Low-Carbon Investment Incentives and Climate Policy*

Preliminary and Incomplete: Please do not cite or circulate

Sarah Armitage[1], Nathan Miller[2], Matthew Osborne[3], and Gretchen Sileo[4]

[1]Boston University
[2]Georgetown University
[3]University of Toronto
[4]Temple University

December 8, 2023

Abstract

Industrial decarbonization requires significant investments in abatement technology. Yet if firms have market power, their incentive to invest may deviate from the socially optimal level – even after carbon externalities have been appropriately priced – because they do not fully internalize the consumer gains from reduced prices or they face preemption motives in deterring rival firms. We explore the implications of market power for technology adoption and efficient policy in the context of the U.S. cement industry. We first estimate a dynamic model drawing on data over 1974-2019, where we endogenize firm decisions to adopt new technology or retire old technology. We then use counterfactual simulations to examine how firms respond alternative investment subsidies.

Keywords: decarbonization, market power, technology adoption, investment, climate change, cement

JEL Codes: L11, L52, Q55, Q58

1 Introduction

A defining characteristic of decarbonization in hard-to-abate industrial sectors such as cement and steel production is that this transition will require substantial sunk cost investment in new plants using low-carbon technologies and upgrades of existing plants to accommodate low-carbon technologies. In cement, for example, it is difficult to envision a largely decarbonized sector that continues to produce Portland cement and that does not make use of some form of carbon capture and storage.[1] Once captured, this CO_2 might be stored as part of broader "carbon management"

*For helpful suggestions, we thank Joseph Aldy, participants at NBER Industrial Decarbonization Pre-Conference, Gautam Gowrisankaran, Ashley Langer, and Tianshi Mu.

[1]Alternative decarbonization pathways would entail developing cement chemistries other than Portland cement, which today accounts for 95% of all cement production in the U.S. About 60% of the CO_2 associated with Portland cement production is a chemical byproduct from the calincation of limestone, which cannot be avoided through fuel switching, electrification, or process efficiencies. See Fennell et al., 2022.

ecosystems, such as using underground aquifers, or added to wet concrete itself. In either case, cement production facilities will need to be constructed or retrofitted to include carbon capture devices, both of which entail substantial sunk costs.

Designing and implementing climate policies to incentivize efficient sunk cost investment have proved challenging. The first-best policy for correcting the unpriced greenhouse gas externality is well understood to be a carbon tax equal to the social cost of carbon (or an equivalent emissions trading system). However, additional complexities arise in settings where firms are able to exercise market power, which includes most industrial sectors. Moreover, many recent policy discussions related to industrial decarbonization are focused not on direct carbon pricing but on various second-best industrial policies. While these alternative policy instruments vary in their specific structure, virtually all are aimed at incentivizing sunk cost investment in low-carbon production methods. Yet very little empirical research in economics has examined the potential impact of these second-best policies on incentives for low-carbon investment, particularly in settings with market power.

In this paper, we study the interactions between market power, lumpy technology adoption decisions, and investment subsidies, using the U.S. Portland cement market as our empirical case study. We explore how market structure influences the effectiveness of technology adoption subsidies; how subsidies might have affected the historical transition path for precalciner kiln adoption over the last 40 years; and how alternative subsidies might affect the transition path for future decarbonization strategies. Cement production is an ideal setting for studying these issues for two reasons. First, as researchers have documented in previous papers, many characteristics of the cement industry allow researchers to gain traction in empirical modeling, such as the fact that it is a homogeneous product or that long-term storage is generally not possible. Second, the industry experienced a significant technological transition from the 1970s to the 2010s, as plants upgraded to precalciner kilns. This technology was far more fuel efficient – 25-35% more efficient than the older kilns that they replaced – but costly to install (Macher et al., 2021).

To answer these questions, we develop and estimate a dynamic model of new technology adoption, production, and exit in the U.S. Portland cement market. Plant-level adoption and retirement decisions depend on their own existing capacity, competitor capacity, and exogenous state variables including coal prices and market demand. We estimate structural parameters for kiln upgrading and decommissioning costs using a two-step estimator, adapted for multi-stage decision-making (Seiler, 2013). We use a gradient boosting algorithm to adapt the static model of cement procurement and plant-level profits estimated in Miller et al. (2023) for dynamic estimation. We then simulate counterfactual scenarios with alternative technology adoption subsidies, using an equilibrium concept that both allows for regional cement markets with a large number of plants and that allows plants to internalize the impact of their decisions on the competitive environment (Gowrisankaran et al., 2022).

Section 2 presents institutional details about the Portland cement industry, including the history of precalciner kiln adoption, and describes potential future CO_2 abatement pathways. Section 3 describes our empirical model of new kiln adoption and old kiln retirement and estimation methods.

Section 4 describes counterfactual simulations, and Section 5 concludes.

2 The Portland Cement Industry

2.1 Background Facts

Portland cement is a finely ground dust that forms concrete when mixed with water and coarse aggregates such as sand and stone. The production of cement involves feeding limestone and other raw materials into rotary kilns that reach peak temperatures of 1400-1450° Celsius. The process produces CO_2 emissions for two main reasons. First, cement plants burn fossil fuels—mostly coal and natural gas—to heat the kilns. Second, the chemical transformation of limestone into cement releases CO_2 as a byproduct. Of these, the combustion of fossil fuels accounts for roughly one-third of emissions and the chemical transformation accounts for two-thirds (Glenk et al., 2023).

Kilns differ in their fuel efficiency. For most of the twentieth century, cement manufacturers relied on "wet" and "long dry" kilns that typically were 100 yards or longer in length.[2] An inefficiency with wet and long dry kilns is that they allow a significant amount of heat to escape the system, both because exhaust gases are not captured and because heat radiates from the kiln. Modern precalciner technology mitigates this inefficiency. The basic idea is that the raw materials can be preheated before they enter the kiln using exhaust gases and heat from a supplementary combustion chamber. As the raw materials then need less time in the kiln for the chemical reactions, the rotary kiln is shorter in length, and this in turn reduces inefficiency due to kiln radiation. A modern precalciner kiln is typically 25-40 yards in length and 25-35% more fuel efficient than a wet or long dry kiln. As a result, they also emit less CO_2 than wet and long dry kilns.

Part of our analysis examines how counterfactual changes in public policy would have affected the speed of the transition to modern precalciner kilns. Figure 1 shows how the transition played out in reality. We plot industry capacity (top panel) and the number of plants (bottom panel) over 1974-2019, both in total and separately for "old technology" (wet and long dry kilns) and "modern technology" (preheater and precalciner kilns). Over the sample period, total industry capacity increases by 20%, from 91 to 109 million metric tonnes, with old technology accounting for nearly all of this capacity at the beginning of the sample and modern technology accounting for nearly all of this capacity at the end. Over the same period, the number of plants falls by 45%, from 163 to 89. This incorporates 13 plants that were constructed during the sample period. As with capacity, nearly all plants use the old technology at the beginning of the sample and the new technology at the end. Modern kilns have far greater capacities than older kilns. Indeed, in 2019 the average annual capacity of a modern technology kiln was nearly double that of a old technology kiln.

The economics of the industry are reasonably straight-forward. Demand is pro-cyclical because cement is used for construction projects. Buyers are large construction firms and ready-mix

[2] Wet kilns process raw materials that are wet-ground into a slurry, whereas long dry kilns process raw materials that are dry-ground into a powder. Modern preheater and precalciner kilns use the dry process. Preheater kilns do not have the supplementary combustion chamber of precalciner kilns.

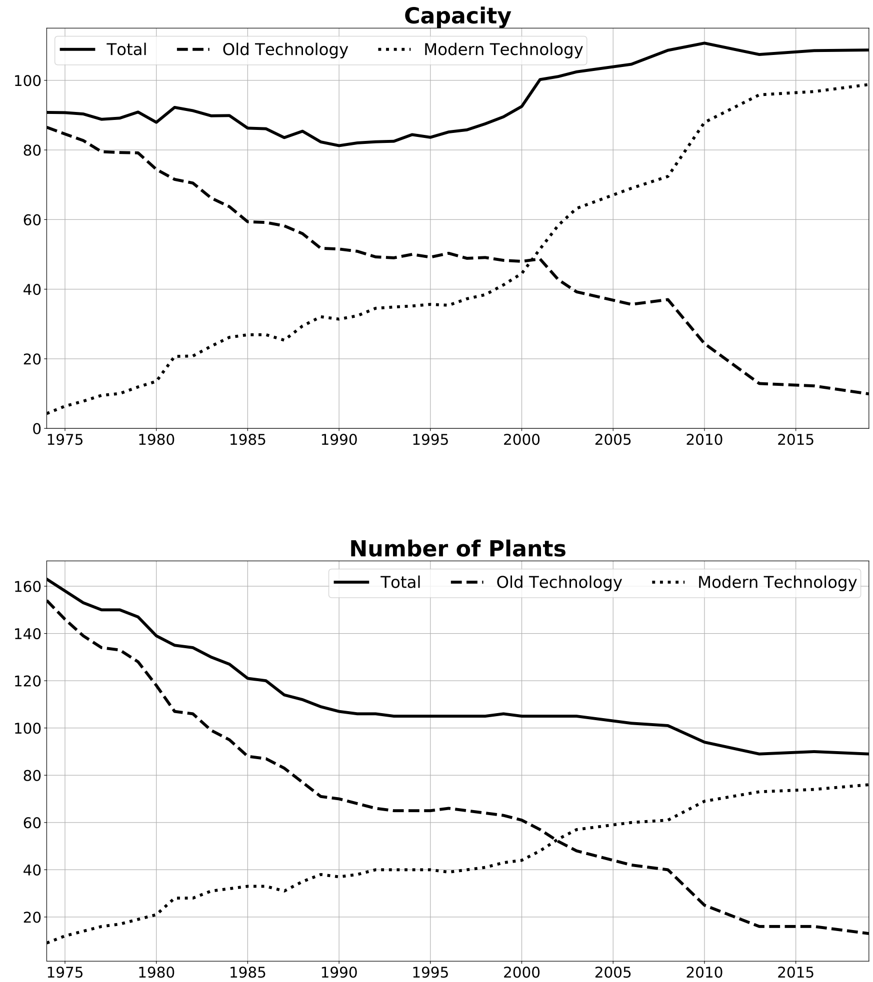

Figure 1: Industry Capacity and the Number of Cement Plants, 1974-2019

Notes: In the top panel, capacity is in millions of metric tonnes. We designate plants as using "Old Technology" if their least efficient kiln is a wet kiln or a long dry kiln, and as using "Modern Technology" if their least efficient kiln uses a precalciner or a preheater. Plants are excluded from the graphs if they are temporarily idled (e.g., due to maintenance or low demand). Data are from the *Plant Information Summary* of the Portland Cement Association.

concrete plants. Most transactions involve short-term contracts. From the perspective of buyers, cement plants are differentiated primarily by their location because transportation costs are a substantial portion of the overall purchasing expense. On the supply-side, the main variable costs are attributable to raw materials, fuel costs, electricity costs, labor, and kiln repair and maintenance (EPA (2009)). The limestone is usually obtained from a quarry adjacent to the plant, and most plants use coal or natural gas for their fuel. Plants run at full capacity except for a single period that is typically 4-6 weeks each year, during which kiln maintenance is conducted. The main way that plants adjust output is by shortening, lengthening, or skipping this maintenance period. Imports are used mainly when demand outstrips domestic capacity. Exports are negligible.

2.2 Abatement Technologies

From the perspective of climate change mitigation, cement is widely considered a "hard to abate" sector. The production of Portland cement yields CO_2 emissions both as a byproduct and from fuel consumption for high-temperature processes. Byproduct CO_2 emissions, accounting for approximately 60% of the total, result from the calcination of limestone to form lime, which is then mixed with sand and clay to form clinker, the key binding agent in cement (Fennell et al., 2022). High temperatures are needed to produce both limestone and clinker; a range of fuels are currently used, coal most commonly (*Mission Possible: Reaching Net-Zero Emissions from Harder-to-Abate Sectors by Midcentury, Sectoral Focus Cement*, 2019).

Decarbonization opportunities fall into several categories, with no single clear path to near full decarbonization. Fuel switching and eventual electrification of kilns can reduce emissions associated with high-temperature processes; existing processes would also benefit from further efficiency improvements, especially in older facilities (*The Net-Zero Industry Tracker: Cement Industry*, 2022; Czigler et al., 2020; *Concrete Future: The GCCA 2050 Cement and Concrete Industry Roadmap for Net Zero Concrete*, 2021). Other decarbonization opportunities exist from using clinker, cement, or concrete more efficiently, or from developing and commercializing new cement chemistries. The ratio of clinker to cement may be reduced by adding substitutes, though scalability may be a challenge as industries providing input materials, such as coal-based electricity generation or steelmaking, undergo their own decarbonization; natural clinker substitutes such as volcanic ash are less well studied (Czigler et al., 2020). Greater efficiency in building or infrastructure design could reduce the total volume of concrete per useful structure. Finally, increased circularity in the cement and concrete supply chain could further reduce emissions, ranging from re-use of non-hydrated cement to producing clinker from decommissioned concrete rather than limestone (*The Net-Zero Industry Tracker: Cement Industry*, 2022; *Mission Possible: Reaching Net-Zero Emissions from Harder-to-Abate Sectors by Midcentury, Sectoral Focus Cement*, 2019).

Perhaps the most challenging aspect of cement decarbonization is addressing the process emissions from the calcination of limestone. Other than novel cement chemistries, there are no meaningful alternatives to some form of carbon capture. The captured carbon may be stored as part of a broader CCUS ecosystem, or may be added to the concrete itself before it hardens, which may have beneficial strengthening properties. Technological innovation may result in higher concentrations of CO_2 in flue gas, reducing the cost of carbon capture. Several large cement manufacturing facilities have announced or undertaken carbon capture pilot projects, and startup companies are actively exploring technologies for adding captured CO2 to concrete (Hook and Dempsey, 2021).

Concrete is the most common building material in the world, used in public and private buildings and infrastructure such as roads, bridges, tunnels, ports, airports, and so forth. Estimates of the additional cost of fully decarbonizing cement production using carbon capture and storage are around $100/ton of concrete, which represents about 100% of the cost of concrete, 15% of the cost of constructing a building, or 1-3% of the typical property value (Fennell et al., 2022; *Mission Possible: Reaching Net-Zero Emissions from Harder-to-Abate Sectors by Midcentury, Sectoral Focus*

Cement, 2019).

2.3 Data

We rely on the data used by Miller et al. (2023). There are two main sources. The first is the *Minerals Yearbook*, an annual publication of the USGS that contains aggregated free-on-board prices, production, consumption, imports, and transportation methods. The USGS collects these data using an annual census of the industry. Response rates are typically well above 90%, and USGS imputes missing values based on other data and their institutional knowledge. The data are aggregated prior to publication in order to preserve confidentiality. As a result, most information is available only at the region-level or national-level. The second main source of data in the *Plant Information Summary* of the Portland Cement Association (PCA), which reports the location, owner, and primary fuel of each cement plant, as well as the age, capacity, and type (wet/dry/precalciner) of each kiln. The PCA collects these data using phone surveys of plants. Data are available annually over 1974-2003 and also for 2004, 2006, 2008, 2010, 2013, 2016, and 2019.[3] We ultimately use data over 1980-2019 in the estimation sample. We rely on a number of sources to impute plant-level and kiln-level data for years in which the *Plant Information Summary* is not published. We refer readers to Miller et al. (2023) for details on these data sets and on a number of auxiliary data sets used in the analysis.

2.4 Defining "Trade Flow" Markets

Transportation costs limit the distance at which cement can be shipped economically from plants to buyers. To simplify the model, we let dynamic competition play out within a number of exogenous "trade flow" markets. Our key geographic assumption is that cement plants make entry, exit, and investment decisions based on their state variables and the aggregated state variables of other plants located in the same market. Thus, plants do not consider the state variables of plants that are outside its market. The assumptions allow us to capture that the incentives of incumbents to adopt abatement technology depends largely on local demand and supply conditions. As the market boundaries place meaningful restrictions on behavior, we draw them with some care.[4]

Our approach, broadly speaking, is to choose markets that are consistent with trade flows in the industry. We start with short run pricing model of Miller et al. (2023), which quantifies equilibrium shipments from each plant to each buyer location (in practice, a county). The short run model is specified to generate realistic shipments. For example, it allows cement to be trucked directly from the plant to the buyer, or to be placed on a barge and shipped along the Mississippi River System. Miller et al. (2023) show that estimated parameters imply reasonable transportation costs, and that equilibrium trade flows in the model match a coarser set of trade flows that are observable in

[3]Most years are available at the Yale University library. We purchased the most recent books from the PCA.

[4]We do not require that short run price competition occurs only among plants in the same region, or that trade flows between regions cannot occur (Section 2.4). Thus, our assumptions are somewhat weaker than those of Ryan (2012) and Fowlie et al. (2016).

data. We provide more details on the short run model later in the paper (Section 3.2).

We then apply a divisive hierarchical clustering algorithm to split counties into groups based on the equilibrium plant-county shipments from the model in the year 1980. In practice, the algorithm starts with all the counties in a single group, and then splits the counties into smaller, heterogeneous groups continuously until every county is on its own. The trade flow markets that we use in our baseline analysis are defined as the groups that exist at the point that the algorithm has defined exactly 10 different groups. Because we specify that the algorithms split counties if the cement shipments they receive are sufficiently dissimilar, we end up with markets in which buyers purchase from similar suppliers in similar proportions.[5]

Figure 2 maps the trade flow markets that we obtain from the clustering analysis. The largest geographical market is the market that provides cement in the area surrounding the Mississippi River. Plants in this area directly compete with one another through barge shipments. The other regions represent contiguous areas where a higher percentage of deliveries are transported via truck. We also provide summary statistics on market concentration and technology adoption for these trade flow markets in 1980 and 2019, the final year in the data, in Table 1. Cluster 1 contains the largest number of plants and has the lowest market concentration. Clusters 9 and 10 are highly concentrated and serve smaller geographical areas. In 1980, few plants have adopted precalciner technology. As of 2019, more than half of every market is comprised of precalciner kilns.

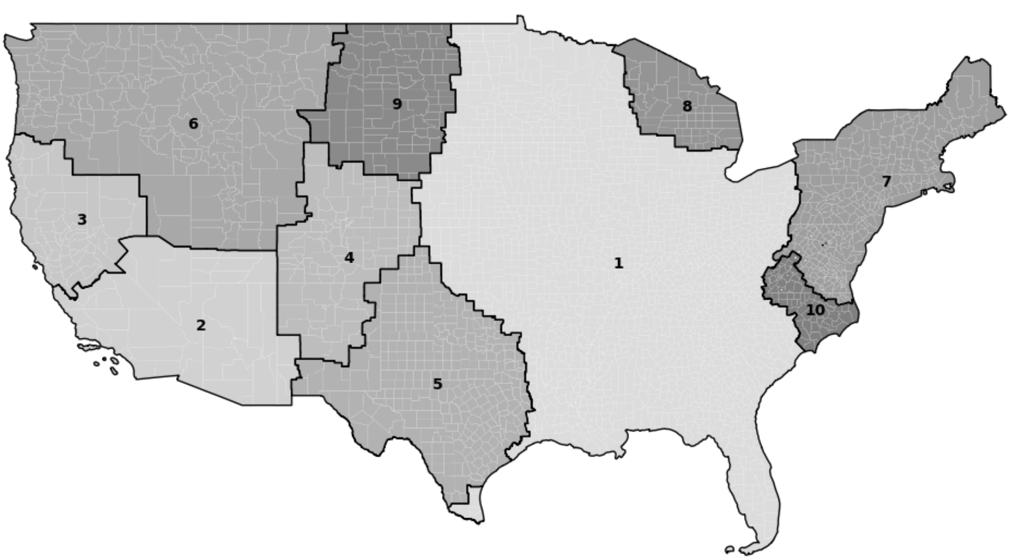

Figure 2: Trade Flow Markets

[5]We implement the clustering algorithm using the diana function in R. The algorithm itself is detailed in chapter 6 of Kaufman and Rousseeuw (1990). Appendix Figure C.1 provides a "dendrogram" that illustrates how the clustering algorithm splits counties in our application.

Table 1: Cluster-Level Statistics in 1980 and 2019

	1980			2019		
	HHI	Number of Firms	Precalciner Capacity Share	HHI	Number of Firms	Precalciner Capacity Share
Cluster 1	512	27	4.76%	1,136	13	78.73%
Cluster 2	2,563	6	7.50%	2,062	7	100.00%
Cluster 3	3,083	4	0.00%	6,916	2	80.87%
Cluster 4	4,489	3	22.50%	2,795	4	76.51%
Cluster 5	1,235	12	23.22%	1,818	8	96.49%
Cluster 6	2,549	6	0.00%	5,793	3	84.02%
Cluster 7	900	13	0.00%	2,489	5	67.91%
Cluster 8	5,396	6	19.25%	7,158	4	66.47%
Cluster 9	10,000	1	0.00%	10,000	1	100.00%
Cluster 10	5,320	2	0.00%	10,000	1	100.00%

Notes: County-level HHIs are obtained from the model. The 2010 *Horizontal Merger Guidelines* classify markets with HHI under 1500 as "unconcentrated," markets with HHI between 1500 and 2500 as "moderately concentrated," and markets with HHI above 2500 as "highly concentrated."

3 Empirical Model

Firms make annual decisions about the future capacity available for production at each plant. We assume that firms observe the competitive landscape within their cluster, including the past behavior of competitors, current fuel prices, and the quantity demanded by consumers. We construct a dynamic discrete choice model to describe the capacity decisions for cement plants in the industry. Capacity decisions take two forms: the adoption of new precalciner kilns (modern technology) and the retirement of wet and dry kilns (old technology). These decisions are made sequentially at the start of each period, and then firms competitively provide cement within a cluster. Below we provide details of the model.

3.1 Model

Each year t firms make capacity decisions for each of the plants that they operate in market m. We treat the model as a non-stationary dynamic problem from 1980-2030, and assume a stationary model thereafter. Plants are denoted as $j \in \mathbb{J}_{f_m(j)}$, where $f_m \in \mathbb{F}_{mt}$ is the set of firms that actively operate in the market as of time t. Firms enter into the period with knowledge of an aggregate measure of competitors' production capacity in old and new technology, $\mathbf{bp_{-fmt-1}} = (bp^{old}_{-fmt-1}, bp^{new}_{-fmt-1})$, as well as their own plant-level production capacity, $\mathbf{bp_{jmt-1}} = (bp^{old}_{jmt-1}, bp^{new}_{jmt-1})$. We denote the set of endogenous state variables as $\mathbf{x_{jmt}} = (\mathbf{bp_{jmt-1}}, \mathbf{bp_{-fmt-1}})$.

All market participants begin the period by observing updated information on the current period's demand (q_{mt}) and market price for coal (p^{coal}_{mt}). To simplify notation, we denote these exogenous state variables as $\mathbf{z_{mt}} = (q_{mt}, p^{coal}_{mt})$. Plants then receive choice-specific upgrade shocks

($\epsilon_{\mathbf{jmt}}^{\mathbf{upgrade}}$) and decide to pay an adoption cost UC to install a new precalciner kiln or to delay the choice until the following period. We denote the adoption decision as a binary variable, $u_{jmt} = \{0, 1\}$. The choice-specific upgrade shocks are *i.i.d.* type 1 extreme value and are private information. For any plant that upgrades to a new precalciner kiln, the size of the capacity increase is drawn from a pre-estimated distribution of new technology capacities, $h_u(bp_{jmt}^{upgrade}|u_{jmt})$. Based on firms' decisions, the aggregate state variable describing the amount of new technology capacity in the market updates.

Next, plants with old technology kilns make retirement decisions. Any plant that operates a kiln using old technology receives choice-specific shocks ($\epsilon_{\mathbf{jmt}}^{\mathbf{retire}}$) and makes decisions to retire anywhere from zero to the maximum number of old kilns currently operating at the plant. We denote the number of kilns retired in the period as $r_{jmt} = \{0, 1\}$. For each old kiln retired at a plant, the plant receives a scrap value of RV. The choice-specific upgrade shocks are *i.i.d.* type 1 extreme value and are private information. For any plant that retires an old technology kiln, the size of the capacity decrease is drawn from a pre-estimated distribution of old technology capacities, $h_r(bp_{jmt}^{old}|r_{jmt})$. Based on these capacity decisions, the aggregate state variable describing the amount of old technology capacity in the market updates.

After making capacity decisions, cement suppliers then compete with one another to provide cement to consumers. Variable profits at each plant are assumed to be functions of a plant's own old and new capacities, the old and new capacities of their competitors, the demand within the market and the market price for coal. Plants also pay a per-period fixed cost to operate old and new technology kilns. Plant-level net profits can therefore be expressed as $VP(\mathbf{x_{jmt}}, \mathbf{z_{mt}}) - FC(\mathbf{x_{jmt}}, \mathbf{z_{mt}})$.

Combining the components of the model, the maximization problem at the plant level can be expressed as follows:

$$\begin{aligned} \max_{\{u_{jmt}, r_{jmt}\}_{t=0}^{\infty}} \mathrm{E}\Bigg[\sum_{t=0}^{\infty} \beta^t \Big(& - UC u_{jmt} + \epsilon_{mt}^{upgrade}(u_{jmt}) \\ & + RV r_{jmt} + \epsilon_{mt}^{retire}(r_{jmt}) \\ & + \lambda VP(\mathbf{x_{jmt}}, \mathbf{z_{mt}}) - FC(\mathbf{x_{jmt}}, \mathbf{z_{mt}}) \Big) | \mathbf{x_{jmt-1}}, \mathbf{z_{mt}}, \epsilon_{\mathbf{mt}} \Bigg] \end{aligned} \tag{1}$$

3.2 Short Run Model of Procurement

Miller et al. (2023) estimate a model of short run competition in which buyers use second-score auctions to procure cement. As buyers typically pay the costs of transportation, the distance between the buyers and cement plants enters through the demand side of the model. The marginal cost functions on the supply-side of the model then are plant-specific and, to incorporate capacity constraints, they are assumed to be continuous and weakly upward-sloping. Given the functional forms used, there exists a Nash equilibrium in which all suppliers bid at marginal cost and those bids produce "market-clearing quantities" that give rise to the same marginal costs.

Summing across each county n, the equilibrium profit of a plant j in period t is

$$\pi_{jt}(\boldsymbol{X}_t, \boldsymbol{\theta}) = \sum_n \overline{p}_{jnt}(\boldsymbol{X}_t, \boldsymbol{\theta}) q_{jnt}(\boldsymbol{X}_t, \boldsymbol{\theta}) - \int_0^{Q_{jt}(\boldsymbol{X}_t, \boldsymbol{\theta})} c_{jt}(Q; \boldsymbol{X}_t, \boldsymbol{\theta}) dQ \tag{2}$$

where $\overline{p}_{jnt}(\boldsymbol{X}_t, \boldsymbol{\theta})$ and $q_{jnt}(\boldsymbol{X}_t, \boldsymbol{\theta})$ are the average price and total quantity that the plant obtains from a given county in equilibrium, $Q_{jt}(\boldsymbol{X}_t, \boldsymbol{\theta}) = \sum_n q_{jnt}(\boldsymbol{X}_t, \boldsymbol{\theta})$ is total plant-level quantity in equilibrium, and $c_{jt}(Q; \boldsymbol{X}_t, \boldsymbol{\theta})$ is the plant's marginal cost function. The functions depend on variables ($\boldsymbol{X}_t$) that summarize the distances between plants and counties, fuel prices, the capacities of plants, fuel prices, and other considerations. They also depend on parameters ($\boldsymbol{\theta}_t$) that are estimated in Miller et al. (2023). We detail the specifications of the functions in Appendix A.

We use this equilibrium profit function to model payoffs in the dynamic model. However, a direct application is infeasible for computational reasons. First, there are more variables in $\boldsymbol{X}_t$ than can accommodated in the state-space of the dynamic model. Second, computing market-clearing quantities for every year in the data typically takes about 15 seconds of computing time, and this is too slow given the dynamics require that payoffs be obtained repeatedly across many counterfactual states. Therefore, in our implementation of the dynamics, we use a simpler approximation of the equilibrium profit function. We describe the approximation in Section 3.3.1.

3.3 Estimation

We recover the upgrade cost, UC, scrap value from retiring an old kiln, RV, and the scale parameter on variable profits, λ, using the two-step forward-simulation approached developed in Bajari et al. (2007) (BBL). In the first step of the BBL approach, one estimates the firm policy functions, as well as the state transition functions, from what is observed in the data. In the second step, one recovers the values of the structural parameters that rationalize the value functions produced by the policy functions that are consistent with the actions taken by firms in the data.

Before estimating the policy functions and recovering the value functions, we first simplify the state space by estimating the parameters that describe the transition for the exogenous state variables. We use observations of market demand and coal prices over 1980-2019 to estimate the transition processes for these state variables. We assume both states evolve according to a deterministic market-level time-trend through 2030 and a stochastic, normally distributed shock. We then discretize the residuals from this process, simplifying the state space further. Additional details on the discretization and the parameters recovered in this process can be found in the Appendix.

In the first step, we separately estimate policy functions for the upgrade decisions and the retirement decisions, because we assume that these decisions occur in separate, sequential stages. The upgrade decision is estimated using a binary logit model. In the data, all plant upgrades include only one kiln, justifying the use of the binary model. The state variables we include in the upgrade function are the focal firm's boilerplate capacity for new and old technology, the aggregate boilerplate capacity of competitors, the firm's boilerplate capacity, and the demand and coal prices. We estimate two model specifications, presented in Table 2. Specification 1 aggregates competitor and own firm boilerplate capacity, while specification 2 splits these capacities into old and new technology. The coefficients on competitor new boilerplate capacity, as well as both own firm capacities, are not statistically significant at the 5% level in specification 2. Currently, we use specification 2 for simulation of the policy functions in recovering the structural parameters.

In these specifications, we find that plants are less likely to adopt a new precalciner kiln when they have higher capacity already – though much less likely to upgrade when that capacity is already precalciner technology. Plants are also less likely to upgrade when they face higher capacity from competitors. They are more likely to upgrade when the face higher demand, and more likely to upgrade – that is, invest in fuel saving technology – when coal prices are higher. Notably, these results are consistent with the findings of Macher et al. (2021), though we have used slightly different ways of constructing the data and estimating equations.

We model the retirement decision using an ordered logit. The options a plant can take here are to retire 0, 1, 2 3 or 4 kilns in a given period, conditional on the retirement decision. This is consistent with over 97% of retirement decisions that are observed in the data. The retirement decision is conditioned on the same set of state variables as the upgrade decision, except that we update the own and competitor new technology capacities to their end of year values, to be consistent with the timing assumptions made in Section 3.1: firms' upgrade decisions are revealed to all rivals prior to them making retirement decisions. Estimation results are presented in Table 3. Details about the conditional ordered logit specification are provided in the Appendix.

The results from the ordered logit model for retirement decisions suggest that plants are more likely to retire old technology kilns when coal prices are higher, when market demand is lower, when competitor capacity is higher, and when the plant has more new technology capacity.

In the second step, we use the estimated policy functions and state transitions to recover the structural parameters that rationalize firms' observed choices. To implement the second stage, we forward-simulate firm value functions at both the estimated policy functions, as well as a set of perturbed, suboptimal policy functions. The structural parameters are chosen to make the value functions conditional on the estimated policies larger than those produced by the perturbed policies. The basic steps we follow here are:

1. Draw $k = 1, ..., N_J$ vectors of plants, states, and perturbed policy functions from the empirical distribution of plants and states observed in each cluster. For the plant $j(k)$ in the plant-state-policy tuple k, we denote the estimated policy as $\hat{\sigma}_{j(k)}$, and the perturbed policy as $\hat{\sigma}'_{j(k)}$.

2. For each k, forward simulate the linear parts of the value functions (upgrade, retirement, variable profits, and the associated choice-specific errors) associated with the policies $\hat{\sigma}_{j(k)}$, and $\hat{\sigma}'_{j(k)}$. Note that when we perform this simulation, we keep the policy functions for plants $-j$ fixed at their estimated values, which we denote as $\hat{\boldsymbol{\sigma}}_{-\boldsymbol{j}}$. We denote the number of simulation draws used to construct the value functions as N_s.

Table 2: Dynamic Estimation: Upgrade Policy Function

	(1)	(2)
Plant New Tech. BP	-1.690^{***}	-1.918^{***}
	(0.409)	(0.428)
Plant Old Tech. BP	-0.161	-0.156
	(0.459)	(0.459)
Competitor BP	-0.079^{***}	
	(0.018)	
Competitor New BP		-0.043^{**}
		(0.021)
Competitor Old BP		-0.103^{***}
		(0.023)
Own Firm BP	-0.206^{***}	
	(0.072)	
Own Firm New BP		-0.252^{***}
		(0.061)
Own Firm Old BP		-0.234
		(0.147)
Market Size	0.033^{***}	0.034^{***}
	(0.007)	(0.009)
Coal Price	0.523^{**}	0.420^{*}
	(0.237)	(0.253)
Constant	-4.992^{***}	-4.628^{***}
	(0.517)	(0.564)
Observations	4,347	4,347
Log Likelihood	-282.755	-279.957
Akaike Inf. Crit.	579.510	577.914

Note: *p<0.1; **p<0.05; ***p<0.01.
SE's clustered at market level. BP in million tonnes.

Table 3: Dynamic Estimation: Retirement Policy Function

	(1)	(2)
Plant New Tech. BP	0.439	
	[0.277, 0.601]	
Plant Old Tech. BP	−0.337	
	[−0.675, −0.079]	
Competitor BP	0.351	0.251
	[−0.221, 0.923]	[0.058, 0.539]
Own Firm New BP	0.250	
	[0.02, 0.48]	
Own Firm Old BP	0.256	
	[−0.028, 0.540]	
Plant + Own Firm New BP		0.382
		[0.310, 0.467]
Plant + Own Firm Old BP		0.090
		[−0.034, 0.219]
Demand	−0.477	−0.423
	[−1.123, 0.169]	[−0.755, −0.181]
Coal Price	0.220	0.219
	[−0.064, 0.504]	[0.087, 0.331]
Observations	3,028	3,028
Log Likelihood	−1252.804	−1306.958

Note: We report bounds on the parameter estimates at the 5th and 95th percentiles. SE's calculated using bootstrapping. BP in 1000s of tons.

3. Find the vector of structural parameters, $\boldsymbol{\theta}$, which minimized the violations of the inequalities implied by step 2.

Below, we describe the details behind steps 1 to 3. In step 1, when we simulate the policy functions at the actual data, for each simulation path s and tuple k, we first draw choice-specific errors associated with the upgrade and retirement decisions. Denote the draws on the error shocks as $\epsilon^{\mathbf{retire}}_{\mathbf{jmt,k,s}}$ and $\epsilon^{\mathbf{upgrade}}_{\mathbf{jmt,k,s}}$, respectively, and the first stage parameter estimates as $\hat{\boldsymbol{\alpha}}^{retire}$ and $\hat{\boldsymbol{\alpha}}^{upgrade}$. We indicate a plants's simulated upgrade decision under the estimated policy in time t, draw s at a state-plant combination k as

$$\hat{u}_{jmt,k,s} = \mathbf{1}\{g^u(\mathbf{x_{mt-1,k,s}}, \mathbf{z_{mt,k,s}}; \hat{\boldsymbol{\alpha}}^{upgrade}) + \epsilon^{upgrade}_{j1t,k,s} > \epsilon^{upgrade}_{j0t,k,s}\},$$

and at the perturbed policy it is

$$\hat{u}'_{jmt,k,s} = \mathbf{1}\{g^u(\mathbf{x_{mt-1,k,s}}, \mathbf{z_{mt,k,s}}; \hat{\boldsymbol{\alpha}}^{upgrade}) + e_{jmt,k,s} + \epsilon^{upgrade}_{j1t,k,s} > \epsilon^{upgrade}_{j0t,k,s}\},$$

where $e_{jmt,k,s}$ is a normally distributed error term,[6] and $g^u(\mathbf{x_{mt-1,k,s}}, \mathbf{z_{mt,k,s}}; \hat{\boldsymbol{\alpha}}^{upgrade})$ is the estimated linear prediction in the binary logit model. For the simulated retirements, we similarly draw the choice-specific error $\epsilon^{retire}_{jt,k,s}$ associated with the retirement decision, compute the predicted payoff from each possible retirement choice, $g^r(\mathbf{x_{mt,k,s}}, \mathbf{z_{mt,k,s}}; \hat{\boldsymbol{\alpha}}^{retire}) + \epsilon^{retire}_{jt,k,s}$, and choose the simulated retirement decision conditional on where the payoff falls conditional on the ordered logit's estimated cutoffs. We denote the simulated retirement decision at the estimated policy as $\hat{r}_{jmt,k,s} \in \{0,1,2,3,4\}$; to construct the perturbed policy, $\hat{r}'_{jmt,k,s}$, we add a normally distributed error to the predicted payoff from retirement.

To construct the simulated value function at a state-firm combination k, we set the number of simulated periods to T, where β^T is small enough it is effectively zero. We thus forward-simulate firm choices, as well as firm and aggregate states, for T periods. The last piece that is necessary for simulating the value functions is the variable profits. These are constructed from a static spatial model of price competition developed in Miller et al. (2023). We describe our procedure for predicting profits at simulated states in Section 3.3.1. We normalize fixed costs to 0, following Aguirregabiria and Suzuki (2014), who show that the levels of fixed costs, entry costs and exit value are not separately identified. This normalization assumption means that scrap values and upgrade costs are measured relative to fixed costs, but we expect fixed costs to be small, given the small bounds on fixed costs estimated in Miller et al. (2023).

Turning to the formulas for the value functions, we first define the present discounted values of upgrade decisions, retirement decisions, their associated errors, and variable profits as:

[6] Half the perturbations use a N(0,1) distribution, and the other half use a N(0,3).

$$
\begin{aligned}
\hat{\mathbf{\Psi}}_{mk}^{upgrade} &= \frac{1}{N_s}\sum_{s=1}^{N_s}\sum_{t=0}^{T}\beta^t \hat{u}_{jmt,k,s} \\
\hat{\mathbf{\Psi}}_{mk}^{upgrade,e} &= \frac{1}{N_s}\sum_{s=1}^{N_s}\sum_{t=0}^{T}\beta^t \epsilon_{j\hat{u}t,k,s}^{upgrade} \\
\hat{\mathbf{\Psi}}_{mk}^{retire} &= \frac{1}{N_s}\sum_{s=1}^{N_s}\sum_{t=0}^{T}\beta^t \hat{r}_{jmt,k,s} \\
\hat{\mathbf{\Psi}}_{mk}^{retire,e} &= \frac{1}{N_s}\sum_{s=1}^{N_s}\sum_{t=0}^{T}\beta^t \epsilon_{jt,k,s}^{retire} \\
\hat{\mathbf{\Psi}}_{mk}^{VP} &= \frac{1}{N_s}\sum_{s=1}^{N_s}\sum_{t=0}^{T}\beta^t \hat{VP}_{jt,k,s}.
\end{aligned} \tag{3}
$$

The value function at the estimated policy can then be expressed as:

$$
\hat{V}_k(\boldsymbol{\theta}) = -UC\hat{\mathbf{\Psi}}_{mk}^{upgrade} + \hat{\mathbf{\Psi}}_{mk}^{upgrade,e} + RV\hat{\mathbf{\Psi}}_{mk}^{retire} + \hat{\mathbf{\Psi}}_{mk}^{retire,e} + \lambda\hat{\mathbf{\Psi}}_{mk}^{VP}, \tag{4}
$$

where the structural parameters are $\boldsymbol{\theta} = (UC, RV, \lambda)$. The value functions and $\hat{\mathbf{\Psi}}$ terms at the perturbed policies, $\hat{V}_k'(\boldsymbol{\theta})$ and $\hat{\mathbf{\Psi}}'$, are similarly defined.

A fundamental assumption in the BBL approach is that the policies played in the data are the result of firms' optimal policies. Under this assumption, at the true values of the structural parameters, it will be the case that $\hat{V}_k > \hat{V}_k'$. The structural parameters are thus chosen to minimize the number of times that previously mentioned the inequality is violated:

$$
\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}} \sum_{k=1}^{N_I} \min\{\hat{V}_k(\boldsymbol{\theta}) - \hat{V}_k'(\boldsymbol{\theta}), 0\}^2 \tag{5}
$$

3.3.1 Payoff Approximations

To incorporate the short-run model of procurement and the associated flow profits into our dynamic model, we build a non-linear predictive model using a gradient boosting algorithm. This model computes predicted payoffs as a function of the dynamic state variables in an order of magnitude less time than is required to solve the underlying procurement model.

To build this predictive model, we first regressed the estimated profits from the procurement model on our dynamic state variables to validate that all variables are economically meaningful predictors and have the expected signs. The results of this exercise are provided in Table 4 below. As expected, profits increase with higher old and new technology capacity, but more so with new technology capacity. Profits decrease with higher competitor capacity, both old and new technology. Plant-level profits also decrease with more capacity owned by the same firm in the market. Lastly,

Table 4: Dynamic State Variables and Estimated Plant Profits

	(1)
Plant New Tech. BP	10.733*** (2.271)
Plant Old Tech. BP	5.868** (2.981)
Competitor New BP	−0.559*** (0.119)
Competitor Old BP	−0.659*** (0.078)
Own Firm New BP	−0.652*** (0.250)
Own Firm Old BP	−0.300* (0.154)
Market Size	0.340*** (0.042)
Coal Price	−3,371.628*** (994.390)
Constant	21,102.500*** (4,135.337)

Note: *p<0.1; **p<0.05; ***p<0.01.
SEs clustered at regional level.
Boilerplate capacity & market size in 1000s of tonnes.

profits increase when demand is greater and decrease when coal prices are higher. Note that all state variables are defined at the regional market (cluster) level.

Next we use the gradient boosting algorithm from the xgboost package in R to estimate a non-linear model of flow profits using the same dynamic state variables. The predictive model that enters into the dynamic estimation and counterfactual simulations is trained on the full data set of cement plants. However, we separately train the model using a 80% sample and subsequently test its predictive ability on the remaining 20% sample to further validate our approach. When we sample observations at the plant level. we find that our predictive model explains 44% of observed variation in profits. When we instead sample observations at the plant-year level, our predictive model explains 89% of the observed variation in profits.

3.4 Preliminary Results

Preliminary estimates of dynamic parameters are presented in Table 5. All coefficients represent hundreds of millions of dollars.

Existing estimates of precalciner entry costs and old kiln decommissioning are scarce in the engineering literature. Nonetheless, one estimate that combines numbers from the European cement association and an environmental group gives costs of $800 million per precalciner kiln (Macher et al., 2021). Our estimates are similar in magnitude. In ongoing work, we are exploring the sensitivity of our results to heterogeneity in upgrade and decommissioning costs.

Table 5: Estimates of Dynamic Parameters

Profit Coefficient (λ)	0.749
Entry Cost (UC)	4.96
Decommissioning Cost (RV)	-1.37

4 Counterfactual Simulations

4.1 Implementation

To implement our counterfactual simulations, we first make some additional simplifications – most notably that all old technology kilns are the same size (183,000 metric tonnes) and that all new technology kilns are the same size (1,073,000 metric tonnes). This assumption means that when a plant decides to upgrade one kiln, we know exactly how their new capacity state variable will change, and it is already discretized into a small number of bins.

We also assume an approximate belief oblivious equilibrium (ABOE), developed in Gowrisankaran et al. (2022). This equilibrium concept represents a hybrid between an oblivious equilibrium on the one hand (Weintraub et al., 2008; Qi, 2013; Benkard et al., 2015), where firms (incorrectly) behave as though they are single agents, and a full dynamic game on the other hand, where firms internalize the impact of their actions on each of their competitors. By contrast, in ABOE, plants only keep track of the aggregate state, not the state of individual competitors, but they internalize how their own actions may affect that aggregate state.

To operationalize this equilibrium, we follow Gowrisankaran et al. (2022) and estimate from simulated data how a plant's competitors' capacity evolves conditional on whether the plant has old technology capacity, new technology capacity, or both. In this way, firms will internalize that if they retire their old technology capacity entirely, for example, competitor capacity will evolve differently going forward. We assume that competitor capacity follows a zero-inflated Poisson process, with lagged competitor capacity and a constant as explanatory variables.

Thus, in our "inner loop," we solve for value functions conditional on state variables and a set of beliefs about the evolution of competitor capacity, using successive approximations to iterate on value functions until convergence. Then, in our "outer loop," we use our estimated dynamic parameters, draws of choice specific errors, and value functions obtained from the convergence of the inner loop to simulate optimal upgrade and retirement choice of each plant in each period. This step yields a simulated dataset from which we re-estimate the zero-inflated Poisson process. We repeat this procedure until value function convergence in both the inner and outer loops. We set the tolerance to 10^{-5} for both the inner and outer loops. We initialize plant-level beliefs using a zero-inflated Poisson process estimated from observed data.

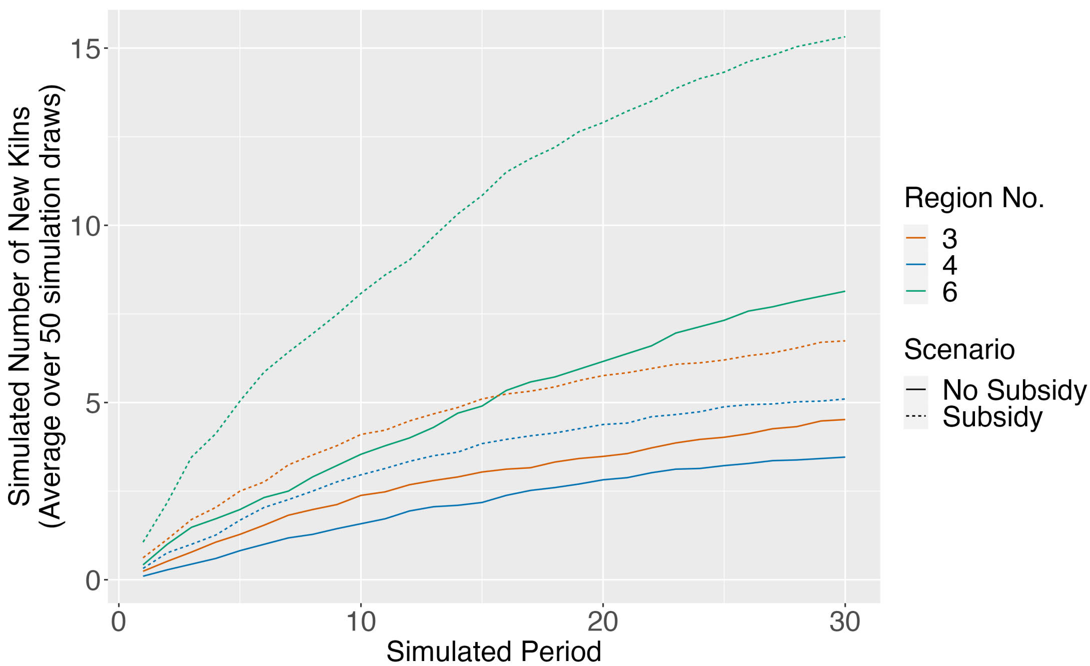

Figure 3: Simulated New Kilns Under Baseline and 30% Investment Subsidy Scenarios

4.2 Preliminary Counterfactual Results

Preliminary counterfactual results for regional markets (clusters) 3, 4, and 6 are presented in Figures 3, 4, and 5.

Preliminary results suggest that the total number of new technology kilns is higher under the 30% investment subsidy scenario than the baseline scenario in every single regional market, though how much higher varies across markets. By contrast, we do not observe a material change in the average number of simulated old technology kilns between the baseline and subsidy scenarios. We do see a slightly lower number of old kilns emerge at the end of the simulated period for regions 4 and 6, but not for region 3. These preliminary results suggest a net capacity expansion effect from the investment subsidy, which has implications both for welfare and for emissions which we are exploring in ongoing work.

5 Conclusion

In this paper, we analyze the dynamics of pollution abatement technology in the U.S. Portland cement industry. The cement industry experienced significant changes due to the introduction of the precalciner kiln—a pollution abating-technology. The empirical setting allows us to study the interaction between market power, high-cost technology adoptions, and investment subsidies. We provide preliminary results that market structure affects the impact of investment subsidies for pollution-reducing technologies. Ongoing work will continue adding richness to the dynamic model

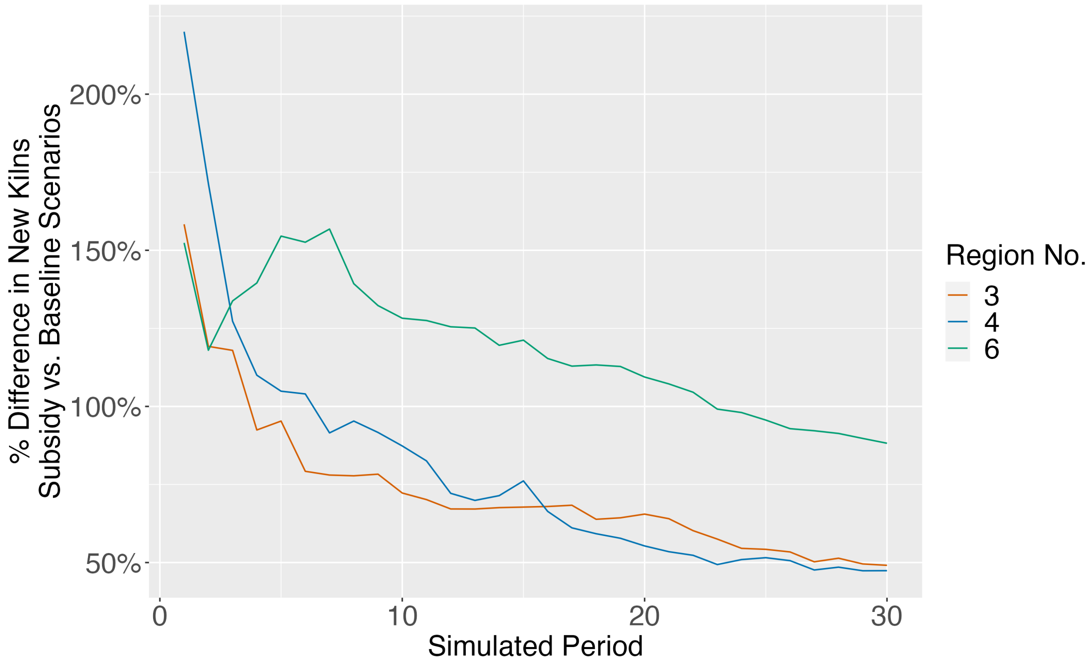

Figure 4: Percentage Difference in Number of New Kilns Under Baseline and 30% Investment Subsidy Scenarios

and exploring alternative subsidy designs.

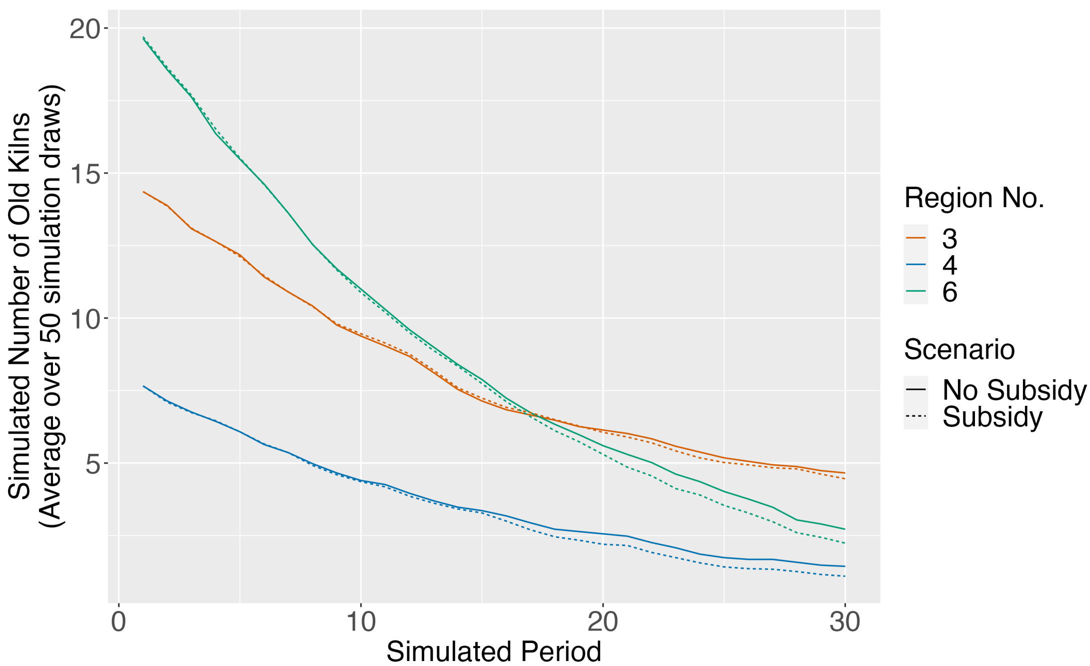

Figure 5: Simulated Old Kilns Under Baseline and 30% Investment Subsidy Scenarios

References

Aguirregabiria, Victor and Junichi Suzuki, "Identification and counterfactuals in dynamic models of market entry and exit," *Quantitative Marketing and Economics*, 2014, *12* (3), 267–304.

Bajari, Patrick, C. Lanier Benkard, and Jonathan Levin, "Estimating Dynamic Models of Imperfect Competition," *Econometrica*, 2007, *75* (5), 1331–1370.

Benkard, C. Lanier, Przemyslaw Jeziorski, and Gabriel Y. Weintraub, "Oblivious equilibrium for concentrated industries," *The RAND Journal of Economics*, 2015, *46* (4), 671–708.

Berry, Steven T., "Estimating Discrete-Choice Models of Product Differentiation," *RAND Journal of Economics*, 1994, *25* (2), 242–262.

Cardell, Scott N., "Variance Components Structures for the Extreme-Value and Logistic Distributions with Applications to Models of Heterogeneity," *Econometric Theory*, 1997, *13*, 185–213. *Concrete Future: The GCCA 2050 Cement and Concrete Industry Roadmap for Net Zero Concrete*

***Concrete Future: The GCCA 2050 Cement and Concrete Industry Roadmap for Net Zero Concrete**, Report, Global Cement and Concrete Association October 2021.*

Czigler, Thomas, Sebastian Reiter, Patrick Schulze, and Ken Somers, *"Laying the foundation for zero-carbon cement," Report, McKinsey Co. May 2020.*

EPA, *"Regulatory Impact Analysis: National Emission Standards for Hazardous Air Pollutants from the Portland Cement Manufacturing Industry," 2009.* `https://www.epa.gov/sites/default/files/2020-07/documents/nonmetallic-minerals_ria_proposal-cement-neshap_2009-04.pdf`.

Fennell, Paul, Justin Driver, Christopher Bataille, and Steven J. Davis, *"Cement and steel – nine steps to net zero.," Nature, March 2022, 603.*

Fowlie, Meredith, Mar Reguant, and Stephen P. Ryan, *"Market-Based Emissions Regulation and Industry Dynamics," Journal of Political Economy, 2016, 124 (1), 249–302.*

Glenk, Gunther, Anton Kelnhofer, Rebecca Meier, and Stefan Reichelstein, *"Cost-Efficient Decarbonization of Portland Cement Production," 2023. Working Paper.*

Gowrisankaran, Gautam, Ashley Langer, and Wendan Zhang, *"Policy Uncertainty in the Market for Coal Electricity: The Case of Air Toxics Standards," 2022. Working Paper.*

Hook, Leslie and Harry Dempsey, *"Quest for 'green' cement draws big name investors to* $300 bn industry,'' Financial Times, July 2021.

Kaufman, Leonard and Peter J. Rousseeuw, *1990. Finding Groups in Data: An Introduction to Cluster Analysis. John Wiley & Sons.*

Macher, Jeffrey, Nathan H. Miller, and Matthew Osborne, *"Finding Mr. Schumpeter: Technology Adoption in the Cement Industry," RAND Journal of Economics, 2021, 52 (1), 78–99.*

Miller, Nathan H., Matthew Osborne, Gloria Sheu, and Gretchen Sileo, *"Technology and Market Power: The United States Cement Industry, 1974-2019," 2023. Working Paper.*
Mission Possible: Reaching Net-Zero Emissions from Harder-to-Abate Sectors by Midcentury, Sectoral Focus Cement

Mission Possible: Reaching Net-Zero Emissions from Harder-to-Abate Sectors by Midcentury, Sectoral Focus Cement, Report, "Energy Transitions Commission" January 2019.

Qi, Shi, "The impact of advertising regulation on industry: the cigarette advertising ban of 1971," *The RAND Journal of Economics*, 2013, *44* (2), 215–248.

Ryan, Stephen, "The Costs of Environmental Regulation in a Concentrated Industry," *Econometrica*, 2012, *80* (3), 1019–1062.

Seiler, Stephan, "The impact of search costs on consumer behavior: A dynamic approach," *Quantitative Marketing and Economics (QME)*, June 2013, *11* (2), 155–203.
The Net-Zero Industry Tracker: Cement Industry

The Net-Zero Industry Tracker: Cement Industry, *Report, "World Economic Forum" July 2022.*

Weintraub, Gabriel Y., C. Lanier Benkard, and Benjamin Van Roy, *"Markov Perfect Industry Dynamics With Many Firms," Econometrica, 2008, 76 (6), 1375–1411.*

Wooldridge, Jeffrey M., *2010. Econometric Analysis of Cross Sectional and Panel Data, Second Edition, MIT Press.*

Appendix Materials

A Short Run Model of Procurement

In this appendix, we provide details on the short run model of procurement that is estimated in Miller et al. (2023). We first specify the functions that enter equilibrium profit (equation (2)), and then provide micro-foundations for the demand functions.

Plants in the model operate one or more kilns. The marginal cost of production at any kiln l (operated by plant j in year t) is given by

$$c_{jt}^{(l)}\left(Q_{jt}^{(l)};\boldsymbol{X}_t,\boldsymbol{\theta}\right)=\boldsymbol{w}_{jt}^{(l)}\boldsymbol{\alpha}+\gamma\left(\frac{Q_{jt}^{(l)}(\cdot)}{CAP_{jt}^{(l)}}-0.5\right)^2\mathbb{1}\left\{\frac{Q_{jt}^{(l)}(\cdot)}{CAP_{jt}^{(l)}}>0.5\right\} \tag{A.1}$$

where $Q_{jt}^{(l)}$ is the output of the kiln, $\boldsymbol{w}_{jt}^{(l)}$ is a vector of kiln-specific cost-shifters, $CAP_{jt}^{(l)}$ is the capacity of the kiln, and the parameters include $(\boldsymbol{\alpha},\gamma)$. The cost-shifters include a constant, a demeaned time trend, and the fuel cost of production, the latter of which captures that modern kilns are more efficient than older wet or long dry kilns. We measure capacity using the boilerplate rating of the kiln, which provides the theoretical maximum amount that could be produced with no downtime for maintenance. Thus, kiln-level marginal costs increase in output once the utilization rate, i.e., the ratio of output to capacity, exceeds 50%. Producing at capacity results in a marginal cost increase of $\gamma/4$ relative to producing at a utilization rate less than 50%.

Each plant allocates output across its kilns to minimize cost. For low-enough production, this entails producing only from the most efficient kiln. However, as that kiln reaches higher levels of utilization, production from less efficient kilns may become economical, and cost minimization dictates that the plant equate the kiln-specific marginal costs of any kiln that it uses. In this manner, we construct a continuous and weakly upward sloping *plant-level* marginal cost function, $c_{jt}(Q_{jt};\boldsymbol{X}_t,\boldsymbol{\theta})$, from the kiln-level marginal cost function of equation (A.1). Figure A.1 shows the marginal cost function that we obtain for one of the multi-kiln plants in our data.

Let the market-clearing plant-level quantities be Q_{jt}^*, and denote the marginal costs that are associated with these market-clearing quantities as $c_{jt}^*(\boldsymbol{X}_t,\boldsymbol{\theta})\equiv c_{jt}(Q_{jt}^*;\boldsymbol{X}_t,\boldsymbol{\theta})$. The equilibrium market share that plant j obtains in county n is given by

$$s_{jnt}(\boldsymbol{X}_t,\boldsymbol{\theta})=\frac{\exp\left(\frac{\overline{u}_{jnt}(\boldsymbol{X}_t,\boldsymbol{\theta})-\phi c_{jt}^*(\boldsymbol{X}_t,\boldsymbol{\theta})}{1-\sigma}\right)}{\sum_{k\neq 0}\exp\left(\frac{\overline{u}_{knt}(\boldsymbol{X}_t,\boldsymbol{\theta})-\phi c_{kt}^*(\boldsymbol{X}_t,\boldsymbol{\theta})}{1-\sigma}\right)} \times\frac{\left(\sum_{k\neq 0}\exp\left(\frac{\overline{u}_{knt}(\boldsymbol{X}_t,\boldsymbol{\theta})-\phi c_{kt}^*(\boldsymbol{X}_t,\boldsymbol{\theta})}{1-\sigma}\right)\right)^{1-\sigma}}{1+\left(\sum_{k\neq 0}\exp\left(\frac{\overline{u}_{knt}(\boldsymbol{X}_t,\boldsymbol{\theta})-\phi c_{kt}^*(\boldsymbol{X}_t,\boldsymbol{\theta})}{1-\sigma}\right)\right)^{1-\sigma}} \tag{A.2}$$

where $\overline{u}_{jnt}(\boldsymbol{X}_t,\boldsymbol{\theta})$ is a function that characterizes the common value that the plant provides to

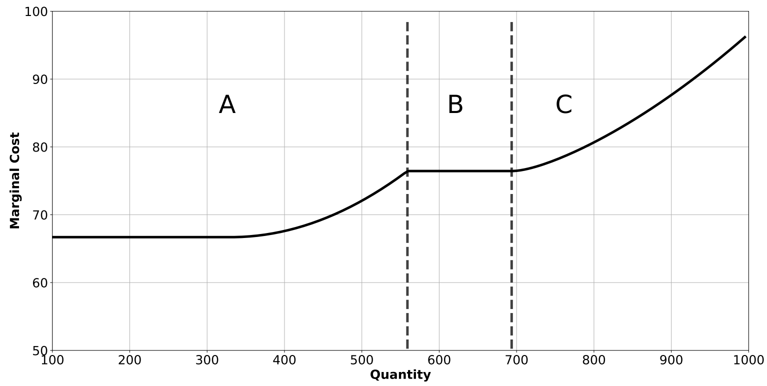

Figure A.1: Plant-Level Marginal Cost per Metric Tonne (Illustrative Example)

Notes: The figure plots the marginal cost function of the Flintkote plant (Kosmodale, Kentucky) in 1974, taking as given our parameter estimates. The plant has two kilns. It initially uses the more efficient kiln to produce marginal output (region A), then it uses the less efficient kiln (region B), and finally it splits marginal output between the kilns (region C). The vertical axis is in dollars per metric tonne and the horizontal axis is in thousands of metric tonnes.

buyers in the county; we specify it momentarily. The first term on the right-hand-side of equation (A.2) is probability that a buyer chooses the plant, conditional on some cement plant being chosen. The second term is the probability that some cement plant is chosen over an outside option (e.g., steel, wood, or asphalt). Thus, the markets shares have a nested logit structure. The equilibrium quantities are given by $q_{jnt}(\boldsymbol{X}_t, \boldsymbol{\theta}) = s_{jnt}(\boldsymbol{X}_t, \boldsymbol{\theta}) M_{nt}$ where M_{nt} is potential demand.

The $\overline{u}(\cdot)$ function that enters the market share function incorporates a disutility of transportation and whether the supplier is a domestic plant or the importer. Shipments can go by truck or rail directly from the plant to the buyer, or to go by barge utilizing the Mississippi River System. The specification is:

$$\begin{aligned}\overline{u}_{jnt}(\boldsymbol{X}_t, \boldsymbol{\theta}) = \min\{\beta_1 d_{j\to n}\ ,\ \beta_1\left(d_{j\to R} + d_{R\to n}\right) + \beta_2\} \\ + \beta_3 TREND_t + \beta_4 IMPORT_j + \beta_0 \quad \text{(A.3)}\end{aligned}$$

The first line on the right-hand-side is the disutility of transportation, where $d_{j\to n}$ is the distance between the plant and the county, $d_{j\to R}$ is the distance between the plant and the Mississippi River System, and $d_{R\to n}$ is the distance between the Mississippi River System and the county. The parameters β_1 and β_2 capture, respectively, the per-mile disutility associated with overland transportation and a fixed disutility associated with barge transportation. This reflects that barge transportation is much more cost efficient than overland transportation on a per-mile basis but requires users to pay loading charges. The preferred form of transportation is used. The second line incorporates a demeaned time trend ($TREND_t$), an indicator for the importer ($IMPORT_j$), and a constant for the inside goods.

The average price that a plant receives in equilibrium conditional on winning an auction in a given county-year can be decomposed into marginal cost and an expected markup:

$$\overline{p}_{jnt}(\boldsymbol{X}_t, \boldsymbol{\theta}) = c^*_{jt}(\boldsymbol{X}_t, \boldsymbol{\theta}) + \overline{m}_{jnt}(\boldsymbol{X}_t, \boldsymbol{\theta}) \tag{A.4}$$

where the expected markup is

$$\overline{m}_{jnt}(\boldsymbol{X}_t, \boldsymbol{\theta}) = -\frac{1}{\phi} \frac{1}{\sum_{k \in \mathbb{J}_{f(j)}} s_{knt}} \log \left[1 - (1 - s_{0nt}) \left(1 - \left(1 - \sum_{k \in \mathbb{J}_{f(j)}} \frac{s_{knt}}{1 - s_{0nt}} \right)^{1-\sigma} \right) \right] \tag{A.5}$$

These equation obtain under a second-score auction of procurement in which buyer i (in county n) scores bidders (plants) according to

$$u_{ijnt} = \overline{u}_{jnt}(\boldsymbol{X}_t, \boldsymbol{\theta}) + \zeta_{int} + (1 - \sigma)\epsilon_{ijnt} \tag{A.6}$$

where each ϵ_{ijnt} is distributed iid type 1 extreme value and ζ_{int} has the unique distribution such that $\epsilon^*_{ijnt} \equiv \zeta_{int} + (1 - \sigma)\epsilon_{ijnt}$ also is type 1 extreme value (Berry, 1994; Cardell, 1997).

B Modified Ordered Logit Model

In this appendix, we describe the modified ordered logit model that we use to estimate the policy function for kiln retirement. The reason we make modifications is because the plants in our model can have a different number of old kilns. Plants with one kiln can retire it or not. The choice for plants with two kilns is different, as they can retire one, both, or neither kiln. Other plants can retire three or four kilns. In principle, we could estimate different policy functions for each group of plants (those with 1, 2, 3, or 4 old kilns) but we find that our estimates are more economically reasonable if the plants are treated together.

We first introduce the notation in chapter 16.3.1 of Wooldridge (2010) for a standard ordered logit model. Suppose that all plants have J kilns, and define y as an ordered response, with values $(0, 1, \ldots, J)$, that represents the number of kilns that are retired. Define the latent variable y^* such that

$$y^* = \boldsymbol{x}\boldsymbol{\beta} + e, \qquad e|\boldsymbol{x} \sim \text{logit} \tag{B.1}$$

where $\boldsymbol{x}$ is a vector of covariates. Let $\alpha_1 < \alpha_2 < \cdots < \alpha_J$ be a set of cut points such that:

$$\begin{aligned} y &= 0 \quad \text{if } y^* \leq \alpha_1 \\ y &= 1 \quad \text{if } \alpha_1 < y^* \leq \alpha_2 \\ &\vdots \\ y &= J \quad \text{if } y^* \geq \alpha_J \end{aligned} \tag{B.2}$$

The conditional probability observing the different values of y are then

$$
\begin{aligned}
Pr(y=0|\boldsymbol{x}) &= \Lambda(\alpha_1 - \boldsymbol{x\beta}) \\
Pr(y=1|\boldsymbol{x}) &= \Lambda(\alpha_2 - \boldsymbol{x\beta}) - \Lambda(\alpha_1 - \boldsymbol{x\beta}) \\
&\vdots \\
Pr(y=J|\boldsymbol{x}) &= 1 - \Lambda(\alpha_1 - \boldsymbol{x\beta})
\end{aligned}
$$

where Λ is the CDF of the logit distribution. The parameters $\boldsymbol{\beta}$ and $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_J)$ can be estimated with maximum likelihood. For each observation i, the contribution to the log-likelihood function is

$$
\begin{aligned}
l_i &= 1[y_i = 0]\log[\Lambda(\alpha_1 - \boldsymbol{x\beta})] + 1[y_i = 1]\log[\Lambda(\alpha_2 - \boldsymbol{x\beta}) - \Lambda(\alpha_1 - \boldsymbol{x\beta})] \\
&\quad + \cdots + 1[y = J]\log[1 - \Lambda(\alpha_1 - \boldsymbol{x}]
\end{aligned}
$$

To modify this standard model for our setting, we assume that the parameters in $\boldsymbol{\beta}$ are the same for every group of plants (1, 2, 3, 4 old kilns). The cut points are allowed to vary across groups. For plants with one old kiln, there is one cut point, for plants with two old kilns, there are two cut points, and so on. In total, we estimate the parameters in $\boldsymbol{\beta}$ and 10 cut points $(10 = 4+3+2+1)$. We construct the log likelihood by combining the observations from plants in all the groups.

C Additional Figures and Tables

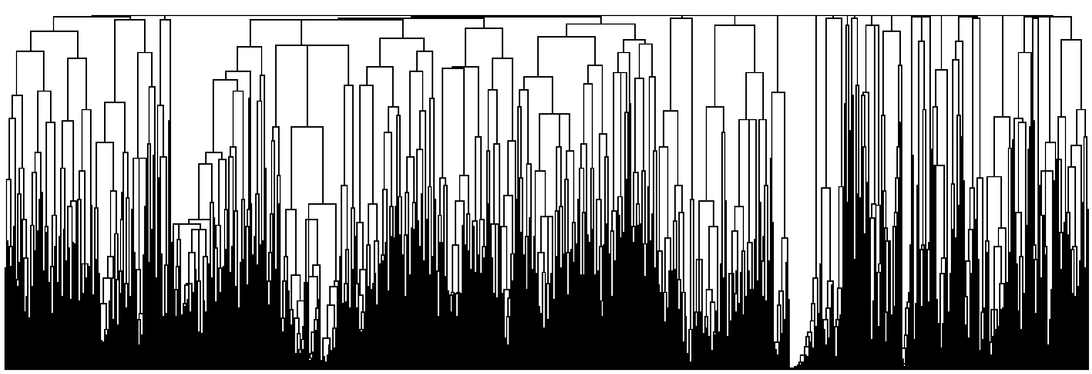

Figure C.1: Dendrogram from the Clustering Algorithm

Notes: The figure illustrates that the hierarchical clustering algorithm starts with all the counties in the same group (at the top) and then splits the counties into smaller, heterogeneous groups continuously until every county is by itself (at the bottom). The "trade flow" markets that we use in our baseline analysis are defined as the groups that exist at the point that the algorithm has defined exactly 10 different groups.