The Costs of Environmental Regulation in a Concentrated Industry*

Stephen Ryan†

November 8, 2004

Abstract

The typical cost analysis of an environmental regulation consists of an engineering estimate of the compliance costs. In industries where fixed costs are an important determinant of market structure this static analysis ignores the dynamic effects of the regulation on entry, investment, and market power. I evaluate the welfare effects of the 1990 Amendments to the Clean Air Act on the US Portland cement industry, accounting for these effects through a dynamic model of oligopoly in the tradition of Ericson and Pakes (1995). Using a recently developed two-step estimator, I recover the entire cost structure of the industry, including the distribution of sunk entry costs and adjustment costs of investment. I find that the Amendments have significantly increased the sunk cost of entry. I solve for the Markov perfect Nash equilibrium (MPNE) of the model and simulate the welfare effects of the Amendments. A static analysis misses the welfare penalty on consumers, and obtains the wrong sign on the welfare effects on incumbent firms.

---

*I would like to thank Pat Bajari, Han Hong, Tom Ahn, Arie Beresteanu, Jane Cooley, Paul Ellickson, Shanjun Li, Chris Timmins, and Justin Trogdon for their helpful comments. All remaining errors are my own.

†Department of Economics, Duke University, Durham, NC 27708.

1 Introduction

In the United States, the Environmental Protection Agency (EPA) is responsible for setting and enforcing regulations broadly consistent with national environmental policies, such as the Clean Air Act (CAA). The CAA gives the EPA a mandate to regulate the emissions of airborne pollutants such as ozone, sulfur dioxide (SO_2), and nitrogen oxides (NO_x), in the hopes of producing a healthier atmosphere. It also requires the Agency to assess the costs and benefits of a regulation before promulgating policy. The cost analysis is typically an engineering estimate of the expenditures on control and monitoring equipment necessary to bring a plant into compliance with the new regulations. However, this type of cost analysis misses most of the relevant economic costs in concentrated industries, in which sunk costs of entry and costly investment are important determinants of market structure. Shifts in the costs of entry and investment can lead to markets with fewer firms and lower production. The resulting increase in market concentration can have far-reaching welfare effects beyond the initial costs of compliance. This is a particularly acute problem for environmental regulators, as many of the largest polluting industries are also highly concentrated.[1]

In this paper, I measure the welfare costs of the 1990 Clean Air Act Amendments on the US Portland cement industry, explicitly accounting for the dynamic effects resulting from a change in the cost structure. Portland cement is the binding material in concrete, a primary construction material found in everything from buildings to highways. The industry is typical of many heavy industries, consuming large quantities of raw materials and generating significant amounts of pollution byproducts. It is a frequent target of environmental activists and has been heavily regulated under the Clean Air Act. In 1990, Congress passed Amendments to the Clean Air Act, adding new categories of regulated emissions and requiring plants to undergo an environmental certification process. It has been the most comprehensive and important new environmental regulation affecting this industry in the last three decades since the original Clean Air Act.

My strategy for evaluating the effects of the Amendments on this industry proceeds in three distinct steps. First, I pose a theoretical model of the cement industry, where

---

[1] For example, the 1997 Economic Census reports that the Herfindahl-Hirschman Index (HHI) for many polluting industries exceeds 1,000, such as manufacturers of paper pulp, petrochemicals, soaps and detergents, tires, ceramic tiles, lime and gypsum, aluminum, and copper, among others. For comparison, the HHI for Portland cement, the industry studied in this paper, is 466. This national measure understates the effective degree of concentration since the industry is spatially segregated into regional markets.

oligopolists make optimal decisions over entry, exit, production, and investment given the strategies of their competitors. Second, using a unique panel data set of the Portland cement industry, I recover parameters which are consistent with the underlying model. Third, I use the theoretical model to simulate economic environments with the cost structures recovered before and after the Amendments. I exploit a specific timing feature of the implementation of the Amendments to identify which changes in the cost structure were due to the regulations. By comparing the predictions of the model under these different cost structures, I evaluate the changes to a number of relevant policy quantities, such as producer profits and consumer surplus, that are the result of the regulation.

The backbone of my analysis is a fully dynamic model of oligopoly in the tradition of Ericson and Pakes (1995). I model the interaction of firms in spatially-segregated regional markets where firms are differentiated by production capacity. Firms are capacity constrained and compete over quantities in homogeneous good markets. Markets evolve as firms enter, exit and adjust their capacities in response to variation in the economic environment. I also incorporate fixed costs of entry, exit, and capacity adjustment. Given a richly-specified state space, I assume that firms optimize their behavior independently across periods, implying a Markov-perfect Nash equilibrium (MPNE).

My model is similar to several other applications of the Ericson-Pakes model.[2] However, I tailor the model to the Portland cement industry in several ways. First, I allow firms to fully adjust their capacity in each period, whereas previous models have looked at investment games where capital accumulates slowly. This is necessary given the infrequent but large adjustments firms make to their capacity levels. Second, following Doraszelski and Satterthwaite (2004), I introduce private information into the model in several ways. Firms have private information about their entry and exit costs, marginal cost of production, and the fixed costs of capacity adjustment. The fixed cost of capacity adjustment induces investment policies where firms are reticent to make small adjustments to their capacity, mirroring the lumpy investment behavior seen in the data. The idiosyncratic shocks help rationalize why, within a group of otherwise ex-ante identical firms, some flourish and grow while others stagnate and fail. Third, I allow for multiple entry and exit in every period. Entrants are not restricted to begin operation at an exogenously imposed capacity level, but rather choose an optimal starting level given the market configuration.

---

[2] See, for example, Fershtman and Pakes (2000), Gowrisankaran and Town (1997), Besanko and Doraszelski (2004), Doraszelski and Satterthwaite (2004), and Benkard (Forthcoming).

The MPNE of the model leads to structural requirements on firm behavior which can be used as the basis of an estimator of the underlying primitives. Previously, the impediment to using these types of models for empirical work has been the computational burden of solving for the MPNE, which makes nested fixed-point estimators (for example, Rust (1987)) impractical. However, a spate of recent papers has shown how to circumvent this problem using a two-step approach in which it is possible to estimate the dynamic model without solving for the equilibrium even once.[3] The first step simply describes *what* the firms do at every state, and the second step recovers parameters of the underlying model that explain *why* the firms behave as they do. In my application, the first step includes flexibly recovering the demand curve, production costs, and the policy functions governing entry, exit, and investment. These reduced-form policy functions describe what actions the firm will undertake given any state vector. The key to understanding the estimator is that these observed policy functions have to be optimal given the underlying theoretical model. Therefore, in the second second step, I recover the remaining unknown parameters best rationalizing the observed policies as optimal given the restrictions of the MPNE. I follow the simulation-based minimum distance estimator proposed by Bajari, Benkard, and Levin (2003) to recover the fixed and variable costs of investment and the distributions of entry and exit costs. I recover these parameters before and after the 1990 Amendments to identify changes in the cost structure due to the regulation.

Once I recover the static and dynamic parameters driving the evolution of the Portland cement industry, I solve for the MPNE of the theoretical model. This step provides policy functions which are used for the simulation of the welfare effects of the Amendments. I evaluate expected producer and consumer welfare, the number and size of firms, and the distribution of costs across incumbents and potential entrants before and after the regulations. In the baseline case of entry into a new market, I find that overall welfare has decreased roughly 15% as a result of the Amendments, due to an increase in the average sunk cost of entry. More importantly, as my estimates show the costs of production are lower after the regulations, the welfare effect on producers depends critically on whether or not the firm is an incumbent. While potential entrants suffer the greatest welfare losses, incumbent firms actually benefit from increased market power due to reduced entry. A static analysis of this industry would preclude changes in barriers to entry, and would obtain the wrong sign for

---

[3] See Bajari, Benkard, and Levin (2003), Hotz and Miller (1993), Aguirregabiria and Mira (2002), Pakes, Ostrovsky, and Berry (2003), Pesendorfer and Schmidt-Dengler (2003), and Jofre-Benet and Pesendorfer (Forthcoming).

the welfare effects of the Amendments on incumbent firms.

I conclude by comparing the MPNE generated by an oligopoly to the social planner's problem. The social planner pursues policies that maximize the expected sum of both consumer and producer welfare, so it is the natural baseline for evaluating the efficiency of the observed equilibrium. The social planner's solution highlights that oligopoly markets are have inefficiently low aggregate capacities. This is due to the fact that oligopolists fail to internalize the full welfare benefits of their investments. As a result, overall market capacity is larger under the social planner in both new and existing markets. In both cases, producers suffer moderate profit losses while consumers enjoy a three- to five-fold increase in surplus which more than offsets those losses.

This paper makes several contributions. First, I recover the entire cost structure of an industry, including the sunk costs of entry and exit, production costs, and investment costs. Recovering these parameters allows the measurement of a regulation's welfare effects in the presence of dynamics and market power for the first time.[4] These welfare cost estimates allow me to determine a lower bound of the value of clean air if the Amendments are to be efficient. One of the most important implications of my findings is that static engineering estimates of compliance costs miss most of the economic penalties associated with a regulation when there are significant sunk entry costs. The static analysis misses the penalty to consumer welfare due to lower production, and obtains a welfare cost of the wrong sign for incumbent producers. I also make a contribution to the investment literature by applying a generalized (S, s) model to a dynamic investment game.[5] My results suggest that both fixed and variable adjustment costs are an important determinant of investment behavior. I also highlight the importance that sunk costs of entry have on industry structure and evolution, since they are a primary determinant of market structure in this industry.

The paper is organized as follows. I give a brief overview of the Portland cement industry and relevant environmental regulations over the last 30 years in Section 2. I discuss the sources of the data and introduce the key variables of the model and estimation in Section 3. Section 4 introduces the theoretical model underpinning the estimation detailed in Section

---

[4] Mansur (2004) also examines the regulation of an industry with market power but is concerned with the effect of concentration on the quantity of pollution emissions. Benkard (Forthcoming) applies many of the ideas formalized in the BBL estimator in his examination of the widebody aircraft industry but does not recover estimates of fixed costs.

[5] Attanasio (2000) and Hall and Rust (2000) apply similar frameworks for modeling automobile purchase decisions and inventory control.

5. I discuss the results in Section 6 and present the results of the counterfactual simulations in Section 7. Section 8 concludes with a summary of my results and a discussion of possible extensions.

2 Portland Cement Industry

Portland cement is a fine mineral dust with useful binding properties that make it the key ingredient of concrete. Water and cement form a paste that binds particulates like sand and stone together and makes a pourable material that hardens over time. The concrete is then used as a fill material, such as in highways and buildings, and in finished products like concrete blocks. I briefly describe the production process for Portland cement and its implications on the economic structure of the industry. I then give an overview of the industry and review the relevant environmental regulations of the last 30 years.

2.1 Economics of Cement Production

Portland cement is composed of calcium, aluminum, silicon, and iron.[6] The cement manufacturing process begins at the quarry where limestone is mined in large chunks. These chunks are pulverized before being sent to the centerpiece of cement operations: an extremely large rotating kiln furnace, as depicted in Figure 1.[7] Under extreme heat, the powder undergoes chemical transformation into clinker, which is cooled and ground into the final product, Portland cement. In its final form, Portland cement is an extremely fine dust, capable of passing through a sieve tight enough to hold water. Once ground, cement is stored for a short time until it is ready to be shipped.[8] Portland cement is of relatively uniform quality since producers in the United States adhere to the American Society for Testing and Materials

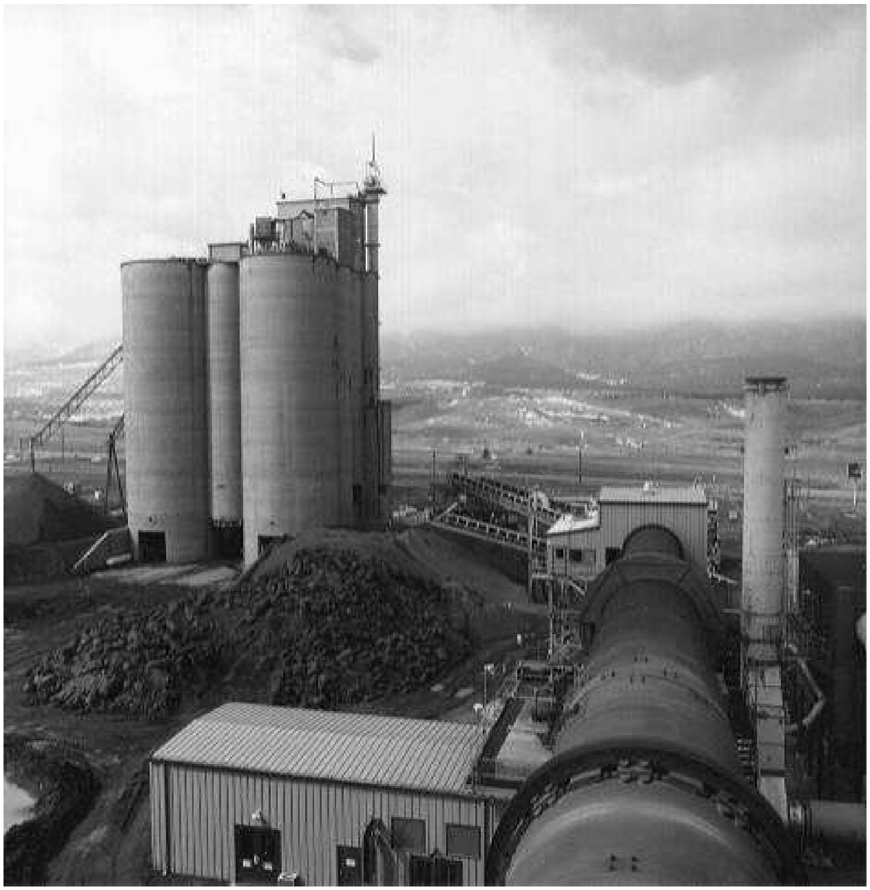
---

[6] The information in this section follows the PCA (2001). The calcium usually comes from limestone, marl, and chalk. Clay, shale, bauxite, and iron ore provide silicon, aluminum, and iron components. Many cement plants also utilize other manufacturing by-products, such as furnace slag, fly ash, and mill scale. The most common mix of ingredients is limestone, clay, and sand.

[7] Cement kilns are the single largest moving piece of industrial equipment in the world. They range in length from 200 to 500 feet and are roughly 15 feet in diameter.

[8] Cement gradually absorbs water in storage, rendering it useless, so producers and distributors do not maintain large stocks.

Figure 1: Cement Kiln



A view of a cement operation, looking down the main kiln with storage towers in the background. The rotating kiln is the single largest moving piece of industrial equipment in the world.

Table 1: Cement Industry Summary Statistics

Year	Production	Imports	Consumption	Price	Capacity	Capacity Per Kiln	Utilization Rate
1980	68,242	3,035	70,173	111.90	89,561	239	76
1981	65,054	2,514	66,092	103.70	93,203	267	70
1982	57,475	2,231	59,572	95.76	89,770	287	64
1983	63,884	2,960	65,838	91.01	92,052	292	69
1984	70,488	6,016	76,186	89.70	91,048	297	77
1985	70,665	8,939	78,836	84.71	88,600	305	80
1986	71,473	11,201	82,837	81.48	87,341	305	82
1987	70,940	12,753	84,204	78.07	86,709	314	82
1988	69,733	14,124	83,851	75.50	86,959	327	80
1989	70,025	12,697	82,414	72.04	84,515	337	83
1990	69,954	10,344	80,964	69.02	83,955	345	83
1991	66,755	6,548	71,800	66.37	84,471	352	79
1992	69,585	4,582	76,169	64.25	85,079	357	82
1993	73,807	5,532	79,701	63.58	84,869	363	87
1994	77,948	9,074	86,476	68.06	85,345	364	91
1995	76,906	10,969	86,003	72.56	86,285	367	89
1996	79,266	11,565	90,355	73.64	85,687	376	93
1997	82,582	14,523	96,018	74.60	86,465	383	96
1998	83,931	19,878	103,457	76.45	87,763	393	96

Summary statistics for the Portland cement industry 1980-1998. The data is from Historical Statistics for Mineral and Materials Commodities in the United States, an online publication of the US Geological Survey. Quantities are in thousands of metric tons, and prices are in 1998 constant dollars.

Specification for Portland Cement.[9]

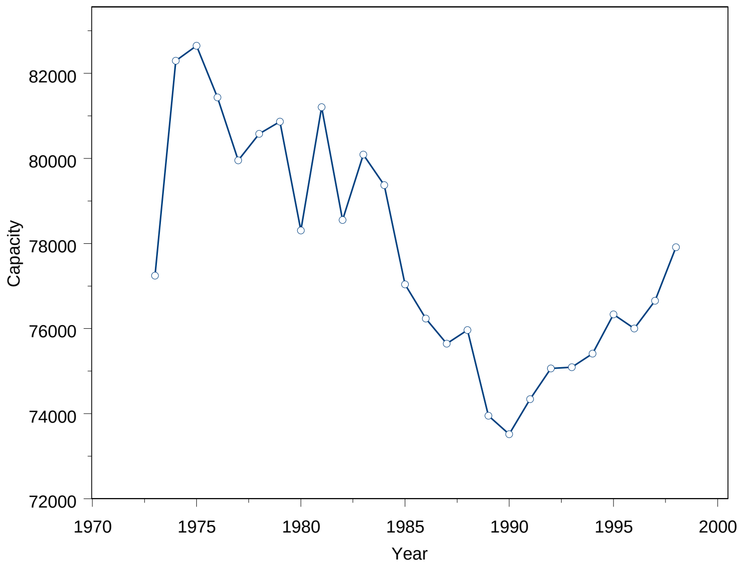
Transportation costs are the most significant factors in determining Portland cement markets. Firms generally locate near population centers or waterways to facilitate shipping their output. For landlocked markets, railway and trucking are the only forms of transportation available. The low value per unit weight and high storage requirements are the principal reasons the majority of cement is shipped locally. Jans and Rosenbaum (1997) quote a Census of Transportation report stating that 82.5% of cement was shipped under 200 miles, with 99.8% being shipped under 500 miles.

In 2000, the domestic Portland cement industry consisted of 116 plants in 37 states, run by one government agency and approximately 40 firms. The industry produced 86 million

---

[9] The most common is Type I, which is intended for general use. There are also a number of minor variants. Type I is also a general-use cement useful when sulfate resistance is needed or a moderate level of hydration heat is required. This is useful because the final application of cement releases heat which may need to be attenuated. Type III has early high-strength properties, Type IV has low heat hydration, and Type V has high sulfate resistance. Information in this section comes from the C150-00 Standard Specification for Portland Cement (2001) code produced by ASTM International. Types I and II accounted for 90% of sales in 1999, and Types I through V were 96% of sales (1973–2001). These factors allow me to concentrate on Portland cement as a single homogeneous good industry.

Figure 2: Clinker production capacity 1973-1998.



tons of Portland cement with a raw value of approximately \$8.7 billion; most of this was used to make concrete, with a final value greater than \$35 billion. Domestic cement production accounted for the vast majority of the cement used in the United States.[10] The construction industry drives demand for cement; since cement expenditures are a small percentage of construction project budgets I treat shocks to cement demand as exogenous.

Table 1 reports summary statistics for the industry over the period 1980-1998. There are two trends worth noting. The first is that capacity utilization rates have risen since the passage of the Amendments. Production has increased while overall productive capacity has remained relatively steady, as shown in Figure 2. The second is that imports became a more important source of cement in the late 1990's. Imports grew as the production of domestic cement reached its maximum level, and firms chose to import instead of build new production facilities.[11] It is worth noting that the import levels of 1998 and 1999 are atypically high

---

[10] The six largest cement production states were California, Texas, Pennsylvania, Michigan, Missouri, and Alabama, in descending order of production, and they accounted for almost 50% of domestic production. About 73% of cement sales were to ready-mixed concrete manufacturers, 12% to concrete product producers, 8% to contractors (primarily for paving), 5% to building materials dealers, and 2% for other uses (2001).

[11] Cement imports come primarily from Canada, China, Korea, Thailand, Spain, and Venezuela. Asian

and have since fallen to less than 2% of apparent consumption, as of 2004.[12] The effects of imports on domestic producers are difficult to quantify due to the idiosyncracies associated with distributing cement from waterborne sources. For most markets, the economic impact is small and indirect, as few regions have the infrastructure and geography to profitably exploit the availability of imports. I assume that these effects are captured by a negative demand shift in the market for domestically-produced cement.

2.2 Environmental Regulation

There have been two major regulatory events of interest to the Portland cement industry in the last 30 years: the Clean Air Act of 1970 and its subsequent Amendments in 1990. The stated purpose of the Clean Air Act was to "protect and enhance the quality of the Nation's air resources so as to promote the public health and welfare and productive capacity of its population." To this end, Congress empowered the EPA to set and enforce environmental regulations governing the emission of airborne pollutants. In 1990, Congress passed the Amendments to the Clean Air Act, which defined new categories of regulated pollutants and required major polluters to obtain a permit for operation.

When the EPA drew up its first set of regulations in 1971, the Portland Cement Association (PCA), the industry's trade group, immediately opposed them in federal court. The verdict of this case, PCA vs. Ruckelshaus[13], was a watershed decision that helped mold the way the EPA crafts regulations. The legal documents surrounding the appeal give a detailed insight into the thought process of both the industry and the EPA.

One particular concern of the industry illustrates the dynamic effects of environmental regulations. There are two different basic processes for producing cement, wet and dry. The wet method is the older method, and wet kilns were typically smaller and less expensive than their dry kiln counterparts. However, the industry had concerns about the economic feasibility of wet process plants after the new Clean Air Act regulations, since this process was more environmentally damaging. As shown in Figure 3 above, these concerns have been justified in the three decades following the initial regulations. In 1997, there were 70 wet

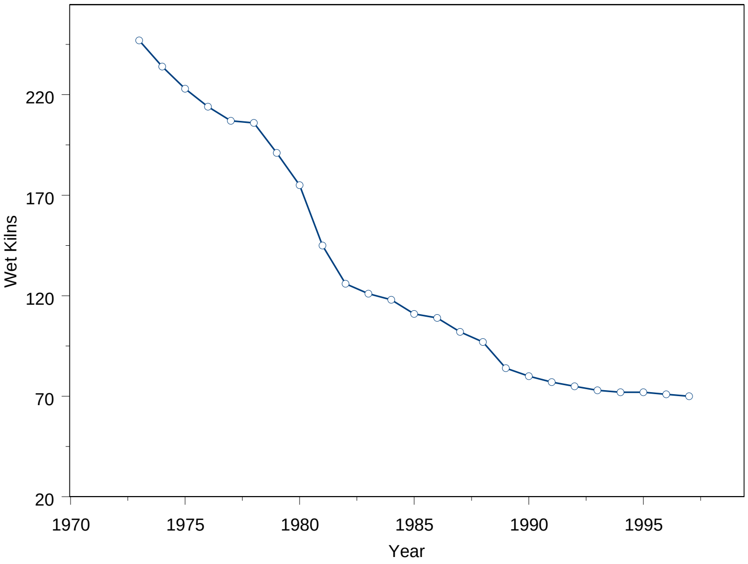
---

sources have become the dominant source of cement imports, with Thailand becoming the single-largest exporter in 2000.

[12] USGS Minerals Yearbook, 2004.

[13] Ruckelshaus was the administrator of the EPA at the time. See PCA versus Ruckelshaus, D.C. Circuit of Appeals, 1973, No. 72-1073.

Figure 3: Number of wet process cement kilns, 1973-1998.

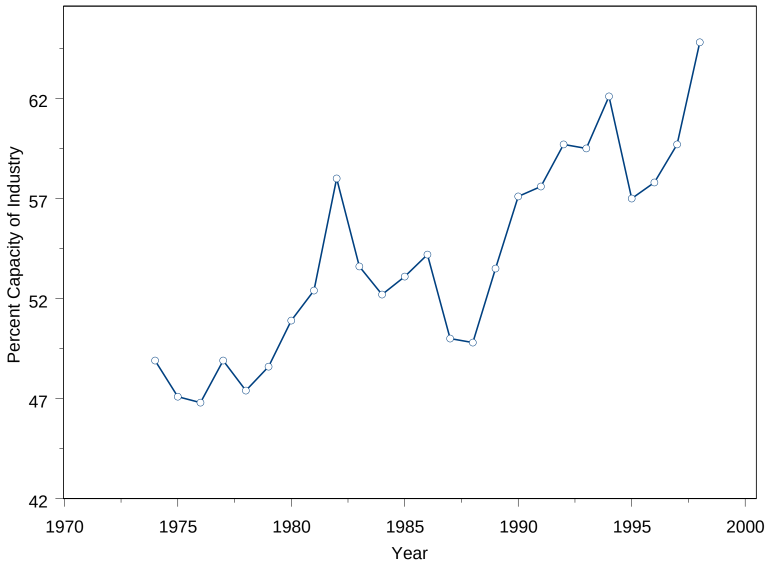


process and 130 dry process kilns, compared to 247 wet process and 198 dry process kilns in 1973. This decrease also had significant distributional effects, as many of the wet kiln establishments that shut down operated in fringe markets. The EPA failed to consider the dynamic effects of the regulations, as they expected the ratio of wet and dry process kilns to remain constant.

The exodus of firms operating wet kilns has coincided with an industry trend towards higher levels of concentration. Figure 4 shows the 10-firm concentration ratio over the last 30 years. This trend has continued into the 1990's, as marginal firms either exit the industry or merge with other firms. The reduction of independent firms has not been offset by entry, particulary after the passage of the Amendments to the Act, as the rate of entry has tended towards zero. This is particularly startling in light of the strong and persistent growth in demand in the later part of the 1990's. In fact, my demand estimates below indicate that the demand for cement was stronger in 1995 onward than any other year since 1980.

The rationale for the EPA's cost-benefit methodology was heavily influenced by the final decision of the D.C. Circuit Court of Appeals, which set the foundation of regulation both in and outside the cement industry for the next three decades. The EPA's metric for assessing economic burden consisted of determining the cost of environmental equipment

Figure 4: Concentration ratio of top ten firms 1974-1998.



as a percentage of the total cost of the plant. In response to the PCA's objections to this methodology, the court wrote:

> Petitioners [PCA] argue that this analysis is not enough—that the Administrator is required to prepare a quantified cost-benefit analysis, showing the benefit to ambient air conditions as measured against the cost of the pollution devices. However desirable in the abstract, such a requirement would conflict with the specific time constraints imposed on the Administrator. The difficulty, if not impossibility, of quantifying the benefit to ambient air conditions, further militates against the imposition of such an imperative on the agency. Such studies should be considered by the Administrator, if adduced in comments, but we do not inject them as a necessary condition of action.

This decision has had far-reaching consequences, as current point-source regulations continue to be set on the basis of the Best Available Control Technology (BACT), which is usually the most expensive.[14] While the federal BACT laws require regulators to show that control

---

[14] In practice, regulators set an emissions limit which the firm can achieve using whatever control technology it wants. However, usually these emissions quotas are set at the limit of current control technology and

technologies are cost-effective before requiring plants to install them, in practice the EPA usually satisfied this requirement by pointing to an existing plant that had successfully adopted the technology.[15]

This paper evaluates the welfare effects of the Clean Air Act Amendments of 1990. These Amendments mandated new monitoring, reporting, and emission requirements for the cement industry. The Amendments created a new class of emission restrictions governing hazardous air pollutants and volatile organic compounds. The EPA did not promulgate final requirements based on the Amendments for 12 years, but the industry anticipated these new standards, as engineering estimates of compliance costs appeared shortly after the passage of regulation. An engineering firm familiar with cement operations released a report detailing the changes necessary for compliance (Grossman, June 1993). It predicted that the new standards would lead to large capital and maintenence expenses, particulary with the control of nitrogen oxide emissions. During this time, the EPA also cited a number of engineering estimates of the cost of compliance in their technical guidance formulation.

Title V of the Amendments also mandated that all firms emitting significant quantities of pollutants would have to apply for operating permits, which very few plants in the cement industry had obtained before 1990. The permits require regular reporting on emissions, which necessitate the installation and maintenence of new monitoring equipment. The Amendments also required firms to draw up formal plans for compliance and undergo certification testing. By 1996, virtually all cement plants had applied for their permits, which they are required to renew every five years.

It should also be noted that the passing of the Amendments coincided with a renewed interest in regulation at the state level. One reason that the EPA's new regulations have taken so long to be promulgated is resistance from 26 states with their own interest in regulating the emissions of the cement industry. Grassroots opposition to the construction and operation of cement plants has grown in the last decade as well, with cement plants being at the top of several state-level environmental groups' list of concerns.[16]

---

require state-of-the-art control equipment. In a related point, the industry had managed to avoid attracting attention over its SO_2 emissions for decades in part because it could claim there was no effective control technology. As of 2004, this is no longer the case, and it is likely that the most expensive control technologies will be required to meet future requirements. See Weiss (2000) for more detail.

[15] My results below suggest some firms may find it profitable to become first-adopters of new, and costly, monitoring and control technology in order to raise barriers to entry.

[16] For example, see www.cementkiln.com and www.ichetucknee.org.

The timing of the implementation of the Amendments is critical to identifying the effects of the regulation in the data. My data runs from 1980 to 1998, covering the period before and after the passage of the Amendments, and ending before the adoption of the new emissions requirements. This means that any changes due to the Amendments during the 1990's are going to enter in through the sunk costs of entry, since there were no changes concerning the emissions of pollutants.[17] This is relevant because my analysis below indicates that the industry has experienced a positive shock to the technology of cement production, increasing the feasible size of large kilns and lowering the costs of production. I use the lag in timing between the passing of the Amendments and the implementation of the new emissions requirements to identify which changes in the cost structure are due to the regulations. One implication of this timing is that the EPA would have assigned an implicit cost of zero to the Amendments during the 1990's.[18]

3 Data

I collect data on the Portland cement industry from 1980 to 1999 using a number of different sources. I require market-level data on prices and quantities to estimate the demand curve for cement. The US Geological Survey collects establishment-level data for all the Portland cement producers in the US and publishes the results in their annual Minerals Yearbook.[19] The USGS aggregates establishment-level data into regional markets to protect the confidentiality of the respondants. The Minerals Yearbook contains the number of plants in each

---

[17] There is the possibility that firms pre-emptively installed control technology in anticipation of new regulations. However, this in unlikely given the uncertainty surrounding the timing and severity of the new emissions.

[18] There are several reasons to assume that all the changes due to the Amendments entered in through a shift in the distribution of sunk entry costs. For one, several industry insiders have indicated that the Amendments are the single largest and most important shift in the cost structure of the industry in the last two decades. Also, the estimation procedure generates reasonable cost structures under this assumption. The implied shift in sunk entry costs is roughly the same magnitude as the sum of the EPA's estimates of the capital costs of new equipment and several industry estimates of the costs of Title V compliance. In combination with the timing restriction, these out-of-sample figures give a check on the validity of the model. Lastly, it is possible to construct control groups in the sample by exploiting differences in state compliance with the Title V permits. However, the relevant state variables move slowly in this industry, and it is unlikely that the changes observed in the 1990's are due to other correlated unobservables outside of the scope of the Amendments.

[19] The Bureau of Mines had this responsibility prior to merging with the USGS in the 1990s. The data was collected by a mail survey, with a telephone follow-up to non-respondants. Typically the total coverage of the industry exceeded 90%; in some years, 100% response was indicated. The USGS attempted to fill in missing observations with data from other sources.

Table 2: Summary Statistics

Variable	Minimum	Mean	Maximum	Standard Deviation
Demand Data				
MARKETQ	186	2,835.84	10,262	1,565.34
PRICE	36.68	67.46	138.99	13.68
PLANTS	1	4.75	20	1.94
WAGE	20.14	31.72	44.34	4.33
COAL	15.88	26.64	42.33	8.13
ELECTRICITY	4.23	5.68	7.6	1.01
POPULATION	689,584	10,224,352	33,145,121	7,416,485
GAS	3.7	6.21	24.3	2.21
Production Data				
QUANTITY	177	699	2348	335
CAPACITY	196	797	2678	386

Demand data are from annual volumes of the USGS's Mineral Yearbook, 1981 to 1999. There are 517 observations in 27 regional markets. The unit for MARKETQ is thousands of tons per year, while PRICE is denoted in dollars per ton. WAGE is denoted in dollars per hour for skilled manufacturing workers, and taken from County Business Patterns. POPULATION is the total populations of the states covered by a regional market. The units are dollars per ton for COAL, dollars per kilowatt hour for ELECTRICITY, and dollars per thousand cubic feet for GAS. All prices are adjusted to 1996 constant dollars. The data on production and capacity are taken from the Portland Cement Association's annual Plant Information Summary, and cover 1980 to 1999. Units on QUANTITY and CAPACITY are in thousands of tons per year.

market and the quantity and prices of shipped cement. There is occassional irregular censoring of data to ensure the confidentiality of individual companies, although this affects only a small number of observations representing a low percentage of overall quantity. Usually the USGS merges a censored region into a larger region in subsequent years to facilitate complete reporting.

I collect data on electricity prices, coal prices, natural gas prices, and manufacturing wages to use as instruments in the demand curve estimation. The data for fuel and electricity prices are from the US Department of Energy's Energy Information Administration.[20] Natural gas and electricity prices are reported at the state level from 1981 to 1999. Coal prices are only available in a full series over that time span at the national average level. I collect skilled manufacturing wages at the state level from the US Census Bureau's County Business Patterns. All prices are adjusted to 1996 constant dollars.

Table 2 shows summary statistics for the demand data. Most markets are characterized by a small number of firms, with the median market contested by four firms. The size

---

[20] http://www.eia.doe.gov.

of the markets varies greatly across the sample: the smallest market is 2% the size of the largest market. Cost factors also varied substantially across markets, with Alaska and Hawaii generally being the most expensive markets to operate a cement facility. The demand for cement is highly dependent on construction, which is regional. I account for these market-specific factors in my analysis by adopting a fixed effect for each market. I also include year-specific dummy variables to account for national-level yearly shocks to cement demand.[21]

Data on the plant-level capacities are from the Portland Cement Association's annual Plant Information Summary (PIS) and cover 1980 to 1998. These trade association databooks have complete coverage of all cement producers in the United States, and give detailed information on grinding and kiln capacity. For each establishment, the PIS reports daily and annual plant capacities. I interpret the daily capacity to be a boilerplate rating, as determined by the manufacturer of the kiln, of how much the kiln produces in a given 24-hour period of operation. I assume the yearly capacity is how much they actually produced in that year. This assumption is supported by the fact that plants operate continuously in runs lasting most of the year. Maintenance is performed during a single shutdown period, generally a month in duration, in which the plant produces nothing. If the firms are assumed to run at perfect efficiency on the days that they operate, then the boilerplate rating multiplied by the length of a year gives the theoretical maximum that a plant could have produced. The yearly capacity numbers never achieve this bound and fluctuate from year to year. Additionally, the yearly numbers add up to the rough market-level quantities reported in the USGS data. Therefore, I take the reported annual capacity of the kiln to be the amount of cement that it actually produced in that year. I emphasize, however, that this quantity is not a fixed percentage of the theoretical maximum capacity. Firms still choose how long to operate their kilns before performing maintenence, and are subject to idiosyncratic shocks affecting the duration of the maintenence period. More productive firms have shorter maintenence periods and therefore can produce more in a given year than less productive firms. Given that firms are at the edge of their maximum productive capacity during the sample period, capacity choice is clearly the most important strategic decision firms have to make, but it should be emphasized that they still face a tradeoff between production and maintenence. The last two rows of Table 2 give the summary statistics for production and capacity levels.

To match the market-level demand data to the establishment data from the PIS, I combine some of the markets in the USGS data to form continuously-reported metamarkets. I

---

[21] The reference market for the dummy variables is Alabama in 1980.

then group all the plants into the appropriate metamarkets for every year of establishment data.[22] The production data consists of an unbalanced panel of 2233 observations.

4 Model

To evaluate the welfare effects of the Amendments, it is necessary to have a theoretical model that captures the salient features of the cement industry. The industry is characterized by simultaneous entry, exit, investment, and production decisions of a small number of firms in each market. The firms behave strategically and anticipate the future when making decisions. The structure within each regional market is primarily determined by the distribution of production capacities among active firms. Firms are also subject to idiosyncratic shocks which can influence the long-run structure of the industry through exit and investment behavior. I build on the work of Ericson and Pakes (1995), who provide an elegant theoretical framework of industry dynamics that can account for these features.

The fundamental idea of the model is that all of the economically-relevant characteristics of the firms in a market can be encoded into a state vector. Firms receive state-dependent revenues from a product market in each period, and can influence the evolution of the state vector through entry and exit and adjustments to their capacity. Equilibrium obtains when firms follow strategies designed to maximize the expected discounted present value of their stream of revenues given the expected strategies of their competitors.

I adapt this general framework to account for the specific features of the cement industry, where the basic building block is a regional homogenous-goods market in which capacity is the most important strategic variable. In each period, incumbents compete over quantities in this market, subject to a private productivity shock which shifts the marginal cost of production. Firms are partially capacity constrained, as they experience smoothly increasing marginal costs as production approaches their theoretical maximum capacity.

---

[22] The markets are: Alabama; Alaska, Hawaii, Oregon, and Washington; Arizona, Nevada, and New Mexico; Arkansas and Oklahoma; California north; California south; Colorado and Wyoming; Florida; Georgia and Tennessee; Idaho, Montana, and Utah; Illinois; Indiana; Iowa, Nebraska, and South Dakota; Kansas; Kentucky, Mississippi, North Carolina, and Louisiana; Maryland, Virginia, and West Virginia; Michigan and Wisconsin; Missouri; New York and Maine; Ohio; Pennsylvania East; Pennsylvania West; South Carolina; Texas north; and Texas south. The USGS provides ZIP code delineations when splitting a state into two markets.

In addition to active incumbents, there is a pool of short-lived potential entrants who must decide whether or not to enter, paying a privately-known sunk cost of entry if they decide to do so.[23] Exiting firms receive revenues from the product market before receiving a privately-known scrap payment and disappearing forever at the end of the period.

Entrants and active firms make capacity adjustment decisions based on the state vector and the expected strategies of their competitors. Capacity adjustments today change the capital stock tomorrow, with firms obtaining a private draw of adjustment costs in each period. The industry vector evolves over time as firms make entry, exit, and investment decisions.

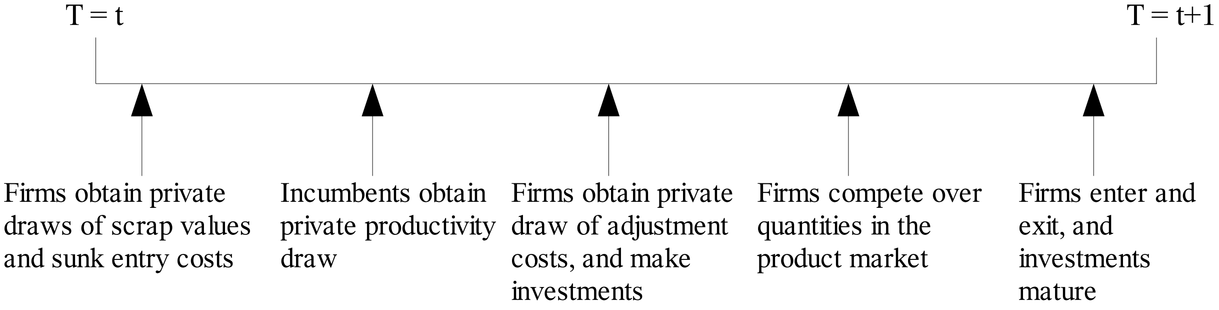
To keep the model tractible, I assume that firm strategies depend only on the current state vector, generating a Markov-perfect Nash equilibrium. Specifically, the MPNE consists of a set of mutual best-response strategies governing entry (for potential entrants), production and exit (for incumbents), and investment (for both entrants and incumbents). In the following sections, I describe each component of the model in detail before deriving the ex-ante value functions for potential entrants and incumbents. The value function will play a crucial role in the counterfactual simulations I use to evaluate the welfare costs of the Amendments.

State Space There are N players who operate in M geographically distinct markets. I restrict firms to behave independently across markets, and drop market-specific notation in what follows. Time is discrete and incumbent firms operate with an infinite horizon. Each market is fully described by a $N \times 1$ state vector, s_t, where s_{it} is the capacity of the ith firm at time t. I divide the firms into two groups: potential entrants ($s_{it} = 0$) and incumbents ($s_{it} > 0$). I assume that the potential entrants are short-lived; if a firm decides not to enter the market in one period, it gives up its slot and is replaced by a new potential entrant in the next period. Firms discount future profits at a constant rate of β. For notational simplicity, I abuse notation slightly in what follows; in the expectations written below that are conditional on s, I assume that this also includes all the information available to a firm in the market. So when a firm projects the expected future market state, they build in the expected effects of entry, exit, and investment policies of all the firms in the market.

---

[23] Conceptually, it is straightfoward to allow long-lived potential entrants, who wait for a low enough draw from the distribution of sunk entry costs before entering. This requires labelling the composition of firms within a market, which becomes an intractible computational problem in the counterfactual simulations, even if it were possible to identify the set of all possible entrants.

Figure 5: Timeline of Decisions



Timing The sequence of events in each period unfolds as follows:

- Incumbent firms receive a draw from the distribution of scrap values and make their exit decision. Potential entrants receive a draw from the distribution of entry values and make their entry decision.
- Incumbent firms and entrants receive a private draw on the fixed cost of capacity adjustment and make investment decisions.
- Firms receive a private productivity shock and compete over quantities in the product market.
- Incumbents chosing to leave the market exit and receive their scrap payment. Entrants pay their entry fee.
- The state vector adjusts as investments mature and firms enter and exit.

Note that firms who decide to exit produce in this period before leaving the market, and that changes in capacity take one period to be implemented. A second feature of the game's timing is that firms make production and investment decisions without knowing the actions of their competitors. Firms observe the state variable at the beginning of each period along with the entry, exit, investment, and production decisions of their rivals in the last period. Since shocks are drawn independently across periods from a common distribution, firms do not update their expectations of future behavior after observing the actions of their rivals.

Product Market Firms compete in quantities in a homogeneous goods product market. I assume that in each market firms face an identical linear demand curve:

$$\text{PRICE}(\text{MARKETQ}) = \alpha_0 - \alpha_1\text{MARKETQ} + \alpha_2\text{REGION} + \alpha_3\text{YEAR}, \tag{1}$$

where MARKETQ is the aggregate market quantity, and REGION and YEAR are a market- and year-specific demand shifters, respectively. Production costs consist of two parts: a constant marginal cost and an increasing function that binds as quantity approaches the capacity constraint. I assume that costs increase as the square of the percentage of capacity utilization, and parameterize both the penalty and the threshold at which the costs bind. Each firm also has a private productivity shock, SHOCK_i, drawn from a common distribution, that shifts the marginal cost of production in each period, giving rise to the following firm profit function:

$$\begin{aligned}\pi_i = \text{Q}_i(\text{PRICE}(\text{MARKETQ}) - \text{MC} + \text{SHOCK}_i)& \\ - 1(\text{UTILPCT}_i > \nu)\left[\text{CAPCOST}\cdot(\text{UTILPCT}_i - \nu)^2\right],&\end{aligned} \tag{2}$$

where Q_i is the firm's output quantity, ν is the threshold at which capacity costs bind, CAPCOST is the capacity penalty, UTILPCT_i is the fraction of capacity utilization, and SHOCK_i is the firm's private productivity shock. The second term accounts for the increasing costs associated with operating near maximum capacity, since firms have to cut into maintenence down time in order to expand production beyond a certain level.

The optimal quantities for each firm are described by the following equation:

$$\begin{aligned}\frac{\partial\pi_i}{\partial\text{Q}_i} = \text{PRICE}(\text{MARKETQ}) - \alpha_1\text{Q}_i - \text{MC} + \text{SHOCK}_i& \\ - 1(\text{UTILPCT}_i > \nu)\left[\frac{2\cdot\text{CAPCOST}}{\text{CAP}_i}(\text{UTILPCT}_i - \nu)\right] = 0.&\end{aligned} \tag{3}$$

Since the term SHOCK_i is private information, the appropriate solution concept for this strategic game is Bayesian Nash equilibrium. Analytic results are difficult to generalize to multiple firms due to the presence of firm-specific state variables entering the best-response functions, so I solve the system of equations numerically.[24]

---

[24] I form a vector of first-order conditions, f, take expectations, and minimize $f'f$ with respect to the vector of quantities, q. The resulting quantities are the BNE quantities, which are used by firms to predict the quantities of their competitors. After matching the observed quantities to the predicted values for each

Per-Period Payoff Function In addition to any profits or losses incurred in the product market, firms can make costly adjustments to their capacity levels. I allow for a fixed cost and constant marginal cost of adjustment, both for positive and negative changes in capacity. The fixed costs capture the idea that firms may have to face significant sunk costs, such as obtaining permits or doing feasibility studies, that accrue regardless of the size of the investment. Divestment sunk costs can be positive as the firm may encounter costs in order to shut down the kiln and dispose of related materials and components. On the other hand, a negative marginal cost of divestment encapsulates the idea that firms can receive revenues from selling their infrastructure. I model these adjustment costs as being private draws from a common, known distribution independent across time periods. This helps capture the idea that firms face varying fixed costs of investment that are dependent on unobservables.

The per-period payoff function is composed of several parts, depending on the firm's status as a new entrant, continuing incumbent, or exiting incumbent. Potential entrants who choose not to enter receive a payment of zero and are never eligible to enter in the future. New entrants do not recoup any profits from the product market; they only pay an entry fee and the costs of their initial investments. The per-period payoff function for new firms in the period they enter is:

$$-\mathrm{SUNK}_i - \mathrm{ADJPOS}_i - \mathrm{INVMCPOS} \cdot \mathrm{INV}_i^e - \mathrm{INVMCPOS2} \cdot (\mathrm{INV}_i^e)^2, \tag{4}$$

where SUNK_i is the entrant's draw from $G(\cdot)$, the distribution of sunk entry costs, INV_i^e is the level of initial investment, ADJPOS_i is the fixed cost of positive investment, INVMCPOS is the marginal cost of positive investment, and INVMCPOS2 is the squared marginal cost of positive investment. Incumbent firms that choose to continue their operations in a market make production and investment decisions, so they receive revenues and have to account for

---

firm, it is possible to invert Equation 3 to find the productivity shock. I posit that the BNE is unique in each market for two reasons. The first is that the marginal revenue of each firm is decreasing in the output of its rivals, which is the most substantive restriction in guaranteeing uniqueness in an asymmetric Cournot game, as discussed in Gaudet and Salant (1991). The second reason is that the solution produced by the numerical minimizer always gave the same answer for a wide range of starting values, including guesses on the boundaries of the parameter space. While not a proof of uniqueness, it is a necessary and easily verified condition.

investment decisions in that period. The per-period payoff function for these firms is:

$$\begin{aligned}\pi_i(s) - 1(\mathrm{INV}_i > 0)(\mathrm{ADJPOS}_i + \mathrm{INVMCPOS} \cdot \mathrm{INV}_i + \mathrm{INVMCPOS2} \cdot \mathrm{INV}_i^2) \\ - 1(\mathrm{INV}_i < 0)(\mathrm{ADJNEG}_i + \mathrm{INVMCNEG} \cdot \mathrm{INV}_i + \mathrm{INVMCNEG2} \cdot \mathrm{INV}_i^2), \quad (5)\end{aligned}$$

where ADJNEG_i is the firm's fixed cost of divestment, INVMCNEG is the marginal cost of divestment, and INVMCNEG2 is the squared marginal cost of divestment. As with positive adjustments, the fixed cost of divestment is also a private draw from a common distribution of negative adjustment costs. If an incumbent chooses to leave the market, it still produces in this period before obtaining a draw, SCRAP_i, from the distribution of scrap values, $F(\cdot)$:

$$\pi_i(s) + \mathrm{SCRAP}_i. \quad (6)$$

Together, these three equations span the payoffs that a firm can achieve in a given state.

Transition Probabilities The last ingredient of the model is the transition process between states. The probability of moving from one state of the system to another is a combination of all the paths that can lead to that state. The key assumption I make regarding these transitions is that a firm's capacity vector is always equal to last period's capacity plus last period's investment.[25] The probability of achieving a state depends on investment, entry, and exit. The probability of each element of a new state vector is the linear combination of three probabilities:

$$\begin{aligned} Pr(s_i \to s_i') = {} & 1(s_i > 0)(1 - Pr(i \text{ exits}|s_i))Pr(s_i'|s_i, \mathrm{INV}_i) \\ & +1(s_i > 0)\left[Pr(i \text{ exits}|s_i)Pr(j \text{ enters}|s_i, i \text{ exits}) \cdot \right. \\ & \left. Pr(s_i'|i \text{ exits}, j \text{ enters}, \mathrm{INV}_j)\right] \\ & +1(s_i = 0)Pr(j \text{ enters}|s_i)Pr(s_i'|j \text{ enters}, \mathrm{INV}_j). \end{aligned}$$

The probability of observing an element of a new state is conditional on whether an incumbent is currently active in that slot. If there is an incumbent, there are two possible ways of obtaining the new state: either the incumbent stays in the market and moves to the new state, or the incumbent exits and is replaced by an entrant at the new state. If there is no

---

[25] This abstracts away from depreciation processes, which do not appear to be significant in the data, and uncertainty in next period's capacity due to random completion times, etc. It is conceptually straightforward to incorporate these extensions in the model.

incumbent, the probability of observing the new state is equal to the probability of a new firm entering at that state. So for any one change in the state vector, I have to account for the entry, exit, and investment decisions of incumbents and potential entrants. To find the probability of the entire state vector shifting to another, I simply multiply out the individual probabilities of each element of the state vector:

$$Pr(s \rightarrow s') = \prod_{i=1}^{N} Pr(s_i' \rightarrow s_i). \tag{7}$$

It is important to note that these transitions are conditionally independent given the choice of s_i, which is going to depend on the actions of the other firms in equilibrium. Since I have a common, known distribution of entry and exit costs in each period, I can write out the probability of entry and exit in terms of their optimal entry and exit strategies:

$$\begin{aligned} Pr(\text{exit}|s_i) &= \int \Phi(s_i, \text{SCRAP}_i)\ dF(\text{SCRAP}_i) & (8) \\ Pr(\text{entry}|s_i) &= \int \Theta(s_i, \text{SUNK}_i)\ dG(\text{SUNK}_i), & (9) \end{aligned}$$

where $\Phi(s_i, \text{SCRAP}_i)$ is the exit rule for the firm given the current state and its draw from the scrap distribution. The analogous entry rule for potential entrants is $\Theta(s_i, \text{SUNK}_i)$. Under weak regularity conditions, Doraszelski and Satterthwaite (2004) prove that there always exists an MPNE in symmetric pure strategies where these policies will take the form of cutoff rules: if and only if the entry cost (scrap value) is below (above) a certain level, then the firm will enter (exit).[26] This boils down the expectation about competitors' entry and exit strategies to a probability that a competitor enters or exits in any given period. I revisit these policies after writing out the value functions describing the present values of both incumbents and potential entrants.

Equilibrium Concept In each time period player i makes entry, exit, production, and investment decisions, collectively denoted by Γ_i. Since the full set of dynamic Nash equilibria is unbounded and complex, I restrict the firms' strategies to be anonymous, symmetric, and Markovian. Therefore, I can write each firm's strategy, σ_{it}, as a mapping from states to

---

[26] In addition to private draws on entry and exit fixed costs, uniqueness and symmetry also requires that the firms have unique investment choices at each state. This is true in the current model since the payoff for investment is continuous and that firms are indifferent to investing or not on a set of measure zero, since the adjustment costs are random.

actions:

$$\sigma_{it} : S_{it} \to \Gamma_{it}.$$

Each firm's strategy maps the current state of the system into a vector of actions. Since the time horizon is infinite, payoffs are bounded, the discount factor, β, is positive and less than one, and firms have Markovian strategies, I drop the time subscript and write the value of being in state $s \in S$ recursively:

$$V_i(s|\sigma(s)) = u_i(\sigma(s)) + \beta \int V_i(s'|\sigma) dP(s'|\sigma(s), s),$$

where $\sigma(s)$ is the vector of firm strategies, $u_i(\sigma(s))$ is the per-period payoff function, and $P(\cdot)$ is the conditional probability distribution governing the transition between states. Markov perfect Nash equilibrium requires each firm's strategy profile to be optimal given the strategy profiles of its competitors:

$$V(s|\sigma_i^*, \sigma_{-i}) \geq V(s|\sigma_i', \sigma_{-i}), \tag{10}$$

for all s and any alternative strategy σ_i'. This equilibrium concept places significant structure on the optimal behavior of firms, which I exploit to construct my empirical estimator. I assume that this MPNE is unique.[27]

Value Functions Given the primitives of the model above, I can write down the ex-ante value functions for both the potential entrant and incumbent. These functions give the expected discounted present value, in dollars, of being at a given state vector. The value can be broken into two components: the per-period payoff function and the continuation value, which is the expected value of next period's state. For example, if there was no entry, exit, or investment, then the value of each state would simply be the expected discounted present value of obtaining the state-specific period payoffs in perpetuity. Firms use the value functions to find their optimal investment, entry, and exit policies. Each firm compares the marginal benefit of being at a new state against costs of achieving that new state when deciding on an investment strategy. Likewise, the firm will compare the draw from the distribution of scrap values against its continuation value when deciding whether or not to

---

[27] I conjecture that the addition of uncertainty at each step of the model (stage game, investment choice, entry and exit choice) helps purify the set of symmetric equilibria. Proving this conjecture is currently beyond the state of the art in this literature. An important contribution of Doraszelski and Satterthwaite is demonstrating the conditions under which a symmetric, anonoymous MPNE exists in pure strategies. Verifying their conditions with the current model is straightforward. However, the uniqueness of this equilibrium remains an important area for future research.

exit. The potential entrant makes a comparison of its draw from the distribution of sunk entry costs against the expected value if it enters.

I integrate out all of the private information in the per-period payoff function when writing out the value functions. To see why, consider the case of a firm making an investment decision conditional on a specific state vector, observable to all players in the market. That firm obtains several private draws from distributions governing shocks to marginal cost, adjustment costs, and the scrap value of exiting in that period. Its optimal strategy is going to depend on its expectation of its competitors strategies; however, its competitors are also forming strategies conditional on its own expected strategy. As every player forms strategies given the expected strategy of its competitors, I have to solve for the expected value function of each state, without conditioning on the specific shocks that an individual firm will experience if it actually operated at that state. Once these ex-ante value functions are solved out, each individual firm uses them to form expectations about the strategies of its competitors for any given state. However, it is worth emphasizing that the each firm will re-optimize its own decisions conditional on the shocks that it receives that period. Thus, the actual outcomes observed in equilibrium will vary substantially from the values predicted by the value function. This is trivially true for discrete actions such as entry and exit. However, if the game was repeated infinitely often the observed stream of payoffs accruing to a firm at a specific state would converge to the value predicted by the value function.

This distinction is not made in the baseline Ericson-Pakes model because there is no private information or uncertainty in the per-period payoff function. For example, in the canonical example of firms investing in product quality the only uncertainty enters in the transition function between states. Given that the ex-ante and ex-post value functions both integrate out the uncertainty arising from the transitions between states, these two functions will be the same. In the present model I have multiple sources of private information, some of which enter the per-period payoff function, which results in differences between the ex-ante and ex-post value functions.

A second point is also in order about why it is necessary to solve for the value function in the first place, given that I have recovered the policy functions and underlying primitives above, and could construct the value function directly. The reason is that the policy functions, and thus the value functions associated with them, are valid only for the specific set of primitives that I recovered in the estimation. To perform any counterfactuals, it is necessary to re-solve for the policy functions, and thus the value functions, for the new set of parameters, as firms will generically alter their optimal strategies in response to changes in the economic environment. If I was interested in questions that did not require changing the underlying primitives, then it would not be necessary to re-solve for the optimal policies analytically.[28]

I first consider the potential entrant, who simply checks the expected value of entering against the draw of entry costs it receives. Since I assumed that these potential entrants live for only one period, I do not have to solve the intertemporal waiting problem, where a firm with a high draw in this period may delay entering until it receives a more favorable draw in the future. Conditional on the current state and the draw from the sunk cost of entry, SUNK_i, the value function for potential entrants who decide to enter in the next period can be written as:

$$\begin{aligned} V_i^e(s, \mathrm{SUNK}_i) = \max_{\mathrm{INV}_i^e} \{ & -\mathrm{SUNK}_i - \mathrm{ADJPOS} - \mathrm{INVMCPOS} \cdot \mathrm{INV}_i^e \\ & -\mathrm{INVMCPOS2} \cdot (\mathrm{INV}_i^e)^2 + \beta E(V(s')|s) \} . \quad (11) \end{aligned}$$

Note that the value function for the entrant includes the optimal choice for an initial investment. The potential entrant is forward-looking and rational, so the expected value of entering accounts for the investments of other firms and their entry or exit decisions. Also note that solving for the optimal investment does not depend on SUNK_i, so firms solve for INV_i^e by finding the optimal investment conditional on entering. For a given state and optimal investment, there exists a draw from the sunk cost distribution such that a firm is indifferent between entering and not. Denoting the optimal investment conditional on entering as INV_i^{e*}, the draw at which a firm is indifferent is:

$$\overline{\mathrm{SUNK}}_i = -\mathrm{ADJPOS} - \mathrm{INVMCPOS} \cdot \mathrm{INV}_i^{e*} - \mathrm{INVMCPOS2} \cdot (\mathrm{INV}_i^{e*})^2 + \beta E(V(s')|s). \quad (12)$$

---

[28] If the policy functions and parameters are recovered precisely enough in the first stage, and the model is not misspecified, then solving for the MPNE policy functions would reproduce those recovered from the data.

Therefore, the value function for a potential entrant in state s is:[29]

$$V_i(s; s_i = 0) = \int_{-\infty}^{\overline{\text{SUNK}}_i} \big(-\text{SUNK}_i - \text{ADJPOS} - \text{INVMCPOS} \cdot \text{INV}_i^{e*} \\ - \text{INVMCPOS2} \cdot (\text{INV}_i^{e*})^2 + \beta E(V(s')|s)\big)\, dG(\text{SUNK}_i). \quad (13)$$

Revisiting the cutoff rule for entry described above, I assume $\Theta(s_i, \text{SUNK}_i)$ is an indicator function that is equal to one when the draw of sunk costs is lower than $\overline{\text{SUNK}}_i$ and zero otherwise.

The derivation of the value function for the incumbent firm is similar to the potential entrant, except it has two parts corresponding to whether or not the firm decides to exit the industry. The incumbent draws a value from the distribution of exit costs, SCRAP_i. If the firm decides to leave the market, it obtains a payoff of:

$$\pi_i(s) + \text{SCRAP}_i. \quad (14)$$

The value function for this firm is simply this payoff, as it receives no payments in the future. If a firm decides not to exit, it also receives the same revenues from the product market but remains active and does not receive SCRAP_i. The value of this state for a continuing firm also contains the expected discounted value of remaining in the market, so the value function can be written as:

$$V_i(s) = \max_{\text{INV}_i} \pi_i(s) - 1(\text{INV}_i > 0)(\text{ADJPOS} + \text{INVMCPOS} \cdot \text{INV}_i + \text{INVMCPOS2} \cdot \text{INV}_i^2) \\ - 1(\text{INV}_i < 0)(\text{ADJNEG} + \text{INVMCNEG} \cdot \text{INV}_i + \text{INVMCNEG2} \cdot \text{INV}_i^2) + \beta E(V(s')|s). \quad (15)$$

Again, the optimal level of investment, INV_i^*, is independent of the draw from the exit cost distribution, so firms solve for this value conditional on staying in the market. As in the potential entrant case, the incumbent also faces a similar cutoff value of SCRAP_i that makes that firm indifferent between continuing and exiting. Solving out equations 14 and 15 for

---

[29] This is equivalent to the amount of money that a firm would be willing to pay to be a potential entrant for one period.

this cutoff value, I find:

$$\overline{\text{SCRAP}}_i = -1(\text{INV}_i^* > 0)(\text{ADJPOS} + \text{INVMCPOS} \cdot \text{INV}_i^* + \text{INVMCPOS2} \cdot (\text{INV}_i^*)^2) \\ - 1(\text{INV}_i^* < 0)(\text{ADJNEG} + \text{INVMCNEG} \cdot \text{INV}_i^* + \text{INVMCNEG2} \cdot (\text{INV}_i^*)^2) + \beta E(V(s')|s). \tag{16}$$

Therefore, I can write the value function for an incumbent firm as:

$$V_i(s) = \int_{-\infty}^{\overline{\text{SCRAP}}_i} \overline{\text{SCRAP}}_i \, dF(\text{SCRAP}_i) + \int_{\overline{\text{SCRAP}}_i}^{\infty} \{\pi_i(s) + \text{SCRAP}_i\} \, dF(\text{SCRAP}_i). \tag{17}$$

Analogously to the entry cutoff rule, I assume the exit cutoff rule, $\Phi(s_i, \text{SCRAP}_i)$, is an indicator function that is equal to one when the draw of the sunk cost of exit is higher than $\overline{\text{SCRAP}}_i$ and zero otherwise.

5 Empirical strategy

5.1 Overview

The empirical goal of this paper is to estimate all of the parameters in the theoretical model described above. I follow the two-step empirical strategy laid out in Bajari, Benkard, and Levin (BBL) (2003). In the first step, I recover the policy functions governing entry, exit, and investment along with the product market profit function. In the second step, I take these functions and impose the restrictions of the MPNE to recover the dynamic parameters governing the costs of capacity adjustment and exit. This then allows me to simulate the value of a new firm entering the market, which can be used to recover the distribution of the sunk costs of entry.

Before getting into the details, it is useful to consider the mapping between the theoretical model above and the empirical estimates below. In the first step the relevant empirical objects that I need to recover are the parameters of the product market profit function and the policy functions that describe what actions the firm will undertake at any state. The model specifies a specific functional form for the parameters that enter into the product market profit function; in this case it is straightforward to write estimators consistent with the underlying assumptions. The theoretical model also provides some guidance for the form of the estimators of the policy functions. Both of the entry and exit decisions are cutoff rules, where a firm will enter (exit) when a private draw from a common distribution is low (high) enough, compared against a dollar amount that is a function of the state variables. This suggests that to recover the policy functions governing entry and exit, I should fit some function of the observable state variables, such as the summed capacity of competitors in the market, against the observed probability of entry and exit. As I assume that these distributions of sunk entry and scrap values are normally distributed, this implies that a probit model of entry and exit is going to be consistent with the underlying theoretical model, given that I characterize the influence of the state variables flexibly enough.

The last empirical object in the first step is the policy function governing firm investment. The theoretical model suggests that the empirical policy function should be a function of the state variables and should be flexible enough to account for lumpy investment behavior. One model that satisfies both of these requirements is the (S, s) rule of investment, introduced by Scarf (1959), where firms tolerate deviation from their optimal level of capacity due to fixed adjustment costs.[30] In the language of the (S, s) rule, firms have a target level bounded on either side by an adjustment band, both of which can be functions of observable variables. When the actual level of capacity hits one of the bands, the firm will make an adjustment to the target level. The target level and bands are only observed when the firm makes adjustments, and are flexibly parameterized to be functions of the underlying state variables. This model also nests the model of continuous investment, and is thus quite flexible in its ability to capture a range of investment behavior.

5.2 First-Stage Estimates

The first step has two goals: recover all parameters independent of a specific dynamic model, and describe the behavior of firms at every state. This means recovering the parameters of the demand curve, production costs, and the policy functions governing firm investment, entry, and exit, in that order. The period profit function from the product market, encompassing the demand curve and production function, is independent from any dynamic considerations in my model and can be estimated as stand-alone objects. The policy functions describe

---

[30] Deriving this rule as the explicit solution to an optimization problem is involved—see Hall and Rust (2000) for an example of the optimality of this rule in an inventory setting.

the empirical behavior of the firms at any given state vector. Under the assumption that the firms in the data play the same equilibrium, these observed policy functions have to be consistent with firms maximizing their outcomes under the theoretical model. In the next step, I find the parameters that rationalize these observed policies as optimal under the model's MPNE. Given a sparse enough state space, the optimal estimator in this first step for the policy functions would be a simple nonparametric description of what the firm does, on average, at every state. While the size of the state space in this industry precludes a completely nonparametric approach, I use flexible methods to characterize the firm behavior, with functional forms suggested by the theoretical model where possible. The first step is estimating the demand curve.

Demand Curve I use a static demand system in my model, so I can recover these parameters from the USGS market-level data independently of any dynamic considerations. I form the following moments:

$$m_1(\alpha) = (nT)^{-1} \sum_{i=1}^{N} \sum_{t=1}^{T} Z'_{it}(\mathrm{PRICE}_{it} - \alpha_0 - \alpha_1 \mathrm{MARKETQ}_{it} + \alpha_2 \mathrm{REGION}_i + \alpha_3 \mathrm{YEAR}_t), \tag{18}$$

where Z_{it} is a vector formed by a constant and cost-shifters: coal prices, gas prices, electricity rates, and wage rates, and REGION and YEAR are vectors of dummy variables. For notational clarity I denote the vector of parameters associated with the region and year fixed effects by α_2 and α_3, respectively. These capture demand shifts for cement and unobserved heterogeneity among markets. The reference level for the dummy variables is Alabama in 1980.

Production Parameters The predicted quantities for each firm in each market, conditional on the vector of production costs, $\hat{Q}_{it}(\alpha)$, are defined by the system of equations in Equation 3. I form a vector of moments from the gradient vector of the difference between the actual and predicted quantities. There are six production parameters: CAPCOST, MC, the level at which capacity costs begin to bind (ν), and late period dummy shifters for each, which generate six associated sample moment conditions:

$$m_2(\alpha) = (nT)^{-1} \sum_{i=1}^{N} \sum_{j=1}^{T} \nabla_\alpha (Q_{it} - \hat{Q}_{it}(\alpha)). \tag{19}$$

I restrict the threshold at which capacity costs bind to be between 0 and 1 with a logit transformation: $\nu = \exp(\tilde{\nu})/(1.0 + \exp(\tilde{\nu}))$. Note that it is possible to back out the productivity shock, SHOCK_i, from the observed quantities and their predicted counterparts, conditional on the estimated parameters, using an inversion of the first-order condition for optimal pricing. I condition on this shock in the investment and exit policies, which I estimate next.[31]

Investment Policy Function I follow Attanasio's (2000) model of the (S, s) rule, with the exception that I only model firms with positive capacity levels at the start and end of each period, treating the entry and exit process separately. This is acceptable in this context because I am only interested in what the investment behavior of a firm will be given a specific state. Firms have a target level of capacity that they adjust to when they make an investment:[32]

$$\text{TARGET}_{it} = \alpha_4' s_1(\text{CAP}_{it}) + \alpha_5' s_2(\text{SUMCAP}_{-it}) + \alpha_6 \text{SHOCK}_{it} + u_{it}^d \tag{20}$$

where the desired level of capacity is a function of the firm's own capacity, the sum of its competitors capacities (SUMCAP_{-it}), its productivity shock from the product market, and a mean zero error term. The functions s_1 and s_2 are approximated using cubic B-splines.[33] For notational simplicity I again denote the vector of parameters associated with these functions as α_4 and α_5.

The critical aspect of the (S, s) rule that generates lumpy investment behavior is that firms only adjust CAP_{it} to TARGET_{it} when it is sufficiently far from the desired level. I model this type of adjustment behavior by assuming that there are upper and lower "bands" which dictate when the firm will make an adjustment. As soon as the actual level of capacity is above the upper band or below the lower band, the firm adjusts to its target level. These

---

[31] Cooper and Haltiwanger (2000) have shown the importance of idiosyncratic shocks as a determinant of investment behavior. See Olley and Pakes (1996), Levinsohn and Petrin (2003), and Ackerberg and Caves (2004) for a recent strand of the literature dealing with the empirical implications and identification of productivity shocks on firm behavior.

[32] Note that this desired level does not necessarily have to coincide with the optimal level in the absence of adjustment costs. The reason for this is that firms may account for the interaction of a depreciation process with adjustment costs and invest to a higher level than in the case without adjustment costs. I do not characterize the optimal level in the absence of adjustment since it does not enter my estimation directly. It is sufficient that the data reveals the level firms adjust to given there are adjustment costs.

[33] B-splines are local polynomials which provide a parsimonious and numerically attractive method for approximating nonlinear functions. I used an order 4 (cubic) B-spline defined over a uniformly spaced knot vector with four interior knots. For a brief discussion of the B-splines and their construction, see Appendix A.1. No constants are used, and one of the B-splines is dropped in the second function to avoid collinearity.

bounds are assumed to be a symmetric function of the same state variables as the target and a mean-zero error term,

$$\text{BAND}_{it} = \text{TARGET}_{it} \pm \exp(\alpha_7' s_1(\text{CAP}_{it}) + \alpha_8' s_2(\text{SUMCAP}_{-it}) + \alpha_9 \text{SHOCK}_{it} + u_{it}^b) \quad (21)$$

This specification ensures that the desired level of adjustment is always in between the bands. This model also nests a model of continuous adjustment in the limit as the bands go to zero. I assume that the residuals in the bands are iid normal with zero mean and equal variance, and are independent of the error in the target.

To derive the likelihood function of observed investments, it is necessary to consider three cases: positive, negative, and no change in capacity. In the derivation that follows, I assume that when firms make adjustments that they reveal both the size of the band and the desired target level.[34] Therefore, the likelihood of a firm making a positive or negative adjustment is simply the joint probability of observing the band and target level:

$$f(u_{it}^d, u_{it}^b) = f(u_{it}^d) f(u_{it}^b) = f(\text{CAP}_{it} - \alpha^T x_{it}^T) f(\log(|\Delta\text{CAP}_{it}|) - \alpha^B x_{it}^B), \quad (22)$$

where ΔCAP_{it} is the change in capacity, and I have economized on notation by collapsing the parameters and covariates in the target and band to $\alpha^T x_{it}^T$ and $\alpha^B x_{it}^B$, respectively. This probability is the product of two normal probabilities due to the independence of the errors in the band and target.

The likelihood of observing no change in capacity is slightly more complicated since I do not observe either the target or band in that period. The probability of observing no change

---

[34] It is a straightforward extension to estimate the model assuming that the size of the band is unobservable.

is:

$$
\begin{aligned}
& Pr(\text{BAND}_{it}^{lower} < \text{CAP}_{it} < \text{BAND}_{it}^{upper}) \\
= \; & Pr\left(\text{TARGET}_{it} - \exp(\alpha^B x_{it}^B + u_{it}^b) < \text{CAP}_{it} < \text{TARGET}_{it} + \exp(\alpha^B x_{it}^B + u_{it}^b)\right) \\
= \; & \int Pr\left(\text{TARGET}_{it} - \exp(\cdot) < \text{CAP}_{it} < \text{TARGET}_{it} + \exp(\cdot)|u_{it}^b\right) dF(u_{it}^b) \\
= \; & \int Pr\left(\alpha^T x_{it}^T + u_{it}^d - \exp(\cdot) < \text{CAP}_{it} < \alpha^T x_{it}^T + u_{it}^d + \exp(\cdot)|u_{it}^b\right) dF(u_{it}^b) \\
= \; & \int \left(\int_{\psi_1}^{\psi_2} dF(u_{it}^d) \right) dF(u_{it}^b) \\
= \; & \int \left[F(\psi_2) - F(\psi_1)\right] dF(u_{it}^b),
\end{aligned}
$$

where $\psi_1 = \text{CAP}_{it} - \alpha^T x_{it}^T + \exp(\cdot|u_{it}^b)$, $\psi_2 = \text{CAP}_{it} - \alpha^T x_{it}^T - \exp(\cdot|u_{it}^b)$, and $F(\cdot)$ is the normal cumulative distribution function. The integral in the last term is easily computed using Gauss-Hermite quadrature.

I estimate the policy function parameters in a two-step procedure. Since I assume that the change in capacity reveals the size of the band, I use a first-stage OLS estimator to recover initial guesses for α in Equations 20 and 21 above. I use these parameters as starting values in a GMM estimator formed from the score vector of the log-likelihood function derived above:

$$m_3(\alpha) = (n(T-1))^{-1} \sum_{i=1}^{N} \sum_{j=2}^{T} \nabla_\alpha \log L(\alpha). \tag{23}$$

Entry and Exit Policy Functions I estimate entry and exit policies conditional on the state vector. As discussed above, I parameterize these probabilities with a probit model. Explanatory variables in both estimations are a constant, the sum of competitors' capacities, and a dummy variable for before and after 1990. I add the firm's capacity and productivity shock to the exit equation. I denote the moments corresponding to the exit probit as $m_4(\alpha)$.

Standard Errors The motivation for using GMM to estimate these initial stages is that I have to correct the variance matrix to account for error introduced by using the results of one estimation as inputs into the next stage. Fortunately, there is a relatively simple and straightforward method to do this, starting with the consistent but inefficient estimates obtained by running each stage separately. Following Newey and McFadden (1994), I stack the moments $m_1(\alpha)$, $m_2(\alpha)$, $m_3(\alpha)$ and $m_4(\alpha)$ and form the following one-step estimator:

$$\tilde{\alpha} = \bar{\alpha} - (\bar{G}'\hat{W}\bar{G})^{-1}\bar{G}'\hat{W}\hat{g}_n(\bar{\alpha}), \tag{24}$$

where $\hat{g}_n(\bar{\alpha})$ is the stacked vector of moments evaluated at $\bar{\alpha}$, an initial parameter vector found by estimating each stage above separately, $\bar{G}$ is a consistent estimator of plim $[\nabla_\alpha \hat{g}_n(\alpha_o)]$, and $\bar{W}$ is a consistent estimator of the inverse variance matrix. I use an efficient weighting matrix, $\bar{\Omega}$, to ensure that $\tilde{\alpha}$ has the same asymptotic variance as the full (iterated) GMM estimator with optimal efficient matrix.[35] Once I have obtained $\tilde{\alpha}$ using the one-step formula, I find a consistent estimate of the covariance matrix using $(G'WG)^{-1}$, where G and W are evaluated at $\tilde{\alpha}$.

5.3 Second-Stage Estimates

The first step has provided functions that describe both how the state vector evolves over time and what product market profits are at each state. The second step is concerned with finding parameters that make these observed policy functions optimal, given the underlying theoretical model. To see how this works, consider the following.

Given a starting state configuration, I simulate the evolution of the state vector forward 100 periods, far enough in the future that payoffs from that period have a very low discounted present value. I update the state vector from period to period by reading off the various policy functions. For example, if the slot is currently empty, then I draw from $U[0, 1]$ and compare it to the probability given the the entry probit conditional on the current state. If the draw is low enough, then the firm makes the appropriate investment, as described by the (S, s) rule, and becomes an active firm at that capacity in the next period. By collecting the actions of the firm through time, I can calculate the present-value payoffs to that path for a given set of parameters. By permuting the policy functions a little bit I generate different paths and different present-value payoffs for a given parameter vector. The key insight of this estimator is that the observed policy functions were generated by profit-maximizing firms

---

[35] There is one complication due to the fact that my moments are defined over data series of differing lengths. Denoting the subvector of moments defined over data set j as $g_j(z_i, \bar{\alpha})$, I construct a block-diagonal covariance matrix, $\bar{\Omega} = \bar{\varpi} \otimes I_3$, where each element of $\bar{\varpi}$ is $\bar{\varpi}_j = \sum_{i=1}^{n_j} g_j(z_i, \bar{\alpha}) g_j(z_i, \bar{\alpha})'$. Similarly, I evaluate $\bar{G}$ piecewise with its sample analogue: $n_j^{-1} \sum_{i=1}^{n_j} \nabla_\alpha g_j(z_i, \bar{\alpha})$. This matrix of derivatives is lower block triangular, as each successive stage has more parameters.

who chose the path with the highest expected discounted stream of payoffs. Therefore, at the true parameters, the payoffs generated by the observed policies should be greater than those generated by any other set of policies. This intuition is the heart of the second step, where I recover of the fixed and variable costs of investment and the distributions of sunk entry costs and exit scrap values. We recover these parameter in two substeps, obtaining estimates of investment costs and the distribution of scrap values before estimating the distribution of sunk entry costs.

5.3.1 Investment Parameters and Distribution of Scrap Values

To derive the estimator for investment costs and the distribution of scrap values, recall the firm's optimal decision, written recursively:

$$\max_{\sigma_i} u_i(\sigma(s), s) + \beta \int V_i(s'|\sigma) dP(s'|\sigma(s), s), \tag{25}$$

Note that, given the parameters estimated in the first step described above, I can decompose u_i into a linear function of its known and unknown components:

$$\begin{aligned} u_i = \pi_i &- 1(x > 0)(\text{ADJPOS} + \text{INVMCPOS} \cdot \text{INV}_i + \text{INVMCPOS2} \cdot (\text{INV}_i)^2) \\ &- 1(x < 0)(\text{ADJNEG} + \text{INVMCNEG} \cdot \text{INV}_i + \text{INVMCNEG2} \cdot (\text{INV}_i)^2) \\ &+ 1(\text{i exits})\text{SCRAP}_i, \end{aligned} \tag{26}$$

where the per-period payoff function, π_i, capacity adjustment, and exit decision have been recovered in previous steps. The unknowns are the costs of capacity adjustment and the distribution of scrap values received upon exit from a market. The unknown parameters, denoted by the vector α, enter linearly into the payoffs of the firm in the current period and all future periods. It is therefore possible to decompose the value function into the vector of parameters and the vector of expected discounted payoffs and actions, $W(s_o; \sigma_i, \sigma_{-i})$:

$$W(s_o; \sigma_i, \sigma_{-i}) = E_{\sigma_i, \sigma_{-i}|s_o} \sum_{t=0}^{\infty} \beta^t \zeta(s_{it}), \tag{27}$$

where $\zeta(s_i)$ is the vector of functions corresponding to the dynamic parameters:

$$\zeta(s_i) = \{\pi_i, -1(\text{INV}_i > 0), -1(\text{INV}_i > 0)\text{INV}_i, -1(\text{INV}_i > 0)\text{INV}_i^2, \\ -1(\text{INV}_i < 0), -1(\text{INV}_i < 0)\text{INV}_i, -1(\text{INV}_i < 0)\text{INV}_i^2, 1(\text{i exits})\}. \quad (28)$$

Note that the α vector contains a 1 in the first position, as the profits from the per-period payoff function enter in for each state irrespective of the unknown parameters. I impose the Markov perfect equilibrium condition (see Equation 10) for all alternative policies σ' to obtain:

$$W(s_o; \sigma_i, \sigma_{-i}) \cdot \alpha \geq W(s_o; \sigma_i', \sigma_{-i}) \cdot \alpha, \quad (29)$$

where the value function has been replaced by the explicit sum defined in Equation 27. At the true parameters the above relation should hold for all alternative policies. Exploiting the linearity of the unknown parameters, I can rewrite the above equation in terms of profitable deviations from the optimal policy:

$$g(x, \alpha) = [W(s; \sigma_i', \sigma_{-i}) - W(s; \sigma_i, \sigma_{-i})] \cdot \alpha. \quad (30)$$

Intuitively, I want to find parameters such that profitable deviations from the optimal policies are minimized. Formally, I draw alternative policies from a distribution H over all policies to generate a set of n_k inequalities, X_k. The true parameter minimizes:

$$\min_{\alpha} \int 1\left(g(X_k, \alpha) > 0\right) g(X_k, \alpha)^2 dH(X_k). \quad (31)$$

To form the estimator, I replace the above with its sample counterpart:

$$Q_n(\alpha) = \frac{1}{n_k} \sum_{i=1}^{n_k} 1\left(g(X_{ki}, \alpha) > 0\right) g(X_{ki}, \alpha)^2. \quad (32)$$

Implementing this estimator proceeds in two separate steps. In the first step, I find W for both the observed and alternative policies.[36] I generate the alternative policies by adding noise to the observed policy functions. For example, to permute the exit policy function I add an error drawn from the standard normal to the terms inside the exit probit. I generate many W to find the terms in 32. The linearity of the unknown parameters becomes useful during the minimization, as I do not have recompute separate outcome paths for each set of

---

[36] See the Appendix for a detailed discussion of computing W.

parameters. Note that the function is not trivially minimized at zero because the profits from the product market enter in each time period. Due to potential flat spots in the objective function, I use the Laplace-type estimator of Chernozhukov and Hong (2003). This estimator is robust to non-smooth functions and also has the nice feature of jointly estimating the mean and variance of the unknown parameters.

5.3.2 Distribution Sunk Entry Costs

Having recovered the policy functions, which describe how the firm will act at each state, and the underlying primitives of the model, which quantify the costs and benefits of those actions, it is possible to find the distribution of sunk costs. Knowing how the firm will act if it enters, along with stream of revenues associated with those behaviors, allows me to compute the expected value of entering a market. If a firm does not enter when these expected profits are positive, it must be because it received a sufficiently large draw on sunk entry costs to make it unprofitable to do so. By matching the cumulative distribution of the sunk costs to the predicted probability of entry I can recover the distribution of sunk costs. Formally, the value of entering at a state is:

$$\begin{aligned} V_i^e(s, \mathrm{SUNK}_i) = \max_{\mathrm{INV}_i^e} \{ & -\mathrm{SUNK}_i - \mathrm{ADJPOS} - \mathrm{INVMCPOS} \cdot \mathrm{INV}_i^e \\ & -\mathrm{INVMCPOS2} \cdot (\mathrm{INV}_i^e)^2 + \beta E(V(s')|s) \} . \end{aligned} \tag{33}$$

The optimal investment is given by the policy function, and I have all the parameters that enter the initial outlay and future stream of revenues save for SUNK_i. Recalling Equation 16, the firm will enter the market when its draw is lower than the value of entering the market, $EV^e(s)$, as defined by the terms to the right of SUNK_i in the above equation. As in the recovery of the dynamic parameters, I simulate many forward paths of possible outcomes given the firm entered. Averaging over these paths gives the expected value of entry, which I then can match against the observed rates of entry at different states. Formally, the probability of entering the market is the probability of receiving a draw that is less than the value of entry:

$$Pr(\mathrm{SUNK}_i \leq EV^e(s)) = G(EV^e(s); \mu_G, \sigma_G^2), \tag{34}$$

where $G(\cdot)$ is the cumulative distribution function of sunk entry costs. Intuitively, I match the observed probability of entry to the probability of a firm receiving a sufficiently small draw from the distribution of sunk costs to make entry profitable. The left-hand side is given by the entry probit. I simulate $EV^e(s)$ for NS different states and match $G(EV^e(s))$ at those values to the observed probability of entry:

$$\min_{\{\mu_G,\sigma_G^2\}} (NS)^{-1} \sum_{i}^{NS} \left[Pr(\text{entry}|s) - G\left(EV^e(s)\right)\right]^2 . \tag{35}$$

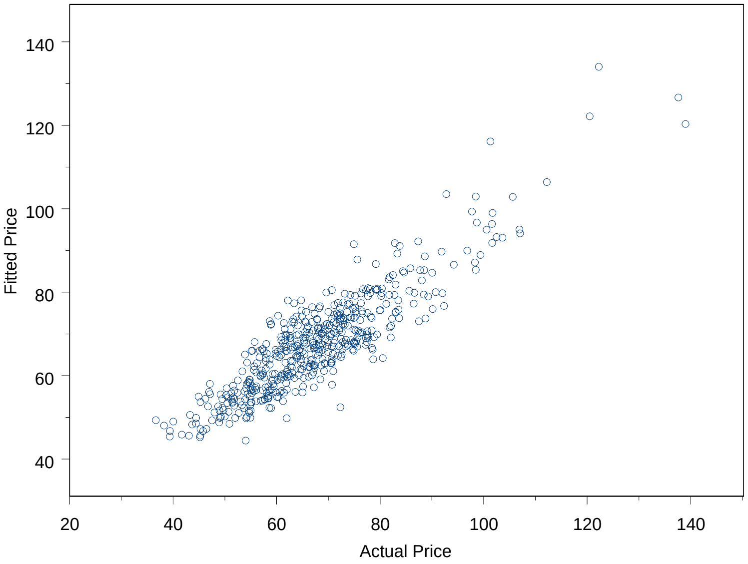
I recover the parameters of the distribution of sunk entry costs before and after the 1990 Amendments.

With the conclusion of this estimation stage, I have recovered all of the parameters of the underlying theoretical model. To summarize, I have used a two-step estimator to recover parameters consistent with an underlying dynamic model of oligopoly. In the first stage, I estimate many parameters without appealing to a specific dynamic model. Under the assumption that firms are maximizing their payoffs, the observed actions of the firms are maps from states to optimal actions. This stage encompasses the recovery of the demand curve, the production cost parameters, and the policy functions governing investment, entry, and exit. Once this stage is complete, I have enough parameters and policy functions in hand to simulate how the firm will behave in the future and what revenues it will receive from the product market. I exploit the optimality of observed policy functions under the assumption of Markov-perfect Nash equilibrium to recover the fixed and variable costs of capacity adjustment and the scrap value of leaving the industry. Once I have recovered the dynamic parameters, I simulate the expected value of entering a market at any given state. I match the CDF of the distribution of sunk costs at these expected values to the observed probability of entry to obtain an estimte of the distribution of sunk entry costs. Next, I present the results of these estimations before making use of the underlying model to perform the welfare calculations and policy counterfactuals.

6 Empirical Results

To faciliate identifying the welfare effects of the Amendments, I estimate the parameters before and after 1990. I follow the empirical strategy laid out above, first estimating the demand curve, production costs, the policy functions governing investment, entry, and exit. Then, imposing the equilibrium condition to the observed policies, I recover the costs of

Figure 6: Pricing Fits



capacity adjustment, the distribution of scrap values, and the distribution of sunk entry costs.

First, I use market-level data on prices and quantities to determine the parameters in Equation 1. The results are presented in Table 10 in the Appendix. The pricing equation fits the observed prices very well, as shown in Figure 6. The dummy variables account for unobservable heterogeneity among the markets; the REGION parameters pick up market-specific demand differences and the YEAR dummy proxies for shocks to cement demand across years. The coefficients have the expected signs, with markets like Alaska and Hawaii having the highest prices. The demand levels in the late 1990's are higher than anything since the early 1980's, reflecting a large boom driven by higher construction spending.

Having estimated the demand curve, I recover the production cost parameters using the supply-side pricing relation in Equation 3. I estimate six parameters: marginal cost, capacity cost, the capacity binding level, and post-1990 dummies for each. The results are shown in Table 3. The estimates indicate that capacity costs become important as firms increase production beyond 87% of their boilerplate capacity. Once firms cross this threshold

Table 3: Production Function Estimation Results

Parameter	Coefficient	Standard Deviation	99% Confidence Interval
CAPCOST	3,922,283	7,814	[3,908,038, 3,939,277]
CAPCOST DUMMY	-2,688,771	8,287	[-2,701,804, -2,666,288]
BINDING LEVEL (ν)	1.883	0.002	[1.881, 1.889]
BINDING LEVEL DUMMY	0.036	0.001	[0.035, 0.037]
MC	42.086	0.03	[42.006, 42.098]
MC DUMMY	-1.976	0.005	[-1.986, -1.963]

This table reports the estimated parameters for the production function. The binding threshold at which the capacity costs become important is bounded to $[0, 1]$ by estimating a logit probability: $\nu = \exp(\tilde{\nu})/\,(1.0 + \exp(\tilde{\nu}))$. At the estimated value of 1.88, this implies that capacity costs start to bind at an approximately 87% utilization rate. All dummy variables refer to period after the passage of the 1990 Amendments.

Table 4: Implied Prices, Revenues, Shocks, and Profits

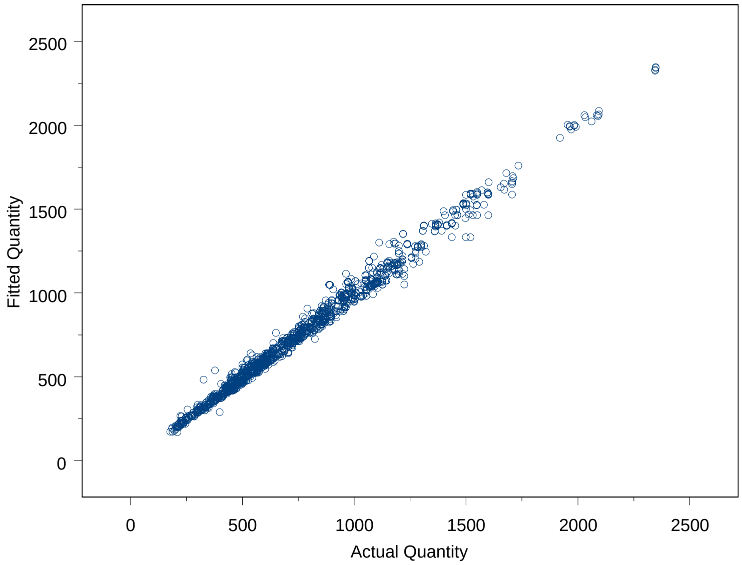
Variable	Min	Mean	Max	Standard Deviation
Price	40	62	102	13
Revenues	9,819	41,938	147,071	19,389
Shock	-3,995	90	8,017	485
Profit	-5,041	13,327	60,470	9,372
Margin	-0.39	0.31	0.60	0.15

This table reports the summary statistics of the implied prices, shocks, revenues, and profits for every firm in my sample at the estimated demand and production parameters. Margin is the implied profit margin calculated as profits divided by revenues.

they experience large, linearly increasing marginal costs as they cut into the normal period of maintenence downtime. The penalty for cutting out your maintenence is significant, preventing most producers from exceeding 90% of their stated production capacity in most periods. Figure 7 shows the overall fit of the predicted production levels against the actual data. The model does a very good job of capturing the production decisions of the firms in this industry.

As a check on the parameters of this model, I compute the market price, revenues, and profit margin for every firm in my sample. The summary statistics for these values are shown in Table 4. The prices are well within the range seen in the data, with the average firm grossing revenues slightly over \$40M a year. Profits average around \$13M a year, which is roughly a 30% profit margin. This is a plausible gross return, as public financials for major cement producer Lafarge North America report an 8% average post-tax profit margin for

Figure 7: Production fits



the four-year period 2000-2003.[37] These findings foreshadow my results below, in that there must be extensive sunk costs in order to sustain such high profit margins. However, before uncovering these costs directly, I first have to describe how firms invest, enter, and exit.

As described previously, I model the investment policy function as an (S, s) rule. Under the assumption that the bandwidth and target level are observable when a firm makes an adjustment, it is possible to obtain consistent estimates for the parameters in Equations 20 and 21 using a simple OLS regression. The bandwidth is determined by a regression of shocks, own capacity, and the capacity of a firm's competitors on the size of the change. The target level coefficients are determined by regressing the same state variables on the post-adjustment capacity. In order to estimate this policy function as flexibly as possible I use B-splines as basis functions for the capacity variables.[38] I use the OLS resuls as starting values for the full maximum likelihood estimator derived above. The results are presented in Table 5.

Interpreting the B-spline coefficients is difficult, since the approximation of the rela-

---

[37] Sales and profit data are from Hoover's Online financial fact sheet for Lafarge North America, 2000-2003. http://www.hoovers.com.

[38] See the Appendix A.1 for details.

Table 5: Policy Function Results

Parameter	Mean	Standard Deviation	95% Confidence Interval
BAND SUMCAP B-spline 1	5.444	1.236	[3.826, 8.376]
BAND SUMCAP B-spline 2	6.239	1.367	[4.068, 8.471]
BAND SUMCAP B-spline 3	6.19	1.551	[4.431, 9.49]
BAND SUMCAP B-spline 4	5.871	1.721	[2.493, 9.008]
BAND SUMCAP B-spline 5	5.99	2.174	[3.033, 8.697]
BAND SUMCAP B-spline 6	8.519	2.852	[5.466, 14.039]
BAND CAP B-spline 1	-3.016	1.769	[-4.77, -0.174]
BAND CAP B-spline 2	-2.485	1.398	[-4.632, -0.448]
BAND CAP B-spline 3	-2.373	1.155	[-4.49, -0.392]
BAND CAP B-spline 4	-0.803	1.664	[-2.894, 2.043]
BAND CAP B-spline 5	-2.887	0.757	[-4.267, -1.855]
BAND Shock	-0	0	[-0, 0]
TARGET SUMCAP B-spline 1	2,247.638	2.007	[2,245.941, 2,251.486]
TARGET SUMCAP B-spline 2	2,203.819	3.294	[2,200.382, 2,212.141]
TARGET SUMCAP B-spline 3	2,256.723	2.392	[2,254.178, 2,262.55]
TARGET SUMCAP B-spline 4	2,202.157	2.156	[2,197.936, 2,205.794]
TARGET SUMCAP B-spline 5	2,293.337	2.623	[2,289.942, 2,295.87]
TARGET SUMCAP B-spline 6	2,190.144	1.129	[2,189.029, 2,191.785]
TARGET CAP B-spline 1	-2,014.84	3.063	[-2,020.113, -2,011.309]
TARGET CAP B-spline 2	-1,756.918	3.661	[-1,762.614, -1,748.606]
TARGET CAP B-spline 3	-1,217.592	2.866	[-1,220.863, -1,210.965]
TARGET CAP B-spline 4	-431.08	2.03	[-433.39, -426.678]
TARGET CAP B-spline 5	222.511	0.602	[222.096, 223.821]
TARGET Shock	0	0.002	[-0.003, 0.005]
σ^2_{BAND}	1.037	0.338	[0.652, 1.733]
σ^2_{TARGET}	213.721	8.96	[206.11, 238.018]

Number of capacity changes = 774. Initial parameters estimates selected through OLS before being estimated by maximum likelihood. SUMCAP refers to the summed capacity of a firm's competitors, while CAP refers to a firm's own capacity, both measured at the time the firm makes an investment decision.

Table 6: Entry and Exit Policy Results

Parameter	Coefficient	Standard Error
Exit Policy		
Constant	-1.306	0.183
CAP	-1.55×10^{-3}	2.81×10^{-3}
SHOCK	-4.60×10^{-5}	8.80×10^{-5}
SUMCAP	4.50×10^{-5}	1.70×10^{-5}
Late Dummy	-0.301	0.081
Entry Policy		
Constant	-1.68	0.210
SUMCAP	3.71×10^{-5}	3.60×10^{-5}
Late Dummy	-0.491	0.242

Sample size for exit policy function = 2233; sample size for entry policy function = 414.

tionship of the covariate to the response variable is a superposition of several piecewise polynomials. Figure 8 shows that the model does a fairly good job of predicting capacity from one period to the next. The productivity shock does not have a statistically significant effect on the size of the band. This is somewhat surprising, as higher levels of efficiency in production is in part due to better administration and organization, qualities which would also translate to efficiencies in investment. On the other hand, this may reflect the fact that more efficient firms are able to make better use of their existing capacity, and so have higher deviation thresholds before making permanent adjustments.

The productivity shock also has no statistically significant effect on the target. This can be due to the fact that being more efficient has two countervailing effects. On one hand, the more efficient firms have less incentive to engage in "precautionary" overinvestment, putting downward pressure on the target level. At the same time, they operate more efficiently, which can have a positive influence on the desired level of capacity. The results suggest that the two effects have roughly equal magnitudes, leading to an empirically indistinguishable effect of productivity on target level.

The entry and exit policy function results are presented in Table 6. For the most part, the marginal effect parameters have the desired sign in the exit equation. As would also be expected, a firm has a marginally lower probability of exiting a market given a higher capacity, which is a measure of the firm's staying power and the strength of market demand. The productivity shock parameter has the correct sign, as more productive firms have a lower incentive to leave the market in any given period. As a firm's competitors become larger it has an increased chance of leaving the market. Firms have a significantly lower probability of leaving the market in the later period.

Interestingly, the capacity of extant firms is not a significant explanatory variable in the entry equation. One explanation for this is that the relationship between extant capacity and the expected value upon entry is complicated by the nonlinear response of competing firms to any entry. For less-capitalized markets, competitors may be more likely to actively respond to the entry of a new firm, leading to a profit-reducing capacity buildup. When faced by a set of larger firms, that entrant may be able to enter as a smaller firm and not face any response from the incumbents. Sometimes this reduced intensity of competition leads to a higher probability of entering the market, just at a lower overall level of capacity, which it is important to emphasize is completely separate from the choice to enter or not.

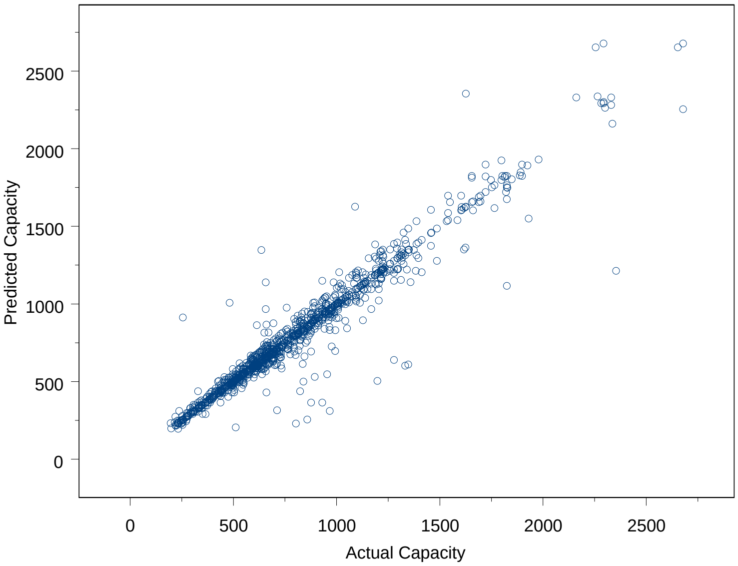
A second set of confounding factors is that larger markets may have higher demand, in which case the capacity of competitors would proxy for this. The addition of market-level fixed effects to account for demand heterogeneity changes the sign on the sum of competitor's capacity. However, the effect was still statistically insignificant and resulted in much less precise estimates of the other parameters. The relatively small size of the relevant data set leads me to use the most parsimonious specification, as the magnitude of the coefficients on the constant and late period dummy are similar across specifications, and more precisely identified in the model with fewer regressors.[39]

The most important variable in the entry model is the dummy for the 1990 Amendments, which is significantly negative. This directly contrasts to the effect of the Amendments on the probability of exit, which is lower after 1990. Couple these facts with the results from the demand and production estimations, and the empirical story becomes clearer. The demand parameters indicate that demand grew and remained strong a few years after the passage of the 1990 Amendments. The production parameters suggest that, due to industrywide technological shocks to the production technology, firms were able to run their kilns longer and cheaper in the later period. Lower costs lead to higher profits, which rationalizes the lower probability of exit in the later period. The one piece of the puzzle which does not fit with these other facts is that there was also a lower probability of entry after the Amendments. The only explanation that resolves this empirical tension is that firms must have faced increased sunk costs to enter and operate in this industry. These costs

---

[39] To check the robustness of the dynamic parameters to these specifications, I ran the estimation scheme using both sets of parameters, with negligible differences.

Figure 8: Investment fits



can have two parts: sunk entry costs and investment costs. I first recover the investment costs before estimating the distribution of sunk entry costs.

The results are presented in Table 7. The first thing to note is both the fixed and variables costs of adjustment playing a significant role in determining the investment behavior of firms in this industry. While the mean of the fixed adjustment costs looks implausibly large at first, recall that firms only make adjustments when they receive sufficiently low draws from this distribution. Firms balance the costs of delaying a change in their capacity, in order to receive a more favorable draw on adjustment costs in the future, against the benefits of investing in the current period. Also, the benefits of selling off capacity are lower than the costs of acquiring the capacity. This is a check on the validity of the model, as there must be no arbitrage in equilibrium. The exit values are fairly low given the size of the profits in this industry. The scrap value distribution only has bite for very marginal firms on the edge of large markets with many competitors. On the other hand, the fixed costs of capacity adjustment are sufficient to induce lumpy investment behavior when I solve for the MPNE of the theoretical model. This is an indication that the empirical parameters are large enough to rationalize the type of investment behavior observed in the data.

Table 7: Dynamic Parameter Results

Parameter	Median	Standard Deviation	95% Confidence Interval
Early Period			
ADJPOS	30,522	146.72	[30,491, 30,963]
INVMCPOS	131	1.98	[125, 131]
INVMCPOS2	0.018	0.001	[0.018, 0.021]
ADJNEG	22,646	597.99	[21,754, 23,562]
INVMCNEG	-1,115.78	114.17	[-1,279, -925]
INVMCNEG2	35.06	4.01	[28.428, 40.742]
SCRAP	84,016	456	[82,640, 84,109]
Late Period			
ADJPOS	27,631	30.32	[27,562, 27,663]
INVMCPOS	70.20	0.75	[69.36, 71.67]
INVMCPOS2	0.015	0	[0.015, 0.015]
ADJNEG	22,216	999	[20,062, 22,996]
INVMCNEG	-1,553	118.88	[-1,645, -1,291]
INVMCNEG2	55.18	2.59	[49.38, 57.25]
SCRAP	54,801	424	[54,423, 55,749]

Point estimates and confidence intervals were obtained using 30,000 simulated outcomes of 2 firms with 100 year lifetimes each. Bootstrapped statistics were generated from 100 bootstraps of 25% subsamples taken with replacement from the simulated data set. Values were sorted and the 5th and 95th percentiles used for determining the confidence intervals.

Interestingly, these parameters show that it is less expensive to invest after the passage of the Amendments. This rules out more expensive investment as a deterrent to entry after the Amendments, and indicates that the regulation must have shifted the distribution of sunk entry costs.

I model the sunk costs of entry as draws from a distribution of costs that is common across markets. To recover the parameters of this distribution, I assume that is approximated by a normal distribution. I match the empirical probability of entry for a given state, given by the probit policy function above, against the cumulative distribution function evaluated at the expected value of entering at that state. The results of the estimation are presented in Table 8. One of the main results of this paper is that the Amendments have increased the mean of the entry cost distribution by roughly 35% while decreasing variance by a small degree. The shift in both the mean and the variance work to decrease the chance of a firm receiving a small enough draw on the sunk cost of entry to warrant building a new facility. These parameters neatly rationalize how it is possible to observe lower entry rates when operation in the industry is more profitable than ever.

The expected value of entering shifts from roughly \$10M to \$35M after the 1990 Amend-

Table 8: Sunk Cost of Entry Distribution Results

Parameter	Median (000 $)	Standard Deviation	95% Confidence Interval
Early Period			
Mean	120,976	11,603	[93,321, 132,865]
Variance	2,688,393,144	518,997,607	[1,546,170,928, 3,278,242,947]
Late Period			
Mean	162,470	7,728	[145,133, 173,115]
Variance	2,046,773,539	273,115,403	[1,456,777,834, 2,433,445,760]

Parameters were estimated by matching the cumulative distribution function of a normal distribution to the empirical probabilities of entry. The expected value of entry at different states were constructed using simulation. 256 runs were conducted at each state; states were varied by the capacity of incumbent firms from 0 to 5000 in 100 unit increments.

ments. The difference in values is due to both the lower costs of production and investment in the later period, and the fact that firms expect a lower amount of competition in all future states. The decrease in the expected number of competitors is caused by the distribution of sunk entry costs increasing by almost 35%, or $40M. I stress that this shift in the distribution of sunk costs is the single most important determinant of market structure in the second period. A static analysis would have only captured the differences in production and investment costs due to technological changes and would have missed the change in the sunk costs. This is particularly true in this industry, since timing of the implementation of the regulations suggests that there should be no changes in the variable cost structure as a result of the Amendments. Without understanding how the distribution of sunk entry costs changes it simply is not possible to rationalize the observed behavior of firms in this industry.

To summarize, I estimate a fully dynamic model of oligopoly before and after the passage of Amendments to the Clean Air Act in 1990. I exploit the timing of the Amendments to isolate the cost stucture changes due to the passage of the Amendments, as EPA did not promulgate the final emissions standards until 12 years after the passage of the Act. So while firms did not face changes to the variable cost structure immediately, they did face changes in their sunk costs. The Amendments imposed a new permitting process which, coupled with increased regulatory attention at the state and local levels, led to an increase in the cost of greenfield plants and a lower probability of operation approval. The parameters I have recovered above are consistent with the underlying regulatory story, and do a good job explaining the behavior of firms in this industry. Estimates in hand, I now turn to synthesizing the empirical and theoretical sections to infer the welfare effects of the Amendments.

7 Policy Experiments

The power of a structural model is the ability to simulate counterfactual policy experiments once a researcher knows the underlying primitives. I have found the primitives which best rationalize the observed policies of the firms as optimal given the posited theoretical model. In fact, solving for the MPNE of the theoretical model with these primitives should generate the policy functions in the data, as long as I have recovered the policy functions flexibly enough in the first stage.[40] My primary interest is to evaluate the welfare effects of the Amendments, so a natural investigation is to determine the differences across cost structures for quantities of economic interest, including welfare measures for both producers and consumers. To achieve this, I compute the MPNE of the theoretical model with two sets of parameters: the observed post-Amendments cost structure, and the post-Amendments cost structure with the distribution of sunk entry costs taken from before the regulation.

The rationale for this exercise is taken from the timing of the regulations. The components of the Amendments dealing with new restrictions on emissions took significantly longer to promulgate than the fixed components dealing with the operation of those plants. As such, the Amendments did not change the variable cost structure of the cement industry until 2002, when new standards were finalized and enforced. However, the 1990 Amendments did change the way that firms operated in the industry, requiring all firms to obtain operating permits and to submit detailed environmental impact studies of both greenfield plants and renovations to existing production processes. The Amendments also coincided with an increased movement at the state and local level to restrict the growth of the cement industry, as many civic groups became increasingly concerned about the environmental impact of cement plant operations in their communities. For these reasons, I assume that the only change in the two periods due to the Amendments enters in through the distribution of entry costs. The timing allows me to separate the influence of positive technological shocks to the production process from the effects of the regulation. Thus, I use the variable cost structure recovered post-Amendments in the counterfactual calculations.

To recover the welfare effects of the Amendments, I solve the theoretical model under the two cost structures described above. Once these policy functions were obtained, it is possible to simulate forward hypothetical markets given some starting configuration. I am primarily interested in evaluating the distribution of producer profits and consumer surplus under two

---

[40] Reference simulations show this is exactly the case.

Table 9: Counterfactual Policy Experiments

	Post 1990 Amendments	Counterfactual	Social Planner
New Market			
Producer profit	56,588.09	89,478.53	-426,805.09
Consumer welfare	308,895.53	337,020.61	2,279,737.55
Periods with no firms	1.01	1.01	1.00
Periods with one firm	4.93	4.24	0.00
Periods with two firms	23.33	17.81	0.00
Periods with three firms	11.01	8.57	0.00
Periods with four firms	59.72	68.36	99.00
Total welfare	365,483.61	426,499.14	1,852,932.45
Average Capacity Active Firms	2,035.36	1,950.10	3,773.55
Average Market Capacity	5,855.98	5,969.45	15,094.20
Incumbent Market			
Producer profit	316,413.99	304,151.53	-181,123.89
Consumer welfare	477,694.68	526,479.59	1,656,169.41
Periods with no firms	0.00	0.00	0.00
Periods with one firm	0.00	0.00	0.00
Periods with two firms	94.02	50.96	2.00
Periods with three firms	5.98	45.93	98.00
Periods with four firms	0.00	3.11	0.00
Total welfare	794,108.66	830,631.13	1,475,045.52
Profits of firm 1	717,325.06	698,447.48	304,738.31
Average Capacity Active Firms	3,337	3,324	4,487.47
Average Market Capacity	6,867	8,169	13,409.91

Industry distributions were simulated along 25,000 paths of length 100 each. All values are present values denominated in thousands of dollars. The new market initially has no firms and four potential entrants. The incumbent market is started with one firm at a capacity of 3,000,000 tons per year, one firm at a capacity of 1,500,000 tons per year, and two potential entrants.

different starting states: a new market with no incumbent firms, and a market with two incumbents and space for two entrants. I am restricted by computational restrictions from considering larger markets, but I find that results are qualitatively similar for market sizes of two, three, or four firms. I restrict the number of active firms to be four, which is the median size of a cement market in the United States.

Table 9 presents the results of the counterfactual simulations. In the case of a new market, where the initial state vector is four empty slots waiting for entrants, overall welfare has decreased significantly due to the Amendments. For producers, this is due to the direct costs associated with paying more on average to enter the market. The decreased probability of observing a full market with four productive firms is the key to understanding the welfare losses for consumers. The Amendments have caused a 10% loss in consumer welfare, primarily due to the probability of observing a market with four active firms is 13% lower in the second period. A curious distributional effect of facing less expected competition after the

Amendments is that the average size of a firm increases by 4.3%, while the overall average market capacity shrinks by 1.8%. This helps explain the curious fact that market capacity increased after the Amendments were passed in 1990, reversing a twenty-year decline. These firms knew that they would be facing a significantly lower probability of competition in the future, and thus had increased incentives to invest. Overall consumer welfare has decreased by 27%, or roughly $60M for a new moderate-sized market.

The effects for the market with two incumbent firms is a good illustration of the heterogeneous effects that the Amendments have had on the industry. While consumer surplus has decreased roughly 10%, the present value of being a firm in the market has actually increased by 4%, as the result of the increased market power associated with lower competition from potential entrants. For the first firm, endowed with a starting capacity of 3,000,000 tons per year, the net present value of being active in the industry has increased by approximately $19M, or almost 3%. So not only does the static cost analysis fail to account for the welfare effects on consumers through reduced entry and increased market power, it actually gives the wrong sign when evaluating the cost to the incumbent firms. This is a direct result of such low entry rates that the probability of observing a market with three or more firms post-Amendments is eight times less likely. So while the average size of the firm in the industry is not much different under either regime, the average size of market capacity is 1,300,000 tons per year lower, a 16% decrease.

An interesting application of the structural model is to examine the differences between the oligopolist's MPNE to that of the social planner. The social planner sums the profits of all firms and consumer welfare for each configuration of capacities. I calculate the social planner's MPNE with the same cost structure as the oligopoly counterfactual. This solution gives an upper bound for the welfare losses under the regulation, as this would have been the best the industry could have done in the absence of the Amendments.

The key characteristic of the social planner's solution is that it exploits lost welfare gains due to inefficiently low investment to increase overall welfare. The social planner is willing to inflict losses on the firms, through costly, expansive investment, in order to drastically increase consumer surplus. The average market size is almost three times larger under the social planner than oligopoly. It is also interesting to note that the social planner solves the maximization problem by having all four firms enter simultaneously and invest heavily, instead of taking the route of one extremely large firm. This is due to the nonlinear costs of investment, which increase quadratically in capacity adjustment. It is the dominant solution for the social planner to start with four firms and grow them to smaller levels than put all of its energy into one large super-firm.

The social planner's problem in a market with the two incumbent firms is qualitatively different. The social planner finds it optimal, given the starting configuration, to restrict market size to three active firms. The most interesting facet of this equilibrium is that total welfare is lower than when the market starts with no active firms. This is more intuitive when considering the market with two starting firms as a constrained subset of the general optimization problem of how to best construct a market for maximum efficiency. With two incumbents, the social planner's solution has fewer firms with larger individual capacities, with a slightly smaller overall market capacity. The starting configuration is far enough from the optimal configuration that the resulting level of overall surplus with two incumbents is less than when starting with a blank slate. This example illustrates the importance of considering starting conditions when making policy analyses.

8 Conclusion

In this paper, I have estimated the welfare costs of the Amendments to the Clean Air Act on the Portland cement industry. My principal finding is that a static analysis of the costs of the regulation will not only underestimate the costs to consumers, but will actually obtain estimates of the wrong sign for incumbent firms. Exploiting the timing structure of the implementation of the Amendments, I identify that the most significant economic change in the Portland cement industry was a large positive shift in the sunk costs of entry. These results highlight the importance of estimating the welfare consequences of regulation using a dynamic model to account for all relevant changes to the determinants of market structure. While market power plays a role in larger markets, welfare losses are primarily driven by a reduction in the mean number of firms serving the market. The additional barriers to entry do lead to active firms with higher capacities, but this effect is dominated by the lower number of total firms.

I find that the (S, s) investment rule is a flexible and powerful method for characterizing the lumpy investment behavior of firms in this industry, since these choices are driven by fixed adjustment costs. The interplay between market power, investment, and production choice is particularly interesting. For smaller markets, firms find it optimal to produce close to marginal cost due to capacity constraints. However, while the majority of firms in the cement industry may produce near the efficient quantity for a given capacity level, they underinvest relative to the social planner. These results further reinforce the need for a fully dynamic model to evaluate the response of the industry to exogenous changes in its cost structure.

I am currently working to provide estimates of the benefits accruing to consumers from cleaner air as a result of lower production. To obtain a first approximation of these benefits, I plan to couple emissions data with production data to estimate a pollution production function. An estimate of the pollution production function would allow me to estimate the reduction of various toxic emissions as a result of the Amendments. The EPA publishes estimates of the benefits of reducing the emissions of these different pollutants, so I can find a rough value of the benefits accruing to consumers from the Amendment's impact on their increased health.

An interesting extension of the present work would be to examine the effects of a "cap-and-trade" market-based emissions control program, similar to the trading program for SO_2 in the electricity industry. In this environment the regulatory authority removes all specific point-source control requirements and instead places an overall cap on the level of emissions in a regional area. Firms are endowed with pollution rights that they are free to trade among each other. This type of policy has the benefit of achieving the most efficient configuration of production within the industry for a given level of pollution. However, it may have other consequences with respect to market power and the concentration to pollution to a subset of firms within the market. By coupling emissions data, reported in many states at the yearly level, to production data I can back out a pollution production function. One question I can then address is whether efficiency obtains in this environment, as some of the more inefficient firms may buy pollution rights in return for additional market power. There are clearly a number of other interesting dynamic questions in this market, from the nonlinear health effects of pollution concentration to the investment incentives of heterogeneous firms in a region. I leave this interesting policy examination for future work.

References

Daniel A. Ackerberg and Kevin Caves. Structural identification of production functions. UCLA Working Paper, 2004.

Victor Aguirregabiria and Pedro Mira. Sequential simulation-based estimation of dynamic discrete games. Technical report, Boston University, 2002.

Portland Cement Association. *Concrete Basics: History and Manufacture of Portland Cement*. Portland Cement Association, 2001. URL `http://www.portcement.org/cb/concretebasics_history.asp`.

Orazio Attanasio. Consumer durables and inertial behavior: Estimation and aggregation of (s, s) rules for automobile purchases. *Review of Economic Studies*, 67:667–96, 2000.

Patrick Bajari, C. Lanier Benkard, and Jonathan Levin. Estimating dynamic models of imperfect competition. Technical report, Stanford University, 2003.

C. Lanier Benkard. A dynamic analysis of the market for widebodied commercial aircraft. *Review of Economic Studies*, Forthcoming.

Steve Berry, Ariel Pakes, and Michael Ostrovsky. Simple estimators for the parameters of dynamic games (with entry/exit examples). Technical report, Harvard University, September 2003.

David Besanko and Ulrich Doraszelski. Capacity dynamics and endogenous asymmetries in firm size. *RAND Journal of Economics*, 35(1):23–49, Spring 2004.

Victor Chernozhukov and Han Hong. A mcmc approach to classical estimation. *Journal of Econometrics*, 115(2):293–346, 2003.

Russell Cooper and John Haltiwanger. On the nature of capital adjustment costs. Technical report, NBER Working Paper #7925, 2000.

Mark Coppejans. On kolmogorov's representation of functions of several variables by functions of one variable. Technical report, Duke University, Forthcoming.

Ulrich Doraszelksi and Mark Satterthwaite. Foundations of markov-perfect industry dynamics: Existence, purification, and multiplicity. Technical report, Northwestern University, 2004.

Richard Ericson and Ariel Pakes. Markov perfect industry dynamics: A framework for empirical work. *Review of Economic Studies*, 62(1):53–82, 1995.

Chaim Fershtman and Ariel Pakes. A dynamic game with collusion and price wars. *RAND Journal of Economics*, 31(2):207–36, 2000.

Gerard Gaudet and Stephen W. Salant. Uniqueness of cournot equilibrium: New results from old methods. *Review of Economic Studies*, 58(2):399–404, April 1991.

G. Gowrisankaran and R. Town. Dynamic equilibrium in the hospital industry. *Journal of Economics and Management Strategy*, 6(1):45–74, 1997.

George Hall and John Rust. The (s,s) rule is an optimal trading strategy in a class of commodity price speculation problems. Technical report, Yale University, 2000.

ASTM International. Specification c150-00 standard specification for portland cement. Technical report, ASTM International, West Conshohocken, PA, 2001.

I. Jans and D. I. Rosenbaum. Multimarket contact and pricing: Evidence from the u.s. cement industry. *International Journal of Industrial Organization*, 15:391–412, 1997.

M. Jofre-Benet and Martin Pesendorfer. Estimation of a dynamic auction game. *Econometrica*, Forthcoming.

Kenneth Judd. *Numerical Methods in Economics*. MIT Press, 1998. ISBN 0-262-10071-1.

James Levinsohn and Amil Petrin. Estimating production functions using inputs to control for unobservables. *Review of Economic Studies*, 70:317–341, 2003.

Erin Mansur. Environmental restructuring in oligopoly markets: A study of electricity restructuring. Technical report, School of Management, Yale University, 2004.

Daniel McFadden and Whitney Newey. *Large Sample Estimation and Hypothesis Testing*, volume 4 of *Handbook of Econometrics*, pages 2111–2245. Elsevier, 1994.

Robert A. Miller and V. J. Hotz. Conditional choice probabilities and the estimation of dynamic models. *Review of Economic Studies*, 60(3):497–531, 1993.

U.S. Bureau of Mines and U.S. Geological Survey. *Minerals Yearbook*. U.S. Government Printing Office, 1973–2001.

G. Steven Olley and Ariel Pakes. The dynamics of productivity in the telecommunications equipment industry. *Econometrica*, 64(6):1263–1297, November 1996.

Ariel Pakes and Paul McGuire. Computing markov-perfect nash equilibria: numerical implications of a dynamic differentiated product model. *RAND Journal of Economics*, 25 (4):555–589, Winter 1994.

Martin Pesendorfer and Philip Schmidt-Dengler. Identification and estimation of dynamic games. Technical report, London School of Economics, 2003.

John Rust. Optimal replacement of gmc bus engines: An empirical model of harold zurcher. *Econometrica*, 55(5):999–1033, September 1987.

H. E. Scarf. The optimality of (s,s) policies in the dynamic inventory problem. In K. J. Arrow, S. Karlin, and P. Suppes, editors, *Chapter 13 in Mathematical Methods in the Social Sciences*. Stanford University Press, 1959.

Hendrik G. van Oss. *Mineral Commodity Survey*. U.S. Geological Survey, January 2001. Cement section.

Stuart J. Weiss. Looking ahead: Environmental issues for the new millennium. *Cement Americas*, 2000.

A Appendix

A.1 B-splines

There are a number of possible bases for forming polynomials of a continuous variable.[41] The most common basis functions are the power functions, whose $n+1$ elements take the form $x = (1, x, x^2, x^3, ..., x^n)$. While easy to implement and estimate, they suffer from two problems. First, they tend to be highly collinear, which can make matrix inversion fail in procedures such as OLS. The second problem is that data points have global influence over the regression parameters. In other words, observations in one partition of the data set have influence on the fit of the function in other partitions. I attenuate both of these problems by using B-splines to approximate nonlinear relationships. B-splines are piecewise polynomials, meaning that data points have local influence on the shaping of the approximating function, and are numerically stable. In general, the researcher partitions the range of the covariate into several adjacent intervals, demarcated by a knot vector. The B-spline fits a local polynomial of degree k over each interval, subject to two constraints: first, the B-splines must be continuous across adjacent cells; second, the first derivatives must be equal across cells at the knot. This ensures that the approximation of a function by a B-spline is continuous and smooth over the range of the explanatory variable. I make use of B-splines of order $k = 4$, or cubic B-splines, as this ensures sufficient smoothness in the approximating function to obtain asymptotic normality. Denote the knot vector as ξ. I choose to use $r = 4$ interior knots uniformly spaced along the domain of the regressor. Ideally, one would find the knot vector through some form of data-driven selection, but given the computational burden in this case I use the simplest approximation. I also add three knots at a and b, where a (b) is arbitrarily slightly less (greater) than the minimum (maximum) element of x. The basis for the B-spline of order 4, $\{B_{l,4}\}_{l=-3}^{r}$, is defined as:

$$\begin{aligned} B_{l,m}(x) &= \frac{x - \xi_l}{\xi_{l+m-1} - \xi_l} B_{l,m-1}(x) + \frac{\xi_{l+m} - x}{\xi_{l+m} - \xi_{l+1}} B_{l+1,m-1}(x), \text{ if } m = 2, 3, 4, \\ B_{l,1}(x) &= 1(x \in [\xi_l, \xi_{l+1})), \end{aligned}$$

---

[41] This section follows the exposition in Coppejans (Forthcoming).

where $1(\cdot)$ is the indicator function. The basis for the B-spline are four functions defined in a piecewise fashion over each interval between knots. The B-spline is defined as:

$$b(x) = \sum_{l=1}^{r+4} \alpha_l B_{l-4,4}(x), \tag{36}$$

where α is a vector of unknown parameters. In each instance a B-spline is used in the estimation above I replace the covariate with the vector of basis functions. The α are recovered in the same manner as if I had used regular polynomials.

A.2 Computational Details

The goal of my structural model is to explain what firms in a market will do in any economic environment. In Section 6, I found primitives which best rationalized the observed policies of firms in the data as optimal given the underlying model. However, in order to evaluate the effects of any changes in these primitives, as I do in Section 7 when I perform my counterfactual experiments, it is necessary to solve for the MPNE at the new parameters. The policy functions recovered from the data are generically valid only for one set of parameters, as profit-maximizing firms will typically change their behavior in response to changes in the underlying primitives. For example, doubling the marginal cost of investment is going to induce firms to change their optimal investment strategies. This section details exactly how I solved out for the MPNE of the theoretical model at a given set of parameters.

A.2.1 Discretization

It is important to realize that backing the primitives of the model out of the observed policies implicitly assumes that the underlying state space is continuous. Conceptually, there is nothing wrong with this, as it is possible to write out the value and policy functions for a firm at any possible configuration of capacities. Practically, however, it is clearly an impossible task to solve for the policy functions of the firm at each point in the state space. The literature deals with this problem by discretizing the state space and solving for the MPNE of that subspace. In the present case, I discretize the capacity levels that any one firm can attain and cap the maximum level. As the state space grows exponentially in the number of firms, and I am interested in capturing the dynamics of typical markets where there may be six active firms at one time, I choose a fairly coarse discretization of the state space: firms may choose their capacity in 750,000 ton per year increments, up to a maximum of 6,000,000 tons per year. The upper limit was chosen high enough to ensure that no firm ever found it optimal to invest to higher levels if given a choice.[42]

A.2.2 Transition Probabilities

One problem with discretizing the state space arises in allowing firms to make lumpy capacity adjustments. The model has to be well-defined for any investment choice, and it vastly simplifies convergence problems to ensure that the best-response functions are smooth in the decisions of competitors. There, I incorporate a smooth mapping between continuous investment choices and discrete states. Intuitively, as a firm increases its level of investment, it has a marginally higher chance of obtaining the next highest discretized capacity level. Formally, I define the transition probability between two states, conditional on the investment of an active firm, as:

$$Pr(s_i'|s_i, \text{INV}_i) = \int_{\phi_1}^{\phi_2} dH(v),$$

where $H \sim U[s_i + \text{INV}_i - \delta, s_i + \text{INV}_i + \delta]$, δ is the distance between two adjacent nodes in the state space, $\phi_1 = s_i + \text{INV}_i - \delta/2$, and $\phi_2 = s_i + \text{INV}_i + \delta/2$. This specification imparts a small degree of uncertainty into the model, which helps ensure that the best response functions of one firm to another are smooth. It also has the attractive feature of nesting the deterministic investment case when the state space is infinitely fine. This is a particular problem in this application, since I depart from previous studies which allowed the firm to make investments which shifted the probability of moving to the next node in the state space.[43]

A.2.3 Solving the MPNE

Solving for the MPNE is conceptually straightforward. The value function for each point in the state space is defined piecewise by Equations 13 and 17, depending on whether or not the

---

[42] I found this bound by solving the monopoly case with extremely large potential capacities. In all the cases I have calculated, no firm ever finds it optimal to exceed the monopolist's maximal capacity choice.

[43] One motivation for allowing lumpy adjustments is simply the problem of mapping the model of jumps between adjacent nodes into observed behavior. The number and timing of investments becomes a function of how fine the state space is discretized, which makes for a more difficult explanation of what exactly firms are doing in the data.

firm is a potential entrant or incumbent, respectively. The solution algorithm exploits that the value function is a recursive function, as $V(\cdot)$ shows up on both sides of the equation. As $V(\cdot)$ maps into itself, the algorithm that I use to find these value functions, and the associated policy functions, is value function iteration. This is a very simple algorithm, suggested by McGuire and Pakes (1994), where the researcher iterates through the entire state space, solving the value and optimal policy at each state given the values and policies of all the other states. This continues until the changes to the value and policy functions reach some very small value. In essence, the solution algorithm has computed an MPNE where no firms want to deviate in their strategies given the strategies of their competitors.

This algorithm has two parts: moving through every point in the state space, and computing the value and policy functions at each point. Moving through the state space is trivial, as a researcher can simply order the states by increasing capacities and move linearly through them. Computing the value at each state is straightforward, as it is composed of two parts: a per-period function and the discounted value of expected future revenues. Given the parameters from the estimation, it is possible to compute the profits accruing from the product market at each state. This is an invariant number that is simply assigned to each state element. The per-period payoff function depends on the strategies of the firm at that state space, and so may change every time through the algorithm. To find the optimal strategies, it is necessary to compute the optimal investment and exit policy for incumbents, and the optimal investment and entry policy for potential entrants.

As described above, the optimal investments are computed conditional on being active in the next period. It is straightforward to solve where marginal benefit equals marginal cost in order to find the optimal level of investment. Finding the marginal benefit is slightly complicated, as it is necessary to compute the expected value of every state possible given the market vector of investments and expected entry/exit strategies. However, one solution to this problem is to simply calculate the expected value of every state in the state space and sum.[44] Once the optimal investment is computed, the solver then computes the probability of receiving a draw on sunk entry costs (scrap values) which would make entry (exit) optimal.

Once these quantities are computed, that state is updated, and the algorithm moves on to the next state and repeats the process. This process continues until the change in the value function is sufficiently small that it is reasonable to conclude that convergence

---

[44] This is slow and inelegant, but guaranteed to work and easy to debug.

has been achieved. In general, the policy functions settle down much faster than the value function. However, as tiny changes in the value function can induce lumpy changes in the policy function, it is necessary to let the value function converge until it the changes are very small. In this paper, I used calculate the norm of the change in the value function and stop the algorithm when it is less than 1E-6.

A.2.4 Convergence Issues

While the algorithm above is easy to implement, convergence can be complicated by several factors. The essential problem is that firms may make lumpy adjustments to their policy functions in response to small changes by their rivals. One way to get around this is to make all of the policy variables in the model continuous functions. The private information at every level of the model assures that firms play against expected policies of their competitors, which helps ensure that firms converge smoothly towards the MPNE. However, it is still possible to end up in what Judd (1998) describes as a "hog cycle", where firms bounce back and forth between optimal responses to each other, while equilibrium lies within a convex combination of these responses. In order to fix these problems, I dampen the responses of the firms to their rivals actions by making the new policy a combination of the optimal reponse and last period's reponse. This has the effect of ensuring the algorithm finds an equilibrium while dramatically slowing down the rate of convergence.[45]

A.3 Computing Simulation Paths

To compute the W terms in the dynamic estimator, I fix the maximum number of active firms and draw an arbitrary starting configuration out of a distribution. The main concern is drawing enough representative starting configurations that various parameters are identified. For example, starting with a market populated by very large firms is not going to help identify scrap values, since exits are such rare events under such circumstances that even permuting the exit policy generates the same observed outcomes. By imposing a degree of diversity on the starting configuration of the capacities of the firms in the market, including firms

---

[45] As an interesting aside, the social planner's solution does not exhibit this behavior, as payoff structure there does not have the feature of one firm winning or losing at another firm's expense. Clearly, the need for dampening depends on the specific problem.

with zero capacity, I observe enough variation in the lifetime values of firms to identify the underlying unknown parameters.

Recall that W is a $k + 1 \times 1$ vector of discounted expected payoffs and actions of the first firm, where k is the dimension of the unknown parameters.[46] Computing these expected payoffs and actions is simple: read the vector of payoffs and policies at each state, discount, add to the previous period's vector, and then update the state vector in response to the firms' actions. To fix these ideas, consider a monopoly market currently without an active firm. The W for this market consists of a vector of zeros until the firm obtains a low enough draw on its sunk entry cost that it finds it optimal to enter the market, say in period n. The amount of investment is recorded, properly discounted by β^n, in the correct slot of W, as is squared investment. In the next period, I begin adding discounted sums to W as the firm receives profits from the product market, makes investments, and, perhaps, exits at some point in the future.

---

[46] The first spot in the vector is a one, in order to sum the stream of discounted profits accruing from the product market. There are no unknown parameters in the product market profit function, so it is just a lump-sum payment firms receive for being active in each state. Conceptually it is posisble to estimate unknown parameters in the profit function by replacing this one with the vector of unknowns.

Table 10: Demand Estimation Results

Parameter	Coefficient	Standard Error
Intercept	82.82	2.58
Slope	-0.003	0.0006
Alaska, Hawaii, Oregon, and Washington	26.72	2.27
Arizona, Nevada, and New Mexico	14.49	2.39
Arkansas and Oklahoma	-2.52	2.40
California North	11.47	2.36
California South	17.44	2.80
Colorado and Wyoming	10.24	2.50
Florida	7.64	2.39
Georgia and Tennessee	0.11	2.34
Idaho, Montana, and Utah	12.67	2.32
Illinois	-4.82	2.37
Indiana	-3.88	2.49
Iowa, Nebraska, and South Dakota	2.83	2.31
Kansas	1.81	2.40
Kentucky, Mississippi, North Carolina, and Louisiana	0.84	2.37
Maryland, Virginia, and West Virginia	-1.33	2.42
Michigan and Wisconsin	3.55	2.35
Missouri	-5.44	2.36
New York and Maine	1.60	2.38
Ohio	3.27	2.37
Oklahoma	-11.58	3.00
Pennsylvania East	7.20	2.62
Pennsylvania West	-1.71	2.37
South Carolina	-4.13	2.62
Texas	39.27	7.52
Texas North	9.53	2.59
Texas South	-2.13	2.54
1982	-4.99	2.02
1983	-6.23	2.01
1984	-5.78	2.02
1985	-8.78	2.14
1986	-12.40	2.00
1987	-15.91	2.00
1988	-18.93	2.00
1989	-19.36	2.08
1990	-20.96	2.05
1991	-22.34	2.03
1992	-23.27	2.03
1993	-17.40	2.13
1994	-12.77	1.84
1995	-7.28	2.13
1996	-6.18	2.11
1997	-5.28	2.12
1998	-3.76	2.09
1999	-3.31	2.10

The instruments for shipped quantity were number of firms in the market, skilled wages, and the prices of coal, electricity, and natural gas. The estimated elasticity of demand is 10.75. Number of observations: 512. Residual standard error: 0.0993 on 466 degrees of freedom. Multiple R-Squared: 0.7568. F-statistic: 32.23 on 45 and 466 degrees of freedom, the p-value is 0. 5 observations deleted due to missing values. Dummy variables are relative to 1980 and Alabama. Texas is a combination of Texas North and Texas South in years where they were not reported separately.