=1}^{S} C_{st}^{\beta_s}$ with $C_{st} = \left[\sum_{i \in \mathscr{S}_s} \xi_{si}^{\frac{1}{\sigma}} c_{sit}^{1-\frac{1}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}}$ where ξ_{si} is a taste shock for the good produced by firm i in sector s, σ is the constant elasticity of substitution across varieties within a sector, and β_s is the expenditure share of household income on sector s, with $\sum_s \beta_s = 1$.

Taking prices of all goods as given, households seek to maximize their utility subject to their budget constraint, given by $\sum_{s=1}^{S} \sum_{i \in \mathscr{S}_s} p_{sit} c_{sit} + B_{t+1} \leq w_t L + T_t + mN_t + \Pi_t + (1+r_t) B_t$ where w is the nominal wage, L is the households' exogenous supply of labor, T denotes transfers net of taxes from the government, B denotes financial assets, r is the one-period interest rate, and Π denotes aggregate profits of all firms rebated in a lump-sum way to households. Natural resources—which can be used to produce energy—belong to the representative household, who sell to firms at an exogenous price m. For simplicity, they are supplied fully elastically.

Utility maximization at time t leads to the following demand for firm i in sector s.

$$c_{sit} = \xi_{si} \left(\frac{p_{sit}}{P_{st}}\right)^{-\sigma} C_{st} \quad \text{with} \quad C_{st} = \beta_s \frac{P_t \mathscr{C}_t}{P_{st}} \tag{3}$$

where the industry-s and the overall consumer price index are given by

$$P_{st} = \left(\sum_{i \in \mathscr{S}_s} \xi_{si} p_{sit}^{1-\sigma}\right)^{\frac{1}{1-\sigma}} \quad \text{and} \quad P_t = \Pi_{s=1}^{S} \left(\frac{P_{st}}{\beta_s}\right)^{\beta_s}. \tag{4}$$

Firm's Production and Emissions Technologies. Each firm $i \in \mathscr{S}_s$ produces its own product, which it sells to households, using a Cobb-Douglas technology that is common across firms within a given sector, s, and country j. Production combines the firm's intangible knowledge ω_{si}, physical capital k_{vsi} of vintage v_{si}, energy n_{si}, and labor ℓ_{si}. The production function in sector s is given by

$$y_{si} = \omega_{si} (v_{si} k_{vsi})^{\kappa_s} n_{si}^{\eta_s} \ell_{si}^{\lambda_s}. \tag{5}$$

The technology for the firm's emissions, which is novel relative to the literature, is informed by our empirical findings. Emissions depend on three terms. First, on how much energy it consumes, n_{si}; we do not distinguish between energy purchased externally or produced within the firm. Second, on the quality of vintage it uses, v_{si}. A higher-quality vintage allows the firm to decrease its emissions by increasing the efficiency of internal production, which impacts scope one emissions, and by using or purchasing cleaner sources of energy, which impacts both scopes of emissions. Finally, on factors common to all firms within a country and sector, such as the energy mix (e.g. emissions are lower in countries where electricity mostly comes from renewable sources).[20] Denoting this third term ϕ^e, a firm's emissions are given by[21]

$$e_{si} = \phi_s^e n_{si} v_{si}^{-\epsilon_v} \tag{6}$$

Importantly, the modeling approach we lay out in Equation 5 and Equation 6 is consistent with the key stylized facts shown in section 4. More productive firms emit less at a given level of output because they consume less energy. And firms with newer capital vintages emit less per unit of energy consumed. We formalize the consistency of the model with our stylized facts at the end of this section.

Firm's Intangible Knowledge and Research. Firms differ in the efficiency with which they combine their inputs. Firm i's intangible knowledge, ω_{si}, is the product of internal research conducted within the firm and spillovers from knowledge accumulated by other firms in the same sector and country. The latter captures the fact that knowledge diffuses across firms within each industry through workers, sharing of information, and the purchase of patents. More specifically, we follow Romer (1986) and assume that ω takes the following form

$$\omega_{si} = A_s^{\rho_s} a_{si} \tag{7}$$

where A_s denotes the aggregate intangible knowledge by all firms in industry s, and $\rho_s \in (0,1)$ is the strength of the externality. The aggregate stock of knowledge is the sum of knowledge capital embedded and produced in each firm, $A_s = \sum_i a_{si}$.

Firms hire workers ℓ_a to conduct research to improve their knowledge with the following

---

[20] While we exclude utilities and energy companies from the sample used for the regressions and estimation of the model, the energy produced by these companies and sold to the corporate sector, and the implied emissions, are captured by the demand of energy and scope two emissions in our counterfactual exercises.

[21] Firms emit greenhouse gases both directly, through production processes within the boundaries of the firm (scope one emissions), and indirectly, through energy purchased from other companies (scope two emissions). Scope one and scope two emissions are both quantitatively important, although with some heterogeneity across industries (see Figure A3), and our analysis encompasses both types.

accumulation technology

$$a'_{si} = (1-\delta_a)a_{si} + \left(\frac{\ell_a}{\gamma_{si}}\right)^{\alpha_s} \tag{8}$$

where the prime superscript refers to the following period, δ_a denotes the depreciation rate of knowledge capital, and γ_{si} is the research efficiency parameter. The latter represents the cost in terms of labor of increasing the productivity of the production process, a. An alternative interpretation is that it represents the cost of increasing the quality of the product. It is firm-specific, capturing differences across firms in productivity for generating ideas that improve their production processes. Firms learn about their research effectiveness after entering the market, and γ is drawn from a cumulative distribution, $G_s(\gamma)$, which is country and industry-specific.

In the rest of the paper we will assume that the elasticity of knowledge creation to labor, α_s, is equal to one minus the sum of the exponents on inputs in the output technology (5), $\alpha_s = 1-(\kappa_s+\eta_s+\lambda_s)$. This assumption is analogous to the more traditional assumption of constant returns to scale.

Firm's Vintage and Investment Decisions. Each period, firms have the opportunity to choose a vintage of capital and invest to expand their stock of physical capital. Vintages are indexed by v and the set of available vintages denoted $\mathcal{V} = \{1, ..., V_s\}$ is assumed to be exogenous for simplicity. We normalize v so that it is equal to the efficiency unit per unit of capital. With this notation, $v_s = V_s$ is the best vintage in sector s. There are markets for each capital vintage. x_v units of a capital vintage v_s cost $x_v q_{v_s}$ on the market. When we turn to the estimation of the model, we will relate the efficiency of the capital stock to its age, to match our empirical finding that newer capital stocks are more emission-efficient.

The decision to upgrade capital vintages is subject to two frictions, capturing realistic obstacles firms face when greening their production processes. First, we assume that it is prohibitively costly to combine vintages. As a result, a firm that updates its vintage needs to replace the entire stock of capital. Second, building on the findings of Kermani and Ma (2022) that most of the value of the capital stock is firm-specific, we assume that if a firm decides to retire its old vintage of capital, upgrade, and invest in a newer capital stock, it cannot recover all the value of its capital on secondary markets. If a firm decides to switch vintage, it recovers only a fraction χ of the value of the capital. When choosing which vintage of capital to use and how much to invest in capital of this vintage, firms therefore trade off the opportunity cost of retiring a productive asset with the profit gains from getting more productive and expanding their business. Taken together, these two assumptions imply that the decision to upgrade one's capital stock is costly and lumpy, two well-documented facts.[22] Finally, every period capital depreciates at rate δ.

Government Interventions: Carbon Taxes, Subsidies and Feebates. We consider four taxes or subsidies. First, there is a carbon tax, with rate τ_e, which is proportional to emissions, e.[23] Second, there is a rebate proportional to sales, τ_y, or a negative sales tax, which we will use in the design of the carbon feebate. Third, there is a subsidy to the best capital vintages, typically the most recent ones. We denote this vintage-specific subsidy τ_v, which results in an after-subsidy unit price of capital of $q_v(1-\tau_v)$. A subsidy to the latest and most productive capital vintage means that $\tau_v = 0$ for $v < V$ and $\tau_V > 0$. Finally, there is a subsidy to research activities that improve internal total factor productivity and quality, which we denote τ_a. The government budget is balanced in every period.

$$T = \sum_{s=1}^{S} \sum_{i \in \mathscr{S}_s} \left(\tau_e e_i - \tau_y p_i y_i - \sum_{v_s} \tau_{v_s} q_{v_s} x_{v_s i} - \tau_a w \ell_{ai} \right) \tag{9}$$

The Incumbent Firm's Problem. After having decided which vintage to use, how much to invest in physical capital, and how much research to conduct to improve intangible knowledge, firms choose how much labor for production ℓ and energy n to use as well as the price of their good. Taking the wage w, the price of energy m, and the prices of capital goods $\{q_v\}_{v=1}^{V}$ as given, a firm entering the period with a stock of capital k_v of vintage v, research productivity

---

[22] See, for example, Winberry (2021) for a recent analysis of the implications of firm-level fixed costs of investment. Closer to our paper, Finkelstein Shapiro and Metcalf (2023) assume a fixed cost of switching to a green technology.

[23] It is possible to replicate the allocation implied by a carbon tax with an equivalent carbon trading scheme in which firms are obligated to buy permits for all emissions, and in which the government earns the receipt of the sale of permits.

γ, and consumers' taste ξ faces the following dynamic profit-maximization problem

$$\begin{aligned}\mathcal{V}_s(a,k_v,v,\gamma,\xi) = \max_{v',k'_{v'},\{x_{v"}\}_{v"=1\ldots V_s},\ell_a,a'} \Big\{ & \pi_s(a,k_v,v,\gamma,\xi) - \sum_{v"=1}^{V_s} q_{v"}x_{v"}(1-\tau_{v"})\left(\mathbb{1}_{x_{v"}\geq 0} + \chi \mathbb{1}_{x_{v"}<0}\right) \\ & - w(1-\tau_a)\ell_a + \frac{1}{1+r}\mathcal{V}_s(a',k'_{v'},v',\gamma,\xi)\Big\} \end{aligned} \tag{10}$$

$$\pi_s(a,k_v,v,\gamma,\xi) = \max_{p,\ell,n}\{py(1+\tau_y) - w\ell - mn - e\tau_e - \kappa_f\} \tag{11}$$

where κ_f is the fixed cost of operating. Equation (10) describes the dynamic decisions taken by the firm—the investment, the vintage and the research decisions—and the resulting value function. This value function is the sum of current profits minus the cost of investing in capital goods and knowledge, plus the recovered retired capital if the firm switches its capital vintage, plus the continuation value discounted at the risk-free rate r. Equation (11) captures the static optimal input decision of the firm, choosing prices and spending on labor and energy to maximize current profits.

The firm solves these dynamic and static problems subject to the demand schedule (3), the production technology (5), the emission technology (6), the law of motion for knowledge (8), and the law of motion for capital (of each vintage) given by:

$$\begin{aligned} k'_{v'} &= (1-\delta)k_{v'} + x_{v'} \\ k'_v &= 0 \qquad\qquad \text{if } \ v' \neq v \end{aligned} \tag{12}$$

Entry and Exit. As further explained later in the section, we assume that relative prices and productivities of vintages are such that new entrants and incumbents that decide to update their capital always choose the newest one, V_s. The discounted sum of profits of a new entrant is thus given by $\mathcal{V}_s(a_s,k_{V_s},V_s,\gamma,\xi)$ and the expected discounted sum of profits before entry is denoted $\mathbb{E}_{G_s}\mathcal{V}_s(a_s,k_{V_s},V_s,\gamma,\xi)$, where the expectation is taken under the sector-specific joint distribution of research productivity and consumers tastes $G_s(\gamma,\xi)$. Potential entrants must pay an entry cost κ_e before learning about their research productivity γ and their consumers' tastes ξ. They enter if they can expect to make positive profits, namely if and only if

$$\mathbb{E}_{G_s}\left[\mathcal{V}_{si}(a_s,k_{V_s},V_s,\gamma,\xi)\right] > \kappa_e. \tag{13}$$

Incumbent firms decide whether to exit each period. Because they have to pay a fixed operating cost every period, κ_f, it is possible for a firm to make negative profits. As a result, an incumbent exits if the present discounted value of profits is negative, even after considering updating its vintage of capital, downsizing or expanding. On the contrary, an incumbent decides to stay if and only if

$$\mathcal{V}_{si}(a, k_v, v, \gamma, \xi) \geq 0 \tag{14}$$

Production Technology for Capital Goods. Capital goods are produced by a competitive sector that uses labor with a linear technology. These price-taking firms maximize profits and solve the following problem

$$\max_{x_{v_s}, \ell_{v_s}} q_{v_s} x_{v_s} - w\ell_{v_s}, \quad \text{subject to} \quad x_{v_s} = z_{v_s}\ell_{v_s} \tag{15}$$

Competitive pricing implies that in equilibrium the price of capital goods is pinned down by the productivity of labor: $q_{v_s} = wz_{v_s}^{-1}$. In addition, the production of capital goods generates emissions. Contrary to the case of final goods, we cannot assume that emissions are a function of energy use and capital vintages, because we abstract from these two inputs. Instead we assume that the emission intensity (the ratio of emissions over revenues) is the same as in the rest of the economy.

The market clearing conditions and the equilibrium definition are given in Appendix A1.1.

5.2 Properties of the Equilibrium

We now characterize the optimal decisions of firms to shed light on the four different margins through which policies affect emissions of firms within a sector. First, in the short-run, firms can adjust their variable inputs to reduce energy consumption. Second, in the medium-run they can change their research efforts to increase their overall efficiency. Third, they can also choose to upgrade to a better vintage of capital, thereby economizing on energy and emissions. Finally, they can make investments and deepen their capital intensity to economize on energy. We end the section by showing how the model qualitatively fits the main empirical findings. Derivations and proofs can be found in Appendix A1.

Optimal Choice of Energy. We start with the optimal consumption of energy given a stock of capital k_v of vintage v. It is given by

$$n = \Lambda_s(\xi) A_s^{\hat{\rho}} \left(\frac{\lambda_s}{w}\right)^{\hat{\lambda}_s} \left(\frac{\eta_s}{m+\tau_e v^{-\epsilon_v}}\right)^{\hat{\eta}_s+1} (v k_v)^{\hat{\kappa}_s} \tag{16}$$

$$\text{with} \quad \Lambda_s(\xi) = \left[\frac{\sigma-1}{\sigma}(1+\tau_y)(\eta_s+\lambda_s) P_s \,(\xi\, Y_s)^{\frac{1}{\sigma}}\right]^{\frac{\sigma}{\sigma+(1-\sigma)(\eta_s+\lambda_s)}}$$

where $\hat{x} = \frac{x}{\frac{\sigma}{\sigma-1}-(\eta_s+\lambda_s)}$ for $x = \rho, \lambda_s, \eta_s$ and $\frac{\sigma}{\sigma-1} > (\eta_s+\lambda_s)$ which ensures that the optimal input decisions are well-defined. It is intuitive that the consumption of energy is increasing in the productive capacity of the firm, as measured by its capital stock k_v, and increasing in the elasticity of output to energy, η_s.

Climate policies affect the consumption of energy first and foremost through the price of energy, $m+\tau_e^{v-\epsilon_v}$, which includes the carbon tax. An increase in the price of energy leads firms to substitute away from energy, to rely relatively more on capital and labor, and to scale down their production.[24] We will see later that it also changes their research effort and investment strategy in the medium run.

But an increase in the price of carbon impacts firms differently. Because they emit more per unit of energy used, firms that operate older and less emission-efficient vintages are impacted more than firms that are more energy and emission-efficient. As a result, firms with less efficient vintages will reduce their energy consumption to a larger extent than more efficient firms. The impact of a carbon tax also depends on sectors: firms operating in sectors where production relies a lot on energy, those with high η_s, are also more affected, which leads to a stronger reduction of energy use in these energy-intensive sectors.

Optimal Research Intensity. The optimal level of labor to conduct research and the implied level of accumulated knowledge in steady state is given by

$$\ell_a = \left(\frac{\Omega_s(v,\gamma,\xi,\tau_e)\hat{\alpha}_s k_v^{\hat{\kappa}_s}}{w(1-\tau_a)(r+\delta_a)\gamma^{\hat{\alpha}}\hat{\delta}_a}\right)^{\frac{1}{1-\hat{\alpha}_s}} \quad \text{and} \quad a = \left(\frac{\ell_a}{\gamma}\right)^{\alpha} \frac{1}{\delta_a} \tag{17}$$

where $\Omega_s(v,\gamma,\xi,\tau_e)$ is the revenue-productivity of (physical and intangible) capitals, since

---

[24] Given the assumption of a Cobb-Douglas production function, the elasticity of substitution across inputs, i.e. the elasticity of the ratio of inputs ℓ/n to their relative price $w/(m+\tau_e v^{-\epsilon_v})$, is one.

by definition $py = \Omega_s(v, \gamma, \xi, \tau_e) k_v^{\hat{\kappa}_s} a^{\frac{1}{\frac{\sigma}{\sigma-1} - (\eta_s + \lambda_s)}}$. Its expression is given in appendix. The revenue-productivity of capital $\Omega_s(v, \gamma, \xi, \tau_e)$ is increasing in v and ξ, decreasing in γ, decreasing in the carbon tax τ_e, and it depends positively on the sectoral price index, P_s, and production Y_s.

As for the optimal consumption of energy, the decision to conduct research depends on the scale of the firm as measured by k_v. Because research uses labor as its only input, research is negatively impacted by an increase in the wage rate w. Moreover, firms that are not efficient at doing research, those with high γ, endogenously conduct less research and use other inputs relatively more, such as energy or raw labor, instead.

It is clear from equation (17) that both research subsidies τ_a and policies that directly target emissions τ_e through the increase in k_v give incentives to firms to scale up their research effort to improve their intangible knowledge, economize on energy and become more productive. As for the consumption of energy, this reaction is especially strong for firms that have a vintage of capital of low quality or for firms that are located in sectors that rely more on energy, those with high η_s.

Optimal Pricing Decision. Firms compete monopolistically and have local market power to set the price of their variety. The optimal price is a constant markup over marginal cost. Given that this is not specific to our setting, we refer the reader to the appendix for more details.

Capital and Vintage Decisions. Firms also decide how much investment to make in physical capital and whether to upgrade to a more productive vintage of capital stock. Let's start with the case of a firm keeping the vintage it has been using so far, and choosing how much to invest. Its optimal steady-state level of capital is given by

$$k_v = \left[\frac{\kappa_s (r + \delta_a) w (1 - \tau_a)}{\alpha_s (r + \delta_k) \delta_a q_v} \right]^{\frac{1 - \hat{\alpha}_s}{1 - \hat{\kappa}_s - \hat{\alpha}_s}} \left[\frac{\Omega_s(v, \gamma, \xi, \tau_e) \hat{\alpha}_s}{w (1 - \tau_a)(r + \delta_a) \gamma^{\hat{\alpha}} \hat{\delta}_a} \right]^{\frac{1}{1 - \hat{\kappa}_s - \hat{\alpha}_s}} \tag{18}$$

where $\hat{\kappa}_s = \frac{\kappa_s}{\frac{\sigma}{\sigma-1} - (\eta_s + \lambda_s)}$.

Each climate policy affects the optimal capital stock of a firm, given by (18), differently. Since the optimal level of capital is decreasing in its price q_v, policies that subsidize the newest vintage of capital $\tau_V > 0$ will give an advantage to firms with this vintage and incentivize them to grow. This expansion will be at the expense of firms with a worse vintage, which will shrink. This reallocation of production to more efficient firms improves the overall productivity and greenness of the sector. But these policies have another important effect: they incentivize firms with a less productive vintage to upgrade by lowering the cost of adopting the newest vintage, a point to which we return below.

Policies that put a price on carbon, such as a carbon tax or a feebate, have ambiguous and heterogeneous effects on capital investment across firms. They incentivize firms to use less energy overall, and to increase their reliance on capital (of any vintage). But they also increase the cost of using capital of older vintages, thereby giving incentives to firms with these vintages to shrink, which leads to the expansion of firms with more emission-efficient vintages.

Finally, subsidies to research lead to more investment in intangibles, and to disinvestment from tangible and physical capital, which leads to a lower stock of capital in equilibrium.

The last margin of adjustment to policies is the possibility for firms to upgrade to a better vintage of capital. This decision amounts to comparing the value of upgrading to the value of continuing production with the older vintage. Using the optimality condition for capital (18), the steady-state value of an incumbent (10) that keeps its old vintage is given by

$$v(a,k,v,\gamma,\xi) = \frac{1}{r} q_v^{\frac{-\hat{\kappa}_s}{1-\hat{\kappa}_s-\hat{\alpha}_s}} \left(\frac{1}{1+r} \frac{(r+\delta_k)}{\hat{\kappa}_s} - \delta_k - \frac{(r+\delta_k)\delta_a}{(r+\delta_a)} \frac{\alpha_s}{\kappa_s} \right) \\ \times \left[\frac{\kappa_s (r+\delta_a) w (1-\tau_a)}{\alpha_s (r+\delta_k)\delta_a} \right]^{\frac{1-\hat{\alpha}_s}{1-\hat{\kappa}_s-\hat{\alpha}_s}} \left[\frac{\Omega_s(v,\gamma,\xi,\tau_e)\hat{\alpha}_s}{w(1-\tau_a)(r+\delta_a)\gamma^{\hat{\alpha}}\hat{\delta}_a} \right]^{\frac{1}{1-\hat{\kappa}_s-\hat{\alpha}_s}}$$

Alternatively, a firm could decide to upgrade its vintage of capital, v_{si}. Let's assume for the moment that firms upgrade to the best available vintage, V_s, and recall that, motivated by empirical evidence, firms that upgrade retire the vintage of capital that they had been using so far and recover only a fraction of its value. The value of upgrading is then given by

$$v^{up}(a,k_v,v,\gamma,\xi) = q_V^{\frac{-\hat{\kappa}_s}{1-\hat{\kappa}_s-\hat{\alpha}_s}} \left[\frac{1}{r} \left(\frac{1}{1+r} \frac{(r+\delta_k)}{\hat{\kappa}_s} - \delta_k - \frac{(r+\delta_k)\delta_a}{(r+\delta_a)} \frac{\alpha_s}{\kappa_s} \right) - 1 \right] \\ \times \left[\frac{\kappa_s (r+\delta_a) w (1-\tau_a)}{\alpha_s (r+\delta_k)\delta_a} \right]^{\frac{1-\hat{\alpha}_s}{1-\hat{\kappa}_s-\hat{\alpha}_s}} \left[\frac{\Omega_s(v,\gamma,\xi,\tau_e)\hat{\alpha}_s}{w(1-\tau_a)(r+\delta_a)\gamma^{\hat{\alpha}}\hat{\delta}_a} \right]^{\frac{1}{1-\hat{\kappa}_s-\hat{\alpha}_s}} + \chi q_v k_v (1-\delta_k) \qquad (19)$$

A firm upgrade if and only if the value (19) exceeds the value $v(a,k,v,\gamma,\xi)$. The following lemma gives a sufficient condition so that a firm that updates always adopts the most productive vintage V_s and establishes that firms that upgrade are the ones with the oldest capital vintage.

Lemma 1 (Vintage decision)**.** *Assume that the elasticity of* z_v *to* v *is strictly higher than* -1.

1. v_s/q_s *is increasing in* v_s.

2. *If a firm updates its vintage, it upgrades to the best vintage* V_s.

3. *There exists a unique* $\underline{v}_s$ *such that a firm with* v *upgrades to* V_s *if and only if* $v < \underline{v}_s$.

Policies that lower the price of the best vintage, q_V, through a subsidy to capital investment τ_V, give direct incentives to all firms with older capital vintages to upgrade since the value of upgrading (19) is increasing in τ_V. Carbon taxes and feebates also give incentives to upgrade but indirectly, through the price of energy. As we have seen in the paragraph on energy use, firms with worse vintages emit more and therefore see their cost rise up quickly when a carbon tax is implemented. These firms therefore have an incentive to upgrade when a carbon tax is put in place.

Partial and General Equilibrium Channels. At the micro-economic level, policies affect emissions by firms through the four channels discussed above: in the short-run, firms can adjust their variable inputs and reduce energy consumption; in the medium-run they can change their research efforts and increase their overall efficiency; they can also choose to upgrade to a better vintage of capital stock, thereby economizing on energy; and can decide to expand the scale of their production by investing in capital. But there are additional general equilibrium channels: market shares are reallocated across heterogeneous firms and production shifts across sectors. For example, sectors and firms that rely more on energy will shrink following an increase in the carbon tax, and those that rely to a larger extent on intangible capital will expand with a research subsidy. In addition, firms enter and exit endogenously. For example, a subsidy to the best vintage increases profitability for new firms, a feebate stimulates entry, and a carbon tax may lead some firms to exit.

5.3 Consistency with Stylized Facts

Importantly, the model described above is qualitatively consistent with our empirical findings: emission intensity at the firm level is driven by the age of the capital stock, knowledge intensity, and size. In addition, productivity and emission-intensity are positively correlated. Through the lens of the model, the firm's vintage quality (v)—which is closely related to the capital's age—and its research efficiency ($1/\gamma$)—which shapes its knowledge intensity—are two underlying drivers of both its emission intensity and productivity. The following proposition formalizes these important properties.

Proposition 1. *In steady-state, holding the capital stock k and the consumers' tastes ξ constant,*

1. *firms with younger vintages (v) or/and higher research efficiency (γ) are greener, i.e. emit less per unit of output:*

$$\ln\frac{e}{y} = \ln c_{st} - (\beta_{v1} + \epsilon_v)\ln v + \beta_{v2}\ln(m + \tau_e v^{-\epsilon_v}) - \beta_\gamma \ln 1/\gamma - \kappa \ln k + \beta_\xi \ln \xi \tag{20}$$

2. *firms with younger vintages (v) or/and higher research efficiency (γ) are more productive, i.e. have higher TFP:*

$$\begin{aligned} \ln TFP = \left(1 - \frac{1}{\sigma}\right)\kappa_s \ln v_i - \left(1 - \frac{1}{\sigma}\right)\eta_s \ln\left(m + \tau_e v^{-\epsilon_v}\right) + \left(1 - \frac{1}{\sigma}\right)\alpha_s \ln\left(1/\gamma\right) \\ + \left(1 - \frac{1}{\sigma}\right)\alpha_s \ln k + \frac{1}{\sigma}\log \xi_i + \ln z_{st} \end{aligned} \tag{21}$$

with $\beta_{v1} = \frac{\kappa}{b_\Omega}, \beta_{v2} = \frac{\eta}{b_\Omega}, \beta_\gamma = \frac{\alpha}{b_\Omega}, \beta_\xi = \frac{1-\eta_s-\lambda_s}{b_\Omega}$ and $b_\Omega = \left(\sigma + (\sigma - 1)(\eta_s + \lambda_s)\right)$, c_{st}, z_{st} are common to all firms within a time period, country and sector. Their full expression is given in Appendix A1.

To understand the first result, recall that emissions over output can be decomposed into the product of emissions over energy consumed and energy over output. The quality of the capital vintage v, and research efficiency $1/\gamma$, are positively related to energy efficiency in equilibrium, and the quality of the capital vintage v also reduces emissions per unit of energy (see equation 6).

As can be seen in the second expression, the model suggests that the way to interpret the positive empirical association we find between TFP (estimated with revenues) and emission intensity is that both are positively shaped by the quality of the vintage v and the research efficiency $1/\gamma$. Note that the relationship between environmental performance and size is more complicated because size depends positively on a firm's vintage quality and research efficiency but also on consumers' tastes.

6 Calibration

We are now ready to turn to the quantitative version of the model, which we will use to investigate the impact of commonly used and discussed mitigation policies. We start by carefully calibrating the model to match important country, sector and firm-level moments. A small subset of the parameters are calibrated externally; the rest are calibrated to match moments from the data.

6.1 External calibration

We begin by calibrating a number of parameters that can be set externally. These parameters are listed in Table 7. The time discount rate, r, the depreciation rate of physical capital, δ_k, and the elasticity of substitution across goods within sector, σ, are all calibrated to standard values used in the literature. The depreciation rate of knowledge δ_a is calibrated following Doraszelski and Jaumandreu (2013) and sources therein. The fraction of the capital stock that can be resold by a firm upgrading its capital stock, χ, is calibrated to match the finding by Kermani and Ma (2022) that on average 65% of the capital stock is firm-specific, which implies $\chi = .35$. We calibrate the difference between the growth rate of vintage productivity and price to match the estimated value of 3.2% in Greenwood et al. (1997).[25] Finally, we assume that the utility function of the representative consumer, U, is linear which implies a constant interest rate in equilibrium. This simplifies the computation of the transition path since the economy allocates all available resources towards the transition before reaching the new steady-state—a process which for most countries and most policies takes only one period.[26]

---

[25] Combined with our own empirical and model-consistent estimate of the growth rate of vintage productivity explained below, we can obtain the growth rate of vintage prices, which is otherwise difficult to observe.

[26] Another advantage of linear utility is that the net present value of consumption measures social welfare and is independent of specific assumptions regarding the intertemporal elasticity of substitution.

Table 7: Externally Calibrated Parameters

Parameter	Description	Value	Source
r	Discount rate	0.04	Standard
δ_k	Depreciation rate of capital	0.05	Standard
δ_a	Depreciation rate of knowledge	0.15	Doraszelski and Jaumandreu (2013)
σ	Elasticity of substitution	8	Standard
χ	Liquidation value	0.35	Kermani and Ma (2022)
$g_v - g_p$	Vintage productivity/price growth	3.2%	Greenwood et al. (1997)

6.2 Internal Calibration

The model's other parameters are jointly calibrated internally and can be split into four categories: those that are common to all firms in the world, to firms within a country, to firms within a sector and a country, and those that are firm-specific. Table A1 gives the list of parameters.

Table 8: Internally Calibrated Parameters and Firms' Characteristics

Parameter	Description	Granularity
g_v	Growth rate of vintage productivity	All countries
ϵ_v	Emissions elasticity to vintage	All countries
κ_f	Cost of operating	Country
ϕ^e	Average emissions per unit of energy	Country × Sector
κ_e	Cost of entry	Country × Sector
β	Expenditure share	Country × Sector
α	Knowledge elasticity to research	Country × Sector
ρ	Knowledge spillover	Country × Sector
κ	Capital elasticity of output	Country × Sector
η	Energy elasticity of output	Country × Sector
λ	Labor elasticity of output	Country × Sector
$G(\gamma, \xi)$	Distribution of initial parameters	Country × Sector
ξ	Consumers' taste	Firm
γ	Research efficiency	Firm
v	Vintage of capital	Firm

We briefly summarize our calibration strategy here and we refer the reader to Appendix A2 for a detailed explanation. At the country and industry level, we calibrate the elasticity of utility to each sector-level good, β_s, to match the share of sales of each sector separately for each country. While our sample of listed firms may not be representative of the whole corporate sector, we re-weight observations in order to match the share of each industry in each country. At the same level, the elasticity of output to its factors is identified using the assumption of constant returns to scale, the average sales and cost of goods sold, and the shares of costs going to labor, energy and research. Building on the literature, in particular Griliches (1992) and Bloom et al. (2013), we assume that the social returns to R&D are equal to its private returns, $\rho_s = \alpha_s$.

At the world level, we assume that the productivity of vintages grows at a constant rate over time, denoted g_v, and that there is a simple relationship between the age of the capital stock and the productivity of vintages given by $v(\text{age of capital}_i) = v_0(1+g_v)^{-\text{age of capital}_i}$ which we estimate by running a regression of firm TFP on the age of the capital stock. Similarly we start from equation (6): $\log\left[\frac{\text{Emission}}{\text{Energy}}\right] = \log(\phi_s^e) - \epsilon_v \log v$ to estimate the elasticity of emissions to capital vintage, ϵ_v.

At the firm level, we estimate the joint distribution of the firm-level variables v, γ, and ξ by matching three firm-level moments: the ratio of intangibles over tangible assets, the age of the capital stock—using $v = (1+g_v)^{-\text{age of capital}_i}$ and our estimates of the growth rate of productivity of vintages g_v—and the relative size of the firm.

6.3 Calibration of Counterfactual Policies

Emissions Target. We calibrate the four policy instruments—the carbon tax τ_e, the feebate (τ_e, τ_y), the subsidy to the best vintage τ_V and the research subsidy τ_a—in each country separately. While countries in practice are likely to consider a mix of carbon pricing and subsidies, we simulate one instrument at a time to isolate the properties of each policy. To ensure comparability across jurisdictions and instruments, we independently set each policy to generate the same 15% decline in each country's total corporate emissions. While such a target may appear below the level of ambition needed to achieve the Paris Agreement goals of reducing emissions by 45% by 2030, we find that more ambitious targets are not always attainable in all countries, especially when using subsidies. To avoid dropping too many countries from our sample, we choose a less ambitious target.[27]

---

[27] However, given that the focus of our analysis on the short to medium-run, the model abstracts from other important margins of climate mitigation and emissions as a ratio to GDP have been declining in the absence of policies, the 15% short-term target may be consistent with the reduction target of 45% by 2030.

Estimated Policies. If J is the number of countries, this calibration generates $J \times 4$ policy instruments, each consistent with a reduction of 15% in their respective country's emissions. Figure A7 in Appendix shows the cross-country distribution of carbon taxes, capital vintage and research subsidies that implement the target. All carbon taxes are at reasonable levels, ranging from about \$8 to \$80 a ton, and below \$50 in most countries. This suggests that a carbon tax is indeed an effective instrument to reduce the dispersion in environmental performance and overall emissions. Subsidies for capital upgrade need to be significantly higher to be successful with more heterogeneity across countries (from 14% to 93%), and an even more extreme result applies to research subsidies (from 56% to 96%).

7 Quantitative Effects of Mitigation Policies

Simulating the general equilibrium effects of each policy allows us to address four important questions. Our model embeds novel channels motivated by our empirical findings that may change our understanding of the trade-offs of different policies relative to existing studies. First, what are the channels of climate policies and can climate policies lower emissions by reducing the dispersion in environmental performance across firms? While the empirical counterfactuals shown in Figure 1 suggest that policies could play an important role, we now assess the extent to which realistically calibrated mitigation policies can effectively lift performance at laggards. Second, what are the costs in terms of output and consumption implied by different policies? Third, how do these costs depend on the heterogeneity across firms? Finally, what are the distributional implications of policies across firms? This question is important for political economy purposes and the acceptability of mitigation policies, which has been a major obstacle to a quick response to the climate crisis by governments.

7.1 Channels of Climate Policies and Firm Heterogeneity

To examine the extent to which each policy lowers emissions by reducing the dispersion in environmental performance across firms, we now examine the channels through which mitigation policies work, and start with a series of decompositions. We first decompose the total reduction in emissions into within- and between-sector contributions. Table 9 shows that the reduction of emissions within industries accounts for the bulk (80-90% or more) of emissions reductions.[28] This is in contrast with typical macroeconomic models used in the literature where cross-sector reallocation account for a much larger share of emissions reductions (Finkelstein Shapiro and Metcalf, 2023).

Table 9: Contribution of Within and Across Sector Emission Reductions To Total

	Carbon Tax	Feebate	Capital Subsidy	Research Subsidy
Within Sector	92%	91%	96%	83%
Across	8%	9%	4%	17%

Notes: the table shows the decomposition of total reduction in emissions into within- and between-sector contributions. Each policy entails a reduction in emissions of 15%.

We next decompose total emissions reductions along the lines of our empirical exercise in section 2, $E = \sum_{s=1}^{S} \sum_{i \in \Omega_s} \frac{e_i}{n_i} \frac{n_i}{y_i} \frac{y_i}{Y} Y + E_K$, into $\frac{e_i}{n_i}$ which denotes the emission-energy intensity, $\frac{n_i}{y_i}$ which denotes the energy-output intensity, $\frac{y_i}{Y}$ which is a measure of output relative to aggregate output, E_K which is total emissions in the capital goods sector. The change in emission-energy intensity e_i/n_i is driven by capital upgrades, since it is the only driver of emission-energy intensity. The change in energy-output intensity $\frac{n_i}{y_i}$, is driven by capital upgrades and changes in the mix of inputs, including research to reduce their energy consumption, which isolates the role played by the optimal mix of inputs and research. Third, we allow emissions in the capital good sectors, E_K, to adjust, which isolates the change in the size of the capital goods sector. Fourth, the change in the market shares of firms y_i/Y isolates the role played by the reallocation of factors across firms. Finally, we let aggregate output Y adjust.

Capital upgrade subsidies lead to a large decrease in the dispersion in environmental performance across firms and in the emission intensity, which is due to the fact that almost 85% of firms choose to upgrade to newer vintages as shown in Figure 5 and Table 10. The subsidy also gives incentives to firms with the best vintage to expand their capital investments

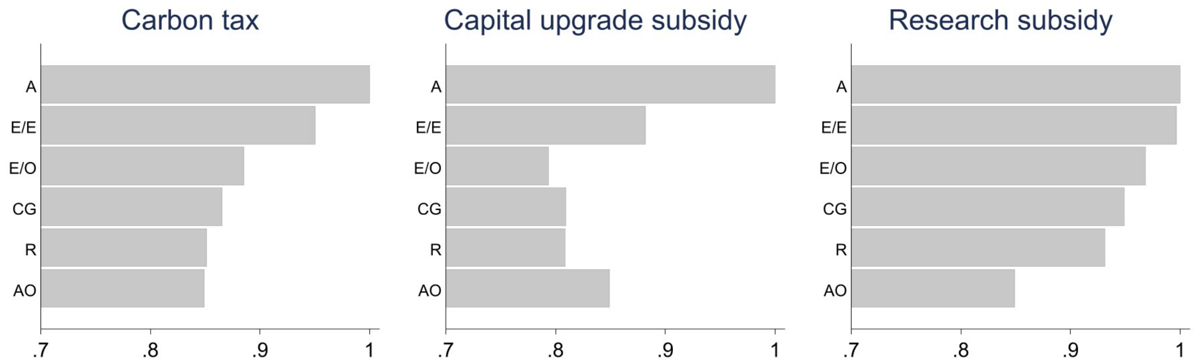
---

[28] The table shows

$$\frac{E^{\text{policy}} - E^0}{E^0} = \underbrace{\sum_s \frac{E_s^0}{E^0} \frac{E_s^{\text{policy}} - \tilde{E}_s^{\text{policy}}}{E_s^0}}_{\text{between sectors}} + \underbrace{\sum_s \frac{E_s^0}{E^0} \frac{\tilde{E}_s^{\text{policy}} - E_s^0}{E_s^0}}_{\text{within}} \quad \text{with} \quad \tilde{E}_s^{\text{policy}} = \frac{E_s^{\text{policy}}}{Y_s^{\text{policy}}} Y_s^0 \frac{Y^{\text{policy}}}{Y^0}$$

where the right-most term in the first equation corresponds to the change in total emissions due to the reduction in emission intensity within sectors and holding the relative size of sectors constant. Given that there is no heterogeneity across firms within the capital goods sector, this decomposition focuses on final goods firms.

Figure 5: Decomposition of emission reduction by channels



A=Actual, E/E=Emission/Energy Intensity, E/O=Energy/Ouput Intensity, CG=Capital Goods, R=Reallocation, AO=Aggregate Output.

Notes: The figure displays bar plots of the decomposition of actual emissions reduction by channels for each policy. The x-axis are total emissions relative to the baseline with no policy. For this reason, actual emissions (A) are equal to one. Subsequent rows show the cumulative effect on emissions incorporating each channel. E/E shows how much emissions decrease due to reductions in emissions per unit of energy used, E/O shows how much emissions decrease due to reductions in energy consumed per unit of output produced, CG shows total emissions coming from the capital goods sector, R the channel of a reallocation of resources across firms and sectors, and AO the channel of changes in total output. Each policy entails a reduction in emissions of 15%. We omit feebates as the decomposition is similar to that of carbon taxes.

to economize on energy. These upgrades and increases in capital intensity lead to a dramatic decrease in the energy-output and emission-energy intensities. As a result, we find that the policy leads to a very large reduction in the dispersion of environmental performance across firms, especially when compared to the carbon tax (Table 10). However, these effects on the reduction of emissions levels are to some extent offset by the increase in emissions coming from the capital goods sector, which expands as a result of the policy, and the increase in aggregate output.

Table 10: Effects of policies on upgrading and distribution of emission intensities

	Carbon Tax	Feebate	Capital Subsidy	Research Subsidy
Share of firms upgrading (in %)	4.6	4.6	84.7	0.0
Δ variance of $\log e/y$ (in %)	-3.7	-4.1	-87.0	0.0

Notes: Changes are given in percentage change relative to the actual economy. The variance is weighted by firm sales. Quantiles are moments of the distribution of emission intensities, also weighted by each firm' sales.

The carbon tax and feebate also reduce emissions through a decrease in emissions-energy intensity and in energy-output intensity, but there is much less reduction in the dispersion in environmental performance across firms. These policies work through a shift in inputs away from energy and in favor of other factors including higher investment in capital and research, and some vintage upgrades. However only five percent of firms worldwide find it optimal to upgrade (Table 10) which explains why we observe significantly less reduction in the variance of emission intensity across firms (Table 10) compared to the case of capital subsidies. As it will become clear in the next section on the costs of each policy, the fact that few firms upgrade is not a bad sign. On the contrary, a carbon tax is a more efficient instrument because it is broad in scope and lead firms to act on all possible margins to reduce emissions. Finally, the reallocation channel—the expansion of more efficient firms and sectors at the expense of less efficient ones—plays a quantitatively modest role.

The research subsidy reduces emissions mainly through the misallocation of inputs and a decline in aggregate production it generates, but does not reduce the dispersion in environmental performance across firms. The increasing demand for researchers pushes real wages up, and crowds out labor for other tasks such as capital and final goods production. It also works through some extent by making firms more productive, and by leading firms to rely less on energy in their mix of inputs e/y. Contrary to the capital subsidy and the carbon tax, it doesn't lead to any upgrades.[29] The research subsidy leads to a proportional improvement in research and productivity at all firms, and therefore the variance of environmental performance is unaffected.

7.2 Macroeconomic Costs and Intertemporal Trade-offs.

Each policy entails different macroeconomic costs and intertemporal trade-offs. The impact of the carbon tax on the net present value of consumption and on long-run consumption are -1.9% and -.5% relative to their actual level, respectively.

The long-run impact of carbon taxes is consistent with estimates from other studies such as Goulder and Hafstead (2017) who find a cost of about 1 percent of GDP, Metcalf and Stock (2020) who review the (small) empirical estimates from the literature and Shapiro and Metcalf (2023) who show that a model with endogenous entry, green technology adoption, frictional labor markets and endogenous labor market participation can rationalize this finding. An important difference between our results and theirs is that in our calibration few firms decide

---

[29] Recall that the subsidy incentivizes better use of existing technologies but does not spur innovation that expands the set of vintages of capital.

to adopt the newest technology with the carbon tax. Firms react to the carbon tax by adjusting along other margins, for example by reducing energy consumption and investing in research, two margins that are not available in their model. The feebate has the same output and consumption cost as the carbon tax but it does better for firms profits, which is why it is appealing in the first place.

Relative to the carbon tax, subsidies to capital and research are both very costly in net present value terms: they lead to a decline in the net present value of consumption of 12.5% and 4.5% respectively because they are too narrow in scope, don't give direct incentives to economize on energy and distort the efficient allocation of inputs. More specifically the increased demand for labor—to produce capital goods in the case of capital subsidies, and for research in the case of the research subsidies—pushes real wages up, and inefficiently crowds out labor from the final good sectors. In addition, too many firms update and the expansion of the capital goods sector implied by the upgrade subsidy generates additional emissions which partially offsets the gains in the final goods sectors, which in turn requires even greater subsidies to reach the 15% target. R&D subsidies stimulate research, which addresses the underinvestment in research related to its externality. However, the size of the subsidies that deliver a 15% reduction in emissions is much larger than the optimal Pigouvian subsidy. We find in unreported simulations that smaller subsidies (around 5% for most countries) can increase output and consumption. Finally, the fiscal costs of subsidies are also substantial: 18% of GDP for capital subsidies and 12% for research subsidies.

Policies that achieve the greatest reduction in dispersion in environmental performance are therefore not the most efficient ones from a net present value of consumption perspective. While capital subsidies lead many firms to adopt the newest technologies and lead to a striking convergence of environmental performance across firms, they are also very costly. Too many firms update, and those that update overaccumulate capital. Carbon taxes in contrast do not lead to such strong convergence across firms, but they are also more efficient, because they incentivize firms to reduce their emissions through all possible margins, including reducing energy use and increasing knowledge intensity.

Policies differ greatly in the intertemporal tradeoffs they entail. Unlike carbon taxes and research subsidies, capital subsidies generate consumption gains in the long run, because firms that update become both greener and more productive, as can be seen in the 1.1% increase in TFP. Quantitatively we find that this positive effect more than offsets the misallocation of inputs: as a result output increases by 4% in the long-run. Intuitively, this increase depends on the extent to which production relies on capital (κ) and we indeed find that the average capital intensity is positively related to the long-term benefits in the cross-section of countries (Figure A8). While the time profile of costs and benefits is flat in the case of the carbon tax and the research subsidies, the costs of capital subsidy are concentrated in the short run as many firms upgrade and invest in better capital vintages. To respond to this increased demand, many new capital goods need to be produced. We find that the equivalent of five years of consumption need to be sacrificed to achieve the transition.

Table 11: Effects of policies on aggregates

In Percentages	Carbon Tax	Carbon Feebate	Capital Subsidy	Research Subsidy
NPV of Consumption	-1.9	-2.0	-12.5	-4.5
Output	-0.5	-0.5	4.0	-7.7
Consumption	-0.5	-0.5	4.0	-7.7
Profits	-0.6	0.0	4.0	-7.8
Labor Productivity	-0.2	-0.2	6.1	2.9
TFP	0.0	0.0	1.1	3.5
Fiscal Cost	-0.6	0.0	18.4	12.4

Notes: In percentage change of the actual economy, except for fiscal cost. Fiscal cost is the sum of the steady-state net subsidies and the costs incurred during the transition, annualized. Fiscal cost is in percent of steady-state GDP in the counterfactual economy. NPV=Net Present Value. We use a 4% time discount factor to compute the net present value of consumption. Output, consumption, profits, labor productivity, TFP refer to their value in the steady-state and are weighted averages across sectors within countries and across countries, where the weights are the country-specific sector shares and countries GDP.

Overall, although our findings show that policies that put a price on carbon emissions, such as carbon taxes or feebates, are more efficient than policies that target a specific input, it also provides a rationale for capital subsidies when the social planner's discount rate is lower than the market interest rate. A lower social discount rate can be justified either on the ethical grounds that society should not discount future generations or on the grounds of sustainability of a non-decreasing level of consumption over generations (Ramsey, 1928; Campbell and Martin, 2023). We find that the social discount rate that would make the capital subsidies as appealing as the carbon tax is 1.3% which is about the same as the discount rate argued for in the Stern report (Stern, 2006).

7.3 Costs of Capital Subsidies and Firm Heterogeneity

How do these costs depend on the nature of heterogeneity across firms? We find that the costs (and long-term benefits) of using capital subsidies to achieve a 15% reduction in emissions crucially depends on the initial distribution of capital vintages in the economy. The reason is that the main channel of transmission of the capital subsidies is through the upgrading of capital vintages. We find that when firms have on average older capital stocks, the short-term cost are smaller, and the net present value of consumption is higher (-3.4% versus -12.5 in the baseline counterfactual, see Table A11).[30]

The reason for this result is that when firms initially have a lower quality capital, fewer firms need to upgrade their vintage and expand their capital investment for the economy as a whole to reach the target (45% instead of 84%, see Table A12).[31] However, a worse initial distribution of vintages leads to smaller long-run aggregate GDP increases (2.8% versus 4%). This result is also driven by the fact that less firms need to update relative to the baseline counterfactual.

7.4 Distributional Effects on Firms' Financial Performance

Policies affect firms differentially and different policies have different distributional implications. A better understanding of these implications is necessary to design policies that can gather broad support, especially from those who are directly affected by it. Our focus is on the distributional effects across firms, which also affect workers. Although we don't model these implications for workers explicitly, in a world in which skills are not fully transferable from one firm to another and transitions are costly, workers' wages and employment opportunities are to some extent tied to their employers' financial performance. In this section, we look at the differential effects along two crucial dimensions: first, along the age of the capital stock, and second, along the energy intensity.

As the first column of Table 12 shows, the impact of a carbon tax on firm profitability is

---

[30] We study this through a simulation in which we reduce the quality of the initial distribution of capital vintages (we divide capital productivity of all firms, v, by 2), keeping everything else constant, and observe how the costs of achieving a certain reduction in emissions change. Results can be found in Appendix.

[31] Can we conclude that countries with older capital stock should use capital subsidies? The answer is yes and no. Yes, if we are comparing countries that differ only in their distribution of vintages. But in practice countries differ in many other dimensions that also shape the macroeconomic costs of policies.

modest: a carbon tax leading to a 15% aggregate emissions reductions decreases the firms' value by only 1 to 5%. This is because the initial negative effect on profits is largely offset by the fact that firms pass through the increase in their cost to consumers. The effect depends on the age of their capital stock. For firms with older, less green, and less efficient physical capital stocks, a carbon tax implies a larger decrease in their profits. Energy intensity, which varies at the sector level, but not at the firm level, also matters. Firms in sectors that rely relatively more on energy are also more strongly impacted by a carbon tax than firms in sectors that rely less on energy.

Feebates with a rebate proportional to sales have similar differential effects as the carbon tax but they can improve overall acceptability. As the first column of Table 12 shows, the decrease in profits implied by carbon taxes is 50% larger than the one implied by feebates. Feebates are appealing because they give incentives to both firms at the frontier and to laggards to converge to leaders while reducing the negative impact for most firms and rewarding a fraction of firms, and thus may be more likely to be supported by broader coalitions than carbon taxes.

Table 12: Effects of policies on firms' value

In % Change to Actual	Carbon Tax	Feebate	Capital Subsidy	Research Subsidy
Newer Capital	-1.8	-1.2	-34.5	-7.3
Older Capital	- 3.5	-3.0	-36.7	-7.9
Lower η	-0.9	-0.3	-38.2	-7.8
Higher η	-5.2	-4.6	-32.5	-7.6

Notes: A firm's value is the present discounted value of current and future profits. The time discount rate is the interest rate used in the calibration.

Subsidies for newer capital vintages have large negative effects on firms' values. This is mostly driven by the general equilibrium increase in real wages, due to the need to allocate an increasing share of the labor force to the capital goods sector when many firms choose to upgrade and invest. This cost is much larger for firms with older capital and that update, because they have to pay a very high cost of investment to upgrade and they lose a large share of the value of the older capital stock.

The research subsidy also has negative effects on the firms' value, mostly because of the implied general equilibrium decrease in output. Note that contrary to the carbon tax and the capital subsidy, it affects all firms to similar degrees because it works mostly through a common decrease in output and demand.

7.5 Technology Transfers to Emerging Markets

Given that climate change is a shared global challenge and the large differences in income per capita between mature and emerging countries, technology transfers are a prominent policy option option for reducing emissions while fostering development. Technology transfers are often discussed in terms of transfers of IP rights from firms in mature market to EMs. From the perspective of firms in EMs that purchase the capital, this is akin to a subsidy. If in addition governments compensate the capital-producing firms that own the IP rights, technology transfers are exactly equivalent to a subsidy by the government for the purchase of newest vintage of capital goods.

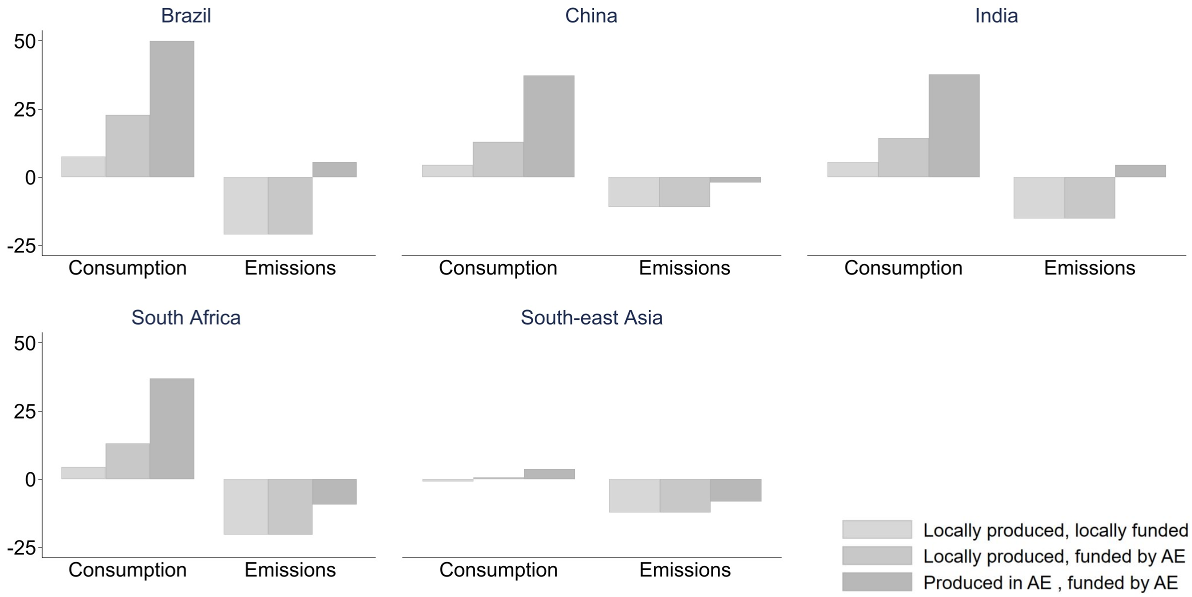
Consistent with this view, our approach is to compute a counterfactual in which all firms in EMs are given a subsidy on their purchase of the newest vintage of capital. The exact percentage covered by the subsidy is an outcome of negotiations across countries, so our baseline counterfactual exercise considers a 50% subsidy. Given the uncertainty in its design, we look at three ways of implementing this transfer: (i) the latest vintage of capital is produced in EMs, and taxpayers in EMs pay for the subsidies; (ii) the latest vintage of capital is produced in EMs, and taxpayers in AEs pay for the subsidies, (iii) the latest vintage of capital is produced in AEs, and taxpayers in AEs pay for the subsidies. Given recent policies in AEs aiming to relocate industrial production locally, option (iii) may be most feasible from a political perspective.[32]

Figure 6 displays the resulting change in aggregate consumption and emissions in each EM in our sample. We find significant output and consumption gains from these policies, especially for the case in which production of the newest vintages occurs in AEs. While output expands by the same amount in case (i) and (ii), consumption increases by more in case (ii) because subsidies are funded through transfers by AEs instead of taxes (case (i)). Output and consumption increase even more in case (iii) because there is no crowding out of inputs (especially labor) and final goods output by the production of capital goods, which

---

[32] Our model does not incorporate labor market frictions—such as partial employment—or domestic spillovers—such as for human capital—in countries that receive FDI.

Figure 6: Impact of Technology Transfers to EMs (50% Subsidy to Newest Vintage)



Notes: The figure shows, for selected EMDE countries or country groups, the percentage change in aggregate consumption and emissions with respect to the baseline after a 50% subsidy on the purchase of the newest vintage of capital. Three types of subsidy designs are considered, depending on whether the subsidy is funded by the local government or by advanced economies, and whether the newest capital vintage is produced locally or in advanced economies.

are produced in and imported from AEs. The decrease of emissions by 10 to 20% in the case of (i) and (ii) is encouraging. However emissions increase in Brazil, India, and in the group of South-east Asia countries in case (iii): while the ratio of emissions over output declines significantly, the expansion of output and consumption of energy offsets some of these gains.

8 Conclusions

This paper brings a firm-level, corporate finance perspective to the literature on mitigating climate change. Firms producing similar goods vary vastly in their environmental performance—how much they emit relative to the scale of their operations. We show that climate "laggards"—firms with higher emission intensity with respect to industry-country peers—are older and have older capital stocks, are less efficient and innovative, and adopt worse management practices. These findings suggest that both better adoption of existing technologies and more innovation can help reduce emissions.

We propose a general equilibrum heterogeneous-firm model with endogenous choice of capital vintages, knowledge accumulation, and emissions. The model rationalizes our empirical findings and helps assess tradeoffs and channels of mitigation policy options. We carefully calibrate the model to aggregate,industry, and firm-level moments and use it to assess the impact of four different policy instruments.

Our counterfactual analysis yields several important insights. First, mitigation policies differ in the extent to which they reduce the dispersion in environmental performance, with capital subsidies having the strongest impact. Second, subsidies for capital upgrades and R&D can help reduce emissions but at significantly larger macroeconomic costs and with different intertemporal tradeoffs relative to carbon taxes and feebates. Third, the degree and the source of heterogeneity across firms matter for the effectiveness of policies, especially for capital upgrade subsidies. Fourth, we also show that different instruments have diverse distributional implications both across firms, and between consumers and firms. While a carbon feebate leads firms to reduce their emissions as effectively as a carbon tax, it also benefits most firms and therefore may be more politically palatable. Finally, we find significant output and consumption gains from technology transfers to EMs, but the decrease in emissions may be limited by the strong increase in output.

References

Acemoglu, Daron, Philippe Aghion, Leonardo Bursztyn, and David Hemous, "The environment and directed technical change," *American Economic Review*, 2012, *102* (1), 131–166.

__ , Ufuk Akcigit, Douglas Hanley, and William Kerr, "Transition to clean technology," *Journal of Political Economy*, 2016, *124* (1), 52–104.

Andrews, Donald WK, Marcelo J Moreira, and James H Stock, "Optimal two-sided invariant similar tests for instrumental variables regression," *Econometrica*, 2006, *74* (3), 715–752.

Annicchiarico, Barbara and Fabio Di Dio, "Environmental policy and macroeconomic dynamics in a new Keynesian model," *Journal of Environmental Economics and Management*, 2015, *69*, 1–21.

Asker, John, Allan Collard-Wexler, and Jan De Loecker, "Dynamic inputs and resource (mis) allocation," *Journal of Political Economy*, 2014, *122* (5), 1013–1063.

Aswani, Jitendra, Aneesh Raghunandan, and Shivaram Rajgopal, "Are carbon emissions associated with stock returns?," *Review of Finance*, 2023.

Barahona, Nano, Francisco A Gallego, and Juan-Pablo Montero, "Vintage-specific driving restrictions," *The Review of Economic Studies*, 2020, *87* (4), 1646–1682.

Beck, Thorsten, Asli Demirguc-Kunt, and Vojislav Maksimovic, "Financial and Legal Constraints to Growth: Does Firm Size Matter?," *The Journal of Finance*, 2005, *60* (1), 137–177.

Becker, Bettina, "Public R&D policies and private R&D investment: A survey of the empirical evidence," *Journal of economic surveys*, 2015, *29* (5), 917–942.

Bloom, Nicholas and John Van Reenen, "Measuring and explaining management practices across firms and countries," *The quarterly Journal of Economics*, 2007, *122* (4), 1351–1408.

__ , Christos Genakos, Ralf Martin, and Raffaella Sadun, "Moden management: Good for the environment or just hot air?," *The Economic Journal*, 2010, *120*, 551–572.

__ , Mark Schankerman, and John Van Reenen, "Identifying Technology Spillovers and Product Market Rivalry," *Econometrica*, 2013, *81* (4), 1347–1393.

__ , Raffaella Sadun, and John Van Reenen, "The organization of firms across countries," *The quarterly journal of economics*, 2012, *127* (4), 1663–1705.

Campbell, John Y. and Ian Martin, "Sustainability in a Risky World," 2023.

Chang, Andrew C, "Tax policy endogeneity: evidence from R&D tax credits," *Economics of Innovation and New Technology*, 2018, *27* (8), 809–833.

Chateau, Jean, Florence Jaumotte, and Gregor Schwerhoff, "Economic and environmental benefits from international cooperation on climate policies," Departmental Paper No 2022/007, International Monetary Fund 2022.

Doraszelski, Ulrich and Jordi Jaumandreu, "Estimating Endogenous Productivity," *The Review of Economic Studies*, 2013, *80* (4), 1338–1383.

Driver, Ciaran and Jair Muñoz-Bugarin, "Financial constraints on investment: Effects of firm size and the financial crisis," *Research in International Business and Finance*, 2019, *47*, 441–457.

Georgieva, Kristalina, "No Time To Waste," IMF Blog 2021.

Giroud, Xavier, "Proximity and investment: Evidence from plant-level data," *The Quarterly Journal of Economics*, 2013, *128* (2), 861–915.

Goetz, Martin Richard, "Financing conditions and toxic emissions," SAFE working paper 2019.

Goulder, Lawrence and Marc Hafstead, *Confronting the Climate Challenge: U.S. Policy Options*, Columbia University Press, 2017.

Greenstone, Michael, John A List, and Chad Syverson, "The effects of environmental regulation on the competitiveness of US manufacturing," Technical Report, National Bureau of Economic Research 2012.

Greenwood, Jeremy, Zvi Hercowitz, and Per Krusell, "Long-Run Implications of Investment-Specific Technological Change," *American Economic Review*, June 1997, *87* (3), 342–362.

Griliches, Zvi, "The Search for R&D Spillovers," *The Scandinavian Journal of Economics*, 1992, *94*, S29–S47.

Haas, Ralph De and Alexander Popov, "Finance and green growth," *The Economic Journal*, 2023, *133* (650), 637–668.

Haller, Stefanie A and Liam Murphy, "Corporate expenditure on environmental protection," *Environmental and Resource Economics*, 2012, *51*, 277–296.

Hoberg, Gerard and Gordon Phillips, "Product market synergies and competition in mergers and acquisitions: A text-based analysis," *The Review of Financial Studies*, 2010, *23* (10), 3773–3811.

_ and _ , "Text-based network industries and endogenous product differentiation," *Journal of Political Economy*, 2016, *124* (5), 1423–1465.

IMF, "Mitigating Climate Change–Growth-and Distribution-Friendly Strategies," World Economic Outlook, International Monetary Fund 2020.

_ , "Near-Term Macroeconomic Impact of Decarbonization Policies," World Economic Outlook, International Monetary Fund 2022.

_ , "Climate Crossroads: Fiscal Policies in a Warming World," Fiscal Monitor October, International Monetary Fund 2023.

Isaiah, Andrews, Stock James, Sun Liyang et al., "Weak Instruments in IV Regression: Theory and Practice," *Annual Review of Economics*, 2018.

Kammer, Alfred, "Europe, And the World, Should Use Green Subsidies Cooperatively," IMF Blog 2023.

Kermani, Amir and Yueran Ma, "Asset Specificity of Nonfinancial Firms*," *The Quarterly Journal of Economics*, 07 2022, *138* (1), 205–264.

Lanteri, Andrea and Adriano A Rampini, "Financing the Adoption of Clean Technology," 2023.

Li, Bo, Fabio Natalucci, and Prasad Ananthakrishnan, "How Blended Finance Can Support Climate Transition in Emerging and Developing Economies," IMF Blog 2022.

Lucking, Brian, "Do R&D tax credits create jobs?," *Working Paper*, 2019.

Lyubich, Eva, Joseph Shapiro, and Reed Walker, "Regulating Mismeasured Pollution: Implications of Firm Heterogeneity for Environmental Policy," *AEA Papers and Proceedings*, 2018, *108*, 136–142.

Metcalf, Gilbert E, "A distributional analysis of green tax reforms," *National tax journal*, 1999, *52* (4), 655–681.

Metcalf, Gilbert E. and James H. Stock, "Measuring the Macroeconomic Impact of Carbon Taxes," *AEA Papers and Proceedings*, May 2020, *110*, 101–106.

Nordhaus, William D, *The spirit of green: the economics of collisions and contagions in a crowded world*, Princeton University Press, 2021.

Pierri, Nicola and Yannick Timmer, "The importance of technology in banking during a crisis," *Journal of Monetary Economics*, 2022, *128*, 88–104.

Popp, David, "International technology transfer, climate change, and the clean development mechanism," *Review of Environmental Economics and Policy*, 2011, *5* (1), 131–152.

Ramsey, F. P., "A Mathematical Theory of Saving," *The Economic Journal*, 1928, *38* (152), 543–559.

Romer, Paul, "Increasing Returns and Long-run Growth," *Journal of Political Economy*, 1986, *94* (5), 1002–37.

Scur, Daniela, Raffaella Sadun, John Van Reenen, Renata Lemos, and Nicholas Bloom, "The World Management Survey at 18: lessons and the way forward," *Oxford Review of Economic Policy*, 2021, *37* (2), 231–258.

Shapiro, Alan Finkelstein and Gilbert E. Metcalf, "The macroeconomic effects of a carbon tax to meet the U.S. Paris agreement target: The role of firm creation and technology adoption," *Journal of Public Economics*, 2023, *218*, 104800.

Shapiro, Alan Finkelstein and Gilbert E. Metcalf, "The macroeconomic effects of a carbon tax to meet the U.S. Paris agreement target: The role of firm creation and technology adoption," *Journal of Public Economics*, 2023, *218*.

Shapiro, Joseph S. and Reed Walker, "Why Is Pollution from US Manufacturing Declining? The Roles of Environmental Regulation, Productivity, and Trade," *American Economic Review*, 2018, *108* (12), 3814–3854.

Stern, Nicholas, *Stern Review on the Economics of Climate Change* Oct 2006.

Syverson, Chad, "What Determines Productivity?," *Journal of Economic Literature*, 2011, *49* (2).

Winberry, Thomas, "Lumpy Investment, Business Cycles, and Stimulus Policy," *American Economic Review*, January 2021, *111* (1), 364–96.

A1 Annex: Model Appendix

A1.1 Market Clearing Conditions

In each country, the markets for all final goods i in all sectors s clear.

$$\forall\ s = 1, \dots, S, \quad \forall\, i \in \mathscr{S}_s, \quad c_{si} = y_{si} \tag{22}$$

Denoting $x_{i v_s}$ the demand of capital goods of vintage v_s by firm i and ℓ_{v_s} the labor used in the production of capital goods of this vintage, the market clearing condition for capital goods of vintage v_s is given by

$$\forall\ v_s = 1, .. V_s \quad \sum_{i \in \mathscr{S}_s} x_{i v_s} = z_{v_s} \ell_{v_s}. \tag{23}$$

The bonds market clears $B_t = 0$ as well as the labor market.

$$\sum_{s=1}^{S} \sum_{i \in \mathscr{S}_s} \left(\ell^a_{si} + \ell_{si} \right) + \sum_{s=1}^{S} \sum_{v_s \in \mathscr{V}_s} \ell_{v_s} = L \tag{24}$$

A1.2 Equilibrium Definition

An equilibrium is a set of value functions $\mathscr{U}, \mathscr{V}_s(a, k, v, \gamma, \xi)$, and policy rules $p_s(a, k, v, \gamma, \xi), k'_s(a, k, v, \gamma, \xi), a'_s(a, k, v, \gamma, \xi), \ell_s(a, k, v, \gamma, \xi), n_s(a, k, v, \gamma, \xi)$, a mass function $M_s(a, k, v, \gamma, \xi)$, a set of capital good prices $\{q_{v_s}\}_{v_S=1}^{V_s}$ for each sector $s = 1 ... S$, and a wage w such that

1. Taking final goods prices, wages, net transfers T and aggregate profits Π as given, households maximize their utility.
2. Given the wage and the capital goods prices, the firm's policy rules and the value function solves the problem (10).
3. Given the wage and the capital goods prices, capital goods firms maximize their profits (15).
4. The mass function $M_s(a, k, v, \gamma, \xi)$ is consistent with the entry condition (13), the exit condition (14), and the policy rules of the final good firms.

5. The final goods (22), bonds, capital goods (23), and labor markets clear (24).

A1.3 Model Solution

In this section, we solve the model and present the main equations of interest, including the optimal variable input choices, pricing, and dynamic decisions (capital, vintage, and R&D).

Firm's choice of variable inputs We first analyze how the firm optimally chooses labor and energy consumption. The operating costs function of the firm is given by:

$$C(\ell, n) = w\ell + mn + \tau_e n v^{-\epsilon_v} \tag{25}$$

Dropping the industry and firm subscript, and denoting μ the lagrange multiplier associated with the production technology in the cost-minimization problem, the first order conditions with respect to ℓ and n are given by:

$$\lambda_s \frac{\mu y}{\ell} = w \tag{26}$$

$$\eta_s \frac{\mu y}{n} = m + \tau_e \frac{e}{n} \tag{27}$$

Combining the two gives:

$$\frac{\eta_s}{\lambda_s}\ell = \frac{m + \tau_e v^{-\epsilon_v}}{w} n$$

Substituting back into the production function, one gets:

$$y = A_s^{\rho} a (\ell)^{\lambda_s + \eta_s} \left(\frac{\eta_s}{\lambda_s} \frac{w}{m + \tau_e v^{-\epsilon_v}} \right)^{\eta_s} (v k_v)^{\kappa_s}$$

We then obtain ℓ and n:

$$\begin{aligned}
\ell &= \Gamma^{\ell}_{sv} y^{\frac{1}{\eta_s+\lambda_s}} \\
\Gamma^{\ell}_{sv} &= A_s^{-\frac{\rho}{\eta_s+\lambda_s}} a^{-\frac{1}{\lambda_s+\eta_s}} \left(\frac{\lambda_s}{\eta_s}\frac{m+\tau_e v^{-\epsilon_v}}{w}\right)^{\frac{\eta_s}{\eta_s+\lambda_s}} (v k_v)^{-\frac{\kappa_s}{\eta_s+\lambda_s}} \\
n &= \Gamma^{n}_{sv} y^{\frac{1}{\eta_s+\lambda_s}} \\
\Gamma^{n}_{sv} &= A_s^{-\frac{\rho}{\eta_s+\lambda_s}} a^{-\frac{1}{\lambda_s+\eta_s}} \left(\frac{\eta_s}{\lambda_s}\frac{w}{m+\tau_e v^{-\epsilon_v}}\right)^{\frac{\lambda_s}{\eta_s+\lambda_s}} (v k_v)^{-\frac{\kappa_s}{\eta_s+\lambda_s}}
\end{aligned}$$

For future reference, we denote

$$\tilde{\Gamma}^{\ell}_{sv} = A_s^{-\frac{\rho}{\eta_s+\lambda_s}} \left(\frac{\lambda_s}{\eta_s}\frac{m+\tau_e v^{-\epsilon_v}}{w}\right)^{\frac{\eta_s}{\eta_s+\lambda_s}} v^{-\frac{\kappa_s}{\eta_s+\lambda_s}}$$

so that $\Gamma^{\ell}_{sv} = \tilde{\Gamma}^{\ell}_{sv}(a k_v^{\kappa_v})^{-\frac{1}{\lambda_s+\eta_s}}$

From the definition of operating costs given by equation (25), we have

$$C(\ell, n) = w\ell\left(1+\frac{\eta_s}{\lambda_s}\right) = \Gamma_{sv} y^{\frac{1}{\eta_s+\lambda_s}} \tag{28}$$

$$\Gamma_{sv} = \Gamma^{\ell}_{sv}\left(w\left(1+\frac{\eta_s}{\lambda_s}\right)\right) \tag{29}$$

For future reference we denote,

$$\tilde{\Gamma}_{sv} = \tilde{\Gamma}^{\ell}_{sv}\left(w\left(1+\frac{\eta_s}{\lambda_s}\right)\right) \tag{30}$$

so that $\Gamma_{sv} = \tilde{\Gamma}_{sv}(a k_v^{\kappa_v})^{-\frac{1}{\lambda_s+\eta_s}}$.

We can then rewrite operating profits as

$$p_{si}^{1-\sigma}\xi_s P_s^{\sigma} Y_s(1+\tau_y) - \Gamma_{sv}\left(p_{si}^{-\sigma}\xi_s P_s^{\sigma} Y_s\right)^{\frac{1}{\eta_s+\lambda_s}}$$

The first order condition for the optimal price is then given by:

$$(\sigma-1)p_{si}^{-\sigma}\xi_s P_s^{\sigma} Y_s(1+\tau_y) = \Gamma_{sv}\frac{\sigma}{\eta_s+\lambda_s} p_{si}^{-\frac{\sigma}{\eta_s+\lambda_s}-1}\left(\xi_s P_s^{\sigma} Y_s\right)^{\frac{1}{\eta_s+\lambda_s}} \tag{31}$$

$$p_{si}^{1-\sigma+\frac{\sigma}{\eta_s+\lambda_s}} = \frac{\Gamma_{sv}}{(1+\tau_y)}\frac{\sigma}{(\sigma-1)(\eta_s+\lambda_s)}\left(\xi_s P_s^{\sigma} Y_s\right)^{\frac{1}{\eta_s+\lambda_s}-1} \tag{32}$$

$$p_{si} = \frac{\sigma}{\sigma-1}\frac{\Gamma_{sv}}{(1+\tau_y)(\eta_s+\lambda_s)} y_{si}^{\frac{1}{\eta_s+\lambda_s}-1} \tag{33}$$

We can rewrite this as a constant markup over marginal cost:

$$\begin{aligned} p_{si} &= \frac{\sigma}{\sigma-1} MC_{si} \\ MC_{si} &= \frac{\Gamma_{sv}}{(1+\tau_y)(\eta_s+\lambda_s)} y_{si}^{\frac{1}{\eta_s+\lambda_s}-1} \end{aligned}$$

We now solve for the equilibrium size (output) of a firm:

$$\left(\frac{y_{si}}{\xi_{si} Y_s}\right)^{-\frac{1}{\sigma}} P_s = \frac{\sigma}{\sigma-1}\frac{\Gamma_{sv}}{(1+\tau_y)(\eta_s+\lambda_s)} y_{si}^{\frac{1}{\eta_s+\lambda_s}-1}$$

$$\begin{aligned} y_{si} &= \left[\left(\frac{\sigma}{\sigma-1}\frac{\Gamma_{sv}}{(1+\tau_y)(\eta_s+\lambda_s)}\frac{1}{P_s}\right)(\xi_{si} Y_s)^{-\frac{1}{\sigma}}\right]^{\frac{1}{1-\frac{1}{\eta_s+\lambda_s}-\frac{1}{\sigma}}} \\ &= \left[\left(\frac{\sigma}{\sigma-1}\frac{\Gamma_{sv}}{(1+\tau_y)(\eta_s+\lambda_s)}\frac{1}{P_s}\right)^{-1}(\xi_{si} Y_s)^{\frac{1}{\sigma}}\right]^{\frac{\sigma(\eta_s+\lambda_s)}{\sigma+\eta_s+\lambda_s-\sigma(\eta_s+\lambda_s)}} \\ &= \left[\left(\frac{\sigma}{\sigma-1}\frac{\tilde{\Gamma}_{sv}}{(1+\tau_y)(\eta_s+\lambda_s)}\frac{1}{P_s}\right)^{-1}(\xi_{si} Y_s)^{\frac{1}{\sigma}}\right]^{\frac{\sigma(\eta_s+\lambda_s)}{\sigma+(1-\sigma)(\eta_s+\lambda_s)}} (k_v^{\kappa_s} a)^{\frac{\sigma}{\sigma+(1-\sigma)(\eta_s+\lambda_s)}} \\ &= \left[\frac{\sigma-1}{\sigma}(1+\tau_y)(\lambda_s+\eta_s)P_s(\xi_{si} Y_s)^{\frac{1}{\sigma}}\right]^{\frac{\sigma(\eta_s+\lambda_s)}{\sigma+(1-\sigma)(\eta_s+\lambda_s)}} \tilde{\Gamma}_{sv}^{-\frac{\sigma(\eta_s+\lambda_s)}{\sigma+(1-\sigma)(\eta_s+\lambda_s)}} (k_v^{\kappa_s} a)^{\frac{\sigma}{\sigma+(1-\sigma)(\eta_s+\lambda_s)}} \end{aligned}$$

In addition, the size of a firm in value is given by

$$p_{si} y_{si} = \frac{\sigma}{\sigma-1}\frac{\Gamma_{sv}}{(1+\tau_y)(\eta_s+\lambda_s)} y_{si}^{\frac{1}{\eta_s+\lambda_s}}$$

Substituting the expression for y_{si} just found into the expressions for ℓ and n, the implied optimal level of labor and energy are given by

$$\begin{aligned}
\ell &= \Lambda_{s\xi} A_s^{\hat{\rho}} \left(\frac{\lambda_s}{w}\right)^{\hat{\lambda}_s+1} \left(\frac{\eta_s}{m+\tau_e v^{-\epsilon_v}}\right)^{\hat{\eta}_s} (v k_v)^{\hat{\kappa}_s} \\
n &= \Lambda_{s\xi} A_s^{\hat{\rho}} \left(\frac{\lambda_s}{w}\right)^{\hat{\lambda}_s} \left(\frac{\eta_s}{m+\tau_e v^{-\epsilon_v}}\right)^{\hat{\eta}_s+1} (v k_v)^{\hat{\kappa}_s} \\
\Lambda_{s\xi} &= \left[\frac{\sigma-1}{\sigma}(1+\tau_y) P_s (\xi_{si} Y_s)^{\frac{1}{\sigma}}\right]^{\frac{\sigma}{\sigma+(1-\sigma)(\eta_s+\lambda_s)}}
\end{aligned}$$

Operating profits before operating cost are given by

$$\begin{aligned}
\pi_{si} + \kappa_s &= py(1+\tau_y) - \Gamma_{sv} y^{\frac{1}{\eta_s+\lambda_s}} && (34) \\
&= py(1+\tau_y)\left(1 - \frac{\sigma-1}{\sigma}(\lambda_s+\eta_s)\right) && (35) \\
&= \left(\frac{\sigma}{\sigma-1}\frac{1}{(\eta_s+\lambda_s)} - 1\right)\Gamma_{sv} y_{si}^{\frac{1}{\eta_s+\lambda_s}} && (36) \\
&= \Omega_{sv\xi}(a k_v^{\kappa_s})^{-\frac{1}{\eta_s+\lambda_s}+\frac{1}{\lambda_s+\eta_s}\frac{\sigma}{\sigma+(1-\sigma)(\eta_s+\lambda_s)}} && (37) \\
&= \Omega_{sv\xi}(a k_v^{\kappa_s})^{\frac{1}{\frac{\sigma}{\sigma-1}-(\eta_s+\lambda_s)}} && (38) \\
&= \Omega_{sv\xi} k_v^{\hat{\kappa}_s} a^{\frac{1}{\frac{\sigma}{\sigma-1}-(\eta_s+\lambda_s)}} && (39)
\end{aligned}$$

with

$$\Omega_{sv\xi} = \left[\frac{\sigma-1}{\sigma}(1+\tau_y)(\lambda_s+\eta_s) P_s (\xi_{si} Y_s)^{\frac{1}{\sigma}}\right]^{\frac{\sigma}{\sigma+(1-\sigma)(\eta_s+\lambda_s)}} \left(\frac{\sigma}{\sigma-1}\frac{1}{(\eta_s+\lambda_s)} - 1\right) \tilde{\Gamma}_{sv}^{\frac{(1-\sigma)(\eta_s+\lambda_s)}{\sigma+(1-\sigma)(\eta_s+\lambda_s)}}$$

We assume $\frac{\sigma}{\sigma-1} > \eta_s + \lambda_s$.

We now go back to our profit function:

$$v(a,k,v,\gamma,\xi) = \max_{\ell_a, x, v', a', k'_v} \left\{ \Omega_{sv\xi} k_v^{\hat{\kappa}_s} a^{\frac{1}{\frac{\sigma}{\sigma-1}-(\eta_s+\lambda_s)}} - \sum_w q_w x_w - w(1-\tau_a)\ell_a + \frac{1}{1+r} v(a', k'_{v'}, v', \gamma, \xi) \right\} \tag{40}$$

$$\begin{aligned}
k'_{v'} &= (1-\delta_k) k_{v'} + x_{v'} && (41) \\
a' &= (1-\delta_a) a + \left(\frac{\ell_a}{\gamma}\right)^{\alpha_s} && (42)
\end{aligned}$$

If μ_k, μ_a denote the Lagrange multipliers associated with the law of motion of capital and knowledge respectively, the FOCs and the E.C. are given by:

$$q_v = \mu_k \tag{43}$$

$$w(1-\tau_a) = \mu_a \alpha \ell_a^{\alpha-1} \gamma^{-\alpha} \tag{44}$$

$$\frac{1}{1+r} v_k(a', k_v', v, \gamma) = \mu_k \tag{45}$$

$$\frac{1}{1+r} v_a(a', k_v', v, \gamma) = \mu_a \tag{46}$$

$$v_k(a,k,v,\gamma) = \Omega_{sv\xi} \hat{\kappa}_s k_v^{\hat{\kappa}_s - 1} a^{\frac{1}{\frac{\sigma}{\sigma-1} - (\eta_s + \lambda_s)}} + \mu_k (1-\delta_k) \tag{47}$$

$$v_a(a,k,v,\gamma) = \Omega_{sv\xi} \hat{\kappa}_s k_v^{\hat{\kappa}_s} \frac{a^{\frac{1}{\frac{\sigma}{\sigma-1} - (\eta_s + \lambda_s)} - 1}}{\frac{\sigma}{\sigma-1} - (\eta_s + \lambda_s)} + \mu_a (1-\delta_a). \tag{48}$$

This results in the following condition:

$$\Omega_{sv\xi} \hat{\kappa}_s k_v^{\hat{\kappa}_s - 1} a^{\frac{1}{\frac{\sigma}{\sigma-1} - (\eta_s + \lambda_s)}} + q_v (1-\delta_k) = (1+r) q_v \tag{49}$$

$$k_v = \left(\frac{\Omega_{sv\xi} \hat{\kappa}_s a^{\frac{1}{\frac{\sigma}{\sigma-1} - (\eta_s + \lambda_s)}}}{q_v (r + \delta_k)} \right)^{\frac{1}{1-\hat{\kappa}_s}} \tag{50}$$

$$\Omega_{sv\xi} \hat{\alpha}_s \frac{k_v^{\hat{\kappa}_s} \ell_a^{\hat{\alpha}_s - 1}}{\gamma^{\hat{\alpha}_s} \hat{\delta}_a} + w(1-\tau_a)(1-\delta_a) = (1+r) w (1-\tau_a) \tag{51}$$

$$\ell_a = \left(\frac{\Omega_{sv\xi} \hat{\alpha}_s k^{\hat{\kappa}_s}}{w(1-\tau_a)(r+\delta_a) \gamma^{\hat{\alpha}} \hat{\delta}_a} \right)^{\frac{1}{1-\hat{\alpha}_s}} \tag{52}$$

where we denote $\hat{\delta_a} = \delta_a^{\frac{1}{\frac{\sigma}{\sigma-1} - (\eta+\lambda)} - 1}$, and where we used the steady-state expression of knowledge $a = \left(\frac{\ell_a}{\gamma}\right)^{\alpha} \frac{1}{\delta_a}$.

We can compute ℓ_a as a function of k_v by taking the ratio between the first and third equations above:

$$\frac{\hat{\kappa}_s}{\hat{\alpha}_s} \frac{\ell_a}{k_v \delta_a} = \frac{(r+\delta_k) q_v}{(r+\delta_a) w (1-\tau_a)} \iff w \ell_a = \frac{(r+\delta_k) \delta_a q_v k_v}{(r+\delta_a)(1-\tau_a)} \frac{\alpha_s}{\kappa_s} \tag{53}$$

Equating equation (53) with the previous one gives:

$$k_v = \left[\frac{\kappa_s(r+\delta_a)w(1-\tau_a)}{\alpha_s(r+\delta_k)\delta_a q_v}\right]^{\frac{1-\hat{\alpha}_s}{1-\hat{\kappa}_s-\hat{\alpha}_s}} \left[\frac{\Omega_{sv\xi}\hat{\alpha}_s}{w(1-\tau_a)(r+\delta_a)\gamma^{\hat{\alpha}}\hat{\delta}_a}\right]^{\frac{1}{1-\hat{\kappa}_s-\hat{\alpha}_s}} \tag{54}$$

In steady-state the value function of an incumbent is given by:

$$\begin{aligned}
v(a,k,v,\gamma,\xi) &= \frac{1}{r(1+r)}\Omega_{sv\xi}k_v^{\hat{\kappa}_s}a^{\frac{1}{\frac{\sigma}{\sigma-1}-(\eta_s+\lambda_s)}} - \frac{1}{r}q_v\delta_k k_v - \frac{1}{r}w(1-\tau_a)\ell_a \\
&= \frac{1}{r(1+r)}\frac{q_v(r+\delta_k)}{\hat{\kappa}_s}k_v - \frac{1}{r}q_v\delta_k k_v - \frac{1}{r}\frac{(r+\delta_k)\delta_a q_v k_v}{(r+\delta_a)}\frac{\alpha_s}{\kappa_s} \\
&= \frac{1}{r}q_v\left(\frac{1}{1+r}\frac{(r+\delta_k)}{\hat{\kappa}_s} - \delta_k - \frac{(r+\delta_k)\delta_a}{(r+\delta_a)}\frac{\alpha_s}{\kappa_s}\right)k_v \\
&= \frac{1}{r}q_v^{\frac{-\hat{\kappa}_s}{1-\hat{\kappa}_s-\hat{\alpha}_s}}\left(\frac{1}{1+r}\frac{(r+\delta_k)}{\hat{\kappa}_s} - \delta_k - \frac{(r+\delta_k)\delta_a}{(r+\delta_a)}\frac{\alpha_s}{\kappa_s}\right) \\
&\qquad \times \left[\frac{\kappa_s(r+\delta_a)w(1-\tau_a)}{\alpha_s(r+\delta_k)\delta_a}\right]^{\frac{1-\hat{\alpha}_s}{1-\hat{\kappa}_s-\hat{\alpha}_s}} \left[\frac{\Omega_{sv\xi}\hat{\alpha}_s}{w(1-\tau_a)(r+\delta_a)\gamma^{\hat{\alpha}}\hat{\delta}_a}\right]^{\frac{1}{1-\hat{\kappa}_s-\hat{\alpha}_s}}
\end{aligned}$$

where the second line uses equation (49).

Alternatively a firm that would switch vintages or that had never operated before would have to pay the full cost of the capital in the first period instead of the just the depreciated capital, which would add $(1-\delta_k)k$ to the cost in the initial period:

$$\begin{aligned}
&q_V^{\frac{-\hat{\kappa}_s}{1-\hat{\kappa}_s-\hat{\alpha}_s}}\left[\frac{1}{r}\left(\frac{1}{1+r}\frac{(r+\delta_k)}{\hat{\kappa}_s} - \delta_k - \frac{(r+\delta_k)\delta_a}{(r+\delta_a)}\frac{\alpha_s}{\kappa_s}\right) - 1\right] \\
&\qquad \times \left[\frac{\kappa_s(r+\delta_a)w(1-\tau_a)}{\alpha_s(r+\delta_k)\delta_a}\right]^{\frac{1-\hat{\alpha}_s}{1-\hat{\kappa}_s-\hat{\alpha}_s}} \left[\frac{\Omega_{sV\xi}\hat{\alpha}_s}{w(1-\tau_a)(r+\delta_a)\gamma^{\hat{\alpha}}\hat{\delta}_a}\right]^{\frac{1}{1-\hat{\kappa}_s-\hat{\alpha}_s}} + \chi q_v k_v(1-\delta_k)
\end{aligned} \tag{55}$$

Vintage decision. We now show that the following lemma

Lemma 2 (Vintage decision)**.** *Assume v_s/q_{v_s} is increasing in v_s.*

1. *If a firm updates its vintage, it upgrades to the best vintage V_s.*

2. *There exists a unique $\underline{v}_s$ such that firm upgrade to V_s if and only if $v_s < \underline{v}_s$.*

A sufficient condition for the first bullet point is that the last value function is increasing in v. After simplification of the profit function, one finds that the first component of the value function $q_V^{\frac{-\hat{\kappa}_s}{1-\hat{\kappa}_s-\hat{\alpha}_s}} \left[\frac{1}{r}\left(\frac{1}{1+r}\frac{(r+\delta_k)}{\hat{\kappa}_s} - \delta_k - \frac{(r+\delta_k)\delta_a}{(r+\delta_a)}\frac{\alpha_s}{\kappa_s}\right) - 1\right] \left[\frac{\kappa_s(r+\delta_a)w(1-\tau_a)}{\alpha_s(r+\delta_k)\delta_a}\right]^{\frac{1-\hat{\alpha}_s}{1-\hat{\kappa}_s-\hat{\alpha}_s}} \left[\frac{\Omega_{sV\xi}\hat{\alpha}_s}{w(1-\tau_a)(r+\delta_a)\gamma^{\hat{\alpha}}\hat{\delta}_a}\right]^{\frac{1}{1-\hat{\kappa}_s-\hat{\alpha}_s}}$ is proportional to $\left(\frac{v}{q_v}\right)^{-\frac{\hat{\kappa}_s}{1-\hat{\kappa}_s-\hat{\alpha}_s}}$. Since $\hat{\kappa} > 0$ by assumption, a sufficient condition for the value function to be increasing in v is that $\frac{v}{q_v}$ increases in v.

A sufficient condition for the second bullet point is that the difference between the value function in case of upgrading and not upgrading be decreasing in v:

$$
\begin{aligned}
& q_V^{\frac{-\hat{\kappa}_s}{1-\hat{\kappa}_s-\hat{\alpha}_s}} \left[\frac{1}{r}\left(\frac{1}{1+r}\frac{(r+\delta_k)}{\hat{\kappa}_s} - \delta_k - \frac{(r+\delta_k)\delta_a}{(r+\delta_a)}\frac{\alpha_s}{\kappa_s}\right) - 1\right] \\
& \qquad \times \left[\frac{\kappa_s(r+\delta_a)w(1-\tau_a)}{\alpha_s(r+\delta_k)\delta_a}\right]^{\frac{1-\hat{\alpha}_s}{1-\hat{\kappa}_s-\hat{\alpha}_s}} \left[\frac{\Omega_{sV\xi}\hat{\alpha}_s}{w(1-\tau_a)(r+\delta_a)\gamma^{\hat{\alpha}}\hat{\delta}_a}\right]^{\frac{1}{1-\hat{\kappa}_s-\hat{\alpha}_s}} \\
& - q_v^{\frac{-\hat{\kappa}_s}{1-\hat{\kappa}_s-\hat{\alpha}_s}} \left[\frac{1}{r}\left(\frac{1}{1+r}\frac{(r+\delta_k)}{\hat{\kappa}_s} - \delta_k - \frac{(r+\delta_k)\delta_a}{(r+\delta_a)}\frac{\alpha_s}{\kappa_s}\right) - \chi(1-\delta_k)\right] \\
& \qquad \times \left[\frac{\kappa_s(r+\delta_a)w(1-\tau_a)}{\alpha_s(r+\delta_k)\delta_a}\right]^{\frac{1-\hat{\alpha}_s}{1-\hat{\kappa}_s-\hat{\alpha}_s}} \left[\frac{\Omega_{sv\xi}\hat{\alpha}_s}{w(1-\tau_a)(r+\delta_a)\gamma^{\hat{\alpha}}\hat{\delta}_a}\right]^{\frac{1}{1-\hat{\kappa}_s-\hat{\alpha}_s}}
\end{aligned}
$$

The first term is independent of the current vintage of the firm, v. By the same reasoning as before, the second term is strictly increasing in v under the assumption that v/q_v is strictly increasing in v. Therefore if this condition is satisfied, the difference is strictly decreasing in v. In addition, note that the decision to upgrade is independent of γ and ξ.

In addition, note that the decision to upgrade is independent of γ and ξ.

A1.3.1 Consistency with Stylized Facts.

Energy intensity is decreasing with vintage and research efficiency (conditional on asset/size):

$$\begin{aligned}
\frac{n}{y} &= \frac{1}{m+\tau_e v^{-\epsilon_v}} \frac{\eta_s}{\eta_s+\lambda_s} \frac{C}{y} = \frac{1}{m+\tau_e v^{-\epsilon_v}} \frac{\eta_s}{\eta_s+\lambda_s} \Gamma_{sv} y_{si}^{\frac{1}{\eta_s+\lambda_s}-1} \\
&= \frac{1}{m+\tau_e v^{-\epsilon_v}} \frac{\eta_s}{\eta_s+\lambda_s} \left(\frac{\Omega_{sv\xi}}{\frac{\sigma}{\sigma-1} - (\lambda_s+\eta_s)} \right)^{1-\eta_s-\lambda_s} \Gamma_{sv}^{\eta_s+\lambda_s} (a k_v^{\kappa_s})^{-\frac{1}{\sigma+(1-\sigma)(\lambda+\eta)}} \\
&= \frac{1}{m+\tau_e v^{-\epsilon_v}} \frac{\eta_s}{\eta_s+\lambda_s} \left(\frac{\Omega_{sv\xi}}{\frac{\sigma}{\sigma-1} - (\lambda_s+\eta_s)} \right)^{1-\eta_s-\lambda_s} \tilde{\Gamma}_{sv}^{\eta_s+\lambda_s} (k_v^{\alpha_s-\kappa_s})^{-\frac{1}{\sigma+(1-\sigma)(\lambda+\eta)}} \\
&\qquad \times \left(\frac{(r+\delta_k)}{(r+\delta_a)\gamma\delta_a^{\frac{1}{\alpha_s-1}}} \frac{\alpha_s}{w\kappa_s} q_v \right)^{-\frac{\alpha_s}{\sigma+(1-\sigma)(\lambda+\eta)}}
\end{aligned}$$

where we have used the optimal share of intangibles over tangibles given by

$$\frac{a^{\frac{1}{\alpha_s}}}{q_v k_v} = \frac{\ell_a}{\gamma \delta_a^{\frac{1}{\alpha_s}} q_v k_v} = \frac{(r+\delta_k)}{(r+\delta_a)\gamma\delta_a^{\frac{1}{\alpha_s-1}}} \frac{\alpha_s}{w\kappa_s} \tag{56}$$

and the expression for operating costs given by $C = \Gamma_{sv\xi} y_{si}^{\frac{1}{\eta_s+\lambda_s}}$.

Finally, note that we used equation (39) to find an expression for y as a function of k and a, and we used equation (56) to find an expression of a as a function of k_v. We also assumed that $\tau_a = 0$.

Hence we have

$$\begin{aligned}
\ln\frac{n}{y} &= \ln c_s - \beta_{v1}\ln v + \beta_{v2}\ln(m+\tau_e v^{-\epsilon_v}) + \beta_\gamma \ln\gamma + \beta_k \ln k + \beta_\xi \ln \xi_{si} \\
\text{with} \quad \beta_{v1} &= \frac{\kappa + \alpha_s \epsilon_{qv}}{\sigma + (1-\sigma)(\eta_s+\lambda_s)} \\
\beta_{v2} &= \frac{\eta}{\sigma + (1-\sigma)(\eta_s+\lambda_s)} - 1 \\
\beta_\gamma &= \frac{\alpha}{\sigma + (1-\sigma)(\eta_s+\lambda_s)} \\
\beta_\xi &= \frac{1-\eta_s-\lambda_s}{\sigma + (1-\sigma)(\eta_s+\lambda_s)} \\
\beta_k &= \frac{\alpha_s-\kappa_s}{\sigma + (1-\sigma)(\lambda_s+\eta_s)}
\end{aligned}$$

and $\ln c_{st}$ is a time and sector-specific variable. By assumption:

$$\ln \frac{e}{n} = -\epsilon_v \ln v$$

Combining both equations give the first result of the proposition.

For TFP, we use the production function, the optimal pricing decision and the expression for the operating cost (28) and (29), and recalling that our measure of TFP is estimated from revenues following Asker et al. (2014), and we obtain:

$$\begin{aligned}
py &= \text{TFP} \times \text{C}^{\left(1-\frac{1}{\sigma}\right)(\lambda_s+\eta_s)} \times k_v^{\kappa_s\left(1-\frac{1}{\sigma}\right)} \\
\ln \text{TFP} &= \left(1-\frac{1}{\sigma}\right)\left[\ln a + \kappa_s \ln v_i + \rho \ln A_s\right] + \frac{1}{\sigma}\log \xi_i + \ln P_s + \frac{1}{\sigma}\ln Y_s \\
&\quad + \left(1-\frac{1}{\sigma}\right)\left[\lambda_s \ln\left(\frac{\lambda_s}{\lambda_s+\eta_s}\right) + \eta_s \ln\left(\frac{\eta_s}{\lambda_s+\eta_s}\right) - \eta_s \ln\left(m + \tau_e v^{-\epsilon_v}\right) - \lambda_s \ln w\right]
\end{aligned}$$

where C is the variable cost. We then use equation 56 to substitute for a and we get:

$$\begin{aligned}
\ln \text{TFP} &= \left(1-\frac{1}{\sigma}\right)\alpha_s \ln\left(\frac{1}{\gamma}\right) + \left(1-\frac{1}{\sigma}\right)\alpha_s \ln k + \left(1-\frac{1}{\sigma}\right)\kappa_s \ln v_i + \frac{1}{\sigma}\log \xi_i \\
&\quad - \left(1-\frac{1}{\sigma}\right)\eta_s \ln\left(m + \tau_e v^{-\epsilon_v}\right) + \ln z_{st}
\end{aligned}$$

where $\ln z_{st}$ which is common to all firms in a time, country and sector specific variable which is common to all firms. This shows that firms with newer vintages and higher knowledge intensity also have higher TFP.

A2 Annex: Estimation

A2.1 Internal Calibration of non firm-specific parameters

A sector correspond to a 2-digit SIC industry to allow for enough firms in each sector for us to target firm-level moments. We explain below which moments we target and which moment is most informative for each parameter.

Table A1: Internally Calibrated Parameters and Firms' Characteristics

Parameter	Description	Granularity
g_v	Growth rate of vintage productivity	All countries
ϵ_v	Emissions elasticity to vintage	All countries
κ_f	Cost of operating	Country
ϕ^e	Average emissions per unit of energy	Country × Sector
κ_e	Cost of entry	Country × Sector
β	Expenditure share	Country × Sector
α	Knowledge elasticity to research	Country × Sector
ρ	Knowledge spillover	Country × Sector
κ	Capital elasticity of output	Country × Sector
η	Energy elasticity of output	Country × Sector
λ	Labor elasticity of output	Country × Sector
$G(\gamma, \xi)$	Distribution of initial parameters	Country × Sector
ξ	Consumers' taste	Firm
γ	Research efficiency	Firm
v	Vintage of capital	Firm

Consistent with the households' optimality condition (3), we calibrate the elasticity of utility to the consumption in each sector, β_s, to match the share of sales of each sector separately for each country: $\beta_s = \frac{\sum_{i\in\Omega_s} p_i y_i}{\sum_s \sum_{i\in\Omega_s} p_i y_i}$. To ensure that our model is consistent with the overall empirical distribution of economic activities, we reweight our shares to match coarser sectoral shares.[33]

The elasticity of output to capital κ_s is identified using the average mark-up, following the following formula: Sales $= \frac{\sigma}{\sigma-1}\frac{\text{COGS}}{\eta_s+\lambda_s}$, where COGS stands for "cost of goods sold". Using the assumption of constant returns to scale, $\kappa_s = 1-\lambda_s-\eta_s-\alpha_s$, and the average sales and cost of goods sold, from our dataset, we start with estimating $\kappa_s+\alpha_s$ as follows: $\kappa_s+\alpha_s =$

---

[33] We use value added shares reported by the OECD (https://data.oecd.org/natincome/value-added-by-activity.htm)

$1-\frac{\sigma}{\sigma-1}\frac{\sum_{i\in\Omega_s}\mathrm{COGS}_i}{\sum_{i\in\Omega_s}\mathrm{Sales}_i}$ where σ is externally calibrated. This captures the share of income going to the owners of the firm, of the stock of physical capital and of the intangible capital. The elasticities of output to labor and energy (λ_s, η_s) are estimated using the first order conditions of the firm, which imply that they are related to the share of costs going to each factor. Using our previous estimate of $(\kappa_s + \alpha_s)$ we obtain each parameter from the following expressions

$$\eta_s = (1-\kappa_s-\alpha_s)\times\frac{A}{A+B} \quad \text{and} \quad \lambda_s = (1-\kappa_s-\alpha_s)\times\frac{B}{A+B}$$

with A the expenditure share of COGS on energy, and B the expenditure share of COGS on labor and other variable inputs. Denoting C the ratio of R&D spending on COGS, we can compute α_s using our estimates of η_s and λ_s by using the first-order condition for research, $\alpha_s = \frac{(\delta_a+r)C}{\delta_a(\eta_s+\lambda_s)}$. Using again the assumption of CRS we obtain κ_s as follows: $\kappa_s = 1-\alpha_s-\lambda_s-\eta_s$.

For the knowledge spillover parameter ρ, we follow Griliches (1992) who reviews the literature and Bloom et al. (2013) who provide estimates of the social and private returns of R&D. Their estimates imply an industry-wide R&D elasticity of output of the same order of magnitude (between half and double) as the elasticity to the firm-specific stock of R&D. We therefore assume that both parameters are equal in each sector and country, i.e. $\rho_s = \alpha_s$.

We then assume that the productivity of vintages grows at a constant rate over time, denoted g_v, and that there is a simple relationship between the age of the capital stock and the productivity of vintages given by

$$v(\text{age of capital}_i) = v_0(1+g_v)^{-\text{age of capital}_i}. \tag{57}$$

We estimate g_v by running a regression of firms' productivity (TFPR), which we construct using our previous parameters, on the age of the capital stock, and other controls such as the share of intangibles, and total log sales. As can be seen from combining Equation 21 and Equation 57 and shown in Appendix A1.3.1, the coefficient on the age of the capital stock is related one-for-one to g_v. We estimate a 11.9% growth rate in the productivity of capital goods every year, which is consistent with values found in the literature analyzing the contribution of improvements in capital goods to long-run growth, Greenwood et al. (1997).

To estimate the elasticity of emissions to capital vintage, ϵ_v, we start from equation (6): $\log\left[\frac{\text{Emission}}{\text{Energy}}\right] = \log(\phi_s^e) - \epsilon_v \log v$. We thus run the following regression (see Table A2) with sector, country, and time fixed effects:

$$\log\left[\frac{\text{Emission}_{it}}{\text{Energy}_{it}}\right] = b_s + b_j + b_t + b_{age} \times \text{Age of Capital}_{it} + \epsilon_{sjti}. \tag{58}$$

We then compute ϵ_v as $\epsilon_v = \frac{\hat{b}_{age}}{\log(1+g_v)}$ and $\phi_s^e = \exp(b_s + b_j)$.

The joint distribution of γ and ξ from which firms draw their initial characteristics upon entering, $G(\gamma,\xi)$, is estimated as follows: we first estimate the pair (γ_i,ξ_i) for each firm within a sector (see below), we then define the sample space as $\{\gamma_i\}_{i\in\Omega_s} \times \{\xi_i\}_{i\in\Omega_s}$ and assume that the probability that a potential entrant draws any pair in this set is uniform: $G(\gamma,\xi) \sim U\left(\{\gamma_i\}_{i\in\Omega_s} \times \{\xi_i\}_{i\in\Omega_s}\right)$.

The cost of operating a firm, κ_f, is common to all firms in a country and is calibrated so that the least profitable firm in a given country is indifferent between staying and exiting the market. The cost of starting a business is common to all firms in a sector and is calibrated so that a potential entrant is indifferent between entering and not, i.e. $\hat{\kappa}_e = \mathbb{E}_{\hat{G}_s}\left[\mathcal{V}_{si}(k_{V_s}, V_s, \gamma, \xi)\right]$ where we use our estimated distribution G_s to compute the expectation and the value function $\mathcal{V}_{si}(.,.,.,.)$ is obtained by solving the model.

A2.2 Internal Calibration of Firm-specific Variables

In the last step of the calibration, we estimate the three firm-level state variables v, γ, and ξ to match three firm-level moments. The research-efficiency parameter, γ, is calibrated to match the ratio of intangibles over tangible assets, and is given by:[34]

$$\gamma = \frac{1}{\text{Ratio of Intangibles over Tangibles}} \frac{(r+\delta_k)}{(r+\delta_a)\delta_a^{\frac{1}{\alpha_s-1}}} \frac{\alpha_s}{w\kappa_s}. \tag{59}$$

As we have explained above, the productivity of vintages is assumed to grow at a constant rate over time, g_v, so that we can estimate v based on the age of the capital stock using $v = (1+g_v)^{-\text{age of capital}_i}$.

To calibrate the consumers' taste, we use the relationship between the relative size of the firm and the relative vintage productivity, research efficiency, and consumers' taste. Given

---

[34] Proofs can be found in Appendix A2.

our estimates of γ and of v from the previous steps, we can recover ξ as follows:

$$\ln\left[\frac{\xi_{si}}{\xi_{sj}}\right] = \ln\left[\frac{\text{COGS}_i}{\text{COGS}_j}\right] - \alpha_s(\sigma_s - 1)\ln\frac{\gamma_j}{\gamma_i} \\ + \kappa_s(\sigma_s - 1)\ln\left[1 + g_v\right](\text{age of capital}_i - \text{age of capital}_j) \tag{60}$$

We normalize the levels of ξ such that the mean across firms within each sector is 1.

A2.3 Research efficiency parameter

We show that the firm research efficiency parameter, γ, is inversely related to the ratio of intangibles over tangibles assets:

$$\gamma = \frac{1}{\text{Ratio of Intangibles over Tangibles}}\frac{(r+\delta_k)}{(r+\delta_a)\delta_a^{\frac{1}{\alpha_s-1}}}\frac{\alpha_s}{w\kappa_s} \tag{61}$$

We set $\tau_a = 0$ and normalize $w = 1$.

Proof. We start from the following equations

$$w\ell_a = \frac{(r+\delta_k)\delta_a q_v k_v}{(r+\delta_a)(1-\tau_a)}\frac{\alpha_s}{\kappa_s}$$
$$a = \left(\frac{\ell_a}{\gamma}\right)^{\alpha}\frac{1}{\delta_a}$$

Combining these two gives the following expression for the ratio of intangibles over tangible capitals $a^{1/\alpha_s}/qk$:

$$\frac{a^{\frac{1}{\alpha_s}}}{q_v k_v} = \frac{\ell_a}{\gamma\delta_a^{\frac{1}{\alpha_s}} q_v k_v} = \frac{(r+\delta_k)}{(r+\delta_a)\gamma\delta_a^{\frac{1}{\alpha_s-1}}}\frac{\alpha_s}{w\kappa_s}$$

□

A2.4 Consumers' taste parameter

We now want to derive the following expression when the carbon tax is 0:

$$\ln\left[\frac{\xi_{si}}{\xi_{sj}}\right] = \ln\left[\frac{\text{COGS}_i}{\text{COGS}_j}\right] - \alpha_s(\sigma_s - 1)\ln\frac{\gamma_j}{\gamma_i} + \kappa_s(\sigma_s - 1)\ln\left[1 + g_v\right](\text{age of capital}_i - \text{age of capital}_j)$$

Proof. We start from the following set of equilibrium conditions and definitions

$$\begin{aligned}
k_v &= \left[\frac{\kappa_s(r+\delta_a)w(1-\tau_a)}{\alpha_s(r+\delta_k)\delta_a q_v}\right]^{\frac{1-\hat{\alpha}_s}{1-\hat{\kappa}_s-\hat{\alpha}_s}}\left[\frac{\Omega_{sv\xi}\hat{\alpha}_s}{w(1-\tau_a)(r+\delta_a)\gamma^{\hat{\alpha}}\hat{\delta}_a}\right]^{\frac{1}{1-\hat{\kappa}_s-\hat{\alpha}_s}} \\
\Omega_{sv\xi}k_v^{\hat{\kappa}_s}a^{\frac{1}{\frac{\sigma}{\sigma-1}-(\eta_s+\lambda_s)}} &= \left(\frac{\sigma_s}{\sigma_s-1}\frac{1}{(\eta_s+\lambda_s)}-1\right)C \\
\Omega_{sv\xi} &= \left[\frac{\sigma-1}{\sigma}(1+\tau_y)(\lambda_s+\eta_s)P_s(\xi_{si}Y_s)^{\frac{1}{\sigma}}\right]^{\frac{\sigma}{\sigma+(1-\sigma)(\eta_s+\lambda_s)}}\left(\frac{\sigma}{\sigma-1}\frac{1}{(\eta_s+\lambda_s)}-1\right)\tilde{\Gamma}_{sv}^{\frac{(1-\sigma)(\eta_s+\lambda_s)}{\sigma+(1-\sigma)(\eta_s+\lambda_s)}}
\end{aligned}$$

Consider a pair of firms i and j within the same country and sector. Combining the expression above gives

$$\begin{aligned}
C &= \frac{\Omega_{sv\xi}k_v^{\hat{\kappa}_s}a^{\frac{1}{\frac{\sigma}{\sigma-1}-(\eta_s+\lambda_s)}}}{\left(\frac{\sigma_s}{\sigma_s-1}\frac{1}{(\eta_s+\lambda_s)}-1\right)} = \frac{\frac{q_v(r+\delta_k)}{\hat{\kappa}_s}k_v}{\left(\frac{\sigma_s}{\sigma_s-1}\frac{1}{(\eta_s+\lambda_s)}-1\right)} \\
C_i/C_j &= \frac{k_{vi}}{k_{vj}} \\
&= \left(\frac{\Omega_{sv\xi j}\gamma_j^{\hat{\alpha}}}{\Omega_{svi}\gamma_j^{\hat{\alpha}}}\right)^{\frac{1}{1-\hat{\kappa}_s-\hat{\alpha}_s}} \\
&= \left(\frac{\xi_{si}^{\frac{1}{\sigma_s+(1-\sigma_s)(\lambda_s+\eta_s)}}\tilde{\Gamma}_{svi}^{\frac{(1-\sigma_s)(\lambda_s+\eta_s)}{\sigma_s+(1-\sigma_s)(\eta_s+\lambda_s)}}\gamma_j^{\hat{\alpha}}}{\xi_{sj}^{\frac{1}{\sigma_s+(1-\sigma_s)(\lambda_s+\eta_s)}}\gamma_j^{\hat{\alpha}}\tilde{\Gamma}_{svj}^{\frac{(1-\sigma_s)(\lambda_s+\eta_s)}{\sigma_s+(1-\sigma_s)(\eta_s+\lambda_s)}}\gamma_i^{\hat{\alpha}}}\right)^{\frac{1}{1-\hat{\kappa}_s-\hat{\alpha}_s}} \\
&= \frac{\xi_{si}\tilde{\Gamma}_{svi}^{(1-\sigma_s)(\lambda_s+\eta_s)}\gamma_j^{\alpha(\sigma-1)}}{\xi_{sj}\tilde{\Gamma}_{svj}^{(1-\sigma_s)(\lambda_s+\eta_s)}\gamma_i^{\alpha(\sigma-1)}} \\
&= \frac{\xi_{si}v_{svi}^{(\sigma_s-1)\kappa_s}\gamma_j^{\alpha_s(\sigma-1)}}{\xi_{sj}v_{svj}^{(\sigma_s-1)\kappa_s}\gamma_i^{\alpha_s(\sigma-1)}}
\end{aligned}$$

where we assume that the carbon tax is zero or negligible relative to the price of energy. Taking logs gives the result.

□

A2.5 Elasticity of emissions to capital vintage

Below we show the regression with sector, country, and time fixed effects:

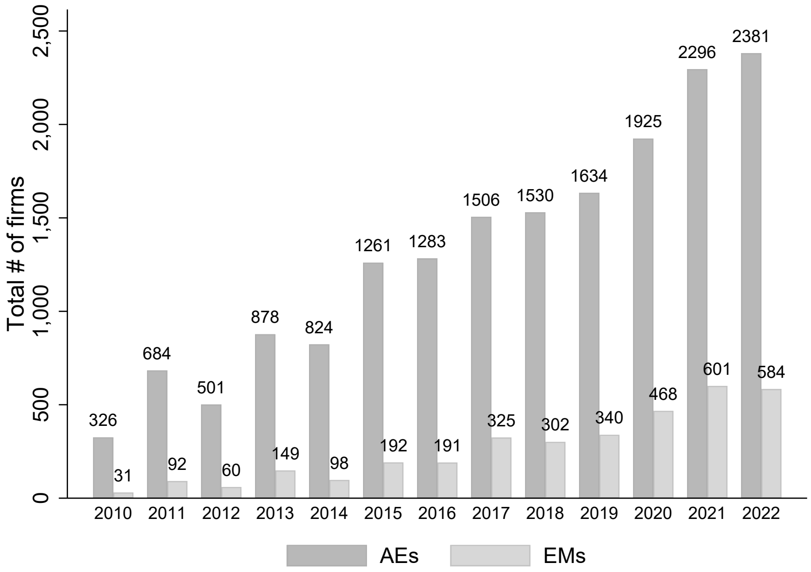
Table A2: Calibration Elasticity of Emissions to Capital Vintages

	(1) log (S1S2 / Energy)
Age of capital	0.03***
	(0.01)
N	5308
R^2	0.12
Adj-R^2	0.11
Industry+country+year FE	Yes

Notes: Industry classification: SIC-2. Standard errors in parentheses, clustered at the country + industry + year. * $p < .1$, ** $p < .05$, *** $p < .01$

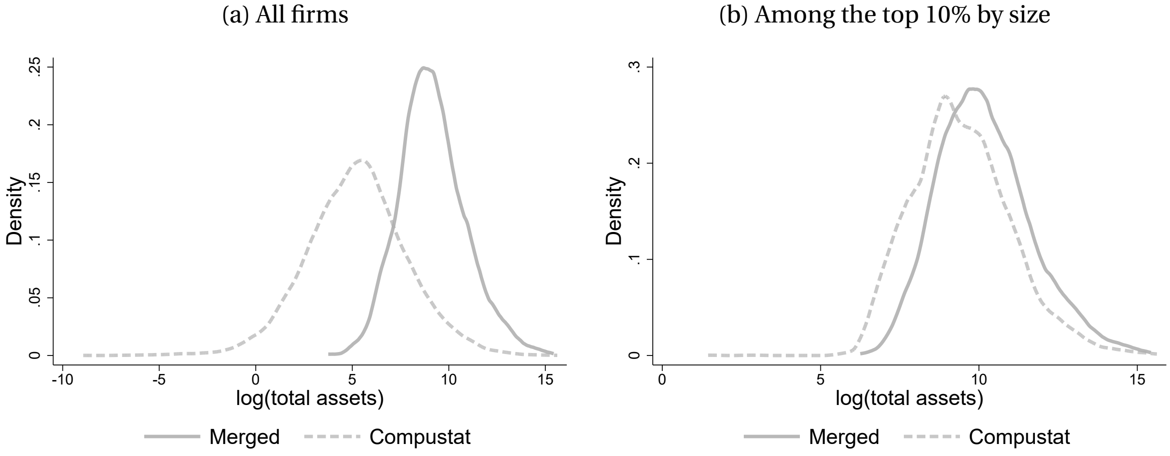
A3 Annex: Figures and Tables

Figure A1: Sample Over Time



Notes: This figure plots the number of firms in the matched Compustat-ICE Data Services sample over time (with non-missing emission intensity) for firms headquartered in advanced economies (AE) and emerging markets (EM).

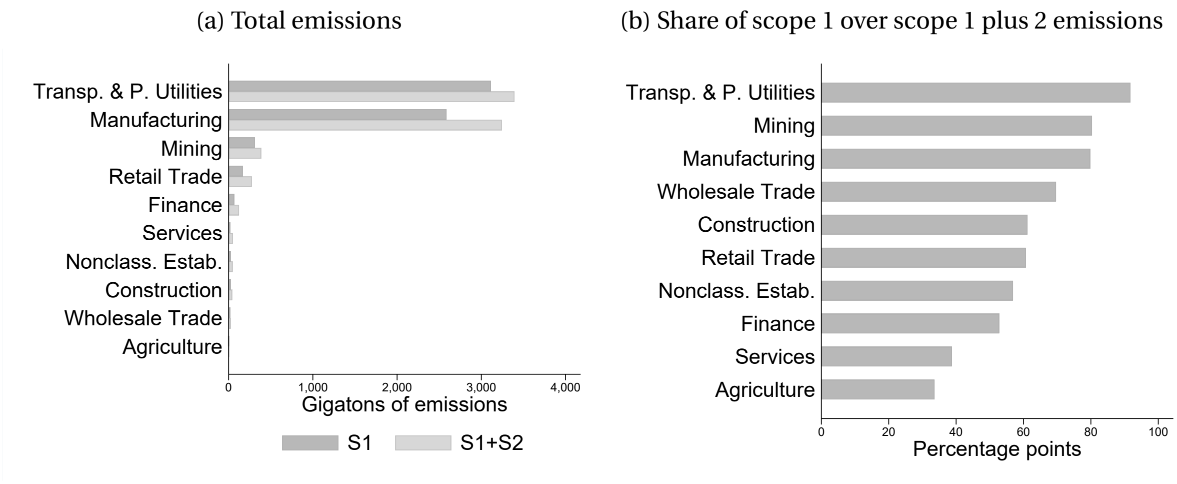
Figure A2: Firm Size: Compustat vs Compustat-ICE Data Services Merged Sample


(a) All firms

(b) Among the top 10% by size

Notes: This figure illustrates the size distribution (log assets) for the Compustat and Compustat-ICE Data Services merged samples.

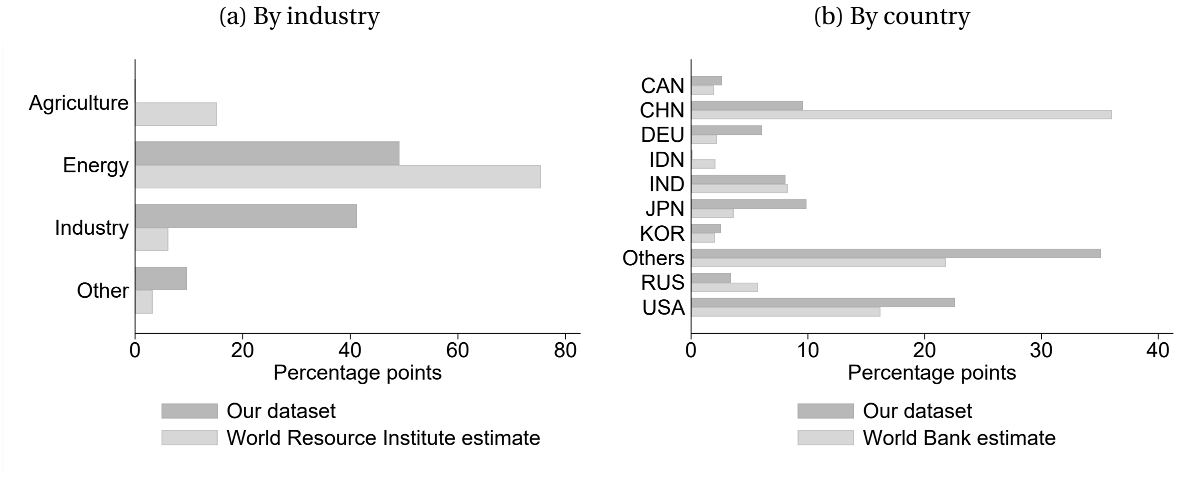
Figure A3: Emissions by Industry


(a) Total emissions

(b) Share of scope 1 over scope 1 plus 2 emissions

Notes: For each industry, panel (a) shows the total emissions generated, while panel (b) plots the shares of scope 1 over scope 1 plus 2 emissions. S1 = scope 1; S1S2 = scopes 1 and 2. Data by all firms in the merged dataset for the year 2019 are used. Agriculture = Agriculture, Forestry, & Fishing; Transp. & P. Utilities = Transportation & Public Utilities; Finance = Finance, Insurance, & Real Estate; Nonclass. Estab. = Nonclassifiable Establishments. 2019 data is used.

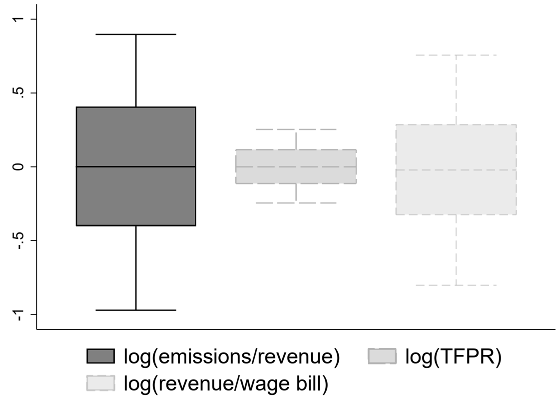
Figure A4: Total Emissions: Compustat-ICE Data Services Merged Sample vs Other Sources


(a) By industry

(b) By country

Notes: Panel (a) compares the shares of total scope 1 emissions emanating from each industry group. Panel (b) compares the shares of total scopes 1 and 2 emissions emanating from each country. 2019 data is used.

Figure A5: Hetereogeneity in Emission Intensity and Productivity

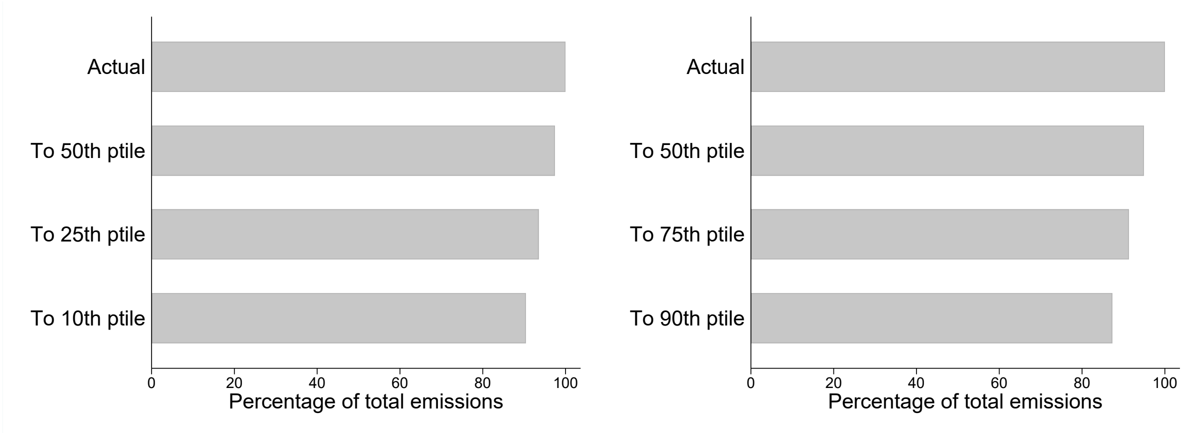


Notes: The figure reports box plots of the log of emission intensity (emissions over revenues), the log of productivity (revenue TFPR), and the log of labor productivity (revenues over wage bill), after controlling for industry × country fixed effects. The boxes illustrate the 25th, 50th, and 75th percentiles of the distributions. Whiskers are drawn to span all data points within 1.5 IQR of the nearer quartile. 2019 data is used. Finance, public administration, and utilities sectors are excluded from the calculation.

Figure A6: Emission Counterfactuals for Specific Channels

(a) Age of Capital

(b) Share of Intangible Capital



Notes: This figure illustrates actual emissions of firms in our sample, together with the counterfactual emissions that we would observe if every firm (a) had at most the same age of capital as the firm in the Xth percentile and (b) had at least the same knowledge intensity (share of intangible capital) as the firm in the percentile. Only industry-country groups with at least 4 firms are included in panel. 4-digit SIC industry classification and 2019 data used. Finance, public administration, and utilities sectors are excluded from the calculation.

Table A3: Emissions over Energy and Firm Characteristics: Alternative Industry Classifications

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Emissions over Energy (STD log emissions / energy)							
STD age of capital	0.10***	0.09***	0.09***	0.03**	0.03**	0.09***	0.07***	0.13***
	(0.02)	(0.02)	(0.02)	(0.01)	(0.01)	(0.02)	(0.02)	(0.04)
STD share intangibles	-0.01	0.00	0.00	0.08***	0.07***	0.01	0.01	0.05
	(0.03)	(0.03)	(0.03)	(0.02)	(0.01)	(0.02)	(0.02)	(0.04)
STD TFPR	0.01	0.01	0.01	-0.08***	-0.09***	0.01	-0.01	0.04
	(0.03)	(0.02)	(0.02)	(0.02)	(0.01)	(0.02)	(0.02)	(0.05)
STD log(assets)	-0.07**	-0.07**	-0.05**	0.03	0.03**	-0.05*	-0.03	-0.07*
	(0.03)	(0.03)	(0.03)	(0.02)	(0.02)	(0.03)	(0.02)	(0.04)
N	2722	2722	2722	6148	6148	3491	3491	942
R^2	0.52	0.49	0.42	0.35	0.29	0.46	0.33	0.37
Adj-R^2	0.23	0.23	0.22	0.17	0.15	0.19	0.19	0.12
Industry × country × year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry classification	SIC4	SIC3	SIC2	GICS4	GICS2	NAICS4	NAICS2	HP

Notes: Industry classification: SIC4 = 4-digit Standard Industrial Classification; GICS2 = 2-digit Global Industry Classification Standard; NAICS6 = 6-digit North American Industry Classification System; HP = Hoberg-Phillips 500. All variables are standardized to have mean zero and standard deviation of one. Standard errors in parentheses, clustered at the country × industry × year. Energy, utilities, finance, and public sectors are excluded from the calculations. * $p < .1$, ** $p < .05$, *** $p < .01$

Table A4: Emissions over Energy and Firm Characteristics: Alternative Indicators

	(1)	(2)	(3)	(4)
	Emissions over Energy (STD log emissions / energy)			
STD Age	0.05***			0.05***
	(0.02)			(0.02)
STD log(RD / assets)		-0.05		-0.07**
		(0.03)		(0.03)
STD log(EBIT/assets)			0.03	0.03
			(0.02)	(0.02)
STD log(assets)	-0.13***	-0.11***	-0.10***	-0.13***
	(0.03)	(0.03)	(0.03)	(0.03)
N	2596	2596	2596	2596
R^2	0.52	0.52	0.52	0.52
Adj-R^2	0.24	0.23	0.23	0.24
Industry × country × year FE	Yes	Yes	Yes	Yes

Notes: All variables are standardized to have mean zero and standard deviation of one. Standard errors in parentheses, clustered at the country × industry × year. Finance, public administration, and utilities sectors are excluded from the calculations. By including only industry × country × year groups with more than one firm, our sample size is reduced from 7,032 to 2,596. * $p < .1$, ** $p < .05$, *** $p < .01$

Table A5: Emission Intensity and Firm Characteristics: Robustness

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
	Emission Intensity (STD log emissions / revenue(t-1))															
	Largest 50% of firms				Adding financial controls				Removing the Covid period				Scope 1 emissions			
STD age of capital	0.05***			0.04**	0.04***			0.03**	0.07***			0.04***	0.04***			0.02*
	(0.02)			(0.02)	(0.02)			(0.01)	(0.02)			(0.02)	(0.01)			(0.01)
STD share intangibles		-0.15***		-0.13***		-0.23***		-0.20***		-0.17***		-0.14***		-0.15***		-0.13***
		(0.02)		(0.02)		(0.02)		(0.02)		(0.02)		(0.02)		(0.01)		(0.01)
STD TFPR			-0.14***	-0.10***			-0.16***	-0.11***			-0.18***	-0.14***			-0.14***	-0.11***
			(0.02)	(0.02)			(0.02)	(0.02)			(0.02)	(0.02)			(0.01)	(0.01)
STD log(assets)	0.03	0.04	0.04	0.04	0.32***	0.46***	0.36***	0.48***	0.05***	0.09***	0.13***	0.14***	0.09***	0.12***	0.15***	0.16***
	(0.03)	(0.03)	(0.03)	(0.03)	(0.05)	(0.05)	(0.04)	(0.05)	(0.01)	(0.01)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)
N	2534	2534	2534	2534	3802	3802	3802	3802	3564	3564	3564	3564	6134	6134	6134	6134
R^2	0.82	0.82	0.82	0.83	0.83	0.85	0.84	0.85	0.80	0.81	0.81	0.82	0.83	0.83	0.83	0.84
Adj-R^2	0.71	0.72	0.72	0.73	0.73	0.76	0.75	0.77	0.69	0.71	0.71	0.72	0.73	0.74	0.74	0.75
Industry × country × year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: All variables are standardized to have mean zero and standard deviation of one. Standard errors in parentheses, clustered at the country × industry × year. Finance, public administration, and utilities sectors are excluded from the calculations. Financial controls include lagged liquidity, leverage, and capitalization ratios, and the market share. The Covid period is defined as 2020 and afterwards. By including only industry × country × year groups with more than one firm, our sample size is reduced from 6,950 to 2,534 for regressions (1-4), from 9,271 to 3,802 for regressions (5-8), from 8,310 to 3,564 for regressions (9-12), and from 13,157 to 6,134 for regressions (13-16). * $p < .1$, ** $p < .05$, *** $p < .01$

Table A6: Emissions over Energy and Firm Characteristics: Robustness

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
	Emissions over Energy (STD log emissions / energy)															
	Largest 50% of firms				Adding financial controls				Removing the Covid period				Scope 1 emissions			
STD age of capital	0.05*			0.05*	0.05			0.05	0.10***			0.11***	0.04**			0.05**
	(0.03)			(0.03)	(0.03)			(0.03)	(0.03)			(0.03)	(0.02)			(0.02)
STD share intangibles		0.05		0.04		0.05		0.03		-0.01		-0.02		-0.00		-0.02
		(0.04)		(0.04)		(0.03)		(0.03)		(0.03)		(0.03)		(0.02)		(0.02)
STD TFPR			0.04	0.03			0.06*	0.05			-0.01	0.01			0.04*	0.05**
			(0.04)	(0.04)			(0.03)	(0.03)			(0.03)	(0.03)			(0.02)	(0.02)
STD log(assets)	-0.09*	-0.10*	-0.09*	-0.10*	-0.11	-0.15	-0.13	-0.15	-0.07**	-0.06*	-0.06*	-0.07**	0.03	0.03	0.02	0.01
	(0.05)	(0.05)	(0.05)	(0.05)	(0.11)	(0.11)	(0.11)	(0.11)	(0.03)	(0.03)	(0.03)	(0.03)	(0.02)	(0.02)	(0.02)	(0.03)
N	1278	1278	1278	1278	1569	1569	1569	1569	2187	2187	2187	2187	2495	2495	2495	2495
R^2	0.47	0.47	0.47	0.47	0.55	0.55	0.55	0.55	0.51	0.50	0.50	0.51	0.66	0.66	0.66	0.66
Adj-R^2	0.14	0.14	0.14	0.14	0.26	0.26	0.27	0.27	0.22	0.21	0.21	0.22	0.45	0.45	0.45	0.45
Industry × country × year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: All variables are standardized to have mean zero and standard deviation of one. Standard errors in parentheses, clustered at the country × industry × year. Finance, public administration, and utilities sectors are excluded from the calculations. Financial controls include lagged liquidity, leverage, capitalization ratios, and the market share. The Covid period is defines as 2022 and afterwards. By including only industry × country × year groups with more than one firm, our sample size is reduced from 4,189 to 1,278 for regressions (1-4), from 4,709 to 1,569 for regressions (5-8), from 6,017 to 2,187 for regressions (9-12), and from 6,711 to 2,495 for regressions (13-16). * $p < .1$, ** $p < .05$, *** $p < .01$

Table A7: Emission Intensity and Firm Characteristics: AEs vs EMDEs

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Emission Intensity (STD log emissions / revenue(t-1))							
	Advanced economies				Emerging markets			
STD age of capital	0.05***			0.02**	0.17***			0.16***
	(0.01)			(0.01)	(0.05)			(0.04)
STD share intangibles		-0.16***		-0.13***		-0.23***		-0.14***
		(0.01)		(0.01)		(0.05)		(0.04)
STD TFPR			-0.18***	-0.15***			-0.32***	-0.27***
			(0.01)	(0.01)			(0.05)	(0.05)
STD log(assets)	0.04***	0.08***	0.12***	0.13***	0.01	0.07	0.22***	0.23***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.05)	(0.05)	(0.06)	(0.06)
N	6089	5917	6027	5832	796	765	792	742
R^2	0.79	0.80	0.80	0.81	0.79	0.79	0.81	0.81
Adj-R^2	0.68	0.70	0.70	0.72	0.64	0.63	0.68	0.67
Industry × country × year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: All variables are standardized to have mean zero and standard deviation of one. Standard errors in parentheses, clustered at the country × industry × year. Finance, public administration, and utilities sectors are excluded from the calculations. * $p < .1$, ** $p < .05$, *** $p < .01$

Table A8: Emission Intensity and Firm Characteristics: Emission Intensity using Assets

	(1)	(2)	(3)	(4)
	Emission Intensity (STD log emissions / assets(t-1))			
STD age of capital	0.06***			0.05***
	(0.01)			(0.01)
STD share intangibles		-0.19***		-0.16***
		(0.01)		(0.01)
STD TFPR			-0.14***	-0.09***
			(0.01)	(0.01)
STD log(assets)	-0.02*	0.02**	0.05***	0.05***
	(0.01)	(0.01)	(0.01)	(0.01)
N	6577	6577	6577	6577
R^2	0.79	0.81	0.80	0.82
Adj-R^2	0.68	0.71	0.69	0.72
Industry × country × year FE	Yes	Yes	Yes	Yes

Notes: All variables are standardized to have mean zero and standard deviation of one. Standard errors in parentheses, clustered at the country × industry × year. Finance, public administration, and utilities sectors are excluded from the calculations. By including only industry × country × year groups with more than one firm, our sample size is reduced from 13,949 to 6,577. * $p < .1$, ** $p < .05$, *** $p < .01$

Table A9: Emission Intensity and Age of Capital (using growth rate of assets)

	Emission Intensity	Age of capital	Emission Intensity	Emission Intensity
Age of capital	0.0198*			0.112***
	(0.0106)			(0.0408)
5-yr growth rate assets		-1.336***	-0.149***	
		(0.0954)	(0.0542)	
N	4,262	4,262	4,262	4,262
R^2	0.849	0.629	0.849	0.086
Industry × Year × Country FE	Yes	Yes	Yes	Yes

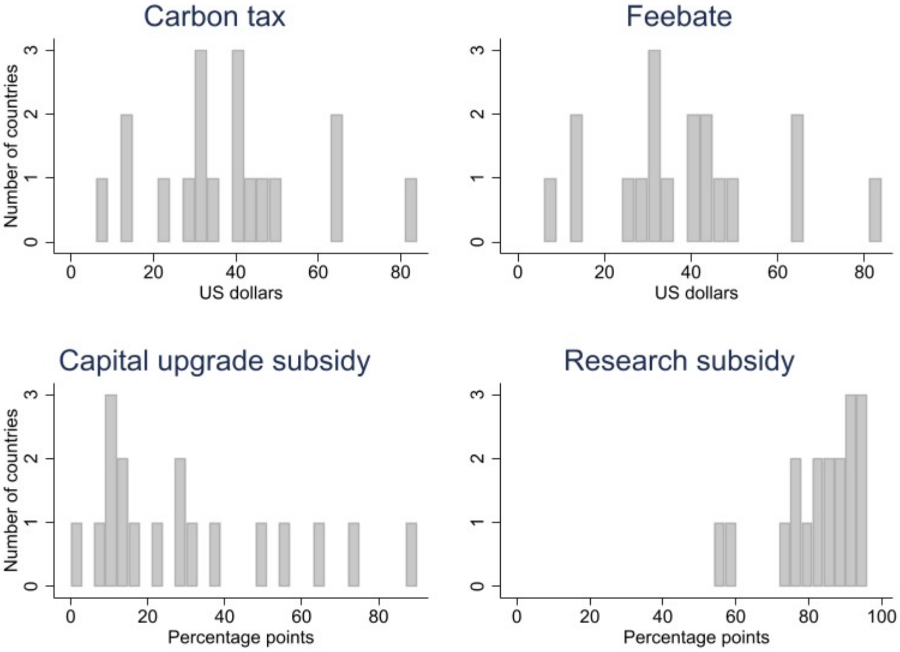
Notes: All regressions include TFP and size of the firm as control (measured with the level of assets, in logs). For the last regression, the F-statistics first stage is 368.06. Standard errors clustered at Industry × year × country (using SIC4 industry classification), * $p < .1$, ** $p < .05$, *** $p < .01$

Table A10: Emission Intensity and Age of Capital (using growth rate of sales)

	Emission Intensity	Age of capital	Emission Intensity	Emission Intensity
Age of capital	0.0203*			0.135***
	(0.0106)			(0.0402)
5-yr growth rate sales		-1.401***	-0.189***	
		(0.102)	(0.0548)	
N	4,254	4,254	4,254	4,254
R^2	0.849	0.627	0.850	0.068
Industry × Year × Country FE	Yes	Yes	Yes	Yes

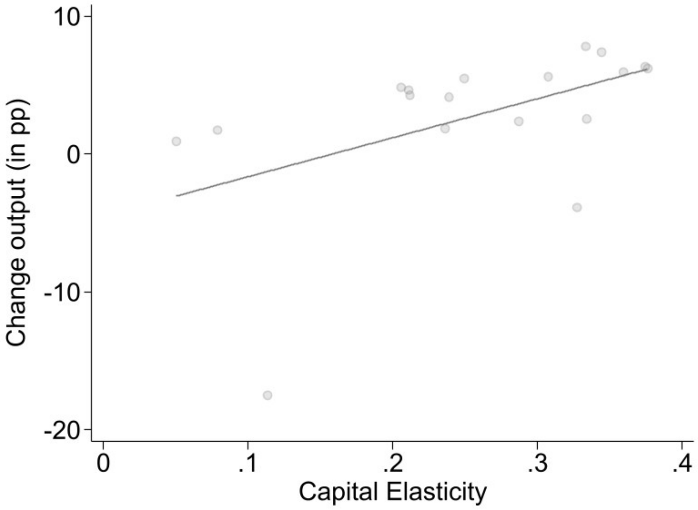
Notes: All regressions include TFP and size of the firm as control (measured with the level of assets, in logs). For the last regression, the F-statistics first stage is 357.42 . Standard errors clustered at Industry × year × country (using SIC4 industry classification), * $p < .1$, ** $p < .05$, *** $p < .01$

Figure A7: Distribution of Taxes and Subsidies in Counterfactuals



Notes: For each policy instrument, we solve for the value it should take in each country (in US dollars or percentage points) to achieve a 15% reduction in emissions relative to the baseline (a no-policy scenario). Each histogram shows, for a different policy, the number of countries (y-axis) for which each policy value (x-axis) is necessary to achieve the emissions reduction target.

Figure A8: Relationship between Capital elasticity and effect on GDP of capital subsidies



Notes: This plot shows a binscatter plot for the cross-section of countries of the average capital elasticity of output and the changes in the steady-state output with respect to the baseline after a capital subsidy calibrated such that emissions decrease by 15%.

Table A11: Effects of policies on aggregates (lower capital quality)

In Percentages	Carbon Tax	Carbon Feebate	Capital Subsidy	Research Subsidy
Output	-0.25	-0.24	2.8	-7.7
Consumption	-0.25	-0.24	2.83	-7.7
Profits	-0.4	0.0	2.7	-7.8
NPV of Consumption	-1.6	-1.6	-3.4	-4.5

Notes: In percentage change of the actual economy. NPV=Net Present Value. We use a 4% time discount factor to compute the net present value of consumption. We reduce the quality of the initial distribution of capital vintages (we divide capital productivity of all firms, v, by 2), keeping everything else constant.

Table A12: Effects of policies on upgrading and distribution of emission intensities (lower capital quality)

	Carbon Tax	Feebate	Capital Subsidy	Research Subsidy
Share of firms upgrading (in %)	4.4	4.4	45.3	0.0

Notes: Changes are given in percentage change relative to the actual economy. The variance is weighted by firm sales. Quantiles are moments of the distribution of emission intensities, also weighted by each firm' sales. We reduce the quality of the initial distribution of capital vintages (we divide capital productivity of all firms, v, by 2), keeping everything else constant.

PUBLICATIONS

Mitigating Climate Change at the Firm Level: Mind the Laggards

Working Paper No. WP/2023/242