# DYNAMIC RESPONSES TO CARBON PRICING IN THE ELECTRICITY SECTOR

Paige Weber*

December 9, 2021

**Abstract**

This paper studies the impact of carbon prices on local air pollution, highlighting two key mechanisms in which a carbon price changes the spatial distribution of air pollution. I develop a model of firm production and efficiency investment decisions to capture these mechanisms and quantify their impacts. I show that the observed carbon price in California led to minimal changes in the spatial distribution of local air pollution emitted by the state's electricity sector. A higher carbon price changes the spatial distribution of local air pollution providing co-benefits from the climate policy, which increase in pre-policy pollution burden.

**Keywords**: Regulation, electricity sector, climate change, air pollution, distributional effects

**JEL classification codes**: L51, L94, Q54, Q53

**Acknowledgements:** This project is supported by an Alfred P. Sloan Foundation Pre-doctoral Fellowship on Energy Economics, awarded through the National Bureau of Economic Research (NBER). I am grateful for discussions with Ken Gillingham, Matt Kotchen, Steven Berry, Ying Fan, Kyle Meng, Gregory Upton, Meera Mahadevan, Danae Hernandez-Cortez, Cailin Slattery, Cuicui Chen, Maura Coughlin, and Sonia Gilbukh about this work, as well as numerous comments from seminar and conference participants at Yale, Columbia, Fordham, East Carolina University, the Center for Economic Studies IFO, the Berkeley Climate Economics Workshop, the Berkeley-Harvard-Yale Virtual Seminar on the Economics of Climate Change and Energy Transition, the Southern Economic Association Annual Meeting, the IAEE session at the ASSA/AEA Annual Meeting, the Triangle Applied Microeconomics Conference, the Wharton Energy Economics and Finance Workshop, and the Virtual Seminar on Climate Economics hosted by the Federal Reserve Bank of San Francisco.

*Weber: Department of Economics, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, email: paigeweber@unc.edu, phone: (310) 897-7983.

# 1 Introduction

When firms create local externalities, the welfare effects of a policy regulating these externalities depend on how the policy changes the geography of production. For instance, carbon prices regulate firms that produce greenhouse gases, many of which also produce co-pollutants that contribute to local air pollution. If firm's responses to carbon prices change production geographies, some regions can experience increases or decreases in local air pollution, even while aggregate greenhouse gas emissions and local air pollution is reduced. Given the disproportionate shares of local air pollution borne by low-income and minority communities (Banzhaf, Ma and Timmins 2019; Colmer et al. 2020), understanding the interaction of carbon prices and local air pollution is critical to assessing the equity impacts of climate change policy (Hernandez-Cortez and Meng 2019; Ed. 2019). In this paper, I study how California's carbon price, implemented via an emissions trading program for carbon emissions, impacts the spatial distribution of local air pollution from firms in California's wholesale electricity sector.

A regulation can change the spatial distribution of firms' production by altering their costs heterogeneously, changing their relative competitiveness. Firms that are relatively more competitive following the regulation would be expected to increase their market share compared to a no-regulation environment. Accordingly, other local externalities of production such as air pollution will also increase from these firms. If firms that increase production in the regulated environment are located in regions that are already heavily polluted, the policy would exacerbate pollution in these regions. On the other hand, if these firms are located in relatively less polluted regions, the regulation could ameliorate disproportionate pollution burdens. Thus, the equity outcomes depend on the spatial distribution of the plants and their cost structures, which determine their responses to the policy.

When emissions trading programs, also known as cap-and-trade programs, cap emissions at efficient levels, they can internalize the global warming externalities of greenhouse gas emissions by requiring that firms pay for their emissions. The price for emissions, the carbon price,[1] is determined in equilibrium based on the behavior of all regulated firms and the total number of tradeable permits created by the regulator. With competitive permit markets, the regulation augments the firm's profit maximization problem by adding a cost, the carbon price, to each unit of emissions. For the firms that I study, carbon pricing impacts the near- and medium- term air pollution distributions through two key mechanisms. First, the carbon price could change the order of units along aggregate supply curves in hourly

[1]Carbon prices are generally defined as a $ per ton of carbon dioxide emissions equivalent. This allows for a single carbon price for any greenhouse gas covered by the regulation, where each greenhouse gas creates a compliance obligation based on its carbon dioxide equivalent.

markets for wholesale electricity by heterogeneously changing firms' costs.[2] In this case, the share of electricity provided by each unit for a given level of demand would change following the carbon price, as well as the local pollution generated by that firm.[3] Second, the carbon price could lead to changes in firms' emissions intensities of both greenhouse gases and local air pollutants from investments in efficiency. The carbon price enters firms' profit maximization decisions analogously to an increase in the cost of inputs, and thus, investments in efficiency should weakly increase following the introduction of the carbon price. Efficiency improvements lead to burning less fuel per unit of output thereby reducing emissions intensities.[4] In summary, in the near- and medium-terms the spatial distribution of local pollution under a carbon price can change from supply curve re-orderings and/or from changes in unit efficiencies.

To capture these mechanisms, I develop a model of electricity generating unit's dynamic production decisions, as well as efficiency investment decisions, and use the estimates from the model to simulate market outcomes across alternative policy scenarios. Firm's make hourly productions, which are dynamic because of the presence of start-up costs incurred to turn units on when not previously operating. Firms also make a one time decision of whether to invest to improve their efficiency or not. The model is designed to capture near- and medium- term outcomes and does not consider firm entry or exit; the investment decision here is whether to improve firm efficiency conditional on having built an electric generating unit.[5] The model is similar to the one developed in Cullen (2015), though I include an additional firm decision, investment in efficiency, which potentially makes it more similar that in Cullen and Reynolds (2017). Unique from Cullen and Reynolds (2017), I include additional dimensions of firm heterogeneity, endogenous heat rates, and a fixed fleet. To estimate the model, I leverage the correspondence between the competitive equilibrium and a cost minimization problem as demonstrated in Cullen and Reynolds (2017).[6]

After estimating the model, I first compare California's carbon price to a no-carbon price world. Next, as many environmental advocates have been concerned about the low carbon prices observed in the state,[7] I simulate outcomes under higher carbon prices that would

[2]The order of units on an aggregate supply curve is called the merit order in electricity markets.

[3]The term market share here refers to the share of residual demand in a given hourly market to be provided by units regulated by the carbon price. Residual demand denotes hourly demand less generation from renewable energy resources and any other preferred generating resources, given priority for cost or regulatory reasons.

[4]Emissions intensities are the amount of emissions released per unit of electricity.

[5]That said, the repeated production decisions in this model can be likened to a repeated entry and exit model: in each period, firms decide whether to "enter" an hourly production market, or not, with fixed costs associated with entry.

[6]Specifically, I use a limiting case of this result when the price elasticity of wholesale electricity demand approaches zero. The necessary assumptions to apply their result in this setting are discussed in Section 2.4.

[7]Carbon prices from 2013-2015 are shown in Figure 1, along with the price floor. It it worth noting that

arise from a more stringent emissions trading program. Finally, I consider that the carbon price is implemented together with a location-specific Pigouvian tax on local air pollution. The results of the model indicate that the cost structure of California's fossil-fuel electricity portfolio leads to minimal reallocation of production and emissions at current carbon prices. Thus, at current carbon prices, the policy has a limited effect on the spatial distribution of air pollution emitted by the electricity sector. While this result quells concerns about the negative unintended consequences of the carbon price's impact on the distribution of local air pollutants, the analysis also reveals that the current carbon price does not meaningfully reduce local air pollution from the state's disproportionately polluted communities. In counterfactual simulations of higher carbon prices, I find larger changes in the spatial distribution of production as a result of the carbon price. These production changes lead to aggregate co-benefits from the climate change policy's impact on local air pollution compared to a no carbon price policy. Further, I find that on average these benefits are larger among counties with higher pre-existing pollution. I compare this to a scenario with a carbon policy at current prices combined with a location-specific Pigouvian tax on local air pollution, which also provide benefits from avoided local air pollution damages in a magnitude similar to the high carbon price scenarios.

A key contribution of this study's analysis of the environmental justice implications of firm's responses to carbon prices is its incorporation of dynamics in the production decision. This is not the first study to include dynamics in electricity firm's production decisions, which is done in Cullen (2015), Cullen (2013), Reguant (2014), Jha and Leslie (2021), and discussed in Mansur (2008).[8] To my knowledge, this paper is the first to include this start-up cost channel to study the equity impacts of environmental policy. Firms in this setting, electricity generating units, incur start-up costs to turn on when previously not producing. These start-up costs render the production decision dynamic, as they require that firms make expectations over future prices to avoid repeatedly turning on and off. Modeling production as a static decision shuts down a key mechanism in which the carbon price could alter the spatial distribution of production. As discussed further in Section 2.5, when firms share the same emissions intensity of fuel, with static decision-marking, the impact of the carbon price can be modeled as an upward shift in the supply curve, together with an increase in the slope of supply curve. In this static analysis, the carbon price preserves the ordering of electricity generating units along the supply curve, and does not lead to re-allocation of production or local externalities across firms. This point is instructive for predicting

during the time period of this study, prices did not reach the price floor, which did happen in 2016.

[8]Reguant (2014) observes firm bidding data, which allows for an alternative approach to estimating start-up costs compared to the approach in **?**, which is closer to what is done here.

anticipated production re-allocation and leakage in other sectors: when production decisions are made largely based on marginal costs from inputs, and a regulation increases the costs of inputs, little production re-allocation should be expected. The presence of costs that are less affected by the regulation creates the opportunity for production re-allocation by re-ordering firms on aggregate supply curves.

This paper contributes to an emerging literature empirically studying the effect of climate change policy on the distribution of air pollution. I build on other work studying the equity impacts of market-based environmental regulation (Fowlie, Holland and Mansur 2012; Hernandez-Cortez and Meng 2019; Grainger and Ruangmas 2018; Walch 2018; Shapiro and Walker 2021) by explicitly modeling the mechanisms in which the regulation could alter these outcomes.[9] While pricing carbon is often the preferred approach to addressing global climate change from an economic efficiency perspective, its impact on the distribution of co-pollutants that create local air pollution is not well understood. Critics of carbon pricing contend that the policy's local impacts lead to worse air quality outcomes for regions that are already heavily polluted. Some research supports this, for example, Cushing et al. (2018) find descriptive evidence that pollution increased in disproportionately polluted communities following the cap-and-trade program. On the other hand, Walch (2018) and Hernandez-Cortez and Meng (2019) find the opposite effect, with the later finding evidence of a reversal of some spatial pollution trends leading to a reduction in pollution among heavily polluted communities. These two papers use reduced form approaches that focus on the policy's effect across all regulated industries. In this paper I develop a model that allows me to elucidate the mechanisms in which the policy could impact equity outcomes from air pollution, as well as to simulate counterfactual equity outcomes under alternative policy scenarios. The model builds on other dynamic competitive equilibrium frameworks such as Jovanovic (1982) and Hopenhayn (1992) and is most similar to Cullen and Reynolds (2017).

Finally, by focusing on the regulation's effect on one regulated industry, the electricity sector, this paper is able to move beyond current literature that asks *whether* market-based regulation led to more local air pollution in disproportionately polluted communities (Walch 2018; Hernandez-Cortez and Meng 2019; Fowlie, Holland and Mansur 2012; Grainger and Ruangmas 2018), to ask *why* the regulation would or would not be expected to have such an effect. This research provides insights in other settings with similar cost structures. In particular, I show that regulations that increase marginal costs would be expected to have little production re-allocation effects among industries where production decisions are

[9] Fowlie, Holland and Mansur (2012); Grainger and Ruangmas (2018) study the equity impacts of market-based policy regulating local air pollution. This paper and Hernandez-Cortez and Meng (2019); Walch (2018) study the cross-effect of market-based mechanisms designed to address climate change on local air pollution.

dominated by marginal costs, due to the preservation of the ordering along the supply curve. Further, I make a theoretical point that when efficiency is decreasing in pollution intensity, under certain conditions on investment costs, investments in efficiency improvements are more likely among the relatively more-efficient and more frequently utilized units, which is further discussed in Weber (2021).

As the fifth largest economy in the world with an environmentally progressive government, California has and continues to serve as an important testing ground for climate change policy. As an early adopter of GHG emissions trading policy,[10] the state was also early to take on the debate around the equity impacts of market-based environmental regulation. The debate has since taken over national climate change policy discussions, playing out in the Biden Administration's selection of the head of the Environmental Protection Agency (Davenport 2020). It is also seen in the current reluctance of policy makers to use market-based mechanisms to meet environmental goals, with a recent preference for platforms that focus on standards and investments (Bushnell 2021). The move away from market-based mechanisms to mitigate greenhouse gases due to the concern that they lead to outcomes that conflict with equity goals is not yet firmly grounded in empirical research. This paper seeks to weigh in on this debate, with the goal of understanding the mechanisms in which market-based climate change policies could change the spatial distribution of local air pollution. In doing so, this paper also contributes to literature regarding the distributional consequences of energy, electricity, and climate change policy (Borenstein 2012, 2013, 2017; Borenstein and Davis 2016; Fullerton, Heutel and Metcalf 2012; Goulder et al. 2018; Knittel and Sandler 2018).

The rest of the paper is organized as follows. Section 2 describes the model of production and efficiency investment decisions, and reviews the theoretical predictions of the model regarding the carbon price's impact on production and investment. Section 3 describes the empirical strategy. Section 4 reviews data sources and the empirical setting. Section 5 presents the results, and Section 6 concludes.

# 2 Dynamic Model of Investment and Production

In this model decisions are made at the electricity-generating unit level, where a unit consists of a heat engine that converts fuel to energy. Units are assumed to act competitively and make decisions as single agents. Accordingly, I treat each unit as acting as an individual

[10] Other countries and regions have also implemented cap-and-trade programs for greenhouse gases, notably the European Union's Emissions Trading Scheme (EU ETS) in 2005, the Regional Greenhouse Gas Initiative in the eastern United States in 2005, and Australia's cap-and-trade program in 2012.

firm and refer to units as firms.[11] Firms make two decisions. First, firms make one-time decisions of whether or not to invest to improve their efficiency. A firm's efficiency is measured by its heat rate, which is the amount of fuel needed to produce one unit of electricity through combustion. Consequently, lower heat rates translate to improved efficiency and lower input costs. This investment decision determines the firm's efficiency in subsequent, repeated hourly operation decisions. I model the investment decision as myopic to market outcomes beyond the the first compliance phase of the carbon policy.[12]

In the production decision, firms make repeated binary operation decisions, deciding whether or not to generate electricity in each hourly market. Operation is more costly if the firm was not operating in the prior period, due to the presence of start-up costs, which are incurred from the additional costs of fuel, auxiliary power, water, additives, chemicals, and wear and tear required to bring a unit online (Kumar et al. 2012).[13] Firms make operation decisions based on their expectations of future demand. Hourly demand is assumed to be inelastic to wholesale electricity prices, and demand shocks are modeled as an AR(1) Markov process, conditional on hour of the day.[14]

## 2.1 Investment Decision

For the investment decision, firms decide whether to invest to improve their efficiency by reducing their current heat rate $\omega'$, measured in Btu per KWh. Firms choose $j$ from a discrete set of activities in set $\mathscr{J}$ of size $J$, which includes $j = 0$ corresponding to no investment. Modeling the investment choices as discrete is supported by the data, which exhibit lumpy investment behavior.[15] Following the investment decision, firm $i$'s heat rate $\omega_i$ is:

[11] I discuss this assumption in section 4.1.

[12] The intuition behind this modeling approach is that at the end of the first compliance period, firm's acquire new information about the regulatory environment for the next period, and make new decisions regarding investments.

[13] Estimates of start-up costs from the National Renewable Energy Laboratory (NREL) show that maintenance costs from the wear and tear of turning generating units on and off composes the large majority of start-up costs (Kumar et al. 2012).

[14] Residential customers pay retail rates, which allows me to treat hourly residential demand as inelastic to the price firms are paid, the wholesale price for electricity. There are some caveats with this approach. One, industrial customers may pay rates closer to wholesale prices. Two, some retail customers may have opted into demand-side management programs that provide incentives to change demand in response to wholesale prices. There is a separate question about changes in demand over time, that is, the question of whether we should expect shifts in the demand curve over time. We could imagine that a growing economy leads to an increase in demand over time; alternatively, we could foresee that an increase in renewable energy reduces the residual quantity demanded by this fossil portfolio. In this paper I do not take a stand on future demand, and consider the response of the fossil portfolio for a given level of demand.

[15] See section A.5 for more information about empirically observed investment. For computational tractability, the estimation collapses the investment decision to a binary choice to invest or not.

$$\omega_i = \omega_i'(1 + \tilde{\delta}) - j_i, \tag{1}$$

where $j_i \leq \omega_i$ and $\tilde{\delta} \in (0, 1)$ is an exogenous depreciation rate that decreases firm efficiency, i.e., increases the firm's heat rate, and corresponds to the time in between the investment decision and the first production decision. Investments to improve efficiency have costs that increase in the size of the efficiency improvement and are denoted $\Gamma$:

$$\Gamma(j_i, v_i, \gamma) = \gamma j_i^{1/\alpha} + v_i, \tag{2}$$

where $\alpha$ governs the rate at which marginal investment costs are increasing in size of desired improvement, and $v_i$ is a private stochastic shock to investment costs, drawn from an extreme value type 1 distribution.

## 2.2 Operating and Production Decision

Firms also make binary operating decisions $a_{it} \in [0, 1]$, conditional on $\omega_i$ chosen in the investment decision. The quantity firm $i$ produces in each hour $q_{it}$ follows from its operating decision, based on the following decision rule:

$$q_{it} = \begin{cases} q_{i,max} & \text{if } P_t \geq mc_i \text{ and } a_{it} = 1 \\ q_{i,min} & \text{if } P_t < mc_i \text{ and } a_{it} = 1 \\ 0 & \text{if } a_{it} = 0. \end{cases} \tag{3}$$

$P_t$ denotes the wholesale equilibrium prices in hour $t$; $mc_i$ denotes the marginal costs of production, defined below; and $q_{i,min}$ and $q_{i,max}$ denote the firm's minimum and maximum operation levels.[16] Modeling production as a discrete choice is supported by the data, which shows that generating units generally operate at one of a discrete set of production quantities, as shown in Figure 17 in Appendix 8.7.[17]

The per period profits for firms are defined:

[16] The simulations set $q_{min} = 0.75 \cdot q_{max}$, where $q_{max}$ is the reported operation capacity of the firm in MW. I compare this approach to estimating $q_{i,min}$ and $q_{i,max}$ from the data using finite mixture models, and find similar estimates when averaging over the firm type categorizations that are used in the empirical analysis.

[17] The number of discrete operating levels varies across generating units; limiting the number of operating levels to two as is done here is a simplifying assumption.

$$\pi_{it} = \begin{cases} q_{it}(P_t - mc_i) & \text{if } a_{i,t-1} = 1 \\ q_{it}(P_t - mc_i) - \kappa_i & \text{if } a_{i,t-1} = 0, \end{cases} \tag{4}$$

where $\kappa_i$ denotes start-up costs. Marginal costs are a function of firm efficiency (heat rate), $\omega_i$; costs of fuel, $f$; emissions intensity of the fuel, $e^f$; and price of GHG permits, $\tau$:[18]

$$mc_i = \omega_i(f + e^f\tau). \tag{5}$$

The above formulation shows that when $f$ and $e^f$ are positive and $\tau$ is non-negative, $mc_i$ is increasing in heat rate, $\frac{\partial mc_i}{\partial \omega_i} > 0$. Further, the formulation demonstrates that investments to reduce heat rate reduce marginal costs by decreasing both fuel costs, $\omega_i \cdot c^f$, and compliance costs, $\omega_i \cdot e^f\tau$. In this sense, the GHG program can be seen as an increase in the cost of inputs to production, analogous to an increase in natural gas prices, as discussed in Mansur (2008).[19] In this model, both fuel costs and carbon prices are known and exogenous, though future research could explore allowing these two input costs to differ in terms of how firms make expectations over future natural gas versus carbon prices. For the purposes of defining the state variables for the model, I define marginal input costs, $c$, as the sum of marginal fuel and GHG prices, $f + e^f\tau$.

## 2.3 The Firm's Problem

This section specifies the state variables and their transitions and formulates the firm's dynamic programming problem. In the production decision, firms observe state variables, $s_{it}$, for the current period demand shock, hour, its lagged operating state, its heat rate, and the input costs, $s_{it} = \{\eta_t, h_t, a_{it-1}, \omega_i, c\}$. Firms make expectations about future period demand shocks and choose $a_{it}$ to maximize the sum of future discounted profits. Firms are assumed to have rational beliefs about future demand, and demand is modeled as an AR(1) process conditional on hour of the day, $h_t$. Hourly demand is assumed to be inelastic to wholesale prices. As such, characterizing firms' expectations of demand shocks is sufficient to characterize their beliefs about wholesale prices, $P_t$. The lagged operating state, $a_{it-1}$, equals 1 when the firm was on in the last period and 0 otherwise. Input costs, $c$, are exogenous, time invariant, and known to firms, and the hour of day evolves as $h_{t+1} = h_t + 1 - \mathbb{1}(h_t = 24) \cdot 24$. Given a price process that is measurable with respect to all possible histories of the demand

[18] The formulation of (5) in units is $\frac{\$}{KWh} = \frac{Btu}{KWH} * \frac{\$}{Btu} + \frac{Btu}{KWH} * \frac{emissions}{Btu} * \frac{\$}{emissions}$.

[19] Mansur (2008) leverages this similarity between incentives from natural gas costs and GHG prices to evaluate carbon abatement costs in electricity markets using changes in natural gas prices.

shocks, the value for firms in time $t$ with state $s_t$:

$$V^{2j}(s_{it}) = \max_{a_{it}\in\{0,1\}} \{\sum_{t=0}^{\infty} \delta^t \, \mathbb{E}[q_{it}(P_t(\eta_t, \omega) - mc(\omega_i)) - \mathbb{1}(a_{it-1} = 0, a_{it} = 1)\kappa_i | \eta_{t-1}]\}, \tag{6}$$

where $a_t$ determines $q_t$ by the decision rule in equation 3 and $\omega$ denotes the efficiencies of all other firms. The second term on the right-hand side reflects start-up costs incurred for every period $t$ that firms operate when not operating in $t-1$.[20] A policy for production decisions is profit maximizing for firms with initial state $s_{it}$ if it satisfies 6.

Next, I formulate the value function for firms in the investment decision. Firms make investment decisions by comparing the costs of investment to the payoffs in production from the reduction in fuel and compliance costs as a result of improved efficiency less the costs of investment:

$$V^1(s_0) = \max_{j\in\mathscr{J}} \{\tilde{\delta} \, \mathbb{E}[V^{2j}(s_0)] - \Gamma(j, v, \gamma, \alpha)\}, \tag{7}$$

where the initial demand shock state in $\mathbf{s_0}$ is the same as the initial demand shock in the production decision, and $\gamma$ and $\alpha$ are investment cost parameters observed by the firm and not by the econometrician. When making expectations about payoffs from production, firms are assumed to have rational expectations about other firms' investment decisions and efficiencies. The optimal policy for production and investment is characterized by the value functions in 6 and 7 above.

## 2.4 Cost Minimization Problem

To evaluate the solution to the firm's two-part dynamic programming problem, I leverage the correspondence between the profit-maximizing choices in a competitive equilibrium and the solution to a production cost minimization problem (henceforth, cost minimization problem). This correspondence is demonstrated to hold in this setting by Cullen and Reynolds (2017) and follows intuition in earlier work in dynamic competitive equilibriums (Lucas and Prescott 1971; Jovanovic 1982; Hopenhayn 1992). Cullen and Reynolds (2017) extend these results to a setting with repeated entry and exit into hourly electricity markets and non-convexities in the aggregate production technology. Establishing that the correspondence holds in hourly wholesale electricity markets requires the additional assumption that firms

[20] Ramp-up and ramp-down costs, which are the costs of the firm to operate at sub-optimal heat rates while ramping up to the preferred operating level, are not explicitly modeled. As a result, these costs are bundled into the start-up costs, since every start-up also requires a ramp-up and ramp-down.

are small relative to the total size of the market, more formally, measure zero. A competitive equilibrium in a model with large firms of positive measure need not exist as start-up costs create non-convexities in costs. The small firms assumption is key to linking the solution to a cost minimization problem with the competitive equilibrium outcome. I solve the cost minimization problem for discrete sized generating units, and interpret the solution as a competitive equilibrium allocation, which is an approximation due to cost non-convexity.

The cost minimization problem is to find the allocations of production and efficiency investments across all firms in the electricity portfolio that minimize the costs of meeting hourly electricity demand. Initially, costs will include carbon prices and exclude local air quality damages as in the empirical setting; subsequently, a counterfactual simulation will modify costs to include local air pollution damages as well. To define the cost minimization problem, I introduce additional notation to collect production quantities and investments across all firms into vectors. Define $\mathbf{j}$ as a vector of investment choices for firms $i \in \{1, ..., \mathrm{U}\}$, where $j_i \in \mathscr{J}\ \forall i$, and denote the set of feasible $\mathbf{j}$ as $\mathcal{J}$. Similarly, let $\mathbf{q}$ be a U-sized vector of production quantities corresponding to each firm, where $q_i \in \{0, q_{i,min}, q_{i,max}\}\ \forall i$, and let Q denote the set of feasible $\mathbf{q}$. Then an allocation $\{\mathbf{j}, \mathbf{q}\}$ is feasible if $\mathbf{j} \in \mathcal{J}$ and $\mathbf{q} \in \mathrm{Q}$. Let $\mathbf{a_t}$ be a U-sized vector with binary elements $a_i \in \{0, 1\}$ indicating whether firm $i$ was operating in period $t$. The set of state variables $\mathbf{s_{it}} = \{\eta_t, h_t, \mathbf{a_{t-1}}, \omega, c\}$ now includes the vector of efficiencies and lagged operating states, $\omega$ and $\mathbf{a_{t-1}}$, respectively. The costs, $G$, associated with electricity demanded for a given demand state, $\eta_t$, are defined as:

$$G(\mathbf{s_t}, \mathbf{a_t}, \mathbf{q}) = \sum_{i=1}^{U} [mc_i q_i - \mathbb{1}(a_{i,t-1} = 0, a_{it} = 1)\kappa_i]. \tag{8}$$

As with the exposition of the firm problem, I formulate the investment choice-specific value functions for the production decision for a given $\mathbf{j}$:

$$\begin{aligned} W^{2\mathbf{j}}(\mathbf{s_t}, \mathbf{q}) = \max_{\mathbf{q} \in \mathrm{Q}} \{ & -G(\mathbf{s_t}, \mathbf{q}) + \delta E[W^{2j}(\mathbf{s_t}, \mathbf{q})] \\ & s.t. \sum_{i=1}^{U} q_i \geq \eta_t \ \forall\ t \end{aligned} \tag{9}$$

The first problem is to choose the optimal investment vector $\mathbf{j}$ where $\Gamma$ now corresponds to the sum of investment costs associated with investment vector $\mathbf{j}$ and where $\mathbf{v}$ denotes the U-sized vector of stochastic shocks to investment costs:

$$W^1(\mathbf{s_0}, \mathbf{q}) = \max_{\mathbf{j} \in \mathcal{J}} \{\tilde{\delta} E[W^{2j}(\mathbf{s_0}, \mathbf{q})] - \Gamma(\mathbf{j}, \mathbf{v}, \gamma, \alpha)\}. \tag{10}$$

## 2.5 Discussion of mechanisms

### 2.5.1 Impact of carbon price on firm market share with static efficiencies

In this section I review the impact of the carbon price on firm market share, denoted $\zeta_i = q_i/q^d$, where $q^d$ denotes hourly residual demand for the fossil portfolio. I review this by holding efficiencies fixed, i.e. without considering investment, so as to isolate the impact of the carbon price all else equal. Understanding whether, and to what extent, market shares change as a result of the regulation is critical to this paper's analysis as changes in the spatial distribution of local air pollutants are driven by production reallocation across firms. Without production reallocation, the relative distribution of pollutants is unchanged even if there are aggregate local air quality co-benefits from reducing total emissions. I demonstrate that with static decision-making the carbon price does not lead to reallocation of production across firms and thus does not provide for a change in the distribution of local air pollutants. Then I show that in a dynamic setting production reallocation can occur. I illustrate that with all else constant, market shares are weakly increasing among the more efficient firms (i.e. decreasing in $\omega$), $\frac{\partial \zeta_i}{\partial \omega_i} \leq 0$.

First, let us review the impact of a carbon price on aggregate hourly supply curves in a static framework. When production decisions are made statically, aggregate hourly supply curves can be constructed with marginal costs alone. Given the technical features of production in this setting, marginal costs can be modeled as a constant function of firm efficiency and fuel cost, so the supply curve reflects a ranking of the firms by efficiencies.[21] To study the impact of the carbon price on the supply curve, we need to evaluate how the carbon price changes firms' marginal costs. Equation 5 shows that with $\tau > 0$, $\frac{\partial^2 mc_i}{\partial \omega_i \partial \tau} = e^f$. Since $e^f > 0$, we see that the increase in marginal costs following the introduction of a non-zero $\tau$ is increasing in the firm's heat rate, $\omega_i$. That is, the increase in marginal costs is larger for the relatively less efficient firms. Then, the carbon price increases the slope and intercept of the supply curve, but does not change the ranking of the firms by marginal costs.

Figure 2 illustrates supply curve changes following the introduction of $\tau > 0$.[22] As earlier, hourly electricity demand is assumed to be inelastic to wholesale electricity prices, so the quantity demanded following the carbon price in the near-term is the same. Then, at a given level of demand, $q_d$, the share of demand provided by each firm, $\zeta_i$, is the same following

[21] The ranking of firms by costs is called the "merit order" in electricity markets.

[22] In the model, firms supply discrete quantities in this market, so the supply curve is a step function composed of discrete quantities of generation from each firm.

the imposition of the carbon price, while the wholesale electricity price is higher. Thus, a static model of electricity production predicts no production reallocation across firms as a result of the carbon price, and hence no redistribution of local air pollution. If we relax the assumption about inelastic demand and introduce some price responsiveness, we still see no reallocation of production across firms; rather, the most costly firms on the margin would reduce production as demand is reduced from the higher wholesale price, which would lower local pollution around the marginal plant.

This result is driven by characteristics of cost structure in this empirical setting and the current set of abatement options. Under current economic conditions for the cost of abatement technologies, the least-cost approach to reduce GHG compliance costs among these natural gas units is to burn less fuel (e.g. become more efficient or produce less). A related outcome of the conditions in this empirical setting is a monotonically decreasing relationship between pollution intensity and firm efficiency. This relationship drives the results of this study and is important to evaluate the external validity of this result. When pollution intensity and firm inefficiency are not monotonically linked, as may be the case in other electricity markets with coal power for example, then re-allocation under a carbon price can occur in a static framework.

For another example of a setting without this aforementioned relationship between pollution intensity and firm inefficiency, consider a world where a technology such as carbon capture and storage (CCS) were of lower cost and observed as an equilibrium abatement choice. CCS provides a way of abating GHG emissions without requiring a reduction in fuel burn, where carbon is captured from pollution sources after production. Suppose firms are heterogeneous in their costs of adopting CCS. A firm that adopts CCS at low cost could re-position itself on the aggregate supply curve. However, absent equilibrium abatement technologies that re-position firms along the supply curve, in a static framework, the carbon price leads to zero reallocation of production across firms as the policy does not change the ordering of firms on an marginal cost (equivalently, efficiency) basis.

Let us now turn to the impact of the carbon price on production when a firm's participation in hourly electricity markets is a function of both marginal costs and start-up costs. Consider two infra-marginal firms, $a$ and $b$, which have the same operating capacity. Suppose that before the carbon price is introduced, the total costs of production are the same for the hour of production when both firms are not previously operating, $(q_a mc_a + \kappa_a) = (q_b mc_b + \kappa_b)$, which implies that in the cost minimization problem presented above, we are indifferent to which firm is producing. Further suppose that $mc_a < mc_b$, which means that $\kappa_a > \kappa_b$ for the above to hold. From the discussion earlier, we know that the carbon price increases marginal costs more for firm $b$ than for firm $a$. On the other hand, start-up costs are predominantly

composed of non-fuel related costs, which means the majority of these costs are unaffected by the carbon price.[23] If the relative change in start up costs is small compared to the change in marginal costs, in the cost minimization problem we now prefer firm $a$ over firm $b$, whereas before the policy we were indifferent. This is the theoretical channel in which production reallocation could occur. The extent to which it occurs in an empirical question, and depends on how much production capacity in the regulated industry is available from firms with relatively higher start-up costs and lower marginal costs.

A broader point can be made here about how environmental regulation impacts individual production decisions and aggregate supply curves. When regulations increase the cost of a polluting input, input costs increase, and firms that efficiently convert inputs to outputs are rewarded. When production decisions include costs unrelated to inputs, a firm that was less competitive because of a high non-input related cost now becomes more competitive. Production reallocation from a regulation stems from the impact of the regulation on the relative competitiveness of firms on the output market. Settings in which post-regulation pollution intensity is monotonic in pre- and post-regulation firm competitiveness do not provide for production reallocation. In these settings, the presence of non-input related costs, such as start-up costs as in this paper, or compliance technologies that re-rank firms' competitiveness on the output market as with the CCS example, create the scope for reallocation. The motivation to study production reallocation in this paper is to predict the spatial reallocation of negative pollution externalities across alternative policy scenarios. However, this question is also highly relevant to other research agendas, for example, to understand a regulation's impact on production leakage, and other local production externalities such as impacts on local employment and wages.

### 2.5.2 Impact on incentives to invest

Here I review two theoretical predictions relating to the impact of the carbon price on investments in efficiency. The first is that investments to improve efficiency (reduce $\omega$) are increasing in $\tau$. This is straightforward from the formulation of the marginal cost function. Since $\frac{\partial mc_i}{\partial \omega_i} = c^f + e^f\tau$, $\tau > 0$ increases the returns to reducing $\omega$, so that investments to improve efficiency will be weakly increasing in $\tau$.

Next, I demonstrate that under certain conditions, private and social returns to investment are increasing in $\zeta_i$. Let the private net returns to investment $j$ be the production cost savings from efficiency improvements, and let the social net returns be avoided damages

[23] Kumar et al. (2012) estimates fuel costs to be around 1 to 2 percent of start-up costs from gas-fired combined single and simple cycle large frame, and 30 percent for gas-fired steam turbines.

from emissions.[24] As shown above in the model, investment $j$ corresponds to new efficiency $\omega_i = \omega_i'(1+\delta) - j_i$. For a given $P$ and $q^d$, private net returns in a given period are $(\omega_i - \omega_i')c^f q_i$ and social net returns are $(\omega_i - \omega_i')\tau e^f q_i$. It is clear then that the discounted sum of both private and social net returns over future periods is increasing in $q_i$ and therefore also increasing in $\zeta_i$ for a given level of demand.

Now, compare the payoffs of investment across firms. If investment costs are not a function of current heat rate, as formulated in this model, then it is clear from the above that higher payoffs are achieved from improving the efficiency of the firms with higher market shares. To further explore this result, now assume that it is more costly to improve the relatively more efficient firms, so that costs are now a function of pre-investment efficiency with $\frac{\partial \Gamma_i}{\partial \omega_i} < 0$. Suppose that firms one and two have $\omega_1 < \omega_2$; investment decisions $j_1'$ and $j_2''$ cost the same, i.e., $\Gamma(\omega_1, j_1) = \Gamma(\omega_2, j_2)$; and $\omega_1 - j_1' < \omega_2 - j_2''$. In this setting it is still optimal to invest in firm one when:

$$\begin{aligned} (\omega_1 - \omega_1')(c^f q_1 + \tau e^f) &> (\omega_2 - \omega_2')(c^f q_2 + \tau e^f) \\ q_1 &> \frac{1}{c^f}\Big(\Big[\frac{\omega_2 - \omega_2'}{\omega_1 - \omega_1'}\Big](c^f q_2 + \tau e^f) - \tau e^f\Big). \end{aligned} \tag{11}$$

This result illustrates that if the market share of a given firm is sufficiently large such that the above holds for the observed distribution of pre-investment efficiencies, other firm market shares, and investment cost parameters, a policy subsidizing efficiency improvements yields the highest private and social returns when targeting higher market share firms.[25]

# 3 Empirical Strategy

The model above is used to characterize the firm's optimal investment decision and recover its optimal policy function for production (dispatch policy function). First, I group firms into $N$ representative firm types. Then I use the optimal investment decision and dispatch policy functions to simulate hourly market outcomes of production for each firm type under alternative policy scenarios. I solve the cost minimization problem in two parts. First, I find the cost minimizing solution to meeting hourly demand across different investment portfolios $\mathbf{j}$. That is, I recover investment choice-specific policy functions for production for

[24] We can incorporate the fact that damages include global and local emissions by thinking of $\tau$ and $e$ as vectors of global and local emissions and damages, respectively, since emissions of local pollutants are also increasing in $\omega$. Distinctions in damage types are omitted here for simplicity.

[25] This result is further discussed in Weber (2021).

$J^N$ investment portfolios.[26]

Next, I use each of these policy functions to simulate different sequences of hourly market outcomes over three years, and I sum the discounted production costs associated with each outcome. Then I use the estimates of investment costs (discussed subsequently), to find the optimal investment decision $\mathbf{j}^*$ given the estimated production costs across each scenario. Next I use simulations from the dispatch policy function associated with $\mathbf{j}^*$ to compare market outcomes across alternative input cost states. I discuss this process in further detail below.

## 3.1 Dispatch Policy Function

The first step is to recover investment choice-specific policy functions for dispatch. I first group firms into representative firm types based on efficiency and size, as these characteristics describe the heterogeneity across firms that enters the cost minimization problem. I use k-means clustering and scree plot analyses to establish firm type groups. Details on this process are described in Appendix 8.4. Let $i$ refer to firm type rather than individual firms, $i \in \{1, ..., N\}$, with $N = 10$. Let $\mathbf{a}$ be a vector with elements $a_i$ indicating the number of firms of type $i$ that were operating in the last period. The optimal dispatch policy function for each investment choice $\mathbf{j}$ is denoted $\sigma^{\mathbf{j}}(\eta, h, \mathbf{a}, \omega^{\mathbf{j}}, ic)$. The policy function maps state variables to $\mathbf{q}$, which has $N$ rows and two columns, where the entry in row $i$ column one corresponds to the number of firms of type $i$ operating at their minimum generation level, and the entry in row $i$ column two corresponds to the number of firms of type $i$ that are operating at their maximum generation level. The total number of firms on at either operating level is constrained by the total number of firms of that type, and the optimal dispatch policy for each state will satisfy condition (7) without violating any constraints.

I use policy function iteration to find the optimal policy function for each $\mathbf{j} \in \boldsymbol{\mathcal{J}}$, using an initial estimate of firm type start-up costs from the literature and an exogenous discount rate.[27] I first do this setting the input costs state to observed input costs; once start-up costs are estimated, I recover unique policy functions for each of the four input costs states discussed in the counterfactual analysis.

[26] To reduce the computational burden, I group firms into five firm investment type groups, setting $N = 5$ here. Further, I set $J = 2$ so that each firm investment type has the option to invest or not. When a given firm investment type invests, all firms of that type invest and reduce their heat rate by 1.5 percent, which is the average heat rate improvement observed in the data among investing firms.

[27] The initial start-up cost guess is \$80 per MW, which was the calibrated estimate used in Cullen and Reynolds (2017). To set a discount rate, I use a one-year interest rate of 4.1 percent, implying an hourly discount rate of 0.99954130.

## 3.2 Optimal Investment Choice

After the optimal dispatch policy $\sigma^{*\mathbf{j}}(\mathbf{s})$ is recovered for each $\mathbf{j}$ investment scenario, the optimal investment choice $\mathbf{j}^*$ is made by solving (7), using the sum of discounted production costs over three years as a measure of the value of each investment scenario.

## 3.3 Identifying and Estimating Structural Cost Parameters

The structural cost parameters in this model that are not observed by the econometrician are the firm's start-up costs and the cost of investment activities.

I estimate start-up costs with the following identification argument and estimation approach. In the context of the firm problem, start-up costs can be identified by the difference in a firm's willingness to operate across two states that differ only in the lagged operating state. In the cost minimization problem, under the assumption that the empirically observed dispatch is cost minimizing, start-up costs can be identified by comparing the difference in dispatch implied by solving the cost minimization problem in a given state with an initial guess of start-up costs, $\kappa_i^0$, versus the empirically observed dispatch for the same state. Accordingly, I estimate start-up costs, $\hat{\kappa_i}$, by evaluating the difference between cost minimizing dispatch implied by the dispatch policy function recovered for some $\kappa_i^0$ and the empirically observed dispatch. Specifically, I use a generalized method of moments approach (GMM) to find the $\hat{\kappa_i}$ that minimizes deviations between the simulated dispatch and empirical counterparts across like states. The estimated start-up costs are provided in Table 3, and range from \$5,885 - \$20,485 across the unit types. Details about this procedure are provided in Appendix 8.3. These estimates are comparable though lower than other start-up estimates in literature – Reguant (2014) estimates start-up costs around \$25,000 for for combined cycle gas plants that are 400-800MW, much larger than those included this data set; the start-up cost estimates in Cullen (2015) for gas plants range from \$14,894 to \$50,829.

To estimate investment costs, I first identify evidence of persistent heat rate improvements that indicate investment are described in Appendix 8.5. Next, I use a simulated method of moments (SMM) approach with the following identification argument. In the cost minimization problem, under the assumption that empirically observed investments are cost minimizing, I find the investment costs that minimize the difference in investment behavior observed empirically and those implied by finding the investment decision that minimizes production costs given an initial guess of investment costs, $\gamma^0$.

To implement the SMM approach, I first estimate investment-conditional choice probabilities (ICCPs) from the observed investment decisions by firm investment type. Then, I simulate investment moments using these ICCPs. I compare these simulated investment

decisions to the investment decisions generated from solving equation 10 with an initial guess of investment costs, $\gamma^0$. The value function in equation 10 is approximated with the three-year production costs associated with the investment scenario, plus the costs of investment. The production cost estimates are calculated from the forward simulation performed with the recovered investment choice-specific dispatch policy functions, $\sigma^{*2j}(\cdot)$. As discussed earlier, I assume the investment cost shock comes from an extreme value type 1 distribution, and I estimate the location and scale distribution parameters for the shock from the SNL data. Then I estimate $\hat{\gamma}$ by finding the parameter that minimizes the distance between the simulated moments using the ICCPs from the data and the investment decisions implied by the model. The advantage of using the SMM approach here is that it provides additional investment moments to match, which is useful in this setting as otherwise there are very few moments.[28] Additional detail about this procedure is provided in Appendix 8.2.

## 3.4 Estimating the Demand Process

The first step to estimating the model is to estimate the demand shock process. Equilibrium prices are then the market clearing prices associated with production allocations that emerge from the cost minimization problem. Given the presence of resources that are strictly preferred for dispatch over the natural gas firms modelled in this paper, we can think of the natural gas firms as satisfying each hour's "residual demand". Residual demand denotes the total hourly consumer demand less energy supplied by resources preferred for dispatch such as zero marginal cost resources like solar and wind energy, and other lower cost resources such as hydro-electric and nuclear energy, and any electricity imports. Thus, the demand shock process estimated and referred to in this paper is more precisely, a residual demand shock process.

To estimate the demand shock process I consider two demand states, high and low. I define a high state in the data as any demand state above the median demand for that hour, and a low state as any below. I estimate the probability that the demand state in a given period is "high" with the following specification:

$$\begin{aligned} \mathbb{1}[\eta_h = high] = \alpha + \xi \mathbb{1}[\eta_{h-1} = high] + \epsilon \\ \forall \quad h \in \{1, ..., 24\} \end{aligned} \tag{12}$$

where $high = 1$ if the demand state is above the median demand in that hour, and $\eta_{h-1}$

[28] On the other hand, GMM is sufficient to recover start-up costs, as in that setting there are many dispatch moments available across different states.

denotes the last period's demand state. This model is estimated separately for each hour of the day, and the coefficient estimates for $\xi$ are the inputs to the transition probability matrix for that hour. This parsimonious specification explains over 70 percent of the variation in high and low demand states over the period 2013 - 2015; the results from a modified version of this specification estimated over all hours of the day and including hourly fixed effects are provided in Table 7. Further details are provided in the Appendix Section 8.6.

An important feature of California's energy market during the time period of this study is the large presence of renewable energy and the state's aggressive renewable portfolio standards. As mentioned above, the demand process estimated here is net of existing levels of renewable resources, and any other preferred resources, available over the time period 2013 - 2015. As discussed earlier, the objective of this paper is to provide a near- and medium-term analysis of the impact of the carbon policy. The demand process estimated here is not expected to hold over longer term horizons in which the quantity of renewable energy is expected to increase. Future research could explore how changing this price process, for example, under a high renewable energy penetration scenario, impacts the results of this model. Over longer term horizons, the direction of the change in levels residual demand is not obvious. For example, a growing California economy would increase residual demand, while increasing renewable energy standards would decrease residual demand overall.[29] Further, longer time horizons would provide time for electricity rate proceedings to adjust in response to wholesale prices, which could lead to an increase or decrease in consumption depending on whether or not renewable energy penetration and other trends have dampened prices.

# 4 Empirical Setting & Data

Estimating this model requires observing size, efficiency, and location, as well as hourly production quantities, fuel inputs, and emissions for all firms regulated by the cap-and-trade program in California.[30] Firm-characteristics and hourly production and emissions quantities are obtained from the subscription-based data provider SNL, which compiles data collected through various federal reporting requirements. SNL data include those from continuous emissions monitoring systems (CEMS), which exhibit some data anomalies related to heat rates reported for combined cycle units. Periods of unrealistically high and low heat rates

[29]The shape of the residual demand curve across hours of the day would also be expected to change with increases in both renewables and demand, which would increase the difference between peak solar energy in the middle the day, and peak consumer demand in the evening. Yet, increasing adoption of storage technologies could also dampen this effect. See Schmalensee (2022) for a discussion of this.

[30]Hourly prices are also obtained from publicly available data maintained by the California Independent System Operator (CAISO) to estimate inferred profits in Figure 3.

indicate data error, likely stemming from periods where these units report generation and/or fuel inputs from one and not both of the steam and gas units. Further detail about how this is addressed is provided in the Appendix 8.9.

To connect changes in firm emissions of local air pollutants to damages from human health impacts, I use marginal damage estimates from AP3, formerly known as the Air Pollution Emission Experiments and Policy Analysis Model (Muller and Mendelsohn 2007), now updated and called AP3 (Clay et al. 2019). AP3 is an integrated assessment model that uses an air quality model to determine how an additional ton of pollutant emissions contributes to pollution concentrations and exposures. The air quality model simulates transport and chemical transformations, including the transformation of volatilie organic compounds ($VOC$), amonia ($NH_3$), nitrogen oxides ($NO_x$), and sulfur dioxides ($SO_2$) to particulate matter. The model also incorporates the impact of emissions of $VOC$s and $NO_x$ on tropospheric ozone ($O_3$), as well as the titration effect, where in environments with sufficiently high $NO_x$, an additional emissions of NOx can reduce ozone. It then uses peer-reviewed dose-response functions to compute the resulting consequences to human health. Finally, it uses standard estimates of the value of mortality and morbidity risks to monetize these effects. Details about the methodology included in the model and subsequent updates are provided in Muller and Mendelsohn (2007) (APEEP), Muller, Mendelsohn and Nordhaus (2011) (AP2), and Clay et al. (2019) (AP3).

Modelling the relationship between pollutant emissions and human health is challenging, and in particular, the contribution of precursor emissions, $SO_2$, $NO_x$, and $NH_3$ into concentrations of $0_3$ and $PM2.5$. While chemical transport models are the most advanced scientific approach to do so, they are cumbersome and computationally intensive to use. AP3 is one of several 'reduced-complexity air quality models' (RCMs) developed over the last couple decades that facilitates researchers access to estimates of the human health costs of pollutant emissions – Gilmore et al. (2019) provides a helpful discussion and comparison of prevalent RCMs.

This paper uses two types of estimates from AP3: one, the estimates of the marginal damages of a pollutant emitted in a given county, and two, county-specific estimates of the marginal damage of an additional pollutant to that county. The latter serve as inputs to the former calculation in AP3 and were made available by the AP3 authors for this paper. The estimates of damages from pollution sources are used in this paper in the Pigouvian tax scenario to set a pollution tax at the economically efficient level, equivalent to the marginal damages of the pollutant. The estimates of the damages to a county from an additional pollutant are used to estimate damages that result in each of the scenarios.

Finally, I use the CalEnviroScreen model to review the distribution of avoided damages

by pre-existing pollution levels. The CalEnviroScreen model was developed by the Office of Environmental Health Hazard Assessment (OEHHA) to develop a numeric rating for the relative pollution burdens and vulnerabilities across census tracts in California (OEHHA and CalEPA 2017). The California Environmental Protection Agency (CalEPA) uses these scores to identify communities in the State with disproportionate pollution burdens, and such communities are identified as "disadvantaged communities." I average these scores at the county-level to match AP3's county-level damage estimates. In using these scores, it is important to note that the distribution of pollution across communities can be thought of as equilibrium outcomes, the result of firm and household sorting processes. Over time, an environmental policy that cleans up certain neighborhoods could induce price changes and sorting. See Banzhaf and Walsh (2013) and Banzhaf (2012, Chapter 2) for a discussion of pollution improvements and environmental gentrification. While the dynamics of population and firm re-sorting following a carbon price are out of scope for this near- and medium-term study, they are important to consider for longer term analyses.[31]

Over the 2012 to 2015 period, the portfolio of fossil-fuel generating units operating in this period in CA consists of around 200 producing units across steam, gas, and combined cycle combustion turbines, with summary statistics available in Table 1.[32] To reduce the computational burden of estimating this model, these units are grouped into 10 firm types as discussed previously, based on firm heat rates and sizes, with summary statistics provided in Table 2. The process to establish firm type groups that explain the most variation in firms along these two dimensions is explained in Appendix 8.4.

The model in this paper assumes that the market is competitive, where the actions of the electricity firms do not impact prices. While the well-known Wolak, Borenstein and Bushnell (2000) paper finds evidence of market power in the California electricity market, the market has undergone significant changes since the study period evaluated in that paper. In particular, in early 2009 the CAISO implemented a host of market reforms intended to improve price transparency and prevent market manipulation as part of its Market Redesign and Technology Upgrade (MRTU). MRTU improved the grid operator's ability to manage real-time congestion with day-ahead generation and transmission schedules, increasingly local marginal pricing by moving from three to 3,000 price nodes, and created an integrated forward market for electricity, transmission capacity, and reserves (CAISO 2009). Through MRTU and other reforms the grid operator in California has evolved significantly in its ability

[31]See Kuminoff, Smith and Timmins (2013) for a review of recent structural spatial equilibrium models that could be used to incorporate household sorting into a study of which households benefit from place-based improvements in environmental amenities.

[32]In 2013, coal units are included in the data; however, these units are located outside of the state and thus not included in this table nor in the estimation.

to detect and prevent market manipulation since BBW's analysis. For example, CAISO currently maintains a process to submit local market power mitigation bids in locations and times that may present opportunities for local market power, such as in heavily congested times and regions. Further, the Herfindahl-Hirschman Index (HHI) for the natural gas units is 0.06, and the majority of units provide less than two percent of total fossil-fuel generation. While these facts do not entirely rule out the presence of non-competitive behavior, the modeling approach in this paper can be understood as predictive of market outcomes when firms behave competitively, which, as the market monitoring reports from California's grid operator suggest, has been the case for the large majority of hours studied in this paper.[33]

In this paper, I do not model transmission constraints; the results of the model and simulation procedure are predictive of what would occur under least-cost dispatch, excluding hours and firms which are dispatched for other reasons not included in the model, such as transmission constraints. In addition, I do not explicitly model imports. The residual demand shock process estimated includes observed imports over the study time period. Over this time period, California's electricity imports remained relatively stable, around 24 and 25 percent of total retail electricity sales in each year 2012 - 2015 (U.S. Energy Information Administration 2014, 2015, 2016, 2017). The carbon price's anticipated effect on imports is not clear. CARB included a provision in the cap-and-trade regulation that intended to prevent emissions leakage, which requires that any entity importing power to California report and obtain carbon allowances to cover the emissions associated with the imported power. For some out of state generation sources, CARB assigns emissions factors by source, for others, CARB assigns a default emissions factor (CARB 2010, 2012, 2013).[34] These provisions do not rule out the potential for a carbon price to impact import quantity. If higher carbon prices such as those simulated in the counterfactuals changed the amount of electricity imported, for example, to increase electricity imports, that would decrease the residual demand needed to be provided by the modelled domestic natural gas firms, reducing total overall emissions. That said, a level shift in imports in all hours would shift residual demand to the left and would not impact the merit order of firms modeled here, which is the focus of this paper given the research questions's emphasis on production and emissions re-allocation.

[33]CAISO's quarterly market monitoring reports for 2014 and 2015 find the overall combined wholesale cost of energy was around (including slightly below, close to equal, and slightly above) their simulated competitive baseline prices under competitive conditions, with some price spikes in a small set of intervals in the second half of 2015 and in Q4 2015 (CAISO 2014, 2015).

[34]See Pauer (2018) for background on the political factors that lead to the inclusion of emissions responsibilities in the regulation.

## 4.1 Descriptive Assessment

Understanding the impacts of the GHG policy on local air quality outcomes requires characterizing changes in market shares and efficiencies as a result of the carbon price, as well as the spatial distribution of these changes. Section 2.5 provides the theoretical prediction that market shares would be weakly increasing among more efficient firms following the introduction of the cap-and-trade program. For a descriptive assessment of what has happened in the program, I estimate changes in firm market share by comparing the average share of total hourly demand provided by each firm in 2015 to that provided before the cap-and-trade program started in January 2013. I plot these market share changes by the firm's average heat rate over 2012 - 2015. The graph in Figure 4 shows a negative relationship between market share changes and heat rate, supporting the theoretical prediction, with a pairwise correlation coefficient significant at the 0.10 level.

To connect changes in market share to the spatial distribution of damages from local air pollutants, I plot market share changes by the AP3 estimate of county-specific $NO_x$ damages where the firm is located. I do not posit a theoretical prediction for the trend. The trend depends on the location of firms that increase (decrease) market share, and the location of counties with high (low) $NO_x$ damages.[35]

The graph in Figure 5 exhibits a negative slope, with a pairwise correlation coefficient significant at the 0.10 level. The implication of these two descriptive results is that the program leads to some market share changes in the direction expected. In addition, there is some evidence that these changes occur in counties where sources emit pollutants with lower damages. Next, I plot the changes in firm efficiencies by the air pollution damage estimate where the firm resides in Figure 6. The figure does not exhibit a trend among heat rate improvements and the local marginal damages from air pollution, and the pairwise correlation coefficient is not significant at the 0.10 level.

The next section presents the results from my model which isolates the impact of carbon prices on market outcomes of production and efficiency changes holding other market features fixed and incorporating dynamics in decision-making.

[35]Further research could explore determinants of the spatial location of the firms, including by characteristics that determine whether a firm is likely to increase or decrease production following a regulation.

# 5 Results

## 5.1 Model Fit

To review model fit, I compare simulated dispatch outcomes from policy functions recovered with heat rates corresponding to firm average heat rates in 2012 to observed dispatch outcomes in 2012. Doing so requires first estimating start-up costs. As discussed above, start-up costs are estimated as the firm type length vector that minimizes the difference between observed dispatch quantities, and dispatch quantities in a simulation with policy functions recovered with zero cost start-ups. I estimate start-up costs using 2012 dispatch outcomes. Details on the estimation and simulation procedure are provided in Appendix 8.1.

I perform T-tests of market share observations in the simulation and those observed empirically, paired by hour of the day and firm type. Table 4 shows that the shares by hour and firm type are not statistically different across simulation and empirical dispatch. Next, I develop hourly average generation quantities and plot them by firm type for the simulation and empirical dispatch, provided in Figure 7. The figure visually shows the model fits most but not all firm types well. For example, firm type 9 is not dispatched in the model, but is dispatched empirically. These differences could reflect that the observed dispatch is not strictly cost-minimizing and/or that there are other constraints that the California electric grid operator faces in scheduling firms for dispatch that the model does not capture. For example, local transmission constraints could lead more costly firms to be scheduled to meet local demand in congested regions and hours. I view the model and simulation procedure as predictive of a least-cost dispatch, excluding hours and firms which are dispatched for other reasons not included in the model.

Before the model can be used to simulate counterfactuals, investment costs must be estimated. To do this, I recover dispatch policy scenarios for a set of investment scenarios. I then use these investment scenarios to simulate market outcomes and calculate the discounted sum of costs associated with each investment scenario. I recover the investment cost that rationalizes the investment conditional choice probabilities observed in the data, given the production cost savings associated with alternative investment scenarios. Further details on this procedure are provided in the Appendix.

## 5.2 Stringent Carbon Policy Scenario

A more stringent climate change policy corresponds to a tighter GHG emissions cap, which increases carbon price and leads to more abatement.[36] To evaluate the market outcomes across climate change policy stringencies, I compare production and emissions outcomes across alternative carbon prices including a permit price equal to zero (henceforth, no carbon policy scenario), and a policy that leads to permit prices of \$36 and \$105 per ton of $CO_2e$ (2007 USD), corresponding to the Obama Administration's Interagency Working Group on the Social Cost of Greenhouse Gases central and high estimate of damages for $CO_2$ emissions emitted in 2015 with a 3 percent discount rate (henceforth, the SCC carbon policy scenario) (EPA 2017). The average fuel cost, $c^f$, over 2012 - 2015 was \$3.6 per MMBtu, and the average permit price over this period translates to \$0.70 per MMBtu (\$13 per ton $CO_2e$). The SCC carbon policy scenario translates to an additional \$2.2 per MMBtu of input cost. As shown in equation 5, in the model the carbon price enters the firm's profit maximization problem (equivalently, the cost minimization problem) by increasing the cost of inputs. Accordingly, I evaluate market outcomes across four input cost states $c = f + \tau$, $\tau = \{0, 0.70, 1.9, 5.7\}$.

In simulating outcomes across policy stringencies, I leverage the dual decision framework of the model to evaluate the impact of carbon prices together with endogenous efficiency investments that respond to these prices. To do so, at the beginning of the cap-and-trade program I find the cost-minimizing investment decision for all firm type efficiencies. Equivalently, in the profit maximization framework firms decide whether to invest to improve their efficiency or not, assuming that all other firm types make optimal investment choices. Investment decisions are made by finding the investment portfolio that is cost minimizing with respect to a market outcomes over a time period $T =$ three years, using the estimated investment costs.

If the California market is stationary, then some $T$ provides a reasonable comparison of the value functions across investment portfolios, and this decision will satisfy equation 10, corresponding to the decision to maximize value over an infinite time horizon. However, I consider that the California market will evolve over longer time horizons, for example, from increased renewable energy penetration, and I view this modeling choice as consistent with the near- and medium- term objectives of this paper's analysis. Further, I view this choice as reasonable in the context of firms' decision-making in this setting, noting that this assumes that firms make efficiency investment decisions myopic to outcomes after $T$. I sum investment costs and discounted production costs for each evaluated investment portfolio

[36] The California cap-and-trade program coverage is broader than the electricity industry, and the equilibrium permit prices are determined by abatement costs across all regulated sectors.

over the simulations and then choose the investment portfolio in each carbon policy that minimizes costs.

Once the optimal investment and dispatch policy functions are identified, market simulations can be performed across input cost states. Figure 8 shows the results of these simulations, comparing the total generation provided by firm type in the current and medium and high carbon price scenarios compared to a no carbon policy scenario including endogenous investment. The figure illustrates that the current carbon price leads to minimal production reallocation across firm types. This result corresponds to minimal reallocation of local air pollutants, and minimal co-benefits (and co-costs), as a result of the policy. Higher carbon prices in the medium high and high carbon price scenarios lead to more noted production reallocation. These results suggest that the current price is small relative to other production costs, such that is doesn't alter firm decision-making. Only carbon prices in the SCC and high SCC scenario are large enough to alter firm decision-making compared to a no carbon policy scenario.

Table 5 summarizes the changes in market share by firm characteristics. Unconditional on fixed costs and size, counterfactuals with higher carbon prices lead firms with higher marginal costs to decrease production relative to a no carbon price scenario. Conditional on marginal cost, firms with higher start-up costs decrease market share. However, conditional on marginal costs and size, firms with higher start-up costs increase market share. This is because the estimated start-up costs are positively correlated with firm size, and column three shows that relatively smaller firms are dispatched more often in counterfactuals with higher carbon prices. Thus, after controlling for size, we see that firms with larger start-up cost are used more often in high carbon price counterfactuals. Overall, column three confirms our theoretical prediction: higher carbon prices lead to a preference for relatively higher start-up and lower marginal cost firms, conditional on firm size.

Figure 8 shows the change in $NO_x$ damages from the production reallocation that occurs in the medium and high carbon policy scenarios. This calculation is done by connecting changes in generation by firm type to the location of firms and the $NO_x$ damages that result from those firms' emissions. The figure shows that generation reductions correspond to avoided $NO_x$ damages, though the relationship is not one-to-one. This stems from the spatial heterogeneity of $NO_x$ damages across the state. Overall, the medium high and high carbon policy reduces $NO_x$ damages by around \$3.5 and \$5.1 million, respectively, over one calendar quarter compared to a no carbon policy scenario, roughly 14 and 20 percent of $NO_x$ pollution damages in one quarter in the no carbon policy scenario.[37] The current carbon

[37]For reference, the no carbon policy scenario is estimated to lead to \$24.7 million damages from $NO_x$ pollution over one quarter.

price leads to minimal production changes compared to the no carbon price counterfactual, decreasing damages by \$0.4 million for 1 calendar quarter.

To understand which communities are receiving the benefits in these policy scenarios, I connect changes in $NO_x$ damages to the communities in which they occur in California. Since the observed carbon price counterfactual showed little variation compared to a no carbon policy, I focus on the medium high (central SCC estimate) carbon price scenario here. Figure 9 plots $NO_x$ damages avoided in a county in the high carbon price scenario, by a measure of counties pre-existing pollution burden, county-average CalEnviroScreen scores.

The figure shows a positive trend in $NO_x$ damages avoided and pre-existing pollution burdens as measured by the CalEnviroScreen score, indicating that avoided damages in the medium high carbon price scenario would happen in counties with relatively higher pre-policy pollution. The map in figure 10 plots these avoided damages on top of a map of California's disadvantaged communities, showing that reductions occur more frequently in these communities.

## 5.3 Efficiency Investment Scenarios

Minimum efficiency standards and other command-and-control type of policy mechanisms are frequently encountered in the energy sector as alternatives to market-based GHG regulation. For example, the Corporate Average Fuel Economy (CAFE) standards for vehicles and the U.S. National Ambient Air Quality Standards (NAAQS) for stationary sources establish maximum thresholds of pollution intensities. In this section I consider the impact of this type of policy in this setting by evaluating the private and social returns from alternative portfolios of efficiency improvements. The portfolios could stem from a policy that, for example, established minimum carbon intensity standards or subsidized efficiency investments for certain types of firms.

To compare alternative investment scenarios, I evaluate the set of heat rate vectors that would result from different combinations of firm type investments. I group the firms into five types based on heat rates and the previous firm type grouping, and I will refer to these five types as investment firm types.[38] I allow each investment firm type to choose to invest or not, where investment results in a 1.5 percent reduction in heat rate, which is the average heat rate reduction observed in the set of firms identified to invest over the 2013 - 2015 time frame.[39] The result is $2^5$ alternative investment decision scenarios $j$, mapping to $J = 32$ different 10 by 1 vectors of firm type heat rates where $j = 1$ corresponds to

[38] I collapse the ten firm type grouping into five firm type groups by sorting on heat rate.

[39] Details on the procedure to identify investment are provided in Appendix 8.5

no investment and $j = 32$ corresponds to all firms investing.[40] The policy iteration that was used in the alternative carbon policy counterfactuals provides the dispatch policies for each of these investment scenarios, and here I evaluate market outcomes simulated over three years for each $j \in J$ investment scenarios. I sum the discounted costs incurred across each $j$ investment scenario, averaged over simulations $S$ , $\frac{1}{S}\sum_{t=1}^{T}\beta^{t-1}G(\mathbf{q_t^{j*}}|\omega^{\mathbf{j}})$, where $\mathbf{q^{j*}}$ is determined by the recovered policy function for investment scenario $\mathbf{j}$.

Figure 11 plots the savings in production costs for each of the alternative investment scenarios compared to no investment (henceforth, gross investment returns) across carbon price scenarios. Changes in gross investment returns are driven by two mechanisms, which I review now in the framework of the cost minimization problem. One, the investment could change a preference for one firm type over another due to the improvement in efficiency (reduction in marginal costs). Two, even if the efficiency improvement does not change preferences across firms, and market shares stay the same following the efficiency improvement, private and social costs would still be reduced if some of the improved firms have non-zero market share.[41] The figure demonstrates that, as expected, the highest gross returns to investment occur in the high carbon price scenario. This illustrates the theoretical point made in Section 2.5 – the carbon price can increases the gross returns to investing to increasing firm efficiency, since the carbon price increases the cost of inputs and thereby increases the costs avoided from improving firm efficiency. The expected returns to investment can be seen as a measure of willingness to pay for efficiency improvements, and the figure shows that higher carbon prices increase some firms' willingness to pay for efficiency investments.

However, we also see variation in gross investment returns across investment scenarios within a given carbon price state. While the scenario with the highest gross returns occurs in the high carbon price, the higher carbon price does not increase the returns for all investment portfolios. The scenarios that do not lead to higher returns in higher carbon prices are those in which investment occurs among firms that are dispatched relatively less often. This is shown in Figure 12, which plots the gross savings from investment compared to the no investment scenario, by the sum of the marker shares of the firms investing in a given scenario. A clear positive correlation emerges among gross returns and market share. This confirms the second theoretical prediction made in Section 2.5; in this industry and empirical setting, policy makers would be better off targeting investment subsidies to the firms with higher

[40] The 5-type grouping is used only for the investment decision; then, for the purposes of dispatch, the investment decision is mapped onto the 10-type grouping used earlier.

[41] Whether or not investment in efficiency improvement changes the ranking of firms in costs depends on both the quantity of efficiency improvement evaluated as well as the distribution of efficiencies. Here all efficiency improvements improve the heat rate by 1.5 percent; whether or not such an improvement changes preferences across firms for dispatch depends on how the firm that improves compares to the firms close to it in terms of costs.

market shares, which in this setting corresponds to the firm with lower pollution intensities.

This result has important implications for environmental regulation, especially in settings where minimum efficiency standards have been used to meet policy objectives. The findings here contradict the claim that policy should focus on improving the dirtiest, least efficient capital to reduce pollution. Rather, here we see larger gains from improving the lower-cost, more frequently utilized capital, which in this setting corresponds to the relatively cleaner firms. Section 2.5 shows the conditions on market shares and investment costs under which this result would hold elsewhere.

## 5.4 Local Air Pollution Tax Scenario

An overarching feature of this empirical setting is the presence of one policy regulating an industry that is responsible for multiple externalities. The efficient way to address a setting with multiple externalities is to implement one policy per externality. I simulate this approach by evaluating a counterfactual scenario with a location-specific Pigouvian tax on $NO_x$ emissions.[42] In this scenario, firms make production decisions as earlier, but with augmented marginal costs. Marginal costs now include two compliance costs, one for GHG emissions and one for $NO_x$ emissions. The $NO_x$ emissions tax is set to the AP3 estimate of the \$ per ton damages for $NO_x$ emitted in county $k$ in which the firm is located. Marginal costs are now calculated as:

$$mc_{ik} = \omega_i(c^f + e^f \tau^{ghg}) + \iota_i \cdot \tau_k^n, \tag{13}$$

where $\iota_i$ is the firm's $NO_x$ emissions intensity (ton per MWh produced), which is observed from the CEMS data, and $\tau_k^n$ is the tax on $NO_x$ in county $k$, distinct from the carbon price, $\tau^{ghg}$. In this scenario I again leverage the correspondence between the competitive equilibrium outcomes and the solution to a cost minimization problem. I find the optimal dispatch policy function in a scenario with currently observed carbon prices as well as the local air pollution tax. The results are summarized in Figure 13, which shows the change in market shares under the local air pollution tax scenario, as well as the resultant changes $NO_x$ damages.

The simulated market outcomes in this counterfactual show a reduction in $NO_x$ damages by \$3.4 million over one calendar quarter, compared to a carbon policy-only scenario at the

[42] A tax on $SO_2$ is not considered as the quantity emitted by these firms is small. It would be desirable to consider a tax on $PM_{2.5}$, but firm PM contributions are not observed. In addition, it's worth noting that $NO_x$'s

observed carbon prices.[43] Taking this result together with the earlier results, we see that adding a local air pollution tax to the observed carbon price leads to a reduction in damages in a similar magnitude as with moving from the observed carbon price to the medium high carbon price at the central SCC estimate.

# 6 Conclusion

The results from the simulations present several important findings. Principally, I find that the structure of the fossil-fuel electric portfolio in terms of marginal costs, start-up costs, and the location of the firms, does not provide for a meaningful redistribution of local air pollutants under a carbon pricing policy at the observed prices in California. While this result quells concerns about the potential adverse equity impacts of the program from electricity sector pollution, it also shows that the carbon price provides minimal improvements in local air pollution. Higher carbon prices do provide for redistribution of air pollution and co-benefits from local air quality improvements, and a carbon price at the central SCC estimate provides a reduction in damages compared to the observed price in a similar magnitude as would be provided by a local air pollution tax together with observed carbon prices. The results indicate that damages to human health from local air pollution emitted by the electricity sector can be reduced by either a local air pollution tax in addition to the carbon price at current carbon prices or by increasing the stringency of the carbon pricing program.

These findings also provide insights for other regulated industries with dynamic production decisions. First, they suggest that in industries where firms use a common fuel and fixed costs are small compared to marginal costs are less exposed to market share changes following an increase in the cost of inputs, which has important implications for analyzing production and emissions leakage potential in other regulatory settings. Second, they are instructive in evaluating alternative investments in efficiency. The imposition of minimum efficiency standards is a common tool in environmental and air quality regulation, yet in this setting, such a standard would lead to far fewer savings in production costs and $NO_x$ damages avoided as compared to a regulation that improved the efficiency among firms with lower pollution intensities. This finding is driven by a key characteristic of the firms studied here — for the portfolio of natural gas firms, pollution externalities are decreasing in production efficiency, and this finding may be generalizable to other industries with this feature.

[43] $NO_x$ damages in the scenario with observed carbon prices and no local air pollution tax are \$24.3 million over one calendar quarter, compared to damages with a carbon price and a local air pollution tax of \$20.9 million.

# 7 Tables and Figures

Table 1: Fossil Unit Summary Statistics, CA 2012-2015

| | 2012 | 2013 | 2014 | 2015 |
|---|---|---|---|---|
| Num. units with non-zero production | 221 | 193 | 207 | 201 |
| Steam turbine | 50 | 37 | 39 | 37 |
| Gas turbine | 90 | 85 | 87 | 87 |
| Combined cycle | 81 | 71 | 81 | 77 |
| Mean capacity MW | 139 | 148 | 134 | 136 |
| Total capacity GW | 30.6 | 28.7 | 27.8 | 27.3 |
| Num. units with capacity change up or down | . | 3 | 2 | 2 |
| Max MW capacity change | . | 6.5 | 0.9 | 10 |
| Mean heat rate (Btu per KWh) | 9183 | 8943 | 8891 | 8966 |
| Median heat rate (Btu per KWh) | 7797 | 7318 | 7511 | 7579 |
| Percent of hours operating | .35 | .30 | .35 | .35 |

Table 2: Unit Characteristics by Type

| Type Num. | Num. Units | Size MW | 2012 HR | MC Rank |
|---|---|---|---|---|
| 1 | 16 | 87 | 17606 | 4 |
| 2 | 10 | 77 | 9327 | 5 |
| 3 | 14 | 88 | 7594 | 3 |
| 4 | 8 | 73 | 8412 | 4 |
| 5 | 33 | 184 | 7012 | 2 |
| 6 | 6 | 201 | 6566 | 1 |
| 7 | 26 | 131 | 13910 | 10 |
| 8 | 40 | 101 | 12330 | 8 |
| 9 | 32 | 191 | 11324 | 7 |
| 10 | 26 | 173 | 10572 | 6 |

Figure 1: California Carbon Price, 2013-2015. Source: California Carbon Dashboard and CARB

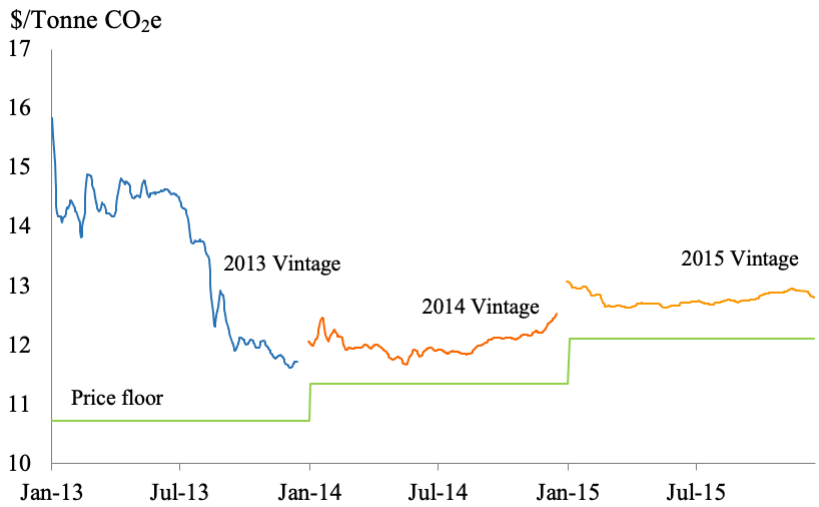



Figure 2: This figure shows the impact of the carbon price on an illustrative supply curve for a given hour $t$. Following the carbon price, the portion of the supply curve composed of natural gas firms shifts up and increases in slope from $mc_t$ to $mc_t'$.

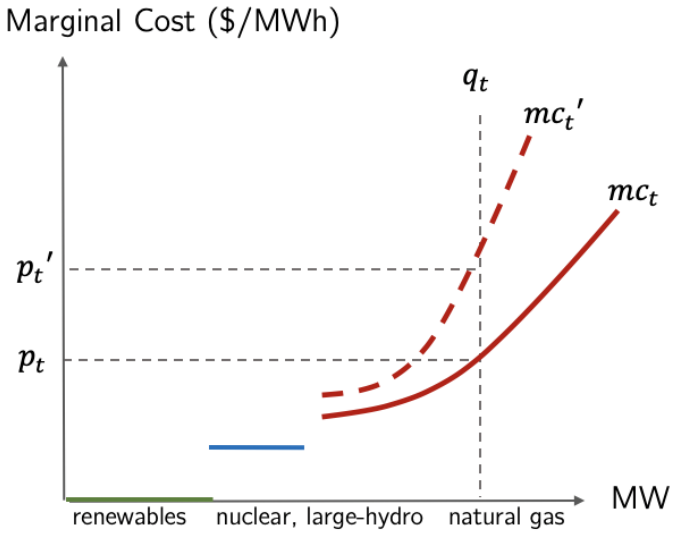


Figure 3: This figure plots average profits per hour without start-up costs, i.e. $q(P - mc)$, for natural gas firms on the y-axis, versus hour of the day on the x-axis.

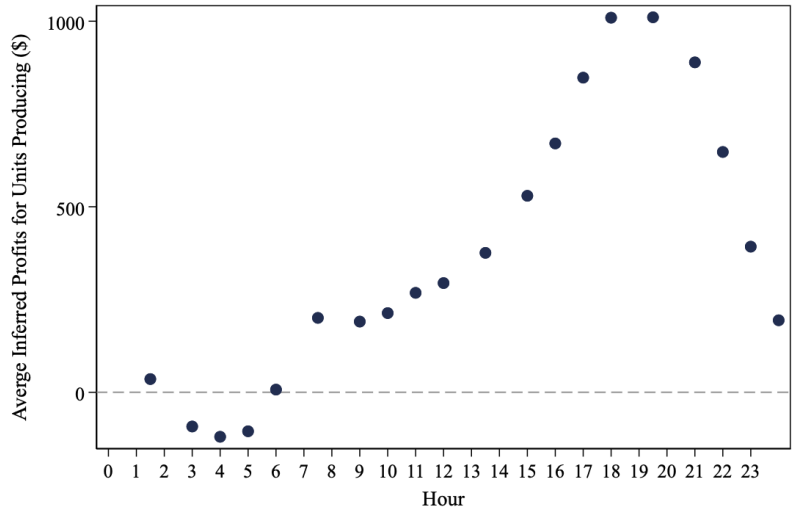



Figure 4: This figure compares market shares, $\zeta_i$, pre- and post- the carbon price (2012 to 2015) on the y-axis, versus the average firm heat rates, $\omega_i$, over the same period on the x-axis. Observations above (below) the 95th (5th) percentiles are removed for visual ease. A linear trend line is shown in blue; the pairwise correlation coefficient is significant at the 0.10 level.

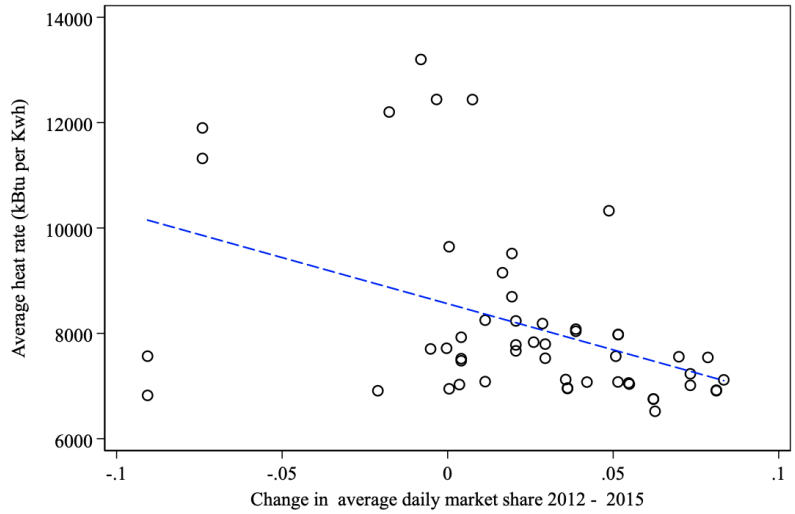


Figure 5: This figure plots each firm's change in market share pre-and post- the carbon price (2012 to 2015) on the y-axis, by the AP3 estimate of marginal damages from $NO_x$ on the x-axis. Observations above (below) the 95th (5th) percentiles are removed for visual ease. A linear trend line in shown in blue; the pairwise correlation coefficient is significant at the 0.10 level.

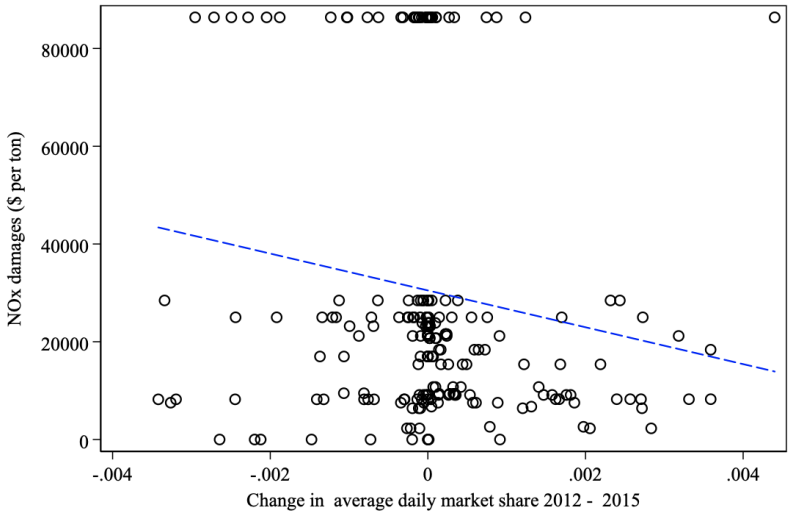



Figure 6: This figure plots each firm's change in efficiency pre- and post- the carbon price (2012 to 2015) on the y-axis, by the AP3 estimate of marginal damages from $NO_x$ on the x-axis. Observations above (below) the 95th (5th) percentiles are removed for visual ease. A linear trend line in shown in blue; the pairwise correlation coefficient is not significant at the 0.10 level.

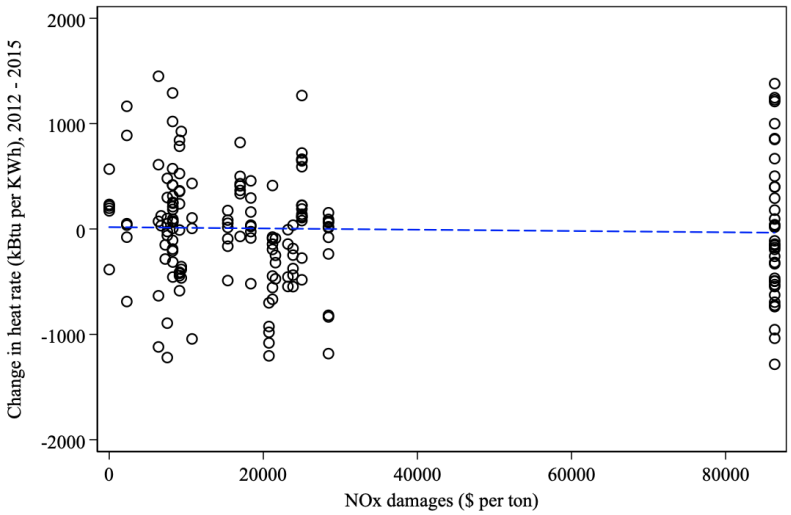


Table 3: Start-up Cost Estimates by Type

| Type Num. | Start-up Cost at $80 per MW | Estimated Start-up Cost | Asymptotic Variance |
|---|---|---|---|
| 1 | 6960 | 7,676 | 0.1267e-08 |
| 2 | 6160 | 5,885 | 0.0651e-08 |
| 3 | 7,040 | 8,977 | 0.0438e-08 |
| 4 | 5840 | 7,063 | 0.06408e-08 |
| 5 | 14,720 | 10,304 | 0.1425e-08 |
| 6 | 16,080 | 20,485 | 0.2345e-08 |
| 7 | 10,480 | 10,799 | 0.4125e-08 |
| 8 | 8,080 | 9,742 | 0.4555e-08 |
| 9 | 15,280 | 11,014 | 0.4248e-08 |
| 10 | 13,840 | 10,463 | 0.5209e-08 |

Given the computational time to perform the estimation, asymptotic variance is reported instead of bootstrapped standard errors.

Figure 7: The figure compares the average hourly dispatch outcomes by firm type in simulations and empirical dispatch in 2012.

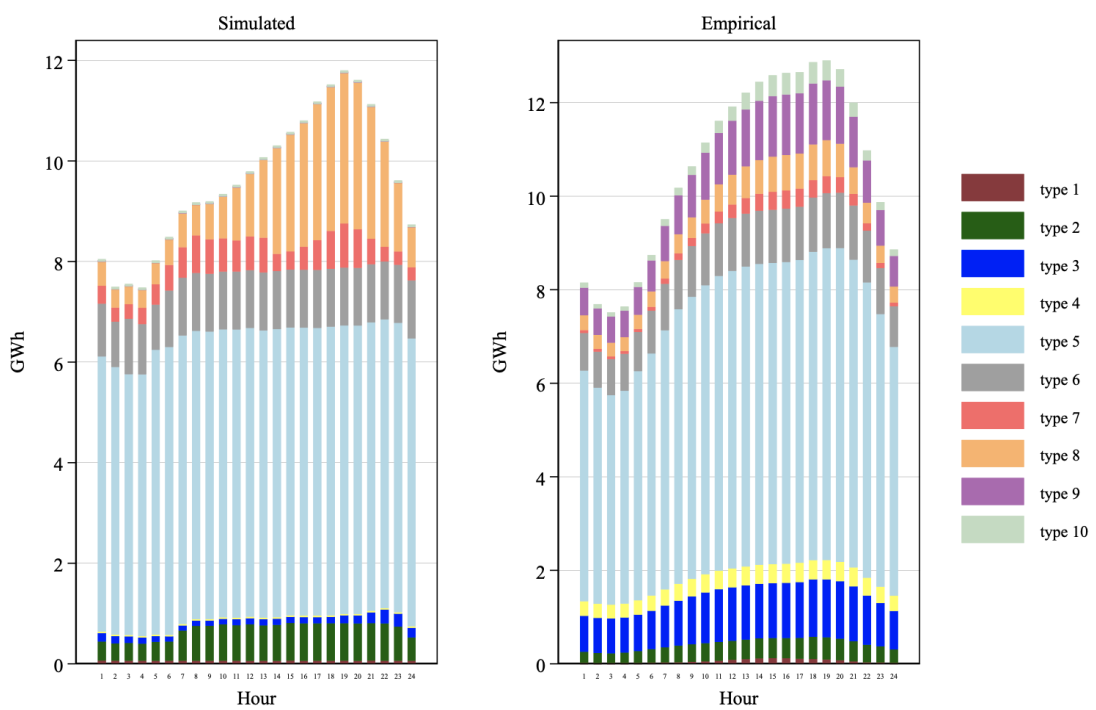


Table 4: Paired T-test for Market Share (MS), Simulated and Empirically Observed

| | Obs | Mean | Std. Err. | Std. Dev. | $H_o$: (diff $< 0$) $\Pr(T < t)$ | $H_o$: (diff$\neq 0$) $\mathrm{P}(\mid T \mid > \mid t \mid)$ | $H_o$:(diff $> 0$) $\Pr(T > t)$ |
|---|---|---|---|---|---|---|---|
| Simulated MS | 21,840 | .10 | .0012 | .1944 | | | |
| Empirical MS | 21,840 | .10 | .0012 | .1825 | 0.4989 | 0.9979 | 0.5011 |

Figure 8: The graph on the left shows the change in generation provided by firm type for current and high carbon price scenarios, compared to generation provided in a no carbon price scenario, inclusive of endogenous investment decisions. The graph on the right shows the change in $NO_x$ damages by firm type at current and high carbon price scenarios compared to a no carbon price scenario. Damages are summed over one calendar quarter of market outcomes.

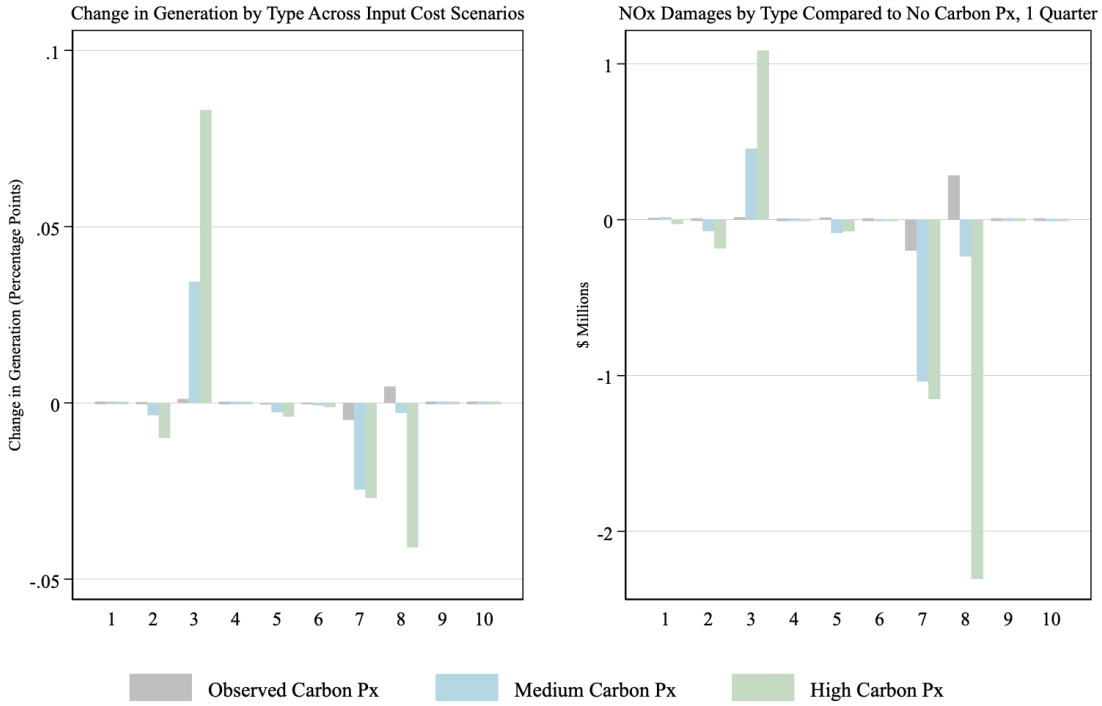


Table 5: Correlations of Counterfactual Results and Unit Characteristics

| | (1) Market Share | (2) Market Share | (3) Market Share |
|---|---|---|---|
| Marginal costs ($10) | -0.0988 | -0.109 | -0.108 |
| | (0.0756) | (0.0778) | (0.0784) |
| Start-up costs ($1000) | | -0.0451 | 0.0302 |
| | | (0.0697) | (0.132) |
| Size (MW) | | | -0.00708 |
| | | | (0.0105) |
| Observations | 40 | 40 | 40 |

Standard errors in parentheses. Number of observations reflects four scenarios times 10 firm types.

Figure 9: This graph shows the change in $NO_x$ damages within a county in the high carbon price scenario, compared to the no carbon price scenario. The y-axis indicates the average CalEnviroScreen3.0 scores of census tracts within a county, where higher values indicate higher pollution burdens.

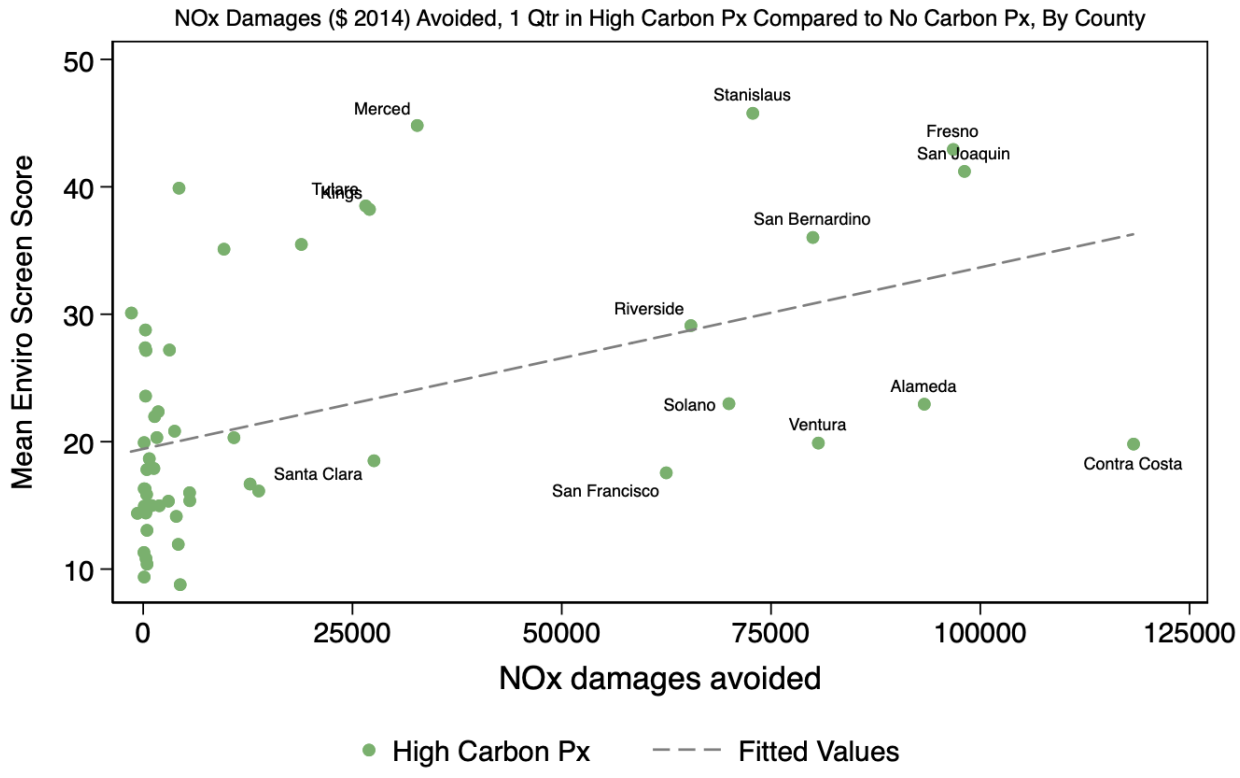


Figure 10: The map shows which areas of California see reductions in $NO_x$ damages under a high carbon price scenario compared to a no carbon price scenario, where darker green indicates larger quantities of avoided damages. These damages are overlaid on the census tracts identified in California as Disadvantaged Communities, indicated with black outline. Disadvantaged communities are defined as those census tracts in California in the top 25 percent of CalEnviroScreen3.0 scores, as well as census tracts that lack an overall CalEnviroScreen score but are in the highest 5 percent of pollution burdens scores.

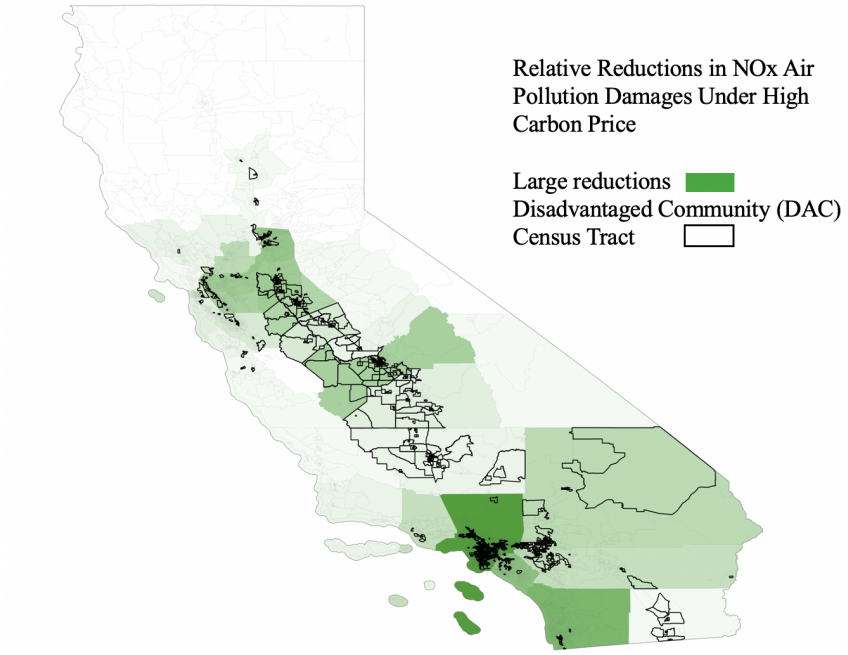


Figure 11: This figure plots the gross savings from production costs avoided from investment compared to no investment across four carbon price scenarios. Each point along the x-axis represents a unique investment scenario modeled.

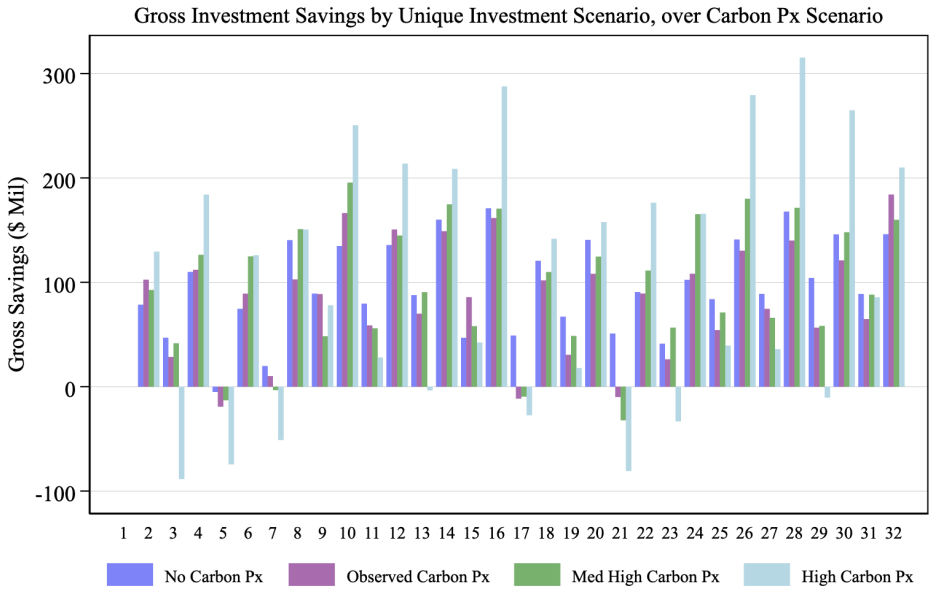



Figure 12: This figure plots gross savings in production costs with investment compared to no investment by the sum of the market shares of invest firms. Each marker represents a unique investment scenario in a unique input cost state. For each scenario, the y-axis represents the sum of the market shares among firms investing in that scenario.

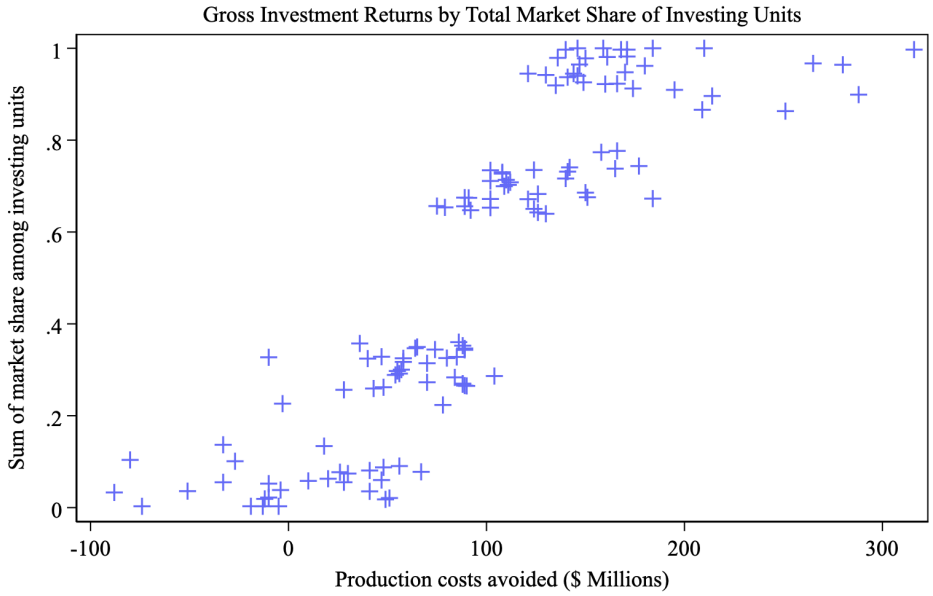
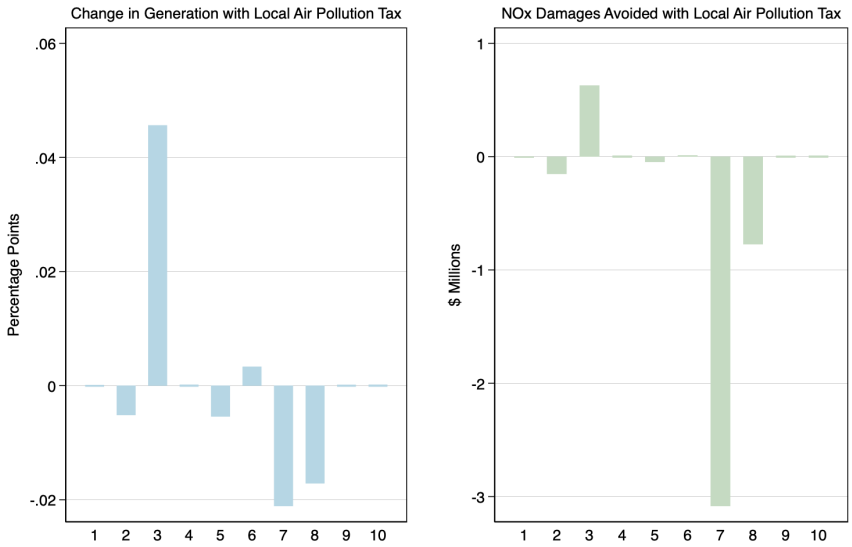


Figure 13: The figure on the left shows how the scenario with a local air pollutant tax in addition to the carbon policy changes market shares across unit types compared to a carbon policy alone. The figure on the right shows the corresponding changes in NOx damages.




# References


**Banzhaf, H Spencer, and Randall P Walsh.** 2013. "Segregation and Tiebout sorting: The link between place-based investments and neighborhood tipping." *Journal of Urban Economics*, 74: 83–98.

**Banzhaf, H. Spencer, ed.** 2012. *The Political Economy of Environmental Justice.* . 1 ed., Stanford University Press.

**Banzhaf, Spencer, Lala Ma, and Christopher Timmins.** 2019. "Environmental Justice: The Economics of Race, Place, and Pollution." *The Journal of Economic Perspectives*, 33(1): 185–208.

**Borenstein, Severin.** 2012. "The Redistributional Impact on Nonlinear Electricity Pricing." *American Economic Journal: Economic Policy*, 4(3): 56–90.

**Borenstein, Severin.** 2013. "Effective and Equitable Adoption of Opt-In Residential Dynamic Electricity Pricing." *Review of Industrial Organization*, 42(2): 127–160.

**Borenstein, Severin.** 2017. "Private net benefits of residential solar PV: The role of electricity tariffs, tax incentives, and rebates." *Journal of the Association of Environmental and Resource Economists*, 4(S1): S85–S122.

**Borenstein, Severin, and Lucas Davis.** 2016. "The distributional effects of U.S. clean energy tax credits." *NBER Tax Policy and the Economy*, 30(1): 191–234.

**Bushnell, James.** 2021. "Can Climate Efforts Be the "Everything Policy Store"?" *Energy Institute at Haas: Energy Institute Blog.*

**CAISO.** 2014, 2015. "Quarterly Report on Market Issues and Performance." *Department of Market Monitoring.*

**CARB, California Air Resources Board.** 2010, 2012, 2013. "Cap-and-Trade Regulation." *https://ww2.arb.ca.gov/our-work/programs/cap-and-trade-program/cap-and-trade-regulation*, Accessed: 2021-12-9.

**Clay, Karen, Akshaya Jha, Randy N.Z. Muller, and Walsh.** 2019. "The External Costs of Shipping Petroleum Products by Pipeline and Rail: Evidence of Shipments of Crude Oil from North Dakota." *The Energy Journal*, 40(1).

**Colmer, Jonathan, Ian Hardman, Jay Shimshack, and John Voorheis.** 2020. "Disparities in PM2. 5 air pollution in the United States." *Science*, 369(6503): 575–578.

**Cullen, Joseph.** 2013. "Measuring the Environmental Benefits of Wind-generated Electricity." *American Economic Journal: Economic Policy*, 5(4): 107–33.

**Cullen, Joseph.** 2015. "Dynamic Response to Environmental Regulation in the Electricity Industry." *Working Paper.*

**Cullen, Joseph, and Stanley Reynolds.** 2017. "Market Dynamics and Investment in the Electricity Sector." *Working paper.*

**Cushing, Lara, Dan Blaustein-Rejto, Madeline Wander, Manuel Pastor, James Sadd, Allen Zhu, and Rachel Morello-Frosch.** 2018. "Carbon trading, co-pollutants, and environmental equity: Evidence from California's cap-and-trade program (2011–2015)." *PLoS medicine*, 15(7).

**Davenport, Coral.** 2020. "Claims of 'Bleak' Environmental Justice Record Appeart to Fell a Biden Favorite." *New York Times.*

**Ed., Chris Costello.** 2019. "Distributional Effects of Environmental Markets: Insights and Solutions from Economics."

**EIA, The U.S. Energy Information Administration.** 2017. "Today in Energy: Natural gas-fired electricity conversion grows as goal remains stable." *https://www.eia.gov/todayinenergy/detail.php?id=32572*, Accessed: 2020-04-10.

**EPA, The Environmental Protection Agency.** 2017. "The Social Cost of Carbon: Estimating the Benefits of Reducing Greenhouse Gas Emissions." *https://19january2017snapshot.epa.gov/climatechange/social-cost-carbon.html*, *Accessed* : 22*October*2020.

**Fowlie, Meredith, Stephen Holland, and Erin Mansur.** 2012. "What do emissions market deliver and to whom? Evidence from Southern California NOx trading program." *American Economic Review*, 102(2): 965–993.

**Fullerton, Don, Garth Heutel, and Gilbert E. Metcalf.** 2012. "Does the Indexing of Government Transfers Make Carbon Pricing Progressive?" *American Journal of Agricultural Economics*, 94(2).

**Gilmore, Elisabeth A, Jinhyok Heo, Nicholas Z Muller, Christopher W Tessum, Jason D Hill, Julian D Marshall, and Peter J Adams.** 2019. "An inter-comparison of the social costs of air quality from reduced-complexity models." *Environmental Research Letters*, 14(7): 074016.

**Goulder, Lawrence H., Marc A. C. Hafstead, Gyurim Kim, and Xianling Long.** 2018. "The Incidence of Carbon Taxes in U.S. Manufacturing: Lessons from Energy Cost Pass-Through." *Working Paper*.

**Grainger, Corbett, and Thanicha Ruangmas.** 2018. "Who wins from emissions trading? Evidence from California." *Environmental and Resource Economics*, 71(3): 703–727.

**Hernandez-Cortez, and Kyle Meng.** 2019. "Do Environmental Markets Cause Environmental Justice?" *Working paper*.

**Hopenhayn, Hugo A.** 1990. "Industry equilibrium dynamics: A general competitive theory."

**Hopenhayn, Hugo A.** 1992. "Entry, Exit, and Firm Dynamics in Long Run Equilibrium." *Econometrica: Journal of the Econometric Society*, 1127–1150.

**Jha, Akshaya, and Gordon Leslie.** 2021. "Start-up Costs and Market Power: Lessons from the Renewable Energy Transition." *Available at SSRN 3603627*.

**Jovanovic, Boyan.** 1982. "Selection and the Evolution of Industry." *Econometrica*, 50(3): 649–670.

**Knittel, Christopher, and Ryan Sandler.** 2018. "The Welfare Impact of Second-Best Uniform-Pigouvian Taxation: Evidence from Transportation." *American Economic Journal*, 10(4): 211–242.

**Kumar, N., P. Besuner, S. Lefton, D. Agan, and D. Hilleman.** 2012. "Power Plant Cycling Costs." *National Renewable Energy Laboratory*.

**Kuminoff, Nicolai V, V Kerry Smith, and Christopher Timmins.** 2013. "The new economics of equilibrium sorting and policy evaluation using housing markets." *Journal of economic literature*, 51(4): 1007–62.

**Lucas, Robert E., and Edward C. Prescott.** 1971. "Investment Under Uncertainty." *Econometrica*, 39(5): 659–681.

**Mansur, Erin T.** 2008. "Measuring Welfare in Restructured Electricity Markets." *Review of Economics and Statistics*, 90(2): 369–386.

**Muller, Nicholas Z., and Robert Mendelsohn.** 2007. "Measuring the damages of air pollution in the United States." *Journal of Environmental Economics and Management*, 54(1): 1–14.

**Muller, Nicholas Z, Robert Mendelsohn, and William Nordhaus.** 2011. "Environmental accounting for pollution in the United States economy." *American Economic Review*, 101(5): 1649–75.

**OEHHA, and CalEPA.** 2017. "CalEnviroScreen 3.0 Factsheet." *CA.gov*.

**Pauer, Stefan U.** 2018. "Including electricity imports in California's cap-and-trade program: A case study of a border carbon adjustment in practice." *The Electricity Journal*, 31(10): 39–45.

**Reguant, Mar.** 2014. "Complementary bidding mechanisms and startup costs in electricity markets." *The Review of Economic Studies*, 81(4): 1708–1742.

**Schmalensee, Richard.** 2022. "Competitive Energy Storage and the Duck Curve." *The Energy Journal*, 43(2).

**Shapiro, Joseph S, and Reed Walker.** 2021. "Where is Pollution Moving? Environmental Markets and Environmental Justice." National Bureau of Economic Research Working Paper 28389.

**U.S. Energy Information Administration, EIA.** 2014, 2015, 2016, 2017. "State Electricity Profiles." *https://www.eia.gov/electricity/data.php*, Accessed: 2021-12-09.

**Walch, Ryan T.** 2018. "The Effect of California's Carbon Cap and Trade Program on Co-Pollutants and Environmental Justice: Evidence from the Electricity Sector." *Environment, PM2 5*, 440–48.

**Weber, Paige.** 2021. "Making Clean Firms Cleaner: Targeting Environmental Regulation to Maximize Returns." 111.

**Wolak, Frank, Severin Borenstein, and James B. Bushnell.** 2000. "Measuring Market Inefficiencies in California's Restructured Wholesale Electricity Market." *American Economic Review*, 5(92): 1376–1405.

# 8 For Online Publication: Appendix

## 8.1 Simulation Procedure

This section outlines the simulation procedure used to generate market outcomes. Simulation uses the recovered policy functions $\sigma^*(\cdot)$ to forward simulate market outcomes over one calendar quarter. Simulation of market outcomes occurs at several points in the estimation process, including as part of the start-up cost estimation, investment cost estimation, and to simulate market outcomes. Here is it discussed assuming one has already recovered the optimal efficiency investment policy, which sets the optimal heat rate vector, $\omega^*$.

### Simulation Steps

1. For a given initial state, set t=0, and draw demand shock $\eta_t$.
2. Let $\omega = \omega^*$ and let $c$ = empirically observed $c$ in 2015 Q1.
3. Use $\sigma^*(\eta_t, \mathbf{a_0}, c, h_t|\omega)$ to find the optimal dispatch for this state $\mathbf{q_t^*}$.
4. Draw a 10 by 1 vector $\mu$ from extreme value distribution with location parameters estimated from dispatch data; add $\mu$ to the dispatch that occurs in this state.
5. Simulate $\eta_{t+1}$ as the next step of a Markov chain with transition matrix $F\eta$.[44]
6. Update the elements $a_{t+1} \in \mathbf{a_{t+1}} = 1$ when elements $q_t \in \mathbf{q_t} > 0$ and 0 otherwise.
7. Update $h_{t+1} = h_t + 1 - \mathbb{1}(h_t = 24) * 24$.
8. Update $\mathbf{q_{t+1}} = \sigma(\eta_{t+1}, \mathbf{a_{t+1}}, ic, h_{t+1})$.
9. Update $t = t + 1$ and repeat until $t = 7160$.

### Simulation of Production Costs for Three-year Market Outcomes

For the simulations that calculate production costs over three years, the procedure above is used to generate quarterly outcomes. For each outcome, total production costs are calculated as the sum of marginal costs and start-up costs based on the firms that are dispatched in

[44] $F\eta$ is estimated from the empirical distribution of demand shocks. The Markov chain simulation proceeds by drawing a random variable $\eta$ with the distribution according to the row in $F\eta$ corresponding to $\eta_t$. Then the discrete inverse transform method is used to simulate $\eta_{t+1}$, which involves drawing a random variable $u$ from a uniform distribution and comparing to the probability entries in row $\eta_t$ of $F\eta$.

each hourly period. The production costs are then stacked over three years, using a quarterly discount rate of 0.0034, which corresponds to an annual interest rate of 1.34 percent. The policy functions are recovered for four different input cost scenarios corresponding to average fuel costs plus no carbon price, a carbon price equal to the average price observed 2012 - 2015, and a carbon price equivalent to a \$42 and \$123 social cost of carbon; in other words setting the permit price $\tau = \{0, 13, 42, 123\}$ in dollar per ton $CO_2e$. The policy scenario-specific estimates of production costs use the optimal dispatch decisions indicated by the recovered dispatch policy function $\sigma^*(\cdot|c)$ for the respective $c$ scenario.

**Simulation and Estimation of $\mathbf{NO_x}$ Damages**

To connect market outcomes to damages to human health, the $NO_x$ damages associated with the simulated market outcomes are estimated.[45] In the simulation above, the number and type of firms that are dispatched in each hour are recorded, $y_t^{sim}$. Each of the 10 firm types represent a set of firms for which I observe each of their locations in longitude and latitude, the $NO_x$ emissions intensities from the firm's continuous emissions monitors (CEMS), and the local marginal damages of air pollutants based on AP3 estimates for the firm's county. I allocate the total production assigned to the firm type in the simulation across the firms of that firm type equally. That is, if there are seven firms in type one, and type one is dispatched to produce 700 MWh of electricity in a given hour, each firm is allocated 100 MWh of production in that hour. Then I multiply each firm's allocated production by its emissions intensity for $NO_x$. I estimate damages to human health by multiplying the firm's $NO_x$ contribution by AP3's estimates of damages from pollution from that firm in any other county. Total damages are the sum of the firm's pollution to all other counties.

## 8.2 Estimation of Investment Costs

I estimate investment costs using a simulated method of moments approach (SMM) as discussed. First, I recover the policy functions for production across $J = 32$ investment scenarios. Then, I use the procedure described in A.1 to forward simulate market outcomes, and I calculate the total gross production costs associated with three years of hourly market outcomes for each scenario, $V^j$. Next, I draw an initial investment cost vector, $\xi^0 = \{\gamma^0, \alpha^0\}$, where $\gamma$ and $\alpha$ are as in equation 2. I select the optimal investment policy based on the simulated production costs, $V^j$, and the investment costs associated with each scenario $\Gamma(\mathbf{j}, \mathbf{v}, \xi)$:

[45] As these firms emit small amounts of $SO_2$, the $SO_2$ results are not reviewed.

$$\mathbf{j}^*(\xi^0) = \arg\max_{j \in J}(V^j + \Gamma(\mathbf{j}, \mathbf{v}, \xi^0). \tag{14}$$

As in the cost minimization problem discussed in the text, $\mathbf{j}$ and $\mathbf{v}$ refer the vector of investment decisions and cost shocks for each firm type, respectively. The investment costs $\Gamma(\mathbf{j}, \mathbf{v}, \xi)$ are calculated as $\sum_{i=1}^{N} \gamma j_i^{1/\alpha} + v_i$, where $j_i \in \{0, 1\}$; $v_i$ is an investment cost shock drawn from a normal distribution with mean and standard error estimated from the sample of investment costs observed in the SNL Energy Platform data.

The SNL Energy Platform has data on gross capital and fixed production costs for a subset of power plants. This subset of power plants corresponds to 14 (28 percent) of the firms flagged as investing based on the criteria outlined in Appendix 8.5. I use these data together with the heat rate improvements observed for these firms over 2012 to 2015 to construct an estimate of mean investment costs per percent heat rate reduction of \$72.2 million, with a standard error of \$49.7 million.

Next, I use the data to estimate the probability of investment across $c$ different firm investment types.[46] I use this probability to simulate $S$ investment decisions, corresponding to $c$-length vectors capturing investment decisions for each firm type. $\mathbf{j_{sim}}$ denotes the matrix with $c$ rows and $S$ columns of simulated moments. $g(\cdot, \xi^0) = (\mathbf{j_{sim}} - \mathbf{j}^*(\xi^0))^2$, a matrix with entries corresponding to the squared deviations from the simulated investment moments and the investment choice made by solving equation 14 given investment costs $\xi 0$. Denote $M = c * S$, which corresponds to the total number of moments. I reshape $g(\cdot, \xi^0)$ into a $M$-sized vector of moments, and estimate $\hat{\xi}$:

$$\begin{aligned} Q(\gamma) &= g(\cdot, \gamma)' \hat{W} g(\cdot, \gamma) \\ \hat{\gamma} &= \arg\min_{\gamma \in \Theta} Q(\gamma), \end{aligned} \tag{15}$$

where $\Theta$ is the set of positive real numbers, and $\hat{W}$ is estimated as $(g(\hat{\gamma})g(\hat{\gamma})')^{-1}$.

I use a bootstrap approach to estimate standard errors. Specifically, I estimate the parameters 500 times, each time taking a random draw of units, and then I calculate the standard error of these estimates. Investment-conditional choice probabilities within an investment firm type are developed by averaging investment outcomes across units within a firm type, conditional on input cost state. In taking random draws for the bootstrap

[46] The use of subscript $c$ here instead of $n$, which in the text corresponds to firm types, is to distinguish between firm investment type groups and firm type groups. The estimation approach collapses the 10 firm type groups to five firm investment type groups for the purposes of investment decisions for computational speed.

estimation, I select ten random units within each firm type and then calculate the investment conditional choice probabilities among those firms.

## 8.3 Estimation of Start-up Costs

The estimation of start-up costs compares production (dispatch) decisions implied by the model for a given start-up cost $\kappa^0$ to empirically observed dispatch. The parameter estimate, $\hat{\kappa}$, is a vector of firm type-specific start-up costs. The simulation procedure in section 8.1 uses the recovered policy function for dispatch $\sigma^*(\cdot)$ to simulate market outcomes over the Q1 2012 period using 2012 average heat rates. Let the $N$-length vector $\mathbf{q}^*$ capture the outcomes for all firm types where $q_n^*$ is the dispatch outcome for a single firm type $n$.

The empirical counterparts are assembled by categorizing each period in the data by the discretized state variables used in the model which include demand shock, lagged operating state, input cost, and hour: $s = \{\eta, \mathbf{a}, h\}$. In constructing the set of moments, dispatch outcomes are not observed empirically for all states in the model. Further, to simplify the linking of lagged states in the model to their empirical counterparts, the moments used are for states where all firms were either on or off in the last period. Denote $S$ as the number of states used for moments. I assemble $N$-length vectors corresponding to empirically observed dispatch by firm type in each state, $\mathbf{q}^{\mathbf{e}}(s)$. I construct a $S$-length vector of moments $g(s, \kappa^0) = \sum_{i=1}^{N}(\mathbf{q}^*(s, \kappa^0) - \mathbf{q}^{\mathbf{e}}(s))^2$, that is, the sum of deviations for all firm types in a given state $s$. I estimate $\hat{\kappa}$ as:

$$\begin{aligned} Z(\kappa) &= g(s, \kappa)' \hat{W} g(s, \kappa) \\ \hat{\kappa} &= \arg\min_{\kappa \in \varkappa} Z(\kappa), \end{aligned} \tag{16}$$

where $\varkappa$ is the set of positive real numbers, and $\hat{W}$ is estimated as $(g(s, \hat{\kappa}) g(s, \hat{\kappa})')^{-1}$.

## 8.4 Unit Type Groups

This section explains the process to group firms into firm types. The goal is to group the firms into a computationally tractable number of firm types, where groups are determined by firm characteristics that are relevant to the cost minimization problem. The cost minimization problem is solved using a vector of firm sizes and efficiencies as inputs. Accordingly, I use k-means to cluster the firms that provide non-zero generation to California in 2015 into groups according to firm size and efficiency (heat rate). I perform clustering with alternative

numbers of groups, from 0 to 20, and then examine the tradeoff between the number of groups and the amount of variation explained by the groups. Intuitively, explained variation should be weakly increasing in the number of groups.

The figure below demonstrates this tradeoff across several metrics. The graph on the top left compares within sum of squares (WSS) across different number of groups $k$, and the top right compares log(WSS) across groups. The graph on the bottom left is the coefficient $\eta^2$, which calculates the proportional reduction in WSS that each $k$ provides, compared to the total sum of squares (TSS):

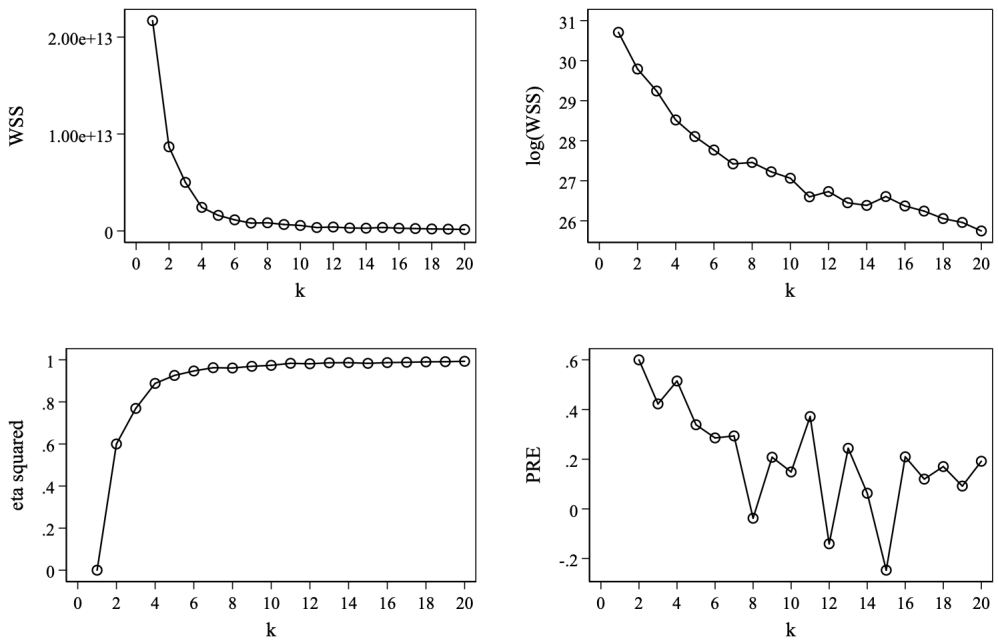
$$\eta_k^2 = 1 - \frac{WSS(k)}{WSS(1)} = \frac{WSS(k)}{TSS} \qquad \forall k \in K. \tag{17}$$

The graph on the bottom right shows the proportional reduction of error (PRE) coefficient, which is defined:

$$PRE_k = \frac{WSS(k-1) - WSS(k)}{WSS(k-1)} \qquad \forall k \geq 2. \tag{18}$$

The WSS plot shows a drop in WSS up to around five groups, with little additional WSS reduction thereafter. The log(WSS) plot shows a steady additional reduction as $k$ increases, with a kink at $k = 10$. The $\eta^2$ plot shows little additional proportional reduction in WSS to TSS after around $k = 6$. Finally, the PRE plot shows that moving from 10 to 11 groups does not explain additional variation, where the negative value indicates that $WSS(11) \geq WSS(10)$. Given this and the kink at $k = 10$ in the log(WSS) plot, I categorize the firms into 10 different firm types based on size and heat rate.

Figure 14: Performance of K-means clustering by number of groups




To apply the model, I calculate the mean operating capacity (MW) and heat rate for each firm type. For the dispatch model, I also need to characterize the minimum and maximum operating levels for each of the firms. To identify theses levels, I model firm generation as a bimodal distribution, and I use a finite mixture model to identify the two means of generation levels, which provides estimates of minimum and maximum operating levels. I also use a more straightforward approach, which assumes that minimum and maximum operating levels are equal to 0.75 and 1.0 times mean capacity (MW) of the firm type, respectively. The included simulation results use this later approach.

## 8.5 Identifying Investment and Investment Levels

This section reviews the process to identify investment in the data. The data provide two measures of firm heat rates, which measure efficiency. One measure is a monthly self-reported heat rate, provided by the firm pursuant to federal reporting requirements. The second measure I call "inferred heat rate," which is a calculation of fuel inputs in one hour, divided by electricity produced in that hour. I assume that heat rate improvements (reduction in heat rate) as indicated by both reported and inferred measures decreasing with persistence provide evidence of an efficiency investment decision. To identify cases of persistent heat rate improvements, I first calculate yearly average heat rates, separately based on reported

and inferred heat rate data. Monthly average heat rates are also calculated for the inferred heat rate measure. Reported heat rates are already at the monthly level, so they do not need to be averaged; missing reported monthly values are filled in based on the last month of available reported data.

Next, I calculate annual heat rate differences by subtracting average annual heat rate in year $t$ from average annual heat rate in year $t+1$, separately for inferred and reported measures of heat rate. Then I identify which firms decreased their heat rates from year $t$ to $t+1$, based on both the inferred and reported heat rate metrics, and I record the year of investment. This approach of calculating heat rate improvements from annual averages identifies evidence of endogenous heat rate improvements while allowing for month-to-month fixed effects that may have exogenous impacts on heat rate, for example, due to seasonal changes in dispatch and weather variation. I find that 50 firms observed in the 2012 - 2015 data set decreased their heat rate by both measures in this annual assessment.

Finally, I identify the percent heat rate improvement among the firms flagged as investing in order to group investments into bins of heat rate improvement. For each firm that invested, I calculate the percent change in heat rate as a result of investment. To apply the model parsimoniously, I collapse the investment decisions into a binary choice to invest or not, where investment occurs around the median investment level, 1.5 percent.

Table 6: Investment Cost Parameter

| Source | Mean ($) | Std.Erro ($) |
|---|---|---|
| SNL Data | 72.2e+06 | 49.7e+06 |
| Simulated Method of Moments $\gamma$ | 1.0449e+06 | 6.9358e+03 |
| Simulated Method of Moments $\alpha$ | 0.9875 | .0034 |

Standard errors for SMM estimation approach are calculated using a bootstrap approach: the standard error is calculated from the parameter estimates using 500 random draws of data samples.

## 8.6 Demand Process and Prices

This appendix provides additional detail and results regarding the construction of the transition probability matrices. First, I show the results from a modified version of equation 8.6 now estimating the model over all hour of the day including hour fixed effects.

$$\mathbb{1}[\eta_h = high] = \alpha + \xi \mathbb{1}[\eta_{h-1} = high] + \epsilon \tag{19}$$

The goal of this specification is to demonstrate that this parsimonious model of demand shocks explains a significant amount of variation in hourly demand states. Indeed, as Table 7 shows, this specification explains over 70 percent of the variation in high and low demand shock states.

Table 7: Demand State Process

| | High demand state |
|---|---|
| High demand state last hour | 0.84*** |
| | (0.00) |
| R-squared | 0.708 |
| N | 26279 |

For the simulation, the level of demand in low and high demand states is set to the 25th and 75th percentiles of hourly residual demand over the period 2013 - 2015, shown in Figure 15 below. Figure 16 then shows average hourly prices during this period to illustrate the correspondence between price and demand fluctuations over the day.

Figure 15: Residual hourly demand in CA, 2013 - 2015

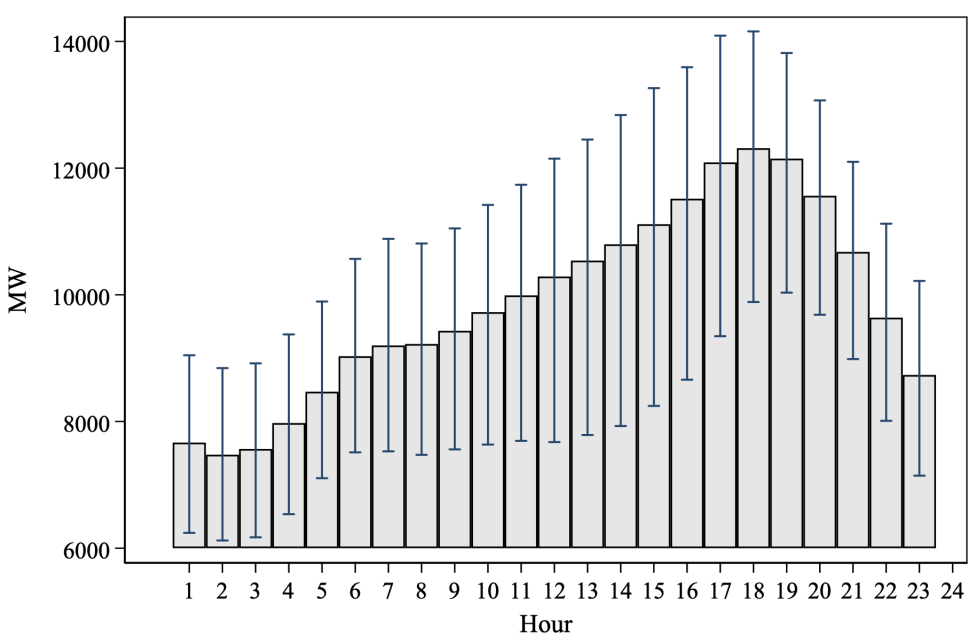
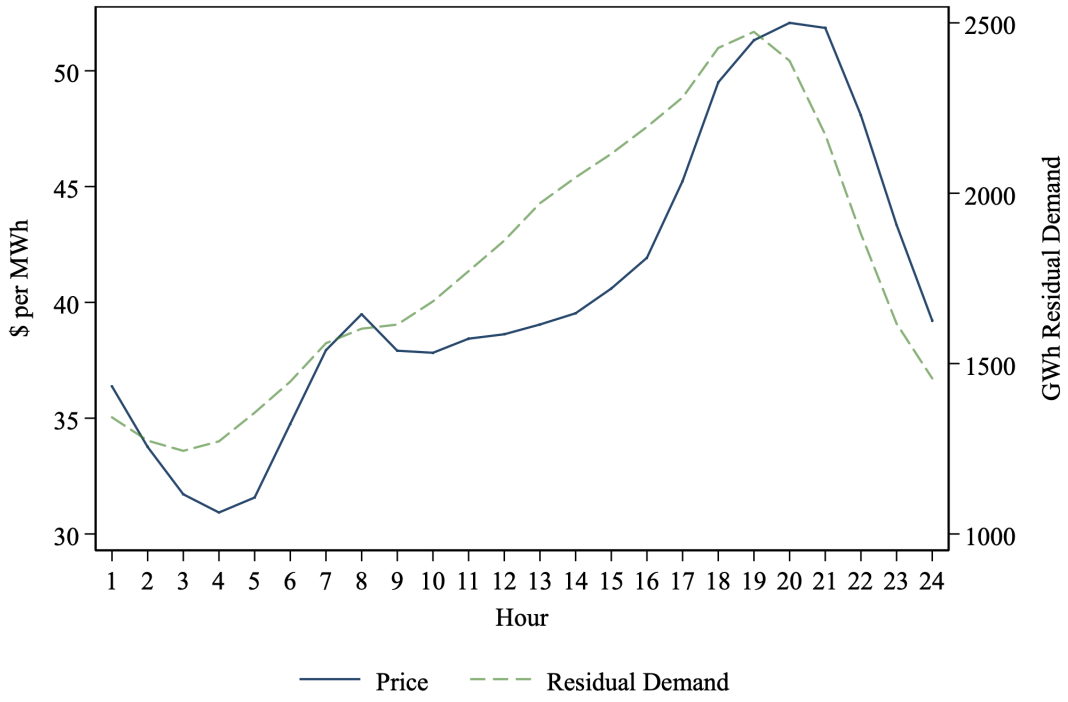


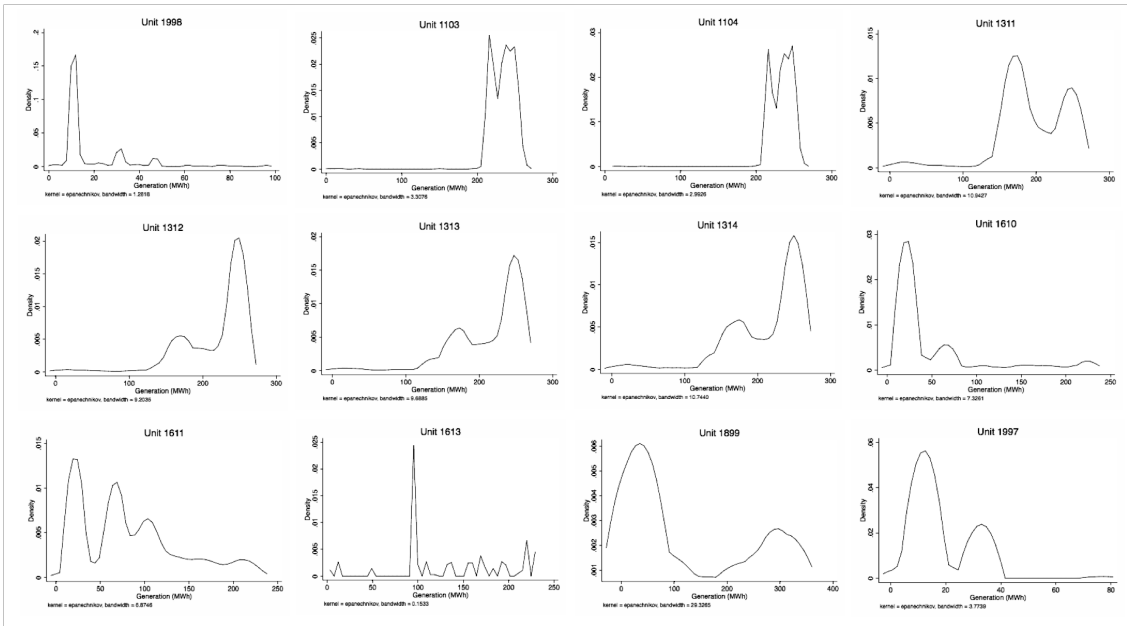
Figure 16: Average Hourly Electricity Prices and Residual Demand in CA, 2013 - 2015




The estimation procedure requires constructing hourly transition probabilities for the demand shocks. To do so, equation is estimated for each hour of the day and estimates of $\xi$ and $1 - \xi$ are used as the probabilities that a high demand state in one hour leads to a high and low demand state in the next hour, respectively.

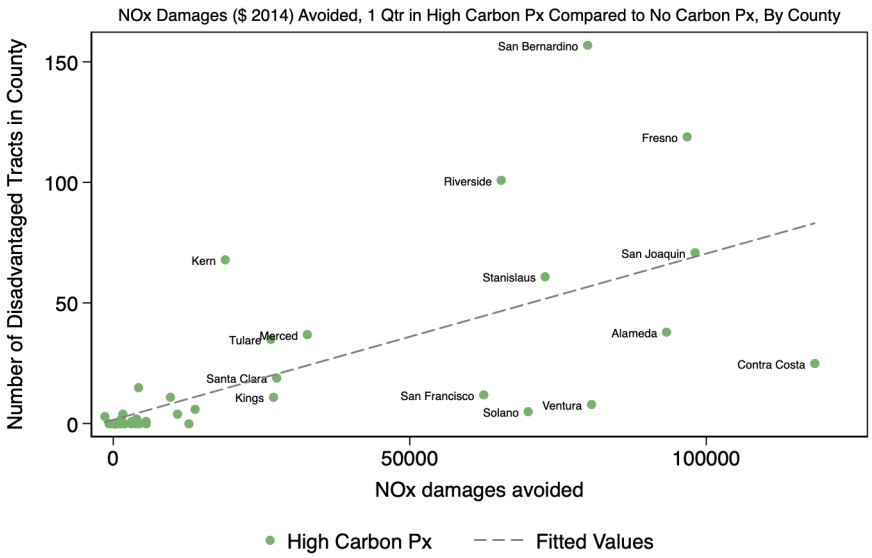
## 8.7 Production Levels

Figure 17: These histograms show frequencies of production levels across a sample of firms. The histograms illustrate the non-continuous nature of production in this setting, as firms appear to operate most frequently across a set of discrete quantities.



## 8.8 Damages Avoided by Number of Disadvantaged Counties

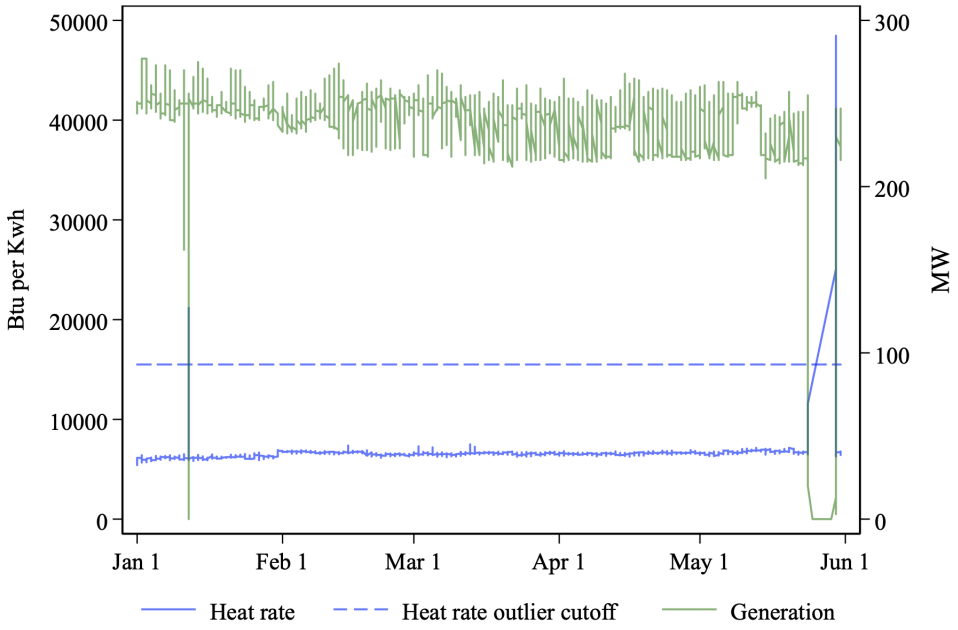
Figure 18: This graph shows the change in $NO_x$ damages within a county in the high carbon price scenario, compared to no carbon price scenario. The y-axis indicates the number of Census tracts within the county identified as Disadvantaged Communities (DACs), indicating community's Census Tract falls in the 25th percentile of highest pollution burdens.




## 8.9 Outliers in CEMS data for Combined Cycle firms

There are hours in which the inferred heat rate in the CEMS data from these firms increases dramatically, stemming from a sharp reduction in reported generation. Figure 19 below provides an example of this in the raw data.

Figure 19: Hourly heat rate and generation in raw CEMS data for sample combined cycle unit




Other empirical researchers have noted this, and attributed this to combined cycle units not reporting generation from both steam and gas turbines in some hours. To account for this behavior, generation is adjusted by a 1.5 multiplier in hours when heat rates exceed what would be reasonably expected (outliers). I define heat rate outliers as those above the 99th percentile for heat rates observed in the data for combined cycle units, which corresponds to heat rates above 15,204 Btu per kWh. For empirical context, the median heat rate observed in the data for combined cycle plants is 7,666 Btu per kWh, which is consistent with other engineering estimates; for example, the EIA estimate the average combined cycle heat rate in 2015 at 7,340 Btu per kWh (EIA 2017).

## 8.10 Equivalence Between the Firm Problem and the Cost Minimization Problem

The estimation approach used in this paper makes use of the equivalence between a competitive market equilibrium and the solution to a social planner's problem, a standard result in dynamic equilibrium models, such as in Lucas and Prescott (1971), Jovanovic (1982), Hopenhayn (1990) and Hopenhayn (1992). While these papers discuss a social planner's problem, I use the term cost minimization problem to denote that the problem may not price all externalities. For example, the paper solves cost minimization problems

with and without pricing GHGs and local air pollution. Since not all externalities are priced in some of the problems, it is clearer to leave out the concept of a social planner.

Cullen and Reynolds (2017) prove that this equivalence demonstrated in earlier work in dynamic equilibrium modeling holds for their model of electricity production and investment. This proof is required as their setting includes several feature that are not all present in earlier proofs of this equivalence. In order for the equivalence to be established, an additional assumption is needed beyond earlier work, which is that firms are 'small' relative to aggregate production quantities. This is required because, as in my model, firm-specific states are not continuous and discontinuous supply functions would otherwise pose a problem in solving a planner's problem. However, under the assumption that firms are 'small' relative to aggregate production, discontinuous supply functions are smoothed out and the transition between aggregate states is continuous. With this assumption, Cullen and Reynolds (2017) demonstrate a market equilibrium exists in their setting, and the social planner's allocation of investment and production is profit maximizing for individual firms. Given the correspondence between the model in Cullen and Reynolds (2017) and my model, I refer to their proof to justify the estimation approach used here.