The Macroeconomic Effects of Carbon Taxes: The Role of Investment and Emission Rate Heterogeneity *

Giuseppe Fiori Nora Traum

September 23, 2019

Abstract

We study the quantitative effects of a carbon tax in a dynamic, general equilibrium model with production heterogeneity and technology adoption. The vintage technology available to a firm determines its emission rate. Adopting newer technology is subject to a non-convex adjustment cost that leads firms to have (S,s) policy functions for technology and capital adjustment. As new and old technologies coexist, the endogenous distribution of vintage technology and capital stocks determines the aggregate effects of a carbon tax. We discipline capital and emission heterogeneity in the model with U.S. microeconomic data. We find that firm heterogeneity in emission rates determines the aggregate effects of a carbon tax, both in the short- and long- run. In the long run, GDP losses from a representative firm model are more than double those with heterogeneous emission rates. Short-run effects depend on the policy implementation, with policies initially exempting older establishments increasing emissions.

Keywords: (S,s) policies, lumpy investment, quantitative general equilibrium, carbon tax.
JEL Codes: E22, E62, H23, H32

*Giuseppe Fiori: Federal Reserve Board, Division of International Finance, 20th and C St. NW, Washington D.C. 20551, United States (website: www.giuseppefiori.net and e-mail: giuseppe.fiori@frb.gov); Nora Traum: HEC Montréal 3000, chemin de la Côte-Sainte-Catherine Montréal (Québec) Canada H3T 2A7 (website: www.hec.ca/en/profs/nora.traum.html and e-mail: nora.traum@hec.ca). First Version: February 15, 2015. We would like to thank Fanny Henriet, Nir Jaimovich, Andrea Lanteri, Eric Leeper, Pietro Peretto, Tony Smith, Laura Taylor, Roger von Haefen, and numerous conference and seminar participants. Financial support from North Carolina State University's Sustainability Initiative Grant is gratefully acknowledged. The views expressed in this paper are solely the responsibility of the authors and should not be interpreted as reflecting the views of the Board of Governors of the Federal Reserve System (or of any other person associated with the Federal Reserve System).

1 Introduction

Substantial differences exist within-industry in the characteristics of producers, e.g. the age of their capital stock, management, and productivity levels. A growing literature in macroeconomics incorporates heterogeneous firms into theoretical models, but the majority are silent as to how economic forces influence a producer's attributes and productivity.[1] This omission is not innocuous as policies often target specific characteristics of firms (e.g. unique tax treatments, preferential contracts to specific producers, and policies aimed at the size or efficiency of producers), giving direct incentives for producers to adapt and change these attributes over time, which in turn affects the aggregate economy. In this paper, we highlight the importance of this issue for a policy that is receiving increasing academic and policy attention: the implementation of a carbon tax. This policy is particularly pertinent as substantial heterogeneity in energy and emissions exists within-industry in the U.S. manufacturing sector, implying the policy's aggregate effects depend on how the underlying distribution of firms respond to the policy.

Our contribution is to develop a general-equilibrium model with production heterogeneity and technology adoption and demonstrate with the model that firm heterogeneity in emission rates—a key attribute related to environmental policy—substantially alters the aggregate economic effects of introducing a carbon tax, both in the short- and long- run. Empirical evidence suggests that heterogeneity in emissions is sizeable and exceeds the degree of heterogeneity in many other measures of productivity (see Lyubich et al., 2018).[2] Plant-level heterogeneity in emissions has two significant effects on design-

[1] Seminal contributions include Thomas (2002), Veracierto (2002), Khan and Thomas (2008), Bachmann and Bayer (2014) and Bloom et al. (2018).

[2] Moreover, several recent empirical works show that the cross-sectional dispersion in plant-level emission intensities correlates with differences in total factor productivity, management, and plant size. Shapiro and Walker (2018) find that U.S. establishments with higher total factor productivity have lower emission intensities. Bloom et al. (2010) find that better managed U.K. establishments have lower energy intensities and CO_2 emissions. Moreover, many studies have found that smaller plants have larger emissions inten-

ing a policy targeting emissions. First, a uniform tax has unequal tax incidence across plants, being more lenient or strict on different producers, depending on their emission rates. Second, this unequal tax incidence creates incentives for plants to change their productivity over time, dynamically altering the within-industry distribution of energy and emissions. To quantify these effects, we develop a general equilibrium framework that explicitly models the timing of firms' decisions to replace old capital and adopt new technology in response to policy measures. Old and new technologies result in different emission levels and efficiencies of capital, labor and energy. To substantiate this channel, we use plant-level data in the U.S. electricity sector and show that the timing of technology adoption through capital accumulation is a significant contributor to emission rate heterogeneity.

We use our model, parameterized to reproduce the pattern of capital accumulation at the plant level, to study the transitional dynamics following the introduction of a tax on emissions. Our findings indicate that the aggregate effects of a carbon tax depend on the underlying microeconomic structure, i.e. the cross-sectional distribution of technology and capital stocks across firms. In the long run, GDP losses predicted from a representative firm model are more than double those from our model accounting for heterogeneous emission rates, as the latter structure explicitly accounts for dynamic changes in the economy's average productivity and energy efficiency. Following a carbon tax enactment, firms with older technology and higher emission rates find it optimal to upgrade to newer technology with lower emission rates to reduce their tax bill. In the short run, the response of emissions depends upon the implementation of the tax. When older establishments are exempt from the carbon tax, emissions *increase* in the short run, as firms with older equipment are not affected by the tax bill initially and have no incentive to

sities within developing countries (see for instance Dasgupta et al. (2002) for air pollution in Mexico and Brazil and Qi et al. (2019) for water pollution in China).

invest in new technology.

Our theoretical model builds on Thomas (2002) in its use of state-dependent investment.[3] Firms combine physical capital, labor and energy to produce a homogenous good. Technological progress is exogenous and evolves deterministically. Adopting the latest technology and adjusting the capital stock is subject to a non-convex adjustment cost. Exercising this option allows the firm to improve its efficiency and to reduce its emission rates. The timing of upgrading technology is endogenous and subject to a non-convex adjustment cost that leads firms to purse (S,s) policy functions. In equilibrium, old and new technologies coexist in the economy by choice.

The implications of a carbon tax on the vintage replacement decision are a priori ambiguous. A carbon tax increases factor prices of producers. On the one hand, higher factor prices make investing in cleaner and more energy-efficient technology desirable. On the other hand, higher factor prices may reduce the size of producers, depressing investment incentives. We find that the cross-sectional distribution of plants is a critical determinant of the short-run effects. Our results show the quantitative relevance of the degree of emission heterogeneity for the cost-benefit analysis of environmental policies.

Related Literature Our work is related to several strands of the literature. First, it relates to studies on the role of firm heterogeneity for investment dynamics. Notable examples include Veracierto (2002), Thomas (2002), Gourio and Kashyap (2007), Khan and Thomas (2008), Bachmann and Bayer (2014) and Bloom et al. (2018). We build on these works with a notable innovation: characteristics, such as productivity or energy efficiency, are chosen by firms. In addition, our quantitative focus distinguishes our work from the large literature on vintage capital. As is the case for Solow (1960), in general, the dynamics of

[3]The approach is related to Caballero and Engel's (1999) generalized (S, s) model in its use of stochastic adjustment costs to simultaneously yield lumpy plant-level investment and smooth aggregates. Also related are Khan and Thomas (2003, 2008), Bachmann et al. (2013), and Bachmann and Bayer (2014), which employ state-dependent investment models to study business cycle properties of aggregate series.

capital vintage models cannot be captured through a representative firm model unless knife-edge conditions are met. As a result, the number of studies that have confronted vintage models with microeconomic data have been limited to non-existent.[4] For a complete list of references and a historical perspective on the evolution of this literature, we refer the reader to Boucekkine et al. (2008).

Our work also relates to the macro literature on environmental policies, see for instance Krusell and Smith (2015).[5] Our study abstracts from the environmental benefits associated with policy options to focus the analysis on the new economic implications that arise from explicitly accounting for the distribution of establishments. Acemoglu et al. (2016) study the role of *innovation* within energy technology.[6] Our work complements their analysis by instead studying the *timing of adoption* of new technology once it is available to current establishments.

Our work advances the growing literature on the effects of environmental policies in heterogeneous-agent general-equilibrium models. Rausch et al. (2010) and Fullerton and Monti (2013) study the distributional effects of green policies on households. Most closely related to our analysis are Shapiro and Walker (2018) and Bosetti and Maffezzoli (2013), which also examine the effects of green policies in heterogeneous agent economies. Shapiro and Walker (2018) use a Melitz-style model to examine the effects of environmental regulation. Bosetti and Maffezzoli (2013) use a standard incomplete markets model with idiosyncratic uncertainty to examine the distributional effects of a carbon tax on households. In contrast to these studies, our framework links the distribution of plants to an endogenous technology adoption decision. Finally, our work is related to the par-

[4]Cooley et al. (1997) study the balanced-growth path and the transitional dynamics of a deterministic model with two sectors and vintage capital and compare it with the neoclassical growth model.

[5]Nordhaus (1994, 2008), Nordhaus and Boyer (2000), Hassler and Krusell (2012) formulate large-scale models to determine optimal carbon prices. Optimal carbon taxes also are considered in Acemoglu et al. (2016) and Golosov et al. (2014) and references therein.

[6]This line of research also relates to the weak version of the Porter (1991) hypothesis, where establishments can respond to a carbon tax by altering their vintage technology.

tial equilibrium analysis of Heutel (2011), which studies an individual vintage plant's response to environmental regulation.

Our paper complements the macroeconomic literature examining how the distribution of heterogeneous firms can alter the aggregate effects of a policy. Restuccia and Rogerson (2008) argue that policies can alter the allocation of resources across producers that have heterogeneous productivity. Miao and Wang (2014) show that the aggregate implications of a corporate tax change depend on how many firms have recently made significant capital adjustments. Our model provides a framework that can be adapted to study these alternative tax policies and assess how production inputs—including total factor productivity, investment, and energy—endogenously respond to policy measures.

Outline The rest of the paper is organized as follows. Using micro data, Section 2 documents that emission rates heterogeneity is related to the age of the equipment available to plants'. Section 3 presents the model, and Section 4 discusses the model's calibration to fit the microeconomic evidence. In Sections 5 and 6, we study the effects of introducing a carbon tax and demonstrate how the effects of a carbon tax depend on the distribution of plant emission rates. Section 7 discusses the sensitivity of our results to various modeling assumptions, and Section 8 concludes.

2 Emission Rates and Equipment Age

In this section, we document how emission rate heterogeneity is related to the timing of capital accumulation. Toward this goal, we use data for U.S. electricity plants to estimate the elasticity of emission rates to age. This elasticity is significantly positive for almost all types of emissions and primary fuel inputs considered, indicating the presence of vintage technology effects.

2.1 Data Sources

The data for power plants comes from two sources. First, the Energy Information Administration's EIA-860 survey contains information about every power plant in the U.S., including age, capacity, and some abatement choices for the generators of a plant. Second, the EPA's Emissions and Generation Resource Integrated Database (eGRID) provides data on emissions and primary fuel sources for all power plants. We merge the EIA and eGRID data at the plant level using a unique plant identifier. Our data coverage is for the years 2004, 2005, 2007, 2009, and 2010.[7] We estimate an emissions function at the plant level.

To create a plant-level age, we take the weighted average of the age of each plant's generators, where we use the summer capacity of each generator for weights.[8] Figure 1 provides a snapshot of this data by plotting the coal power plant's emission rates by age in 2010 and suggesting a negative relationship between the age of plant equipment and emissions exists.

2.2 Empirical Analysis

We follow the approach in Heutel (2011) to quantify the relationship between emissions and the age of plant equipment. We estimate via ordinary least squares (OLS) the following specification:

$$\ln emr_{it} = \alpha + \beta \ln age_{it} + X_{it} + \epsilon_{it}, \quad (1)$$

where *emr* is the plant's output emission rate (measured in pounds per megawatt), *age* is the weighted average of the plant generator's years, and X_{it} denotes controls for the type of primary fuel input and state and year fixed effects. We estimate this function for

[7] eGRID data for 2006 and 2008 are not available.

[8] The Energy Information Administration has employed similar weights in its publications of electric power generators age. Our results are robust to alternative weighting using winter capacity of generators, an average of winter and summer capacity, or by a generator's net electricity generation (for cases where this information is reported).

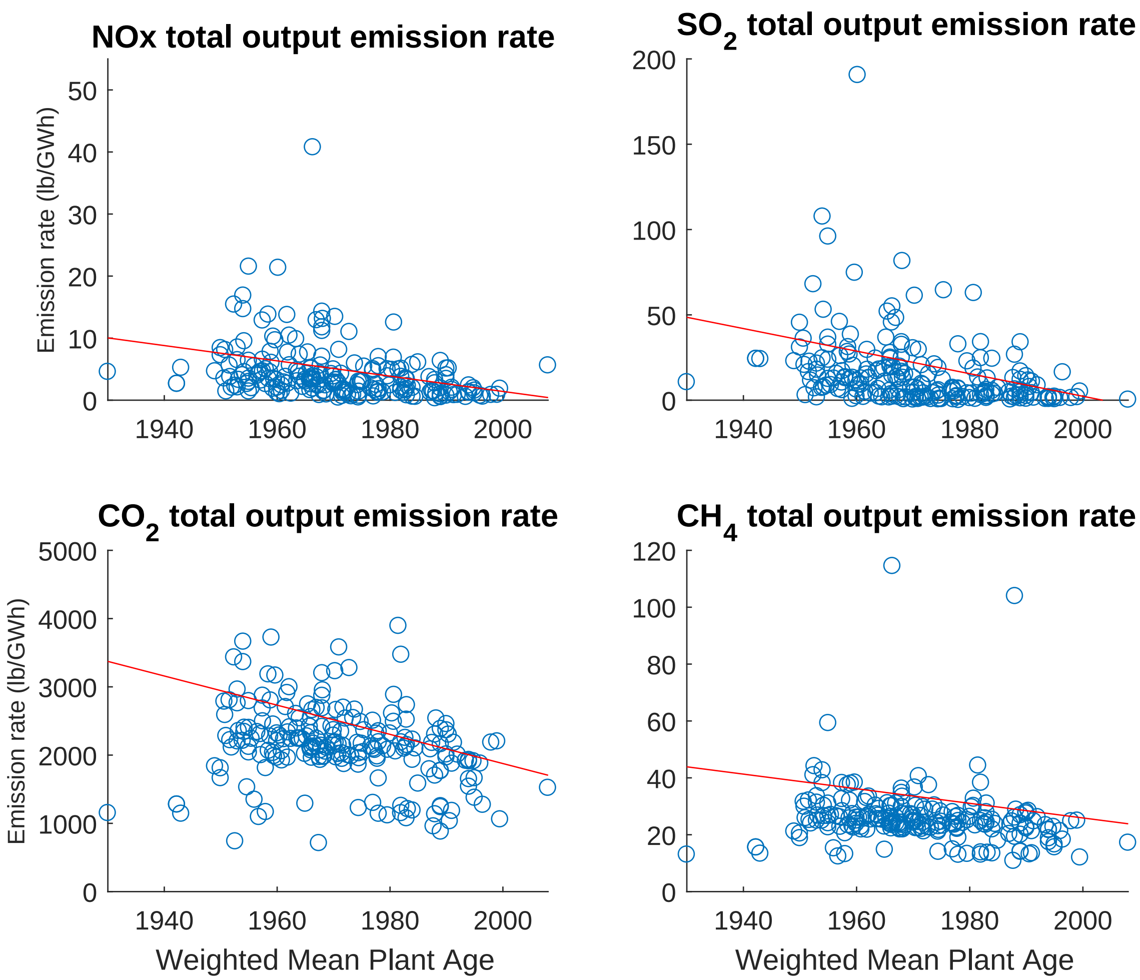

Figure 1: Emission Rates

Notes: Emission Rates for coal power plants in the U.S by plant age for 2010. The plant age is constructed as the weighted mean of the plant's generators' age, weighted by each generators net electricity generation. Each scatter plot is accompanied by its linear fit.

four distinct emissions: sulfur dioxide SO_2, nitrous oxide NO_x, carbon dioxide CO_2, and methane CH_4. Our coefficient of interest is β.

Table 1 reports the estimation results for plants with coal as the primary fuel input separately from alternative fuel inputs. Unreported state and year dummies, as well as controls for fuel input types, are included in all regressions to account for unobserved heterogeneity and time effects. In almost every specification, the elasticity of emissions with respect to age is positive and significant at the 1% level. The elasticity for sulfur

Table 1: Emissions Function Estimates

	Non-coal			
	SO_2 rate	NO_x rate	CH_4 rate	CO_2 rate
ln *age*	0.567***	0.574***	0.103***	0.107***
constant	-1.721***	1.170***	4.682***	6.698***
observations	10,559	12,165	9,384	11,423
Adj R-squared	0.71	0.57	0.65	0.45
	Coal			
	SO_2 rate	NO_x rate	CH_4 rate	CO_2 rate
ln *age*	0.907***	0.438***	0.045	0.059**
constant	-1.737***	-0.002	2.909***	7.154***
observations	2,809	2,809	2,193	2,809
Adj R-squared	0.28	0.22	0.40	0.44

Notes: Dependent variable is the plant's output emission rate (measured in pounds per megawatt), age is the weighted average of the plant generator's years online as of 2010. Each specification includes controls for the type of primary fuel input and state and year fixed effects. Asterisks denote (*) 10%, (**) 5% or (***) 1% significance. Each specification is estimated using ordinary least squares (OLS).

dioxide to age is the highest, estimated to be almost unitary for coal power plants (0.91). The elasticity for nitrous oxide also is quantitatively significant. The elasticity for carbon dioxide is lower (0.11 and 0.06 for non-coal and coal plants) but is still significantly higher than zero.

Our results point to the importance of vintage effects: Part of the observed heterogeneity in emission rates is endogenous to the firms' behavior and attributable to the timing of capital installation. Our evidence supports the ideas put forth in Johansen (1959) and Solow (1960)—newer vintages of better quality result in fewer emissions.[9]

[9] A large literature emphasizes the role of investment as a vehicle to introduce new technology in the production process. Using detailed firm-level Italian data, Fiori and Scoccianti (2018) provide evidence on the role of capital accumulation for productivity dynamics at the firm-level. Sakellaris (2004) and Power (1998) investigate the same question using U.S. plants while Licandro et al. (2005) use Spanish data. Using

To quantify the aggregate effects of carbon taxes, we develop a general equilibrium model that links emission heterogeneity to equipment age using the discipline imposed by the estimates in Table 1. We turn to this issue next.

3 Theoretical Framework

To study the effects of a tax on emissions we develop a general equilibrium framework of production heterogeneity and technology adoption that accounts for the link between the equipment age and the heterogeneity in emissions rates discussed in Section 2.

As in Solow (1960), firms introduce the latest technology in the production process by acquiring new equipment. The key innovation in our setup consists of the explicit modeling of the choice of technology upgrade through capital accumulation.[10]

Our framework builds on the investment literature that studies the role of plant-level non-convexity in a general equilibrium framework. Notable examples include Thomas (2002), Khan and Thomas (2008), and Bachmann et al. (2013).

In the model, old and new technologies coexist as upgrading to the latest vintage is subject to a non-convex adjustment cost. This assumption leads firms to pursue (S,s) policies, which allow the model to reproduce the pattern of capital accumulation at the plant level discussed by Doms and Dunne (1998).[11] Moreover, the model accounts for the evidence in Fiori and Scoccianti (2018) where firms with newer equipment are more productive than firms with older ones.

product data, similar evidence is provided by Gordon (1990) and using aggregate data by Greenwood et al. (2000).

[10]It is worth noticing that, because of the C.E.S. production function, the dynamics of our model cannot be represented by a stand-in aggregate firm as in Solow (1960) and Greenwood et al. (2000). Thus, the distribution of capital stocks and vintage technology enters the state space of the model.

[11]The lumpy nature of the process of capital accumulation is a feature of the data in many countries. Haltiwanger et al. (1999) provide further evidence for the U.S. economy, Bachmann and Bayer (2014) for Germany, Licandro et al. (2005) for Spain, Nilsen and Schiantarelli (2003) for Norway, and Gourio and Kashyap (2007) for Chile.

In the model, the introduction of a carbon tax impacts the timing of technology adoption and capital accumulation. The age of equipment affects the mapping between energy usage and emissions and, in turn, the burden of the carbon tax for a given plant. Thus, a carbon tax alters the economic incentives that determine the plant's optimal size and its decision to upgrade technology through capital accumulation.

We describe the economic environment in which individual plants operates in Sections 3.1, 3.2, 3.3, and 3.4. We then focus on the aggregation of the plants' decisions in Section 3.5. In Sections 3.6 and 3.7, we describe the households' and the government's problems, respectively. Sections 3.8 and 3.9 discuss the equilibrium of the model. In Section 3.10, we show that our model nests as a special case the neoclassical growth model with a C.E.S. production function.

3.1 Production at the Plant Level

There is a continuum of units that produce a homogenous good using as inputs capital, labor, and energy. Labor and energy can be freely adjusted in each period while upgrading technology and the capital stockd is costly.

Output is produced by perfectly competitive plants via an increasing and concave production function F defined over capital (K), labor (L), energy (E) and the technology available to the firm (Z). As in Krusell and Smith (2015), F takes the form of a constant elasticity of substitution (CES) production function that aggregates energy E and a Cobb-Douglas composite consisting of capital K and labor L. The production function is given by

$$Y_{j,t} = \left[\omega^{\frac{1}{\epsilon}} \left(Z_{j,t} K_{j,t-1}^{\gamma} L_{j,t}^{\nu} \right)^{\frac{\epsilon-1}{\epsilon}} + (1-\omega)^{\frac{1}{\epsilon}} (Z_{j,t} E_{j,t})^{\frac{\epsilon-1}{\epsilon}} \right]^{\frac{\epsilon}{\epsilon-1}} \quad \forall j = 0 \ldots J, \tag{2}$$

where $\gamma + \nu < 1$, ϵ is the price elasticity of energy and $1 - \omega$ is a quasi-share parameter for energy. The efficiency of inputs of production depends upon Z, the technology vintage commonly acquired by all plants that chose to upgrade their capital in period j. Thus, the subscript j denotes a production unit that has invested and adopted the newest technology j periods ago. By paying the adjustment cost, plants can adopt the latest technology and optimally choose their size. The leading vintage ($j = 0$) of technology, $Z_{0,t}$, evolves deterministically, increasing every period at the gross growth rate of Θ_Z. Subsequent vintages inherit the level of technology from the last time of capital replacement: $Z_{j,t} = Z_{0,t-j}$.

In every period, the plant chooses its current level of employment and energy usage, produces output, and pays labor, energy and emission taxes. We indicate with W the real wage and with P^e the real price of energy.

3.2 Energy Usage and Emissions

Plants are subject to a carbon tax, τ_t^{em}, on their total emissions. Emissions (EM) are proportional to energy usage, but potentially vary with plant vintages: $EM_{j,t} = \Omega_j E_{j,t}$. The standard assumption in the literature is to assume units of emissions such that $\Omega_j = 1 \; \forall \; j$. The empirical evidence presented in the previous section indicates that emission rates vary with equipment age. In light of this, we consider the case where Ω_j satisfies $\frac{\partial \Omega_j}{\partial j} > 0$, so that older plants have lower levels of productivity and higher emission rates.[12] In Section 4.1 we discuss the parameterization of the emission function.

3.3 Vintage Upgrading Decision at the Plant Level

Upgrading technology and the capital stock is costly and is subject to a non-convex cost denominated in units of labor. As in Thomas (2002), these adjustment costs, denoted by

[12] Recent micro evidence suggests that emissions vary with other plant characteristics. Shapiro and Walker (2018) find that U.S. establishments with higher total factor productivity have lower emission intensities. Bloom et al. (2010) find that better managed U.K. establishments have lower energy intensities and CO_2 emissions. Several studies have found that, within developing countries, small plants have higher emissions intensities than large plants, see for instance Dasgupta et al. (2002) and Qi et al. (2019).

ξ, are independently and identically distributed across establishments and across time with a known cumulative distribution of $G(\xi)$ and finite upper support B. The fixed cost is denominated in units of labor to ensure that plants cannot outgrow adjustment costs along the balanced growth path, as is common in the literature (see for instance Thomas, 2002).

In each period, plants receive a current realization of the fixed adjustment cost. After observing this realization, plants decide whether to upgrade their vintage technology and choose the optimal size. If a plant decides to pay the fixed cost, ξW_t, in units of labor, in the following period it upgrades its efficiency to the latest vintage $Z_{0,t+1}$, and chooses the optimal size $K_{0,t+1}$:

$$Z_{0,t+1} = Z_{0,t}\Theta_{Z_0} \; \forall j = 0...J, \tag{3}$$

$$K_{0,t+1} = (1 - \delta - \delta_j^s)K_{j,t} + I_{j,t} \; \forall j = 0...J, \tag{4}$$

where δ is the physical rate of depreciation of capital and δ_j^s measures capital obsolescence. δ_j^s captures the notion that when an investing plant upgrades its technology, part of its current capital stock may not be compatible with the new technology.

If an establishment postpones upgrading its technology, it keeps its current vintage Z. It is worth noting that the gap between the current Z available to the firm and the frontier Z_0 measures the degree of technological obsolescence. The stock of capital next period is the depreciated current level of capital:

$$Z_{j+1,t+1} = Z_{j,t} \; \forall j = 0...J, \tag{5}$$

$$K_{j+1,t+1} = (1 - \delta)K_{j,t} \; \forall j = 0...J. \tag{6}$$

Each plant's current flow of profit is determined by its total revenues less wage payments, investment, energy usage and emissions taxes, and adjustment costs. Given diminishing returns, plants make profits that are rebated to households in lump-sum fashion.

3.4 Technology Dynamics

The level of input efficiency $Z_{0,t}$ evolves deterministically over time at the rate $\Theta_{Z_{0,t}}$:

$$Z_{j,t} = Z_{0,t-j} = \frac{Z_{0,t}}{\prod_{s=t-j}^{0} \Theta_{Z_{0,s}}}. \tag{7}$$

As a result, delaying upgrading technology increases the distance from the technological frontier.

The aggregate variables of the economy inherit the deterministic trend in technology. A stationary representation of the model implies the following restrictions in the growth rates of individual variables: $\Theta_{Y,t} = \Theta_{C,t} = \Theta_{G,t} = \Theta_{W,t} = \Theta_{K,t} = \Theta_{Z_{0},t}^{\frac{1}{1-\gamma}}$, $\Theta_{E,t} = \Theta_{Z_{0},t}^{\frac{\gamma}{1-\gamma}}$, $\Theta_{P^e,t} = \Theta_{Z_0,t}$.

3.5 Plant Aggregation

All plants that choose to invest obtain in the next period the leading vintage of technology and face the same distribution of adjustment costs. As a result, they share the same expected stream of future marginal revenues for any given choice of future capital. Thus, investors obtaining $Z_{0,t+1}$ choose a common size with target capital $K_{0,t+1}$, and plants adjusting at a given time are identical immediately after investing. The cross-sectional distribution of establishments over capital and technology levels is therefore summarized by the distribution of plants across j, i.e. *time-since-adjustment*. Each vintage shares the same time since the last capital adjustment and has the same capital stock, productivity, and energy-emissions linkage (i.e., energy efficiency).

Because of the continuity of the adjustment cost distribution, the choice of whether to adjust follows a reservation policy. Within each vintage, there exists a marginal plant whose draw of the fixed cost makes it just worthwhile to invest, i.e. whose fixed cost implies the benefit and cost of investment are equal. All plants of the same group that draw costs at or below this group-specific threshold also invest, implying that the investing fraction for each group, $\alpha_{j,t}$, is retrievable from the fixed cost's cumulative density function (cdf).

At each date, the cross-sectional distribution of time-since-adjustment can also be represented using alternative dimensions of heterogeneity. The vectors $\mathbf{K}_t = \{K_{j,t}\}$ and $\mathbf{Z}_t = \{Z_{j,t}\}$ summarize the heterogeneity in capital holdings and input efficiency. The fraction of plants associated with each group is given by the predetermined vector $\boldsymbol{\vartheta}_t = \{\vartheta_{j,t}\}$. Each $\vartheta_{j,t}$ describes the number of firms owning vintage j capital stock and technology. The evolution of the distribution is determined as follows. Let $\boldsymbol{\alpha}_t = \{\alpha_{j,t}\}$ denote the fraction of adjustment rates. The evolution of the distribution of plants $\boldsymbol{\vartheta}_{t+1}$ is determined by the following equations:

$$\vartheta_{0,t+1} = \sum_{j=0}^{J} \alpha_{j,t}\vartheta_{j,t}, \tag{8}$$

$$\vartheta_{j+1,t+1} = (1-\alpha_{j,t})\vartheta_{j,t} \;\; \forall j = 1, 2...J. \tag{9}$$

The group of plants that have upgraded their technology in the current period is the weighted sum of adjusters in each group. If a plant decides not to adjust at date t, it becomes vintage $t+1$ in the subsequent period.

The average efficiency of inputs in the economy is the weighted average of the individual efficiency across plants and depends on the upgrading choice observed at the plant-level. The total stock of capital in the economy is the weighted sum of the stock of capital across plants.

3.6 Households

The economy features a continuum of identical households who have access to a complete set of state-contingent claims. As there is no heterogeneity across households, these assets are in zero net supply in equilibrium.[13]. Moreover, they own shares in the portfolio of plants and supply labor. Each household values consumption and leisure and maximizes the intertemporal utility function:

$$\sum_{t=0}^{\infty} \beta^t \left[\ln C_t - \psi \frac{L_t^{1+\eta}}{1+\eta} \right]. \tag{10}$$

Consumption is taxed at the rate τ^c, and financed by labor income, taxed at the rate τ^l, lump-sum transfers from the government, TR, and profits received from the plants, Π. The household's flow budget constraint is given by

$$(1+\tau_t^c)C_t = (1-\tau_t^l)W_t L_t + TR_t + \Pi_t. \tag{11}$$

3.7 Government

The government's budget constraint is given by

$$G_t + TR_t = \tau_t^l W_t L_t + \tau_t^c C_t + \tau_t^{em} P_t^e EM_t, \tag{12}$$

where G_t denotes government expenditure and EM_t is aggregate emissions. Fiscal instruments are set exogenously. In our baseline experiments, when the government legislates a carbon tax ($\tau^{em} > 0$), the lump-sum transfer TR_t is assumed to adjust to ensure a balanced budget (i.e. that equation (12) holds). In addition, we consider alternative

[13]Since these assets are in zero net supply, we refrain from explicitly including them in the budget constraint

"revenue-recycling" schemes, allowing either the labor tax rate, consumption tax rate, or level of government expenditures to adjust to satisfy the government budget constraint following the carbon tax introduction.

3.8 Aggregate Variables

The cross-sectional distributions of $\mathbf{K}_t$, $\mathbf{Z}_t$, and $\boldsymbol{\vartheta}_t$ enter the state space of the model and determine aggregate variables in the economy. Because the model does not admit an aggregate production function, total output, energy usage, emissions and the capital stock are the weighted sum of the respective plant-level variables:

$$Y_t = \sum_{j=0}^{J} \vartheta_{j,t} Y_{j,t}, \quad E_t = \sum_{j=0}^{J} \vartheta_{j,t} E_{j,t}, \quad EM_t = \sum_{j=0}^{J} \vartheta_{j,t} \Omega_j E_{j,t}, \quad K_t = \sum_{j=0}^{J} \vartheta_{j,t} K_{j,t}.$$

Aggregate investment is determined by the weighted sum of plant-level investment, where the weights are given by the fraction of plants that decide to invest ($\vartheta_{j,t}\alpha_{j,t}$):

$$I_t = \sum_{j=0}^{J} \vartheta_{j,t}\, \alpha_{j,t} I_{j,t}.$$

The economy is subject to a set of aggregate constraints. Energy is imported from abroad at a given international price P_t^e, and its supply is perfectly elastic. Trade is balanced by assumption: energy imports are financed by final good exports.[14] Thus, the sum of household consumption, aggregate investment and the cost of energy cannot exceed total production:

$$C_t + G_t + I_t + P_t^e E_t \leq Y_t. \tag{13}$$

Total hours worked by the household must satisfy the weighted sum of employment in

[14]This modeling convention is standard and has a long history in the literature, see Hassler and Krusell (2012) for a recent example.

production and adjustment activities in each sector:

$$\sum_{j=0}^{J} \vartheta_{j,t} \, L_{j,t} + \sum_{j=0}^{J} \vartheta_{j,t} \, \Xi(\alpha_{j,t}) \leq L_t, \tag{14}$$

where the average adjustment cost for each group is defined as

$$\Xi(\alpha_{j,t}) = \int_0^{G^{-1}(\alpha_{j,t})} x dG(x). \tag{15}$$

3.9 Stationary Equilibrium

In this section, we report the critical equations of the model in stationary form. We denote by lower case variables that are in deviation from their trend. The production function expressed in stationary form is:

$$y_{j,t} = \left[\omega^{\frac{1}{\epsilon}} \left(z_{j,t} \frac{k_{j,t-1}}{\Theta_K}^{\gamma} L_{j,t}^{\nu} \right)^{\frac{\epsilon-1}{\epsilon}} + (1-\omega)^{\frac{1}{\epsilon}} (z_{j,t} e_{j,t})^{\frac{\epsilon-1}{\epsilon}} \right]^{\frac{\epsilon}{\epsilon-1}} ; \; \forall j = 0 ... J. \tag{16}$$

$z_{j,t}$ is the level of efficiency available to the firm relative to the technological frontier, and it measures the technological gap of the plant.[15] A competitive equilibrium satisfies the following efficiency conditions derived from the profit optimization of each plant. We first discuss the optimality conditions for labor and energy, the inputs of production that can be freely adjusted every period. Each plant equates the cost of each input to its marginal product (denoted by MPL and MPE for labor and energy, respectively). Plant-level employment satisfies the static condition for labor, for $j = 0, ..., J$:

$$w_t = MPL = \nu (\omega y_{j,t})^{\frac{1}{\epsilon}} \left(\frac{1}{\Pi_{s=t-j}^{0} \Theta_{Z_0,s}} \right)^{\frac{(\epsilon-1)}{\epsilon}} \Theta_{K,t}^{-\frac{\gamma(\epsilon-1)}{\epsilon}} k_{j,t-1}^{\frac{\gamma(\epsilon-1)}{\epsilon}} L_{j,t}^{\frac{\nu(\epsilon-1)}{\epsilon}-1}. \tag{17}$$

[15] In the stationary representation of the model $z_{j,t}$ is the ratio between the current technology of the firm, $Z_{j,t}$, and the technological frontier, $Z_{0,t}$.

Plant-level energy usage satisfies, for $j = 0, ..., J$:

$$p^e(1 + \tau_t^{em}\Omega_j) = MPE = \frac{1}{\Pi_{s=t-j}^{0}\Theta_{Z_0,s}} \left[\frac{(1-\omega)y_{j,t}}{\frac{e_{j,t}}{\Pi_{s=t-j}^{t}\Theta_{Z_0,s}}} \right]^{\frac{1}{\epsilon}}. \tag{18}$$

Although plants face the same price of energy p^e, the *effective* price that determines the energy usage depends upon the mapping between energy usage and emissions. When Ω_j is increasing in j, the effective cost of energy increases with the plant's distance from the technological frontier. For a given value τ_t^{em}, plants with older technology face a steeper tax schedule. As a result, the introduction of a carbon tax reduces the size of a plant, if energy is an input of production that is less than perfectly substitutable by labor or capital. This is the familiar *static* effect of a tax.

The conditions for the optimal capital allocation states that the marginal cost of installing capital (which is one) equals the expected marginal value of installing an additional unit of capital:

$$\mu_{0,t} = 1, \tag{19}$$

which for each $j = 0, ..., J$ is defined by:

$$\mu_{j,t} = \mathbb{E}_t \left\{ \frac{\beta_{t,t+1}}{\Theta_{C,t+1}} \left[\begin{array}{c} \gamma(\omega y_{j,t+1})^{\frac{1}{\epsilon}} \left(\frac{1}{\Pi_{s=t+1-j}^{0}\Theta_{Z_0,s}} \right)^{\frac{(\epsilon-1)}{\epsilon}} \Theta_{K,t+1}^{-\frac{\gamma(\epsilon-1)}{\epsilon}} k_{j,t}^{\frac{\gamma(\epsilon-1)}{\epsilon}-1} L_{j,t+1}^{\frac{\nu(\epsilon-1)}{\epsilon}} \\ +(1-\delta-\delta_j^s)\alpha_{j,t+1} + (1-\delta)(1-\alpha_{j,t+1})\mu_{j+1,t+1} \end{array} \right] \right\}. \tag{20}$$

The introduction of the carbon tax decreases the marginal product of capital for a given value of k_j reducing the optimal size of the plant.

We now turn to characterize the upgrading decision of the plant that, in turn, determines the hazard rates $\boldsymbol{\alpha}$, and thus the cross-sectional distribution represented by the vector $\boldsymbol{\vartheta}$.

The finite upper support for the fixed cost cdf, combined with the increasing technological gap, makes upgrading increasingly valuable across vintages and, in turn, capital adjustment more likely. As in Thomas (2002), this simplifies the state space of the model, as the economic history is redundant beyond a finite number of lags. Once the gap from the technological frontier has increased enough, the value of investing offsets the highest possible fixed cost. In equilibrium, there exists an endogenously chosen vintage J at which full adjustment occurs: $\alpha_J = 1$. For $j < J$, the optimal fractions of adjusting are interior solutions equating the anticipated value of adjusting one additional plant from group j to the additional costs entailed, consisting of the investment required and the adjustment cost $w_t G^{-1}(\alpha_{j,t})$. That is, for each $j = 0, ..., J-1$:

$$v_{0,t} - v_{j+1,t} = i_{j,t} + w_t G^{-1}(\alpha_{j,t}), \tag{21}$$

where

$$v_{j,t} = \mathbb{E}_t \left\{ \beta_{t,t+1} \left[\begin{array}{c} y_{j,t+1} - w_{t+1}(L_{j,t+1} + \Xi(\alpha_{j,t+1})) - p^e(1 + \tau^{em}_{t+1}\Omega_j)e_{j,t+1} \\ -\alpha_{j,t+1}(i_{j,t+1} - v_{0,t+1} + v_{j+1,t+1}) + v_{j+1,t+1} \end{array} \right] \right\}, \tag{22}$$

where $\beta_{t,t+1}$ is equal to the stochastic discount factor $\beta\frac{\lambda_{t+1}}{\lambda_t}$. $v_{j,t}$ represents the expected value of a plant with capital $k_{j,t}$ in the next period. There is a marginal firm for which equation (21) holds with equality that identifies $\alpha_{j,t}$, the fraction of plants for the given vintage j that decides to invest, which can be interpreted as the hazard rate. $\alpha_{j,t}$ is increasing in j, the number of periods of inaction of a plant. The higher the gap between the productivity of the latest vintage and its current vintage, the higher the probability that a plant invests. The latter prediction holds in the data, see Haltiwanger et al. (1999). It is worth noticing that the introduction of a carbon tax has a *dynamic* effect in that it impacts both the value of upgrading/postponing investment. A higher effective carbon tax makes a given vintage j, more likely to invest in new technology to escape the tax burden, especially when the emission rate is increasing in j.

Finally, household optimization implies that the marginal rate of substitution between leisure and consumption equates to the real wage w_t:

$$(1 - \tau_t^l)w_t = \psi L_t^\eta (1 + \tau_t^c)c_t. \tag{23}$$

3.10 Representative Firm Model

Our model nests as a special case the standard one-sector neoclassical growth model with a C.E.S. production function defined over capital, labor and energy. When the upper support of the distribution of the idiosyncratic shock B is set to zero, all plants can continuously adopt the latest technology and adjust their stock of capital in each period. As a result, the dynamics of the economy are characterized by a representative firm. We briefly outline this version of the model, as we find it useful to compare the results of our environmental policy experiments below to this representative agent benchmark. A representative household and the government behave as in sections 3.6 and 3.7. Emissions are assumed to be proportional to energy usage: $EM_t = \Omega E_t$. The representative firm maximizes expected, discounted profits given by

$$\sum_{t=0}^{\infty} \beta_{t,t+1} \left[Y_t - W_t L_t - P_t^e (1 + \tau_t^{em} \Omega) E_t - K_{t+1} + (1 - \delta) K_t \right], \tag{24}$$

where $\beta_{t,t+1} = \beta^t \frac{\lambda_t}{\lambda_0}$ is the stochastic discount factor of the household, and output is defined as in (2). Appendix A lists the stationary equilibrium conditions for this version of the model.

4 Model Parameterization

In this section, we parameterize the model to U.S. data. To quantify the aggregate effects of introducing a carbon tax, there are two critical aspects of the model to measure. First, the model needs to reproduce the salient features of the pattern of capital accumulation at the micro-level; through investment, plants upgrade their technology and potentially reduce their tax burden. Second, the model needs to be consistent with the plant-level evidence on the relationship between emission rates and vintage technology available to a firm, as discussed in Section 2. We focus our discussion in this section on these two key elements.

4.1 Parameterization

Following the business cycle literature, we calibrate the model to fit key first-order moments of the U.S. economy. Table 2 summarizes parameter values, targeted moments, and data sources. We assign values to 15 parameters related to individual preferences (β, ψ, and η), characteristics of the production process (γ, ν, ϵ, ω, δ, δ_j^S, Θ_{Z_0}, B and Ω), and government policies (G/Y, τ^l and τ^c).

One period in the model represents one quarter. The discount factor β is set to 0.99 to target a real annual interest rate of 4%. As in King and Rebelo (1999), the preference parameter that governs the disutility of labor (ψ) is chosen to target a steady-state labor effort of 0.2. We select the Frisch elasticity of the labor supply to 1, the upper bound of the recent estimates discussed in Keane and Rogerson (2015).

The world price of energy in steady state is normalized to one, and the energy price elasticity ϵ is set to 0.25 to reflect the low short-run substitutability of energy with capital and labor inputs estimated by Kilian and Murphy (2014).[16] We calculate the ratio of the

[16]We take this value as representative of the estimates in the literature for the short-run price elasticity of

U.S. average energy usage by firms to GDP over the period 1970-2012 to be 0.05.[17] Based on this, we calibrate ω to imply steady state energy usage is 5% of gross output.

The rate of capital depreciation (δ) is set to 0.015, a standard choice in this literature (see Fiori (2012) for a discussion). In absence of empirical guidance, for our benchmark calibration we set $\delta_j^s = 0$ for all j, and explore the implications of having $\delta_j^s > 0$ in Section 7.2.

The long-run growth rate of technology is set to 1.0025 to target the yearly average growth rate of output of 1.6% observed in the United States during the post-WWII sample.

We set steady-state fiscal parameters to match post-WWII averages in the United States. The government consumes 20% of output in steady state, roughly equivalent to the average U.S. total government consumption to GDP ratio. The consumption tax rate is 0.07, reflecting an average rate across all levels of U.S. government. The labor tax is set to 0.2, which equals the average labor income tax rate calculated using the method of Jones (2002) (see Leeper et al. (2010) for more details).

We now turn to the parameterization of the emission function and the non-convex adjustment cost. We discipline the emission function, that links energy usage (E_j) to emissions EM_j, $EM_j = \Omega_j E_j$ using the empirical estimates in Section 2. We note that all that is needed is $\frac{\partial \Omega_j}{\partial j}$, i.e. the slope. Since we did not find evidence of non-linear effects of equipment age for the emission rates in the data, we choose a constant $\frac{\partial \Omega_j}{\partial j}$. Moreover, the intercept of the emission function is inconsequential, as it amounts to a rescaling of the carbon tax rate. We set $\frac{\partial \Omega_j}{\partial j}$ equal to 0.5, to agree with the midpoint of the estimated

oil demand, see Kilian and Murphy (2014) and references therein.

[17] Table 3.6 of the Energy Information Administration's (EIA) Annual Energy Review (AER) provides annual nominal expenditures on energy products, defined as the total of oil, gas, and electricity, by residential, commercial, industrial, and transportation sectors. We define total firm spending on energy products as the sum of commercial and industrial spending on energy products, plus spending in the transportation sector on energy products due to firms.

Table 2: Benchmark Calibration

Parameter		Value	Target
Discount Factor	β	0.99	Annual real interest rate = 4%
Disutility of Labor	ψ	21.3	Steady state labor= 0.2
Inverse of Frisch elasticity	η	1	Literature
Production function parameter	γ	0.32	Literature
Production function parameter	ν	0.58	Labor share
Energy price elasticity	ϵ	0.25	Kilian-Murphy (2014)
Production distribution parameter	ω	0.95	Energy/Output = 5%
Depreciation Rate	δ	0.015	Literature
Capital obsolescence	δ_j^s	0	Authors' choice
Growth rate of productivity	Θ_{Z_0}	1.0025	Data
Gov spending-output ratio	G/Y	0.2	Data
Labor tax rate	τ^l	0.20	Data
Consumption tax rate	τ^c	0.07	Data
Upper support adj. cost distribution	B	0.075	Fraction of lumpy investors
Emission function	$\frac{\partial \Omega_j}{\partial j}$	0 or 0.5	Data

coefficients. In addition, we consider the case when $\frac{\partial \Omega_j}{\partial j} = 0$, reproducing the standard assumption in the literature that emissions do not vary with vintage.

We rely on the investment literature to discipline with micro evidence the distribution for the adjustment cost. As in Thomas (2002), the cumulative distribution for adjustment costs is set to imply a uniformly distributed cost between 0 and B. The upper support of the distribution (B) is set to agree with evidence on investment spikes reported by Doms and Dunne (1998). In the average year, only 8% of plants raise their capital stocks by 30% or more, giving rise to lumpy investors. Setting $B = 0.075$ matches the fraction of lumpy investors, while their investment activities in the model account for 31% of aggregate investment (25% in the data).[18] The dashed line of Figure A.1 reports the steady-state distribution of firms $(\boldsymbol{\vartheta})$, while the dotted line reports the hazard rates (α_j). The maximum time-since-adjustment J is equal to 35; in other words, the maximum time-

[18]Given the quarterly calibration of the model, the target for lumpy investors is 2% at a quarterly frequency: plants that are lumpy investors in period t (i.e. they experience an investment rate above 30%) along the balanced growth path will likely be lumpy investors again beyond period $t+4$.

since-adjustment for a plant is about 10 years. The resulting steady-state distribution is reported in Figure *A*.1. The algorithm employed to solve the model adapts the one in Thomas (2002) and is described in Appendix B.

5 The Aggregate Effects of Carbon Taxes

We use our theoretical framework to study the macroeconomic effects of introducing a carbon tax on emissions in the U.S. economy. In this section, we show the importance of accounting for plant heterogeneity in emission rates to quantify the aggregate effects of the carbon tax. In Section 5.1, we discuss the details of our experiment, while in Section 5.2, we characterize the dynamic responses of the economy.

5.1 Introducing a Carbon Tax

The government introduces in the current period a carbon tax, so that $\tau^{em} > 0$, which is known to be permanent. At the same time, the government announces how carbon tax revenues will be utilized. For our benchmark experiments, the revenues are assumed to be rebated to the households in lump-sum fashion. Alternative financing scenarios are discussed in Section 6.2. In all the policy experiments that follow, for comparability we calibrate the carbon tax to achieve a long term reduction in emissions of 1%. Following the approach in Gavazza et al. (2018), our experiments characterize the transitional dynamics of the model under perfect foresight. We assume that the economy starts in a steady state with no carbon tax, i.e. $\tau_{em} = 0$. Thus, the steady state of the model is independent of the assumption about emission rates. At time one, the economy is surprised by the government enacting a carbon tax. As this policy action is permanent, eventually the economy moves to a new steady state.

5.2 Firm and Emission Rates Heterogeneity

We start by discussing the dynamics obtained when $\frac{\partial \Omega_j}{\partial j} = 0$, that is when all plants' emissions are equal to energy usage, the common assumption in the literature. This case is useful as it allows us to disentangle the role of firm heterogeneity and emission rates heterogeneity. We assume that tax revenues are rebated to the households in lump-sum fashion. Figure 2 shows the impulse responses. The responses of output, emissions, investment, the target capital stock, the number of plants investing (i.e. investors), and the average efficiency in the economy are plotted as percentage deviations from the initial steady state.[19]

As emissions are tied to energy usage, the introduction of a carbon tax effectively increases the factor price of energy. Given the low substitutability between energy and other inputs of production, the representative firm finds it optimal to reduce its size by lowering its capital stock. Across all horizons, output, emissions, and the target capital stock decrease. These factors create opposing incentives on a plant's capital accumulation decision. On the one hand, a higher effective energy price makes investment in technology and capital with lower emission rates desirable. On the other hand, a higher factor price reduces the size of producers and depresses the expected value of a plant at the technological frontier ($v_{0,t}$ in equation 22), which reduces investment incentives. The responses in Figure 2 show that the latter effect dominates in the short run. On impact, plants delay investment, and the number of investors decreases. Since fewer establishments are upgrading their capital and technology, the average energy efficiency in the economy declines and the average age of the capital stock (not pictured) increases. Over time, the number of investors rises as plants that have delayed investment and whose capital stock is depreciating find it more profitable to replace capital. In the long run, a

[19]The carbon tax necessary to achieve a 1% reduction in energy in the long-run is 3.30%.

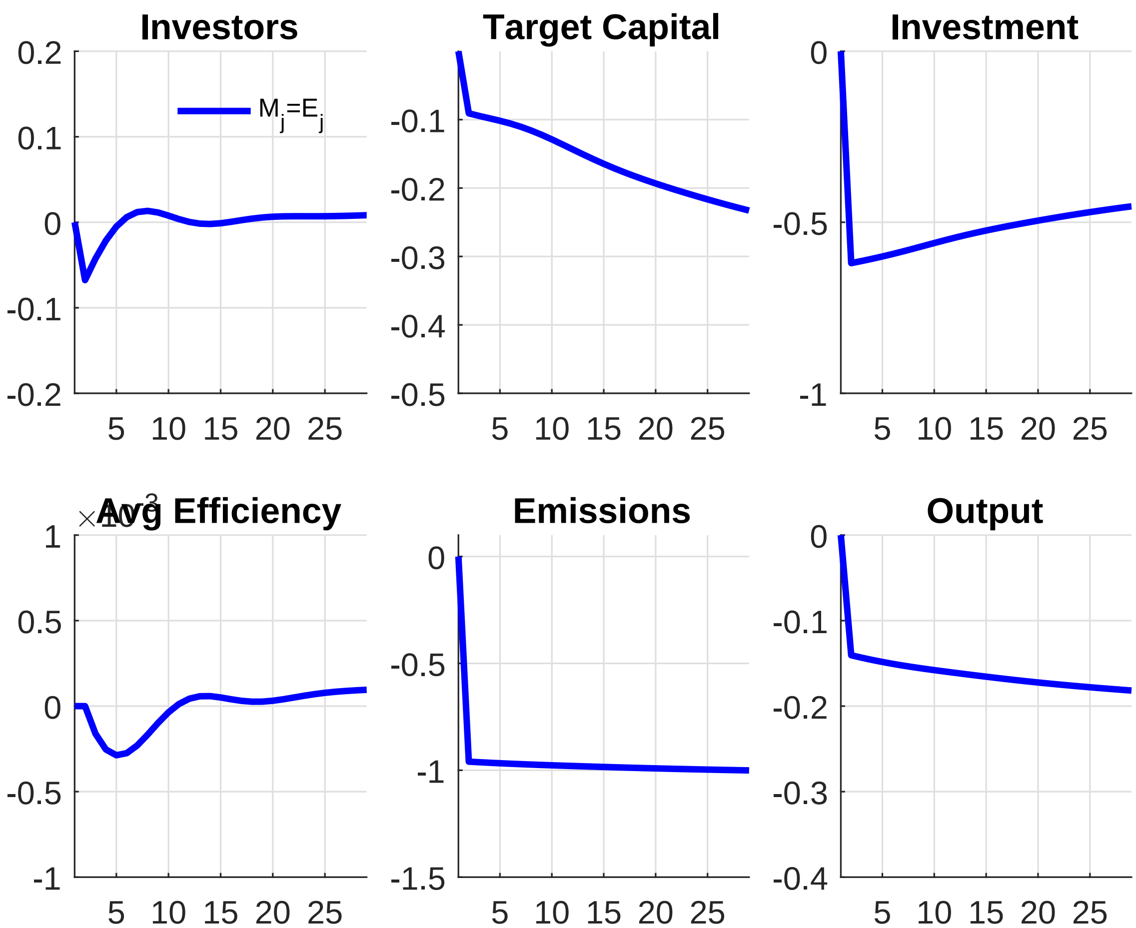

Figure 2: Carbon Tax

Notes: Responses when $\frac{\partial \Omega_j}{\partial j} = 0$ to a permanent carbon tax set to achieve a long term reduction in emissions of 1%. X-axis denotes quarters while y-axis measures percentage deviation from initial steady state.

permanently higher effective price on energy creates a permanent shift in the distribution of plants, with a higher number of establishments investing each period and a higher economy-wide energy efficiency. However, these long-run effects are rather small.

A plant's investment decision depends on how the future expected path of the carbon tax affects the plant's future, effective energy costs. When emission rates vary with plant vintage, a flat carbon tax has unequal tax incidence across plants of different vintages. As a result, a carbon tax alters the timing of capital replacement and the target capital stock. Figure 3 illustrates this issue by assessing the carbon tax policy when $\frac{\partial \Omega_j}{\partial j} > 0$ (dashed lines). In this case, plants with older capital face a higher effective tax as they have lower levels of productivity and higher emission rates, consistent with empirical evidence.[20] For comparison, the initial scenario when $\frac{\partial \Omega_j}{\partial j} = 0$ also is plotted (solid lines). The re-

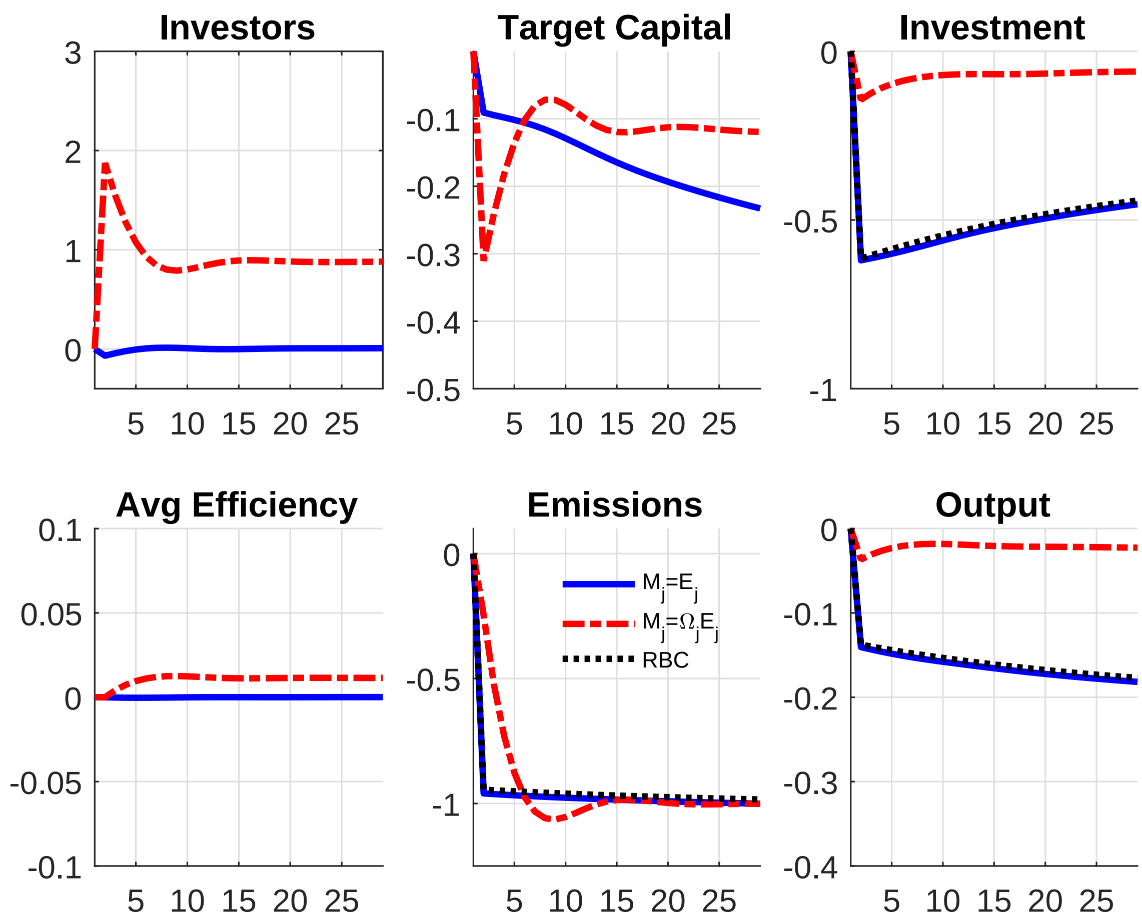

Figure 3: Carbon Tax

Notes: Comparison of heterogeneous firms and representative firm models following a reduction in emissions of 1%. Solid blue lines: plants with homogenous emission rates ($\frac{\partial \Omega_j}{\partial j} = 0$); dashed red lines: plants with heterogenous emission rates ($\frac{\partial \Omega_j}{\partial j} > 0$). Black dotted lines: representative firm model. X-axis denotes quarters while y-axis measures percentage deviation from initial steady state.

sponses in the economy with heterogeneous emission rates are quantitatively different. This difference can be traced to the shift in the cross-sectional distribution of vintage and

[20]See, for instance, Shapiro and Walker (2018) and Bloom et al. (2010). For this experiment, all that is needed is for $\frac{\partial \Omega_j}{\partial j} > 0$. Assuming this partial derivative is constant, its value is inconsequential. Since we calibrate the carbon tax τ^{em} to achieve a long-run reduction in energy of 1%, alternative values imply different τ^{em} necessary to achieve the reduction.

capital stocks induced by the carbon tax. When $\frac{\partial \Omega_j}{\partial j} > 0$, the burden of the carbon tax falls more heavily on older vintages, as they have relatively higher emission rates. This induces plants to have a lower target capital stock. At the same time, plants have higher incentives to invest, as their emission rates decline with newer capital and technology. In the short run, the number of investors increases – 2 percentage points above the initial steady state – and remains substantially higher in the long run. In turn, the average energy efficiency increases and the average age of the capital stock declines. Despite the spike in investors, the contraction in the target capital stock causes aggregate investment and output to decrease in the short and long run. Relative to the case where $\frac{\partial \Omega_j}{\partial j} = 0$, accounting for heterogeneous emission rates dampens the drop in aggregate activity as the tax induces firms to invest in cleaner and more energy-efficient capital. Indeed, to achieve a 1% decrease in emissions in the long run, the legislated carbon tax is only 0.47% in this case, compared to 3.30% when $\frac{\partial \Omega_j}{\partial j} = 0$.

5.3 Comparison with Representative Firm Model

We conclude by discussing the dynamics relative to the existing benchmark in the literature, the representative-firm RBC model presented in Section 3.10. This exercise disentangles the role of firm heterogeneity from emission rate heterogeneity. The dotted lines in Figure 3 report the impulse responses for the RBC model.[21] The RBC responses are indistinguishable from the $\frac{\partial \Omega_j}{\partial j} = 0$ economy, i.e. the economy where emission and energy usage rates are proportional. This result is not surprising in the investment literature and is in line with the so-called 'irrelevance result' in Thomas (2002), who shows firm heterogeneity per se does not matter for aggregate economic dynamics. What we stress is the role of heterogeneous emission rates for the aggregate effects of a carbon tax rather than firm heterogeneity per se. Adding a dimension of heterogeneity that is endogenous to the

[21] As in the other scenarios, we calibrate the carbon tax to produce a long-run decline in emissions of 1%.

firm's investment behavior alters the propagation of the tax policy.

6 Alternative Policy Scenarios

We now consider alternative ways to implement carbon taxes that have received attention in the existing literature. In Section 6.1 we focus on "grandfathering", a provision that exempts existing production units from the carbon tax. In Section 6.2, we study the aggregate dynamics in the presence of alternative "revenue recycling" scenarios. The main message of our analysis is that allowing for grandfathering may substantially backfire in the short-run by increasing emissions. Vintage-specific policies like grandfathering also can alter markedly the government's success in raising revenues through various revenue recycling programs.

6.1 Vintage-Differentiated Policy

As discussed in Stavins (2006), environmental policies often are vintage-differentiated. The most common application is "grandfathering" or exempting production units installed before a specific date from the policy, a practice customarily applied to new regulations. Reasons for such implementation normally are attributed to short-run fairness, so that the rules-of-the-game for establishments are not changed midstream, or to political pressures. Heutel (2011) and Shapiro and Walker (2018) argue that modeling a carbon tax captures well the implementation of various U.S. regulations. To examine the economic consequences of such reforms, we consider a variant of the carbon tax policy where the tax gradually affects establishments. We assume that in period 0, a government enacts a carbon tax that applies to the newest vintage ($j = 0$) in period 0, but exempts all older vintages. The tax is assumed to apply to all future vintages of higher technological progress as well, that is vintage $j = 0$ at times $t = 1, 2, ...$ In this way, in period 0, vintages of $j \geq 1$

are exempted from the policy and only are affected when they upgrade their capital and technology.

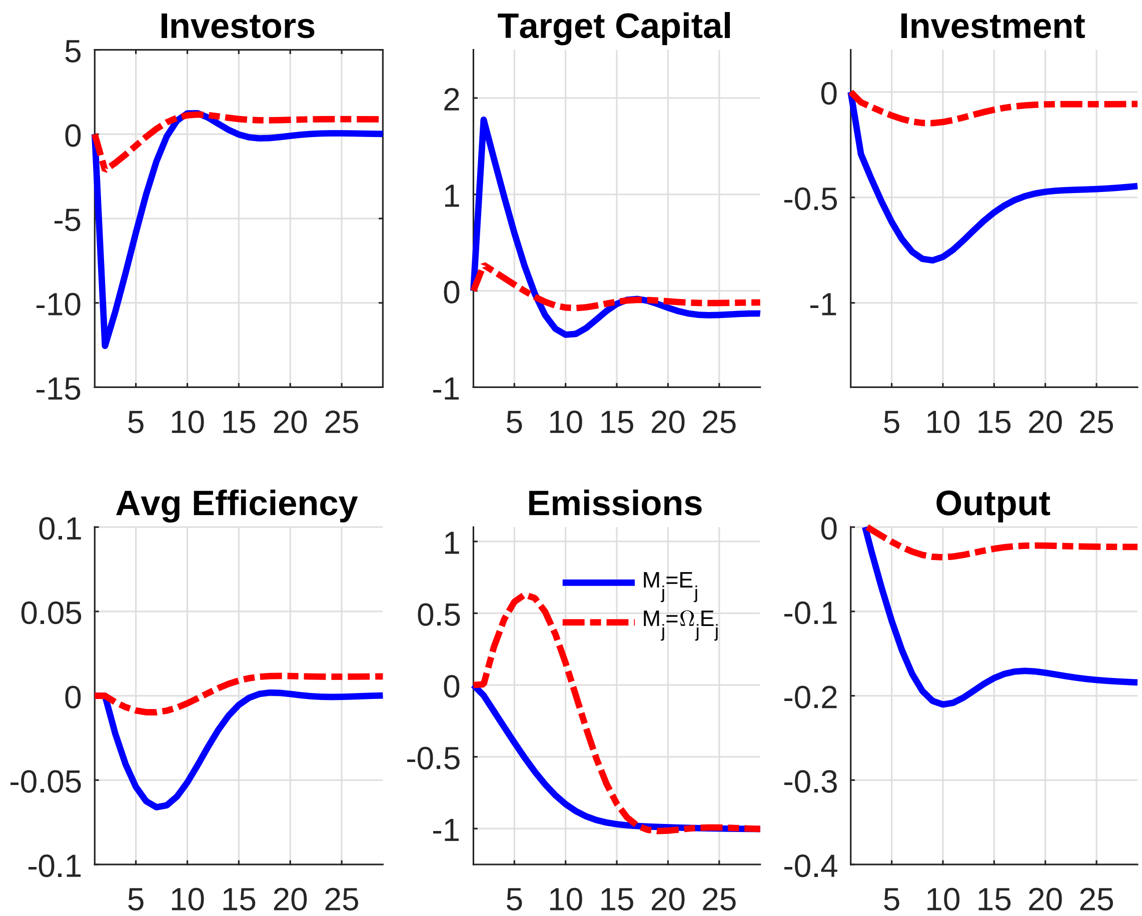

Figure 4: Carbon Tax - Grandfathering

Notes: Responses to a permanent carbon tax that gradually affects plants, as they update their capital stock, and is set to achieve a long term reduction in emissions of 1%. Solid blue lines: $\frac{\partial \Omega_j}{\partial j} = 0$; dashed red lines: $\frac{\partial \Omega_j}{\partial j} = 0$. X-axis denotes quarters while y-axis measures percentage deviation from initial steady state.

Figure 4 displays the responses of variables to the gradual tax reform in the model with plant heterogeneity in emissions (dashed lines) and common emissions (solid lines). In the short-run, emissions increase under heterogeneous emission rates, while they decrease with homogeneous emission rates. Following the tax enactment, when $\frac{\partial \Omega_j}{\partial j} > 0$ the number of investors declines by over 10%. Since plants are only subject to the policy when they update their capital and technology, the expected value of a plant at the technological frontier decreases. Thus, the effective price of new capital rises, and older capital is retired later. This leads the average age of the capital stock and aggregate emissions to increase markedly in the short run. These results are consistent with empirical work on the effects of grandfathering regulation. The most cited example of such practice in the U.S. is the New Source Review Program from the Clean Air Act of 1970, which the empirical literature shows reduced investment and increased the age of capital of exempted plants (see for instance Bushnell and Wolfram (2012) and Nelson et al. (1993)). Concerning the model with homogeneous emission rates, emissions instead decline as the drop in output offsets the increase in the average age of capital. Our findings suggest that exempting production units with old and dirtier capital conflicts, in the short run, with emission targets.

6.2 Alternative "Revenue Recycling" Scenarios

As the fiscal use of carbon tax revenues often is cited as a motivation for the tax's implementation, we consider the effects of alternative "revenue recycling" scenarios. The top panel of figure 5 repeats the carbon tax policy when it affects all plants simultaneously in the economy with homogeneous emission rates, $\frac{\partial \Omega_j}{\partial j} = 0$. The solid lines repeat our benchmark case where tax revenues are rebated back to households in lump-sum fashion whereas the dotted lines use the revenues to reduce the consumption tax and dashed lines use revenues to reduce the labor tax.[22]

Offsetting the carbon tax revenues with reductions in either the consumption or labor distortionary tax results in a smaller aggregate contraction since both create offsetting expansionary effects. A reduction in consumption taxes makes households more willing to

[22]For the perfect foresight experiments for this economy, an increase in government spending produces the same effects as a reduction in consumption taxes, as both create the same demand-side effects. We present results for the uniform emission rate case only as the heterogeneous emission rate economy exhibits the same qualitative patterns.

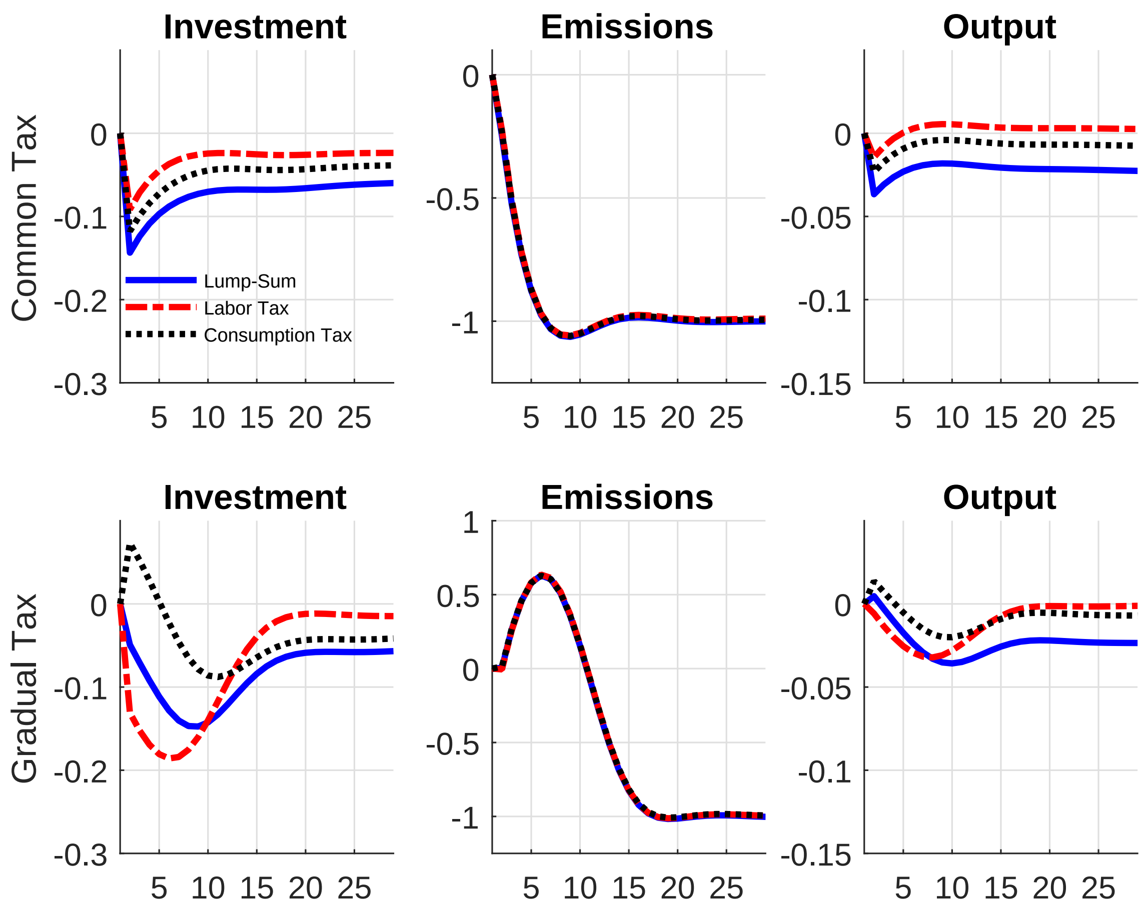

Figure 5: Carbon Tax - Revenue Recycling

Notes: Various revenue recycling scenarios following a carbon tax introduction. Solid blue lines: rebating proceeds lump-sum to households; dashed red lines: reducing labor taxes. Black dotted lines: reducing consumption taxes. X-axis denotes quarters while y-axis measures percentage deviation from initial steady state.

consume goods, and in turn, more willing to supply labor. Likewise, a reduction in labor taxes makes households more willing to work. The labor effect in both cases alleviates the contractionary supply-side implications of the higher effective energy price.

The bottom panel of Figure 5 repeats the financing scenarios for the gradual carbon tax, in which some plants are initially exempted, as in Section 6.1. In the long run, the revenue recycling effects are the same as those from the uniform tax case, as all plants eventually are subject to the tax even with gradual reform. The short-run dynamics, however, can differ markedly. On impact, offsetting the carbon tax revenues with the labor tax is more contractionary than the lump-sum rebates to households. In this case, carbon tax revenues are initially very low, as only plants of the newest vintage are subject to the tax and the number of investors decreases. When coupled with the overall contractionary effects in the economy, which depress government revenues, the labor tax is forced to *rise* to balance the government budget. Thus, the policy leads to two contractionary fiscal actions on impact, further dampening output and investment in the short run. These results provide a cautionary warning, suggesting that the revenue gains from a carbon tax could appear gradually, depending on the policy implementation.

7 Robustness

In this section, we investigate the sensitivity of our results to alternative policy experiments and model specifications. We present results for the economy with plant heterogeneous emissions, $\frac{\partial \Omega_j}{\partial j} > 0$.

7.1 Policy Size

For the baseline experiment, we calibrated the carbon tax shock to deliver a long-run reduction in emissions of 1%. To determine how sensitive our inferences are to the shock size, figure *A.2* repeats the experiment when the government alternatively targets a 5% reduction in emissions with the carbon tax (dashed lines) or a 10% reduction in emissions (dotted lines). The respective sizes of the carbon tax in each experiment are 0.47%, 2.50%, and 5.52%. More substantial carbon tax reforms deliver more significant responses from macroeconomic aggregates, with the change in responses roughly proportional to the change in the tax reform.

7.2 Irreversible Investment

For our benchmark specification, we assumed that the non-depreciated portion of older capital stocks could be converted freely to the latest vintage of capital. Alternatively, one could argue that a portion of older vintage capital is irreversible. $\delta_j^s > 0$, from equation 4, captures the notion that when an investing plant upgrades its technology, part of its current, vintage capital stock may not be operable with the new technology and must be scrapped.

To investigate the effects of such irreversible investment, we conduct an experiment where we set $\delta_j^s = \delta^s \alpha_j$ and $\delta^s = 0.5$. In this case, the degree of capital irreversibility varies across vintages, with older vintages scrapping larger portions of capital because the capital is irreconcilable with the new technology. We recalibrate the upper support of the fixed cost distribution (B) to continue to match empirical evidence on investment spikes (see Section 4.1 for more details) and calibrate the carbon tax to target a long-run reduction in emissions of 1%.[23] We find that irreversibility tends to reduce the response of average energy efficiency, inducing larger output losses in the long run. Figure $A.3$ reports the results.

8 Conclusion

As climate concerns are increasingly discussed, policymakers weigh options to reduce emissions. Understanding how such policy options affect investment behavior is of central importance, as encouraging more energy-efficient and less pollutant technology adoption is a key policy consideration. In this paper, we develop a general-equilibrium framework to study how a carbon tax affects the decision to replace old technologies with new

[23] Under this scenario B is equal to 0.14, and the tax rate is equal to 1.67%.

ones. Our model can fit a variety of firm-level stylized facts of the U.S. economy – newer, more productive plants emit less per unit energy than older, unproductive plants; productive plants produce more than unproductive plants; and plants make large and infrequent investments.

We find that the quantitative predictions following a carbon tax reform depend on the cross-sectional distribution of capital stock and vintage technology across plants. Accounting for heterogeneous emission rates reduces long-run output losses by about 50% relative to a representative firm model, as the economy's average productivity and energy efficiency rises following a carbon tax. Moreover, the size of the tax necessary to achieve the same reduction in long-run emissions varies in the model with heterogeneous plant emissions. The short-run response of the economy depends on details about the implementation of the tax. A policy that initially exempts older plants *increases* emissions in the short run. Overall, the results suggest that the characteristics of the production structure at the microeconomic level are informative about the aggregate effects of a carbon tax. Taking into account these characteristics could have sizable effects for many policy considerations, including the design of an optimal carbon tax.

Our results have immediate implications for empirical research as well, calling attention to the need to understand the plant-level characteristics of production that are endogenous to the plant's investment behavior. Distinguishing how plant dispersion varies across industries, regions, and countries is essential for predicting how environmental policies will affect technology adoption.

References

ACEMOGLU, D., U. AKCIGIT, D. HANLEY, AND W. KERR (2016): "Transition to Clean Technology," *Journal of Political Economy*, 124, 52–104.

BACHMANN, R. AND C. BAYER (2014): "Investment Dispersion and the Business Cycle," *American Economic Review*, 104, 1392–1416.

BACHMANN, R., R. CABALLERO, AND E. ENGEL (2013): "Aggregate Implications of Lumpy Investment: New Evidence and a DSGE Model," *American Economic Journal: Macroeconomics*, 5, 29–67.

BLOOM, N., M. FLOETOTTO, N. JAIMOVICH, I. SAPORTA-EKSTEN, AND S. J. TERRY (2018): "Really Uncertain Business Cycles," *Econometrica*, 86, 1031–1065.

BLOOM, N., C. GENAKOS, R. MARTIN, AND R. SADUN (2010): "Modern Management: Good for the Environment or Just Hot Air?" *Economic Journal*, 120, 551–572.

BOSETTI, V. AND M. MAFFEZZOLI (2013): "Taxing Carbon under Market Incompleteness," Fondazione Eni Enrico Mattei Working Papers No. 831.

BOUCEKKINE, R., D. DE LA CROIX, AND O. LICANDRO (2008): "Vintage Capital," in *New Palgrave Dictionary of Economics, Second Edition*, ed. by S. Durlauf and L. Blume, 628–631.

BUSHNELL, J. B. AND C. D. WOLFRAM (2012): "Enforcement of Vintage Differentiated Regulations: The Case of New Source Review," *Journal of Environmental Economics and Management*, 64, 137–152.

CABALLERO, R. J. AND E. M. ENGEL (1999): "Explaining Investment Dynamics in U.S. Manufacturing : A Generalized (S,s) Approach," *Econometrica*, 67, 783–826.

COOLEY, T. F., J. GREENWOOD, AND M. YORUKOGLU (1997): "The replacement problem," *Journal of Monetary Economics*, 40, 457–499.

DASGUPTA, S., R. E. LUCAS, AND D. WHEELER (2002): "Plant Size, Industrial Air Pollution, and Local Incomes: Evidence from Mexico and Brazil," *Environment and Development Economics*, 7, 365–381.

DOMS, M. AND T. DUNNE (1998): "Capital Adjustment Patterns in Manufacturing Plants," *Review of Economic Dynamics*, 1, 409–429.

FIORI, G. (2012): "Lumpiness, Capital Adjustment Costs and Investment Dynamics," *Journal of Monetary Economics*, 59, 381–392.

FIORI, G. AND F. SCOCCIANTI (2018): "Aggregate Productivity and Microeconomic Heterogeneity: The Role of Vintage Technology," *mimeo*.

FULLERTON, D. AND H. MONTI (2013): "Can Pollution Tax Rebates Protect Low-Wage Earners?" *Journal of Environmental Economics and Management*, 66, 539–553.

GAVAZZA, A., S. MONGEY, AND G. L. VIOLANTE (2018): "Aggregate Recruiting Intensity," *American Economic Review*, 108, 2088–2127.

GOLOSOV, M., J. HASSLER, P. KRUSELL, AND A. TSYVINSKI (2014): "Optimal Taxes on Fossil Fuel in General Equilibrium," *Econometrica*, 82, 41–88.

GORDON, R. J. (1990): *The Measurement of Durable Goods Prices*, no. gord90-1 in NBER Books, National Bureau of Economic Research, Inc.

GOURIO, F. AND A. KASHYAP (2007): "Investment Spikes: New Facts and A General Equilibrium Exploration," *Journal of Monetary Economics*, 54, 1–22.

GREENWOOD, J., Z. HERCOWITZ, AND P. KRUSELL (2000): "The Role of Investment-Specific Technological Change In The Business Cycle," *European Economic Review*, 44, 91–115.

HALTIWANGER, J., R. COOPER, AND L. POWER (1999): "Machine Replacement and the Business Cycle: Lumps and Bumps," *American Economic Review*, 89, 921–946.

HASSLER, J. AND P. KRUSELL (2012): "Economics and Climate Change: Integrated Assessment in a Multi-Region World," *Journal of the European Economic Association*, 10, 974–1000.

HEUTEL, G. (2011): "Plant Vintages, Grandfathering, and Environmental Policy," *Journal of Environmental Economics and Management*, 61, 36–51.

JOHANSEN, L. (1959): "Substitution versus Fixed Production Coefficients in the Theory of Economic Growth: A Synthesis," *Econometrica*, 27, 157–176.

JONES, J. (2002): "Has Fiscal Policy Helped Stabilize the Postwar U.S. Economy?" *Journal of Monetary Economics*, 49, 709–746.

KEANE, M. AND R. ROGERSON (2015): "Reconciling Micro and Macro Labor Supply Elasticities: A Structural Perspective," *Annual Review of Economics*, 7, 89–117.

KHAN, A. AND J. THOMAS (2003): "Nonconvex Factor Adjustments in Equilibrium Business Cycle Models: Do Nonlinearities Matter?" *Journal of Monetary Economics*, 50, 331–360.

——— (2008): "Idiosyncratic Shocks and the Role of Nonconvexities in Plant & Aggregate Investment Dynamics," *Econometrica*, 76, 395–436.

KILIAN, L. AND D. P. MURPHY (2014): "The Role of Inventories and Speculative Trading in the Global Market for Crude Oil," *Journal of Applied Econometrics*, 29, 454–478.

King, R. and S. Rebelo (1999): "Resuscitating Real Business Cycles," *Handbook of Macroeconomics*, 1, 927–1007.

Krusell, P. and A. A. Smith (2015): "A Global Economy-Climate Model with High Regional Resolution," Yale University manuscript.

Leeper, E. M., M. Plante, and N. Traum (2010): "Dynamics of Fiscal Financing in the United States," *Journal of Econometrics*, 156, 304–321.

Licandro, O., M. R. Maroto Illera, and L. Puch (2005): "Innovation, Machine Replacement and Productivity," *C.E.P.R. Discussion Papers*, 1.

Lyubich, E., J. S. Shapiro, and R. Walker (2018): "Regulating Mismeasured Pollution: Implications of Firm Heterogeneity for Environmental Policy," *AEA Papers and Proceedings*, 108, 136–42.

Miao, J. and P. Wang (2014): "Lumpy Investment and Corporate Tax Policy," *Journal of Money, Credit and Banking*, 46, 1171–1203.

Nelson, R. A., T. Tietenberg, and M. R. Donihue (1993): "Differential Environmental Regulation: Effects on Electric Utility Capital Turnover and Emissions," *The Review of Economics and Statistics*, 75, 368–373.

Nilsen, Ø. A. and F. Schiantarelli (2003): "Zeros and Lumps in Investment: Empirical Evidence on Irreversibilities and Nonconvexities," *The Review of Economics and Statistics*, 85, 1021–1037.

Nordhaus, W. D. (1994): *Managing the Global Commons: The Economics of Climate Change*, MIT Press.

——— (2008): *A Question of Blance: Weighing the Options on Global Warming Policies*, Yale University Press.

NORDHAUS, W. D. AND J. BOYER (2000): *Warming the World: Economic Models of Global Warming*, MIT Press.

PORTER, M. (1991): "America's Green Strategy," *Scientific American*, 264, 168.

POWER, L. (1998): "The Missing Link: Technology, Investment, And Productivity," *The Review of Economics and Statistics*, 80, 300–313.

QI, J., X. TANG, AND X. XI (2019): "The Size Distribution of Firms and Industrial Water Pollution: A Quantitative Analysis of China," Fudan University manuscript.

RAUSCH, S., G. METCALF, J. REILLY, AND S. PALTSEV (2010): "Distributional Implications of Alternative U.S. Greenhouse Gas Control Measures," *The B.E. Journal of Economic Analysis & Policy*, 10, Symposium.

RESTUCCIA, D. AND R. ROGERSON (2008): "Policy Distortions and Aggregate Productivity with Heterogeneous Establishments," *Review of Economic Dynamics*, 11, 707–720.

SAKELLARIS, P. (2004): "Patterns of plant adjustment," *Journal of Monetary Economics*, 51, 425–450.

SHAPIRO, J. S. AND R. WALKER (2018): "Why is Pollution from U.S. Manufacturing Declining? The Roles of Environmental Regulation, Productivity and Trade," *Americaon Economic Review*, forthcoming.

SOLOW, R. (1960): *Investment and technological progress*, Stanford University Press, chap. 7.

STAVINS, R. N. (2006): "Vintage Differentiated Environmental Regulation," *Stanford Environmental Law Journal*, 25, 29–66.

THOMAS, J. K. (2002): "Is Lumpy Investment Relevant for the Business Cycle?" *Journal of Political Economy*, 110, 508–534.

VERACIERTO, M. L. (2002): "Plant-Level Irreversible Investment and Equilibrium Business Cycles," *American Economic Review*, 92, 181–197.

APPENDIX

A Equilibrium Conditions for Representative Firm Model

The stationary equilibrium conditions for the standard neoclassical growth model with a C.E.S. production function are as follows:

$$
\begin{aligned}
y_t &= \left[\omega^{\frac{1}{\epsilon}} \left(\left(\frac{k_{t-1}}{\Theta_{K_t}}\right)^{\gamma} L_t^{\nu}\right)^{\frac{\epsilon-1}{\epsilon}} + (1-\omega)^{\frac{1}{\epsilon}} (e_t)^{\frac{\epsilon-1}{\epsilon}}\right]^{\frac{\epsilon}{\epsilon-1}} \\
p_t^e(1+\tau_t^{em}\Omega) &= (1-\omega)\frac{y_t}{e_t} \\
w_t &= \nu(\omega y_t)^{\frac{1}{\epsilon}} \left(\frac{k_{t-1}}{\Theta_{K_t}}\right)^{\frac{\gamma(\epsilon-1)}{\epsilon}} L_t^{\frac{\nu(\epsilon-1)}{\epsilon}-1} \\
\psi L_t^{\eta} &= \frac{(1-\tau_t^l)w_t}{(1+\tau_t^c)c_t} \\
\lambda_t &= \frac{c_t^{-1}}{1+\tau_t^c} \\
\lambda_t &= E_t \frac{\beta\lambda_{t+1}}{\Theta_{K_{t+1}}} \left(\gamma(\omega y_{t+1})^{\frac{1}{\epsilon}} \left(\frac{k_t}{\Theta_{K_{t+1}}}\right)^{\frac{\gamma(\epsilon-1)}{\epsilon}-1} L_{t+1}^{\frac{\nu(\epsilon-1)}{\epsilon}} + (1-\delta)\right) \\
y_t &= c_t + g_t + i_t + p_t^e e_t \\
k_t &= (1-\delta)\frac{k_{t-1}}{\Theta_{K_t}} + i_t \\
\Theta_{K_t} &= \Theta_{Z_t}^{\frac{1}{1-\gamma}}
\end{aligned}
$$

B Model Solution

The model solution procedure follows Thomas (2002) and Gourio and Kashyap (2007). The computation of the steady state requires a numerical procedure because J, the maximum time-since-adjustment, is endogenously determined. We first guess a value for J. Conditional on this value, we guess the target capital k_0, the fraction of plants investing $\boldsymbol{\alpha}_j$, the energy levels of firms e_j, and the real wage w. After using these values to solve for the remaining variables, equations (17), (18), (19), and (22) are used to verify and update the guess. This procedure continues for values of J until α_J is endogenously equal to 1 (i.e., plants with capital k_J adjust with probability 1). We consider perfect foresight exercises that give a permanent change in policy parameters. Transition dynamics from the initial equilibrium to the final equilibrium are found by solving the model as a nonlinear forward-looking deterministic system.

C Cross-Sectional Distribution - Steady State

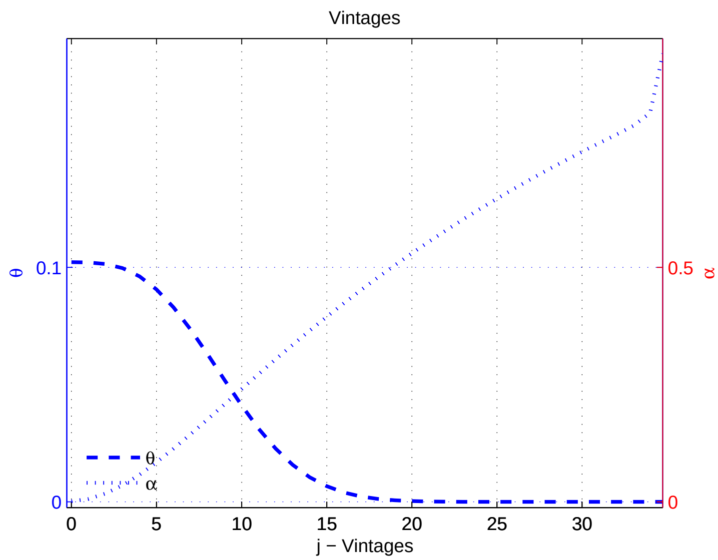

Figure A.1: Steady State Distribution

Notes: The initial steady state distribution of firms (θ_j) and hazard rates (α_j).

D Robustness

Below, we report the impulse responses of aggregate variables following the introduction of a carbon tax, assuming that the government wants to reduce emissions, by 1%, 5%, and 10% respectively.

Here, we report the impulse responses of aggregate variables following the introduction of a carbon tax, comparing the case with $\delta^s = 0.5$ and the baseline when $\delta^s = 0$.

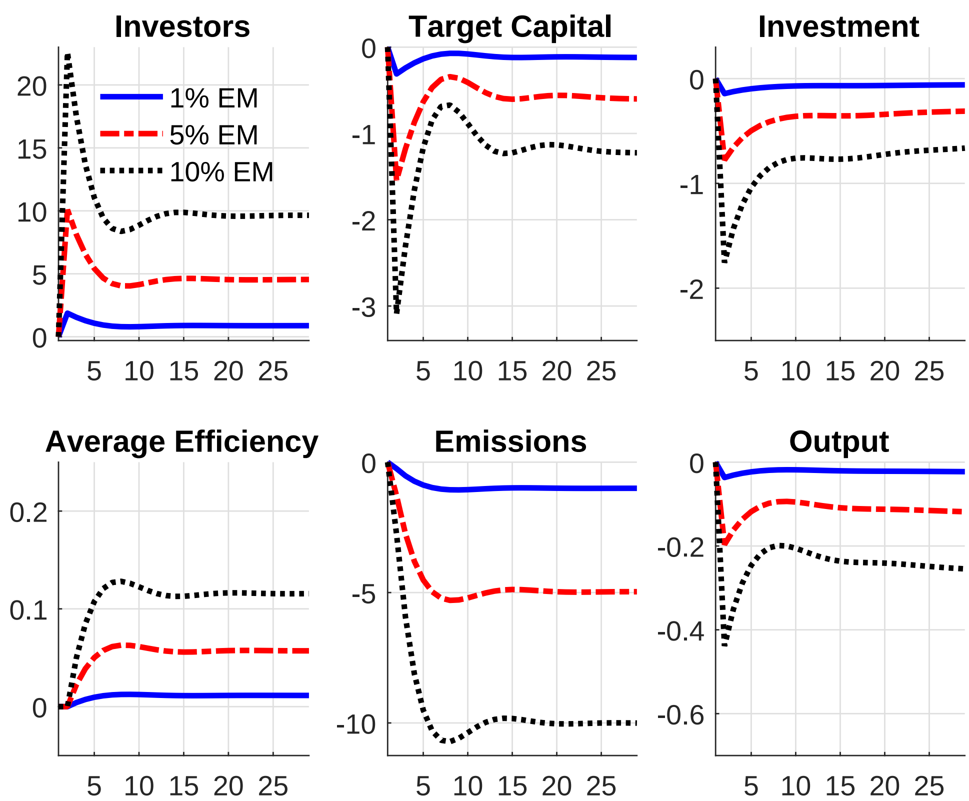

Figure $A.2$: Carbon Tax - Different Size

Notes: Various size carbon tax reforms. Solid blue lines: tax to induce 1% long-run drop in emissions; dashed red lines: tax to induce 5% long-run drop in emissions; Black dotted lines: tax to induce 10% long-run drop in emissions. X-axis denotes quarters while y-axis measures percentage deviation from initial steady state.

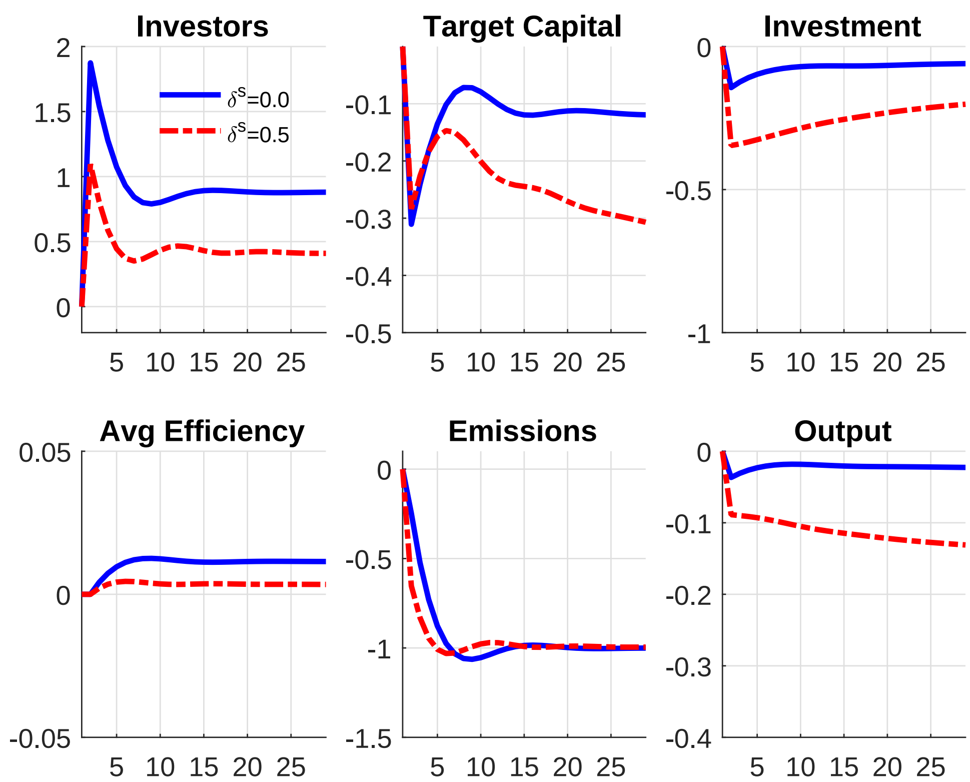

Figure $A.3$: Carbon Tax

Notes: Carbon tax reforms under different δ^s. Solid blue lines: $\delta^s = 0.0$; dashed red lines: $\delta^s = 0.5$. X-axis denotes quarters while y-axis measures percentage deviation from initial steady state.