The Consequences of Missing the Trees for the Forest*

Michael B. Nattinger
University of Wisconsin - Madison
Alexander H. von Hafften
University of Wisconsin - Madison

August 18, 2025

Abstract

We document that the heterogeneity of carbon emission intensity across U.S. public firms is substantial and persistent. Using a tractable heterogeneous-firm GE model, we show that heterogeneity in carbon dependence is quantitatively relevant for determining the socially optimal carbon tax. With heterogeneity in this dimension and endogenous entry and exit, the firm distribution becomes endogenously greener in terms of production technology in response to carbon taxes - a composition effect which is absent from models without this heterogeneity. When we shut down this channel, the model-implied Pigouvian tax which implements the socially optimal allocation is 86 percent higher than when the composition effect is present, suggesting that integrated assessment models abstracting from this form of heterogeneity may recommend carbon taxes that are larger than optimal. Nevertheless, we find that the welfare consequences of implementing the too-large carbon taxes are much smaller than the consequences of ignoring climate change altogether.

*We would like to thank Manuel Amador, Carter Braxton, Dean Corbae, Dmitry Orlov, Lubos Pastor, Martin Schneider, Ananth Seshadri, and Ken West for their helpful comments as well as seminar participants at the Society for Economic Dynamics 2024 Annual Meeting, Spring 2023 and Fall 2022 UW-Madison Finance Brownbags, Fall 2023 UW-Madison Finance Brownbag, and Fall 2023 Wisconsin-Minnesota Joint Workshop for their helpful comments. This material is based upon work supported by the National Science Foundation Graduate Research Fellowship under Grant No. DGE-2137424. Any opinion, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

1 Introduction

The effects of climate change, in large part driven by carbon emissions, are increasingly visible across the globe. 2023 marked the hottest year on Earth ever recorded,[1] resulting in record-breaking heat waves across Asia, extreme wildfires in Canada, California, and Hawaii, and destructive flooding in Libya. Spearheaded by William Nordhaus, economists have long explored the socially optimal policy response to the interaction between the global macroeconomy and the Earth's climate. Important contributions to the climate-macro literature, such as Golosov, Hassler, Krusell, and Tsyvinski (2014) and Krusell and Smith (2022), assume homogeneous production technologies across firms. Yet, little evidence suggests that this assumption is in line with reality. As we show in this paper, there is strong evidence for great heterogeneity in emissions generation across firms. This observation motivates our question: Can this aspect of heterogeneity be safely assumed away?

To answer this question, we first document enormous heterogeneity in both the quantity of emissions and emissions intensity of production across firms. We show that this heterogeneity is almost entirely driven by permanent across-firm variation. Next, we write down an Integrated Assessment Model (IAM) where firms vary heterogeneously in carbon dependence of production and solve for a carbon tax scheme which implements the constrained first-best allocation. We calibrate the model, nonparametrically matching the distribution of carbon emission intensity observed in the data. We then validate that the model is consistent with both the aggregate response to the price of carbon, comparing our model to the empirical estimates of Känzig (2023), and cross-sectional relationships between misallocation, productivity, and emissions intensity as measured by Kim (2023). Using the calibrated and validated model, we compute the laissez-faire and constrained efficient response to a green transition as in the literature (Golosov, Hassler, Krusell, and Tsyvinski (2014), Krusell and Smith (2022)). We then address the impact of model misspecification when setting carbon taxes by shutting down the relevant dimension of heterogeneity in our model. The implied Pigouvian tax path within our model is 85.7 percent higher without this dimension of heterogeneity than with, but we find that implementing these too-high carbon taxes in the model with-heterogeneity still results in substantial gains in welfare relative to the laissez-faire allocation. We finally show that these takeaways are globally robust to the key model parameter informing our results, the substitutability of green and brown capital. We conclude that the costs to society of ignoring climate change are substantially higher than the costs to society of ignoring this dimension of heterogeneity.

[1] "2023 was the world's warmest year on record, by far," `https://www.noaa.gov/news/2023-was-worlds-warmest-year-on-record-by-far`, published on January 12, 2024.

Our empirical contribution is to document the existence of vast heterogeneity of emissions among U.S. public companies. Using data on emissions from Bloomberg merged with Compustat, we find both substantial heterogeneity in terms of carbon emissions and in terms of emissions per unit of capital, which we define to be emissions intensity. The firm at the 90th percentile creates over 330 (20) times as many emissions (per unit of capital) than the firm at the 10th percentile. We then break this heterogeneity down into two components, permanent across-firm heterogeneity and temporary within-firm heterogeneity. We find that the variation across-firms accounts for the vast majority (95 percent) of the variation in emissions intensity. This fact persists even within-sector and within-industry. Hence, this evidence strongly suggests that some firms have more pollutive production technologies than others, which we later show is relevant for climate-related policy.

To evaluate the implications of this form of heterogeneity on carbon policy recommendations, we develop an IAM in which firms vary in the carbon dependence of their production process. Firms choose to invest in green and brown capital, and endogenously enter and exit. In response to carbon policy changes, such as a carbon tax, firms adjust both the greenness and quantity of their capital stock on the intensive margin but also exit on the extensive margin. As more carbon dependent firms face larger tax burdens, they exit more frequently and the distribution of production technologies endogenously becomes greener. This composition channel is absent from models with homogeneous production technologies across firms, and is the mechanism through which ignoring the heterogeneity of carbon dependence has the potential to impact policy recommendations. In order to quantitatively assess the magnitude of this effect on policy recommendations, we proceed in several steps. We first characterize the constrained first-best allocation where the social planner is constrained to the technologies available to the firms, utilizing methods developed by Lucas and Moll (2014), Moll and Nuno (2018), and Ottonello and Winberry (2023) and extend these methods to handle discrete choices, which in our case are endogenous firm entry and exit decisions. We then derive a sequence of carbon taxes which implements the constrained first-best allocation in a decentralized equilibrium.

Next, we ensure that the economic forces within our model match empirical estimates. We do so by calibrating our model to match both our estimated distribution of carbon emission intensity in the data and validate that the model matches the aggregate macroeconomic response to a carbon price shock as identified by Känzig (2023), as well as the joint relationship between misallocation, productivity, and emissions as measured by Kim (2023). With the calibrated and validated model in hand, we solve for both the laissez-faire equilibrium and the constrained efficient allocation in response to a green transition scenario as motivated by Golosov, Hassler, Krusell, and Tsyvinski (2014) and Krusell and Smith (2022). We illustrate the composition channel of economic adjustment within the scenario, as well as the intensive margin channels.

We then are able to consider the impact of the heterogeneity of production technologies on policy recommendations by considering a misspecified model which reduces the relevant dimension of heterogeneity to a singleton and repeat the entire analysis. Without the heterogeneous production technologies, the compositional channel is shut down and the marginal social cost of carbon rises. As a result, the derived optimal path of carbon taxes is 86 percent higher on average over the green transition. We ask what the welfare consequences of implementing the suboptimally large carbon taxes as implied by the misspecified model within the baseline model with heterogeneity. While it is trivial that the suboptimal path of taxes will not achieve the constrained first-best level of welfare, it still achieves 70 percent of the gain in welfare from moving from the laissez-faire allocation to the constrained first-best allocation. Finally, we demonstrate robustness of this pattern holds globally for the key parameter governing substitutability of brown and green capital. We conclude that ignoring heterogeneity of carbon dependence of firms can result in derived carbon taxes that are too large, but also that the welfare consequences of ignoring climate change altogether dwarf that of ignoring this aspect of heterogeneity when combating climate change.

Our paper highlights the heterogeneous response to carbon policy that firms will take, and quantifies the importance of taking this into account when considering setting carbon policy. Along the way, we make several contributions. Empirically, we document that the variation in emissions intensity among U.S. public firms is dominated by across-firm rather than within-firm variation, even when looking within-sector or within-industry. Theoretically, we contribute by solving for constrained efficient allocations and deriving tax schemes that decentralize the constrained efficient allocation in a decentralized equilibrium in a class of quantitatively relevant heterogeneous-firm models with discrete entry/exit decisions. We further contribute to the climate macro literature by considering the welfare impact of model misspecification in terms of policy response to climate change. This approach can be used to quantify the consequences of other forms of model misspecification in responding to climate change.

Related Literature. This paper contributes to the literature on the optimal policy response to the interaction between the macroeconomy and the Earth's climate. Nordhaus (1992a, 1992b) pioneered this literature by developing the Dynamic Integrated model of Climate and the Economy (DICE) and in subsequent revisions, including the most recent in Barrage and Nordhaus (2023). The related class of models, often referred to as "Integrated Assessment Models" or IAMs, notably Golosov, Hassler, Krusell, and Tsyvinski (2014) (hereafter GHKT), recast DICE using modern macroeconomic tools. Hassler, Krusell, and Smith

(2016) provide a basic introduction to IAMs, and, relative to our paper, these models largely use representative agents when determining the optimal policy response.

In recent literature, researchers have started to explore the implications of heterogeneity along several dimensions for optimal carbon policy. Much of this literature highlights both that additional policy instruments, in conjunction with a uniform Pigouvian carbon tax, are often needed to achieve the first-best allocation in the presence of heterogeneity and the qualitative or quantitative relationship between the uniform Pigouvian carbon tax and the second-best carbon tax without additional instruments.

The first strand of this literature focuses on geographic heterogeneity, starting with the Regional Integrated model of Climate and Economy (RICE) in Nordhaus 1996. Hassler and Krusell (2012) highlight heterogeneity in welfare implications of carbon taxes depending on regional fossil fuel endowments. Krusell and Smith (2022) find that the damages from climate change are unevenly distributed across space with some colder areas, like Canada and Russia, potentially benefiting from warmer temperatures. Imposing a uniform globally-optimal carbon tax without interregional transfers results in regions with welfare gains and other with losses. Furthermore, since policy is typically regionally implemented in practice, this literature considers the design of unilateral carbon policy and the influence of carbon leakage and endogenous changes in trade and migration (Kortum and Weisbach, 2021, Kortum, Wang, and Yao, 2022, Conte, Desmet, and Rossi-Hansberg, 2023).

A second strand of the literature focuses on the implications of household heterogeneity both over time and cross-sectionally. Departing from the perspective of the social planner who maximizes the welfare of infinitely lived households, Kolikoff, Kubler, Polbin, Sachs, and Scheidegger (2019) evaluate optimal carbon taxes within an OLG framework with selfish generations and highlight the importance of intergenerational transfers in conjunction carbon taxes for all generations to be better off. Additionally, since poor households consume relatively more emission-intensive consumption bundles than rich households, this strand of the literature also argues that the second-best uniform carbon tax without lump-sum transfers is lower than the Pigouvian uniform carbon tax due to regressivity (Belfiori, Carroll, and Hur, 2024, Bourany, 2024).

A third strand of the literature focuses on the implications of firm heterogeneity. Empirically, Lyubich, Shapiro, and Walker (2018) document substantial heterogeneity in within-narrow-industry energy and emissions intensity at the plant level in the entire U.S. manufacturing sector. In concurrent work to ours, Kim (2023) also documents substantial heterogeneity in the emissions intensity of U.S. public companies as well as a negative relationship between emissions intensity and productivity within industries.[2] In a model where hetero-

[2] Kim (2023) defines emission intensity as emissions per unit of sales, while we use emissions over capital.

geneity in emission intensity across firms is primarily driven by misallocation from financial frictions, Kim (2023) finds that optimal carbon taxes given misallocation are higher than optimal carbon taxes without such misallocation because carbon taxes secondarily improve allocative efficiency in the economy. Importantly, the composition effect that is the highlight of our paper is absent from Kim (2023) because, in his model, production technologies do not differ between firms.

We make both empirical and theoretical contributions relative to the existing literature. Empirically, we document the substantial and persistent heterogeneity in the emission intensity of U.S. public companies. While the extent of this heterogeneity for U.S. public companies is concurrently documented by Kim (2023) and for U.S. manufacturing plants by Lyubich, Shapiro, and Walker (2018) in the cross-section, we are the first paper, to our knowledge, to document that this dispersion is dominated by permanent, across-firm variation. This persistence is key to our modeling approach where firms differ by permanent carbon dependence of production.

Theoretically, we extend a standard IAM to a setting with heterogeneous firms a la Hopenhayn and Rogerson (1993), Gomes (2001), Khan and Thomas (2013), and Clementi and Palazzo (2015). We solve for the constrained socially optimal allocation by applying and extending techniques developed by Lucas and Moll (2014), Moll and Nuno (2018), and Ottonello and Winberry (2023). We illustrate the composition effect along the quantified transition. To quantify the impact of misspecifying this form of heterogeneity, we collapse the firm-level heterogeneity in dependence on brown capital and resolve for the optimal path of carbon taxes, which are on average 86 percent higher than the baseline. Finally, in the spirit of the quantitative analysis by Barrage (2020) on fiscal policy distortions, we evaluate the cost of this suboptimal climate policy and find that these too-high carbon taxes recover 70 percent of welfare benefit of the optimal carbon taxes.

The remainder of our paper is organized as follows. We document persistent heterogeneity of carbon emission intensity in Section 2. In Section 3, we introduce our GE heterogeneous-firm model. We characterize the socially optimal allocation and decentralize the allocation in competitive equilibrium in Section 4. We quantify the socially optimal green transition in Section 5. Finally, we quantify the implications of ignoring firm-level heterogeneity in Section 6, evaluate robustness of this quantitative result in Section 7, and conclude in Section 8.

As we show in Section 5.2.1, the choice of denominator is not innocuous and our model is able to recover the main empirical results of Kim (2023) without requiring factor misallocation due to financial frictions.

2 The Carbon Emissions of U.S. Public Firms

We begin with an empirical analysis of the distribution of carbon emissions across U.S. public firms. In this section, we document two stylized facts. First, we document substantial cross-sectional heterogeneity in the carbon emission distribution. Second, we show that the vast majority of the variation in the carbon emissions in the data originates from across-firm variation, rather than within-firm. These facts together imply that firms are persistently heterogeneous in their emissions intensity to a degree that has the potential to influence the relative social costs from carbon emissions that a social planner would internalize. This social cost enters the planner's decision-making when deciding whether to exit firms, and hence the planner is more likely to exit a highly pollutive firm than a green firm, all else equal. This extensive margin effect endogenously induces a change in the composition of the production technology in use in the economy.

To this aim, we construct a balanced panel of firm-level data over a five-year period from 2016 to 2020 using two primary data sources.[3] Firstly, we utilize the Scope 1 and Scope 2 greenhouse gas emissions data obtained from firms' annual reports, which are collected by Bloomberg (referred to as GHG_SCOPE_1 and GHG_SCOPE_2). Scope 1 emissions are thousand metric tonnes of carbon-equivalent emissions (COe) generated by sources under the control or ownership of the firm. This category includes carbon emissions resulting from the combustion of fossil fuels in boilers, furnaces, or vehicles, for instance. Scope 2 emissions are indirectly caused by the firm's activities, such as the emissions associated with the firm's purchase and use of energy.[4] In our variable of interest, we use Scope 1, Scope 2, and the sum of Scope 1 and Scope 2 emissions. This sample of firms covers between 30 and 32 percent of the annual total U.S. greenhouse gas emissions. Compustat provides us with standard variables including gross property, plants, and equipment (PPEGT). We adjust nominal variables for inflation using a GDP deflator to express these in 2015 dollars. For industry classification, we also rely on Compustat for Standard Industrial Classification (SIC) division codes and 2-digit SIC major groups codes.[5]

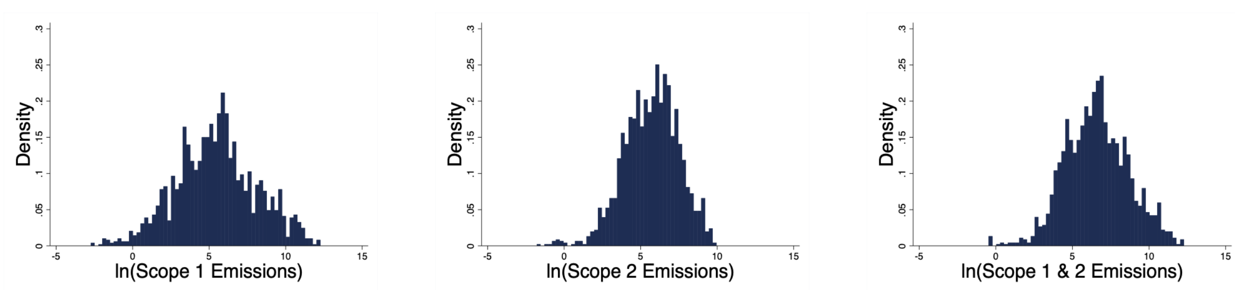
[3]For robustness, we report our main results using a 10-year balanced panel containing fewer firms in Appendix A. The results are all qualitatively and quantitatively similar in the longer sample.

[4]Greenhouse gas emissions are also available at Scope 3, which are emissions that are produced upstream or downstream within the firm's value chain. To avoid double-counting emissions up and down stream, we do not use Scope 3 emissions.

[5]As is standard practice in corporate finance, we filter out financial firms (SIC Division H), utilities (2-Digit SIC Major Group 49), and public administration and nonclassifiable firms (SIC Division J). Filtering out utilities is particularly important for our analysis as we include Scope 2 emissions in our analysis. Emissions generated by a power plant in the process of generating power utilized by another company would count as Scope 1 emissions for the power plant and Scope 2 emissions for the other company; dropping utilities prevents this double-counting. To ensure this does not impact our main empirical results, in Appendix A we repeat our analysis including utilities companies, only including Scope 1 emissions so as to continue avoiding

After imposing the selection criteria, the balanced panel with Scope 1 has 389 firms, the balanced panel with Scope 2 emissions has 365 firms, and the balanced panel with Scope 1 plus Scope 2 has 362 firms. As shown in Appendix A, the firms with available emissions data are on average four-to-five times larger than the average firm in the full Compustat sample in terms of number of employees, earnings, and capital. These differences may arise due to more scrutiny from ESG-focused investors demanding that larger firms release information about their emissions. These firms are similarly more capital intensive with a lower employees to capital ratio. In terms of sectoral representation, firms providing services are underrepresented compared to full Compustat sample and mining, manufacturing, and transportation firms are overrepresented. The profitability of firms with emissions data and the full Compustat sample are not statistically different. We plot histograms of the log emissions in Figure 1.

Figure 1: Distribution of Carbon Emissions



(a) Scope 1 Emissions (b) Scope 2 Emissions (c) Scope 1 & 2 Emissions

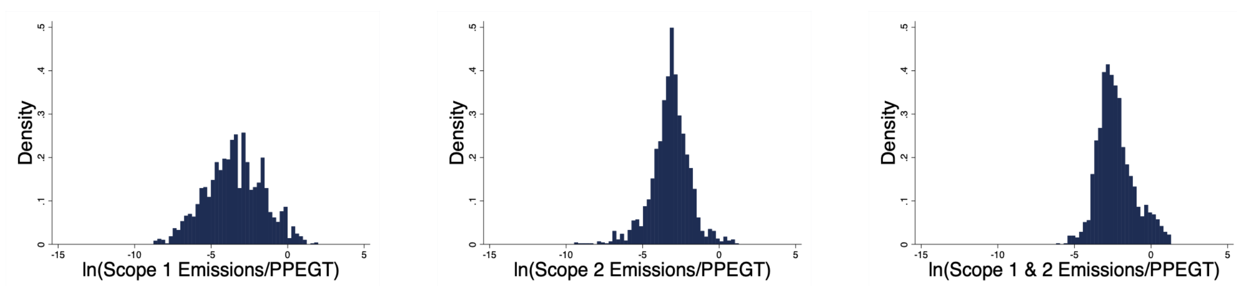
Notes: This figure shows the unconditional pooled distribution of the carbon emissions, in units of metric tonnes of CO2e, for Compustat firms in a five-year balanced panel between 2016-2020. Panel (a) shows Scope 1 emissions, which are those created by the operations directly controlled by the firm including subsidiaries and affiliates. Panel (b) shows Scope 2 emissions, which are those resulting indirectly from the firm's activity in particular energy generation. Panel (c) shows Scope 1 emissions plus Scope 2 emissions.

As is evidenced by Figure 1, there is a wide dispersion in the level of emissions produced by U.S. public firms. To further illustrate this point, we calculate that the 90th percentile of the emissions distribution is 2340×, 248×, and 333× larger than the 10th percentile for Scope 1, Scope 2, and Scope 1 + Scope 2, respectively. This is perhaps not surprising for a number of reasons. One such reason is that the U.S. firm distribution is well understood to have dispersion in scale. For example, in our sample the 90th percentile of PPEGT is 2285× larger than the 10th percentile. While the total level of emissions produced by

double counting. The results remain quantitatively and qualitatively identical to our baseline results. We thank Lubos Pastor for this suggestion.

society ultimately drives global climate change, we argue that whether one firm should be considered 'greener' in terms of their production technology than another firm should be based not on the total level of emissions that the firms produce, but rather on the amount of emissions that the firms produce *relative to their scale of production*.[6] Hence, we define *emissions intensity* as the (log of) emissions per unit capital, $y = \log\left(\frac{\text{GHG}}{\text{PPEGT}}\right)$, where $\text{GHG} \in \{\text{GHG\_SCOPE\_1}, \text{GHG\_SCOPE\_2}, \text{GHG\_SCOPE\_1} + \text{GHG\_SCOPE\_2}\}$. Moving forwards, we will consider emissions intensity as our measure of the 'greenness' of a firm's production technology, but the rest of our results are effectively unchanged if the (log of) level of emissions were used in place of emissions intensity. Similarly, our results are also robust to different measures of firm scale - we choose PPEGT as our baseline measure of scale as the quantity $\frac{\text{emissions}}{\text{capital}}$ directly maps in to our model when bringing our model to the data as in Section 6. We illustrate the distribution of emissions intensity in Figure 2.

Figure 2: Distribution of Carbon Emission Intensity



(a) Scope 1 Emissions (b) Scope 2 Emissions (c) Scope 1 & 2 Emissions

Notes: This figure shows the unconditional pooled distribution of carbon emission intensity—defined as metric tonnes of CO2e per thousands of 2015 dollar of property, plants, and equipment for Compustat firms in a five-year balanced panel between 2016-2020. Panel (a) shows Scope 1 emissions, which are those created by the operations directly controlled by the firm including subsidiaries and affiliates. Panel (b) shows Scope 2 emissions, which are those resulting indirectly from the firm's activity in particular energy generation. Panel (c) shows Scope 1 emissions plus Scope 2 emissions.

As is clear from Figure 2, there is a noticeably large dispersion in carbon emissions in-

[6] For example, consider three firms named A, B, and C. These firms operate identical production technologies except firm C's production technology generates twice as much carbon emissions than that of firms A and B. Clearly, firm C operates a production technology that is less green than the production technologies operated by firms A and B. Now suppose firms A and B merge to form firm AB. Firm AB generates the same level of emissions as firm C, but firm AB generates twice as much production, and hence twice as much societal benefit from production, as firm C. If we rank how 'green' a production technology is by the level of emissions, we would correctly label the production technologies used by firms A and B as equally green, and firm C as less green. However, we would rank the production technologies of firm AB as equally green as firm C, despite the fact that the societal benefit from producing those emissions is different between the two firms. If we instead adjust for scale in our ranking, the ranking of societal impact relative to the amount of emissions produced is preserved.

tensity among U.S. public firms. For emission intensity, the 90-10 ratios are 132×, 19×, 22× for Scope 1, Scope 2, and Scope 1 + Scope 2, respectively. While scale is a key factor in the emissions created by firms, there is clearly substantial variation in emission intensity itself. However, the degree to which the variation in emissions intensity variation is permanent across-firm variation vs temporary within-firm variation is obscured by these histograms and simple statistics. To separate permanent across-firm variation from temporary within-firm variation, we regress emissions intensity on firm-level fixed effects α_i, as in equation (1):

$$y_{i,t} = \alpha_i + \eta_{i,t}, \tag{1}$$

where the firm fixed effects α_i reflects variation across-firms and the residual term $\eta_{i,t}$ varies within-firm over time. Further assuming that the fixed effect α_i and residual $\eta_{i,t}$ are independent, it is immediate that:

$$\underbrace{\mathbb{V}[y_{i,t}]}_{\text{total}} = \underbrace{\mathbb{V}[\alpha_i]}_{\text{across-firm}} + \underbrace{\mathbb{V}[\eta_{i,t}]}_{\text{within-firm}},$$

and hence the total variation in emissions intensity $\mathbb{V}[y_{i,t}]$ is the sum of across- and within-firm variation $\mathbb{V}[\alpha_i]$ and $\mathbb{V}[\eta_{i,t}]$, respectively.

Further, it is simple to show that the (unadjusted) total R^2 from regression 1 is equal to the ratio of estimated across-firm variation to estimated total variation $\hat{\mathbb{V}}[\alpha_i]/\hat{\mathbb{V}}[y_{i,t}]$. In words, the total R^2 is the estimated fraction of total variation originating from across-firm variation, where the rest of the variation originates from within-firm across-time variation. We report the total R^2 statistics from Regression (1) in Table 1.

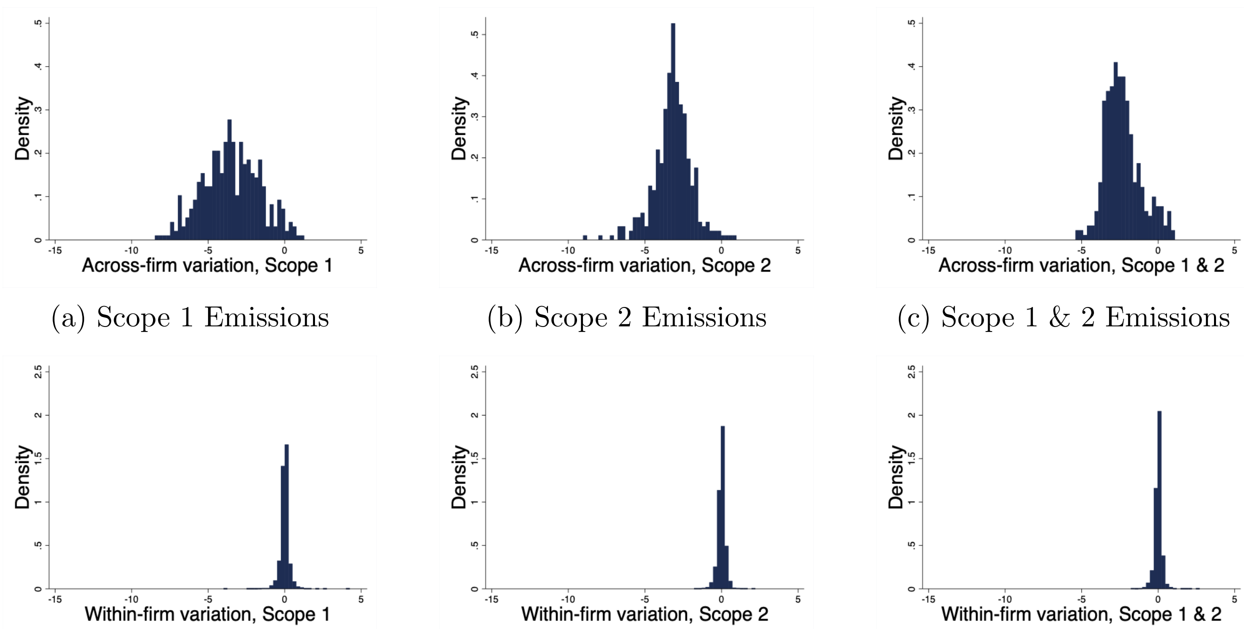
Table 1: Total R^2 Values from Regression 1

Emission Scope	R^2	95% CI
Scope 1	0.970	(0.958,0.980)
Scope 2	0.954	(0.939,0.966)
Scope 1 & 2	0.956	(0.943,0.967)

Notes: This table reports the total R^2 values from OLS estimates of Regression 1, along with 95% confidence intervals. The confidence intervals are bias corrected and accelerated bootstrap intervals as in DiCiccio and Efron (1996).

Table 1 shows that the overwhelming majority of the variation in emissions intensity data is explained by across-firm variation, and within-firm variation is comparatively tiny (between 3-4.5 percent depending on the emission measure). This is robust across the Scope of emissions, given by the rows of Table 1. We further illustrate this empirical fact via histograms of estimated firm-level emission intensity $\hat{\alpha}_i$ and residuals $\hat{\eta}_{i,t}$ in Figure 3.

Figure 3: Across-Firm vs Within-Firm Variation



(a) Scope 1 Emissions (b) Scope 2 Emissions (c) Scope 1 & 2 Emissions

(d) Scope 1 Emissions (e) Scope 2 Emissions (f) Scope 1 & 2 Emissions

Notes: This figure visualizes the distribution of firm-level emission intensity $\hat{\alpha}_i$ OLS estimates (top row) and residuals (bottom row) from Equation 1. The distribution of $\hat{\alpha}_i$ is visually very similar to the distribution of the raw data as reported in Figure 2. By comparison, the distribution of residuals is much tighter, concentrated closely around zero.

Figure 3 visualizes the distributions of the estimated firm-level fixed effects $\hat{\alpha}^c + \hat{\alpha}_{i,t}$ and within-firm variation $\hat{\eta}_{i,t}$ from regression (1) across emission scopes. The distributions are starkly different from one another, with the distribution of within-firm variation tightly distributed around zero whereas the firm fixed effects are widely varied. The interpretation is clear: firms vary substantially from one-another in their emissions intensity, but vary far less across years within a given firm. While the results presented thus far show that across-firm variation is a strong feature of the data, they do little to explain what drives this permanent heterogeneity across firms. One strong candidate for a driver of such heterogeneity are industry and sectoral differences. In Appendix A we perform a decomposition of the emissions intensity data across- and within-industries. We do find correlation of emissions intensity within industries as one may expect, but somewhat more surprising is the fact that across-firm variation is almost as dominant a feature of the within-sector and within-industry emissions intensity data as it is the raw emissions intensity data. We show this using a 2-stage approach. First, we regress off sector (or industry) fixed effects from the emissions intensity data, and store the within-sector (industry) residuals $r_{i,t}$. Then, in a second stage we regress the within-sector (industry) residuals on firm-level fixed effects, given by equation (2).

$$r_{i,t} = \beta_i + \epsilon_{i,t} \tag{2}$$

If we again assume β_i and $\epsilon_{i,t}$ are independent, we can again decompose the within-sector (industry) residual emission intensity into across-firm permanent variation and within-firm temporary variation. And just as before, the (unadjusted) total R^2 from regression (2) has the interpretation of the fraction of within-sector (industry) residual emission intensity that is driven by permanent across-firm variation. We report these R^2 estimates in Table 2.

Table 2: Explanatory Power of Firm Fixed Effects Within Sector and Industry

Within-Sector			Within-Industry		
Scope	R^2	95% CI	Scope	R^2	95% CI
Scope 1	0.964	(0.948,0.976)	Scope 1	0.911	(0.879,0.943)
Scope 2	0.947	(0.930,0.961)	Scope 2	0.926	(0.900,0.946)
Scope 1 & 2	0.953	(0.939,0.966)	Scope 1 & 2	0.905	(0.882,0.931)

Notes: This table reports R^2 from estimating Equation 2 where the residual from estimating Equation 10 is regressed on firm-level fixed effects. In Equation 10, firm-level emission intensity is regressed on sector (or industry fixed effects) for each sector (or industry) separately. Sectors are defined as SIC Divisions and industries are defined as 2-Digit SIC Major Groups. The confidence intervals are bias corrected and accelerated bootstrap intervals as in DiCiccio and Efron (1996).

Table 2 shows that persistent across-firm variation dominates the within-industry emissions intensity variation, just as Table 1 shows that persistent across-firm variation dominates the raw data itself. So, while emissions intensity are correlated within sectors and industries, substantial within-sector (industry) variation exist which overwhelmingly driven by across-firm variation.

In Appendix A, we demonstrate that all of our results are robust to defining $y_{i,t}$ as log(GHG) and $\log\left(\frac{\text{GHG}}{\text{Output}}\right)$, the definition of emission intensity used by Kim (2023). None of our results change much either qualitatively or quantitatively. We also show that our results are again both qualitatively and quantitatively similar if we report all R^2 values as adjusted

R^2, which lack the clear interpretation of unadjusted R^2 in this context but penalizes based on the number of regressors included in the regressions. Further, we show that our results are not purely driven by the short panel length by repeating our baseline exercises with a 10-year panel of 115 firms. Again, our results remain qualitatively and quantitatively consistent. We then show that our results are robust to including utilities firms in our sample, in this case focusing on Scope 1 only to avoid double-counting. Finally, we consider adding additional controls to regression (1). We find very little added explanatory power from adding such controls, indicating that the variation in emissions intensity that we observe in the data is overwhelmingly (1) across-firm variation, and (2) a very small amount of unexplained within-firm variation which we interpret as measurement error.

Moving forwards, we will focus on the across-firm persistent heterogeneity in emissions intensity in our theoretical work as opposed to industrial variation. We do so to remain as close as possible with the literature that we compare ourselves to such as Nordhaus (2007), Golosov, Hassler, Krusell, and Tsyvinski (2014), and Krusell and Smith (2022), all of which ignores substitution across industries of final goods produced across the green transition. We thus will assume the goods produced by different industries are perfectly substitutable with one another, but allow the carbon emissions intensities across the firm distribution to vary across firms but not within-firm across-time.

3 Heterogeneous-Firm Integrated Assessment Model

In our structural model, time is discrete and infinite-horizon. The model consists of four main interconnected blocks: a detailed firm block, the focus of the heterogeneous-firm model; a climate block that broadly follows the climate-macro literature; a simple representative consumer; and a government. The model is in general equilibrium, with the wage and risk-free interest rate adjusting and labor and risk-free asset markets clearing. Agents in the model are atomistic and do not internalize their contributions to climate damages, the key externality which provides room for policy to improve on the laissez-faire allocation.

3.1 Model Economy

Firms vary heterogeneously across four dimensions. Firm i operating in date t has exogenously evolving TFP $z_{i,t}$, permanent carbon dependence of production a_i, and endogenously-evolving brown capital $k_{b,i,t}$ and green capital $k_{g,i,t}$. Brown capital is pollutive and creates carbon emissions $\xi_{i,t} = \gamma k_{b,i,t}$ where $\gamma \in \mathbb{R}^+$ is a constant of proportionality.[7] The vector

[7]In Appendix C.1, we relax the assumption that carbon emissions are deterministic. The planner's solution is unchanged and remains decentralizable through Pigouvian taxes on emissions in a competitive

$s_{i,t} = [z_{i,t}, a_i, k_{b,i,t}, k_{g,i,t}]'$ summarizes firm i's idiosyncratic state. Firms produce with a constant elasticity of substitution (CES) aggregate of green and brown capital with elasticity parameter ρ and where its carbon dependence a_i is the weight on brown capital. Given a market-clearing wage w_t, firm i employs labor $L_{i,t}$ and produces decreasing return-to-scale Cobb-Douglas production of labor and effective capital with coefficients ν and α, respectively. The stock of carbon emissions S_t in the atmosphere destroys fraction $D(S_t)$ of output of all firms; the climate damage and the carbon cycle are discussed further in Subsection 3.2. Firms pay a proportional tax τ_t on brown capital, which in the laissez-faire allocation is zero for all periods $\tau_t = 0\ \forall t$; we derive the socially optimal path of brown capital taxes τ_t in Section 4. A firm with state vector s generates cash flow $\pi_t(s)$, which is production net of climate damage, wages paid to workers, and an unavoidable fixed cost of c_f units of labor required to operate,

$$\pi_t(s) = \max_L [1 - D(S_t)] \exp(z) A_t^{1-\alpha} [a^{1-\rho} k_b^\rho + (1-a)^{1-\rho} k_g^\rho]^{\frac{\alpha}{\rho}} L^\nu - w_t L - w_t c_f - \tau_t k_b,$$

where $A_{t+1} = (1+\iota)A_t$ is the economy-wide component of TFP that grows exogenously at rate ι. As $\pi_t(s)$ can become negative, firms endogenously exit in equilibrium.

After production, firms both exogenously and endogenously exit. With independent probability $\lambda \in [0,1]$, firm i is forced to exit and, with complementary probability $1-\lambda$, firm i chooses whether to endogenously exit. A nonexiting firm chooses investments in brown and green capital, $x_{b,i,t}$ and $x_{g,i,t}$, subject to adjustment costs $\psi(x_{j,i,t}, k_{j,i,t})$ for $j \in \{b, g\}$ and depreciation δ. An exiting firm eats its nondepreciated capital net of the adjustment costs $\psi^X(k) \equiv \psi[-(1-\delta)k, k]$ to drive both its capital stocks to zero. Given the value from exiting $V^X(k_b, k_g)$ and continuing (nonexiting) $V^C(s)$, the ex-ante value function of the firm with state vector s entering period t is

$$V_t(s) = \pi_t(s) + \lambda V^X(k_b, k_g) + (1-\lambda) \max\{V^X(k_b, k_g), V_t^C(s)\}, \tag{3}$$

where the exiting value is

$$V^X(k_b, k_g) = (1-\delta)(k_b + k_g) - \psi^X(k_b) - \psi^X(k_g),$$

and the continuing value is

$$V_t^C(s) = \max_{x_b, x_g} -x_b - x_g - \psi(x_b, k_b) - \psi(x_g, k_g) + \frac{1}{R_t} \mathbb{E}_{z'}[V_{t+1}(s')], \tag{4}$$

equilibrium.

$$\text{s.t. } k'_j = (1-\delta)k_j + x_j, \quad \text{for } j \in \{b, g\}.$$

Let $k_{b,t}(s), k_{g,t}(s), X_t(s)$ as the investment and exit policy functions for an operating firm with state vector s at period t where $X_t(s) = 1$ indicates that the firm exits.

A large mass of potential firms can enter competitively by paying an fixed cost κ denominated in units of labor. After an entrant pays $w_t\kappa$, they independently draw a permanent carbon dependence $a_i \sim Q_a$, and a signal of their initial TFP $q_i \sim Q_q$. Then the entrant chooses whether to exit before setting up their firm or invest $k_{b,i,t}$ in brown capital and $k_{g,i,t}$ in green capital. The entrant then draws initial shock to TFP $\epsilon_{q,i,t} \sim N(0, \sigma_q^2)$ and becomes an operating firm with value as defined in Equation (3). Mathematically, the entrant problem boils down to a free entry condition:

$$w_t\kappa \geq \mathbb{E}_{a,q}[V_t^E(a,q)],$$

where an entrant's value $V_t^E(a,q)$ after observing carbon dependence a and signal q is

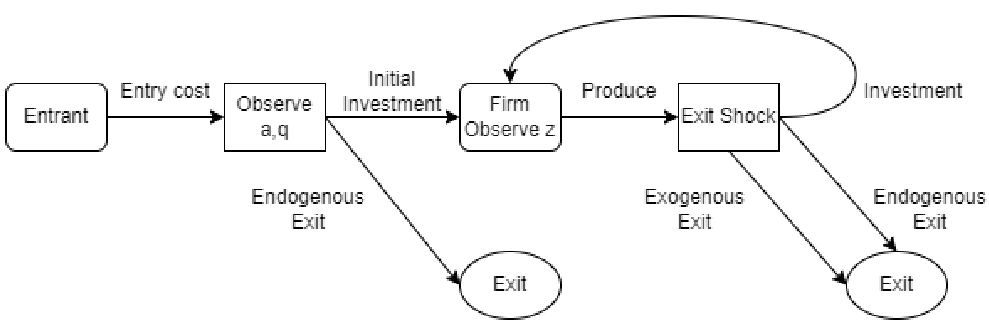
$$V_t^E(a,q) = \max\{0, \max_{k_b,k_g} -k_b - k_g + \mathbb{E}_{z|q}[V_t(s)]\},$$

and where the initial TFP z draw is

$$z = \rho_q q + \epsilon_q.$$

Let $k_{b,t}^E(a,q), k_{g,t}^E(a,q), X_t^E(a,q)$ be the investment and exit policies for an entrant firm that draws a and q in period t and B_t as the endogenous measure of potential firms chooses to enter in period t. Figure 4 summarizes the timing associated with firms' decisions.

Figure 4: Timing of Firm Problem



Notes: This figure describes the timing of a model period for the IAM described in Section 3.

At the beginning of period t, the distribution μ_t of incumbent firms are inherited from the previous period $t-1$ where $\mu_t(s)$ is the measure of firms in state s. After potential firms enter, the endogenous distribution $\Phi_t(\mu_t)$ of firms operate. Given exit shocks, endogenous continuation and investment decisions by operating firms, and the law of motion of idiosyncratic productivity, the distribution $\mu_{t+1} = T_t^*(\Phi_t(\mu_t))$ of firms that were operating in period t will continue into the next period.[8] We then define the total mass of brown capital utilized in production as $K_t^b = \int k_b d\Phi(\mu_t)(s)$.

A unit mass of representative consumers inelastically supplies a unit of labor $L^S = 1$ and invests in a risk-free asset b_{t+1} in zero net supply with gross interest rate R_t. Households also own shares in firms where a share represents a divisible claim on the dividends of the firm and the number of shares of each firm is normalized to one. Households buy shares $\Theta_{i,t+1}$ in all operating or potential firms i at price $p_{i,t}$ that pay dividend $d_{i,t}$.[9] Finally, the household pays a lump sum tax (or, if negative, receives a lump sum transfer) T_t.[10] The household optimization problem at t is

$$\max_{\{C_m, b_m, \{\Theta_{i,m}\}\forall i\}_{m=t}^{\infty}} \sum_{m=t}^{\infty} \beta^m U(C_m)$$
$$\text{s.t. } C_m + \frac{1}{R_m} b_{m+1} + \int p_{i,m}\Theta_{i,m+1} di + T_m = w_m L^S + b_m + \int (p_{i,m} + d_{i,m})\Theta_{i,m} di \qquad \forall m \geq t.$$

A standard Euler equation prices the risk-free bond:

$$U'(C_t) = \beta R_t U'(C_{t+1}) \implies \frac{1}{R_t} = \beta \frac{U'(C_{t+1})}{U'(C_t)}.$$

[8]Given firm policy functions and the mass of entrants, we can write these measure operators as

$$\Phi_t(\mu)(s) = \mu(s) + B \int Q_q(q) Q_a(a)[1 - X_t^E(a,q)] \mathbb{1}_{k_b = k_{b,t}^E(a,q)} \mathbb{1}_{k_g = k_{g,t}^E(a,q)} p_q(z - \rho_q \log(q)) dq,$$
$$T_t^*(\Phi)(s') = \int (1-\lambda)[1 - X_t(s)] \mathbb{1}_{k_b' = (1-\delta)k_b + x_{b,t}(s)} \mathbb{1}_{k_g' = (1-\delta)k_g + x_{g,t}(s)} \mathbb{1}_{a'=a} Q_z(z'|z) d\Phi(s)$$

where $p_q(\epsilon_q)$ is the density of the distribution of initial productivity shocks ϵ_q.

[9]The dividend of firm i depends on the firm's status as operating or entrant and its exit policies

$$d_{i,t} = \begin{cases} \pi_t(s_{i,t}) - x_b(s_{i,t}) - x_g(s_{i,t}) - \psi[x_b(s_{i,t}), k_{b,i,t}] & \\ \qquad -\psi[x_g(s_{i,t}), k_{g,i,t}], & \text{if } i \text{ is operating and } X_t(s_{i,t}) = 0, \\ \pi_t(s_{i,t}) + V^X(k_{b,i,t}, k_{g,i,t}), & \text{if } i \text{ is operating and } X_t(s_{i,t}) = 1, \\ -w_t\kappa - k_b^E(a_i) - k_g^E(a_i) + \pi_t(s_{i,t}) - x_b(s_{i,t}) - x_g(s_{i,t}) & \\ \qquad -\psi[x_b(s_{i,t}), k_b^E(a_i)] - \psi[x_g(s_{i,t}), k_g^E(a_i)], & \text{if } i \text{ is entrant and } X_t^E(a_i) = X_t(s_{i,t}) = 0, \\ -w_t\kappa - k_b^E(a_i) - k_g^E(a_i) + \pi_t(s_{i,t}) + V^X(k_b^E(a_i), k_g^E(a_i)), & \text{if } i \text{ is entrant, } X_t^E(a_i) = 0 \text{, and } X_t(s_{i,t}) = 1, \\ -w_t\kappa, & \text{if } i \text{ is entrant and } X_t^E(a_i) = 1. \end{cases}$$

[10]This lump sum tax/transfer clears the government budget constraint given any revenue generated by other implemented policies. In the case of taxes on brown capital, the government budget clearing condition implies that $T_t = -\tau_t K_t^b$.

The gross interest rate R_t is also used by firms to discount dividend streams. Due to climate damage dynamics as well as technological growth, aggregate consumption is non-stationary so households value consumption differently at the margin in different states of the world. Thus, as firms discount using R_t then they inherit the household's stochastic discount factor to value dividend streams.

The first-order condition with respect to shares in firm i implies

$$U'(C_t)p_{i,t} = \beta U'(C_{t+1})(p_{i,t+1} + d_{i,t+1})$$

Substituting in the risk-free rate, iterating forward, and assuming appropriate transversality condition implies that the stock price of firm i is the present value of its future dividends:

$$p_{i,t} = \sum_{m=t}^{\infty} \left[\prod_{n=t}^{m} \frac{1}{R_n} \right] d_{i,m+1}.$$

Note that, if a firm exits, it is clear from this expression that its stock price falls to zero.

Finally, clearing in the labor market, the risk-free asset market, and the stock market for all operating and potential firms i at time t implies

$$\begin{aligned} \int L^d(s) d\Phi_t(\mu_t)(s) &= 1, \qquad \forall t, \\ b_t &= 0, \qquad \forall t, \\ \Theta_{t,i} &= 1 \qquad \forall i, t. \end{aligned}$$

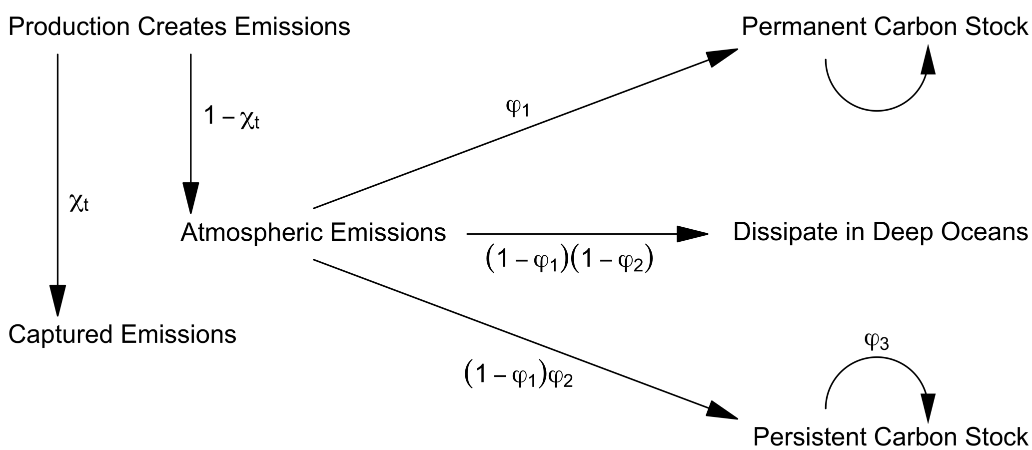
3.2 Carbon Cycle

We broadly follow Nordhaus (2007), Golosov, Hassler, Krusell, and Tsyvinski (2014), and Krusell and Smith (2022) in defining and calibrating the carbon cycle in the model economy. Figure 5 graphically represents the carbon cycle within our economy. We track two carbon stocks as state variables for the economy: permanent emissions S_t^1 and persistent emissions S_t^2. The total emissions in the atmosphere are the sum of these stocks, $S_t = S_t^1 + S_t^2$. A fraction φ_1 of the flow of carbon emissions remain in the atmosphere permanently and do not ever dissipate adding to S_{t-1}^1, while a fraction $(1 - \varphi_1)\varphi_2$ of emissions dissipate with decay rate φ_3 adding to S_{t-1}^2. Finally, the remaining fraction $(1 - \varphi_1)(1 - \varphi_2)$ of emissions are immediately captured by environmental reservoirs like the deep oceans and do not create damage. We further follow Krusell and Smith (2022) in modeling expected improvements in emissions reduction technology through an exogenously evolving carbon-absorption parameter χ_t. The explicit law-of-motion for the carbon stocks are given by the following:

$$S_t^1 = S_{t-1}^1 + (1-\chi_t)\varphi_1\gamma K_t^b$$
$$S_t^2 = \varphi_3 S_{t-1}^2 + (1-\chi_t)(1-\varphi_1)\varphi_2\gamma K_t^b$$

The total carbon stock $S_t = S_t^1 + S_t^2$ is then assumed to have a damaging effect on the environment. As mentioned in Section 3, the carbon stock causes damages to output, $D(S_t)Y_t$ where $D(S_t)$ is an increasing function of the carbon stock.

Figure 5: The Carbon Cycle



Notes: This figure details the carbon cycle as described in Subsection 3.2. Atmospheric emissions enter two carbon stocks, one permanent and the other persistent but not permanent. Some atmospheric emissions dissipate immediately and do not enter either stock.

An equilibrium definition in this economy simply includes the climate block in an otherwise standard equilibrium definition.

Definition 1. *A* ***Carbon-Cycle Competitive Equilibrium*** *is a set of allocations* $\{k_{b,m}^E(a,q), k_{g,m}^E(a,q), x_{b,m}(s), x_{g,m}(s), b_m, \Theta_{i,m}, L_m^d(s), S_m^1, S_m^2, \mu_m\}_{m=t}^\infty$*, prices* $\{w_m, R_m, p_{i,m}\}_{m=t}^\infty$*, continuation rules* $\{X_m^E(a,q), X_m(s)\}_{m=t}^\infty$*, mass of entrants* $\{B_m\}_{m=t}^\infty$*, and government policy* $\{\tau_m, T_m\}_{m=t}^\infty$ *such that firm decisions solve their problems, household decisions solve their problem, the free entry condition holds, the law of motion of environmental carbon stocks hold, the labor, risk-free asset, and stock markets clear in each period, and the government budget constraint is satisfied.*

4 The Constrained First-Best Allocation

We refer to the economy with brown capital taxes set to zero $\tau_t = 0\ \forall t$ as the "Business-As-Usual" (BAU) laissez-faire economy. A natural point of comparison is the economy run optimally by a constrained social slanner. The planner is constrained to the same technologies as the agents in the decentralized economy including the firm-level adjustment cost functions, the exogenous supply of labor to the economy, the law-of-motion for the environment as described by Subsection 3.2, and resource feasibility constraints. The key difference between the planner's allocation and the BAU competitive equilibrium is that the planner internalizes the climate externality. The planner's objective is to maximize the consumer's discounted utility stream subject to these constraints. The constrained planner's problem can be defined using the following recursive formulation:

$$\mathcal{W}_t(\mu, S^1, S^2) = \max_{x_b(\cdot), x_g(\cdot), X(\cdot), x_b^E(\cdot), x_g^E(\cdot), X^E(\cdot), L^d(\cdot), B} U(C_t) + \beta \mathcal{W}_{t+1}(\mu', S^{1\prime}, S^{2\prime}) \tag{5}$$

subject to

$$\begin{aligned}
C_t &= (1 - D(S^{1\prime} + S^{2\prime}))Y_t - I_t - \Psi_t \\
1 &= \int (L^d(s) + c_f)\Phi(\mu, x_b^E(\cdot), x_g^E(\cdot), X^E(\cdot), B)(s)ds + B\kappa \\
K^b &= \int k_b \Phi(\mu, x_b^E(\cdot), x_g^E(\cdot), X^E(\cdot), B)(s)ds \\
\mu' &= T^*(\mu, x_b(\cdot), x_g(\cdot), X(\cdot), x_b^E(\cdot), x_g^E(\cdot), X^E(\cdot), B) \\
S^{1\prime} &= S^1 + (1 - \chi_t)\varphi_1 \gamma K^b \\
S^{2\prime} &= \varphi_3 S^2 + (1 - \chi_t)(1 - \varphi_1)\varphi_2 \gamma K^b \\
B &\geq 0
\end{aligned}$$

where

$$\begin{aligned}
Y_t &= \int \exp(z) A_t^{1-\alpha}[a^{1-\rho}k_b^\rho + (1-a)^{1-\rho}k_g^\rho]^{\frac{\alpha}{\rho}}(L^d(s))^\mu \Phi(\mu, x_b^E(\cdot), x_g^E(\cdot), X^E(\cdot), B)(s)ds \\
I_t &= \int (x_b(s) + x_g(s))(1-\lambda)\Phi(\mu, x_b^E(\cdot), x_g^E(\cdot), X^E(\cdot), B)(s)ds \\
&+ \int (-(1-\delta)k_b - (1-\delta)k_g)\lambda\Phi(\mu, x_b^E(\cdot), x_g^E(\cdot), X^E(\cdot), B)(s)ds \\
&+ \int (x_b^E(a) + x_g^E(a))\mathbb{1}_{X^E(a)=0} B Q_a(a)da \\
\Psi_t &= \int (\psi[x_b(s), k_b] + \psi[x_g(s), k_g])(1-\lambda)\Phi(\mu, x_b^E(\cdot), x_g^E(\cdot), X^E(\cdot), B)(s)ds \\
&+ \int [\psi^X(k_b) + \psi^X(k_g)]\lambda\Phi(\mu, x_b^E(\cdot), x_g^E(\cdot), X^E(\cdot), B)(s)ds,
\end{aligned}$$

and where $T^*(\cdot)$ is the measure operator from one period to the next and $\Phi(\cdot)$ is the measure of operating firms.[11] The first constraint the planner faces is the resource constraint. The next constraint is the labor supply constraint, and the third constraint is the "adding-up" constraint relating the "big-K" to "little-k" for the context of brown capital. The next three constraints ensure that the planner internalizes that the law of motion of the firm distribution and carbon stocks depend on the choices made at the firm-level. Finally, the planner must obey the non-negativity constraint on the mass of newly entering firms.

As is notated by the dynamic program (5), the planner's state variables include the infinite-dimensional distribution of incumbent firms μ, as well as the permanent and persistent carbon stocks from period $t-1$, S^1 and S^2 respectively. The planner chooses the firm-level investment and continuation policies for entrants and operating firms and the mass of entrant firms.

Problem (5) can be written in Lagrangian form by attaching multiplier for the constraints:

$$
\begin{aligned}
\mathcal{L}_t = U(C_t) &+ \lambda_t^L \underbrace{\left[1 - \int (L^d(s) + c_f)\Phi(\mu, x_b^E(\cdot), x_g^E(\cdot), X^E(\cdot), B)(s)ds - B\kappa\right]}_{\text{Labor supply constraint}} \\
&+ \lambda_t^k \underbrace{\left[K^b - \int k_b \Phi(\mu, x_b^E(\cdot), x_g^E(\cdot), X^E(\cdot), B)(s)ds\right]}_{\text{Definition of total brown capital}} + \lambda_t^1 \underbrace{(S^{1\prime} - S^1 - (1-\chi_t)\varphi_1\gamma K^b)}_{\text{Law of motion of permanent stock}} \\
&+ \lambda_t^2 \underbrace{(S^{2\prime} - \varphi_3 S^2 - (1-\chi_t)(1-\varphi_1)\varphi_2\gamma K^b)}_{\text{Law of motion of persistent stock}} + \lambda_t^B \underbrace{(B-0)}_{\text{Nonnegative entrant mass}} \\
&+ \beta \mathcal{W}_{t+1}(T^*(\mu, x_b(\cdot), x_g(\cdot), X(\cdot), x_b^E(\cdot), x_g^E(\cdot), X^E(\cdot), B), S^{1\prime}, S^{2\prime})
\end{aligned}
$$

Proposition 2 relates the Lagrange multiplier on aggregate brown capital λ_t^k to the Lagrange multipliers on the permanent and persistent carbon stocks. Since an additional unit of aggregate brown capital creates additional uncaptured emissions that contribute to the permanent and persistent stocks of carbon emissions, the shadow value of that additional unit of aggregate brown capital is the discounted sum of the future marginal climate damages.

Proposition 2. *The Lagrange multiplier on aggregate brown capital, λ_t^k, can be written*

$$\lambda_t^k = (1-\chi_t)\gamma\varphi_1\lambda_t^1 + (1-\chi_t)\gamma(1-\varphi_1)\varphi_2\lambda_t^2,$$

[11] Both measure operators are formally defined for the context of the planner problem in Appendix B, the definitions closely match those from the competitive equilibrium.

where the Lagrange multipliers on the permanent and persistent carbon stocks are given by

$$\lambda_t^1 = \sum_{s=t}^{\infty} \beta^{s-t} U'(C_s) D'(S_s^1 + S_s^2) Y_s$$

$$\lambda_t^2 = \sum_{s=t}^{\infty} (\varphi_3 \beta)^{s-t} U'(C_s) D'(S_s^1 + S_s^2) Y_s.$$

Proof. See Appendix B.2. □

We then proceed as in Lucas and Moll (2014), Moll and Nuno (2018), and Ottonello and Winberry (2023) by defining the *augmented Bellman equation* and *augmented entry problem* in Definition 3.

Definition 3. *Let* $\hat{\tau}_t \equiv \frac{\lambda_t^k}{U'(C_t)}, \hat{w}_t \equiv \frac{\lambda_t^L}{U'(C_t)}, \frac{1}{\hat{R}_t} \equiv \beta \frac{U'(C_{t+1})}{U'(C_t)}$. *Then, the augmented Bellman equation is defined as*

$$\begin{aligned} \hat{\omega}_t(s, \mu, S^1, S^2) = \max_{L, x_b, x_g, X} & (1 - D(S^{1\prime} + S^{2\prime})) \exp(z) A_t^{1-\alpha} [a k_b^\rho + (1-a) k_g^\rho]^{\frac{\alpha}{\rho}} L^\nu \\ & + (1-\lambda) \left[-x_b - x_g - \psi(x_b, k_b) - \psi(x_g, k_g) \right] \\ & + \lambda \left[(1-\delta) k_b + (1-\delta) k_g - \psi^X(k_b) - \psi^X(k_g) \right] \\ & - \hat{w}_t (L + c_f) - \hat{\tau}_t k_b \\ & + (1-X)(1-\lambda) \frac{1}{\hat{R}_t} \mathbb{E}[\hat{\omega}_{t+1}(s', \mu', S^{1\prime}, S^{2\prime})] \end{aligned} \tag{6}$$

subject to:

$$\begin{aligned} x_i &= -(1-\delta) k_i \text{ if } X = 1 \\ k_i' &= (1-\delta) k_i + x_i, \end{aligned}$$

Then, we define the augmented entry problem:

$$\hat{\omega}_t^E(a, q, \mu, S^1, S^2) = \max_{\hat{X}^E(a,q), \hat{x}_b^E(a,q), \hat{x}_g^E(a,q)} (1 - \hat{X}^E(a,q)) [-\hat{x}_b^E(a,q) - \hat{x}_g^E(a,q) + \mathbb{E}[\hat{\omega}_t(s, \mu, S^1, S^2)]] \tag{7}$$

subject to

$$\hat{x}_i^E(a,q) = -(1-\delta) k_i \text{ if } \hat{X}^E(a,q) = 1 \; \forall i \in \{b, g\}.$$

Next, we prove that the policies induced by the social planning problem match the policies induced by the augmented firm problem (Proposition 4) and that the augmented firm problem can be interpreted as the marginal social value of firms at s (Proposition 5).

Proposition 4. *The augmented Bellman equation and augmented entry problem defined in (6) and (7) induces firm-level and entrant-level policies that match those of the social planner, whose problem is defined by (5).*

Proof. See Appendix B. □

Proposition 5. *The augmented Bellman equation defined in (6) induces values* $\hat{\omega}_t(s, \mu, S^1, S^2)$ *equal to the Gateaux derivative of social welfare with respect to the measure of firms in a state s, denoted in units of the consumption good. That is,*

$$\hat{\omega}_t(s, \mu, S^1, S^2) = \frac{\partial \mathcal{W}_t(\mu, S^1, S^2)}{\partial \mu(s)} \frac{1}{U'(C_t)}$$

Proof. See Appendix B. □

Finally, we can characterize the first-order condition with respect to the mass of entering firms using the augmented Bellman equation (Proposition 6).

Proposition 6. *The planner's first-order condition with respect to B can be written as the following:*

$$\mathbb{E}[\hat{\omega}_t^E(a, \mu, S^1, S^2)] + \frac{\lambda_t^B}{U'(C_t)} = \hat{w}_t \kappa, \tag{8}$$

Proof. See Appendix B.1. □

The authors note that, while the objects $\mathcal{P}_t \equiv \{\hat{\tau}_s, \hat{w}_s, \frac{1}{\hat{R}_s}\}_{t=s}^{\infty}$ are written notationally to be reminiscent of taxes and prices, in the planner's problem they are more accurately described as internalized damage of emissions, the marginal value of labor in units of the consumption good given the exogenous supply, and the SDF of the household under the optimal policy. We next show that $\mathcal{P}_t$, interpreted as taxes and prices, implements the planner's solution in a competitive equilibrium.

4.1 Decentralization through Brown Capital Taxes

We show that the definition of an equilibrium is satisfied under the set of taxes $\{\hat{\tau}_s\}_{s=t}^{\infty}$ and prices $\{\hat{w}_s, \hat{R}_s\}_{s=t}^{\infty}$ given by the planner's problem, and that the equilibrium policies and mass of entrants match those of the social planner (Proposition 7).

Proposition 7. *Under the set of taxes* $\{\hat{\tau}_s\}_{s=t}^{\infty}$ *matching those defined in Definition 3, the set of wages and risk free interest rates given by* $\{\hat{w}_t, \hat{R}_t\}_{s=t}^{\infty}$*, also as defined in Definition 3, induces a competitive equilibrium. Moreover, the firm-level policies and mass of entrants within the induced competitive equilibrium match that of the social planner.*

Proof. See Appendix C. □

The intuition is simple: When the distortionary taxes are levied on the firms in the economy, their dynamic program at the appropriate prices is identical to the dynamic program that characterizes the planner's solution, and hence the solutions are identical. The firm-level policies and values match exactly those of the planner. Then, Equation (8) ensures the competitive entry condition is satisfied, the labor market clearing is ensured by the labor resource constraint of the planner under the planner's chosen mass of entrants B_t, and Walras' law ensures market clearing in the risk-free asset market. Hence, a competitive equilibrium can be constructed from the appropriate objects within the planner's problem, namely the elements of $\mathcal{P}_t$.

In Appendix C.1, we demonstrate that the decentralization of the planner's allocation holds unchanged under a model extension which aligns the unobservability of brown capital by the taxation authority with reality. More specifically, we assume that the taxation authority does not observe brown capital directly, but can observe emissions. If emissions remained deterministic as in our baseline model, the taxation authority would easily be able to back out a firm's brown capital from their emissions so to make the problem more interesting we allow for some idiosyncratic mean preserving noise in the emissions generation process. This noise then rationalizes the within-firm variation in emissions intensity that we observe as a small part of the variation in our data in Section 2. Under this model extension, the idiosyncratic noise in emissions washes out from the planner's perspective and the constrained optimal allocation remains unchanged. A tax on emissions directly, as opposed to the unobserved brown capital, is able to implement the constrained optimal allocation. As the constrained optimal allocation, decentralized laissez-faire allocation, and implementability of constrained optimal allocation via a carbon tax remain unchanged under this model extension, the economies are effectively equivalent.

Finally, in Appendix C.2 we show that the constrained optimal allocation can be implemented not only through a carbon tax, but also via a carbon credit system where the quantity of allowed emissions is chosen by the government and the price of the emissions credits clears a competitive market with the government as an inelastic seller. The price of the credit then plays a similar role to the carbon tax and the price that clears the market is the Pigouvian tax from the baseline model. While this result may be unsurprising in the context of Pigouvian taxation, we lever the equivalency for our validation in Section 6.

5 Quantifying the Green Transition

We next quantify the transition to the long-run, through the climate-economy transition we face today, given the model outlined in Section 3. We first describe the calibration of the model and then compute the transition to the long-run over a horizon of 10 thousand years. We compare this "green transition" under the BAU laissez-faire allocation versus the constrained-optimal allocation.

5.1 Calibration

We present our calibrated parameters in Table 3. The selected values for parameters are largely based on commonly used parameters and functional forms in existing literature to allow comparability, we then internally calibrate parameters for novel model ingredient using simulated method of moments (SMM). We initiate transition paths in 1990 where the assume the economy is on its balanced growth path with zero brown capital taxes. We use an exogenous productivity growth rate of $\iota = 0.01$ following Krusell and Smith (2022) and calibrate to an annual risk-free rate of 4 percent in line with estimates of very-long discount rates from Giglio, Maggiori, Rao, and Stroebel (2021); this corresponds to an annual discount factor of $\beta = 0.971$. We set household preference to be log-utility.

In the firm block, we set standard parameters outside the model following the firm dynamics literature and calibrate both the estimated distribution of carbon dependence of entrants and fixed cost of production using GMM. For externally calibrated parameters, we take standard parameters for the capital $\alpha = 0.3$ and labor $\nu = 0.65$ share as well as capital depreciation $\delta = 0.12$. Idiosyncratic productivity follows a log AR(1) process. We use estimates from Khan and Thomas (2013) for the persistence and volatility of log-productivity ($\rho = 0.659$ and $\sigma_z = 0.118$, respectively). Following Clementi and Palazzo (2016), we assume that entrant productivity signals are Pareto distributed with tail parameter $\Xi = 2.69$, with the log-TFP parameters also governing the transition from signal to initial productivity $\rho_z = \rho_q, \sigma_z = \sigma_q$. We assume quadratic adjustment costs

$$\psi(x, k) = \hat{\psi} \left(\frac{x}{k} \right)^2 k$$

and set curvature of the capital adjustment costs $\hat{\psi} = 0.297$ following Corbae and D'Erasmo (2021). For the exogenous exit rate λ, we use 0.025, which is the employment-weighted firm exit rate between 1979 and 2012 from Decker, Haltiwanger, Jarmin, and Miranda (2016). At this rate, firms are expected to receive an exit shock every 40 years. As a baseline, we assume that production is Cobb-Douglas in brown and green capital with $\rho = 0$. In Section

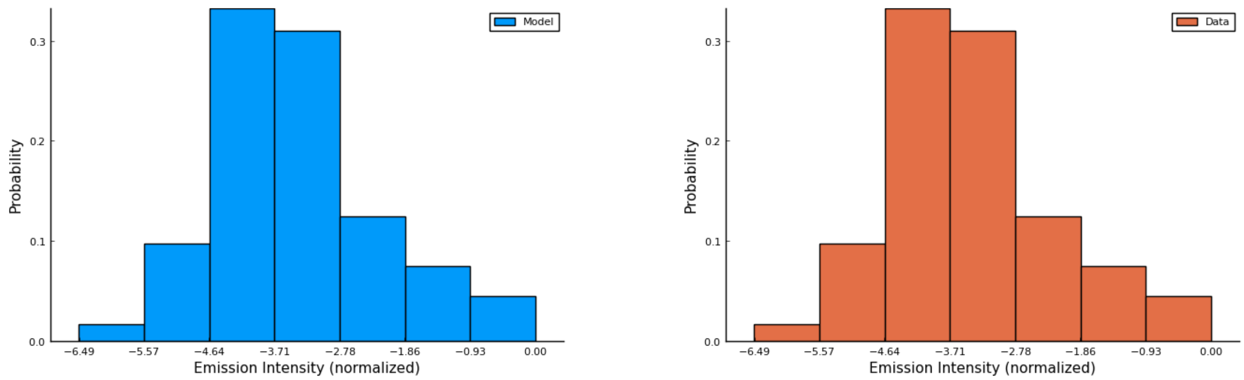
7, we evaluate the sensitivity of our results to this choice of ρ by computing green transitions using the entire range of possible values of ρ from brown and green capital being perfect substitutes to perfect complements and find that our quantitative results are quite robust. Finally, we start each transition path on the balanced growth path with no carbon taxes and with environmental damage fixed at the value implied by the 1989 level of carbon stock, then back out the implied entry cost from the entry condition normalizing the wage in the BAU case to 1.[12]

Next, we estimate the distribution of carbon dependence of entrants $Q(a)$ and the fixed cost of production c_f using simulated method of moments. We target the distribution of emission intensity of operating firms and the exit rate. The first panel of Figure 6 shows the firm-level fixed effects from the emission intensity regression that measure emissions as the sum of Scope 1 and Scope 2 from Section 2. In order to map into the model, we need to make a normalization; in particular, we assume that the firm with the highest observed emission intensity uses only brown capital. The second panel shows the analogous distribution from the model; it is the distribution of emission intensity of operating firms in the BAU balanced growth path with constant environmental damage when $Q(a)$ equals its estimate. The flexibility in setting $Q(a)$ allows us to target the model distribution to the data distribution quite closely. We outline the estimation procedure in Appendix E. In addition, we target an unconditional exit rate of 9 percent from Decker, Haltiwanger, Jarmin, and Miranda (2016), which results in a fixed cost of production c_f of 0.021.

[12]This class of models is well-known to feature a one-to-one relationship between the equilibrium wage and entry cost κ. Hence, different calibrations of the entry cost κ only affect the level of the equilibrium wage. It is very convenient in the computation of the model to fix the wage at one and recover the entry cost consistent with this choice, see for example Corbae and D'Erasmo (2021).

Figure 6: Calibration of Entrant Emission Dependence

(a) Emission Intensity in Data (b) Emission Intensity in Model



Notes: This figure shows the GMM targets for the estimation of entrant emission dependence. The first panel shows the firm-level fixed effects from the OLS regression of Scope 1 and 2 emission intensity from Section 2. The second panel shows the emission intensity distribution in the BAU balanced growth path with constant environmental damage.

Regarding the climate block of the model, we follow Golosov, Hassler, Krusell, and Tsyvinski (2014) in using the following damage function:

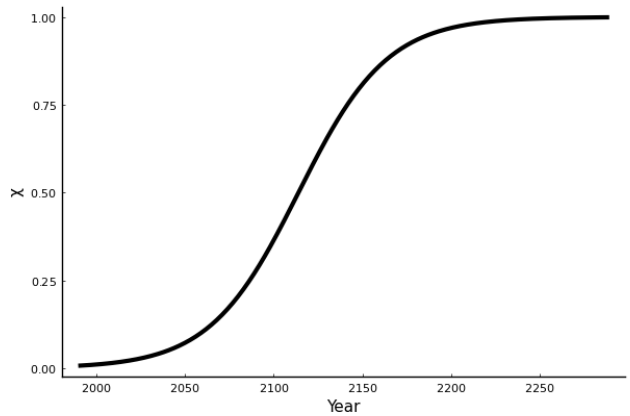
$$D(S) = 1 - \exp(-\Delta(S - \bar{S}))$$

where $\bar{S}$ is the pre-industrial stock of carbon emissions. Furthermore, we follow Krusell and Smith (2022) in how we model the rest of the carbon cycle. In particular, we initialize the stock of permanent and transitory carbon emissions to their levels in 1989, and adopt how Krusell and Smith (2022) model new emissions' contributions toward permanent and transitory stocks of carbon emissions, and the persistence of the transitory component of carbon emissions. We assume the same functional form for expected exogenous improvements in emission capture technologies as Krusell and Smith (2022):

$$\chi_t = \begin{cases} 1 - \left(1 + \exp\left(\log\left(\frac{0.01}{0.99}\right)\right)\frac{t - t_{\chi=0.5}}{t_{\chi=0.01} - t_{\chi=0.5}}\right)^{-1}, & \text{if } t < t_{\chi=1} \\ 1, & \text{if } t \geq t_{\chi=1} \end{cases} \tag{9}$$

This functional forms results in one-percent of emissions are captured at $t_{\chi=0.01}$, one-half of emissions are captured at $t_{\chi=0.5}$, and all emissions after $t_{\chi=1}$. We plot this function in Figure 7. Finally, we use the amount of emissions in 1989 reported by Krusell and Smith (2022) to back out the quantity of emissions per unit of brown capital (γ) that makes the level of brown capital in 1990 consistent with the level of emissions produced in the data.

Figure 7: Captured Fraction of Emissions



Notes: Figure 7 shows the fraction of emissions that are captured before they enter the atmosphere and the carbon cycle. We follow Krusell and Smith (2022) in assuming that one percent of emissions are captured before entering the atmosphere by 2000, half of the emissions are captured by 2115, and all emissions are captured by 2291. This functional form is given by Equation (9).

Table 3: Parameters

(a) Preferences

Parameter	Value	Source	Description
β	0.971	$R^{BGP} = 1.04$ from Giglio et al. (2021)	Time preference
ι	0.01	Krusell and Smith (2022)	Economy growth rate

(b) Production

Parameter	Value	Source	Description
α	0.3	Standard parameter	Capital share
ν	0.65	Standard parameter	Labor share
λ	0.025	Decker et al. (2016)	Exogenous exit rate
$\hat{\psi}$	0.297	Corbae and D'Erasmo (2021)	Capital adjustment cost
δ	0.12	Standard parameter	Depreciation rate
ρ	0.0	Cobb-Douglas baseline	Substitutability between capital types
ρ_z, ρ_q	0.659	Khan and Thomas (2013)	Idiosyncratic productivity persistence
σ_z, σ_q	0.118	Khan and Thomas (2013)	Idiosyncratic productivity volatility
Ξ	2.69	Clementi and Palazzo (2016)	Pareto tail of entrant productivity signal
κ	0.0071	Consistent with unit wage	Entry cost

(c) Climate

Parameter	Value	Source	Description
Δ	0.000053	Golosov et al. (2014)	Damage function parameter
φ_1	0.2	Krusell and Smith (2022)	Fraction of permanent emissions
φ_2	0.398	Krusell and Smith (2022)	Fraction of dissipated persistent emissions
φ_3	0.998	Krusell and Smith (2022)	Persistence of persistent emission stock
$\bar{S}$	581	Golosov et al. (2014)	Pre-industrial level of emissions
$S_{9,1}$	684	Krusell and Smith (2022)	Stock of persistent emission in 1999
$S_{9,2}$	118	Krusell and Smith (2022)	Stock of permanent emission in 1999
E_9	8.741	Krusell and Smith (2022)	Emissions in 1999
$t_{\chi=0.01}$	10	Krusell and Smith (2022)	Years until 1% of emissions are captured
$t_{\chi=0.5}$	125	Krusell and Smith (2022)	Years until half emissions are captured
$t_{\chi=1}$	301	Krusell and Smith (2022)	Years until all emissions are captured
γ	39.137	Consistent with total emissions in 1990	Emissions per unit of brown capital

Note: This table outlines the parameterization for the quantitative model. See text for discussion of the calibration of model parameters.

5.2 Model Validation

After calibrating the model, we next turn to demonstrating that our model is consistent with the climate-economy in two dimensions. First, we show that our model is consistent with the cross-sectional relationship between emissions intensity and firm-level outcomes, and second we show that the model's response to fluctuations in the price of carbon matches that of the data.

5.2.1 Cross-Sectional

Using data on U.S. public firms, Kim (2023) documents a negative relationship between firm-level emissions per unit of sales to various measures of firm-level productivity and factor misallocation as in Hsieh and Klenow (2009). Here, we validate our model and calibration by replicating the empirical findings of Kim (2023). Kim (2023) focuses on four measures: (1) marginal revenue product of capital (MRPK) defined as sales over capital, (2) marginal revenue product of labor (MRPL) defined as sales over labor, and (3) revenue productivity (TFPR) defined as the geometric mean of MRPK and MRPL using capital share parameter estimated using a control function approach as in Olley and Pakes (1996), and (4) productivity, which is the residual from an estimated production function. To replicate these measures, we first simulate a large cross-section of operating firms from the initial balanced growth path of our model. Constructing the model-equivalents of MRPK and MRPL is immediate. To construct TFPR, we use the true capital share α of the data-generating process. Similarly, to construct productivity, we use the true productivity z_j of the data-generating process. We then perform the identical OLS regressions as Kim (2023):

$$y_j = \beta_{0,i(j)} + \beta_1 \log\left(\frac{Emissions_j}{Output_j}\right) + \varepsilon_j$$

where y_j is either MRPK, MRPL, TFPR, or productivity. Kim (2023) includes fixed effects at the 4-digit industry level. To bound the precision of these fixed effects in our setting without explicit industries, we run both specifications with a common intercept for all firms (analogous to the coarsest industry definition) and intercepts that depend on the carbon dependence of production a_j (analogous to the finest industry definition).

Table 4 shows the regression results. Our model delivers negative coefficients for MRPK, TFPR, and productivity found by Kim (2023). Indeed, the coefficients from Kim (2023), -0.152 for MRPK, -0.094 for TFPR, and -0.091 for productivity, even sit comfortably between our estimates without and with fixed effects for carbon dependence.[13] Our model does not deliver any dispersion in MRPL across firms (omitted from table) because labor is employed on the spot market in Cobb-Douglas production, so labor is a constant fraction of output. Concerns that our model does not deliver this negative relationship may be somewhat alleviated because the MRPL coefficient from Kim (2023) is -0.075, which is the

[13] Kim (2023) does not report the regression coefficient on productivity; we used WebPlotDigitizer to back it out from Figure 3(b).

weakest across the four measures.

Table 4: Relationship of Emissions over Output to Measures of Productivity

	log(MRPK)		log(TFPR)		Productivity	
	Estimates from Kim (2023)					
$\log(\frac{\text{Emissions}}{\text{Output}})$	-0.152		-0.094		-0.091	
	Model Simulation					
$\log(\frac{\text{Emissions}}{\text{Output}})$	-0.123	-1.000	-0.037	-0.300	-0.040	-0.338
Carbon Dependence FEs?	No	Yes	No	Yes	No	Yes

Note: This table presents the comparison between regressions ran on model-simulated firm data and the empirical estimates of Kim (2023). The coefficients are untargeted. We compare the Kim (2023) coefficients to model coefficients including and not including fixed effects for carbon dependence a. We view the industry fixed effects included by Kim (2023) as being a level of industry definition in-between not including carbon dependence fixed effects and including them. Reassuringly, the coefficients of Kim (2023) fall between the coefficients estimated on model-simulated data.

Kim (2023) uses these negative relationships to motivate financial frictions driving variation in emission intensity. Despite our model featuring no financial frictions, it can replicate these empirical results. An alternative way to explain these regression results in the data is that using sales or output as a denominator for emissions intensity may not be innocuous when examining the relationship between emissions and productivity. Since productivity raises output, this regression specification mechanically incorporates simultaneity bias.[14]

5.2.2 Time-Series

As further model validation, we recreate the impulse response function estimates from Känzig (2023) within our model. Känzig (2023) uses structural vector-autoregression techniques to estimate the effects of shocks in the price of emissions permits in the European Union's emissions-trading system (ETS) using policy announcements. As we show in Appendix C.2, a carbon credit system such as the ETS is an alternative implementation mechanism for the

[14]More concretely, consider a simplified model where operating capital produces emissions linearly $Emissions = \gamma \cdot Capital$ and output is productivity times capital $Output = z \cdot Capital$. Then regressing emissions over output on productivity mechanically results in a negative relationship:

$$\frac{Emissions}{Output} = \beta_0 + \beta_1 z \implies \frac{\gamma}{z} = \beta_0 + \beta_1 z$$

In our model, these linear relationships do not hold but the mechanical negative correlation induced by simultaneity does hold, as evidenced by Table 4.

constrained-efficient allocation. In this system, changes to the price of carbon are analogous to changes to carbon taxes in our baseline decentralization. Therefore, we construct an analogous exercise to that of Känzig (2023) within our model by feeding in the path of deviations of the cost of carbon from steady-state from Känzig (2023) as a path of carbon taxes in our model and compute the response of the macroeconomy to this path of taxes.

More specifically, we take the path of harmonized index of consumer prices (HICP) energy from Figure 3. Känzig (2023) plots results for 40 months, so we feed in the estimate for month 0, 12, 24, and 36 and then a constant geometric decay over a further sixteen years equal to the decay between months 24 and 36 in our model, which is calibrated to an annual frequency. We then solve the resulting transition allowing general equilibrium changes in wages, the mass of entrant, and the interest rate. We hold climate damage constant to damage on the initial BGP.

Table 5 reports analogous measures within the model to what Känzig (2023) reports. In particular, the first and second column compares percent change in aggregate output in the model, which corresponds to industrial production in the data. Our model predicts a smaller drop in output, but well within the confidence intervals of Känzig (2023)'s estimates. The third and fourth columns compare percent changes in the general price level. In our model, consumption is the numeraire, so here we compare percent changes in one over the wage, which is the price of the consumption good under a change in numeraire to labor. Our model does quite well and is close to the point estimates of Känzig (2023). Finally, we compare interest rate changes in the fifth and sixth column. Our model predicts that interest rates drop when taxes are imposed because aggregate consumption falls. Our model-generated impulse response function is quite close to the point estimates of Känzig (2023) except at one year. Känzig (2023) finds that interest rates increase at one year, but the confidence interval is quite wide and includes zero.

Table 5: Impulse Response Function Comparison

Years	Output/IP		HICP		2-Year Interest Rate	
	Känzig	Model	Känzig	Model	Känzig	Model
0	-0.247	-0.106	0.173	0.169	-0.038	-0.068
1	-0.602	-0.102	0.196	0.126	0.255	-0.048
2	-0.765	-0.064	0.156	0.089	-0.022	-0.031
3	-0.448	-0.017	0.084	0.059	-0.038	-0.016

Note: This table outlines results from recreating impulse responses from Känzig (2023). Output and price level changes are in percent change from balanced growth path level, while interest rate changes are in percentage points.

5.3 The Green Transition

After validating our model's cross-sectional relationship between carbon dependence and misallocation, as well as the time-series response of the economy to the price of carbon, we present the green transition as implied by the model. Throughout the figures in this section, the blue lines represent the transition under the laissez-faire allocation while the orange lines represent the transition under the carbon taxes which implement the constrained-efficient allocation in a decentralized equilibrium.

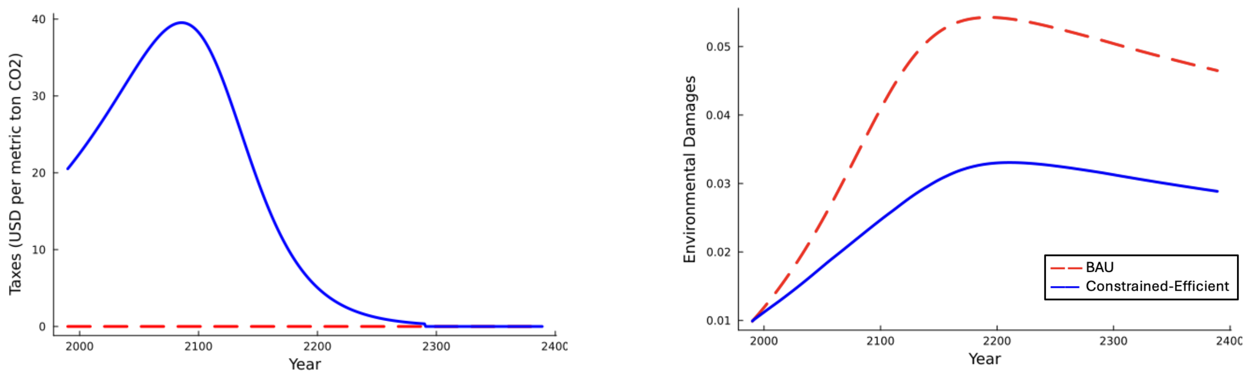
In Figure 8, we present key climate objects from the computed green transitions. The left panel illustrates the tax on brown capital, whereas the right panel displays environmental damage. The optimal taxes on brown capital are forward-looking, internalizing the marginal cost of carbon for firms. Firms then respond to this tax in two ways. Firstly, the carbon tax directly incentivizes all firms to shift their capital allocation toward green capital. Secondly, it encourages firms with higher carbon intensities (high values of a) to endogenously exit. To prevent the accumulation of permanent emissions and mitigate the long-term damage to output, brown capital taxes are high during the early stages of the transition path. As the technology to capture emissions improves and all emissions are captured $\chi_t \to 1$, brown capital taxes no longer prevent future climate damage, so the optimal carbon tax falls to zero $\tau \to 0$.

Environmental damages begin at relatively lower levels and gradually increase as carbon emissions accumulate in the atmosphere. Once all new carbon emissions are captured (χ_t reaches unity), climate damages experience a gradual decline as the persistent component of the stock of carbon emissions dissipates. Ultimately, climate damages stabilize at a level solely caused by the permanent stock of carbon emissions. The optimal taxes result in lower climate damage throughout the entire transition path.

Figure 8: Brown Capital Taxes and Environmental Damage

(a) Brown Capital Taxes

(b) Environmental Damage

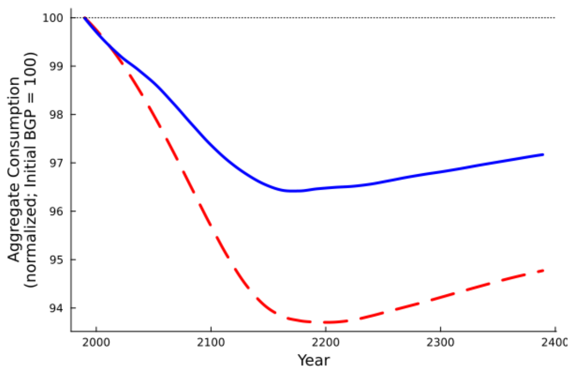


Notes: Figure 8 describes the climate objects (brown capital taxes and environmental damage) computed at two equilibrium allocations. The first equilibrium is the BAU allocation, defined by zero taxes on firms. The second equilibrium is the constrained-efficient allocation, implemented in competitive equilibrium through carbon taxes.

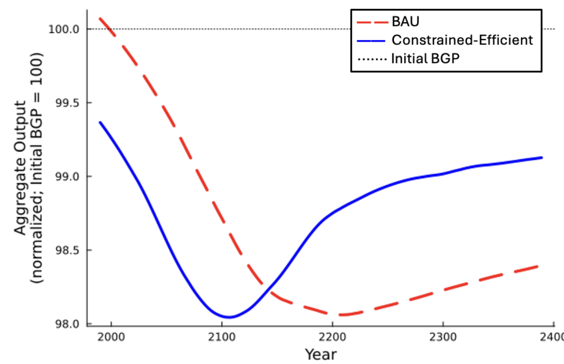
Figure 9 presents detrended aggregate consumption and output under the two allocations. In the absence of taxes, aggregate consumption and output experiences a decline due to climate damages from accumulated carbon emissions. The consumption equivalent of moving from the BAU allocation to the constrained-efficient allocation is 0.244%. Turning to aggregate output, this disinvestment leads to a drop in aggregate output as the measure of operating firms is lower. Aggregate consumption with optimal taxes is higher throughout the transition path despite lower aggregate output due to lower climate damages. Once all emissions are captured ($\chi_t = 1$), aggregate output (net of climate damages) under optimal taxes starts to exceed aggregate output under zero taxes because climate damages with optimal taxes are lower.

Figure 9: Aggregate Quantities

(a) Consumption



(b) Output Net of Environmental Damage



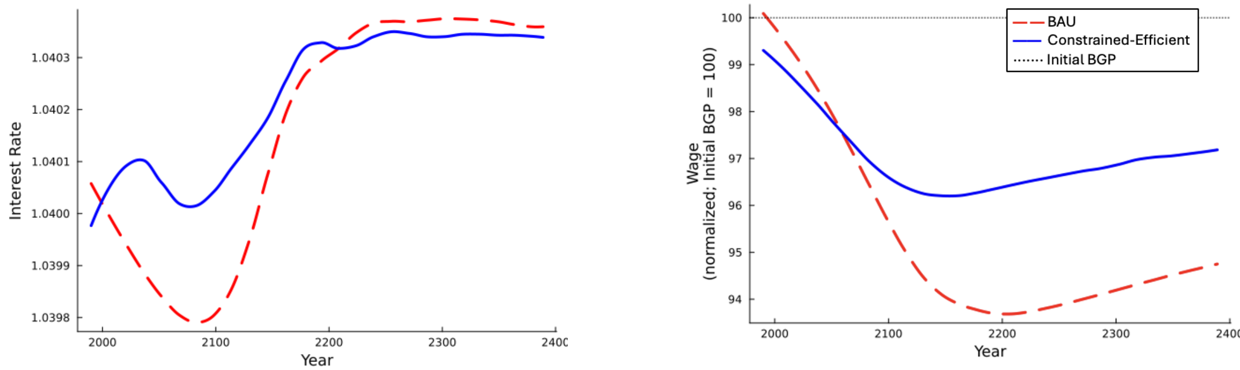
Notes: Figure 9 describes the evolution of macroeconomic aggregates in the two computed equilibria. The first equilibrium is the BAU allocation, defined by zero taxes on firms. The second equilibrium is the constrained-efficient allocation, implemented in competitive equilibrium through carbon taxes.

In Figure 10, we illustrate the endogenous movements of interest rates and wages over the transition path. Interest rates are determined by the consumption Euler equation of the representative household. With zero taxes, as climate damages reduce aggregate output, interest rates experience a slight decline due to the corresponding decrease in aggregate consumption. In the constrained-efficient allocation, there is an initial jump in aggregate consumption as firms respond to the higher tax burden by disinvesting, which in-turn causes the interest rate to drop. As brown capital taxes gradually decrease and climate damage converges to a constant, the interest rate returns to its balanced growth path level.

In the BAU case, as the climate damage worsens the marginal product of labor in the economy falls. Under the constrained-efficient allocation, the marginal product of labor falls immediately as investment is lower, reducing the marginal product of labor. In the long run, climate damage is lower than it is in the BAU case, and the marginal product of labor is higher.

Figure 10: Prices

(a) Interest Rates (b) Wages



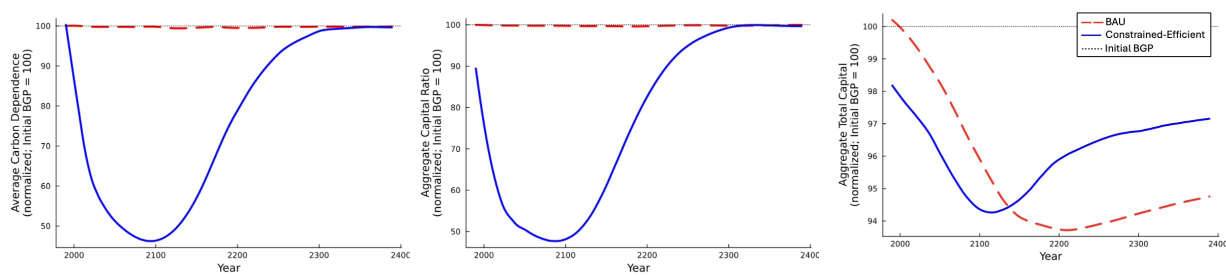
Notes: Figure 10 describes the equilibrium prices in the two computed equilibria. The first equilibrium is the BAU allocation, defined by zero taxes on firms. The second equilibrium is the constrained-efficient allocation, implemented in competitive equilibrium through carbon taxes. The equilibrium interest rate falls immediately under the constrained-efficient allocation as firms optimally invest less in brown capital and less in capital overall in response to the internalized environmental externalities. As the emissions capturing technology improves, firms increase investment and the interest rate rises. The marginal product of labor falls in the constrained-efficient allocation due to the reduction in investment in capital. Hence, the wage falls. Wages eventually increase as climate damages fall.

In Figure 11, we examine the intensive and extensive margins of capital in the constrained-efficient allocation relative to the BAU allocation. We can examine the composition effect of brown capital taxes on the measure of firms, as measured by the average carbon dependence a of firms in the economy. Without Pigouvian taxes, the average carbon dependence remains unchanged. However, with the implementation of carbon taxes, we observe a higher rate of endogenous exits among firms with highly carbon-intensive production processes. As a result, the average carbon dependence decreases, indicating a shift towards greener production technologies. As emission capturing technology progresses and eventually captures all emissions, leading to the phase-out of brown capital taxes, the average carbon dependence returns to its initial level.

We next examine the intensive margin as through two main effects. First, we illustrate the *substitution effect* through the ratio of aggregate brown capital to aggregate green capital. In the laissez-faire economy, this ratio stays at unity. In the constrained-efficient allocation, the capital ratio drops as the carbon taxes incentivize firms to use more green capital relative to brown capital. Once $\chi_t \to 1$, the capital ratio returns to unity. Secondly, we examine the *scale effect* through the aggregate total capital. Whether the constrained planner's choice of total capital at the firm-level is higher or lower than the laissez-faire allocation is generally ambiguous due to two offsetting forces. Holding climate damages fixed, the planner lowers total investment because the planner internalizes the capital choices on climate damage (analogous to a "partial equilibrium" change). However, the social planner also lowers climate damage through their improved allocation, so firms are effectively more productive which increases the marginal product of capital (analogous to a "general equilibrium" change). In our calibrated model, the partial equilibrium effect dominates before $\chi_t \to 1$ and the planner lowers total capital relative to the BAU allocation. The general equilibrium effect dominates after $\chi_t \to 1$ and the planner raises total capital relative to the BAU allocation.

Figure 11: Composition, Substitution, and Scale Effects

(a) Composition Effect (b) Substitution Effect (c) Scale Effect



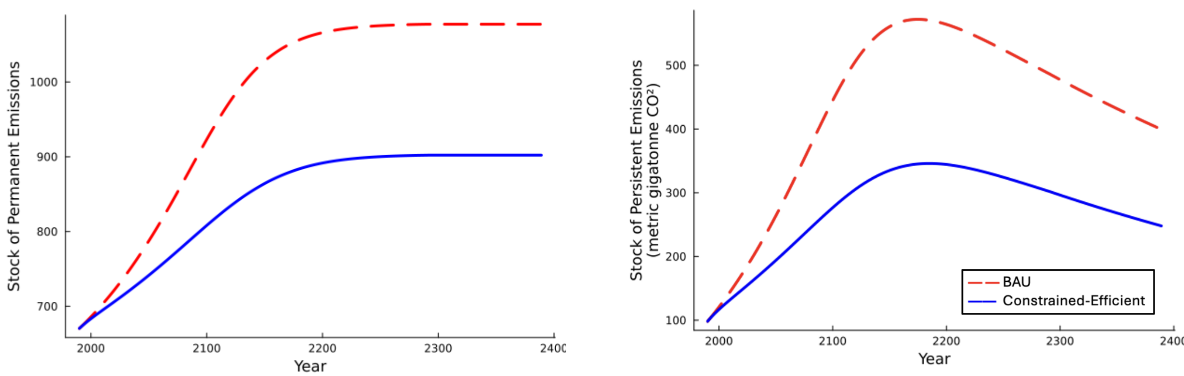
Notes: Figure 11 describes the composition, substitution, and scale effects by comparing the average a level, the aggregate capital ratio, and aggregate total capital in the two computed equilibria. The first equilibrium is the BAU allocation, defined by zero taxes on firms. The second equilibrium is the constrained-efficient allocation, implemented in competitive equilibrium through carbon taxes. The average a in the economy falls under the constrained-efficient allocation because the social planner exits more brown firms than green firms, whereas the firms are symmetric in the BAU economy. The aggregate capital ratio in the constrained-efficient allocation falls (more green capital relative to brown capital), while it stays at one in the BAU allocation. Aggregate capital falls, initially more for the constrained-efficient allocation relative to BAU but later the constrained-efficient allocation has a higher level of aggregate capital than the BAU allocation.

Figure 12 shows the climate outcomes under both sets of taxes. When zero taxes are implemented, aggregate brown capital remains unchanged, resulting in a consistent flow of emissions. These emissions contribute to both the permanent and persistent stock of emissions, which in turn cause environmental damage to output.

Under the optimal tax scheme, the level of brown capital decreases leading to a reduction in the flow of emissions. This decline in emissions has two significant effects: it lowers the level of the permanent stock of emissions in the atmosphere and reduces the peak of the persistent stock of emissions. Consequently, environmental damages are mitigated both in the short run, when the persistent stock of emissions causes damages, and in the long run, after the persistent stock of emissions has fully dissipated.

Figure 12: Climate

(a) Permanent Stock of Emissions (b) Persistent Stock of Emissions



Notes: Figure 12 describes the climate variables in the two computed equilibria. The first equilibrium is the BAU allocation, defined by zero taxes on firms. The second equilibrium is the constrained-efficient allocation, implemented in competitive equilibrium through carbon taxes. Emissions are substantially lower in the short run in the constrained-efficient allocation than in the BAU allocation. This results in a lower permanent and persistent stock of emissions over time, and a substantially lower environmental damage under the constrained-efficient allocation.

6 Missing the Trees for the Forest

What is missed when carbon taxes are set without considering heterogeneity in carbon dependence across firms? In this section, we evaluate the effect of this heterogeneity on the model-implied Pigouvian tax. We do so by considering an alternative model where we collapse set of carbon dependence a to singleton, which eliminates the composition effect in our model. Under this change, the social planner is no longer able to affect the distribution of carbon dependence through firm entry and exits, but the intensive margin adjustments both in terms of the ratio of green-to-brown capital and total capital are retained.

6.1 Pigouvian Taxes absent Heterogeneous Carbon Dependence

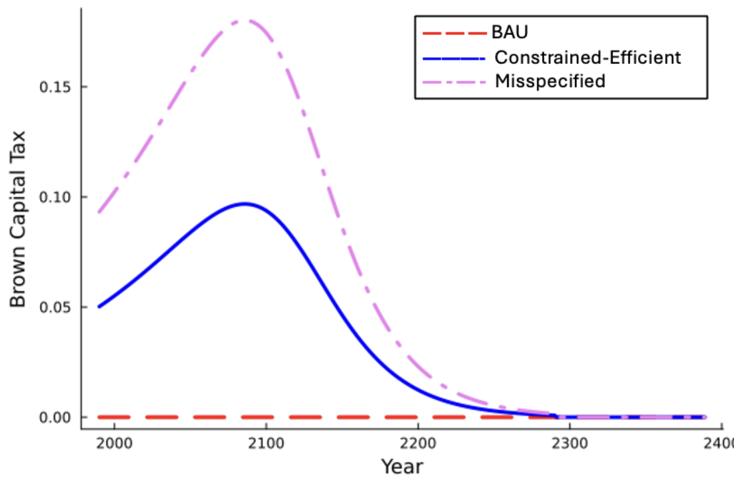
The first question we wish to answer with this experiment is, how different would the computed optimal tax path be if the modeler chose to ignore heterogeneity in carbon dependence? We answer this question by first recalibrating our model without heterogeneity to ensure it remains consistent with the data, excluding of course matching the distribution of carbon emission intensity. We instead match the singleton $a_{\text{single}} = 0.0624$ to the mean emissions intensity in our data. We recalibrate the fixed cost of production c_f, the entry cost κ, and emissions per unit of brown capital γ, by targeting the exit rate of 9 percent (Decker, Haltiwanger, Jarmin, and Miranda, 2016) and initial emissions level E_9. Table 7 shows the recalibrated parameters. In this alternative model, we then recompute optimal green transition and the associated Pigouvian carbon taxes. Figure 13 shows that the model-implied Pigouvian tax from the model without heterogeneity is much larger than the Pigouvian tax from the model with heterogeneity. Over the first 100 years of the green transition, the Pigouvian tax rate from the model without heterogeneity (the "misspecified" model) is 85.7 percent higher than the tax rate from the model with heterogeneity. The Pigouvian tax is higher under the model without heterogeneity because the marginal social damage of emissions are higher when the composition effect of carbon taxes are absent.

Table 6: Recalibration in Counterfactual Model

Parameter	Description	Target	Baseline	Alternative
κ	Entry Cost	Unit wage in initial BGP	0.007098	0.007134
γ	Emissions per unit of brown capital	Total emissions in 1990	39.137	72.422
c_f	Fixed cost of production	Exit rate = 9%	0.02118	0.02115

Note: This table describes the recalibration of the alternative model without heterogeneity. The baseline model is our model with heterogeneity in firms' carbon dependence a. In the alternative model, we shut down heterogeneity in this dimension by collapsing the distribution of carbon dependence a to a singleton.

Figure 13: Optimal Taxes Without Heterogeneity in Carbon Dependence



Note: This figure displays the path of the Pigouvian tax on carbon that implements the constrained-efficient allocation in the decentralized competitive equilibrium, under the baseline model (blue line) and alternative model without heterogeneity in carbon dependence (pink dash-dotted line). The Pigouvian tax is much higher in the model without heterogeneity as the marginal social damage from emissions are higher absent the composition effect.

6.2 Welfare Costs of Model Misspecification

We next aim to measure how costly it is to ignore heterogeneity in firms' carbon dependence when computing optimal carbon taxes. To this aim, we feed in the implied Pigouvian tax from the model without heterogeneity into the model with heterogeneity and solve for the green transition under this tax scheme. Since carbon taxes are higher in the model absent heterogeneity, consumption and output are lower in the short-run, but with fewer emissions and lower climate damage they are higher in the long-run. Despite the tax path being suboptimally large for the model with heterogeneity, we find that there are still substantial welfare gains to implementing this path of taxes as opposed to no taxes at all. The consumption equivalent of moving from the BAU and the transition with too-large taxes, backed out from the misspecified model, is 0.1718%, which is 70.3% percent of the consumption equivalent of moving from the BAU allocation and the socially optimal allocation.

Table 7: Welfare Consequences of Model Misspecification in Setting Carbon Taxes

Implemented Tax Path	CE (%)	CE / Optimal Tax CE (%)
Baseline Model with Optimal Taxes from Baseline Model	0.2444	100
Baseline Model with Optimal Taxes from Misspecified Model	0.1718	70.3

Note: This table describes the welfare consequences of model misspecification in setting carbon taxes. We perform this comparison by first collapsing the relevant dimension of heterogeneity in our model, firms' carbon dependence a, to a singleton and compute the Pigouvian tax which is optimal under the alternative model. This carbon tax path is much larger than the tax path that is optimal under the baseline model with heterogeneity. We then solve for the competitive equilibrium in the model with heterogeneity, where we feed in the larger-than-optimal path of taxes, optimal in the alternative model, into the baseline model where they are suboptimal. Despite the fact that this path of carbon taxes are larger than optimal, there is still a welfare gain of implementing this tax path relative to implementing no carbon taxes at all. We find that the welfare gain in moving from the BAU allocation to the suboptimally large carbon tax path is 0.1718 percent whereas the welfare gain in moving from the BAU allocation to the socially optimal allocation is 0.2444. Hence, the suboptimal path of taxes still achieves over 70 percent of the welfare gain accomplished by the optimal path of taxes.

From this exercise, we make two conclusions. First, by ignoring heterogeneity in carbon dependence a modeler may compute optimal carbon taxes that are substantially higher than those that would be computed had heterogeneity been considered, as in the real world. Second, the welfare costs associated with implementing the too-high of carbon taxes are much lower than not implementing any tax at all. The costs of inaction seem to vastly dwarf the costs of model misspecification along this dimension.

7 Parameter Sensitivity

In this section, we evaluate the robustness of our result that the implied Pigouvian taxes from models that do not take into account firm heterogeneity are larger than optimal taxes as implied by models that do take firm heterogeneity into account, but the welfare improvement of imposing these suboptimal taxes from misspecified models, relative to not setting any carbon taxes, is large. In this section, we rigorously evaluate this finding with respect to the influence of the elasticity of substitution between brown and green capital ρ. As brown and green capital are not observed in the data, the elasticity of substitution between these forms of capital is difficult to measure empirically. In the baseline calibration, we set $\rho = 0.0$, which corresponds to Cobb-Douglas and is roughly inline with estimates of the substitution parameter between fuel sources from Stern (2012) at -0.058.

Here, we repeat our entire modeling exercise for each choice of ρ by calibrating the fixed cost of production c_f and the entrant distribution of carbon dependence Q_a to match exit rate and distribution of the emission intensities of operating firms. We evaluate the business-as-usual green transition and then compute the socially optimal transition. We then collapse heterogeneity in carbon dependence and recalibrate the model to ensure we continue to match our model's moments with the data, and compute the path of taxes under the alternative model. We then feed the path of suboptimal taxes, optimal for the alternative model without heterogeneity, into the model with heterogeneity. We compare optimal taxes and suboptimal taxes by looking at the average difference in the first 100 years and we compare the welfare implications by comparing the consumption equivalent relative to the BAU of the optimal and suboptimal transition.

In the CES production function, ρ ranges from $-\infty$, which is Leontief perfect complements, to 1, which is perfect substitutes. For expositional purposes, we evaluate transitions and make plots indexed by θ between -1 and 1, which translates ρ using the following function:

$$\rho = \begin{cases} -\infty, & \text{if } \theta = -1 \\ \frac{\theta}{\theta+1}, & \text{if } -1 < \theta < 0 \\ \theta, & \text{if } 0 \leq \theta \leq 1 \end{cases}$$

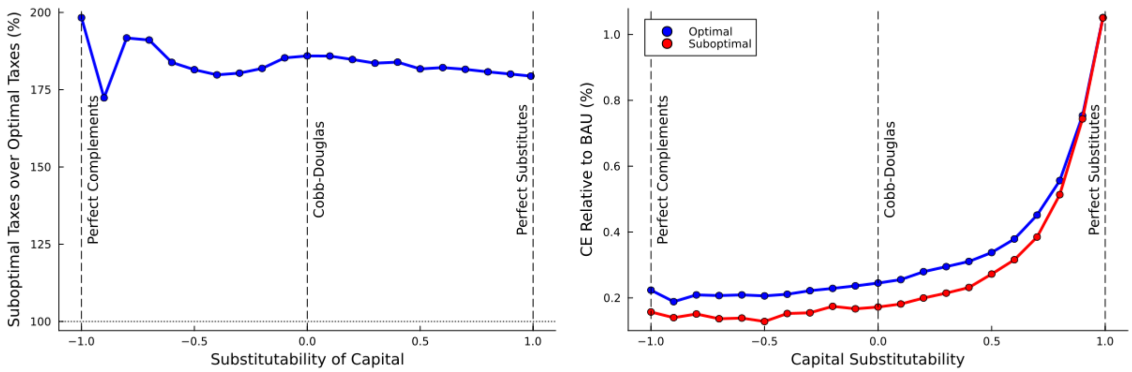
This translation is convenient because $\theta = -1$ corresponds to perfect complements and $\theta = 1$ corresponds to perfect substitutes, while preserving $\theta = 0$ as Cobb-Douglas (our main calibration) and continuity throughout. We evaluate the model along an equally-spaced grid of θ values every 0.1 increment between -1 and 1.

Figure 14a compares the average tax rate, from the models with and without heterogeneity, for the first 100 years of green transition across different values of capital substitutability. Throughout the entire range of capital substitutability, the model-implied optimal tax path without heterogeneity remains substantially larger than the model with heterogeneity. This demonstrates that the importance of the composition effect in computing optimal taxes does not depend on any particular assumption about the substitutability of capital. Across all values of the substitution parameter ρ, we find that the model without heterogeneity implies a much higher Pigouvian carbon tax.

Figure 14b shows the consumption equivalent associated with moving from the BAU to the constrained-efficient allocation and the consumption equivalent associated with moving from the BAU to the world with taxes set without considering heterogeneity in carbon dependence. As brown and green capital becomes more substitutable, the welfare improvement associated with both tax systems improves. This is intuitive, as if brown and green capital are highly substitutable in production, more damages from emissions can avoided without sacrificing as much production.

Figure 14: Parameter Sensitivity

(a) Tax Gap (b) Consumption Equivalents



Note: This figure displays robustness of our results of the impact of firm heterogeneity in carbon dependence on model-implied optimal paths of taxes, with respect to the substitutability of capital. We globally find that the optimal carbon tax from models without heterogeneity in carbon dependence is much larger than from models with heterogeneity in firms' carbon dependence. We also globally find that there is still a substantial gain in welfare from implementing these larger carbon taxes in the models with heterogeneity, suggesting that the welfare cost of model misspecification along this dimension is far less than the welfare cost of ignoring climate change altogether.

8 Conclusion

In conclusion, we document substantial persistent heterogeneity in carbon emission intensity among U.S. public firms. We introduce a general equilibrium framework to examine the implications of this heterogeneity on carbon policy. We solve for the socially optimal allocation in our model, and derive a carbon tax scheme which decentralizes the optimal allocation in a competitive equilibrium. The tax scheme that achieves this allocation varies across time but is cross-sectionally constant, despite substantial firm heterogeneity across several dimensions.

We find that the socially optimal allocation of our model features a composition effect where the distribution of production technologies utilized in the economy changes relative to the laissez-faire baseline. Specifically, the social planner exits more polluting firms than in the BAU equilibrium, and these exits contribute to a reduction in aggregate emissions. This composition effect is, by definition, absent in representative-firm economies.

Next, we use our framework to investigate the implications of ignoring heterogeneity in carbon dependence of production. Shutting down carbon dependence heterogeneity in our model, we find that the computed optimal carbon tax scheme rises by over 85 percent relative to the model with such heterogeneity.

We then quantitatively measure the welfare implications of model misspecification along this dimension. We measure this by implementing the too-large carbon tax in the version of our model with carbon dependence heterogeneity. We find that the too-large carbon tax path is still able to achieve over 70% of the welfare gains relative to laissez-faire compared to the welfare gains achieved by the optimal allocation.

Finally, we show that this result is globally robust to the substitutability of green and brown capital.

Overall, this research contributes to the existing literature by offering a comprehensive framework to analyze the interplay between firm-level carbon dependence, environmental costs, and optimal policy design. We conclude that ignoring heterogeneity of carbon dependence in computing optimal carbon taxes can result in getting the level of carbon tax wrong, but the welfare consequences of such model misspecification are far lower than the consequences of policy inaction.

References

[1] Arkolakis, C. and Walsh, C. (2023). "Clean Growth," Working Paper.

[2] Barrage, Lint, and William D. Nordhaus (2023). "Policies, Projections, and the Social Cost of Carbon: Results from the DICE-2023 Model," NBER Working Paper No. 31112.

[3] Bourany, Thomas (2024). "Optimal Energy Policy and the Inequality of Climate Change," working paper.

[4] Corbae, D. and D'Erasmo, P. (2021). "Reorganization or Liquidation: Bankruptcy Choice and Firm Dynamics," *Review of Economic Studies*. 88 (5): 2239-2274.

[5] Crane, L., Decker, R., Flaaen, A., Hamins-Puertolas, A., Kurz, C. (2022). "Business exit during the COVID-19 pandemic: Non-traditional measures in historical context," *Journal of Macroeconomics* 72: 103419.

[6] DiCiccio, T. and Efron, B. (1996). "Bootstrap Confidence Intervals," *Statistical Science*. 11 (3): 189-228.

[7] Giglio, S., Maggiori, M., Rao, K., Stroebel, J. (2021). "Climate Change and Long-Run Discount Rates: Evidence from Real Estate," *The Review of Financial Studies*, 34: pp. 3527-3571.

[8] Golosov, M., Hassler, J., Krusell, P., and Tsyvinski, A. (2014). "Optimal Taxes on Fossil Fuel in General Equilibrium," *Econometrica*. 82 (1): 41-88.

[9] Gomes, J. (2001). "Financing Investment," *The American Economic Review*, 91(5), pp.1263-1285.

[10] Hassler, J., Krusell, P., Olovsson, C., and Reiter, M. "On the effectiveness of climate policies," Working Paper.

[11] Hassler, John, Per Krusell, and Anthony A. Smith Jr. (2016). "Environmental Macroeconomics," Handbook of Macroeconomics, Volume 2B, Chapter 24.

[12] Hassler, J., Krusell, P., 2012. Economics and climate change: integrated assessment in a multi-region world. J. Eur. Econ. Assoc. 10 (5), 974–1000.

[13] Hopenhayn, H. (1992). "Entry, Exit, and Firm Dynamics in Long Run Equilibrium," *Econometrica*, 60(5), pp.1127-1150.

[14] Känzig, D. (2023). "The Unequal Economic Consequences of Carbon Pricing," NBER Working Paper 31221.

[15] Känzig, D., and Konradt, M. (2023). "Climate Policy and the Economy: Evidence from Europe's Carbon Pricing Initiatives," NBER Working Paper 31260.

[16] Khan, A., and Thomas, E. (2013). "Credit Shocks and Aggregate Fluctuations in an Economy with Production Heterogeneity," *Journal of Political Economy*. 121 (6): 1055-1107.

[17] Kim, S. (2023). "Optimal Carbon Taxes and Misallocation across Heterogeneous Firms," working paper.

[18] Kortum, S. and D. Weisbach (2021). "Optimal Unilateral Carbon Policy." Working Paper 2021-138, Becker Friedman Institute.

[19] Krusell, P., and Smith, A. (2022). "Climate Change Around the World," NBER Working Paper 30338.

[20] Lyubich, E., J. Shapiro, and R. Walker (2018). "Regulating Mismeasured Pollution: Implications of Firm Heterogeneity for Environmental Policy," AEA Papers and Proceedings 108: 136-142.

[21] Lemmon, M., Roberts, M., and Zender J. (2008). "Back to the beginning: Persistence and the cross-section of corporate capital structure," *Journal of Finance*, 63: 1575-1608.

[22] Lucas, R. and Moll, B. (2014). “Knowledge Growth and the Allocation of Time,” *Journal of Political Economy* 122(1): 1-51.

[23] Moll, B. and Nuno, G. (2018). “Social Optima in economics with Heterogeneous Agents,” *Review of Economic Dynamics* 28: 150-180.

[24] Nordhaus, W. (2007). “To tax or not to tax: The case for a carbon tax.” *Review of Environmental Economics and Policy* 1: 26-44.

[25] Nordhaus, William (1992a). “The ”Dice” Model: Background and Structure of a Dynamic Integrated Climate-Economy Model of the Economics of Global Warming,” Cowles Foundation discussion paper no. 1009).

[26] Nordhaus, William (1992b). ”An Optimal Transition Path for Controlling Greenhouse Gases”. Science. 258 (5086): 1315–1319.

[27] Ottonello, P. and Winberry, T. (2023). “Investment, Innovation, and Financial Frictions,” Working Paper.

[28] Stern, D. I. (2012): “Interfuel Substitution: A Meta-Analysis,” *Journal of Economic Surveys*, 26, pp. 307–331.

[29] Tauchen, G. (1986). “Finite State Markov-Chain Approximations to Univariate and Vector Autoregressions,” *Economics Letters*, 20(2), pp. 177-181.

[30] Weisbach, D., S. Kortum, M. Wang, Y. Yao (2022). “Trade, Leakage, and the Design of a Carbon Tax” NBER Working Paper 30244.

A Data Appendix

Table A1: Scope 1 Sample Selection

Variable	Mean (all)	SD (all)	N (all)	Mean (Scope 1)	SD (Scope 1)	N (Scope 1)	T-Statistic
EBITDA (real)	845.36	3477.96	19279	4722.73	8229.06	1965	-20.70
EBITDA/PPEGT	-1.65	135.51	19279	0.39	0.82	1965	-2.10
Employment	15.27	60.00	19279	61.93	137.52	1965	-14.90
Employment/PPEGT (real)	0.04	0.46	19279	0.01	0.03	1965	8.64
PPEGT (real)	4776.78	24032.17	19279	25724.36	55777.58	1965	-16.49

Table A2: Scope 2 Sample Selection

Variable	Mean (all)	SD (all)	N (all)	Mean (Scope 2)	SD (Scope 2)	N (Scope 2)	T-Statistic
EBITDA (real)	845.36	3477.96	19279	4891.34	8481.78	1830	-20.25
EBITDA/PPEGT	-1.65	135.51	19279	0.42	0.88	1830	-2.12
Employment	15.27	60.00	19279	64.35	141.67	1830	-14.70
Employment/PPEGT (real)	0.04	0.46	19279	0.01	0.03	1830	8.41
PPEGT (real)	4776.78	24032.17	19279	25605.26	55276.59	1830	-15.98

Table A3: Scope 1 + Scope 2 Sample Selection

Variable	Mean (all)	SD (all)	N (all)	Mean (Scope 1+2)	SD (Scope 1+2)	N (Scope 1+2)	T-Statistic
EBITDA (real)	845.36	3477.96	19279	4863.87	8426.99	1810	-20.13
EBITDA/PPEGT	-1.65	135.51	19279	0.40	0.84	1810	-2.11
Employment	15.27	60.00	19279	64.48	142.25	1810	-14.60
Employment/PPEGT (real)	0.04	0.46	19279	0.01	0.03	1810	8.51
PPEGT (real)	4776.78	24032.17	19279	25631.08	55421.38	1810	-15.87

Table A4: Sample Selection by Sector

SIC Division	Percent (all)	Percent (Scope 1)	Percent (Scope 2)	Percent (Scope 1+2)
A: Agricultural, Forestry, and Fishing	0.35	0.00	0.00	0.00
B: Mining	7.38	11.09	9.23	9.34
C: Construction	1.30	0.25	0.55	0.28
D: Manufacturing	50.19	59.08	59.07	59.72
E: Transportion, Communications, Electric, and Gas	7.75	11.20	11.48	11.33
F: Wholesale Trade	3.44	2.09	2.19	2.21
G: Retail Trade	6.44	4.83	4.92	4.97
I: Services	23.13	11.45	12.57	12.15

A.1 Contribution of Sector and Industry to Across-Firm Variation

To assess the role of sector-level and industry-level differences in driving emissions intensity, we run the following fixed effect regression separately for each sector and industry

$$y_{i,t} = \beta^c + \beta_{j(i)} + r_{i,t}, \tag{10}$$

where $\beta_{j(i)}$ is the fixed effect for the SIC Division (to capture sector-level variation) or the 2-Digit SIC Major Group (to capture industry-level variation) that firm i is in. By assuming $\beta_{j(i)}$ is independent of $r_{i,t}$, we can decompose the variance of the data:

$$\underbrace{\mathbb{V}\mathrm{ar}(y_{i,t})}_{\text{total}} = \underbrace{\mathbb{V}\mathrm{ar}(\beta_{j(i)})}_{\text{across-sector (or industry)}} + \underbrace{\mathbb{V}\mathrm{ar}(\varepsilon_{i,t})}_{\text{within-sector (or industry)}} .$$

Note that the total R^2 in this case is equal to the ratio of the across-sector (or across-industry) variation over the total variation. Hence, it is a convenient and easy to interpret R^2 as the fraction of variation driven by sector (or industry). We report the R^2 of Regression (10) in Table A5.

Table A5: Sector and Industry Fixed Effects

	Sector			Industry	
Scope	R^2	95% CI	Scope	R^2	95% CI
Scope 1	0.160	(0.094449,0.22752)	Scope 1	0.657	(0.5847,0.69021)
Scope 2	0.137	(0.065866,0.21516)	Scope 2	0.383	(0.26429,0.46955)
Scope 1 & 2	0.070	(0.023326,0.11401)	Scope 1 & 2	0.537	(0.45185,0.58812)

Notes: This table reports R^2 values from estimating Regression (10) where firm-level emission intensity is regressed on sector (or industry) fixed effects. Sectors are defined as SIC Divisions and industries are defined as 2-Digit SIC Major Groups. Confidence intervals are 95% block bootstrapped bias-corrected and accelerated intervals as in DiCiccio and Efron (1996).

Table A5 shows that sector has very little power in terms of explaining the variation in the emissions intensity data but industry has explanatory power, explaining between one-third and two-thirds of the variation depending on the emission scope. As we show in the main text, the overwhelming majority of the emissions intensity variation not accounted for by sector and industry is across-firm. As our model does not feature multiple industries and industry classification is a way that firms persistently vary from one another, we calibrate our model to match the estimates of the distribution of permanent firm emission intensities $\hat{\alpha}^c + \hat{\alpha}_i$ from equation (1), and consider sector/industry to be one source of the variation in $\hat{\alpha}_i$ that we observe in the data.

A.2 Empirical Robustness - Log Levels of Emissions

Table A6: R^2 Values from Regression (1)

Scope	R^2	95% CI
Scope 1	0.986	(0.97836,0.99191)
Scope 2	0.983	(0.97961,0.98794)
Scope 1 & 2	0.989	(0.98457,0.99227)

Notes: This table reports the R^2 values from OLS estimates of Regression (1) [with LHS in log levels, not log ratios as in the baseline], along with 95% confidence intervals. Confidence intervals are 95% block bootstrapped bias-corrected and accelerated intervals as in DiCiccio and Efron (1996).

Table A7: Firm Fixed Effects Within Sector and Industry - Levels

	Sector			Industry	
Scope	R^2	95% CI	Scope	R^2	95% CI
Scope 1	0.983	(0.97346,0.99027)	Scope 1	0.965	(0.94792,0.98029)
Scope 2	0.982	(0.97773,0.98725)	Scope 2	0.975	(0.96921,0.98298)
Scope 1 & 2	0.987	(0.98205,0.99147)	Scope 1 & 2	0.977	(0.96909,0.98541)

Notes: This table reports R^2 from Regression (2) where the residual from Regression (10) [LHS in log levels, not log ratio as in the baseline] is regressed on firm-level fixed effects. In Regression (10), firm-level emissions are regressed on sector (or industry) fixed effects for each sector (or industry) separately. Sectors are defined as SIC Divisions and industries are defined as 2-Digit SIC Major Groups.Confidence intervals are 95% block bootstrapped bias-corrected and accelerated intervals as in DiCiccio and Efron (1996).

A.3 Empirical Robustness - Alternative Emission Intensity $\log\left(\frac{\textbf{GHG}}{\textbf{Output}}\right)$

Table A8: R^2 Values from Regression (1)

Scope	R^2	95% CI
Scope 1	0.983	(0.97680,0.98905)
Scope 2	0.970	(0.96216,0.97741)
Scope 1 & 2	0.982	(0.97742,0.98728)

Notes: This table reports the R^2 values from OLS estimates of Regression (1) [with LHS descaled by employment, not capital], along with 95% confidence intervals. Confidence intervals are 95% block bootstrapped bias-corrected and accelerated intervals as in DiCiccio and Efron (1996).

Table A9: Firm Fixed Effects Within Sector and Industry - Descaled by Employment

	Sector			Industry	
Scope	R^2	95% CI	Scope	R^2	95% CI
Scope 1	0.977	(0.96717,0.98520)	Scope 1	0.943	(0.92080,0.96566)
Scope 2	0.965	(0.95564,0.97392)	Scope 2	0.944	(0.92678,0.95987)
Scope 1 & 2	0.977	(0.96939,0.98369)	Scope 1 & 2	0.950	(0.93619,0.96730)

Notes: This table reports R^2 from Regression 2 where the residual from Regression 10 [descaled by output, not capital as in the baseline] is regressed on firm-level fixed effects. In Regression 10, firm-level emission intensities are regressed on sector (or industry) fixed effects. Sectors are defined as SIC Divisions and industries are defined as 2-Digit SIC Major Groups. Confidence intervals are 95% block bootstrapped bias-corrected and accelerated intervals as in DiCiccio and Efron (1996).

A.4 Empirical Robustness - Adjusted R^2

Table A10: Adjusted R^2 Values from Regression (1)

Scope	adjusted R^2	95% CI
Scope 1	0.962	(0.94691,0.97462)
Scope 2	0.943	(0.92348,0.95738)
Scope 1 & 2	0.945	(0.92826,0.95919)

Notes: This table reports the adjusted R^2 values from OLS estimates of Regression (1), along with 95% confidence intervals. Confidence intervals are 95% block bootstrapped bias-corrected and accelerated intervals as in DiCiccio and Efron (1996).

Table A11: Firm Fixed Effects Within Sector and Industry

Scope	Sector adjusted R^2	95% CI	Scope	Industry adjusted R^2	95% CI
Scope 1	0.955	(0.93516,0.97015)	Scope 1	0.889	(0.84839,0.92936)
Scope 2	0.934	(0.91239,0.9509)	Scope 2	0.907	(0.87463,0.93299)
Scope 1 & 2	0.941	(0.92432,0.95734)	Scope 1 & 2	0.882	(0.85274,0.9137)

Notes: This table reports the adjusted R^2 from Regression (2) where the residual from Regression (10) is regressed on firm-level fixed effects. In Regression (10), firm-level emission intensities are regressed on sector (or industry) fixed effects for each sector (or industry) separately. Sectors are defined as SIC Divisions and industries are defined as 2-Digit SIC Major Groups. Confidence intervals are 95% block bootstrapped bias-corrected and accelerated intervals as in DiCiccio and Efron (1996).

A.5 Empirical Robustness - 10 Year Balanced Panel

Table A12: R^2 Values from Regression (1)

Scope	R^2	95% CI
Scope 1	0.961	(0.95,0.97135)
Scope 2	0.924	(0.8766,0.95043)
Scope 1 & 2	0.929	(0.89104,0.95683)

Notes: This table reports the R^2 values from OLS estimates of Regression (1), along with 95% confidence intervals. Uses the 10-year balanced panel, as opposed to the 5 year balanced panel in the baseline. Confidence intervals are 95% block bootstrapped bias-corrected and accelerated intervals as in DiCiccio and Efron (1996).

Table A13: Firm Fixed Effects Within Sector and Industry

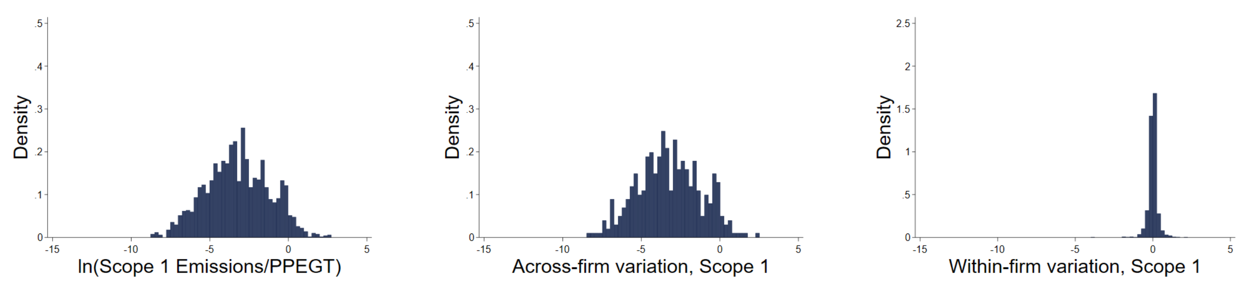
	Sector			Industry	
Scope	R^2	95% CI	Scope	R^2	95% CI
Scope 1	0.953	(0.93806,0.96682)	Scope 1	0.870	(0.83708,0.91376)
Scope 2	0.901	(0.84835,0.93624)	Scope 2	0.843	(0.78647,0.90488)
Scope 1 & 2	0.922	(0.88341,0.95386)	Scope 1 & 2	0.826	(0.75598,0.90237)

Notes: This table reports the R^2 from Regression (2) where the residual from Regression (10) is regressed on firm-level fixed effects. Uses 10 year balanced panel, vs 5 year balanced panel in baseline. In Regression (10), firm-level emission intensities are regressed on sector (or industry) fixed effects for each sector (or industry) separately. Sectors are defined as SIC Divisions and industries are defined as 2-Digit SIC Major Groups. Confidence intervals are 95% block bootstrapped bias-corrected and accelerated intervals as in DiCiccio and Efron (1996).

A.6 Scope 1, including Utilities

In the baseline text, we exclude utilities so as to avoid double counting when we use Scope 1 + 2 as our measure of firm-level emissions. This effectively distributes the emissions generated by utility companies as Scope 1 emissions to the firms that use that energy in their production. In this section we provide robustness of our results to another method of avoiding double-counting, by including utilities back into our sample but then only focusing on Scope 1 emissions.

Figure A1: Distribution of Carbon Emission Intensity, Including Utilities



(a) Emissions Intensity (b) Across-Firm Variation (c) Within-Firm Variation

Notes: This figure replicates Figure 2 and Figure 3 from Section 2, but for Scope 1 emissions using a firm sample that includes utilities. We find extremely similar results to our baseline.

Table A14: R^2 Values from Regression (1)

Scope	R^2	95% CI
Regression (1)	0.977	(0.969,0.983)
Within-Sector	0.973	(0.963,0.980)
Within-Industry	0.931	(0.912,0.954)

Notes: This table reports the R^2 values from OLS estimates of Regression (1), along with 95% confidence intervals, as well as Regression (10). Uses Scope 1 emissions, with a panel that includes utilities. Confidence intervals are 95% block bootstrapped bias-corrected and accelerated intervals as in DiCiccio and Efron (1996).

A.7 Added Controls

In this appendix, we add controls to our baseline regressions to determine the degree to which within-firm variation is explainable by firm-level controls. We find very small changes to the R^2 values of these regressions. We conclude that firm-level controls do very little to explain the within-firm variation in the emission intensity data. This motivates the structure of emissions generation in our model. In our model, emissions are generated by dirty capital which is fixed in the short-run, and hence the emissions/capital ratio does not correlate with productivity shocks.

Table A15: Explanatory Power and Firm-Level Controls

	Scope 1		
	R^2(controls)	R^2(no controls)	R^2(controls) - R^2(no controls)
emp/capital	0.9706	0.96955	0.0010475
	(0.95782,0.98051)	(0.95753,0.9797)	(0.00026975,0.0027009)
ebitda/capital	0.97	0.96955	0.00045343
	(0.95752,0.9798)	(0.95753,0.9797)	(8.8674e-05,0.0014472)
sales/capital	0.97117	0.96955	0.0016252
	(0.95865,0.98099)	(0.95753,0.9797)	(0.00069716,0.0029637)
all	0.97137	0.96955	0.001823
	(0.95882,0.98116)	(0.95753,0.9797)	(0.00072343,0.0031675)
	Scope 2		
	R^2(controls)	R^2(no controls)	R^2(controls) - R^2(no controls)
emp/capital	0.95851	0.95418	0.0043323
	(0.94361,0.96924)	(0.93879,0.96591)	(0.0017869,0.009948)
ebitda/capital	0.95544	0.95418	0.0012609
	(0.9397,0.96723)	(0.93879,0.96591)	(0.00017504,0.0034894)
sales/capital	0.95711	0.95418	0.0029333
	(0.94141,0.96832)	(0.93879,0.96591)	(0.0019204,0.0051788)
all	0.95974	0.95418	0.005558
	(0.94421,0.97002)	(0.93879,0.96591)	(0.0025677,0.008772)
	Scope 1 & 2		
	R^2(controls)	R^2(no controls)	R^2(controls) - R^2(no controls)
emp/capital	0.96016	0.95617	0.0039895
	(0.9463,0.97071)	(0.94261,0.96736)	(0.0013596,0.0089023)
ebitda/capital	0.95698	0.95617	0.00080703
	(0.94347,0.96805)	(0.94261,0.96736)	(7.9374e-05,0.003049)
sales/capital	0.961	0.95617	0.0048307
	(0.94831,0.97165)	(0.94261,0.96736)	(0.0021603,0.0082169)
all	0.96208	0.95617	0.0059087
	(0.94838,0.97216)	(0.94261,0.96736)	(0.0023974,0.009394)

Notes: This table reports the difference in R^2 with and without controls for employment/capital, ebitda/capital, and sales/capital. The added explanatory power by adding these controls is extremely small relative to the across-firm variation (R^2(no controls)) and unexplained variation ($1 - R^2$(controls)). Confidence intervals are 95% block bootstrapped bias-corrected and accelerated intervals as in DiCiccio and Efron (1996).

B Characterization of Planner's Solution

In this appendix, we characterize the planner's solution.[15] We start by characterizing the full social planner's problem as well as its Lagrangian. Then we derive the marginal social value of firms at particular location in the state-space. We derive optimally conditions from the full planner's problem including first-order conditions for continuous choices, like labor and investment, and inequalities (or equivalently thresholds) for discrete choices, like exit. Then we show that marginal social value is equal to the value function induced by an augmented firm Bellman equation written similar to the marginal social value.

B.0.1 Planner's Problem

Let (S^1, S^2) be last period's atmospheric carbon stock and μ the measure of incumbent firms continuing from the prior period. The constrained planner's objective is

$$\begin{aligned}\mathcal{W}_t(\mu, S^1, S^2) = &\max_{x_b(\cdot),x_g(\cdot),X(\cdot),x_b^E(\cdot),x_g^E(\cdot),X^E(\cdot),L^d(\cdot),B} U(C_t)\\ &+ \beta\mathcal{W}_{t+1}(T^*(\mu, x_b(\cdot), x_g(\cdot), X(\cdot), x_b^E(\cdot), x_g^E(\cdot), X^E(\cdot), B), S^{1\prime}, S^{2\prime})\end{aligned}$$

where $(S^{1\prime}, S^{2\prime})$ follow environmental laws of motion

$$\begin{aligned}S^{1\prime} &= S^1 + (1-\chi_t)\varphi_1\gamma K^b\\ S^{2\prime} &= \varphi_3 S^2 + (1-\chi_t)(1-\varphi_1)\varphi_2\gamma K^b,\end{aligned}$$

K^b is the aggregate capital of operating firms

$$K^b = \int k_b \Phi(\mu, x_b^E(\cdot), x_g^E(\cdot), X^E(\cdot), B)(s)ds,$$

Φ is the measure of operating firms

$$\begin{aligned}\Phi(\mu, x_b^E(\cdot), x_g^E(\cdot), X^E(\cdot), B)(s) = &\underbrace{\mu(s)}_{\text{incumbents}}\\ &+ \underbrace{\int \mathbb{1}_{k_b=x_b^E(a,q)}\mathbb{1}_{k_g=x_g^E(a,q)}[1-X^E(a,q)]Q_a(a)Q_q(q)Bp_q(z-\rho_q q)dq}_{\text{entrants}},\end{aligned} \tag{11}$$

[15] We use a similar approach to Ottonello and Winberry (2023). The major departure is firms in our framework exit both endogenously and exogenously while all exits in Ottonello and Winberry (2023) are exogenous.

and T^* operator defines how the measure of firms evolves

$$
\begin{aligned}
T^*(\mu, x_b(\cdot), x_g(\cdot), & X(\cdot), x_b^E(\cdot), x_g^E(\cdot), X^E(\cdot), B)(s') \\
&= \int \mathbb{1}_{k_b'=(1-\delta)k_b+x_b(s)} \mathbb{1}_{k_g'=(1-\delta)k_g+x_g(s)} Q_z(z'|z) \mathbb{1}_{a'=a} \\
&\quad \times [1-X(s)](1-\lambda) d\Phi(\mu, x_b^E(\cdot), x_g^E(\cdot), X^E(\cdot), B)(s). \qquad (12)
\end{aligned}
$$

The planner's problem is subject to three constraints. First, the planner's problem is subject to the resource constraint of the economy:

$$
C_t = (1 - D(S^{1\prime} + S^{2\prime}))Y_t - I_t - \Psi_t
$$

where aggregate output (net of environmental damage) Y_t, aggregate investment I_t, and aggregate capital adjustment costs Ψ_t are

$$
\begin{aligned}
Y_t &= \int \exp(z) A_t^{1-\alpha} [a k_b^\rho + (1-a) k_g^\rho]^{\frac{\alpha}{\rho}} (L^d(s))^\nu \Phi(\mu, x_b^E(\cdot), x_g^E(\cdot), X^E(\cdot), B)(s) ds \\
I_t &= \underbrace{\int (x_b(z, k_b, k_g, a) + x_g(z, k_b, k_g, a))(1-\lambda)\Phi(\mu, x_b^E(\cdot), x_g^E(\cdot), X^E(\cdot), B)(s) ds}_{\text{(dis)investment for operating firms}} \\
&\quad + \underbrace{\int (-(1-\delta)k_b - (1-\delta)k_g)\lambda\Phi(\mu, x_b^E(\cdot), x_g^E(\cdot), X^E(\cdot), B)(s) ds}_{\text{disinvestment from exogenously exiting firms}} \\
&\quad + \underbrace{\int (x_b^E(a) + x_g^E(a)) \mathbb{1}_{X^E(a)=0} B Q_a(a) da}_{\text{investment for entrants}} \\
\Psi_t &= \underbrace{\int (\psi[x_g(z, k_b, k_g, a), k_g] + \psi[x_b(z, k_b, k_g, a), k_b])(1-\lambda)\Phi(\mu, x_b^E(\cdot), x_g^E(\cdot), X^E(\cdot), B)(s) ds}_{\text{capital adjustment costs from firms w/o exogenous exit shock}} \\
&\quad + \underbrace{\int [\psi^X(k_b) + \psi^X(k_g)]\lambda\Phi(\mu, x_b^E(\cdot), x_g^E(\cdot), X^E(\cdot), B)(s) ds}_{\text{capital adjustment costs from exogenously exiting firms}}.
\end{aligned}
$$

Plugging in these definitions, we can rewrite the resource constraint as

$$
\begin{aligned}
C_t = (1 - D(S^{1\prime} + S^{2\prime})) &\int \exp(z) A_t^{1-\alpha} [a^{1-\rho} k_b^\rho + (1-a)^{1-\rho} k_g^\rho]^{\frac{\alpha}{\rho}} (L^d(s))^\nu \Phi(\mu, x_b^E(\cdot), x_g^E(\cdot), X^E(\cdot), B)(s) ds \\
&- \int (x_b(z, k_b, k_g, a) + x_g(z, k_b, k_g, a))(1-\lambda)\Phi(\mu, x_b^E(\cdot), x_g^E(\cdot), X^E(\cdot), B)(s) ds \\
&- \int (-(1-\delta)k_b - (1-\delta)k_g)\lambda\Phi(\mu, x_b^E(\cdot), x_g^E(\cdot), X^E(\cdot), B)(s) ds
\end{aligned}
$$

$$
\begin{aligned}
&- \int (x_b^E(a) + x_g^E(a)) \mathbb{1}_{X^E(a)=0} B Q_a(a) da \\
&- \int (\psi[x_b(z,k_b,k_g,a),k_b] + \psi[x_g(z,k_b,k_g,a),k_g])(1-\lambda)\Phi(\mu, x_b^E(\cdot), x_g^E(\cdot), X^E(\cdot), B)(s) ds \\
&- \int [\psi^X(k_b) + \psi^X(k_g)] \lambda \Phi(\mu, x_b^E(\cdot), x_g^E(\cdot), X^E(\cdot), B)(s) ds \qquad (13)
\end{aligned}
$$

Second, the planner's problem is subject to the labor supply constraint:

$$
1 = \underbrace{\int (L^d(s) + c_f)\Phi(\mu, x_b^E(\cdot), x_g^E(\cdot), X^E(\cdot), B)(s) ds}_{\text{labor used by operating firms}} + \underbrace{B\kappa}_{\text{labor to cover entry costs}}
$$

Third, the planner must choose a nonnegative mass of entrants $B \geq 0$.

B.0.2 Lagrangian of Planner's Problem

We can write the planner's Lagrangian as

$$
\begin{aligned}
\mathcal{L}_t = & \, U(C_t) + \lambda_t^1[S^{1\prime} - S^1 - (1-\chi_t)\varphi_1\gamma K^b)] + \lambda_t^2[S^{2\prime} - \varphi_3 S^2 - (1-\chi_t)(1-\varphi_1)\varphi_2\gamma K^b] + \lambda_t^B B \\
& + \lambda_t^L \left[1 - \int (L^d(s) + c_f)\Phi(\mu, x_b^E(\cdot), x_g^E(\cdot), X^E(\cdot), B)(s) ds - B\kappa \right] \\
& + \lambda_t^k \left[K^b - \int k_b \Phi(\mu, x_b^E(\cdot), x_g^E(\cdot), X^E(\cdot), B)(s) ds \right] \\
& + \beta \mathcal{W}_{t+1}(T^*(\mu, x_b(\cdot), x_g(\cdot), X(\cdot), x_b^E(\cdot), x_g^E(\cdot), X^E(\cdot), B), S^{1\prime}, S^{2\prime})
\end{aligned}
$$

where Φ, T^*, and C_t are defined by (11), (12), and (13), respectively.

B.1 Marginal Social Value

We start by computing the Gateaux derivative of the social welfare function with respect to the mass of firms at a particular location $(z, k_b, k_g, a) = s$ in the state-space at time t, which we follow Moll and Nuno (2018) in assuming exists:

$$
\begin{aligned}
\frac{\partial \mathcal{W}_t(\mu, S^1, S^2)}{\partial \mu(s)} = & \, U'(C_t) \Big[(1 - D(S^{1\prime} + S^{2\prime})) \exp(z) A_t^{1-\alpha} [a^{1-\rho} k_b^\rho + (1-a)^{1-\rho} k_g^\rho]^{\frac{\alpha}{\rho}} (L^d(s))^\nu \\
& + (1-\lambda)\left(-x_b(s) - x_g(s) - \psi[x_b(s), k_b] - \psi[x_g(s), k_g]\right) \\
& + \lambda \left((1-\delta)k_b + (1-\delta)k_g - \psi^X(k_b) - \psi^X(k_g) \right) \Big] \\
& - \lambda_t^L (L^d(s) + c_f) - \lambda_t^k k_b
\end{aligned}
$$

$$
\begin{aligned}
&+ (1 - X(s))\beta \int \frac{\partial \mathcal{W}_{t+1}(\mu', S^{1\prime}, S^{2\prime})}{\partial \mu'(s')} \frac{\partial T^*(s')}{\partial \mu(s)} ds' \\
&= U'(C_t)\Big[(1 - D(S^{1\prime} + S^{2\prime})) \exp(z) A_t^{1-\alpha} [a^{1-\rho} k_b^\rho + (1-a)^{1-\rho} k_g^\rho]^{\frac{\alpha}{\rho}} (L^d(s))^\nu \\
&\qquad + (1 - \lambda) \left(-x_b(s) - x_g(s) - \psi[x_b(s), k_b] - \psi[x_g(s), k_g]\right) \\
&\qquad + \lambda \left((1-\delta)k_b + (1-\delta)k_g - \psi^X(k_b) - \psi^X(k_g)\right)\Big] \\
&- \lambda_t^L (L^d(s) + c_f) - \lambda_t^k k_b \\
&+ (1 - X(s))(1 - \lambda)\beta \int \frac{\partial \mathcal{W}_{t+1}(\mu', S^{1*}, S^{2*})}{\partial \mu'(\rho z + \epsilon, k_b', k_g', a)} p(\epsilon) d\epsilon
\end{aligned} \tag{14}
$$

B.2 Planner's Optimality Conditions

B.2.1 Environmental Aggregate Variables

The planner's first-order conditions with respect to tomorrow's carbon stock $S^{1\prime}$ and $S^{2\prime}$

$$
\begin{aligned}
\frac{\partial \mathcal{L}_t}{\partial S^{1\prime}} = 0 &\implies \lambda_t^1 = U'(C_t) D'(S^{1\prime} + S^{2\prime}) Y_t - \beta \frac{\partial}{\partial S^{1\prime}} \mathcal{W}_{t+1}(T^*(\mu, \cdot), S^{1\prime}, S^{2\prime}) \\
\frac{\partial \mathcal{L}_t}{\partial S^{2\prime}} = 0 &\implies \lambda_t^2 = U'(C_t) D'(S^{1\prime} + S^{2\prime}) Y_t - \beta \frac{\partial}{\partial S^{2\prime}} \mathcal{W}_{t+1}(T^*(\mu, \cdot), S^{1\prime}, S^{2\prime}).
\end{aligned}
$$

Taking the envelope conditions with respect to S^1 and S^2

$$
\begin{aligned}
\frac{\partial}{\partial S^1} \mathcal{W}_t(\mu, S^1, S^2) &= -\lambda_{t+1}^1 \\
\frac{\partial}{\partial S^2} \mathcal{W}_t(\mu, S^1, S^2) &= -\varphi_3 \lambda_{t+1}^2
\end{aligned}
$$

and iterating forward, we can write λ_t^1 and λ_t^2 as

$$
\begin{aligned}
\lambda_t^1 &= \sum_{s=t}^{\infty} \beta^{s-t} U'(C_s) D'(S_s^1 + S_s^2) Y_s \\
\lambda_t^2 &= \sum_{s=t}^{\infty} (\varphi_3 \beta)^{s-t} U'(C_s) D'(S_s^1 + S_s^2) Y_s
\end{aligned}
$$

assuming appropriate transversality condition.[16] The planner's first-order condition with respect to aggregate brown capital utilized by operating firms is

$$\frac{\partial \mathcal{L}_t}{\partial K^b} = 0 \implies \lambda_t^k = \lambda_t^1 (1-\chi_t)\varphi_1 \gamma + \lambda_t^2 (1-\chi_t)(1-\varphi_1)\varphi_2 \gamma.$$

This proves Proposition 2 in the main text.

B.2.2 Operating Firm Labor Allocation

Next, the planner's first-order condition with respect to firm-level labor allocation $L^d(s)$ at particular location in the state-space $(z, k_b, k_g, a) = s$ is

$$\begin{aligned}
&\frac{\partial \mathcal{L}_t}{\partial L^d(s)} = 0 \\
\implies &\lambda_t^L \Phi(\cdot)(s) = \Phi(\cdot)(s) U'(C_t)(1 - D(S^{1\prime} + S^{2\prime}))\nu \exp(z) A_t^{1-\alpha} [a^{1-\rho} k_b^\rho + (1-a)^{1-\rho} k_g^\rho]^{\frac{\alpha}{\rho}} (L^d(s))^{\nu-1} \\
\implies &\frac{\lambda_t^L}{U'(C_t)} = (1 - D(S^{1\prime} + S^{2\prime}))\nu \exp(z) A_t^{1-\alpha} [a^{1-\rho} k_b^\rho + (1-a)^{1-\rho} k_g^\rho]^{\frac{\alpha}{\rho}} (L^d(s))^{\nu-1}. \quad (15)
\end{aligned}$$

B.2.3 Incumbent Firm Investment

Next, we turn to the planner's first-order condition with respect to firm-level investment policies $x_b(s), x_g(s)$ at a particular location $(z, k_b, k_g, a) = s$ conditional on not exiting. This first-order condition for firm-level investment in green capital is

$$\begin{aligned}
&\frac{\partial \mathcal{L}_t}{\partial x_g(s)} = 0 \\
\implies &U'(C_t)(1-\lambda)(1 + \psi'[x_g(s), k_g])\Phi(\mu, \cdot)(s) \\
&\qquad = \beta \int \frac{\partial W(\mu', S^{1*}, S^{2*})}{\partial \mu'(s')} \frac{\partial T^*(\mu, \cdot)(s')}{\partial k_g'} \underbrace{\frac{\partial k_g'(s)}{\partial x_g(s)}}_{=1} ds' \\
&\qquad = \beta \int \frac{\partial}{\partial k_g'(s)} \frac{\partial W(\mu', S^{1*}, S^{2*})}{\partial \mu'(s')} T^*(\mu, \cdot)(s') ds' \\
&\qquad = \beta(1-\lambda)\Phi(\mu, \cdot)(s) \\
&\qquad \times \int \frac{\partial}{\partial k_g'(s)} \frac{\partial W(\mu', S^{1*}, S^{2*})}{\partial \mu'(\rho z + \epsilon, k_b'(z), k_g'(s), a)} p(\epsilon) d\epsilon \\
\implies &U'(C_t)(1 + \psi'[x_g(s), k_g]) = \beta \frac{\partial}{\partial k_g'(s)} \int \frac{\partial W(\mu', S^{1*}, S^{2*})}{\partial \mu'(\rho z + \epsilon, k_b'(z), k_g'(s), a)} p(\epsilon) d\epsilon \quad (16)
\end{aligned}$$

[16] We can explicitly assume $\lim_{t\to\infty} \beta^t U'(C_t) D'(S_t^1 + S_t^2) Y_t = 0$. Because $\varphi_3 < 1$, this condition implies $\lim_{t\to\infty} (\varphi_3 \beta)^t U'(C_t) D'(S_t^1 + S_t^2) Y_t = 0$ by the squeeze theorem.

assuming the order of integration can be swapped. Similarly, the first-order condition with respect to firm-level investment in brown capital can be written as

$$\frac{\partial \mathcal{L}_t}{\partial x_b(s)} = 0$$
$$\implies U'(C_t)(1 + \psi'[x_b(s), k_b]) = \beta \frac{\partial}{\partial k_b'(s)} \int \frac{\partial W(\mu', S^{1*}, S^{2*})}{\partial \mu'(\rho z + \epsilon, k_b'(z), k_g'(s), a)} p(\epsilon) d\epsilon \tag{17}$$

B.2.4 Incumbent Firm Continuation

The planner strictly prefers not to exit the firm in state s if the marginal welfare benefit of firm s continuing is greater than the marginal welfare benefit of firm s exiting,

$$\underbrace{U'(C_t)\Big[(1-\delta)[k_b + k_g] - \psi^X(k_b) - \psi^X(k_g)\Big]}_{\text{exit}} \tag{18}$$
$$< \underbrace{U'(C_t)\Big[- x_b(s) - \psi[x_b(s), k_b] - x_g(s) - \psi[x_g(s), k_g]\Big] + \beta \int \frac{\partial \mathcal{W}_{t+1}(\mu', S^{1\prime}, S^{2\prime})}{\partial \mu'(\rho z + \epsilon, k_b'(s), k_g'(s), a)} p(\epsilon) d\epsilon}_{\text{operate}},$$

and in this case the planner chooses for the firm not to exit. Relatedly, if the exiting welfare benefit strictly exceeds the operating welfare benefit then the planner will exit the firm. In general, it is also possible for the planner to be indifferent. In our case, this will hold for a measure zero of firms, due to the absolute continuity of the distribution over a in our state space. However, more generally there may be a point mass at s for which the planner optimally chooses for some positive measure to exit and the rest of the mass to not exit. If this is the case, the optimality condition of the planner's choice of measure of firms within s to exit is indifference between the exit and operating value of the firms in s.

B.2.5 Mass of Entrants

Next, the planner's first-order condition with respect to the mass of entrants B is

$$\frac{\partial \mathcal{L}_t}{\partial B} = 0$$
$$\implies U'(C_t)\frac{\partial C_t}{\partial B} + \lambda_t^B + \beta \frac{\partial \mathcal{W}_{t+1}(\cdot)}{\partial B} = \lambda_t^L \left(\int (L^d(s) + c_f) \frac{\partial \Phi(\cdot)(s)}{\partial B} ds + \kappa \right)$$
$$+ \lambda_t^k \int k_b \frac{\partial \Phi(\cdot)(s)}{\partial B} ds. \tag{19}$$

We evaluate derivatives of the measure of operating firms Φ, consumption C_t, and welfare tomorrow $\mathcal{W}_{t+1}$ with respect to the mass of entrants B, separately. First, the derivative of the measure of operating firms with respect to B is

$$\frac{\partial \Phi(\cdot)(s)}{\partial B} = \int \mathbb{1}_{k_b = x_b^E(a,q)} \mathbb{1}_{k_g = x_g^E(a,q)} [1 - X^E(a,q)] Q_a(a) Q_q(q) p_q(z - \rho_q q) dq. \tag{20}$$

Second, using (13), the derivative of consumption with respect to B is

$$\begin{aligned}
\frac{\partial C_t}{\partial B} &= (1 - D(S^{1\prime} + S^{2\prime})) \int \exp(z) A_t^{1-\alpha} [a^{1-\rho} k_b^\rho + (1-a)^{1-\rho} k_g^\rho]^{\frac{\alpha}{\rho}} (L^d(s))^\nu \frac{\partial \Phi(\cdot)(s)}{\partial B} ds \qquad (21) \\
&- \int (x_b(z, k_b, k_g, a) + x_g(z, k_b, k_g, a))(1 - \lambda) \frac{\partial \Phi(\cdot)(s)}{\partial B} ds \\
&- \int (-(1-\delta) k_b - (1-\delta) k_g) \lambda \frac{\partial \Phi(\cdot)(s)}{\partial B} ds \\
&- \int (x_b^E(a) + x_g^E(a)) \mathbb{1}_{X^E(a)=0} Q_a(a) da \\
&- \int (\psi[x_b(z, k_b, k_g, a), k_b] + \psi[x_g(z, k_b, k_g, a), k_g])(1 - \lambda) \frac{\partial \Phi(\cdot)(s)}{\partial B} ds \\
&- \int (\psi[-(1-\delta) k_b, k_b] + \psi[-(1-\delta) k_g, k_g]) \lambda \frac{\partial \Phi(\cdot)(s)}{\partial B} ds
\end{aligned}$$

where (20) defines $\frac{\partial \Phi(\cdot)(s)}{\partial B}$. Finally, the derivative of welfare tomorrow with respect to B is

$$\begin{aligned}
\frac{\partial \mathcal{W}_{t+1}(\mu', \cdot)}{\partial B} &= \int \frac{\partial \mathcal{W}_{t+1}(\mu', \cdot)}{\partial \mu'(s')} \frac{\partial \mu'(s')}{\partial B} ds' \\
&= \int \frac{\partial \mathcal{W}_{t+1}(\mu', \cdot)}{\partial \mu'(s')} \frac{\partial T^*(\mu, \cdot)(s')}{\partial B} ds' \\
&= \int \frac{\partial \mathcal{W}_{t+1}(\mu', \cdot)}{\partial \mu'(s')} \Bigg[\int \mathbb{1}_{k_b' = (1-\delta)k_b(s) + x_b(s)} \mathbb{1}_{k_g' = (1-\delta)k_g(s) + x_g(s)} \mathbb{1}_{z' = \rho z + \epsilon} \mathbb{1}_{a' = a} \mathbb{1}_{X(s)=0} p(\epsilon) \\
&\qquad\qquad \times (1 - \lambda) \frac{\partial \Phi(\cdot)(s)}{\partial B} ds \Bigg] ds' \\
&= \int \int \frac{\partial \mathcal{W}_{t+1}(\mu', \cdot)}{\partial \mu'(s')} \mathbb{1}_{k_b' = (1-\delta)k_b(s) + x_b(s)} \mathbb{1}_{k_g' = (1-\delta)k_g(s) + x_g(s)} \mathbb{1}_{z' = \rho z + \epsilon} \mathbb{1}_{a' = a} \mathbb{1}_{X(s)=0} p(\epsilon) \\
&\qquad \times (1 - \lambda) \frac{\partial \Phi(\cdot)(s)}{\partial B} ds' ds \\
&= \int (1 - \lambda) \mathbb{1}_{X(s)=0} \frac{\partial \Phi(\cdot)(s)}{\partial B} \int \frac{\partial \mathcal{W}_{t+1}(\mu', \cdot)}{\partial \mu'(s')} \mathbb{1}_{k_b' = (1-\delta)k_b(s) + x_b(s)} \mathbb{1}_{k_g' = (1-\delta)k_g(s) + x_g(s)} \\
&\qquad\qquad \times \mathbb{1}_{z' = \rho z + \epsilon} \mathbb{1}_{a' = a} p(\epsilon) ds' ds \\
&= \int (1 - \lambda) \mathbb{1}_{X(s)=0} \frac{\partial \Phi(\cdot)(s)}{\partial B}
\end{aligned}$$

$$\times \left(\int \frac{\partial \mathcal{W}_{t+1}(\mu', \cdot)}{\partial \mu'(\rho z + \epsilon, (1-\delta)k_b + x_b(s), (1-\delta)k_g + x_g(s), a)} p(\epsilon) d\epsilon \right) ds$$

$$= \int (1-\lambda) \mathbb{1}_{X(s)=0} \frac{\partial \Phi(\cdot)(s)}{\partial B} \left(\int \frac{\partial \mathcal{W}_{t+1}(\mu', \cdot)}{\partial \mu'(\rho z + \epsilon, k_d'(s), k_g'(s), a)} p(\epsilon) d\epsilon \right) ds, \tag{22}$$

where the first equality holds by the law of total differentiation (in integral form), the last equality holds with $k_i'(s) = (1-\delta)k_i + x_i(s)$, and $\frac{\partial \Phi(\cdot)(s)}{\partial B}$ is defined in (20). Substituting (20), (21), and (22) into (19) and rearranging terms, we get

$$\lambda_t^L \kappa = \lambda_t^B + \int \Bigg(U'(C_t) \Big[(1 - D(S^{1\prime} + S^{2\prime})) \exp(z) A_t^{1-\alpha} [a^{1-\rho} k_b^\rho + (1-a)^{1-\rho} k_g^\rho]^{\frac{\alpha}{\rho}} (L^d(s))^\nu \tag{23}$$

$$+ (1-\lambda)\left(-x_b(s) - x_g(s) - \psi[x_b(s), k_b] - \psi[x_g(s), k_g]\right)$$

$$+ \lambda\left((1-\delta)k_b + (1-\delta)k_g - \psi[-(1-\delta)k_b, k_b] - \psi[-(1-\delta)k_g, k_g]\right) \Big]$$

$$- \lambda_t^k k_b - \lambda_t^L L^d(s) + \mathbb{1}_{X(s)=0}(1-\lambda)\beta \int \frac{\partial \mathcal{W}_{t+1}(\mu', S^{1*}, S^{2*})}{\partial \mu'(\rho z + \epsilon, k_b', k_g', a)} p(\epsilon) d\epsilon$$

$$- U'(C_t)(x_b^E(a) + x_g^E(a)) \Bigg) \int \mathbb{1}_{k_b = x_b^E(a,q)} \mathbb{1}_{k_g = x_g^E(a,q)} [1 - X^E(a,q)] Q_q(q) Q_a(a) p_q(z - \rho_q q) dq c$$

Substituting in the marginal social value (14), we can simplify to

$$\lambda_t^L \kappa = \lambda_t^B$$

$$+ \int \int \int \left[\frac{\partial \mathcal{W}_t}{\partial \mu(\rho_q q + \epsilon_q, x_b^E(a,q), x_g^E(a,q), a)} - U'(C_t)(x_b^E(a,q) + x_g^E(a,q)) \right] Q_q(q)[1 - X^E(a,q)] Q_a(a) p_q(\epsilon_q) d\epsilon_q da dq, \tag{24}$$

B.2.6 Entrant Firm Investment

As a preliminary, note that for any Lebesgue-integrable function $\mathcal{G}$ on the state space where its derivative with respect to entrant investment $\frac{\partial \mathcal{G}(\rho_q q + \epsilon_q, x_b^E(a,q), x_g^E(a,q), a)}{\partial x_b^E(a,q)}$ exists and is bounded,

$$\frac{\partial}{\partial x_b^E(a,q)} \int \mathcal{G}(s) \Phi(s) ds = \frac{\partial}{\partial x_b^E(a,q)} \int \mathcal{G}(s) \left[\mu(z, k_b, k_g, a) + \int \mathbb{1}_{k_b = x_b^E(a,\hat{q})} \mathbb{1}_{k_g = x_g^E(a,\hat{q})} [1 - X^E(a,\hat{q})] Q_a(a) B p_q(z - \rho_q \hat{q}) d\hat{q} \right] ds$$

$$= \frac{\partial}{\partial x_b^E(a,q)} \int \int \mathcal{G}(s) \mathbb{1}_{k_b = x_b^E(a,\hat{q})} \mathbb{1}_{k_g = x_g^E(a,\hat{q})} [1 - X^E(a,\hat{q})] Q_a(a) B p_q(z - \rho_q \hat{q}) d\hat{q} ds$$

$$= \frac{\partial}{\partial x_b^E(a,q)} \int \mathcal{G}(s) \mathbb{1}_{k_b = x_b^E(a,q)} \mathbb{1}_{k_g = x_g^E(a,q)} [1 - X^E(a,q)] Q_a(a) B p_q(z - \rho_q q) ds$$

$$= [1 - X^E(a,q)] Q_a(a) B \frac{\partial}{\partial x_b^E(a,q)} \int \mathcal{G}(s) \mathbb{1}_{k_b = x_b^E(a,q)} \mathbb{1}_{k_g = x_g^E(a,q)} Q_q(q) p_q(z - \rho_q q) ds$$

$$= [1 - X^E(a,q)] Q_a Q_q(q)(a) B \frac{\partial}{\partial x_b^E(a,q)} \int \mathcal{G}(\rho_q q + \epsilon_q, x_b^E(a,q), x_g^E(a,q), a) p_q(\epsilon_q) d\epsilon_q$$

$$= [1 - X^E(a,q)] Q_a(a) B \int \frac{\mathcal{G}(\rho_q q + \epsilon_q, x_b^E(a,q), x_g^E(a,q), a)}{\partial x_b^E(a,q)} p_q(\epsilon_q) d\epsilon_q,$$

where the sixth inequality applies Leibniz integral rule and requires the conditions on $\mathcal{G}$. We assume all equilibrium functions are Lebesgue-integrable with bounded partial derivatives.

The planner's first-order condition with respect to investment in brown capital by entrant firm at given carbon dependence $x_b^E(a,q)$, for firms that the planner does choose to continue $X^E(a,q)=0$, is

$$\begin{aligned}
&\frac{\partial \mathcal{L}_t}{\partial x_b^E(a)} = 0\\
\implies U'(C_t)\frac{\partial C_t}{\partial x_b^E(a)} + \beta\frac{\partial W_{t+1}(\cdot)}{\partial x_b^E(a)} &= \lambda_t^L \frac{\partial}{\partial x_b^E(a)}\int (L^d(s)+c_f)\Phi(\cdot)(s)ds\\
&\quad + \lambda_t^k \frac{\partial}{\partial x_b^E(a)}\int k_b\Phi(\cdot)ds,
\end{aligned}$$

where

$$\begin{aligned}
\frac{\partial C_t}{\partial x_b^E(a)} &= (1-D(S^{1\prime}+S^{2\prime}))\frac{\partial}{\partial x_b^E(a)}\int \exp(z)A_t^{1-\alpha}[a^{1-\rho}k_b^\rho + (1-a)^{1-\rho}k_g^\rho]^{\frac{\alpha}{\rho}}(L^d(s))^\nu\Phi(\cdot)(s)ds\\
&- \int (x_b(z,k_b,k_g,a)+x_g(z,k_b,k_g,a))(1-\lambda)\frac{\partial\Phi(\cdot)(s)}{\partial x_b^E(a)}ds\\
&- \frac{\partial}{\partial x_b^E(a)}\int(-(1-\delta)k_b-(1-\delta)k_g)\lambda\Phi(\cdot)(s)ds\\
&- BQ_a(a)Q_q(q)\\
&- \frac{\partial}{\partial x_b^E(a)}\int(\psi[x_b(z,k_b,k_g,a),k_b]+\psi[x_g(z,k_b,k_g,a),k_g])(1-\lambda)\Phi(\cdot)(s)ds\\
&- \frac{\partial}{\partial x_b^E(a)}\int(\psi[-(1-\delta)k_b,k_b]+\psi[-(1-\delta)k_g,k_g])\lambda\Phi(\cdot)(s)ds.
\end{aligned}$$

Applying Leibniz integral rule simplifies the first-order condition to

$$\begin{aligned}
&U'(C_t)\\
&= \int \frac{\partial}{\partial x_b^E(a,q)}\Bigg(U'(C_t)\Big[(1-D(S^{1\prime}+S^{2\prime}))\exp(z)A_t^{1-\alpha}[a^{1-\rho}x_b^E(a,q)^\rho + (1-a)^{1-\rho}x_g^E(a,q)^\rho]^{\frac{\alpha}{\rho}}\\
&\qquad\qquad \times (L^d(z,x_b^E(a,q),x_g^E(a,q),a))^\nu\\
&\qquad - (x_b(z,x_b^E(a,q),x_g^E(a,q),a)+x_g(z,x_b^E(a,q),x_g^E(a,q),a))(1-\lambda)\\
&\qquad + (1-\delta)(x_b^E(a,q)+x_g^E(a,q))\lambda\\
&\qquad - (\psi[x_b(z,x_b^E(a,q),x_g^E(a,q),a),x_b^E(a,q)]+\psi[x_g(z,x_b^E(a,q),x_g^E(a,q),a),x_g^E(a,q)])(1-\lambda)\\
&\qquad - [\psi^X(x_b^E(a,q))+\psi^X(x_g^E(a,q))]\lambda\Big]\\
&\quad - \lambda_t^L(L^d(z,x_b^E(a,q),x_g^E(a,q),a)+c_f) - \lambda_t^k x_b^E(a,q)\Bigg)p_q(z-\rho_q q)dz
\end{aligned}$$

$$+\beta\frac{1}{BQ_a(a)Q_q(q)}\frac{\partial W_{t+1}(\cdot)}{\partial x_b^E(a,q)}.$$

Now, consider the derivative of tomorrow's welfare with respect to entrant investment, under $X^E(a,q)=0$

$$\begin{aligned}
\frac{\partial W_{t+1}(\cdot)}{\partial x_b^E(a,q)} &= \int \frac{\partial W(\mu', S^{1*}, S^{2*})}{\partial \mu'(s')}\frac{\partial \mu'(s')}{\partial x_b^E(a)}ds' \\
&= \int \frac{\partial W(\mu', S^{1*}, S^{2*})}{\partial \mu'(s')}\frac{\partial T^*(\mu,\cdot)(s')}{\partial x_b^E(a)}ds' \\
&= BQ_a(a)Q_q(q)\int \frac{\partial}{\partial x_b^E(a,q)}(1-X(z,x_b^E(a),x_g^E(a),a))(1-\lambda)p_q(z-\rho_q q) \\
&\qquad \times \int \frac{\partial W(\mu', S^{1*}, S^{2*})}{\partial \mu'(\rho z+\epsilon, k_g'(z,x_b^E(a,q),x_g^E(a,q),a), k_d'(z,x_b^E(a,q),x_g^E(a,q),a),a)}p(\epsilon)d\epsilon dz
\end{aligned}$$

by switching the order of differentiation, plugging in the definition of T^*, and switching the order of integration. Plugging in the marginal social value (14), the first-order condition simplifies to

$$U'(C_t) = \frac{\partial}{\partial x_b^E(a,q)}\int \frac{\partial W_t(\mu, S^{1\prime}, S^{2\prime})}{\partial \mu(z, x_b^E(a,q), x_g^E(a,q), a)}p_q(z-\rho_q q)dz, \tag{25}$$

Following an identical process, the first-order condition with respect to $x_g^E(a,q)$ is

$$U'(C_t) = \frac{\partial}{\partial x_g^E(a,q)}\int \frac{\partial W_t(\mu, S^{1\prime}, S^{2\prime})}{\partial \mu(z, x_b^E(a,q), x_g^E(a,q), a)}p_q(z-\rho_q q)dz. \tag{26}$$

B.2.7 Entrant Firm Continuation

Finally, the planner's decision of $X^E(a,q)$ is similar to the incumbent firm continuation decision. The planner chooses $X^E(a,q)=1$ iff the marginal value of exiting an entrant firm (zero) is greater than the marginal value of the entrant operating:

$$\begin{aligned}
&\underbrace{0}_{\text{exit}} > \underbrace{BQ_a(a)Q_q(q)\left[U'(C_t)\left(-x_b^E(a,q)-x_g^E(a,q)\right)+\int \frac{\partial \mathcal{W}_t(\mu, S^{1\prime}, S^{2\prime})}{\partial \mu(z, x_b^E, x_g^E, a)}p_q(z-\rho_q q)dz\right]}_{\text{operate}} \\
&\implies 0 > U'(C_t)\left(-x_b^E(a)-x_g^E(a)\right)+\int \frac{\partial \mathcal{W}_t(\mu, S^{1\prime}, S^{2\prime})}{\partial \mu(z, x_b^E, x_g^E, a)}Q_z(z)dz
\end{aligned} \tag{27}$$

B.3 Augmented Firm Bellman Equation

Now define $\omega_t(s,\mu,S^1,S^2) \equiv \frac{\partial \mathcal{W}_t(\mu,S^1,S^2)}{\partial \mu(s)}$ as the marginal social value of firms at a particular location in state-space s at time t,

$$\begin{aligned}
\omega_t(s,\mu,S^1,S^2) = U'(C_t)\Big[&(1-D(S^{1\prime}+S^{2\prime}))\exp(z)A_t^{1-\alpha}[a^{1-\rho}k_b^{\rho}+(1-a)^{1-\rho}k_g^{\rho}]^{\frac{\alpha}{\rho}}(L^d(s))^{\nu} \\
&+(1-\lambda)\left(-x_b(s)-x_g(s)-\psi[x_b(s),k_b]-\psi[x_g(s),k_g]\right) \\
&+\lambda\left((1-\delta)(k_b+k_g)-\psi^X(k_b)-\psi^X(k_g)\right)\Big] \\
&-\lambda_t^L(L^d(s)+c_f)-\lambda_t^k k_b \\
&+(1-X(s))(1-\lambda)\beta\mathbb{E}[\omega_{t+1}(s',\mu,S^1,S^2)] \\
\implies \bar{\omega}_t(s,\mu,S^1,S^2) = &(1-D(S^{1\prime}+S^{2\prime}))\exp(z)A_t^{1-\alpha}[a^{1-\rho}k_b^{\rho}+(1-a)^{1-\rho}k_g^{\rho}]^{\frac{\alpha}{\rho}}(L^d(s))^{\mu} \\
&+(1-\lambda)\left(-x_b(s)-x_g(s)-\psi[x_b(s),k_b]-\psi[x_g(s),k_g]\right) \\
&+\lambda\left((1-\delta)k_b+(1-\delta)k_g-\psi^X(k_b)-\psi^X(k_g)\right) \\
&-\hat{\lambda}_t^L(L^d(s)+c_f)-\hat{\lambda}_t^k k_b \\
&+(1-X(s))(1-\lambda)\beta\frac{U'(C_{t+1})}{U'(C_t)}\mathbb{E}[\bar{\omega}_{t+1}(s',\mu',S^{1\prime},S^{2\prime})]
\end{aligned}$$

where $\bar{\omega}_t(s,\mu,S^1,S^2) \equiv \frac{\omega_t(s,\mu,S^1,S^2)}{U'(C_t)}$, $\hat{\lambda}_t^k \equiv \frac{\lambda_t^k}{U'(C_t)}$, and $\hat{\lambda}_t^L \equiv \frac{\lambda_t^L}{U'(C_t)}$.

B.3.1 Incumbent Firm

We now show that the marginal social value $\bar{\omega}_t(s,\mu,S^1,S^2)$ achieves a value function $\hat{V}_t(s,\mu,S^1,S^2)$ induced by an augmented Bellman equation written similarly. In particular, consider value function

$$\begin{aligned}
\hat{V}_t(s,\mu,S^1,S^2) := \max_{\hat{L}^d(s),\hat{x}_b(s),\hat{x}_g(s),\hat{X}(s)} &(1-D(S^{1\prime}+S^{2\prime}))\exp(z)A_t^{1-\alpha}[a^{1-\rho}k_b^{\rho}+(1-a)^{1-\rho}k_g^{\rho}]^{\frac{\alpha}{\rho}}(\hat{L}^d(s))^{\nu} \\
&+(1-\lambda)\left(-\hat{x}_b(s)-\hat{x}_g(s)-\psi[\hat{x}_b(s),k_b]-\psi[\hat{x}_g(s),k_g]\right) \\
&+\lambda\left[(1-\delta)(k_b+k_g)-\psi^X(k_b)-\psi^X(k_g)\right] \\
&-\hat{\lambda}_t^L(\hat{L}^d(s)+c_f)-\hat{\lambda}_t^k k_b \\
&+(1-\hat{X}(s))(1-\lambda)\beta\frac{U'(C_{t+1})}{U'(C_t)}\mathbb{E}[\bar{\omega}_{t+1}(s',\mu',S^{1\prime},S^{2\prime})]
\end{aligned} \tag{28}$$

subject to $\hat{x}_b = -(1-\delta)k_b$ and $\hat{x}_g = -(1-\delta)k_g$ if $\hat{X}(s) = 1$. The first-order condition with respect to $\hat{L}^d(s)$

$$\hat{\lambda}_t^L = (1 - D(S^{1\prime} + S^{2\prime}))\nu \exp(z) A_t^{1-\alpha} [a^{1-\rho} k_b^\rho + (1-a)^{1-\rho} k_g^\rho]^{\frac{\alpha}{\rho}} (\hat{L}^d(s))^{\nu-1}$$

matches labor allocation optimality condition from full planner's problem (15). The first-order condition with respect to $\hat{x}_b(s)$ conditional on choosing $\hat{X}(s) = 0$ is

$$1 + \psi'[x_b(s), k_b] = \beta \frac{U'(C_{t+1})}{U'(C_t)} \frac{\partial}{\partial k_b'(s)} \mathbb{E}[\bar{\omega}_{t+1}(s', \mu', S^{1\prime}, S^{2\prime})]).$$

matches brown investment optimality condition from full planner's problem (16). The first-order condition with respect to $\hat{x}_g(s)$ conditional on choosing $\hat{X}(s) = 0$

$$1 + \psi'[x_g(s), k_g] = \beta \frac{U'(C_{t+1})}{U'(C_t)} \frac{\partial}{\partial k_g'(s)} \mathbb{E}[\bar{\omega}_{t+1}(s', \mu', S^{1\prime}, S^{2\prime})]$$

matches green investment optimality condition from full planner's problem (17). Finally, firm s exits $\hat{X}(s) = 1$ if

$$\begin{aligned} &\underbrace{(1-\delta)(k_b + k_g) - \psi^X(k_b) - \psi^X(k_g)}_{\text{exit}} \\ > \; &\underbrace{-\hat{x}_b(s) - \psi[\hat{x}_b(s), k_b] - \hat{x}_g(s) - \psi[\hat{x}_g(s), k_g] + \beta \frac{U'(C_{t+1})}{U'(C_t)} \mathbb{E}[\bar{\omega}_{t+1}(s', \mu', S^{1\prime}, S^{2\prime})]}_{\text{operate}}, \end{aligned}$$

continues if operating value strictly exceeds exiting, and is indifferent if the values are equal. We assume that an indifferent firm breaks the tie by following the exit policy assigned by the planner. As a reminder, this holds for a measure zero of firms in our model. The exit decision rules thus matches those of the planner in (18).

Hence, the policies induced by the optimization problem given by Equation (28) satisfy the optimality conditions from the full constrained planner's problem. Furthermore, $\bar{\omega}_t(s, \mu, S^1, S^2) = \hat{V}_t(s, \mu, S^1, S^2)$, so we can now write $\bar{\omega}_t(s, \mu, S^1, S^2)$ as an augmented Bellman equation:

$$\begin{aligned} \bar{\omega}_t(s, \mu, S^1, S^2) = \max_{\hat{L}^d(s), \hat{x}_b(s), \hat{x}_g(s), \hat{X}(s)} & (1 - D(S^{1\prime} + S^{2\prime})) \exp(z) A_t^{1-\alpha} [a^{1-\rho} k_b^\rho + (1-a)^{1-\rho} k_g^\rho]^{\frac{\alpha}{\rho}} (\hat{L}^d(s))^\nu \\ & + (1-\lambda)\left(-\hat{x}_b(s) - \hat{x}_g(s) - \psi[\hat{x}_b(s), k_b] - \psi[\hat{x}_g(s), k_g]\right) \\ & + \lambda\left[(1-\delta)(k_b + k_g) - \psi^X(k_b) - \psi^X(k_g)\right] \end{aligned} \tag{29}$$

$$
\begin{aligned}
&- \hat{\lambda}_t^L(\hat{L}^d(s) + c_f) - \hat{\lambda}_t^k k_b \\
&+ (1 - \hat{X}(s))(1 - \lambda)\beta \frac{U'(C_{t+1})}{U'(C_t)} \mathbb{E}[\bar{\omega}_{t+1}(s', \mu', S^{1\prime}, S^{2\prime})]
\end{aligned}
$$

subject to $\hat{k}_i' = (1 - \delta)k_i + \hat{x}_i$ and $\hat{x}_i = -(1 - \delta)k_i$ if $\hat{X}(s) = 1$.

B.3.2 Entrant Firm

Consider the following optimization problem for entrants:

$$
\bar{\omega}_t^E(a, q, \mu, S^1, S^2) = \max_{\hat{X}^E(a,q), \hat{x}_b^E(a,q), \hat{x}_g^E(a,q)} (1 - \hat{X}^E(a, q)) \left[-\hat{x}_b^E(a, q) - \hat{x}_g^E(a, q) + \mathbb{E}[\bar{\omega}_t(s, \mu, S^1, S^2)] \right] \tag{30}
$$

subject to $\hat{x}_b^E = 0$ and $\hat{x}_g^E = 0$ if $\hat{X}(s) = 1$. Conditional on not exiting, the first-order condition with respect to $\hat{x}_b^E$ is

$$
1 = \frac{\partial}{\partial \hat{x}_b^E(a, q)} \mathbb{E}[\bar{\omega}_t(s, \mu, S^1, S^2)],
$$

which matches the planner's first-order condition (25). Similarly, the first-order condition with respect to $\hat{x}_g^E(a, q)$

$$
1 = \frac{\partial}{\partial \hat{x}_g^E(a, q)} \mathbb{E}[\bar{\omega}_t(s, \mu, S^1, S^2)]
$$

matches the planner's first-order condition (26). Finally, entrant a exits $\hat{X}^E(a) = 1$ iff

$$
\underbrace{0}_{\text{exit}} > \underbrace{-x_b^E(a, q) - x_g^E(a, q) + \mathbb{E}[\bar{\omega}_t(s, \mu, S^1, S^2)]}_{\text{operate}},
$$

which matches the planner's optimality condition (27).

Let $\tau_t \equiv \frac{\lambda_t^k}{U'(C_t)}$, $\hat{w} \equiv \frac{\lambda_t^L}{U'(C_t)}$, and $\frac{1}{\hat{R}_t} \equiv \beta \frac{U'(C_{t+1})}{U'(C_t)}$. Since the dynamic program given by (29) matches that of (6), it must be that $\bar{\omega}_t(s, \mu, S^1, S^2) = \hat{\omega}_t(s, \mu, S^1, S^2)$, and then (7) is an identical dynamic program to (30), so $\bar{\omega}_t^E(a, \mu, S^1, S^2) = \hat{\omega}_t^E(a, \mu, S^1, S^2)$, and the policies induced by the programs are also equal. Hence, Propositions 4 and 5 hold.

We finally note that Equation (24) can be written:

$$
\mathbb{E}[\hat{\omega}_t^E(a, q, \mu, S^1, S^2)] + \frac{\lambda_t^B}{U'(C_t)} = \hat{w}_t \kappa,
$$

proving Proposition 6.

C Decentralization of the Planner's Solution

Let us rewrite the firm's Bellman equations in the decentralized economy with the appropriate tax wedges added:

$$\mathcal{V}_t(s) = \pi_t^D(s) - \tau_t k_{b,t} + \lambda \mathcal{V}^X(k_b, k_g) + (1-\lambda)\max\{\mathcal{V}^X(k_b, k_g), V_t^C(s)\}, \tag{31}$$

where $\pi_t(s)$ is the cash flow function:

$$\pi_t(s) = \max_L (1 - D(S^{1\prime} + S^{2\prime}))\exp(z) A_t^{1-\alpha}[a^{1-\rho}k_b^\rho + (1-a)^{1-\rho}k_g^\rho]^{\frac{\alpha}{\rho}} L^\nu - w_t L - w_t c_f,$$

and where the exiting value is given by:

$$\mathcal{V}^x(k_b, k_g) = (1-\delta)(k_b + k_g) - \psi^X(k_b) - \psi^X(k_g),$$

and nonexiting value defined by the solution to the following optimization problem:

$$\mathcal{V}_t^C(s) = \max_{x_b, x_g} -x_b - x_g - \psi(x_b, k_b) - \psi(x_g, k_g) + \frac{1}{R_t}\mathbb{E}[\mathcal{V}_{t+1}(s')], \tag{32}$$

subject to:

$$k_j' = (1-\delta)k_j + x_j, \quad \text{for } j \in \{b, g\}.$$

Free entry implies:

$$w_t \kappa \geq \mathbb{E}[\mathcal{V}_t^E(a, q)],$$

where an entrant's value after observing carbon dependence a is given by:

$$\mathcal{V}_t^E(a, q) = \max\{0, \max_{k_b, k_g} -k_b - k_g + \mathbb{E}[\mathcal{V}_t(s)]\}. \tag{33}$$

Now, for the set of taxes and prices as defined by the set $\mathcal{P}_t$, the dynamic program given by (31) and (33) in the competitive economy are identical to the analogous dynamic programs from the augmented Bellman equation and augmented entry problem given by (6) and (7). Hence, it follows that the values and induced policies are the same, as the dynamic programs are the same. That is, $\mathcal{V}_t(s) = \hat{\omega}_t(s, \mu_t, S_{t-1}^1, S_{t-1}^2)$ and $\mathcal{V}_t^E(a) = \hat{\omega}_t^E(a, \mu_t, S_{t-1}^1, S_{t-1}^2)$.

Now, note that by definition $\frac{1}{\hat{R}_t} = \beta \frac{U'(C_{t+1})}{U'(C_t)}$ so the Euler of the representative household holds at that risk-free rate, and market clearing of the risk-free asset is achieved.

Next, note that Equation (7) implies that the free entry condition, (33) is satisfied. Moreover, the labor market clears under B_t from the planner's solution,[17] and as the firm-level labor demands are equal under both problems and the supply of labor is exhausted in the planner's problem with mass of entrants B_t.

Finally, note that the firm's policies solve their problems, and thus a competitive equilibrium that implements the planner's problem is achieved under a path of taxes $\{\tau_t\}_{t=t_0}^{\infty}$, prices $\{\hat{w}_t, \hat{R}_t\}_{t=t_0}^{\infty}$, and mass of entering firms $\{B_t\}_{t=t_0}^{\infty}$.

C.1 Decentralization through Tax on Noisy Emissions

Here, we consider implementation of the planner's solution under a model extension. In this subsection, we consider a case where the level of brown capital each firm owns $k_{b,i,t}$ is not observed and hence is not taxable. However, the taxation authority is allowed to observe the level of emissions $\xi_{i,t}$. Of course, if emissions are still deterministic then the taxation authority trivially observes brown capital by inverting the linear relationship between the observed emissions and brown capital level. To make the case more interesting, we hence further assume that emissions are no longer deterministic for a given firm, but instead allow for a mean-preserving spread, making emissions a noisy signal of unobserved brown capital:

$$\xi_{j,t} = \eta_{j,t} \gamma k_{b,j,t};\ \eta_{j,t} \sim_{iid} F; \mathbb{E}[\eta_{j,t}] = 1\ \eta_{j,t} \geq 0\ a.s. \tag{34}$$

We prove that the constrained social planner's solution does not change (Proposition 8) and that it can be decentralized with a tax on emissions instead of a tax on brown capital (Proposition 9).

Proposition 8. *The constrained social planner's solution is unchanged under the noisy emissions as given by Equation 34.*

Proof. See Appendix D. □

Proposition 9. *The planner's solution remains implementable, with a tax on emissions replacing a tax on brown capital.*

Proof. See Appendix D. □

[17] Note that the labor market clearing condition must trivially bind for the planner, else consumption can be strictly increased by providing the remaining labor.

The intuition behind these proofs is straightforward. We first show that the planner's solution is unchanged. The uncertainty over firm-level emissions washes out in the planner's problem when integrating across the firm measure during aggregation. Next, we show that the planner's solution remains implementable through a linear tax on (noisy) emissions. The intuition here is that the risk-neutral firms integrate out the mean-preserving noise in their future profits coming from the realization of their emissions (since the tax is linear) when making their investment decisions, and after realization of the emissions shock the carbon tax is sunk from the perspective of the firm and hence does not impact the chosen policies. Hence the policies of the firm remain identical to those of the constrained planner.

C.2 Decentralization of the Planner's Allocation through Emissions Credits

In this section, we prove by construction that the constrained planner's allocation can be implemented by the government through commitment to a carbon credit scheme. Specifically, suppose the government credibly announces that they will supply a mass of carbon credits $\mathcal{C}_t$ in each period. The government will sell carbon credits in period t at the moment of operation to the measure of firms, in fixed supply $\mathcal{C}_t$ and with market clearing price p^c_t. An equilibrium in this setting thus corresponds to our prior definition in 1 with additional requirement of market clearing in the carbon credit market in each period. Any operating firm must procure an equal amount of carbon credits per each unit of brown capital they utilize, or else face a sufficiently strong penalty such that all firms comply with the requirement.

It is simple to construct the equilibrium. First, we set the supply of credits in each period equal to the total brown capital level in the planner's allocation $\mathcal{C}_t = K^b_t$ and show that it implements the equilibrium. Consider the set of prices that constitutes the equilibrium in our baseline decentralization case, and append further the set of carbon credit prices $\{p^c_s\}_{s=t}^{\infty}$ where $p^c_s = \hat{\tau}_s$. Under these prices, the Bellman equations for the firms exactly equal those of the planner, and hence the firm-level policies correspond as well. Then, the aggregates also correspond, including crucially that the level of brown capital equals that of the planner. As this was the level of supply set by the government, the carbon credit market clears and the new definition of equilibrium is satisfied.

Hence, the planner's allocation in our model can be implemented through carbon credits, i.e. by fixing the quantity of carbon emissions, as opposed to setting the price of emissions.

D Alternative Decentralization of the Planner's Solution

Suppose the taxation authority cannot observe k_b, and hence implementing the tax on brown capital such as in section 4.1 is infeasible. In this section, we will show that the socially efficient allocation is achievable through a tax on emissions so long as emissions are observable to the taxation authority, even if the emissions are noisy.

Suppose firm-level emissions are given by Equation (34). From the perspective of the social planner, since the law of large numbers washes out the idiosyncratic uncertainty over $\eta_{i,t}$, the problem remains the same as in 5, and hence the socially optimal policies are identical. We will now show that a decentralized equilibrium that implements the planner's policies can be achieved using a tax on emissions at the firm-level rather than brown capital.

Consider the following decentralized problem with emissions taxes $\tau_t^E = \frac{\tau_t}{\gamma}$, for a realization ξ of emissions:

$$
\begin{aligned}
\hat{\mathcal{V}}_t(s,\xi,\mu,S^1,S^2) = &\max_{x_b(s,\xi),x_g(s,\xi),L^d(s,\xi),X(s,\xi)} (1-D(S^{1\prime}+S^{2\prime}))\exp(z)A_t^{1-\alpha}[a^{1-\rho}k_b^{\rho}+(1-a)^{1-\rho}k_g^{\rho}]^{\frac{\alpha}{\rho}}(L^d(s,\xi))^{\mu} \qquad (35)\\
&+(1-\lambda)\left(-x_b(s,\xi)-x_g(s,\xi)-\psi[x_b(s,\xi),k_b]-\psi[x_g(s,\xi),k_g]\right)\\
&+\lambda\left((1-\delta)k_b+(1-\delta)k_g-\psi^X(k_b)-\psi^X(k_g)\right)\\
&-\tau_t^E\xi-\hat{w}(L^d(s,\xi)+c_f)\\
&+(1-X(s,\xi))(1-\lambda)\frac{1}{\hat{R}_t}\mathbb{E}[\hat{\mathcal{V}}_{t+1}(s',\mu',S^{1\prime},S^{2\prime})]\\
&=-\tau_t^E\xi\\
&+\max_{x_b(s,\xi),x_g(s,\xi),L^d(s,\xi),X(s,\xi)} (1-D(S^{1\prime}+S^{2\prime}))\exp(z)A_t^{1-\alpha}[a^{1-\rho}k_b^{\rho}+(1-a)^{1-\rho}k_g^{\rho}]^{\frac{\alpha}{\rho}}(L^d(s,\xi))^{\mu}\\
&+(1-\lambda)\left(-x_b(s,\xi)-x_g(s,\xi)-\psi[x_b(s,\xi),k_b]-\psi[x_g(s,\xi),k_g]\right)\\
&+\lambda\left((1-\delta)k_b+(1-\delta)k_g-\psi^X(k_b)-\psi^X(k_g)\right)\\
&-\hat{w}(L^d(s,\xi)+c_f)\\
&+(1-X(s,\xi))(1-\lambda)\frac{1}{\hat{R}_t}\mathbb{E}[\hat{\mathcal{V}}_{t+1}(s',\xi',\mu',S^{1\prime},S^{2\prime})].
\end{aligned}
$$

Hence, the policies are independent of the stochastic ξ, and are only a function of s. Moreover, since $\mathbb{E}_t[\tau_t^E\xi] = \mathbb{E}[\frac{\tau_t}{\gamma}\gamma k_{b,t}]$,

$$
\begin{aligned}
\mathbb{E}_{t-1}[\hat{\mathcal{V}}_t(s,\xi,\mu,S^1,S^2)] = \mathbb{E}_{t-1}\Bigg[&\max_{x_b(s),x_g(s),L^d(s),X(s)} (1-D(S^{1\prime}+S^{2\prime}))\exp(z)A_t^{1-\alpha}[a^{1-\rho}k_b^{\rho}+(1-a)^{1-\rho}k_g^{\rho}]^{\frac{\alpha}{\rho}}(L^d(s))^{\mu}\\
&+(1-\lambda)\left(-x_b(s)-x_g(s)-\psi[x_b(s),k_b]-\psi[x_g(s),k_g]\right)
\end{aligned}
$$

$$
\begin{aligned}
&+ \lambda \big((1-\delta)k_b + (1-\delta)k_g - \psi^X(k_b) - \psi^X(k_g)\big) \\
&- \tau_t^E \xi - \hat{w}(L^d(s) + c_f) \\
&+ (1 - X(s))(1-\lambda)\frac{1}{\hat{R}_t}\mathbb{E}[\hat{\mathcal{V}}_{t+1}(s', \mu', S^{1\prime}, S^{2\prime})]\Bigg], \\
&= \mathbb{E}_{t-1}\Bigg[\max_{x_b(s), x_g(s), L^d(s), X(s)} (1 - D(S^{1\prime} + S^{2\prime}))\exp(z)A_t^{1-\alpha}[a^{1-\rho}k_b^\rho + (1-a)^{1-\rho}k_g^\rho]^{\frac{\alpha}{\rho}}(L^d(s))^\mu \\
&+ (1-\lambda)\left(-x_b(s) - x_g(s) - \psi[x_b(s), k_b] - \psi[x_g(s), k_g]\right) \\
&+ \lambda \big((1-\delta)k_b + (1-\delta)k_g - \psi^X(k_b) - \psi^X(k_g)\big) \\
&- \tau_t k_b - \hat{w}(L^d(s) + c_f) \\
&+ (1 - X(s))(1-\lambda)\frac{1}{\hat{R}_t}\mathbb{E}[\hat{\mathcal{V}}_{t+1}(s', \mu', S^{1\prime}, S^{2\prime})]\Bigg].
\end{aligned}
$$

i.e. the ex-ante values and ex-post policies induced by (35) equal those of the augmented Bellman (6), and hence the planner's allocation can be decentralized by taxation of noisy, but observed, emissions in lieu of a direct tax on brown capital.

E Entrant Carbon Dependence Estimation

This appendix describes how we estimate the entrant distribution over carbon dependence $Q(a)$ to match the distribution in the data. First, we describe a normalization assumption to map the data into the model. Second, we describe how we use a simplified model to choose a grid points and choose an initial guess for $Q(a)$. A simpler version of the model that ignores depreciation and capital adjustment provides an excellent initial guess that we refine with numerical optimization.

In the data, we observe emissions and total capital for each firm and, in the model, one unit of brown capital creates γ emissions, so emission intensity can be separated into $\log \gamma$ and the firm's log capital ratio. We assume that the firm with the highest observed emission intensity only uses brown capital, so its emission intensity equals $\log \gamma$. Thus, by the subtracting the highest firm's emission intensity, we get the log capital ratio for each firm

$$
\log \frac{\text{Emissions}_i}{\text{PPEGT}_i} - \log \frac{\text{Emissions}_{(n)}}{\text{PPEGT}_{(n)}} = \log \frac{k_{b,i}}{k_{b,i} + k_{g,i}}
$$

A simplified version of the firm problem in BAU is

$$
\max_{k_b, k_g} -k_b - k_g + \mathbb{E}_z\Big[z\big[1 - D(K^b)\big]\left[a_i k_b^\rho + (1-a_i)k_g^\rho\right]^{\alpha/\rho}\Big]
$$

Combining investment first-order conditions together, we get a mapping between carbon dependence and intensity.

$$\frac{k_b}{k_b + k_g} = \frac{a_i^{\frac{1}{1-\rho}}}{a_i^{\frac{1}{1-\rho}} + (1 - a_i)^{\frac{1}{1-\rho}}}$$
$$\implies a_i = \frac{[\exp(\frac{k_b}{k_b+k_g})]^{1-\rho}}{[\exp(\frac{k_b}{k_b+k_g})]^{1-\rho} + [1 - \exp(\frac{k_b}{k_b+k_g})]^{1-\rho}} \tag{36}$$

We use seven a-grid points to correspond to the seven bars in the first panel of Figure 6. The initial guess of $Q(a_i)$ is height of each bar and we take the midpoint of each bar and compute the corresponding a_i from Equation 36. We then use a numerical optimization routine to refine probability of each grid point $Q(a)$ to minimize the sum of the squared deviation over the seven bars.

F Computation of Quantitative Model

We solve the problem numerically using value function iteration. We first detrend the firm problems by dividing the Bellman equations and augmented Bellman equations by A_t. We then discretize the state-space: we approximate the z process following Tauchen (1986), discretize both k_b/A_t and k_g/A_t linearly in logs.

To compute a transition path, we first solve for the firm distribution associated with the balanced growth path with no climate damages nor taxes. At this point, the interest rate is at the balanced growth path level. In this balanced growth path, we set the wage to one and solve for the implied entry cost κ, which is held constant for the rest of the transition path while wages move to satisfy the entry condition. We take as the initial firm distribution μ the invariant distribution over the detrended balanced growth path. We then guess the path of interest rates, output taxes, and brown capital taxes over the transition path. Given these guesses, we move backward along the transition path solving for value and policy function. The value functions imply the value of entry, which pins down the wage along the transition path. We then move forward along the transition path updating the measure of firms consistent with the policy functions. We solve the full transition path over 5,000 years, which allows for the convergence of the firm distribution. This extended timeframe is largely driven by the strong persistence of the transitory component of the stock of carbon emissions. The stock of emissions determines damages from climate change which then pin down optimal output taxes. We provide sufficient time for the temporary carbon stock to dissipate for the optimal output taxes to converge and then the firm distribution converge. After 2,000 years, we extend the aggregate variable for an additional 5,000 years assuming that are on the balanced growth path and grow by exogenous growth rate of TFP. We then update the guess of the interest rate based on the consumption Euler equation of the representative household and the guesses of the optimal output and brown capital taxes based on the solutions to the planner's problem.