Size-Dependent Exemptions in Cap-and-Trade and Aggregate Productivity*

Moudachirou Oumarou†

This version: Août, 2024

Click here for latest version

Abstract

This paper examines the long-term effects of size-dependent exemptions in cap-and-trade systems, which allow firms below a specific emissions threshold to avoid regulation. Using data from California's Cap-and-Trade program, I find that 40% of firms qualify for these exemptions, with significant clustering near the threshold, evidenced by a discontinuity test in emissions distribution post-regulation. This pattern suggests that firms strategically adjust emissions to avoid regulatory costs. By extending an industry dynamics framework with heterogeneous firms, I quantify the policy's impact on aggregate productivity, finding that removing exemptions raises productivity by 0.2% as resources shift to more productive firms. Additionally, I show that the policy causes misallocation, as intermediate-productivity firms near the threshold exhibit higher average Total Factor Productivity (TFPR) than larger, more productive firms subject to emissions costs.

Keywords: emission, exemption, productivity, misallocation
JEL Classification: H32, L11, Q58

*I am grateful to my advisor Guillaume Sublet for his guidance and support. I have also benefited from comments by Rui Castro and Emanuela Cardia. I thank seminar participants at University of Montreal's macroeconomics workshop, the CIREQ Interdisciplinary PhD Student Symposium on Climate Change, CIREQ Ph.D. Students' Conferences for their invaluable contributions and insights. I acknowledge the financial support of Fonds de Recherche du Québec société et Culture (FRQSC).

†Economics department, Université de Montreal: moudachirou.oumarou@umontreal.ca

1 Introduction

Modern environmental policies often include exemptions for specific groups of regulated entities, typically based on criteria like firm size or sector, allowing some to bypass the costs of regulatory compliance.

This paper specifically explores size-dependent exemption criteria that involve specific emissions thresholds, where only firms exceeding a specific emission level face regulatory costs. This raises a critical question: Are these exemptions widespread and, if so, what are their implications for policy effectiveness? The answer is a clear yes for the first part of the question considering the list of jurisdictions allowing for such exemption in their environmental regulation. I document in Table 1 an overview of jurisdictions allowing for this type of exemption. The adoption of such exemption policy is common and independent of the regulatory carbon pricing systems. For instance, the Alberta's Specified Gas Emitters Regulation applies only to plants emitting over 100 kt of CO2 equivalent annually, while the Quebec and California Cap-and-Trade, as well as the Chile Carbon tax set the threshold at 25 kt. Facilities emitting below these thresholds are exempt from the regulatory program, avoiding the need to internalize their emission impact. For the second part of the question, the answer is unclear. This paper provides an answer by quantifying the effects of these exemption policies on aggregate productivity.

Why would policymakers design exemptions that seemingly undermine the financial intake and emissions control objectives of these regulations? While the literature offers several reasons—such as reducing administrative cost for monitoring the regulation (Keen and Mintz, 2004; Dharmapala et al., 2011), alleviate economic burdens for smaller firms or encouraging the entry of new businesses (Berman and Bui, 2001; Tombe and Winter, 2015) —there are also significant economic concerns about how these exemptions influence firm behavior and overall efficiency. A size dependent policy may create disparities in marginal emission costs, distorting the incentives for pollution reduction and the adoption of clean technologies. The question becomes, what if exempted firms are more carbon-intensive than their regulated counterparts?

In these regards, this paper uses firm-level data from California's Cap-and-Trade program to reveal that many firms cluster around the emission threshold to avoid regulatory costs, resulting in efficiency documented with a discontinuity test in the distribution of emission in the post-regulation years. This result suggest firms self-selection to avoid paying emission costs. I calibrate an industry dynamic frameworks including differences in production and pollution efficiency across heterogeneous firms to replicate the distribution of firms, emissions, and employment under California's Cap-and-Trade. I show that eliminating the exemption policy would lead to an increase in output (0.05%) and productivity (0.2%), without sacrificing environmental goals. In doing so, this study suggests that treating all firms equally under environmental regulations could yield better economic outcomes while achieving the same emission reduction objective.

To further characterize the inefficiency implied by the exemption policy, I use the model to calculate the extent to which variations in firms' Total Factor Productivity Revenue (TFPR) are influenced by differences within and across firm categories. These categories are based on their emissions relative to the exemption threshold. Firms with intermediate productivity levels, which bunch at the emission threshold, exhibit significantly higher average and dispersion in TFPR compared to large, productive firms that incur emission costs.

Related Literature.- The paper contributes to three growing areas of literature. Firstly, it addresses the optimal choice between tax and intensity standards, considering market structure. Intensity standards often outperform taxes because implementing precise Pigouvian taxes, reflecting the first-best, is challenging, and they are applicable only under fully competitive markets (Buchanan, 1969). Emission taxes worsen underproduction for firms with market power (Li and Sun, 2015; Tombe and Winter, 2015; Li and Shi, 2017) or facing leakages (Fowlie, 2009; Holland, 2012; Fowlie et al., 2016a). In contrast, intensity standards create an implicit output subsidy, partially compensating firms and offsetting market failures or incomplete markets.

Secondly, the paper aligns with the literature on size-dependent policies, where input factors are misallocated, resulting in inefficient output production and a loss in aggregate productivity (Guner et al., 2008; Restuccia and Rogerson, 2008; Garicano et al., 2016). The impact of such policies is substantial as they distort the average size of establishments (Guner et al., 2008) and can contribute to differences in Total Factor Productivity (TFP) across countries (Hsieh and Klenow, 2009) such that any deregulation equalizing the marginal product of inputs improves output. Empirically, the paper explores non-linear budget sets and bunching estimation induced by size-dependent policies. Previous studies (Burtless and Hausman, 1978; Hausman, 1985) estimate labor supply under a kink budget set, revealing that reported data on labor income from taxed units do not exhibit bunching at the kink point, contrary to model predictions. Recent efforts in the literature aim to reconcile model predictions and survey data using rich administrative data (Saez, 2010; Chetty, 2012; Kleven and Schultz, 2014).

Finally, the paper contributes to the literature on optimal taxation, exploring the exemption of small firms in the presence of externalities. Examining countries with value-added tax (VAT) policies applied to firms above a certain threshold level, Keen and Mintz (2004) demonstrates that the optimal threshold involves a trade-off between raising government tax revenue (lower threshold) and the distortion caused by tax asymmetry on firms, along with compliance and administrative costs.

This paper is similar to Kaplow (2019) and Fowlie (2009) in that both explore an exogenous exemption of small firms. However, there are key differences. Firstly, I model the exemption criteria to be threshold-specific, fully incorporating emissions as an endogenous variable in a general equilibrium model. In contrast, Kaplow (2019) distinguishes between exempted and regulated firms exogenously, and Fowlie (2009) considers firms' exemption to be randomly assigned in a partial equilibrium. Secondly, I delve into the quantitative macroeconomics implications of exemption by considering differences in firms' productivity and emission rates. On the other hand, Fowlie (2009) primarily studies its implications on emissions leakage.

The remaining sections of this paper are organized as follows. In Section 2, I present summary statistics on the distribution of firms' emissions. I also provide empirical evidence regarding firms' relative emissions concerning the exemption threshold and the selection decision. Section 3 introduces a basic primer model to build intuition and derive key properties related to the exemption threshold. In Section 4, I present a more comprehensive conceptual framework. Moving to Section 5, I conduct policy simulations, quantitatively analyzing the implications of the exemption by comparing it with alternative policies through counterfactual analysis.

2 Empirical Motivation

In this section, I delve into an empirical investigation of the link between firms' emissions and their selection based on the exemption threshold. The primary dataset for this analysis is the Greenhouse Gas Reporting data from the California Air Resources Board. Under the Cap-and-Trade system, all establishments are mandated to report their annual emissions. This dataset provides comprehensive information on facility-level annual emissions and sector classification from 2008 to 2021. Notably, the California regulation, implemented during 2012-2013, features a fixed exemption threshold of 25 kt of CO2 equivalent, which remains unchanged over time. Meanwhile, the average permit price from the auction rose from \$10 to \$22 US.

Table 1. Jurisdictions with size-dependent exemption in GHG regulation

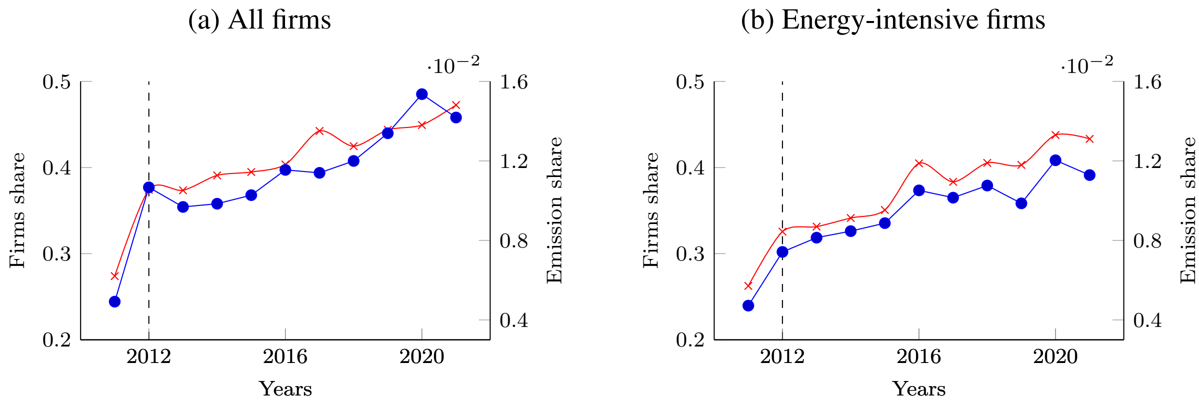
Jurisdiction	Regulation type	Implementation year	Threshold[a] (in tCO2 equivalent)
Alberta	TPS[b]	2007	100000
Beijing	ETS[c]	2013	5000
California	ETS	2013	25000
Chile	Carbon tax	2017	25000
China	ETS	2021	26000[d]
Manitoba	ETS, Carbon tax	2020[e]	50000
Mexico	ETS	2020	100000
Quebec	ETS	2013	25000
Saskatchewan	ETS	2019	25000
Singapore	Carbon tax	2019	25000
Tamaulipas	Carbon tax	2021	25 (monthly)
UK	ETS	2021	2500

[a] The threshold refers to annual emissions unless otherwise specified
[b] Tradable Performance Standard
[c] Emission Trading Scheme
[d] Covered only power sector plants
[e] Expected to be voted in March 19, 2020 but postponed due to the pandemic

Between 2008 and 2021, there is a notable increasing trend in the share of firms and emissions below the exemption threshold, as illustrated in Figure 1. The left axis represents the share of firms below the exemption threshold, while the right axis shows the fraction of emissions from these exempted firms. On average, four out of ten firms emit below the exemption threshold, accounting for 1.12% of the overall emissions. There is a distinct and significant jump in both these margins from 2011 to 2012, marking the beginning of the regulation. The proportion of firms operating below the threshold increased by more than 10 percentage points in 2012. Moreover, these proportions consistently show an upward trajectory over time, reaching 50% and 1.5% in 2020, respectively. This implies that nearly half of the companies are small emitters. Although they are exempt from regulatory costs, their combined emissions constitute only 1% of the total emissions.

At the cross-sectional level, there is evidence of clustering around the exemption threshold in the distribution of emissions after regulation. This phenomenon is absent in the distribution before regulation. This suggests that firms manipulate their emissions to fall below the exemption threshold, thereby avoiding regulatory costs. Figure 2 illustrates the difference in emission distributions between the first year after the regulation's implementation (2013) and one year before (2011). The comparison reveals clear signs of clustering. First, there is a substantial increase in the number of firms just below the exemption threshold (25,000 tonnes of CO_2) [see Figure 2a] compared to the counterfactual distribution

Figure 1. Firms below the exemption threshold

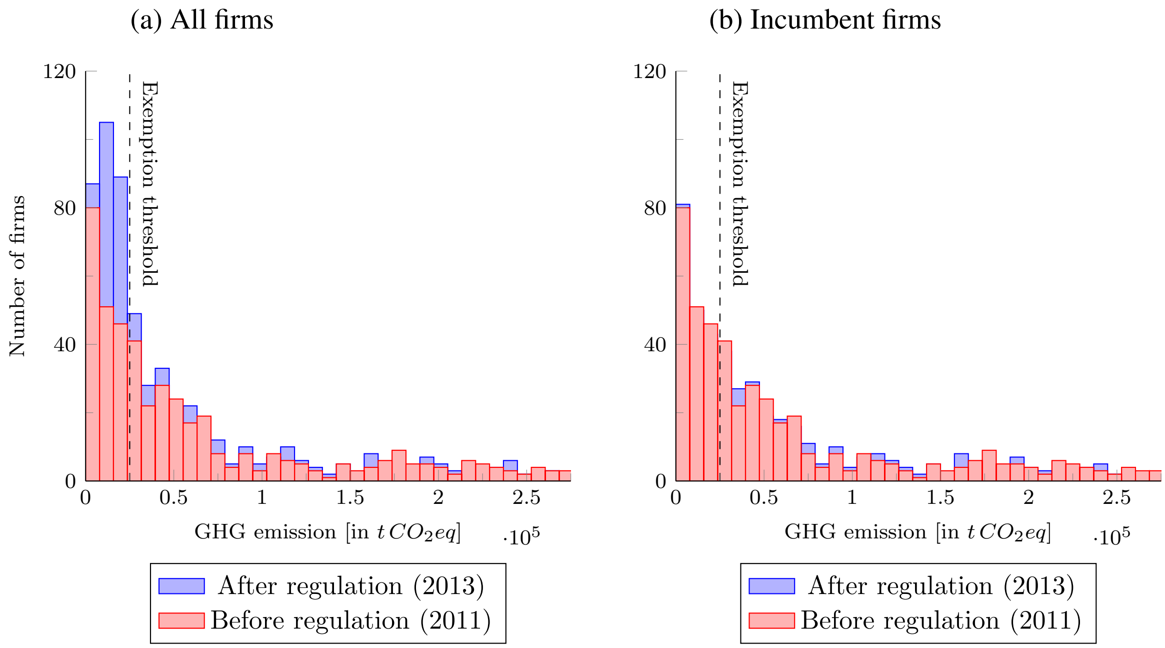


Note.- In each figure, the red curve represents the dynamics of the share of firms polluting below the exemption threshold and the blue curve depicts the emission share of those firms exempted relative to the overall emission. In the right panel, these two statistics are specifically shown for firms operating in energy-intensive sectors, which comprise firms falling within NAICS sectors 21, 22, 31-33, and 41-49.

(before regulation). Second, there is a significant drop in the after-regulation distribution precisely at the cutoff, a feature absent in the distribution before regulation. Consequently, a positive mass of firms deliberately adjusts their emissions to hover around the kink point, strategically avoiding regulatory costs. In panel 2b, the pool of firms in the after-regulation distribution is limited to the same firms as in the before-regulation distribution, excluding new entrants. The previously noted evidence of clustering disappears, and the two distributions become almost identical. This result suggests that the clustering around the kink point is primarily driven by new entrant firms. This can be attributed to the fact that incumbent firms, operating before the regulation, are less likely to alter their entry-year technology of emission to a less emission-intensive one, given the high associated costs.

Did firms cluster around the emission threshold to avoid regulatory costs? Is this clustering a result of firms' self-selection? To address these questions, I use the approach from Cattaneo et al. (2018) to test for discontinuities in the emission density around the threshold. This method uses local polynomial density estimation to assess whether self-selection occurs near the threshold where policy assignment varies. Additional details about the procedure can be found in McCrary (2008), Cattaneo et al. (2018), and Cattaneo et al. (2020). The results of the discontinuity test are presented in Table A2. The findings indicate that the p-value of the test is statistically significant for multiple years after the regulation (2015, 2017, 2020, and 2021), while it is not significant before the regulation. Moreover, when using a restricted sample containing only the same incumbent firms as in 2011, the test is not significant for almost all the years after the regulation, except for 2017 and 2021.

Figure 2. Evidence of bunching in the distribution of emission



Note.- The after-distribution plot in panel (b) comes from the restricted sample consisting of the same firms present before the regulation (in 2011), meaning that the new entrant firms are excluded.

The next section explains the economics of a size-dependent policy in the context of environmental regulation. A simplified static model is used to establish a basic understanding. Subsequently, a more comprehensive general equilibrium model of firms' dynamics is employed to justify the various empirical findings. This includes a quantitative assessment of the impacts on firms' size distribution, aggregate output, and emissions, aligning with the documented facts.

3 A Simple Model of Emission and Exemption

In this section, I develop a straightforward model of heterogeneous GHG emitters to provide insight into the implications of exemption.

3.1 Economic Environment

The model represents a static economy with a continuum of firms producing a homogeneous dirty good using a dirty resource[1] owned by identical households. These households

[1]This can be interpreted as any fossil fuel used for production.

supply their endowment inelastically to firms. The firms' production process generates GHG emissions by utilizing the only available resource. Firms' productivity, denoted by ω, is drawn from an exogenous distribution[2], represented by $G(\cdot)$. The regulator is concerned with firms' pollution and implements a carbon tax policy. Firms emitting above the emission threshold $(\bar{e})$ pay a carbon tax (τ) for each unit of emission, and the carbon tax revenue is redistributed to households as a lump sum transfer. Larger polluters emitting beyond the threshold pay the carbon tax, while those operating below the threshold are exempt.

3.2 Technology

Firms learn about their productivity ω and the regulatory instruments $(\tau, \bar{e})$ before producing. Denoting $e(\omega)$ as the emission of firm ω and $q(\omega)$ as the associated level of production, the emission technology, jointly with production, is expressed as follows:

$$\begin{cases} q(\omega) = & \omega f(R(\omega)) \\ e(\omega) = & h\big(q(\omega)\big) \end{cases} \tag{1}$$

Here, R is the amount of resources used, and f and h are functions with the following properties: $f(0) = 0$, $f'_R(.) > 0$, $f''_R(.) < 0$; $h(0) = 0$, $h'_q > 0$.

Firms aim to maximize their total profit, net of the cost of regulation, which depends on their emission level relative to the exemption threshold $\bar{e}$. Let P_R be the price of the resource relative to the numeraire, and $\pi_\omega(\bar{e}, \tau)$ represents the profit of firm ω given the instruments:

$$\max_{R(\omega)} \pi_\omega(\bar{e}, \tau) = \begin{cases} q(\omega) - P_R R(\omega) & \text{if } e(\omega) \leqslant \bar{e} \\ q(\omega) - P_R R(\omega) - \tau e(\omega) - C_f & \text{otherwise} \end{cases} \tag{2}$$

Only firms operating above the exemption threshold pay the effluent charge τ. They also bear an additional fixed cost C_f related to regulatory compliance for services like measuring and reporting yearly emissions. The optimal dirty resources demand for firms, derived from the first-order condition, is given by:

[2]The literature often favors a Pareto distribution (Melitz, 2003; Tombe and Winter, 2015; Egger et al., 2021) for productivity.

$$R^*(\omega) = \begin{cases} f'^{-1}\left[\dfrac{P_R}{\omega}\right] & \text{if } e(\omega) < \bar{e} \\ f^{-1}\left[\dfrac{h^{-1}(\bar{e})}{\omega}\right] & \text{if } e(\omega) = \bar{e} \\ f'^{-1}\left[\dfrac{P_R}{\omega(1-\tau h_q^{'})}\right] & \text{if } e(\omega) > \bar{e} \end{cases} \tag{3}$$

Note that firms' demand for resources is an increasing convex function of productivity level for large and small emitters but a decreasing convex function for those operating at the exemption threshold. The input demand for two firms with intermediate level of productivity ω_1 and ω_2 clustered at the exemption threshold is reversed, with the former being more efficient than the latter. Moreover, the asymmetry in the emission cost borne by firms with the differential treatment induced by the exemption in the regulation creates a kink in firms' total cost and, therefore, in the emission distribution. Denoting the subscript "b" and "a" respectively for firms "below" and "above" the exemption threshold, and "c" for the distorted emitters "constrained" at the $\underline{e}$, the result is formalized by the following proposition:

Proposition 1. *There exist two productivity thresholds $\underline{\omega}$ and $\bar{\omega}$ defined by*

$$\underline{\omega} \equiv: \pi_b(e_b(\underline{\omega}), 0) = \pi_c(\bar{e}, 0), , \textit{and}, , \bar{\omega} \equiv: \pi_a(e(\bar{\omega}), \tau) = \pi_c(\bar{e}, 0) \tag{4}$$

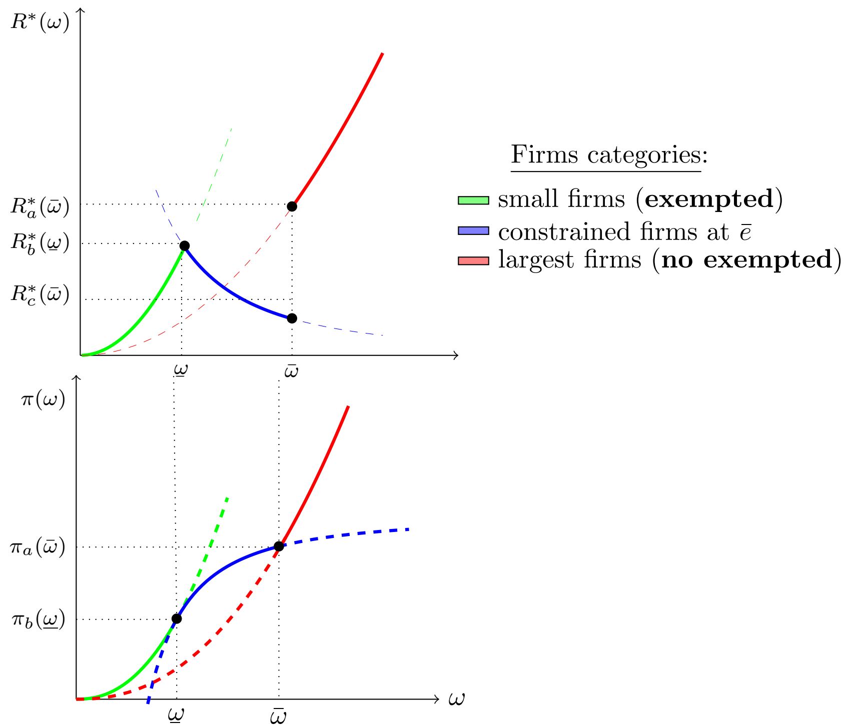
such that: i) firms with low productivity ($\omega < \underline{\omega}$) emit below $\bar{e}$ and become exempted; ii) large productive firms ($\omega > \bar{\omega}$) produce above $\bar{e}$; iii) intermediate productive firms $\omega, \in, (\underline{\omega}, \bar{\omega})$ produce at $\bar{e}$ to avoid paying the emission tax.

The condition for inequality, $e(\omega) < \bar{e}$, allows firms to legally avoid the carbon tax. This establishes a productivity threshold[3] denoted as $\underline{\omega}$. Small emitters operate below this threshold, while intermediate and constrained firms operate above it. Large emitters, polluting beyond $\bar{e}$, are those with productivity surpassing a specific threshold $\bar{\omega}$, allowing them to easily cover the regulatory cost. The productivity cut-off $\bar{\omega}$ is determined by the indifference condition between bunching at $\bar{e}$ or increasing emissions to a higher level with productivity above $\bar{\omega}$.

Given the policy instruments $(\tau, \bar{e})$, the equilibrium is defined by a system of three equations with three unknowns: $\underline{\omega}$, $\bar{\omega}$, and P_R. For simplicity, I assume negligible govern-

[3]The same solution can be found by intersecting the emission equation of small firms with the exemption threshold: $h\bigg(\underline{\omega} f(R_b^*(\omega))\bigg) - \bar{e} = 0$.

Figure 3. Implication of exemption and firms categories



ment administrative costs associated with monitoring the regulation. Appendix 5.3 shows the existence of a solution and evaluates the sign (and magnitude) of changes in the endogenous variables when regulatory instruments vary. Results derived from the implicit function theorem are summarized in the following proposition:

Proposition 2.

1. *An increase in the carbon tax decreases $\underline{\omega}$ and P_R^* in equilibrium. Consequently, the mass of firms below the threshold decreases. The effect on $\bar{\omega}$ is ambiguous.*

2. *An increase in the exemption threshold increases $\bar{\omega}$ and P_R^* in equilibrium. Consequently, the mass of firms above the exemption threshold increases. The effect on $\underline{\omega}$ is ambiguous.*

The proof is in Appendix 5.3. Intuitively, the overall impact of a variation in each instrument on endogenous variables consists of both direct effects and general equilibrium effects. The direct effects are measured under constant prices. When regulatory costs rise, efficient firms operating above the exemption threshold reduce their demand for inputs. This causes an increase in $\bar{\omega}$ (the *direct effect* on $\bar{\omega}$) and a drop in resource prices due to inelastic supply. Consequently, input demand for small firms increases while everything else remains constant. This causes previously exempted firms near $\underline{\omega}$ to enter the category of firms clustering at the exemption threshold, leading to higher production and emissions due to the decrease in production costs P_R. Furthermore, previously constrained firms near $\bar{\omega}$ take advantage of P_R to transition into the largest firms category, resulting in a decrease in $\bar{\omega}$ (the *general equilibrium effect* on $\bar{\omega}$). Consequently, the overall effect of an increase in the emission price on $\bar{\omega}$ is the combination of these two opposing effects. The outcome depends on which effect is stronger. The aggregate effect is considered positive (i.e., $\frac{d\bar{\omega}}{d\tau} > 0$) when the price motive alone is insufficient to counterbalance the direct effect.

Similarly, an increase in the exemption threshold follows the same logic. The direct effect is a surge in the mass of exempted firms and a decline in those above, as both $\underline{\omega}$ and $\bar{\omega}$ increase. This results from the output increase of previously constrained firms near $\bar{e}$ (now becoming exempted) and previously large firms near $\bar{\omega}$ (now constrained at $\bar{e}$). In turn, the increase in the general equilibrium price associated with the motionless supply curve has an indirect opposite effect on small firms, potentially offsetting the positive direct effect: $\frac{d\underline{\omega}}{d\bar{e}} > 0$ when the direct effect counterbalances the general equilibrium effect.

3.3 Comparison of Policies

To thoroughly evaluate the impact of exemptions in environmental regulation, I compare their effects on aggregate variables with two alternative regulatory policies.

The first alternative policy, termed the *no-exemption regime*, involves all firms bearing the full cost for each unit of emitted pollution. The second alternative, known as the *full-exemption regime*, represents a laissez-faire scenario with no regulatory measures in place. Under the *no-exemption regime*, the exemption is eliminated, and all firms' emissions follow the distribution depicted by the red curve in Figure 3 ($\bar{e} = 0$, $\tau > 0$). Conversely, the *full-exemption regime* represents a laissez-faire situation with no regulatory measures ($\tau = 0$), resulting in all firms following the distribution illustrated by the green curve. In both cases, the levels of $\underline{\omega}$ and $\bar{\omega}$ are set to zero.

Firms with an intermediate level of productivity, which operate at the threshold emission level under the baseline scenario with an exemption, experience a decrease (or increase) in production levels under the *no-exemption regime* (or *full-exemption regime*). For simplicity, let's define the production function $f(R)$ as having decreasing returns to scale, characterized by the specification $f(R) = R^{\theta}$, where $0 < \theta < 1$. The firms' emissions are modeled as a linear function of output, denoted as $h\big(q(\omega)\big) = \sigma(\omega)q(\omega)$, where $\sigma(\omega) > 0$

represents the firm-specific emission rate.

In the analysis, the labels *fe* and *ne* represent the first and second alternative policies, corresponding to full-exemption and no-exemption, respectively. The baseline policy, where larger firms comply with regulations by covering the entire cost of emissions, is denoted as *be*. Based on the equilibrium, the following properties are derived:

Proposition 3. *Aggregate variables under the different regimes are characterized by:*

1. $P^*_{R;fe} \geqslant P^*_{R;ne}$; $P^*_{R;fe} \geqslant P^*_{R;be}$; $P^*_{R;be} \lesseqqgtr P^*_{R;ne}$
2. $R^*_{fe}(\omega) = R^*_{ne}(\omega) = \omega^{1/(1-\theta)} / \left(\int_0^\infty \omega^{1/(1-\theta)} dG(\omega) \right)$
3. $Q^*_{fe} = Q^*_{ne} \geqslant Q^*_{be}$;

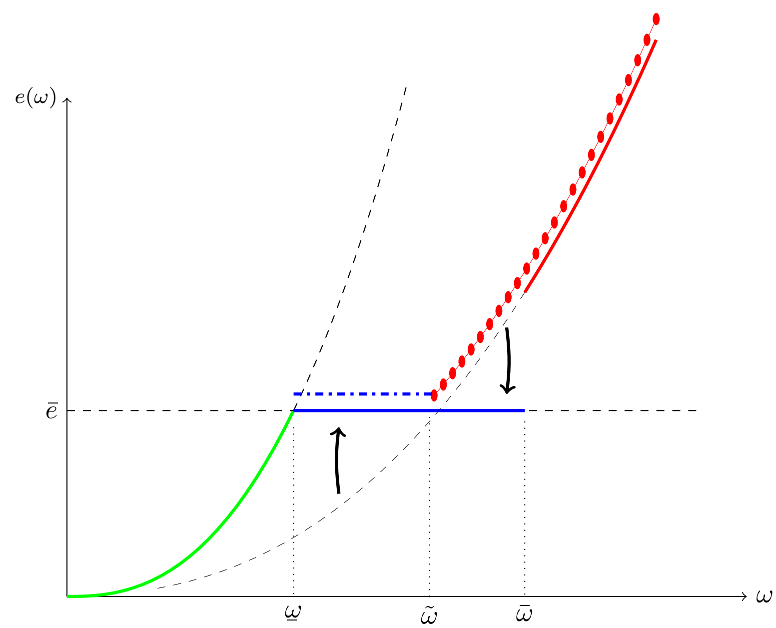
The proof is provided in Appendix 5.3. The rationale behind the first point of the proposition is simple. Firms' demand for inputs is maximal under the *full-exemption regime* and minimal under the *non-exemption regime*, given that no firms (or all firms, respectively) bear the burden of the carbon tax. Therefore, comparing the general equilibrium price that satisfies the resource condition for each case is straightforward, considering the inelastic labor supply.

However, comparing input prices between the baseline and non-exemption cases is more intricate. Transitioning from a policy scenario without an exemption to one with an exemption leads to an increase in output production (resulting in an increase in input prices) for small productive firms operating below the exemption threshold ($\omega \leqslant \bar{\omega}$). This positive effect on the equilibrium price is what is represented by the first term of the equations in (5). Simultaneously, firms with a level of emission located near $\bar{e}$ from below will increase their resource and thus, output demand (resulting in an increase in input prices), while those located near $\bar{e}$ from above will reduce their input and output demand (resulting in a decrease in input prices). Both types of firms previously operating near $\bar{e}$ under the non-exemption regime will cluster at $\bar{e}$ to circumvent the regulation under the be. The overall effect on input prices induced by these firms is ambiguous due to the two conflicting effects. This is what is represented by the first term of the equations in (5) for which the sign is unknown. Figure 4 displays a visual representation of the change in firms' emissions, output, and resource demand when exemptions are added to the regulation, where previously all firms were covered. It illustrates how the categories of these figures change when transitioning from a carbon policy without exemptions to a carbon tax with exemptions.

$$
\begin{aligned}
P_{R;be}^{1/(1-\theta)} - P_{R;ne}^{1/(1-\theta)} &= \int_0^{\underline{\omega}} \left(\underbrace{\left(\omega\theta\right)^{1/(1-\theta)} - \left(\omega\theta(1-\tau\sigma(\omega))\right)^{1/(1-\theta)}}_{\geqslant 0} \right) dG(\omega) \\
&\quad + \underbrace{\int_{\underline{\omega}}^{\bar{\omega}} (\omega\theta)^{-1/\theta} \left((\underline{\omega}\theta)^{1/\left(\theta(1-\theta)\right)} - (1-\tau\sigma(\omega))^{1/(1-\theta)} \left(\omega\theta\right)^{1/\left(\theta(1-\theta)\right)} \right) dG(\omega)}_{\gtreqless 0}
\end{aligned} \tag{5}
$$

The last two points of Proposition 3 are outcomes of the price analysis. When firms receive uniform treatment, either through full exemption or full taxation, there is no alteration in resource allocation, and therefore, in aggregate production.

Figure 4. Transition from carbon without exemption to a carbon tax with exemption



Notes.- This figure displays how figures categories changes when moving from carbon without exemption to a carbon tax with exemption.The exemption is threshold-specific

3.4 Optimizing Emission Tax in Second-Best Scenario

Previous section shows that the exemption provision leads to output inefficiency. In this section, I derive the second-best optimal carbon tax under incomplete regulation given an exemption-threshold. The planner determines the carbon tax to maximize the indirect social welfare function. This function combines both firm profits and government revenue while considering emission externality. The process involves substituting the equilibrium competitive allocation, functions of the policy instruments $(\tau, \bar{e})$, into each component of the welfare function. Following insights from optimal taxation literature (Diamond and Mirrlees, 1971; Yitzhaki, 1979), I assume that changes in taxes do not affect general equilibrium prices.

$$\max_{\tau} W(\tau, \bar{e}) = \max_{\tau} \left\{ \underbrace{\int_0^{\infty} \pi(\tau, \bar{e}) dG(\omega)}_{\text{Firm profits}} + \underbrace{\tau \int_0^{\bar{\omega}} e(\tau, \bar{e}) dG(\omega)}_{\text{Government revenue}} - \underbrace{\lambda \int_0^{\infty} \sigma(\omega) q(\tau, \bar{e}) dG(\omega)}_{\text{Emission externality}} \right\} \tag{6}$$

Here, the positive parameter λ denotes the social marginal damage related to environmental quality. Denoting $\mathcal{E}_\tau = \dfrac{d \log \big(1 - G(\bar{\omega})\big)}{d \log(\tau)}$ as the elasticity of the mass of large productive firms paying the emission cost with respect to a change in the emission tax τ, the optimal rule for the carbon tax is as follows:

Proposition 4.

1. *Given $\bar{e}$, the optimal carbon tax τ satisfies:*

$$\frac{\theta}{1-\theta}(\lambda - \tau^*) \mathbb{E}_\omega \left(\frac{\tau^* \sigma(\omega)}{1 - \tau^* \sigma(\omega)} e(\omega) \,\middle|\, \omega > \bar{\omega} \right) = \Big(\lambda \bar{e} - (\lambda - \tau^*) e_a(\bar{\omega}) \Big) \mathcal{E}_\tau \tag{7}$$

2. *Particularly, $\tau^* = \lambda$ under "no-exemption regime"*

The details of the proof can be found in Appendix 5.3. The optimal tax rate is characterized by equation (7). Here, the left-hand side represents the social benefit of increasing the carbon tax, leading to large emitters to reduce emission. The right-hand side represents the excess burden caused by the exemption threshold. Given a specific exemption level, any deviation from the first-best carbon tax level by the regulator results in increased deadweight loss. This is due to a higher proportion of constrained firms.

In the absence of an exemption, the equation suggests setting the tax rate to the marginal damage of emissions, denoted as λ (Pigouvian tax), because the right-hand side of equation (7) is zero. This necessitates $\tau^* = \lambda$.

4 An Extended Industry Model with Emission and Exemption

In this section, I embed the mechanism of the simplest model in a quantitative model by adding additional features of emission-intensive sectors. The model builds on industry dynamics models from Hopenhayn (1992) and Hopenhayn and Rogerson (1993). In this model, only large emitters above the exemption threshold pay the carbon tax. The economy comprises heterogeneous firms, a representative consumer, and a government that collects the carbon tax income and redistributes it to households in a lump-sum manner. Firms produce homogeneous goods using labor and capital. Time is discrete, and the model horizon is infinite. The two policy instruments remain unchanged. However, in this setting, firms' emission rate is drawn from a distribution G independent of the distribution of productivity. The heterogeneity in emission rate captures firms' heterogeneity in emission technologies, which is not necessarily correlated with firms' productivity[4].

4.1 Firms and Technology

Firms differ in productivity and emission rate. The productivity shock follows a persistent AR(1) process $\log(\epsilon_{t+1}) = \rho_\epsilon \log(\epsilon_t) + \nu_t$ with $\nu_t \sim \mathcal{N}(0, \sigma_\epsilon)$. Firms' emission rate, σ, is drawn from the uniform distribution $\mathcal{U}([\Sigma_{lb}, \Sigma_{ub}])$. After production takes place, firms pay for the associated emission if that emission is higher than the exemption threshold.

The joint emission-production technology follows Copeland and Taylor (1994, 2004) and Shapiro and Walker (2018), allowing firms to reduce emissions by allocating an endogenous fraction $\chi \in [0, 1]$ of inputs to abatement:

$$\begin{cases} y = (1-\chi)\epsilon n^\alpha k^\gamma, \\ e = (1-\chi)^{1/\upsilon}\sigma y \end{cases} \tag{8}$$

The production function exhibits decreasing returns to scale with $\alpha + \gamma < 1$. Emissions (e) are represented as an increasing linear function of production and a decreasing convex function of the proportion (χ). Without abatement, one unit of output generates σ units of emission, and less with abatement. The pollution elasticity (υ) reflects the effectiveness of

[4]While several earlier studies suggest a negative relationship between firms' productivity and emission rates, more recent papers, such as Dardati and Saygili (2020) and Fowlie et al. (2016b), propose an independent distribution for both variables when conducting their joint distribution analysis. In reality, productive firms can exhibit varying emission rates depending on the quality of their emission technology.

abatement technology: higher υ implies less pollution reduction. By substituting $(1-\chi)$ from one expression into the other, production can be rewritten as:

$$y = \left[\frac{e}{\sigma}\right]^{\upsilon} \left[\epsilon n^{\alpha} k^{\gamma}\right]^{1-\upsilon}$$

where emission becomes an input to the production function with a constant return to scale, a concept widely used in the literature since Pethig (1976) and Copeland and Taylor (1994).

4.2 Incumbent Firms

Incumbents' static problem consists of hiring labor at wage rate w_t, buying capital at interest rate r, and paying for emissions generated if above the threshold. Conditional on the production and pollution efficiency representing the state variables, the firms' value function is:

$$\begin{aligned} V(\epsilon,\sigma) &= \max_{n,k,\chi} \left\{ \pi(\epsilon,\sigma) + (1-\kappa)\beta \int_{\epsilon'|\epsilon} V(\epsilon',\sigma) dH(\epsilon'|\epsilon) \right\} \\ \text{s.t.} \quad \pi(\epsilon,\sigma) &= py - wn - rk - \tau e \mathbb{1}_{e>\bar{e}} \\ y &= (1-\chi)\epsilon n^{\alpha} k^{\gamma} \\ e &= (1-\chi)^{1/\upsilon} \sigma y \end{aligned} \tag{9}$$

Here, p_t denotes the price of the output good and $(1-\kappa)$ is the exogenous probability of firms' survival. Once production occurs, firms learn their emissions generated and have to comply with the regulation by paying the emission cost depending on the relative emission to the exemption threshold. Only firms polluting above the exemption threshold $\bar{e}$ are constrained by the regulation and pay for the total emission.

Optimal choices.- Solving the firm's problem involves considering a nonlinear carbon taxation plan that distinguishes between three types of firms: small emitters below the threshold, exempt from the carbon tax; constrained firms emitting at $\bar{e}$; and large emitters exceeding $\bar{e}$. The disparity in regulation costs, due to exemptions, creates differential incentives for firms to adopt abatement technology. For a positive carbon tax, the optimal abatement is given by:

$$\chi^* = \begin{cases} 0 & \text{if } (e_t \leqslant \bar{e}) \\ 1 - \left(\frac{\upsilon}{\tau\sigma}\right)^{\upsilon/(1-\upsilon)} & \text{if } (e_t > \bar{e}) \wedge \left(\sigma \geqslant \frac{\upsilon}{\tau}\right) \end{cases} \tag{10}$$

where $\wedge$ is the logical conjunction operator. Small polluters below $\bar{e}$ lack incentive for abatement ($\chi^* = 0$). Larger emitters above the threshold opt for abatement when their emissions surpass a certain level, which decreases with the carbon tax and increases with pollution elasticity. In this case, the share of resources used for production purposes $(1-\chi^*)$ decreases with the tax. The capital demand for each firm category is:

$$k(\epsilon_t, \sigma_t) = \begin{cases} A_1 \epsilon_t^{1/(1-\psi)} & \text{if } (e_t < \bar{e}) \\ A_2 \epsilon_t^{-1/\psi} & \text{if } (e_t = \bar{e}) \\ A_1 \left[(1-\upsilon)(1-\chi^*)\epsilon_t\right]^{1/(1-\psi)} & \text{if } (e_t > \bar{e}) \wedge \left(\sigma \geqslant \dfrac{\upsilon}{\tau}\right) \\ A_1 \left[(1-\sigma_t\tau)\epsilon_t\right]^{1/(1-\psi)} & \text{if } (e_t > \bar{e}) \wedge \left(\sigma < \dfrac{\upsilon}{\tau}\right) \end{cases} \tag{11}$$

where $\psi = \alpha + \gamma < 1$ and A_1, A_2 are functions of input prices w and r[5]. The capital demand shows a distinct pattern concerning firms' productivity: it decreases concavely for firms operating at the exemption threshold, while it increases convexly for firms operating below or above the threshold. Firms above the exemption threshold experience a decrease in optimal capital demand in response to the carbon tax or permit price. Labor demand is derived from the ratio of the marginal product of the two inputs, which equals the ratio of input prices at equilibrium. In the absence of regulation ($\tau = 0$) or when firms above the threshold are much cleaner ($\sigma \to 0$), both capital and labor remain constant for small and large polluters.

The Bellman equation in (9) increases and remains continuous in productivity ϵ. Therefore, firms' optimal category choice can be characterized by two pairs of productivity cut-off rules $\underline{\epsilon}$ and $\bar{\epsilon}$, both functions of firms' emission rate. Specifically, given the emission rate, firms with productivity levels below $\underline{\epsilon}$ are exempted from pollution charges. Firms with intermediate productivity levels falling between the two cut-offs, $\underline{\epsilon}(\sigma) \leqslant \epsilon \leqslant \bar{\epsilon}(\sigma)$, cluster around the emission level of $\bar{e}$. On the other hand, regulated firms are those with higher productivity levels given the emission rate.

4.3 Entrant Firms

A significant number of firms enter the industry as long as their expected discounted value exceeds the fixed entry cost c_e. The entry condition is:

---

[5] $A_1 = \left(\dfrac{\gamma}{r}\right)^{(1-\alpha)/(1-\psi)} \left(\dfrac{\alpha}{w}\right)^{\alpha/(1-\psi)}$, $A_2 = \left(\dfrac{\bar{e}}{\sigma}\right)^{1/\psi} \left(\dfrac{r\alpha}{w\gamma}\right)^{-\alpha/\psi}$

$$\int_{\epsilon,\sigma} V(\epsilon,\sigma)dH_0(\epsilon)dG(\sigma) - pc_e = 0 \tag{12}$$

After covering the initial entry cost, new entrants randomly select an efficiency level from the stationary density of the productivity shock, denoted as H_0. The emission rate is drawn from the same distribution G as that of incumbent firms.

4.4 Cross-Sectional Distribution

The law of motion for the cross-sectional distribution of firms Γ evolves as follows:

$$\Gamma(\epsilon',\sigma) = (1-\kappa)\int_{\epsilon}\Gamma(\epsilon,\sigma)dH(\epsilon'|\epsilon) + \mathcal{M}\int_{\epsilon} dH_0(\epsilon)dG(\sigma) \tag{13}$$

Here, $\mathcal{M}$ represents the mass of new entering firms. Equation (13) shows the evolution of active establishments between two consecutive periods. The first term represents the mass of survived incumbents transitioning from states ϵ to ϵ', and the second term represents the total mass of new entrants.

4.5 Households

There is a representative household that supplies labor inelastically to firms and derives utility from output consumption. Given the aggregate emission in the economy, E, households maximize lifetime utility subject to a budget constraint where total revenue consists of labor income, w, capital rental income, $r_t K_t$, dividends from aggregate profits, and lump-sum transfers from the regulator:

$$\max_{C_t, K_{t+1}} \sum_{t=0}^{\infty} \beta^t \left[U(C_t) - \nu E_t\right] \tag{14}$$

$$\text{s.t.} \quad p_t\left(C_t + K_{t+1} - (1-\delta)K_t\right) = w_t + r_t K_t + Tr_t + \Pi_t \tag{15}$$

The total labor supply is normalized to 1, and ν captures the marginal disutility of pollution. The long-run interest rate, derived from the first-order condition, is a function of the discount rate β and the capital depreciation rate δ:

$$r = \frac{1}{\beta} - (1-\delta) \tag{16}$$

4.6 Stationary Equilibrium

Let $s = \{\epsilon, \sigma\} \in S$ be a generic state vector. Given policy instruments $(\bar{e}, \tau)$, a stationary recursive equilibrium consists of a value function $V(s)$, labor demand $n(s)$, capital demand $k(s)$, abatement choice $\chi(s)$, two productivity cut-offs $(\underline{\epsilon}(s), \bar{\epsilon}(s))$, a wage rate w, an interest rate r, a distribution of firms $\Gamma(s)$, and a mass of entrants $\mathcal{M}$ such that:

(i) $V(s)$, $n(s)$, $k(s)$, $\chi(s)$, $\underline{\epsilon}(s)$, and $\bar{\epsilon}(s)$ solve the incumbent's problem.

(ii) The free entry condition (12) holds.

(iii) Households optimize.

(iv) $\Gamma(s)$ is stationary: $\Gamma'(s) = \Gamma(s)$.

(v) Market clearing conditions hold:

(a) Labor market clears: $1 = \int_{s \in S} n(s) d\Gamma(s)$

(b) Goods market clears: $C + \int_{s \in S} \Big[\delta k(s) - y(s) \Big] d\Gamma(s) + c_e \mathcal{M} = 0$

5 Quantitative Exploration of Size Dependent Exemption

In this section, I calibrate the model to replicate key facts about the California economy. Then, I use the model as a laboratory to conduct experiments and simulate policy reforms.

5.1 Calibration

The model parameters are categorized into two groups. The first group consists of parameters obtained from existing literature, which remain fixed. The second group includes parameters that are calibrated to minimize the distance between model-generated moments and observed data.

Table 2a lists the first group of parameters, their values, and sources. The discount factor β is set at 0.96, consistent with an annual gross interest rate of 4%. The price of output goods is normalized to 1. Following Restuccia and Rogerson (2008), labor and capital shares α and γ are set at 0.283 and 0.567, respectively. The depreciation rate is set at 0.10, as in Spencer (2022). The lower bound of the emission rate, Σ_{lb}, is normalized to zero.

The remaining parameters (σ_ϵ, ρ_ϵ, Σ_{ub}, τ, υ) are calibrated within the model. Table 2b shows these parameters alongside the comparison of data and model moments. The selected moments effectively identify the parameters. Here's the intuition: productivity parameters σ_ϵ and ρ_ϵ influence the distribution of firm size and employment. I target the share of firms with fewer than 5 and more than 50 employees to determine these parameters. τ is identified to match the dispersion measure $P90/P50$ ratio of emissions. Based on the 2005 Pollution Abatement Costs and Expenditure (PACE) survey, the average air pollution abatement share is 1.9%. Thus, the abatement elasticity υ is set to achieve this value.

Table 2. Calibration

(a) Preset parameters

Parameters	Meaning	Value	Sources/Targets
β	Discount factor	0.96	4% interest rate
δ	Depreciation rate	0.10	US data
α	Labor income share	0.283	Restuccia and Rogerson (2008)
γ	Capital income share	0.567	Restuccia and Rogerson (2008)
κ	Exit rate	0.10	US data
c_e	Entry cost	0.01	Normalization
p	Price of final output	1	Normalization
Σ_{lb}	Lower bound of emission rate	0	Normalization

(b) Internally calibrated parameters

Parameters	Meaning	Value	Targets	Data	Models
Σ_{ub}	Emission rate upper bound	1154.35	Average emission[a]	0.517	0.510
σ_ϵ	Productivity volatility	0.0290	Share of firms with $\leqslant 5$ employees[b]	0.698	0.590
ρ_ϵ	Productivity persistence	0.9863	Share of firms with $\geqslant 50$ employees	0.041	0.045
τ	Carbon tax	0.0002	Ratio P90/P50 of emission	19.35	19.12
υ	Abatement elasticity	0.0088	Average share of emission abatement	0.019	0.019

[a] Emission is in million tonnes of CO2 equivalent. As result, the exemption threshold is 0.025.
[b] Data on employment and firms size comes from the Employment Development Department of California

Overall, the model reproduces the targets successfully, closely aligning data and model moments. The estimated abatement elasticity is 0.009, within the range used in the literature. Shapiro and Walker (2018) suggest a range of 0.001 to 0.0557, and Dardati and Saygili (2020) adopts the upper bound value in a recent study.

To validate the model, I assess its ability to replicate additional, non-targeted moments and salient features. Specifically, I compare the distributions of establishments, employment, and emissions implied by the model with their counterparts in the data. Table 3a presents results for the first two distributions, while Table 3b shows the emission distribution by quintile. The calibration targeted only the share of establishments with fewer than

5 employees and over 50 employees. Yet, it effectively replicates the overall establishment and employment distributions across firm sizes. The share of emissions by quintile generated by the model also aligns closely with the data. These consistent matches instill confidence in using the model to evaluate the quantitative implications of the exemption policy and to conduct counterfactual experiments.

Table 3. Untargeted Moments

(a) Distribution of Establishment and Employment

	Firms size	< 5	5-9	10-19	20-49	$\geqslant 50$
Data	Employment	0.081	0.067	0.098	0.163	0.591
	Establishments	0.698	0.116	0.083	0.062	0.041
Calibrated model	Employment	0.092	0.099	0.155	0.257	0.396
	Establishments	0.590	0.154	0.120	0.091	0.045

(b) Distribution of Emission (Share of emission by emission size)

Quintiles of emission	Quintile1	Quintile2	Quintile3	Quintile4	Quintile5
Data	0.003	0.007	0.016	0.053	0.919
Calibrated model	0.000	0.021	0.151	0.385	0.440

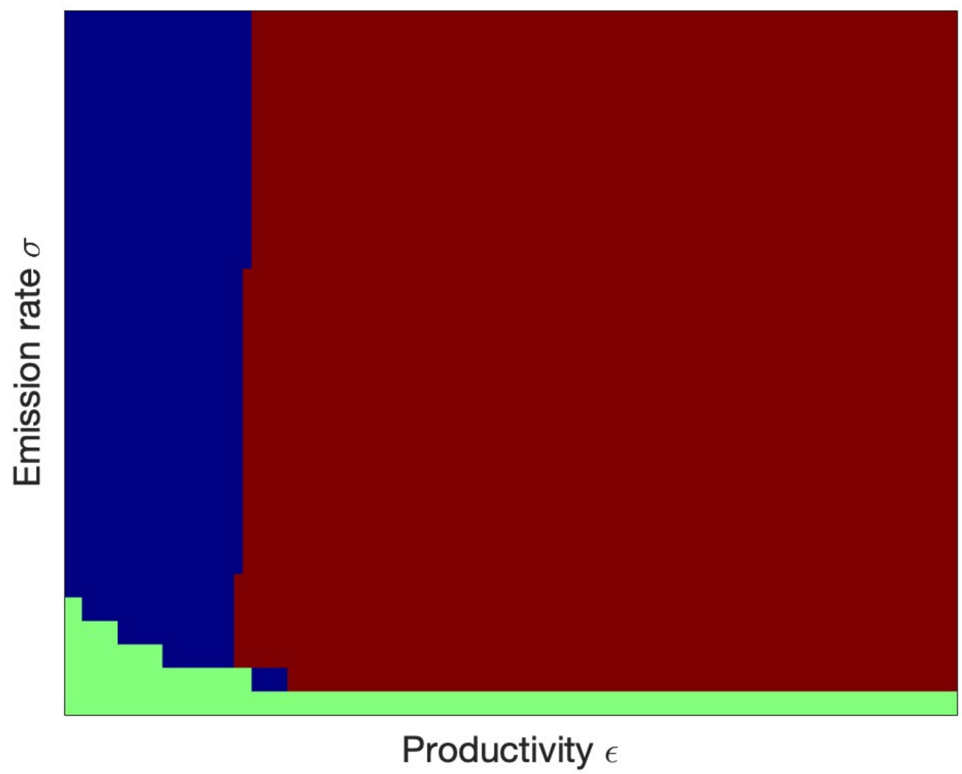
5.2 Distribution of Firms Categories

Figure 5 shows the distribution of firm categories based on their emissions relative to the exemption threshold. It highlights differences between each category under the baseline policy scenario with size dependent exemption for small emitters. Green areas represent exempted firms. Blue areas show constrained firms near $\bar{e}$. Brown areas depict large emitters exceeding $\bar{e}$. Each area indicates the significance of each category.

The distribution of firm categories aligns with earlier qualitative findings from the simpler model. For a given emission rate, firms constrained at the exemption threshold have intermediate productivity, and the share of firms concentrated at that emission level is about 6.5%. Consequently, the proportion of firms operating above (below) the threshold is 89.7% (3.7% respectively).

Full-exemption (no-exemption) cases, characterized by only green (blue) areas across productivity and emission levels, are intentionally omitted from the distribution display.

Figure 5. Distribution of firms category



Note: The figure illustrates whether the firms are below, exactly at or above the exemption threshold depending on their productivity and emission rate. I refer to that as the distribution of firms categorized. The green surface represents exempted firms, the blue surface depicts constrained firms that cluster around $\bar{e}$, and the brown color indicates large emitters exceeding $\bar{e}$.

5.3 Implications of Exemption Under Alternative Policy

This section evaluates the macroeconomic impact of a size dependent exemption in regulatory costs by comparing changes in aggregate variables when the exemption threshold is eliminated (*no-exemption*).

To ensure a fair comparison and assess the impact of each policy, aggregate variables in the baseline model are normalized to 100. The carbon tax under each alternative is re-calibrated to achieve the same emission level as the baseline. This re-calibration ensures the comparability of policies in terms of aggregate variables with the same emission reduction target.

In Appendix 5.3, I present the results before the re-calibration exercise to validate the findings. Interestingly, aggregate emissions are lower under the no-exemption policy, where all firms bear the emission costs. Other aggregate variables show similar trends, as under the same aggregate emission in Table 4. The effects of exemptions on aggregate variables after re-calibration increase in the same direction as before, reaching the same emission objective.

Consider first the elimination of the exemption provision (*no-exemption*) with unchanged aggregate pollution. In this scenario, all firms pay their emission costs, eliminating the previously observed distorted bunching behavior. As a result, aggregate productivity increases by 0.02% compared to the baseline with exemptions. Firms that were previously exempt now face regulatory costs, leading them to reduce their labor demand. This reduction in labor demand increases the mass of new entrants, which clears the labor market.

The cost of labor from the entry condition increases due to three forces: first, a decrease in firms' values due to the additional emission cost for previously exempted firms; second, an increase in firms' value from better input allocation of productive firms induced by removing the exemption; and finally, a decrease in the carbon tax to reach the same emission level as in the baseline. Removing the exemption increases the size of firms with over 50 employees, resulting in a 0.03% increase in their share. This highlights the significant distortion caused by the exemption threshold instrument, as resources are inefficiently allocated towards less productive firms.

Table 4. Percentage variation from benchmark to no-exemption

Aggregate variables	$\Delta(\%)$
Emission	0
Output	+ 0.05
Capital	+ 0.06
TFP	+ 0.02
Mass of new entrant	+ 0.01
Wage rate	+ 0.06

Dispersion within and across categories.- To assess the distortion caused by the exemption policy, Table 5a quantifies the variation in individual Total Factor Productivity-revenue (TFPR) by calculating both weighted and unweighted standard deviations from the average TFPR across firms. Following Hsieh and Klenow (2009), a firm's "Revenue Productivity" ($TFPR$) is expressed as:

$$TFPR = \frac{y^{\alpha+\gamma}}{(wn)^{\alpha}\,(rk)^{\gamma}} \tag{17}$$

In the absence of distortion, TFPR is constant across all firms. However, when firms are treated differently, especially when productive firms face higher taxes than unproductive ones, TFPR is no longer constant. This variation may result in the misallocation of resources.

Let $\omega_{i,j}$ represent the weight of firm i in category j, where $j = 1, 2, 3$ denotes firms below, at, or above the exemption threshold, respectively. The firm's weight is given by $\omega_{i,j} = \dfrac{\Gamma_{i,j}}{\sum_{i,j} \Gamma_{i,j}}$, with $\Gamma_{i,j}$ being the mass of active firms derived from equation (13).

In Table 5a, in addition to the standard deviation, the interdecile and interquartile ranges are computed to gauge heterogeneity in firms' TFPR. Results reveal volatility in TFPR across firms, particularly among those in the second category that leak emission costs by operating at the exemption threshold. Essentially, high TFPR volatility in constrained firms at $\bar{e}$ contributes to the overall volatility under the baseline policy scenario with exemption, indicating misallocation of inputs[6]. The average TFPR for firms operating at the exemption threshold is higher than that of firms emitting strictly above the emission threshold, with the former group consisting of firms with an intermediate level of productivity and the latter comprising those with a much higher productivity level.

I also decompose, in Table 5b, the total variance of TFPR, denoted by V, into the variance within (W) and between (B) categories. The former measures the average dispersion across firms within categories, and the latter measures the dispersion between the average TFPR of the three categories. The variance decomposition is given by:

$$V(TFPR_{i,j}) = \underbrace{\sum_{j=1}^{3} \omega_j \big(TFPR_j - TFPR\big)^2}_{B \equiv \text{volatility between-categories}} + \underbrace{\sum_{j=1}^{3} \omega_j \Bigg[\overbrace{\sum_{i \in j} \omega_{i,j} \big(TFPR_{i,j} - TFPR_j\big)^2}^{W_j \equiv \text{within category } j \text{ volatility}} \Bigg]}_{W \equiv \text{average of,} W_j} \tag{18}$$

where $\omega_j = \sum_{i \in j} \omega_{i,j}$ is the weight of each category of firms.

[6] Dispersion within firms strictly below and above the cut-off is respectively zero and small, with a standard deviation of 0.0006. The dispersion within non-exempted firms arises because firms with large emission rates adopt abatement technology while those with smaller emission rates do not.

Table 5. Differences in firms TFPR

(a) Dispersion of TFPR

	All firms	Firms below threshold	Firms constrained	Firms above threshold
Weighted StD	0.19	0	0.31	0
Unweighted StD	0.28	0	0.29	0
P90-P10	0.61	0	0.80	0
P75-P25	0	0	0.41	0
Weighted mean	2.03	1.97	2.68	1.99
Unweghted mean	2.10	1.97	2.61	1.99

(b) Variance decomposition

	Share of variance	
	Between categories	Within categories
Weighted variance	82.69	17.31
Unweighted variance	79.52	20.48

Notes.- Firms weight is computed by $\frac{\Gamma_i}{\sum_i \Gamma_i}$. The variance is decomposed following equation (18).

The main takeaway of the decomposition is that the volatility in TFPR is explained mainly by the volatility between different categories. 83% (80% respectively) of the weighted (unweighted respectively) variance of TFPR occurs between categories. A small percentage of the overall dispersion comes from differences in firms' TFPR within the same category.

Discussion and Conclusion

This study investigates how firms respond to environmental regulations that use a size dependent exemption, where only large emitters face regulatory taxes. I identify negative effects on aggregate measures such as output and Total Factor Productivity (TFP), and a discontinuity in emission and employment distribution. Using a calibrated model with heterogeneous emitters, I analyze outcomes under different policy scenarios.

Results indicate that the exemption imposes firm-specific emission costs, disproportionately affecting productive firms and causing resource misallocation. Removing this provision and treating all firms equally enhances output and TFP while achieving the same emission reduction goal.

These findings question the appropriateness of flexibility in environmental regulation and highlight the need for policymakers to be aware of potential distortions, especially when accommodating small firms in externality correction. How flexible should regulations be, and which firms should they target? This paper initiates a discussion on these questions and underscores the impact on firms' incentives to adopt environmentally friendly technology, in addition to the effects on macroeconomic variables.

References

Berman, Eli and Linda TM Bui, "Environmental regulation and productivity: evidence from oil refineries," *Review of Economics and Statistics*, 2001, *83* (3), 498–510.

Buchanan, James M, "External diseconomies, corrective taxes, and market structure," *The American Economic Review*, 1969, *59* (1), 174–177.

Burtless, Gary and Jerry A Hausman, "The effect of taxation on labor supply: Evaluating the Gary negative income tax experiment," *Journal of political Economy*, 1978, *86* (6), 1103–1130.

Cattaneo, Matias D, Michael Jansson, and Xinwei Ma, "Manipulation testing based on density discontinuity," *The Stata Journal*, 2018, *18* (1), 234–261.

_ , _ , and _ , "Simple local polynomial density estimators," *Journal of the American Statistical Association*, 2020, *115* (531), 1449–1455.

Chetty, Raj, "Bounds on elasticities with optimization frictions: A synthesis of micro and macro evidence on labor supply," *Econometrica*, 2012, *80* (3), 969–1018.

Copeland, Brian R and M Scott Taylor, "North-South trade and the environment," *The quarterly journal of Economics*, 1994, *109* (3), 755–787.

_ and _ , "Trade, growth, and the environment," *Journal of Economic literature*, 2004, *42* (1), 7–71.

Dardati, Evangelina and Meryem Saygili, "Aggregate impacts of cap-and-trade programs with heterogeneous firms," *Energy Economics*, 2020, *92*, 104924.

Dharmapala, Dhammika, Joel Slemrod, and John Douglas Wilson, "Tax policy and the missing middle: Optimal tax remittance with firm-level administrative costs," *Journal of Public Economics*, 2011, *95* (9-10), 1036–1047.

Diamond, Peter A and James A Mirrlees, "Optimal taxation and public production I: Production efficiency," *The American economic review*, 1971, *61* (1), 8–27.

Egger, Hartmut, Udo Kreickemeier, and Philipp M Richter, "Environmental policy and firm selection in the open economy," *Journal of the Association of Environmental and Resource Economists*, 2021, *8* (4), 655–690.

Fowlie, Meredith L, "Incomplete environmental regulation, imperfect competition, and emissions leakage," *American Economic Journal: Economic Policy*, 2009, *1* (2), 72–112.

Fowlie, Meredith, Mar Reguant, and Stephen P Ryan, "Market-based emissions regulation and industry dynamics," *Journal of Political Economy*, 2016, *124* (1), 249–302.

—, —, and —, "Market-based emissions regulation and industry dynamics," *Journal of Political Economy*, 2016, *124* (1), 249–302.

Garicano, Luis, Claire Lelarge, and John Van Reenen, "Firm size distortions and the productivity distribution: Evidence from France," *American Economic Review*, 2016, *106* (11), 3439–79.

Guner, Nezih, Gustavo Ventura, and Yi Xu, "Macroeconomic implications of size-dependent policies," *Review of economic Dynamics*, 2008, *11* (4), 721–744.

Hausman, Jerry A, "The econometrics of nonlinear budget sets," *Econometrica: Journal of the Econometric Society*, 1985, pp. 1255–1282.

Holland, Stephen P, "Emissions taxes versus intensity standards: Second-best environmental policies with incomplete regulation," *Journal of Environmental Economics and management*, 2012, *63* (3), 375–387.

Hopenhayn, Hugo A, "Entry, exit, and firm dynamics in long run equilibrium," *Econometrica: Journal of the Econometric Society*, 1992, pp. 1127–1150.

Hopenhayn, Hugo and Richard Rogerson, "Job turnover and policy evaluation: A general equilibrium analysis," *Journal of political Economy*, 1993, *101* (5), 915–938.

Hsieh, Chang-Tai and Peter J Klenow, "Misallocation and manufacturing TFP in China and India," *The Quarterly journal of economics*, 2009, *124* (4), 1403–1448.

Kaplow, Louis, "Optimal regulation with exemptions," *International Journal of Industrial Organization*, 2019, *66*, 1–39.

Keen, Michael and Jack Mintz, "The optimal threshold for a value-added tax," *Journal of Public Economics*, 2004, *88* (3-4), 559–576.

Kleven, Henrik Jacobsen and Esben Anton Schultz, "Estimating taxable income responses using Danish tax reforms," *American Economic Journal: Economic Policy*, 2014, *6* (4), 271–301.

Li, Zhe and Jianfei Sun, "Emission taxes and standards in a general equilibrium with entry and exit," *Journal of Economic Dynamics and Control*, 2015, *61*, 34–60.

— and Shouyong Shi, "Emission taxes and standards in a general equilibrium with productivity dispersion and abatement," *Macroeconomic Dynamics*, 2017, *21* (8), 1857–1886.

McCrary, Justin, "Manipulation of the running variable in the regression discontinuity design: A density test," *Journal of econometrics*, 2008, *142* (2), 698–714.

Melitz, Marc J, "The impact of trade on intra-industry reallocations and aggregate industry productivity," *econometrica*, 2003, *71* (6), 1695–1725.

Pethig, Rüdiger, "Pollution, welfare, and environmental policy in the theory of comparative advantage," *Journal of environmental economics and management*, 1976, *2* (3), 160–169.

Restuccia, Diego and Richard Rogerson, "Policy distortions and aggregate productivity with heterogeneous establishments," *Review of Economic dynamics*, 2008, *11* (4), 707–720.

Saez, Emmanuel, "Do taxpayers bunch at kink points?," *American economic Journal: economic policy*, 2010, *2* (3), 180–212.

Shapiro, Joseph S and Reed Walker, "Why is pollution from US manufacturing declining? The roles of environmental regulation, productivity, and trade," *American Economic Review*, 2018, *108* (12), 3814–3854.

Spencer, Adam Hal, "Policy effects of international taxation on firm dynamics and capital structure," *The Review of Economic Studies*, 2022, *89* (4), 2149–2200.

Tombe, Trevor and Jennifer Winter, "Environmental policy and misallocation: The productivity effect of intensity standards," *Journal of Environmental Economics and Management*, 2015, *72*, 137–163.

Yitzhaki, Shlomo, "A note on optimal taxation and administrative costs," *The American Economic Review*, 1979, *69* (3), 475–480.

Appendices

Additional facts

To determine whether new and existing firms emit closer or farther from the threshold compared to their incumbent peers, I perform the following regressions:

$$\log\left(\frac{Emission_{j,c,s,t}}{\bar{E}}\right) = \beta_1 \text{Entrants}_{j,c,s,t} + \lambda_t + \lambda_s + \lambda_c + \epsilon_{j,c,s,t} \tag{19}$$

$$\log\left(\frac{Emission_{j,c,s,t}}{\bar{E}}\right) = \beta_2 \text{Exiters}_{j,c,s,t} + \lambda_t + \lambda_s + \lambda_c + \epsilon_{j,c,s,t} \tag{20}$$

$$\log\left(\frac{Emission_{j,c,s,t}}{\bar{E}}\right) = \beta_1 \text{Exiters}_{j,c,s,t} + \beta_2 \text{Entrants}_{j,c,s,t} + \lambda_t + \lambda_s + \lambda_c + \epsilon_{j,c,s,t} \tag{21}$$

In these equations, $\text{Exiters}_{j,c,s,t}$ is a binary variable that equals 1 if establishment j located in city c and operating in sector s exits at time t. A sector is defined as the first 6 digits of the NAICS code. $\bar{E}$ represents the exemption threshold, while λ_t, λ_s, and λ_c control for year, sector, and city fixed effects, respectively. The explained variable captures the extent to which a firm's emission is close to the cutoff, where a negative value indicates that the plant emits below the threshold.

The first two regressions (Equations (19) and (20)) separately estimate the effects of entry and exit, while Regression 3 (Equation (21)) jointly estimates the effects, accounting for potential correlation between entrants and exiters. The estimation results of the coefficients β are reported in Table A1, where columns (1), (2), and (3) correspond to regressions (19), (20), and (21), respectively.

Table A1. Regressions

	(1)	(2)	(3)
Entrant	-0.57^{***}		-0.58^{***}
	(0.15)		(0.15)
Exiter		-2.07^{***}	-2.08^{***}
		(0.58)	(0.58)
R^2	0.54	0.55	0.56
N	7050	7050	7050

Note.-Entrant and Exiter variables are dummy taking 1 if firm is entrant (exiter respectively) or zero otherwise. Robust standard errors are reported in parentheses. Significance levels are denoted as follows: *, **, and *** represent statistical significance at the 10%, 5%, and 1% levels, respectively

On average, new firms entering the market are about 0.57 units closer to the exemption threshold compared to their established counterparts in the same city and sector. Conversely, firms exiting the industry tend to have significantly lower emissions, approximately 2.07 units closer to the exemption threshold, on average, than the established firms in the same sector and city. When contrasting the emissions of exiting firms with those of new entrants, it is evident that exiting firms generally have lower emissions, while new entrants are much closer to the threshold. These findings remain consistent when considering both entry and exit indicators in the same regression.

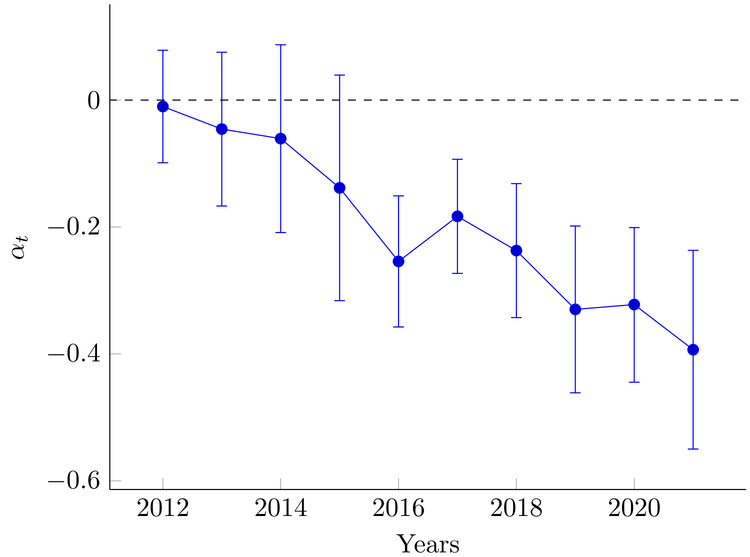
The analysis depicted in Figure A1 highlights a notable decrease in firms' proximity to the exemption cut-off over time compared to the relative distance observed in 2011, pre-regulation. To discern this pattern, I gauge the relative degree of firms' emission proximity using year indicators and fixed effects for sector and city, using 2011 as the reference year:

$$\log\left(\frac{Emission_{j,c,s,t}}{\bar{E}}\right) = \sum_{t=2012}^{2021} \alpha_t \text{Year}_t + \lambda s + \lambda_c + \epsilon_{j,c,s,t}$$

In Figure A1, each estimated coefficient α_t is compared to the reference year 2011, and the vertical bar denotes the 95% confidence interval for each point estimate with robust standard errors. The graph unmistakably illustrates a declining trend in coefficient values over time, becoming particularly pronounced and statistically significant at the 1% level starting from 2016. In simpler terms, firms are increasingly reducing their emission to be near the exemption cut-off compared to their emissions before-emission.

In summary, my findings reveal that new entrant firms tend to bunch near the exemption cut-off, while the emission distribution of incumbent firms remains relatively stable in response to the regulation. Furthermore, both exiting and newly entering firms have emissions closer to the exemption threshold compared to incumbents. Overall, 40% to 50% of firms are exempted for having emissions below the exemption limit.

Figure A1. Evolution of the distance of firms' emission to exemption threshold



Note.- The figure plots α_t from regression (5.3) and gives how close firms (emission) are to the threshold overtime. The vertical bar denotes the 95% confidence interval for each point estimate with robust standard errors.

Discontinuity test

Table A2. Test of discontinuity around the exemption cut-off

Years	2013	2014	2015	2016	2017	2018	2019	2020	2021
	All sample								
Statistics test	0.20	0.58	1.71*	1.08	2.12**	1.25	-0.70	1.93*	2.03**
	Restricted sample								
	-0.11	-0.05	0.8603	-0.07	1.89*	0.50	0.87	-0.18	2.64***

Note: This table reports the discontinuity test around the exemption threshold by following Cattaneo et al. (2018). *, **, *** indicate significance at the 10%, 5%, and 1% level, respectively.

Comparative statics with respect to regulation instruments

The equilibrium is fully determined by the following system of equations:

$$\begin{cases} h\Big(\underline{\omega} f(R_b^*(\underline{\omega})\Big) - \bar{e} = 0 \\ \bar{\omega} f(R_a^*(\bar{\omega})) - P_R R_a^*(\bar{\omega}) - \tau h\Big(\bar{\omega} f(R_a^*(\bar{\omega}))\Big) - C_f - h^{-1}(\bar{e}) + P_R R_c^*(\bar{\omega}) = 0 \\ \int_0^{\underline{\omega}} R_b^*(\omega) g(\omega) d\omega + \int_{\underline{\omega}}^{\bar{\omega}} R_c^*(\omega) g(\omega) d\omega + \int_{\bar{\omega}}^{\infty} R_a^*(\omega) g(\omega) d\omega - 1 = 0 \end{cases} \tag{22}$$

where the subscript b and a refer respectively to 'below' and 'above' the exemption threshold. c refers to distorted emitters 'constrained' at the threshold. In vectorial notation, let denote $\mathbf{x} = (\underline{\omega}, \bar{\omega}, w)$ and the equations system as $J(\mathbf{x}\,;\,\bar{e}, \tau) = 0$. For a given couple of instruments $I_0 = (\bar{e}_0, \tau_0)$, and as long as the associative matrix of J is not singular ($|J((\mathbf{x}\,;\,\bar{e}, \tau)| \neq 0$), the implicit function theorem implies that there exist a function $\Lambda : I \to \mathbb{R}$ in the neighborhood of $I_0 \subset I = \{(\bar{e}, \tau)\}$ such that: $\mathbf{x} = \Lambda(\bar{e}, \tau)$ for every couple of instrument $(\bar{e}, \tau) \in I$, the regulation instruments set.

Using equations (22) and the envelop theorem, the derivatives matrix D of equilibrium equations J with respect to $\mathbf{x}$ is :

$$D_x J(x;\,\bar{e}, \tau) = \begin{pmatrix} C & 0 & B \\ 0 & A & -R_a^*(\bar{\omega}) + R_c^*(\bar{\omega}) \\ \underbrace{R_b^*(\underline{\omega}) - R_c^*(\underline{\omega})}_{=0} & \Big(R_c^*(\bar{\omega}) - R_a^*(\bar{\omega})\Big) g(\bar{\omega}) & D \end{pmatrix} \tag{23}$$

where $A = (1 - \tau h_q^{'}) f\big(R_a^*(\bar{\omega})\big) - P_R \dfrac{\partial R_c(\bar{\omega})}{\partial \bar{\omega}} = \dfrac{P_R}{\bar{\omega}} \left[\dfrac{f\big(R_a^*(\bar{\omega})\big)}{f'\big(R_a^*(\bar{\omega})\big)} - \dfrac{f\big(R_c^*(\bar{\omega})\big)}{f'\big(R_c^*(\bar{\omega})\big)} \right] > 0$, $B = h_q^{'} \underline{\omega} f' \dfrac{\partial R_b(\underline{\omega})}{\partial P_R} = h_q^{'} \dfrac{f'(R_b^*(\underline{\omega}))}{f''(R_b^*(\underline{\omega}))} < 0$ and $C = h_q^{'} \left(f(R_b^*(\underline{\omega})) - \dfrac{P_R f'(R_b(\omega))}{\underline{\omega} f'' R_c^*(\underline{\omega})} \right) > 0$. The second equality in A is derived firm the FOC and the inequality from the increasing of the function $u(.) = \dfrac{f(.)}{f'(.)}$.

$$D = \int_0^{\underline{\omega}} \frac{\partial R_b^*(\omega)}{\partial P_R} g(\omega) d\omega + \int_{\underline{\omega}}^{\bar{\omega}} \underbrace{\frac{\partial R_c^*(\omega)}{\partial P_R}}_{=0 \text{ as } R_c^* \perp\!\!\!\perp P_R} g(\omega) d\omega + \int_{\bar{\omega}}^{\infty} \frac{\partial R_a^*(\omega)}{\partial P_R} g(\omega) d\omega$$

$$= \int_0^{\underline{\omega}} \underbrace{\left(\frac{1}{\omega f''(R_b^*)}\right)}_{<0} g(\omega) d\omega + \int_{\bar{\omega}}^{\infty} \underbrace{\left(\frac{1}{\omega(1-\tau h_q^{'}) f''(R_a^*)}\right)}_{<0} g(\omega) d\omega < 0$$

and $|D_{\mathbf{x}} J(\mathbf{x};\, \bar{e}, \tau)| = C\Big[DA - \big(R_c^*(\bar{\omega}) - R_a^*(\bar{\omega})\big)^2 g(\bar{\omega})\Big] < 0$ showing the non-singularity of the jacobian $D_x J(.)$, thus, the existence of a solution.

Using the implicit function theorem, we compute $\dfrac{d\Lambda}{d([\bar{e}, \tau])} = -\Big(D_x J(x;\, \bar{e}, \tau)\Big)^{-1} D_{(\bar{e},\tau)} J(\mathbf{x}, \bar{e}, \tau))$. where

$$D_{\bar{e}} J(\mathbf{x}, \bar{e}, \tau)) = \begin{pmatrix} -1 \\ \dfrac{-1}{h_q^{'}(q_c(\omega))} + P_R \dfrac{\partial R_c^*(\bar{\omega})}{\partial \bar{e}} \\ \displaystyle\int_{\underline{\omega}}^{\bar{\omega}} \frac{1}{h_q^{'}(q_c(\omega)) \omega f'(R_c^*(\omega))} g(\omega) d\omega \end{pmatrix} = \begin{pmatrix} -1 \\ \dfrac{1}{h_q^{'}(q_c(\bar{\omega}))} \left[\dfrac{P_R}{\bar{\omega} f'(R_c^*(\bar{\omega}))} - 1\right] < 0 \\ \displaystyle\int_{\underline{\omega}}^{\bar{\omega}} \frac{1}{h_q^{'}(q_c(\omega)) \omega f'(R_c^*(\omega))} g(\omega) d\omega > 0 \end{pmatrix} \tag{24}$$

The sign of the second element of the vector $D_{\bar{e}} J(\mathbf{x}, \bar{e}, \tau))$ comes from the fact that $R_a(\bar{\omega}) - R_c(\bar{\omega}) > 0$ which imply $1 > 1 - \tau h_q^{'} > \dfrac{P_R}{\bar{\omega} f'(R_c^*(\bar{\omega}))}$ because $f'(.) \downarrow$

Therefore,

$$\frac{d\underline{\omega}}{d\bar{e}} = \underbrace{\frac{1}{-|D_{\mathbf{x}} J(\mathbf{x};\, \bar{e}, \tau)|}}_{>0} \Bigg[\underbrace{-\Big(AD - (R_c(\bar{\omega}) - R_a(\bar{\omega}))^2 g(\omega)\Big)}_{>0} + \underbrace{\frac{1}{h_q^{'}(q_c)} \left(\frac{P_R}{\underline{\omega} f'(R_c^*(\bar{\omega}))} - 1\right) B(R_c \bar{\omega} - R_a(\bar{\omega})) g(\omega))}_{<0}$$
$$\underbrace{-AB\left(\int_{\underline{\omega}}^{\bar{\omega}} \frac{1}{h_q^{'}(q_c) \omega f'(R_c^*(\omega))} dG(\omega)\right)}_{>0} \Bigg]$$
$$> 0.$$

$$\frac{d\bar{\omega}}{d\bar{e}} = \underbrace{\frac{C}{-|D_{\mathbf{x}} J(\mathbf{x};\, \bar{e}, \tau)|}}_{>0} \Bigg[\underbrace{D \frac{1}{h_q^{'}(q_c)} \left(\frac{P_R}{\bar{\omega} f'(R_c^*(\bar{\omega}))} - 1\right)}_{>0} - \underbrace{(R_c(\bar{\omega}) - R_a(\bar{\omega}))) \int_{\underline{\omega}}^{\bar{\omega}} \frac{1}{h_q^{'}(q_c) \omega f'(R_c^*(\omega))} dG(\omega)\Big)}_{<0} \Bigg]$$
$$> 0$$

$$
\begin{aligned}
\frac{dP_R}{d\bar{e}} &= \underbrace{\frac{C}{-|D_{\mathbf{x}}J(\mathbf{x};\,\bar{e},\tau)|}}_{>0}\Bigg[-\underbrace{(R_c(\bar{\omega})-R_a(\bar{\omega})))g(\bar{\omega})\frac{1}{h_q'(q_c)}\left(\frac{P_R}{\bar{\omega}f'(R_c^*(\bar{\omega}))}-1\right)}_{<0}+\underbrace{A\int_{\underline{\omega}}^{\bar{\omega}}\frac{1}{h_q'(q_c)\omega f'(R_c^*(\omega))}dG(\omega)\Big)}_{>0}\Bigg]\\
&> 0
\end{aligned}
$$

$$
D_\tau J(x;\,\bar{e},\tau) = \begin{pmatrix} 0 \\ -h(\bar{\omega}f\big(R_a^*(\bar{\omega})\big)) + \underbrace{\left((1-\tau h_q^{'}(q_a(\bar{\omega}))\bar{\omega}f^{'} - P_R\right)}_{=0}\dfrac{\partial R_a^*(\bar{\omega})}{\partial\tau} \equiv E < 0 \\ \displaystyle\int_{\bar{\omega}}^{\infty}\frac{\partial R_a^*(\omega)}{\partial\tau}dG(\omega) < 0 \end{pmatrix}
$$

$$
\text{Therefore}\begin{pmatrix} \dfrac{d\underline{\omega}}{d\tau} \\ \dfrac{d\bar{\omega}}{d\tau} \\ \dfrac{dP_R}{d\tau} \end{pmatrix} = \begin{pmatrix} \dfrac{BE\Big(R_c^*(\bar{\omega})-R_a^*(\bar{\omega})\Big)g(\bar{\omega}) - AB\displaystyle\int_{\bar{\omega}}^{\infty}\frac{\partial R_a^*(\omega)}{\partial\tau}dG(\omega)}{-|D_{\mathbf{x}}J(\mathbf{x};\,\bar{e},\tau)|} < 0 \\ \dfrac{CDE - C\Big(R_c^*(\bar{\omega})-R_a^*(\bar{\omega})\Big)\displaystyle\int_{\bar{\omega}}^{\infty}\frac{\partial R_a^*(\omega)}{\partial\tau}dG(\omega)}{-|D_{\mathbf{x}}J(\mathbf{x};\,\bar{e},\tau)|} \lesseqgtr 0 \\ \dfrac{-CE\Big(R_c^*(\bar{\omega})-R_a^*(\bar{\omega})\Big)g(\bar{\omega}) + AC\displaystyle\int_{\bar{\omega}}^{\infty}\frac{\partial R_a^*(\omega)}{\partial\tau}dG(\omega)}{-|D_{\mathbf{x}}J(\mathbf{x};\,\bar{e},\tau)|} < 0 \end{pmatrix}
$$

Alternatives policies

Proof Proposition. 1. Under the alternative scenarios, except for the partial-exemption, $\underline{\omega} = \bar{\omega} = 0$ and the first two equations of the equilibrium disappear and the last one is reduced to the following: $\int_0^\infty (\frac{P_R}{\omega\theta})^{1/(\theta-1)} dG(\omega) = 1$ and to $\int_0^\infty \left(\frac{P_R}{(1-\tau\sigma(\omega))\omega\theta}\right)^{1/(\theta-1)} dG(\omega) = 1$ respectively for the full-exemption and no-exemption. Thus, we derived the expression of P_R in each case and the first inequality of the proposition is trivial:
$P^*_{R;fe} = \left(\int_0^\infty (\omega\theta)^{1/(1-\theta)} dG(\omega)\right)^{1-\theta}$ and $P^*_{R;ne} = \left(\int_0^\infty \left(\omega\theta(1-\tau\sigma(\omega))\right)^{1/(1-\theta)} dG(\omega)\right)^{1-\theta}$

The expressions of the aggregate productivity Φ are derived using the expression of P_R.

When $\bar{e} > 0$ and by substituting first equation of the system (22) into the second equation of the system and into the resources constraint (third equation), we have respectively

$$(1-\theta)(1-\tau\sigma(\omega))^{1/(1-\theta)}\left(\frac{\bar{\omega}}{\underline{\omega}}\right)^{1/(\theta(1-\theta))} - \left(\frac{\bar{\omega}}{\underline{\omega}}\right)^{1/\theta} + \theta - C_f = 0 \tag{25}$$

and

$$\begin{aligned}
P_{R;be}^{1/(1-\theta)} &= \int_0^{\underline{\omega}} \left(\omega\theta\right)^{1/(1-\theta)} dG(\omega) + (\underline{\omega}\theta)^{1/\theta(1-\theta)} \int_{\underline{\omega}}^{\bar{\omega}} (\omega\theta)^{-1/\theta} dG(\omega) \\
&\quad + \int_{\underline{\omega}}^{\infty} \left((1-\tau\sigma(\omega))\omega\theta\right)^{1/(1-\theta)} dG(\omega) \\
&\leqslant \int_0^{\underline{\omega}} \left(\omega\theta\right)^{1/(1-\theta)} dG(\omega) + \int_{\underline{\omega}}^{\bar{\omega}} (\underline{\omega}\theta)^{1/(1-\theta)} dG(\omega) + \int_{\underline{\omega}}^{\infty} \left((1-\tau\sigma(\omega))\omega\theta\right)^{1/(1-\theta)} dG(\omega) \\
&\leqslant \int_0^{\underline{\omega}} \left(\omega\theta\right)^{1/(1-\theta)} dG(\omega) + \int_{\underline{\omega}}^{\bar{\omega}} (\omega\theta)^{1/(1-\theta)} dG(\omega) + \int_{\underline{\omega}}^{\infty} \left((1-\tau\sigma(\omega))\omega\theta\right)^{1/(1-\theta)} dG(\omega) \\
&\leqslant P_{R;fe}^{1/(1-\theta)}
\end{aligned}$$

The first and the second inequality from the latter derive from the fact that the second term is bounded as $\omega \in (\underline{\omega}, \bar{\omega})$. The last inequality stands from $\sigma(\omega) \leqslant 1/\tau$. Therefore, $P^*_{R;be} \leqslant P^*_{R;fe}$.

Equation (25) is a function in $\frac{\bar{\omega}}{\underline{\omega}} \geqslant 1$ but independent of P_R which decrease for $\frac{\bar{\omega}}{\underline{\omega}} \leqslant \frac{1}{1-\tau\sigma(\omega)}$ in which case the function is negative but increase for $\frac{\bar{\omega}}{\underline{\omega}}$ above and the function is positive. Therefore there is a unique root $\dfrac{\bar{\omega}^*}{\underline{\omega}^*}$ solution of the equation.

From $P_{R;ne}^{1/(1-\theta)} = \displaystyle\int_0^\infty \Big(\omega\theta(1-\tau\sigma(\omega))\Big)^{1/(1-\theta)} dG(\omega)$, we have,

$$
\begin{aligned}
P_{R;be}^{1/(1-\theta)} - P_{R;ne}^{1/(1-\theta)} = &\int_0^{\underline{\omega}} \Bigg(\underbrace{\big(\omega\theta\big)^{1/(1-\theta)} - \big(\omega\theta(1-\tau\sigma(\omega))\big)^{1/(1-\theta)}}_{\geqslant 0} \Bigg) dG(\omega) \\
&+ \int_{\underline{\omega}}^{\bar{\omega}} (\omega\theta)^{-1/\theta} \Bigg((\underline{\omega}\theta)^{1/\big(\theta(1-\theta)\big)} - (1-\tau\sigma(\omega))^{1/(1-\theta)} \big(\omega\theta\big)^{1/\big(\theta(1-\theta)\big)} \Bigg) dG(\omega)
\end{aligned}
$$

where the sign of the second term depends on the sign of $k(\omega) = (\underline{\omega}\theta)^{1/\big(\theta(1-\theta)\big)} - (1-\tau\sigma(\omega))^{1/(1-\theta)} \big(\omega\theta\big)^{1/\big(\theta(1-\theta)\big)}$ which is a continuous and decreasing function between $(\underline{\omega}, \bar{\omega})$ such that

$k(\underline{\omega}) = (\underline{\omega}\theta)^{1/\big(\theta(1-\theta)\big)} \Big(1 - (1-\tau\sigma(\omega))^{1/(1-\theta)} \Big) > 0$ and

$k(\bar{\omega}) = (\underline{\omega}\theta)^{1/\big(\theta(1-\theta)\big)} \left(1 - (1-\tau\sigma(\omega))^{1/(1-\theta)} \left(\dfrac{\bar{\omega}}{\underline{\omega}}\right)^{1/\big(\theta(1-\theta)\big)} \right) \gtreqless 0 \iff 1 \leqslant \dfrac{\bar{\omega}}{\underline{\omega}} \lesseqgtr (1-\tau\sigma(\omega))^{-\theta}$

Using the equation (25) solution requirement, $k(\bar{\omega}) \leqslant 1$ and the sign of $P_{R;be} - P_{R;ne}$ is ambiguous. However, $P_{R;be} - P_{R;ne} \geqslant 0$ if $\dfrac{\bar{\omega}}{\underline{\omega}} = (1-\tau\sigma(\omega))^{-\theta}$ i.e 'be'='pe'. In that case, the third term cancel out.

2. Proof of production inequality: The aggregate productivity under the baseline after inserting the first equation of the equilibrium system is

$$
\begin{aligned}
Q_{be} = &\left(\frac{P_R}{\theta}\right)^{\theta/(\theta-1)} \Bigg[\int_0^{\underline{\omega}} \omega^{1/(1-\theta)} dG(\omega) + \int_{\underline{\omega}}^{\bar{\omega}} \underline{\omega}^{1/(1-\theta)} dG(\omega) \\
&+ \int_{\bar{\omega}}^{\infty} \big(1-\tau\sigma(\omega)\big)^{\theta/(1-\theta)} \omega^{1/(1-\theta)} dG(\omega) \Bigg] \\
\leqslant &\left(\frac{P_R}{\theta}\right)^{\theta/(\theta-1)} \Bigg[\int_0^{\underline{\omega}} \omega^{1/(1-\theta)} dG(\omega) + \int_{\underline{\omega}}^{\bar{\omega}} \omega^{1/(1-\theta)} dG(\omega) + \int_{\bar{\omega}}^{\infty} \omega^{1/(1-\theta)} dG(\omega) \Bigg] \\
= &\left(\frac{P_R}{\theta}\right)^{\theta/(\theta-1)} \Bigg[\int_0^{\infty} \omega^{1/(1-\theta)} dG(\omega) \Bigg]
\end{aligned}
$$

where from the last equation,

$$\begin{aligned}\left(\frac{P_R}{\theta}\right)^{1/(1-\theta)} &= \int_0^{\underline{\omega}} \omega^{1/(1-\theta)} dG(\omega) + (\underline{\omega}\theta)^{1/\left(\theta(1-\theta)\right)} \int_{\underline{\omega}}^{\bar{\omega}} \omega^{-1/\theta} dG(\omega) \\ &\quad + \int_{\bar{\omega}}^{\infty} \left(1-\tau\sigma(\omega)\right)^{\theta/(1-\theta)} \omega^{1/(1-\theta)} dG(\omega) \\ &\leqslant \int_0^{\underline{\omega}} \omega^{1/(1-\theta)} dG(\omega) + \int_{\underline{\omega}}^{\bar{\omega}} \omega^{1/(1-\theta)} dG(\omega) + \int_{\bar{\omega}}^{\infty} \omega^{1/(1-\theta)} dG(\omega) \\ &= \int_0^{\infty} \omega^{1/(1-\theta)} dG(\omega)\end{aligned}$$

. Therefore, $Q^*_{be} \leqslant \left(\frac{P_R}{\theta}\right)^{\theta/(\theta-1)} \left[\int_0^{\infty} \omega^{1/(1-\theta)} dG(\omega)\right]^{1-\theta} = Q^*_{ne}$

$\square$

Optimal tax instrument

$$
\begin{aligned}
\max_{\tau} W(\tau,\bar{e}) &= \max_{\tau}\Big\{Tr+\Pi-\lambda\bar{E}\Big\}\\
&= \max_{\tau}\Big\{\tau\int_{\bar{\omega}}^{\infty} e(\tau,\bar{e})+\int_{0}^{\infty}\pi(\tau,\bar{e})-\lambda\int_{0}^{\infty}\sigma(\omega)q(\tau,\bar{e})dG(\omega)\\
&= \max_{\tau}\Big\{\tau\int_{\bar{\omega}(\tau)}^{\infty}\sigma\omega\Big(\frac{\omega\theta(1-\tau\sigma)}{P_R}\Big)^{\theta/(1-\theta)}\\
&\quad+\Big[\int_{0}^{\underline{\omega}}\omega(1-\theta)\Big(\frac{\omega\theta}{P_R}\Big)^{\theta/(1-\theta)}dG(\omega)+\int_{\underline{\omega}}^{\bar{\omega}(\tau)}\Big(\frac{\bar{e}}{\sigma}-P_R\Big(\frac{\bar{e}}{\sigma\omega}\Big)^{1/\theta}\Big)dG(\omega)\\
&\quad+\int_{\bar{\omega}(\tau)}^{\infty}[\omega(1-\tau\sigma)]^{1/(1-\theta)}(1-\theta)\Big(\frac{\theta}{P_R}\Big)^{\theta/(1-\theta)}dG(\omega)\Big]-\lambda\Big[\int_{0}^{\underline{\omega}}\sigma\omega\Big(\frac{\omega\theta}{P_R}\Big)^{\theta/(1-\theta)}dG(\omega)\\
&\quad+\int_{\underline{\omega}}^{\bar{\omega}(\tau)}\bar{e}dG(\omega)+\int_{\bar{\omega}(\tau)}^{\infty}\omega\Big(\frac{\omega\theta(1-\tau\sigma)}{P_R}\Big)^{\theta/(1-\theta)}dG(\omega)\Big]\Big\}\\
&= \max_{\tau}\Big\{\int_{\bar{\omega}(\tau)}^{\infty}\omega\Big(\sigma\tau-\sigma\lambda+(1-\tau\sigma)(1-\theta)\Big)\Big(\frac{\omega\theta(1-\tau\sigma)}{P_R}\Big)^{\theta/(1-\theta)}dG(\omega)\\
&\quad+\int_{0}^{\underline{\omega}}\big(1-\theta-\lambda\sigma\big)\omega^{1/(1-\theta)}\Big(\frac{\theta}{P_R}\Big)^{\theta/(1-\theta)}dG(\omega)+\int_{\underline{\omega}}^{\bar{\omega}(\tau)}\big(1-\lambda\sigma\big)\frac{\bar{e}}{\sigma}dG(\omega)\\
&\quad-\int_{\underline{\omega}}^{\bar{\omega}(\tau)}P_R\Big(\frac{\bar{e}}{\sigma\omega}\Big)^{1/\theta}dG(\omega)\Big\}
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial W(\tau,\bar{e})}{\partial\tau}=0 \iff &\int_{\bar{\omega}(\tau)}^{\infty}\sigma\theta\omega\Big(\frac{\omega\theta(1-\tau\sigma)}{P_R}\Big)^{\theta/(1-\theta)}dG(\omega)+\Big[-\frac{\theta}{1-\theta}\int_{\bar{\omega}(\tau)}^{\infty}(\alpha+\tau\sigma\theta)\frac{\sigma\theta\omega}{P_R}\omega\Big(\frac{\omega\theta(1-\tau\sigma)}{P_R}\Big)^{\theta/(1-\theta)-1}\\
&-\underbrace{\bar{\omega}\Big(\frac{\bar{\omega}\theta(1-\tau\sigma)}{P_R}\Big)^{\theta/(1-\theta)}}_{e_a(\bar{\omega})/\sigma}g(\bar{\omega})\frac{\partial\bar{\omega}}{\partial\tau}\Big]+\underbrace{\Big(\frac{\bar{e}}{\sigma}-P_R\big(\frac{\bar{e}}{\sigma\bar{\omega}}\big)^{1/\theta}\Big)}_{=\pi_c(\bar{\omega})=\pi_a(\bar{\omega})=(1-\theta)(1-\tau\sigma)e_a(\bar{\omega})/\sigma}g(\bar{\omega})\frac{\partial\bar{\omega}}{\partial\tau}-\lambda\bar{e}g(\bar{\omega})\frac{\partial\bar{\omega}}{\partial\tau}
\end{aligned}
$$

where $\alpha=1-\theta-\sigma\lambda$. By denoting $E_a=\int_{\bar{\omega}(\tau)}^{\infty}\sigma\omega\Big(\frac{\omega\theta(1-\tau\sigma)}{P_R}\Big)^{\theta/(1-\theta)}dG(\omega)$ the total emission of productive firms that pays for the regulatory cost, $e_a(\bar{\omega})$ the emission of firm $\bar{\omega}$, and $\mathcal{E}_\tau=d\,log(1-G(\bar{\omega}))/d\,log\tau=-g(\bar{\omega})\frac{\partial\bar{\omega}}{\partial\tau}\frac{\tau}{1-G(\bar{\omega})}$ the elasticity of large emitters above the threshold with respect to carbon tax, we derive from the previous equation result of the proposition.

Quantitative implication of threshold-based exemption

This section shows the aggregate variables under the different policy scenario prior the re-calibration exercise to reach the same emission level.

Table A3. Policies simulation and comparison

	Baseline exemption	No-exemption
Emission cost	$\tau e \mathbb{1}_{e>\bar{e}}$	τe
Emission	100	99.77
Output	100	100.05
Capital	100	100.05
TFP	100	100.02
Mass of entrant	100	99.99
Wage rate	100	100.05
Firms below $\bar{e}$ (%)	3.72	–
Firms constrained at $\bar{e}$ (%)	6.54	–
Firms above $\bar{e}$ (%)	89.74	–
Share of firms with $\geqslant 50$	4.53	4.56
Emission share in top quintile	44.02	44.08
Share of firms abating (%)	93.33	93.33
Average abatement share	1.88	1.88

Eliminating the exemption threshold results in a more efficient allocation of resources, yielding an aggregate Total Factor Productivity (TFP) gain of 0.5%. The proportion of firms in the top size category increases when all firms are treated equally under a no-exemption policy. Interestingly, transitioning from the baseline model with an exemption to a policy without exemption leads to a decrease in aggregate emissions but an increase in output by 0.23% and 0.05%, respectively. Conversely, transitioning to a Partial-exemption policy is associated with an increase in both variables.

The number of new entrant firms decreases, allowing for the clearing of the labor market, where firms in the top size category demand more employment. The cost of labor rises to offset the decline in firm values, a result of additional costs incurred by previously exempted firms when transitioning from a policy with exemption to a policy without exemption.