NORTHWESTERN UNIVERSITY

Market Power in Markets for Power

A DISSERTATION

SUBMITTED TO THE GRADUATE SCHOOL

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

for the degree

DOCTOR OF PHILOSOPHY

Field of Economics

By

James Myatt

EVANSTON, ILLINOIS

June 2018

ABSTRACT

Market Power in Markets for Power

James Myatt

I study how electricity generation firms exert market power, raising price above marginal cost. Typical studies of market power in electricity markets focus on how firms sustain markups in the short-run energy market. I explore other channels electricity generation firms use to strategically maximize profits. First, I analyze strategic investments and disinvestments. I find that firms with market power choose a portfolio of generators which significantly departs from a least-cost planner's optimal portfolio. Next, I test if firms that have common owners, which gives them less incentive to compete, sustain higher markups. I find evidence of this effect using a reduced-form approach, but a structural approach has less convincing conclusions. Finally, I study the capacity market, a relatively new market where generators can earn extra revenue. In a stylized model, I find that current capacity market designs result in overpaying for capacity. I then use a regulatory feature of the PJM electricity market to analyze if higher capacity prices do indeed drive higher net investment, and I find no convincing evidence.

Acknowledgements

I would like to thank the many people who helped and supported me along the way to completing this dissertation. First, I owe a debt of gratitude to my dissertation committee: Mar Reguant, Rob Porter, and Igal Hendel. I am thankful that Mar arrived at Northwestern during the early stages of my research career and helped shape my disparate ideas into a set of well-defined research questions that reflect my interests in informing public policy. More than just the "big picture," she helped me with the day-to-day tasks in research which helped reduce weeks of debugging code to just days. Rob has also been a great source of guidance in my research. I am grateful that he is almost always available for a meeting, and he can quickly understand my questions and offer insightful advice. I appreciate his commitment to his graduate students through detailed editing and nearly perfect attendance at student workshops. Finally, Igal has helped me see the important, big economic ideas in research. He pressed me to concisely describe the market distortions I study and often found the problems I had missed when staring at the same problem for too long. All three have guided me along the way to becoming an economist and made me feel like a colleague.

I am also thankful to have completed my dissertation at the Northwestern Economics Department, which strongly supports research in empirical industrial organization. I formed most of my research questions while listening to the many talented speakers invited for workshops and conferences. I was also able to complete a large part of my research while working as a research assistant for the Center for the Study of Industrial Organization. Northwestern and the Economics Department have provided so many resources for me to get started in my research career, and I am proud to be an alumnus.

Of course I would not be able to complete my dissertation without the love and support of my parents and my sister, Alice. They have given me a love of reading and learning. They also inspire me to do the hard work to get both done! I am thankful they supported my decision to go to graduate school right after college, moved me halfway across the country, and bought some plane tickets for me to get back home! Alice has helped me through this dissertation as only an older sister can. She has picked up the phone on my lowest research days, but also helped me experience some of the "good life" while living as a grad student. I also want to thank my girlfriend, Lily Bruno. From listening to me endlessly describe my research to buying a suit for me while I finished my last slides, I am happy she is on my team!

Finally I would like to thank all my friends in graduate school and back home. Celebrating successes and helping me get through the many frustrations while researching this dissertations was invaluable. The Thursday Night Dinner group helped me blow off steam, talk about the high-level "fun stuff" in economics, and was like having a family even being away from home. I think we have almost solved all the world's problems. I want to especially thank my roommate, David Benson. While we both agree separating home life from work life is important, I believe jumping through all the hoops of the PhD and venting about it together has made it all possible.

Table of Contents

List of Tables

List of Figures

CHAPTER 1

Introduction

Market power, the extent to which firms can charge prices higher than marginal cost, is a significant distortion in electricity markets. Prices can regularly be over ten times the average price and allegations of market manipulation are common. Mitigating market power is important because the industry is large, with almost $400 billion in sales across the United States in 2016,[1] and electricity is a major input in the rest of the economy. Because market power abuse continues to be a problem in electricity markets, there has been a large body of economic research in electricity spot markets. I add to this literature by analyzing other channels electricity generation firms with market power use to increase markups.

In the second chapter, I analyze strategic investment and disinvestment decisions made by electricity generation firms. In a competitive market, the long-run generation portfolio of firms should replicate the least-cost generation portfolio, minimizing long-run investment costs as well as variable costs of production. However, theory predicts that firms with market power may depart from a planner's first-best portfolio. Empirically quantifying this effect is difficult because most plants were built decades ago and it is challenging to separately identify the effect of market power on investment from all the other factors that go into planning a power plant. Instead of analyzing the entire generation portfolio in the United States, I study the large amount of recent coal plant retirements. Coal plants

[1] https://www.eia.gov/electricity/annual/html/epa_02_06.html

that were expected to last for many more years or decades have difficulty competing with natural gas generators benefitting from low gas prices. I employ a structural model of long-run oligopoly investment and short-run production to quantify how many of these coal plants retired efficiently, in-line with a planner, and how many retired for strategic reasons. I find that about two-thirds of coal capacity retired for strategic reasons. I also model a bailout of coal plants, a policy several states and the Department of Energy are considering, and I find that the costs of a bailout outweighs any benefits for consumers when firms can still exert market power in the energy market. This is one of the first empirical studies of strategic investment in electricity markets.

Next, in the third chapter, I test whether electricity generation firms that share a common owner charge higher prices. Increasingly, institutional investors own major stakes in competing firms across the whole economy, but also in the electricity generation industry. Firms that share a common owner have less incentive to compete because competition usually lowers total industry profits, hurting investors that own some of all firms in the industry. I follow the recent literature on detecting and quantifying the effect of common ownership on price by using a reduced-form approach. I contribute to the literature by complementing this analysis with a structural model of imperfect competition in the energy spot market, taking common ownership into account. Because marginal costs are known in the electricity industry, I can test the behavioral assumption that electricity generation firms take their common owners into account without biasing estimates of cost. Using the reduced-form approach, I find that firms with more owners in common charge higher prices. However, the structural model fits the data better when I assume that firms maximize their own profits, rather than their common owners' profits. I then use the structural model to explore the effect of a proposed law to limit investments in competing firms. In an environment where firms do take their common owners' interests into account, the law lowers producer surplus, but increases consumer surplus enough so that total surplus increases.

Finally, in the fourth chapter, I analyze capacity markets. Capacity markets are a relatively new way for electricity generators to earn revenue. Today, 20-30% of consumers' bills go towards the capacity market. Capacity payments awarded in the market are supposed to compensate generators for reliability and they should prevent power outages. High prices in the capacity market are intended to attract investment or retain existing capacity. However, many electricity markets across the world do not have a capacity market. Generators earn revenue from energy sales alone. These markets do not have drastically different reliability problems compared to markets with opportunities for capacity payments. This paper adds to the growing literature that analyzes the necessity and outcomes of capacity markets. I develop a stylized model of a typical capacity auction. I find that awarding payments to multiple generators increases equilibrium bids. Within the stylized model, simply covering the sunk costs of new investment procures the desired level of capacity and saves consumers money. Then, I analyze investment outcomes in PJM, the largest electricity market in the United States. I do not find convincing evidence that higher capacity prices are driving higher net investment in PJM.

CHAPTER 2

Market Power and Long-Run Technology Choice in the U.S. Electricity Industry

2.1. Introduction

Long-run investment in electricity generation is a challenging problem. Demand fluctuations, inelastic consumers, and lack of storage necessitate using a mix of baseload and peaking generation technologies to minimize the cost of supplying electricity. Economic theory suggests that firms in a perfectly competitive market will invest in the least-cost mix of these generating technologies. However, market distortions are known to exist. I examine how market power distorts the portfolio of technologies away from the least-cost mix that a planner would choose, and I evaluate second-best policies to correct the distortion resulting from market power.

Distortions in technology choices are important as the U.S. electricity industry is currently undergoing major changes in its generation portfolio. Developments in horizontal drilling technology to extract oil and gas in previously unreachable shale deposits have led to a dramatic decrease in natural gas prices since 2008. The reserves of oil and gas are expected to last for several decades. The natural gas price decrease has had a significant effect on the electricity market where generators fueled by natural gas are typically the marginal, price-setting technology. Savings from fuel costs have passed through to wholesale prices of electricity, which makes it more difficult for baseload technology, such as coal and nuclear, to earn sufficient revenue to recoup their fixed costs. Additionally, the small but growing presence of renewable technologies undercuts baseload technologies and lowers wholesale prices. As a result, there has been a dramatic and unexpected increase in the amount of coal plants retiring. Regulators are concerned about the loss of these reliable, low marginal-cost baseload plants.

The literature on market power and long-run investment in electricity markets is largely theoretical, and it suggests that oligopolists will hold less baseload capacity than a least-cost planner. Early work by [**von der Fehr and Harbord, 1997**] shows that firms with market power will underinvest in aggregate capacity to push up prices. Moreoever, there is underinvestment in the baseload technology rather than the peaking technology. Later work by [**Arellano and Serra, 2007**] shows that when firms are incentivized to cover peak demand (so there is no underinvestment in aggregate capacity), firms exert market power by profitably distorting their mix of capacity towards more peaking and less baseload capacity than a least-cost planner would choose. A key feature of these models is that, conditional on capacity built, firms must supply electricity using their lowest marginal cost generators. That is, they cannot strategically withhold baseload capacity that has already been built, a common regulation in electricity markets where marginal costs are public information. Holding less baseload technology increases the amount of hours in a year when peaking units are marginal, so that remaining baseload units command high markups more often.

I employ a structural model to quantify the distortion that market power creates in the US electricity industry. In an ideal setting, I could randomly assign symmetric oligopolists to a cross section of markets with similar demand conditions. I could then see how market power affects the mix of technologies that they choose to build. Instead, I account for the heterogeneity in firms and endogenous capacity choices in a two-stage investment model. In the first stage, firms observe the drop in natural gas prices and can make costly adjustments to their pre-2008 generation capacity portfolios. In the second stage, they use their updated generation portfolios to compete in a series of hourly spot markets for a full year. I use the actual retirements and building decisions of firms to identify cost parameters in the model.

I find that firms' market outcomes depart significantly from the choices of a hypothetical least-cost planner. In MISO and PJM, two of the most coal-reliant markets in the US, firms have retired about five and three times as much as a planner would choose, respectively. Second-best policies to prevent coal retirements have modest benefits. Preventing all coal retirement in both markets would save consumers less than 1% on their total energy bills. Further, incentivizing firms with fixed-cost subsidies to prevent retirements is expensive and the cost outweighs any gains from lower electricity prices. My results indicate that preventing coal retirement does not greatly lower costs for consumers. However, regulators may also consider other costs or benefits not in my model such as reliability or pollution.

My paper adds to the literature evaluating the performance of liberalized electricity markets. Much of the literature has focused on short-run gains in efficiency: [**Cicala, 2015**], [**Bushnell and Wolfram, 2005**], [**Davis and Wolfram, 2012**]. My paper is one of the first empirical papers to evaluate distortions in the long-run outcomes of electricity markets. The closest related paper is [**Bushnell and Ishii, 2007**] which employs a dynamic framework to study investment in electricity markets and simulates outcomes.

Also, [Ishii and Yan, 2007] analyze building decisions of entrants just after restructuring. My paper is also related to the growing literature analyzing long-run outcomes in capital-intensive and concentrated industries [Ryan, 2012], [Fowlie et al., 2014], [CollardWexler, 2013]. This literature shows that failing to account for long-run responses to policy changes can result in misleading welfare calculations.

The main contribution of my paper is to quantify investment distortions in the electricity market. Distortions have been studied in the US before restructuring, but earlier analysis was largely qualitative and theoretical. Mine is the first empirical study, to my knowledge, that addresses the long-run economic efficiency of restructured electricity markets. I also analyze capacity payments or fixed-cost subsidies in electricity markets. Capacity payments are becoming part of electricity markets world-wide, but there is little research done on their effects.

2.2. Conceptual Framework

Investment in electricity markets requires tradeoffs between sunk costs, annual operating fixed costs, marginal costs, and timing. Generation technologies are ranked from lowest to highest marginal cost, called the "merit order" in the industry. There is an inverse relationship between technologies' marginal costs and sunk and annual operating fixed costs. Lower marginal cost technologies have higher sunk and annual operating fixed costs. Because demand fluctuates on an hourly and seasonal basis, minimizing the long-run cost of supplying electricity necessarily requires a mix of technologies. Low marginal cost capacity is too expensive to build and reserve for just a few peak hours. The optimal mix of technologies depends on their cost structures and the variability of demand.

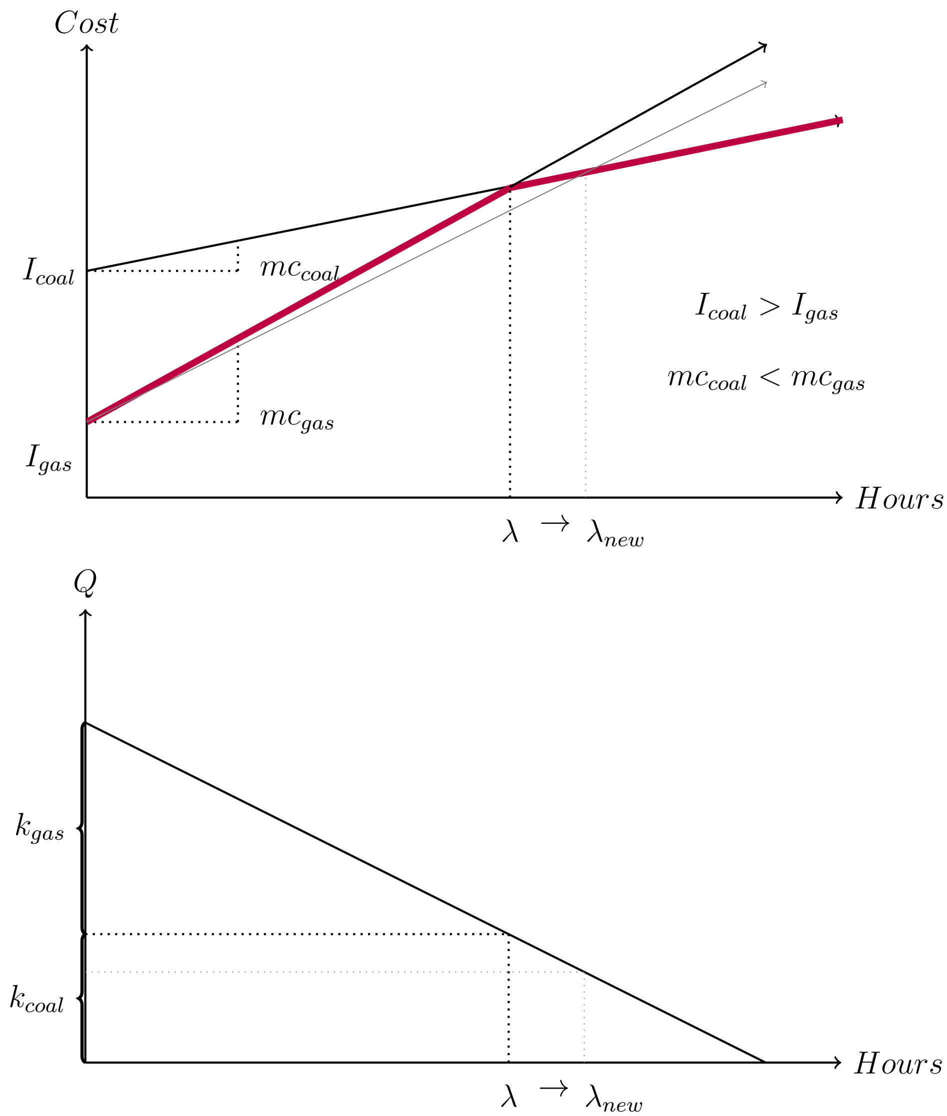
I demonstrate the tradeoffs and incentives of a planner and a firm with market power in a simple setting before detailing the data, industry background, and my structural model. I narrow the technology space to coal and gas generators. Coal is the "baseload" technology, with a low marginal cost and a high annual operating fixed cost. Gas is the "peaking" technology with a high marginal cost and a low annual operating fixed cost. There are no sunk costs in this example.

2.2.1. Long-Run Planner

The top panel of Figure 2.1 shows the tradeoff between marginal costs, fixed costs, and timing. Each dark curve represents the total cost of supplying 1 unit of electricity for a given amount of hours. The coal generator has a high annual operating fixed cost, I_{coal}, but it has a low marginal cost, mc_{coal}. Conversely, the gas generator has a low annual operating fixed cost, I_{gas}, but a higher marginal cost, mc_{gas}. Fixed costs are paid per unit of capacity. The lower envelope of the cost curves represents the least-cost technology choice given the amount of hours the generator will run. The threshold λ is the cutoff point where a least-cost planner switches from choosing a gas generator to choosing a coal generator. If a generator only needs to be used for less than λ hours per year, the high fixed annual operating cost of a coal generator outweighs the benefits of its low marginal cost. Similar tradeoffs are made in energy efficient appliances and vehicles. Typically, energy efficiency technologies cost more when purchasing (higher fixed cost), and savings come from lower energy expenses while in prolonged use.

The threshold λ is important when a planner minimizes the cost of serving fluctuating demand throughout the year. The bottom panel of Figure 2.1 shows a "load curve,"

Figure 2.1. Least cost capacity mix



Notes: The top panel shows the total costs of operating a coal generator and a gas generator for a given number of hours. The threshold λ is the cutoff number of hours where a coal generator is lower cost than a gas generator. When the marginal cost of gas decreases, the optimal cutoff shifts to λ_{new} as the marginal cost advantage of coal over gas has decreased, it takes a higher threshold of time for coal to dominate gas. The bottom panel shows an example load curve, which describes how many hours of the year demand is *at least* a given quantity on the vertical axis. The optimal capacity technology mix for this market is given by the curly braces along the vertical axis. Demand levels above k_{coal} occur less than λ hours making gas the optimal technology. Lower mc_{gas} corresponds to cutoff λ_{new} and lower optimal coal capacity.

which, when interpreted from the vertical axis, maps how many hours demand is *at least* a given level of Q. The vertical axis of the graph ranges from the lowest level of demand in a given year to the peak level of demand. The horizontal axis ranges from the 0th hour of the year to the total number of hours in a year: 8,760. The optimal amount of coal capacity, k_{coal} is where λ, which drops directly down from the top panel, maps through the load curve to the quantity axis. Any level of demand above k_{coal} happens for less than λ hours per year. The marginal cost benefit of coal compared to gas is not experienced long enough to overcome the fixed cost disadvantage. Gas plants in this setting are called "peaking" plants because they run less often; they are turned on only during high demand, peaking periods.

The Shale Revolution dramatically lowered the price of natural gas. Fuel price makes up the majority of the marginal costs for fossil fuel generators, so the marginal cost of gas has decreased. The top panel of Figure 2.1 shows the total cost of running a gas generator with decreased mc_{gas} as the light grey curve. With lower mc_{gas}, the cutoff time for switching from gas to coal moves from λ to the right at λ_{new}. The marginal cost advantage of coal over gas has decreased, so the amount of hours a coal generator needs to run to be favored to gas has increased. In the bottom panel of Figure 2.1, a higher λ_{new} implies less coal capacity. Retiring some coal generation in response to the Shale Revolution is economically efficient.

2.2.2. Long-Run Monopolist

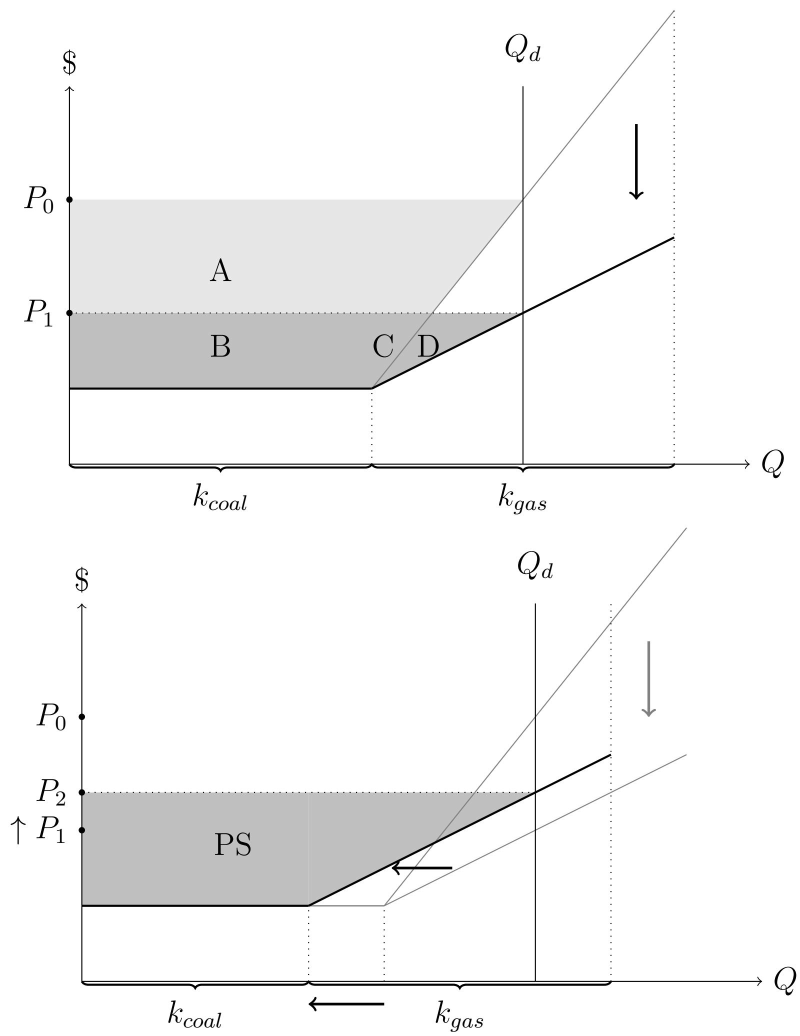
Like the planner, a firm with market power faces similar tradeoffs between marginal costs, fixed costs, and timing. However, the firm uses capacity to earn profits rather than serve demand at least cost. Similar to the planner, the firm uses a mix of technologies, but the reasoning differs. The firm will not serve higher levels of demand with only coal capacity as the rents from these hours are exceeded by the annual operating fixed costs of capacity. The firm, anticipating peak hours where coal capacity will be inframarginal, strategically underinvests in coal capacity. Underinvesting in coal capacity increases the amount of hours remaining coal generators are inframarginal and earn high rents. Also, for a given level of demand, underinvesting means more expensive generators will be called on, increasing the market price and inframarginal rent. Firms with market power tradeoff building coal capacity and earning extra rents per unit versus extra units lowering the market price in peak hours and reducing the number of peak hours.

Figure 2.2 shows a monopolist's response to the Shale Revolution in a stylized setting. The upper panel shows the firm's marginal cost curve initially consists of a flat portion from coal capacity, k_{coal}, and an upward sloping portion, in light grey, from gas capacity k_{gas}. In a given hour with quantity demanded, Q_d, and price, P_0,[1] the monopolist earns producer surplus $A + B + C$. Producer surplus is used to recoup the fixed costs of capacity and earn profits. The precise levels of k_{coal} and k_{gas} resulted from the monopolist maximizing long run profits.

The Shale Revolution lowered the price of natural gas and therefore the marginal cost of the gas generators. In Figure 2.2, the darker, upward sloping segment represents the new marginal cost of the firm's gas capacity. With the same levels of capacity and same demand, Q_d, the market price decreases to P_1, and the firm earns less producer surplus: $B + C + D$. The bottom panel of Figure 2.2 shows how the firm can increase its profits

[1]In this example, demand is perfectly inelastic and price is set equal to marginal cost for illustrative simplicity.

Figure 2.2. Oligopolist response to Shale Revolution



Notes: The top panel shows the producer surplus earned by a firm for coal capacity, k_{coal}, and gas capacity, k_{gas}, before and after the Shale Revolution. With high gas marginal costs, the market price is P_0 and producer surplus is $A + B + C$. With lower gas marginal costs, the price is P_1 and producer surplus decreases to $B + C + D$. The bottom panel shows that the firm can increase producer surplus, in the long-run, by retiring coal capacity. For the same demand conditions, price is higher at P_2. The firm also increases profits by forgoing some fixed costs of capacity.

in the long run. Retiring coal capacity allows the firm to forgo some fixed costs. It also means that higher marginal cost natural gas generators must be called on to serve demand, increasing the market price and inframarginal rents. The example shows the advantage of committing to less coal capacity for a given hour. Less coal capacity also increases the amount of hours that coal plants are inframarginal. In summary, the responses to the Shale Revolution of firms with market and a planner differ because a planner retires coal capacity just to recoup fixed costs and firms with market power amplify retirement to increase inframarginal rents, in addition to saving on some fixed costs.

2.3. Data

I use data from several sources to analyze the effects of market power on coal generator retirements. The Energy Information Administration (EIA), the independent statistics arm of the US Department of Energy, provides information on generator characteristics. EIA Form 860 collects information on every generator in the United States, and it reports each generator's capacity, technology type, operating status, initial year in service, and final year in service. EIA Form 923 reports monthly generator operations and data to construct marginal costs for almost every fossil fuel generator in the country. I accessed this data through SNL Financial, an independent data company which verifies and cleans up the original EIA data.[2] SNL Financial also constructs marginal cost data and annual operating fixed costs for each generator, and SNL provides the ownership structure, at a parent-company level, for each generator over time. The EIA also provides commodity price data on coal and natural gas. The Environmental Protection Agency's (EPA) Continuous Emissions Monitoring System (CEMS) provides hourly gross production data for

[2] The EIA actually uses the SNL version of the data when conducting its own analysis.

almost all fossil fuel generators in the US. Finally, each electricity market, called Regional Transmission Organizations (RTOs), provides hourly price and quantity data.

2.4. Industry Background

Historically in the United States, regulated utilities made long-run capacity investment decisions. Utility companies were vertically integrated: generating, transmitting, and delivering electricity to all consumers in their service areas. Because of the costly infrastructure investment needed in this industry, there has always been concern over its tendency to be a natural monopoly. A series of legislative and judicial decisions led to the modern regulated utility.[3] In exchange for a guaranteed monopoly charter over a geographic service area, a state regulator was allowed to set prices and veto infrastructure investments. Utilities, controlling all aspects of production and distribution, had all the information about demand and costs, and they could minimize both short-run and long-run costs.

There are many well-known economic problems with the traditional regulated utility model. Prices regulators set, which guarantee a specified rate of return, do not incentivize utilities to minimize cost [**Abito, 2017**]. Similarly, regulated rates of return tied to assets incentivized firms to overinvest in capital intensive projects, not necessarily minimizing cost [**Cicala, 2015**]. The political ideology of state governments influenced the generosity of regulators, distorting investment [**Lim and Yurukoglu, 2016**]. After decades of the regulated utility model, dissatisfaction with ex-post uneconomic and expensive capacity decisions inspired a wave of restructuring across the United States in the late 1990s [**Borenstein and Bushnell, 2014**].

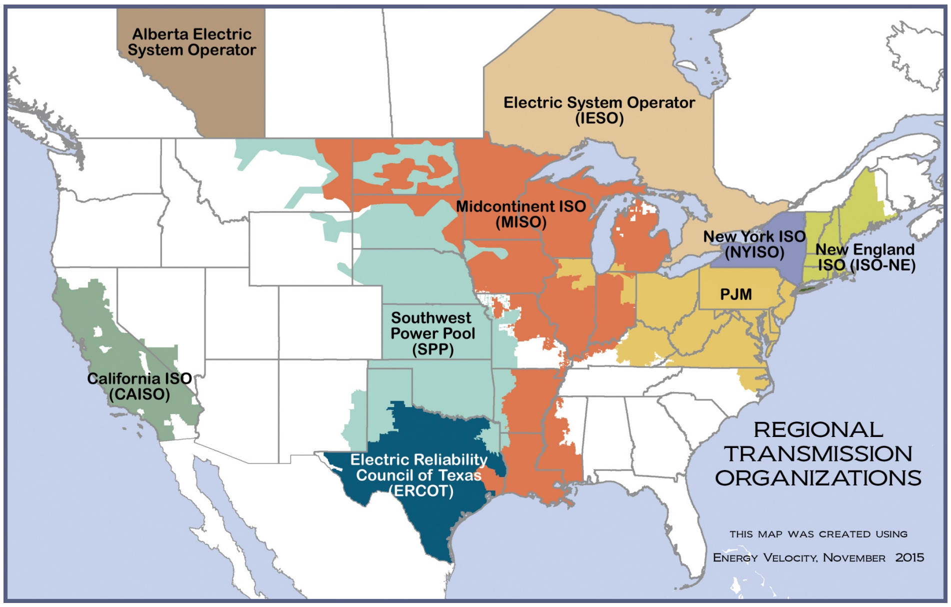
[3]The Public Utility Holding Company Act of 1935, Federal Power Act of 1935

Today, liberalized wholesale electricity markets are characterized by upstream competitive generation firms that supply electricity to consumers via their local regulated utility. Transmission and distribution networks are still owned by utilities and subject to state regulators. However, control of the networks was given to independent system operators (ISOs).[4] They are independent in that they are not market participants and often are set up as non-profit organizations. The collection of many regional utility networks under an ISO is considered a wholesale market for electricity. Figure 2.3 displays the major ISOs in North America. California (CAISO), New York (NYISO), Texas (ERCOT) are almost entirely within one state. New England (ISONE), the Midcontinent ISO (MISO), and the Pennsylvania-New Jersey-Maryland Interconnection (PJM) span across multiples states.[5] While this paper models almost all these ISOs for estimation, the counterfactual focuses on MISO and PJM which are the most coal reliant. Not all regions of the United States liberalized their electricity markets, and they are still under the traditional utility model. These regions are in white in Figure 2.3. Analyzing these regions is beyond the scope of this paper because of the unique incentives of these firms described above.

The Federal Energy Regulatory Commission (FERC) tasks ISOs with a variety of objectives to ensure competitive and reliable electricity markets. In the short run, ISOs administer auctions to procure enough electricity from suppliers to meet demand at the lowest cost for each hour of the year. Increasingly, FERC and the ISOs are focused on firms' long-run investments and divestments. Every ISO in the United States, except for

[4]These can also be Regional Transmission Authorities (RTOs). The two classifications are under the authority or recommendation of the Federal Energy Regulatory Commission. For the purposes of this paper, the distinction between the two is unimportant and I refer to both interchangeably.
[5]The Southwest Power Pool (SPP) is excluded from the analysis because of data availability

Figure 2.3. Liberalized Wholesale Electricity Markets in the United States



Source: Federal Electricity Regulatory Commission (FERC)

ERCOT in Texas, provides some mechanism for additional sources of revenue for generators that are not tied to electricity production. These subsidies, or capacity payments, are intended to influence the long-run decisions of firms to incentivize a more competitive market. However, their overall economic efficiency is still debated [**Bushnell et al., 2017**]. Regulators should consider the market-specific trends of demand, technology, and competitive incentives of firms making long-run decisions.

2.4.1. Demand

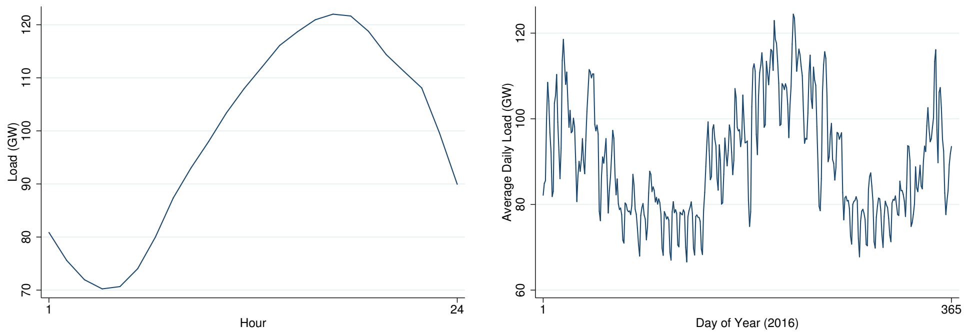
Demand for electricity, called "load" in the industry, is thought of as inelastic [6]. While there have been some advances,[7] most consumers do not know the true price of electricity when consuming it. Because consumers do not respond to price in the short-run, the level of demand fluctuates throughout the day and seasonally according to people's and businesses' activity. The left panel of Figure 4.3 shows that demand peaks in the evening on a typical day in PJM, one of the largest electricity markets in the US. It is lowest right before most people awake. Daily demand patterns have systematic differences across weekdays and weekends. Throughout the year, demand increases with needs for heating and cooling. The right panel of Figure 4.3 shows demand in winter can be nearly twice the demand in spring or fall. The usual daily usage of electricity in an ISO is a major factor in firms' long-run technology choices.

While there are hourly and seasonal demand fluctuations to consider when planning how to supply electricity, aggregate consumer behavior has been stable on an annual basis for years. [**Davis, 2017**] finds that after decades of increasing per capita energy use, there have been significant declines recently. So even though population is increasing, aggregate energy use is relatively flat. Indeed, Figure 2.5 shows the density of hourly loads New York across a 13 year span, and there is no clear trend of increasing or decreasing demand. So within year demand schedules matter for producers, but they can expect a relatively stable environment when planning for the future.

[6] [**Ito, 2014**] finds that consumers have almost no response to their most recent monthly electricity bills. The medium-run (a few months lag time) elasticity is -0.088

[7] Some consumers can sign contracts where they know they will pay a peak rate at certain hours and get discounts at other hours. Also, energy-intensive industrial producers can sign contracts where they can be paid to reduce electricity consumption during peak times, which is called "demand response"

Figure 2.4. PJM Load Schedules



(a) Load Schedule 6/1/2016 **(b)** Average Daily Loads 2016

Notes: Load data is taken from PJM's website.

Figure 2.5. Density of Loads NYISO

Notes: Kernel Densities of hourly loads schedules for certain years in New York. Data taken from the NYISO website.

2.4.2. Generation Technologies

Firms meet fluctuating, inelastic demand with a variety of generation technologies which can be broadly classified as fossil fuel, nuclear, or renewable. Fossil fuel generation is the most significant category in the United States accounting for 65% of all production in 2016.[8] Fossil fuel generators are reliable and flexible in terms of production. Baseload coal plants supply electricity almost constantly throughout the year. More flexible peaking plants like oil and gas plants come online for sometimes just a few hours each year. Nuclear is the second largest category of generators accounting for 20% of production. Nuclear generators produce at low marginal costs and are largely baseload power. Starting up or shutting down can take days, so all nuclear generators run continuously at full capacity except for a few scheduled maintenance days. Renewable energy is the third largest category making up 15% of generation with equal parts hydro, solar, and wind. While having low marginal costs, renewables are subject to the variability of their energy sources.

Generation technologies have an innate tradeoff between annual operating fixed costs and marginal costs. Low marginal cost technologies have high annual operating fixed costs. Table 2.1 displays this relationship across nuclear generators and the four main subcategories of fossil fuel generation: coal, combined-cycle gas turbines (CCGT),[9] conventional gas, and oil. Categories of generators are in order of increasing marginal cost, called the "merit-order" in the industry. Marginal cost is in units of megawatt-hours (MWh) where one MWh is enough electricity to power 750-1000 homes for one hour.[10] For fossil

[8] https://www.eia.gov/tools/faqs/faq.php?id=427&t=3

[9] Combined-cycle gas turbines (CCGT) differ from conventional steam turbine gas generators. While both types burn natural gas, CCGT can generate 50% more electricity with the same gas input (https://www.gepower.com/resources/knowledge-base/combined-cycle-power-plant-how-it-works).

[10] http://www.energy.ca.gov/glossary/ISO_GLOSSARY.PDF

Table 2.1. Cost Structure of Nuclear and Fossil Fuel Generators

Technology	Nuclear	Coal	CCGT	Gas	Oil
Marginal Cost	11.74	26.40	27.48	62.71	258.05
Fixed Cost	109,586	29,038	12,513	8,580	9,576

Notes: Costs are averages across all generators in the markets of interest. Marginal costs ($/MWh) were constructed as in equation (1) with heat rates from the 2016 data. The prices of fossil fuels were assumed to be their annual averages from 2012-2016 from the Energy Information Administration (EIA). Variable O&M costs and annual fixed costs ($/MW-year) are from estimates in the data

fuels, fuel expenses account for the majority of marginal cost, but it is also composed of small variable operating and maintenance costs. Marginal costs can be constructed from Equation 2.1. Each generator has a heat rate, HR, specified by the manufacturer, which is the rate at which the unit can convert thermal energy from fossil fuels to 1 MWh of electricity. Interacting the heat rate with the price of fuel P_{fuel}, yields the fuel portion of the marginal cost. Variable operating and maintenance costs, $VO\&M$, include treating and pumping water. Nuclear and renewables have the lowest marginal costs. Annual operating fixed costs are in units of megawatt per year (MW-year) and include annual payrolls and permitting. Importantly, these costs do not include sunk investment costs. Firms weigh the trade-offs between marginal costs, annual operating fixed costs, and sunk investment costs to make an optimal portfolio of technologies.

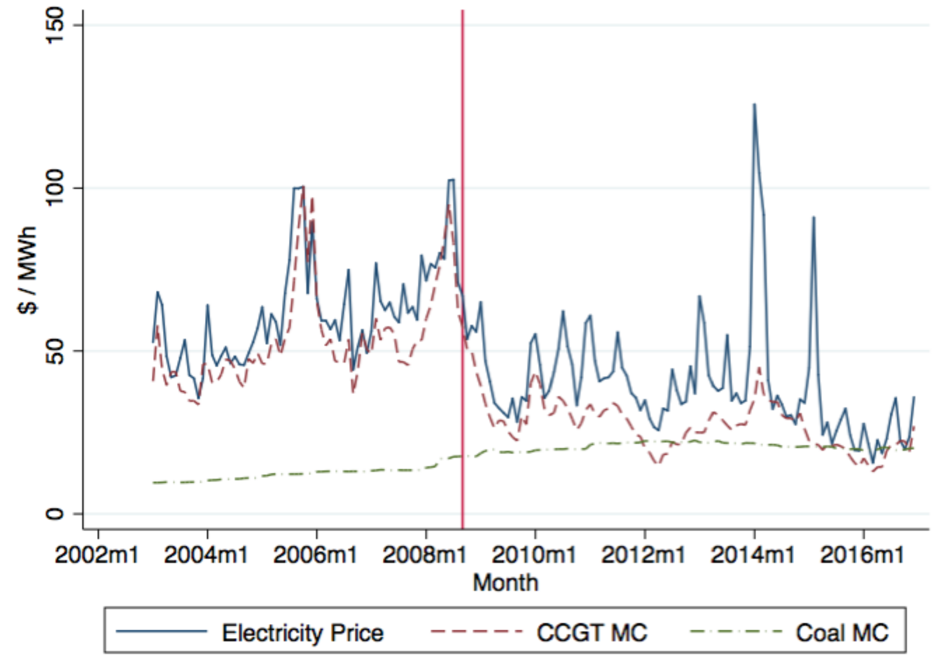
$$MC = VO\&M + HR \cdot P_{fuel} \tag{2.1}$$

2.4.3. The Shale Revolution and Electricity Industry Dynamics

Starting in the mid-2000s, technological advances in drilling have allowed firms to extract previously inaccessible deposits of oil and gas in shale formations. The large supply of cheap natural gas is greatly affecting the electricity industry. Figure 4.5 shows the price of electricity is closely tied to the price of natural gas. Natural gas generators are often the marginal, price setting generator in a market so changes in gas prices pass through to the electricity markets. Between 2008 and 2009, the effects of the Shale Revolution fed into the electricity industry lowering prices by almost one half. Shale oil and gas reserves are expected to last the United States for decades.

Cheaper natural gas has both short-run and long-run effects in the electricity industry. Natural gas generators can now bid into electricity markets at a lower price increasing the likelihood that they are called on to produce in a given hour. This affects coal plants in several ways. Some coal plants with higher marginal costs are less likely to be called on to produce. Also, coal plants that continue to produce receive less inframarginal rent: Figure 4.5 shows the difference between market price and coal marginal cost decreased signficantly after 2008. Finally, coal plants may be inframarginal less often. Figure 2.7 illustrates the reduced rents using data from PJM. The left panel shows the actual market marginal cost curve in 2016. If the level of inelastic demand is 150 GW, in a competitive market, almost all coal plants would produce and earn inframarginal rent from a market price around $40. The right panel shows the market supply curve if the price of natural gas doubled, which is closer to gas prices before the Shale Revolution. Almost no coal plants directly compete with CCGT generators, and their inframarginal rent is much higher, earning a price around $80, for the same level of demand. Therefore, the technology

Figure 2.6. Prices and Marginal Costs

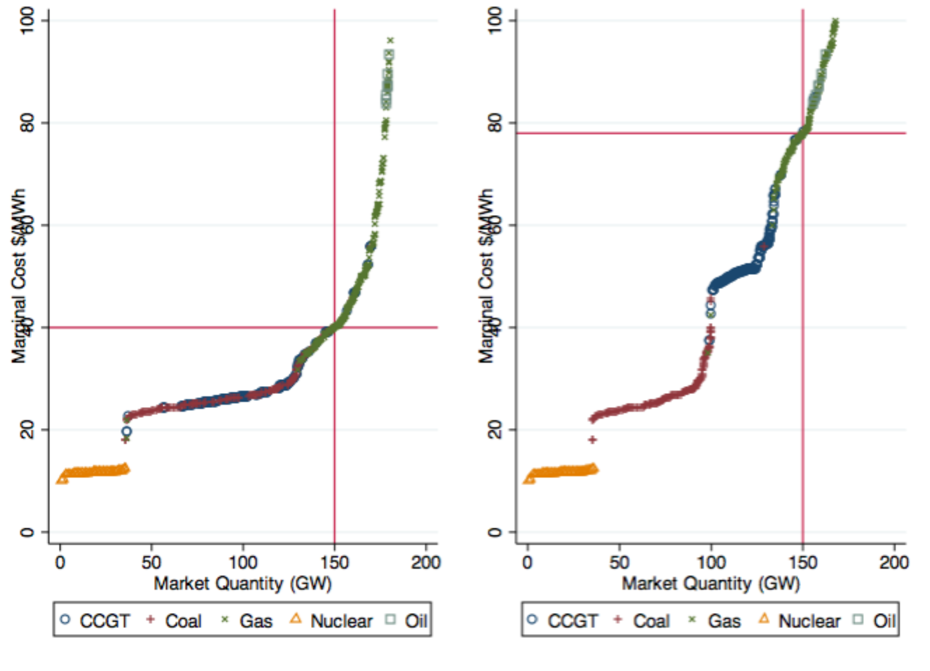


Notes: Electricity price data is the average hourly price by month in the New York ISO (NYISO). CCGT marginal cost data is from the Henry Hub price series available through the EIA. Prices have been converted from mmBTU units to $/MWh using the average heat rate in the data. Coal marginal cost data is from the St. Louis FRED price index of short tons of coal. The index was converted to short ton prices using the EIA price of a short ton for the base year and then converted to $/MWh using the average heat rate of coal generators in the data.

mix of any market might have been ideal for a long run equilibrium with higher natural gas prices. But the sudden impact of the shale revolution necessitates long-run capacity adjustments.

Indeed, firms have adjusted their portfolios in response to the unforeseen drop in natural gas prices. Figure 2.8 shows the percentage change of total capacity for CCGT and coal technologies over time. The series is normalized so the 2004 capacity is 100 for each technology. Coal generators had little change before the 2010s. Very few were retired

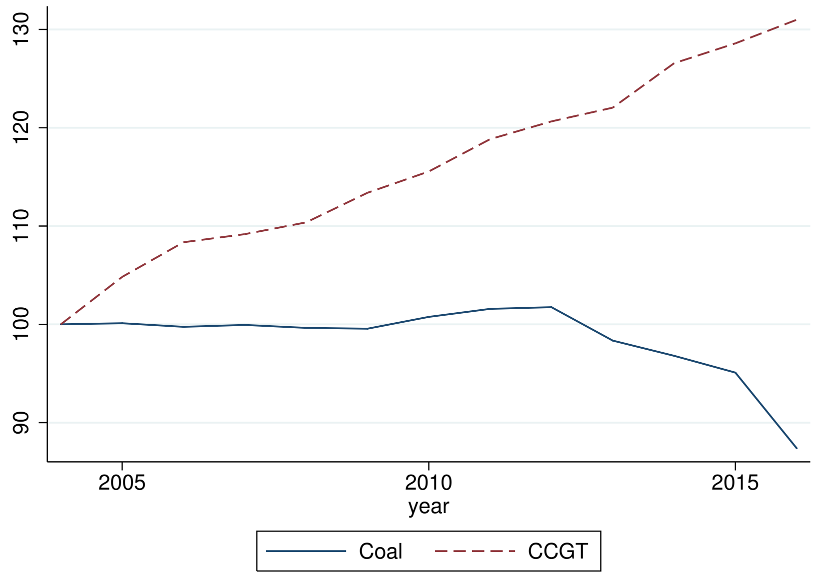
Figure 2.7. Market Marginal Cost Curves



Notes: Constructed from non-renewable generators in PJM Interconnection. The left panel is for natural gas prices in 2016. The right panel has the same set of generators but when natural gas prices are doubled. Generators smaller than 50 MW or with marginal costs above 100 have been removed for clarity.

or built. But after the drop in natural gas prices many firms requested permission from their ISOs to retire plants. Retiring generators is a heavily regulated process and can take years, so the effect is lagged and starts in 2013. Before the drop in natural gas prices, few of these generators were expected to retire. CCGT capacity has grown over the period. While CCGT plants have followed a steady trend of building, the trend was expected to slow without the Shale Revolution. [**Brehm, 2017**] analyzed long-run building plans that are required by regulators and found that without a drop in natural gas prices, the growth of CCGT capacity would have been significantly lower.

Figure 2.8. Capacity Changes



Notes: Total coal and CCGT capacity for all ISOs analyzed over time. Capacities are normalized so that 2004 levels are 100. From 2004 to 2016 coal capacity fell by almost 15% and CCGT capacity grew over 30%

In response to coal plant retirements, state legislatures are considering subsidizing fixed costs for these generators. Most notably, Ohio has considered giving coal plants between \$3-5 billion over the course of a decade to prevent early retirement. Additionally, in October 2017, the Department of Energy proposed a revenue guarantee to coal and nuclear plants estimated to cost between \$800 million and \$3.8 billion annually. Fixed-cost capacity-payments are already a feature of many electricity markets which aim to influence long-run outcomes. Proposed technology specific capacity payments would be an additional revenue source for coal generators on top of any existing capacity payments. There is debate over how effective these subsidies are at influencing long-run outcomes and if they "crowd out" future generation that would have been built.

2.5. Model

I model the long-run electricity industry in a finite-horizon, two-stage investment game. Each firm in a market begins with a predetermined, unmodeled portfolio of generators. In the first stage they make costly adjustments to their portfolios. In the second stage, they use their portfolios to compete in a year long spot market to sell electricity. Firms expect to use their new portfolios for the next T years and the conditions of spot market, demand schedules and costs, will remain the same each year. I will describe the details of each stage of the game starting with the second stage.

2.5.1. The Second Stage Spot Market

The framework for the second stage spot electricity market follows [**Bushnell et al., 2008**]. A grid-operator must procure electricity for every hour, t, for a full year at the lowest possible price p_t. Each hour, there is inelastic and perfectly forecastable demand y_t. All demand is met from two sources of suppliers. The first source is a competitive fringe which supplies electricity according to the function $q_t^{fringe} = \beta p_t$. The second source is a set of n oligopolists in the market who face a residual demand curve given in Equation 2.2. They set their quantities in a Cournot game $Q_t(p_t) = \sum_i q_i$.

$$Q_t(p_t) = y_t - \beta p_t \tag{2.2}$$

Firms can be characterized by the generators they own. Each firm's capacity is described by the vector $k_i \in \mathbb{R}_+^{J_i}$ where J_i is the number of generators a firm owns plus any generators it might build. Each generator also has a constant marginal cost and a per-MW annual fixed cost given by $c_i \in \mathbb{R}_+^{J_i}$ and $I_i \in \mathbb{R}_+^{J_i}$, respectively. Marginal costs and annual fixed costs vary across and within technology groups. A firm's generation capacity and respective marginal costs combine to make a firm-level marginal cost function $mc_i(q_{it}; k_i, c_i)$. Marginal cost functions are increasing step functions as each generator can have an idiosyncratic marginal cost. Each hour of the year, each firm chooses quantities to maximize profits in a capacity constrained Cournot Equilibrium and earns annual profits $\Pi_i(k_i; k_{-i}, c_i, c_{-i})$.

2.5.2. The First Stage Capacity Choice

Oligopolists are endowed with a set of generators $k_i^0 \in \mathbb{R}_+^{J_i}$. Generators are used in the spot market to produce energy and earn annual profits $\Pi_i(k_i; k_{-i}, c_i, c_{-i})$, but firms must also pay annual operating fixed costs, regardless of production, $I_i' \cdot k_i \in \mathbb{R}$. Each individual generator has an idiosyncratic per-MW fixed cost I_{ij}. Each firm's annual operating fixed costs are organized in the vector $I_i \in \mathbb{R}_+^{J_i}$. The initial set of generators can be interpreted as the equilibrium capacities with the "pre-Shale Revolution" marginal cost structure. After there is a decrease in the price of natural gas, the marginal costs of all natural gas plants change. In response, the oligopolists can make costly adjustments, either building generators or retiring existing generators, $x_i \in \mathbb{R}^{J_i}$, resulting in a new portfolio of generators $k_i^0 + x_i = k_i^1 \in \mathbb{R}_+^{J_i}$, given the choices x_{-i} of all other oligopolists. They expect this choice to be their only chance to change portfolios, and the resulting portfolios, demand schedules, and cost structure will remain the same for the next T years. Their long-run problem can be summarized as

$$\begin{aligned}
&\max_{x_i} \frac{1-\delta^T}{1-\delta}\left(\Pi_i(k_i^1; k_{-i}^1, c_i, c_{-i}) - I_i' \cdot k_i^1\right) - C_i(x_i) \\
&k_i^1 = k_i^0 + x_i
\end{aligned}$$

where δ is a discount factor and Π_i is the reduced form profits on the annual spot market for the given choices of capacities k_i^1, k_{-i}^1 and corresponding marginal costs c_i, c_{-i}. Firms that build capacity trade off the net benefit of extra capacity in the spot market with the extra fixed cost of holding capacity I_{ij} and the cost of building the additional capacity C_i. Firms that retire coal capacity trade off losing revenue from capacity, increasing inframarginal rents, reducing annual fixed costs I_{ij}, and paying costs (or scrap value) of retiring capacity $C_i(x_i)$.

2.6. Estimation and Results

The goal of estimation is to recover the unobserved parameters in the model. There are parameters governing (1) the short-run, hourly spot market; (2) annual fixed costs; and (3) sunk investment and retirement costs. Many of these parameters have been reliably estimated by other researchers, and they are used throughout the industry and academic studies. I utilize their estimates, and I use common economic assumptions to identify and estimate the remaining parameters.

In the short-run, the outcomes of hourly spot markets are governed by the marginal costs of oligopolists' generators, the levels of inelastic hourly demands, and the responsiveness of the competitive fringe to market prices. The marginal costs of oligopolists generators, $c_{ij} \in \mathbb{R}$, are observed in the data. When I consider hypothetical generators that have not been built and therefore are not in the data, I use the average marginal cost of similar, new generators in the same market. Inelastic hourly demands are provided by ISO websites. The only parameter to estimate in the short-run market is the responsiveness of the competitive fringe, β, in each ISO. Given the primitives and a behavioral assumption that firms compete within a capacity-constrained Cournot-Nash equilibrium, I predict the quantities, prices, and profitability of each firm for each hour in each ISO as a function of their capacity holdings.

Long-run profitability is governed by annual fixed costs and one-time sunk costs. Each generator has a per-MW fixed cost, $I_{ij} \in \mathbb{R}$. I utilize estimates constructed by the data provider, SNL, which are similar to many estimates of the same costs the EIA makes available.[11] Sunk costs of building a CCGT generator or retiring a coal plant are crucial to calculating my counterfactual. The EIA provides estimates for the "overnight cost of capital" of CCGT plants, which I utilize. However, there is less known about the scrap value or cost of coal generators. I identify this value using a long-run profit maximization assumption.

I estimate all unknown parameters in sequence rather than jointly. The two-stage nature of the model and complicated calculations favor estimating spot-market primitives first, and then using those results to estimate the long-run capacity adjustment costs.

2.6.1. Competitive Fringe Supply Relation

Many of the smaller firms in the data have complex incentives making it difficult to predict their production. For example, small municipal utilities are contracted by their

[11] https://www.eia.gov/analysis/studies/powerplants/capitalcost/pdf/capcost_assumption.pdf

local governments to meet demand before selling to the grid. Also large industrial steel or paper plants may generate electricity for their own production needs, but they can sell electricity to the grid if it is more profitable. There are even small firms with just a few plants or even just a single plant that may not be able to affect market prices. Additionally, firms may export outside their ISO or electricity may be imported from firms in other ISOs for reasons I do not observe. I follow the framework developed in [**Bushnell et al., 2008**] to estimate a supply relation for production from these firms which are collectively called the competitive fringe.

There is no clear definition of which firms are strategic versus which firms are fringe players. I follow the "small fish swim free" rule of the Public Utility Commission of Texas, which regulates the ERCOT market. The rule gives an absolute defense against market power abuse for firms owning less than 5% of installed capacity. I further restrict the 5% rule to only consider fossil fuel capacity because it can be used strategically. Nuclear generation is nearly constant throughout the year, just stopping for maintenance. Renewable technology runs according to its energy source, sun or wind. Later in the analysis, I add nuclear generation to firms' strategic capacity portfolios at a marginal cost of 0. I leave renewables outside of my analysis as they make up only a small fraction of capacity in the markets of interest: MISO and PJM. Across all ISOs I analyze, the 5% rule defines many categories of firms as fringe players. For example all industrial plants (those that primarily produce something other than electricity), all municipal generators, and all single plant firms are classified as fringe players. These firms arguably respond to market prices and are small enough to not warrant serious attention for market power. However, the arbitrary threshold does define some large, legacy utilities as fringe players. For example, Pacific Gas & Electric in California is defined as a fringe player. In [**Bushnell et al., 2008**], they used a firm-level aggregate capacity theshhold of 800 MW of fossil fuel capacity to divide the two categories of firms, and Pacific Gas & Electric was also ruled out as a strategic firm. The 800 MW rule is not as good of a definition today as it was for the 1999 data used in [**Bushnell et al., 2008**]. ISOs are vastly different sizes and some firms that should be categorized as fringe surpass that threshold.

Ideally, I could estimate a supply relation for the competitive fringe directly, but I am missing some data. Fringe generation is any non-oligopolist generation that can respond to price: small fossil fuel generators and net imports. However, I lack data on imports into each ISO, so I construct a fringe generation variable from an accounting identity. In each market, hourly demand is supplied by oligopolists' fossil fuel generators, the fringe's fossil fuel generators, nuclear power, renewables like solar and wind, and net imports from neighboring ISOs. I use this fact to construct the fringe generation variable in Equation 2.3. I have data on the sum total of oligopolists' fossil fuel generation for each hour, q_t^{olig}, from the EPA CEMS data. Because nuclear generators run at nearly full capacity all hours of the year, I proxy for their generation with the total capacity of all nuclear generators within the ISO, k_t^{nuke}. Finally, I have data on hourly wind generation in some ISOs, q_t^{wind}. I lack data on hourly wind generation in CAISO, NYISO, and PJM. I am also missing data on hourly hydroelectric and solar generation in all ISOs. As a result, my variable for fringe generation will be too large in some hours. However, the approximation error should be small as renewables, especially solar, make up only a small portion of generation capacity in most markets and particularly markets where coal is prominent.

$$q_t^{\text{fringe}} = q_t^{\text{fringe fossil}} + \text{net imports}_t = load_t - q_t^{olig} - k_t^{nuke} - q_t^{wind} \tag{2.3}$$

For each market, I construct a panel of the 8,784 hours in 2016[12] with data on fringe generation and prices. The supply relation to be estimated is given in Equation 2.4. In addition to price, I have month, day-of-the-week, weekend, and hour fixed effects to control for persistent demand fluctuations that come from seasonal temperatures, daily, and weekly usage patterns. Because price and quantity are jointly determined by supply and demand, price is endogenous. Therefore, I estimate Equation 2.4 via two-stage least squares with the actual hourly system-wide demand as an instrument. Since demand is inelastic, at least in the short-run, the actual level is the demand shifter that can identify the slope of the supply relation.

$$q_t^{fringe} = \beta p_t + \sum_{i=1}^{12} \alpha_i Month_{it} + \sum_{j=1}^{7} \delta_j Day_{jt} + \omega Weekend_t + \sum_{1}^{24} \phi_h Hour_{ht} + \varepsilon_t \tag{2.4}$$

Table 2.2 shows the estimates for the coefficient on price for each market. Coefficients should be interpreted as the extra generation from the fringe from a one dollar increase in the market price for electricity. Naturally, MISO and PJM have the largest coefficients because of their large size. In comparison to [**Bushnell et al., 2008**], the estimates

[12]Most years have 8,760 hours, but 2016 was a leap year with one extra day. I use the extra hours for additional data points assuming that leap years do not have unusual generation patterns compared to normal years. In some cases, data was missing because of an error in collecting by the ISO. In these cases, I treat the hours as missing at random and make no correction in my estimation.

Table 2.2. Fringe Supply Curve Estimates

VARIABLES	(1) CAISO	(2) ERCOT	(3) ISONE	(4) MISO	(5) NYISO	(6) PJM
price	196.3***	435.5***	54.24***	757.1***	113.5***	625.8***
	(13.91)	(21.55)	(1.264)	(15.99)	(3.275)	(11.02)
Observations	8,478	8,782	8,784	8,784	8,425	8,782

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Notes: The supply relation for each ISO was estimated by two-stage least squares with actual hourly load as an instrument. First stage results are very strong. Regressions with less than 8,784 observations had missing data from the ISO and are treated as missing at random.

have grown larger because the increased size of the markets and increased demand.[13] For example, for 1999 data, CAISO had an estimated coefficient on price of 124.8 in comparison to the 2016 estimate of 196.3. ISONE had an estimate of 10.8 in comparison to 54.24 today. The largest difference is in PJM, where the actual size of the market has grown substantially. In 1999, the market consisted of the namesake states: Pennsylvania, New Jersey, and Maryland. The 1999 estimate of supply response to price was 8.5. Today the market has added upper Appalachia, Virginia, Ohio, Indiana, and northern Illinois, including Chicago, and the estimated supply curve slope is 625.8. Another difference between the 1999 estimates and today is that the competitive fringe is bigger. In 1999, restructuring was new and many legacy utilities still retained generating assets. Today, there are many firms that individually hold less than 5% of total capacity. These firms make a responsive competitive fringe.

[13][**Bushnell et al., 2008**] only analyzed California (CAISO), New England (ISONE), and the Pennsylvania-New Jersey-Maryland Interconnection (PJM)

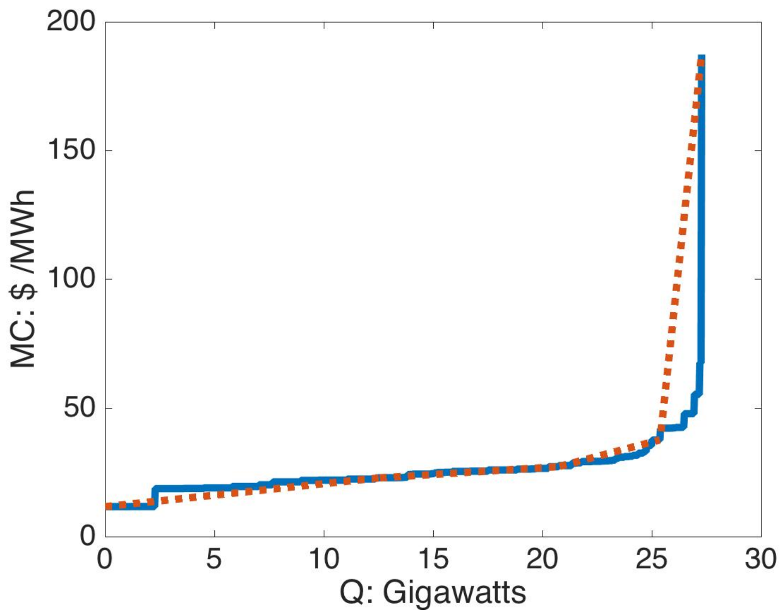
2.6.2. Spot Market Simulation

Estimating the sunk costs of retiring coal generators and much of the counterfactual requires knowing firms' spot-market profitability for different portfolios of generators. Given an estimate of the residual demand curve for each hour and the marginal costs of each firms' generators, I impose a behavioral assumption that firms compete in a capacity-constrained Cournot-Nash equilibrium. I then calculate hourly production, prices, and profitability. Annual profitability of a firm's portfolio, given the portfolios of other firms, is simply the sum of hourly profits for the entire year.

On the supply side of the market, firms' generators have marginal costs given in the data. For computational reasons, I smooth each firm's step-function marginal cost function into a piece-wise continuous function with four segments. I find the quartiles of each firm's 2010 marginal cost relation in the data to divide the segments. Then I use the data on capacity and generators' marginal costs to create a continuous piece-wise function for the firm-level marginal cost function. Figure 2.9 shows the actual marginal cost function for AES, the largest firm in PJM, in 2010, and the figure also plots the continuous, piece-wise approximation of the marginal cost function.

Hourly demand is inelastic, but oligopolists face a residual demand curve taking into account fringe generation. The slopes of the fringe generation for each ISO are estimated above in Table 2.2. I estimate the intercept of the residual demand curve using Equation 3.7 where $q_{i,t}$ is the actual hourly oligopoly production, $\hat{\beta}$ is from the estimate in Table 3.7 , and p_t^{actual} is the actual observed hourly price. The resulting estimates of hourly intercepts allow a better fit of the residual demand curve around the observed prices.

Figure 2.9. Example firm-level marginal cost function



Notes: Step function, in blue solid line, is the actual marginal cost function of AES, the largest firm in PJM, in 2010. The dotted orange line is the piece-wise, continuous approximation of the marginal cost function used in calculating sport market outcomes.

$$\hat{\alpha}_t = \sum_{i=1}^{N} q_{i,t}^{\text{actual olig}} + \hat{\beta} p_t^{actual} \tag{2.5}$$

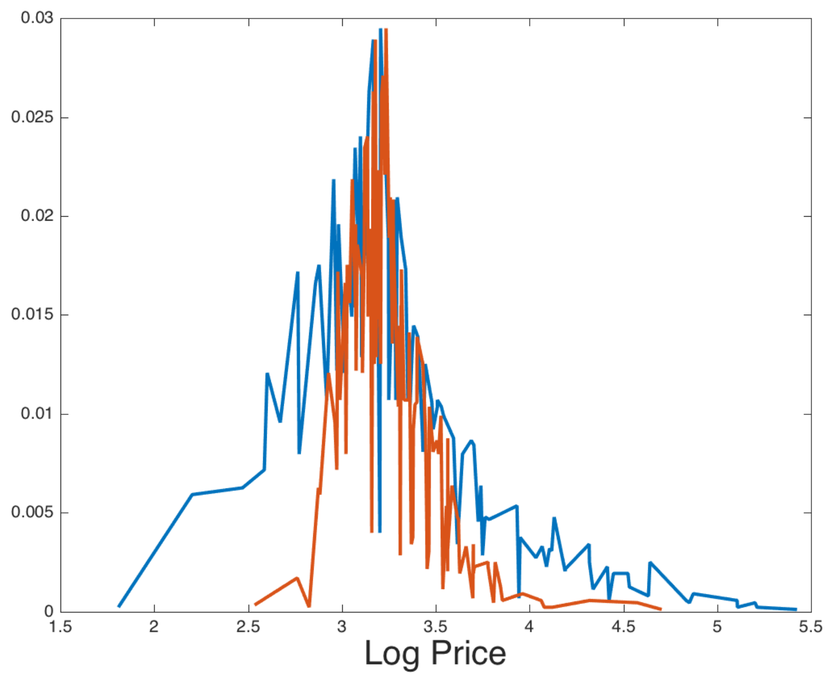
Given the estimate of the residual demand curve in each market and in each hour, $Q(p_t) = \hat{\alpha}_t - \hat{\beta} p_t$, and each firm's marginal cost function, I calculate firm level quantities and spot-market prices as the result of a capacity-constrained Cournot game. I follow [**Ito and Reguant, 2016**] and formulate the Nash Equilibrium of this game as a mixed-integer program, which simplifies calculation[14]. To further reduce the computational burden, instead of calculating the profits for all 8,784 hours of the year for each market, I

[14]See the Appendix for computational details

take a sample of 100 hours and rescale each hour to match the distribution of loads and add up to 8,784. I use a k-means clustering algorithm to find a representative sample of hours. The panel of spot market data includes hourly variables on price, total demand, and fringe generation. Of the 8,784 observations, the k-means clustering algorithm finds k = 100 hundred clusters to subdivide the data and cluster observations to the nearest mean across variables. It is similar to the k-nearest neighbor classification. I find the profits of each firm for each hour and scale by the number of true hours each sample hour represents. The result is an estimate for annual profits $\Pi_i(k_i^{2016}; k_{-i}^{2016}, c_i, c_{-i})$.

The spot market formulation relies on the assumption that oligopolists' behavior can be approximated by a capacity-constrained Cournot-Nash game. Results also depend on the designation of oligopolists and fringe players as well as the estimation of the fringe's supply relation. A good metric of model fit is how closely the predicted hourly prices match the distribution of actual hourly prices in each market. I use the actual 2016 oligopolists' generators in the data and the model to simulate outcomes for a subsample of 100 hours and compare with the actual prices. Figure 2.10 shows the comparison in PJM. In general, the distribution of predicted prices has less variance than the true distribution. At the lower end of the distribution, I am not able to predict prices as low as prices observed in the data. In reality, prices can be lower than marginal costs because fossil fuel generators will produce below marginal cost when shutting down is too costly. My model treats each hour independently and cannot replicate that behavior. I also cannot replicate very high peak prices. One reason is because the fringe supply relation is linear in prices. In reality, at higher prices, extra fringe generation may have to be imported from very far away or will be scarce. A better fit would come from adding a

Figure 2.10. PJM predicted log price distribution (orange) vs actual log price distribution (blue)



Notes: The distributions come from the sample of 100 representative hours found using a k-means algorithm. I use log prices for clarity as some prices are very high compared to the mean.

quadratic price term in the supply relation. However this would come at a computational expense because the mixed-integer program used to solve the market requires a linear set-up.

In addition to the the price distribution, another measure of model fit is examining the markups that the model predicts. Table 2.3 has the weighted averages of the true price, predicted price and predicted markups for each ISO[15]. Predicted markups are at

[15]These are estimates of hourly prices and markups, dependent on the estimate of $\hat{\beta}$ in each ISO. Reporting standard errors is computationally burdensome because resulting prices are the outcome of the structural model, which takes considerable time calculate. I do not report standard errors since $\hat{\beta}$ is precisely estimated in each market.

Table 2.3. Spot Market Simulation Fit

ISO	CAISO	ERCOT	ISONE	MISO	NYISO	PJM
Average Hourly Price: Data	29.82	21.44	29.59	24.27	25.47	27.53
Average Hourly Price: Model	28.53	26.80	33.61	25.53	25.31	26.04
Average Hourly Markup: Model	4.39%	18.73%	6.96%	10.37%	7.43%	18.37%

Notes: Prices are in MWh and are the weighted by hours of the year. Markups are the standard index of $\frac{p-c}{p}$ where c is the marginal generator producing. Markups are also weighted by hours of the year.

reasonable levels. ERCOT and ISONE have predicted prices significantly higher than actual prices. [**Bushnell et al., 2008**] finds that a model that fails to account for forward contracts will have predict higher prices. Use of forward contracts varies across markets, and absent data on contracts positions there will be some upward bias.

2.6.3. Long-Run Capacity Adjustment Costs

Long-run retirement costs and investment costs are crucial for understanding the incentives of oligopolists when considering counterfactual policies. Because most of the strategic long-run choices of firms involve coal and CCGT, I limit the choice set of the firm to just retiring coal capacity or building CCGT capacity. In the data, I aggregate all CCGT building decisions and all coal retirement decisions between 2010 and 2016, and I model them as if they happened all at once. The individual timing may have mattered for strategic reasons, but generation capacity is sticky and will last for decades, which is the focus of my analysis. Additionally, the limited strategy space is not innocuous. In reality, firms in the data make adjustments to other technologies. Firms retired many conventional natural gas and oil generators during this period, but they were expected retirements before the Shale Revolution. I model all other changes outside of building CCGT and retiring coal as exogenous to each firm.

I use an estimate for the sunk cost of building CCGT capacity provided by the EIA called the "overnight cost of capital" which is about $1 million per MW of new CCGT capacity. Less is known about the economic cost of retiring coal generators. I parameterize the cost of retiring coal using Equation 2.6. The notation $\tilde{x}_i^{coal}$ denotes the sum of all coal capacity retired. The variable cost of adjusting capacity rationalizes why firms hold onto very unprofitable coal capacity which have high operating fixed costs.

$$C(\tilde{x}_i^{coal}) = \mathbb{1}\{\tilde{x}_i^{coal} > 0\}(\varepsilon_i^{coal} + \gamma_1^{coal}\tilde{x}_i^{coal} + \gamma_2^{coal}(\tilde{x}_i^{coal})^2) \tag{2.6}$$

To identify the cost of retiring coal, I assume firms retired generators to maximize long run profits. Estimated variable adjustment parameters will best rationalize the choices made. I further assume a continuous choice set for capacity. While true adjustment is lumpy, this approximation allows for a unique optimal adjustment given by the first order necessary conditions in Equation 2.7. I add in a structural error to the first order necessary condition because I cannot perfectly predict each firms' adjustment.

$$-\frac{\partial \Pi_i(k_0^1 - \tilde{x}_i; k_{-i}^1, c_i, c_{-i})}{\partial x_i} + I_{j,coal} - \gamma_1^{coal} - 2\gamma_2^{coal}\tilde{x}_i^{coal} + u_i^{coal} = 0 \tag{2.7}$$

Since observations are at the firm-level, I have low power for estimating these parameters. In particular, there are only 15 firms in my data that retired coal capacity during the period. Table 2.4 shows the results for two specifications: a quadratic cost function and a linear cost function. Standard errors in Table 2.4 are the usual OLS standard errors,

Table 2.4. Variable Cost Estimates

VARIABLES	(1) OLS	(2) Constant
x	-13.69	
	(11.44)	
Constant	203,732***	162,107***
	(56,066)	(44,661)
Observations	15	15
R-squared	0.099	0.000

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

conditional on estimates of $\hat{\beta}$ from the second stage of the model. Since those are precisely estimated I do not correct for the two stage estimation in inference. Given the small sample size, I prefer to use the simpler linear cost specification. In the counterfactual, I model firms' decisions under the actual regulatory environment they faced, and the linear coal retirement cost matches the data well. Similarly, I use a linear capacity adjustment cost for building CCGT which is provided by the EIA. They estimate the overnight cost of capital as a linear cost of capacity.

Fixed costs, ε_i^j are not estimated. They are needed to rationalize why my model predictions do not accurately predict all investment. In particular, when I model the long-run decisions in each market, some firms make no adjustments while in reality there was some adjustment. The difficulty with estimating the distribution of fixed costs here has to do with multiple equilibria and multiple forms of heterogeneity across firms. A moment inequality approach to estimating a parametric distribution of these costs could be feasible. Regardless, a per-MW variable cost of capacity adjustment fits the data well in the counterfactual suggesting these costs fixed costs may be small in the framework.

2.7. Counterfactual

I use the two-stage oligopoly model and estimates to find the impact of coal retirements on consumers in the two largest coal markets: MISO and PJM. First, I find the capacity adjustment decisions a least-cost planner, pricing at marginal cost, would have made in response to the drop in natural gas prices from the Shale Revolution, and I contrast those with what my model predicts the oligopolists would choose. The "wedge" between these two outcomes shows how well liberalized markets are performing compared to a perfectly competitive benchmark. Second, I quantify the value of coal capacity by showing how the cost of buying electricity on the spot market decreases with more coal capacity. This is a lower bound for policy makers on any benefit to consumers of incentivizing delaying retirement of coal plants. They may also care about the reliability of coal plants or supporting jobs in the industry, which would increase the benefits of retaining coal. Finally, I find the optimal subsidy for a "second-best" planner who can incentivize oligopolists to retain coal generators with fixed-cost subsidies and wants to minimize the cost consumer's electricity bills. This would be a lower bound on the cost of retaining coal plants. Including environmental costs of added pollution would increase costs.

2.7.1. Planner's Problem

As a baseline, I consider the counterfactual long-run decisions of a cost-minimizing planner. In this setting, a benevolent planner owns all oligopolists' generators in 2010. After observing a drop in the marginal cost of natural gas, she can make costly adjustments to her portfolio. The goal of the planner is minimize the total cost of supplying electricity as in Equation 2.8

(2.8)

$$\min_x VarCosts(k^{2010} + x, y, \beta) + \quad AnnualFixedCosts(k^{2010} + x; I) + AdjSunkCosts(x)$$

Each hour, the planner utilizes her own capacity and the competitive fringe's capacity to meet the inelastic level of demand y_t at the lowest cost. Just like the oligopolists, the planner operates over the next 20 years and discounts the future according to δ. The planner must also pay annual fixed costs for her generators and any one-time sunk costs to make adjustments to her portfolio.

The planner, like oligopolists, may only retire coal capacity or add CCGT capacity. Building a CCGT generator amounts to a rightward shift of the short-run supply curve; decreasing the variable costs of producing electricity for many hours of the year. However, the short-run variable cost savings come at the expense of a sunk cost of building a new generator and paying more annual fixed costs to operate the new generator. Conversely, retiring a coal generator amounts to a leftward shift of the short-run supply curve; increasing the variable costs of meeting hourly demands. The planner must also pay the one-time sunk cost of retiring any coal generators. However, retiring coal generators allows the planner to forgo their annual fixed costs. The planner considers these tradeoffs when minimizing the total cost of meeting electricity demand.

For each market, I contrast the planner's adjustments, the true adjustments from the data, and the adjustments my model predicts. Table 2.5 shows the results from this exercise. The markets of most interest are MISO and PJM, those with the most coal capacity and retirement. The solution to the oligopolists' game of capacity adjustment

Table 2.5. Capacity Adjustments

	CAISO	ERCOT	ISONE	MISO	NYISO	PJM
Initial 2010 Coal	0	16,179	1,389	41,552	1,027	64,326
Coal Retired. - Data	0	0	264	5,325	400	16,812
Coal Retired - Model	0	0	1,389	5,198	400	13,692
Coal Retired - Planner	0	0	1,389	1,054	0	5,095
Initial 2010 CCGT	6,857	19,177	8,309	12,476	250	16,917
CCGT Built - Data	1,335	947	0	650	0	4,135
CCGT Built - Model	400	1,200	400	0	0	2,400
CCGT Built - Planner	0	800	0	0	0	0

Notes: Capacity is in MW. The top panel shows initial capacities and adjustments for coal and the bottom panel shows the analogous capacities for CCGT. The first line shows the initial aggregate capacities held by oligopolists in each market. The second line shows the true adjustments from the data. The third line shows the equilibrium adjustments the model predicts in the two-stage game. The final line shows the adjustments a least-cost planner would make with the same initial capacities.

is another measure of the model's fit. The predicted aggregate coal capacity retired fits the data well in MISO and is slightly below the data in PJM. The model predicts that oligopolists retire more coal than a planner would choose. In MISO, oligopolists retired five times more coal than a planner would choose. In PJM, oligopolists retired just under three times as much coal as a planner.

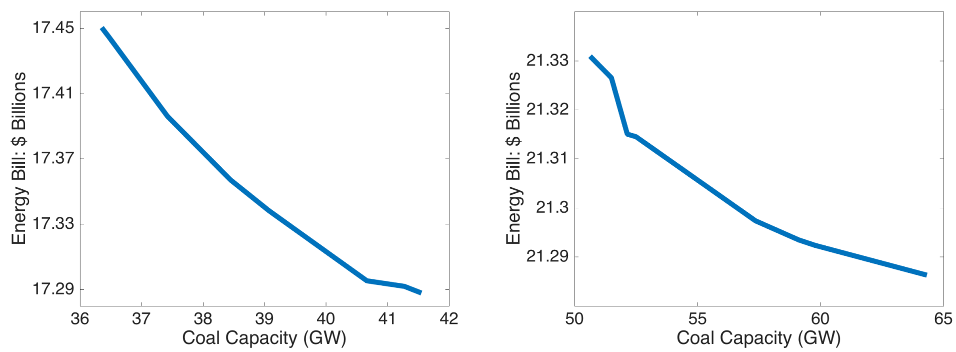
In general, the results reflect the comparative static of the theoretical model: firms with market power distort their technology mix to less baseload coal and more peaking CCGT. In all markets, the model predicts oligopolists retire at least as much coal the planner chooses. Also, the planner builds much less CCGT technology than oligopolists choose to build. The differences between planner choices and firm choices show that long run responses to the Shale Revolution are more than just adjusting technology portfolios to a change in input prices; oligopolists are distorting their long run portfolios strategically to raise prices.

2.7.2. Value of Baseload Coal Capacity

Policy makers are concerned about the unprecedented amount of coal retirements for a variety of reasons. Firms with market power will retire coal to mitigate some of the low spot market prices caused by cheap natural gas. As a low marginal cost technology, coal generators are typically inframarginal and run almost all the time. Retiring them will raise prices for all consumers, but policy makers are not sure by how much. The value of my model is that it takes into account endogenous equilibrium investment decisions and short-run production decisions as the amount of coal changes. As coal is retired, my model allows for firms to invest in CCGT or change their strategic spot-market behavior which will affect the total energy sales. This allows policy makers to see the total change in consumer surplus when considering what amount of coal retirement is reasonable.

Figure 2.11 shows the total of consumers' energy bills as a function of coal capacity for MISO and PJM, respectively. The horizontal axis ranges from the amount of coal capacity my model predicts without a subsidy up to the initial 2010 coal capacity. In MISO, where my model predicts a 13% decrease in coal capacity, retaining all capacity will save consumers $163 million dollars or about 0.93% of their total energy expenses. In PJM, where my model predicts a decrease of 21% of coal capacity, retaining all coal capacity will save consumers about $45 million dollars or about 0.21% off electricity bills. While relative savings are small, the absolute numbers put an upper bound on expenses policy makers might utilize to mitigate coal retirement driven by market power.

Figure 2.11. Consumer Energy Bill vs. Coal Capacity



(a) MISO **(b)** PJM

Notes: The total energy bill is the hourly load times hourly price, summed across all hours of the year. Energy bills drop with more coal capacity because of coal's low marginal cost.

2.7.3. Policy Tool: Coal Capacity Subsidy

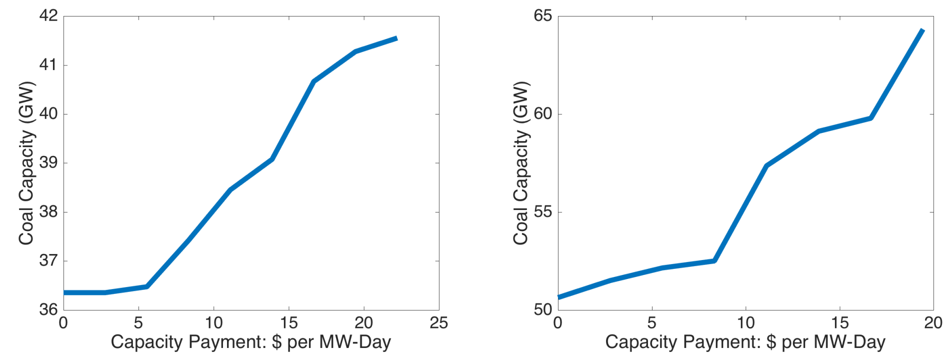
The primary policy counterfactual I study is a coal-specific fixed-cost subsidy. In the model, this decreases the annual, per-MW fixed costs that firms must pay for each of their coal generators. While proposed policies across states may not design their subsidies exactly in this form, it is still of general interest in the electricity industry. Generally, proposed subsidies of baseload plants are designed to guarantee revenue rather than be tied to production. It is a parsimonious way to account for the government's cost of incentivizing firms' retirement decisions. Also, per-MW subsidies for capacity are already used in wholesale electricity markets, and they are called capacity payments. I take these already existing subsidies as given in my model, and they apply to all fossil fuel and nuclear generators. My counterfactual is about adding an additional coal-specific capacity payment. This is one of the first empirical analyses of the efficacy of electricity market capacity payments in an economics framework.

In MISO and PJM, I consider a variety of coal-specific capacity payments. This amounts to changing the per-MW fixed cost of holding coal capacity to $I'_{ij} = I_{ij} - s\ \forall$ i, j where generator j must be a coal generator. For each s, I allow firms to start with their initial 2010 generation portfolios. They may build CCGT generators in 400 MW increments which is a standard size, and pay the sunk investment cost estimated by the EIA. They can also retire coal generators and pay the retirement cost in 2.6. Retiring a coal generator is done by setting its capacity to 0 in the model. Firms make adjustments strategically in a Nash Equilibrium: they maximize long-run profits given the adjustments of other firms. While there is a possible problem of multiple equilibria in this complete information set up, I simply solve for a fixed point and ignore the possibility of other equilibria by iterating on best response functions.

Figure 2.12 shows the resulting coal capacity as a function of the subsidy for MISO and PJM, respectively. To simplify calculations, I solve for Nash Equilibria for only a few candidate levels of s. In MISO, a coal capacity payment of around \$15 per MW-day would halve coal retirements, and all coal could be retained with a payment of \$23 per MW-day. These payments would be on top of capacity payments that all generators, including coal, can get which ranged from \$2.99 to \$72 in 2016. In PJM, a coal capacity payment of \$12 per MW-day would halve coal retirements, and all coal could be retained with a payment around \$20 per MW-day. In the most recent capacity price auction, generators in PJM earned payments between \$76.53 and \$187.87. While these payments are expensive, they are effective as policy tools for shaping long-run portfolios.

When considering incentivizing firms to retain coal plants, regulators should consider both the costs and the benefits. Figure 2.13 plots the spot market savings of retaining

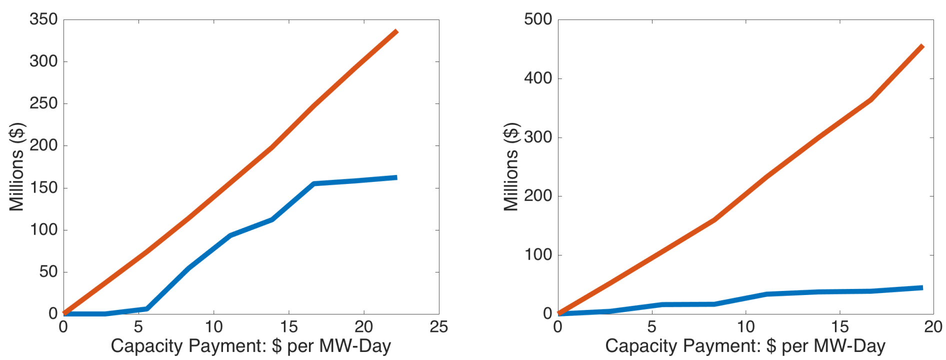
Figure 2.12. Coal Capacity vs. Subsidy



(a) MISO **(b)** PJM

Notes: Each panel shows the equilibrium aggregate coal capacity of oligopolists after the Shale Revolution for different coal capacity payments.

Figure 2.13. Costs and Benefits of Capacity



(a) MISO **(b)** PJM

Notes: The orange line shows the annual cost of paying the capacity payment to all coal generators in the market. The blue line shows the spot market savings of the additional coal capacity that was incentivized to forgo retirement

coal plants with a subsidy and the cost of the subsidy. At all levels of the subsidy, the costs of retaining capacity are more expensive than the spot market savings from that capacity. While there is no ideal subsidy in this case, the magnitudes of the graphs are still informative for policy makers.

The benefit of retaining coal capacity only includes spot market savings from the low marginal cost technology, and including other factors may raise the curve. One reason these savings are low is because the difference between the marginal cost of coal and the marginal cost of CCGT is very small right now. If the cost of natural gas were to increase, the spot market benefit of coal would increase (also the necessary subsidy to retain coal would decrease). Natural gas prices could increase if the United States relaxed regulations in exporting natural gas and domestic prices elevated closer to the world price. Restrictions on hydraulic fracturing, the drilling method to extract shale gas, would also limit the supply and raise prices. Reliability is another benefit excluded from the analysis. The Department of Energy recently directed FERC to ensure that reliability is adequately compensated in energy markets.[16] Coal generators are one of the most reliable technologies and retaining capacity for resource adequacy would also increase the benefit curve.

The cost of retaining coal capacity is the cost of paying the technology-specific subsidy to all coal generators. While controversial, regulators could lower the cost by paying higher subsidies to generators more likely to retire and lower subsidies to generators less likely to retire. Of course a discriminatory subsidy is complicated by information asymmetries between state regulators and firms' knowledge of their own costs. However, the cost of retaining coal capacity may already be too low because I have excluded environmental externalities from my analysis.

[16] http://insidelines.pjm.com/pjm-responds-to-doe-proposal/

2.8. Conclusion

I have analyzed the impact of the Shale Revolution on long-run outcomes in electricity markets and how market power distorts firms' reactions to those of a least cost planner. Overall, firms with market power are retiring more coal generators than a planner would choose. They do this to strategically limit some baseload capacity and earn higher inframarginal rents on their remaining units.

I utilized a two-stage model of strategic interaction to analyze this effect. Firms are able to exert market power by committing to long term capacity decisions and utilizing their generators strategically in the short run. Intervening in the long-run outcomes by incentivizing firms to change their capacity decisions has modest savings for consumers.

From a policy perspective, "bailing" out coal plants cannot be justified as a tool to lower consumers' energy bills in the current environment. More research is needed on the technology's value as a reliable resource. Also, other policy channels such as antitrust may be more effective and inexpensive.

CHAPTER 3

Anti-Trust and Common Ownership: Evidence from the New England Electricity Market

3.1. Introduction

Institutional investors, companies that manage mutual funds, index funds, and other assets on behalf of their customers, have become a major part of the equity market. Typical investment products offered by institutional investors pool customers' money together to buy many different stocks or replicate an index such as the S&P 500. These investment products benefit customers because they are low-cost, capture the growth of the overall market, and diversify against the idiosyncratic risk of any particular firm's performace. Institutional investors have succeeded with these products, and their market share has risen from about 5% in 1950 to over 50% in 2010 [**Rydqvist et al., 2014**]. By managing so many assets, they have become the largest owner in nearly 90% of public companies in the S&P 500 meaning they hold board seats in many firms and can influence management [**Azar, 2017**].

As institutional investors increasingly become the largest owner of many publicly traded firms, they own major stakes in *rival* firms within an industry. For example, Berkshire Hathaway recently bought major stakes in the four largest airlines: American Airlines, Delta, Southwest, and United. There are several reasons to invest in competing firms rather than pick the most undervalued stock. In long-term investing, diversifying across firms benefits investors by mitigating risk. However, scholarship suggests that firms with common owners have less incentive to compete, harming consumers ([**Bresnahan and Salop, 1986**] and [**O'Brien and Salop, 2000**]). Models of imperfect competition, such as the classic Cournot model, typically predict that total industry profits under oligopoly competition (duopoly, triopoly, etc.) are less than the monopoly profit. Therefore, managers of firms with common owners would serve their owners' financial interests by cooperating with rivals rather than competing.

Antitrust authorities and researchers are giving more attention to the effects of common ownership. European antitrust authorities have considered factoring in common owners when considering a merger.[1] In academic research, there is a growing literature about the effects of common ownership. [**Azar et al., 2017**] found evidence that increased common ownership has increased average prices on airline tickets by about 4%, and [**Azar et al., 2016**] found similar evidence in the banking industry. [**Posner et al., 2016**] brings evidence to policy, and recommends limiting portfolios of institutional investors to mitigate the effects of common ownership. However, [**O'Brien and Waehrer, 2017**] argue that current research results, which identify the effects of common ownership on price with reduced form approaches, may not be robust enough yet to recommend intervention. Model misspecification, unobserved confounders, and endogeneity of ownership may lead to an incorrect conclusion that common ownership does indeed increase prices. The argue that structural modeling approaches, which can incorporate richer detail on cost, demand, and firms' incentives, may have more robust results.

[1] http://www.oecd.org/daf/competition/common-ownership-and-its-impact-on-competition.htm

In this paper, I test for evidence that common ownership increases prices in the New England electricity market. The wholesale electricity market is a well-suited setting for finding evidence that common ownership affects prices. Electricity markets have well known market power problems and characteristics that facilitate collusive behavior. I follow previous literature and find reduced form evidence that common ownership has increased the average electricity price by 5.5%. I compare these reduced form results with results from a novel structural model. Importantly, many of the economic primitives of the model, such as cost, are in the data and do not need to be estimated. I use the model to test the *behavioral* assumption that firms account for common ownership when making strategic decisions. If firms do indeed account for common ownership, the average price of electricity is 8.6% higher than if firms maximize their own profits in the model. However, the model's predictions are closer to the true data under the assumption that firms simply maximize their own profits and *do not* account for common ownership. In a setting where firms do account for common ownership and investors control firms in proportion to their ownership stakes, I find that capping ownership in rival firms to 1% saves consumers $3.3 billion over the 6 year time frame analyzed or 6.3% of their total expenditures. Total surplus also increases under this regulation.

This paper adds to two distinct literatures. First, it adds to the growing literature on the effects of common ownership. My reduced form results find that common ownership increased prices in electricity markets on a similar magnitude as [**Azar et al., 2017**] finds in airline markets and [**Azar et al., 2016**] in banking. [**O'Brien and Waehrer, 2017**] have valid critiques of the reduced form approaches to identify how much common ownership raises prices. The contribution of this paper is to address their criticism by analyzing the problem using a structural model. The structural model explicitly maps how demand, cost, and firm incentives lead to market outcomes. This paper also adds to the literature on modeling electricity spot markets. I generalize the framework developed in [**Bushnell et al., 2008**], which models the electricity market as a Cournot game, to include the possibility that firms put weights on the interests of their different owners. The paper is also similar to [**Hortascu et al., 2017**] and [**Hortaçsu and Puller, 2008**] because it exploits the rich data on economic primitives in electricity markets to identify behavioral parameters.

3.2. Theories of Common Ownership

Typical models of imperfect competition, such as Cournot, assume that a firm's manager and the firm's owners have the aligned incentive to maximize firm profits. However, if two rival firms have the same owner, they have less incentive to compete. In the extreme case of duopolists merging to form a monopoly, the usual symmetric Cournot model predicts that the joint firm charges the monopoly price, which is higher than the pre-merger duopoly price. Also, the monopoly profit of the merged firm is greater than the sum of the two duopolists profits. Most models of imperfect competition predict that any merger where one firm takes 100% ownership of a rival firm lessens competition. But the change in the competitive environment is less clear when a firm takes *partial* ownership in a rival firm.

Early work on competition and ownership control over rival firms was done in the context of joint ventures. [**Bresnahan and Salop, 1986**] derive a concise metric to measure competition where firms embark on a joint venture to produce output. The optimal output produced and the market price depend on both the financial interests the firms have in the joint venture as well as the control each firm has over the joint venture. The theoretical framework results in a generalization of the common Herfindahl-Hirschman Index (HHI) which they call Modified HHI (MHHI). [**O'Brien and Salop, 2000**] generalized the framework for measuring the degree of competition when third party investors own parts of many firms in the same industry.

The theoretical framework is driven by how much managers weigh the payoffs of their different owners. Equation 3.1 shows the objective function of the manager of one of the output producing firms, j. Instead of simply maximizing firm j's profits, the manager maximizes a weighted sum of all the profits of institutional investors, $i = 1 \ldots M$, that own firm j. The weight γ_{ij} represents the relative control owner i has over manager j to favor i's profits over other partial owners. The objective function is a generalization of typical models where $\gamma_{ij} > 0$ for investors who only have a stake in j and no other rival firms.

$$\max_{q_j} \sum_i \gamma_{ij} \pi^i = \max_{q_j} \sum_i \gamma_{ij} \Big(\sum_k \beta_{ik} \underbrace{[P(Q)q_k - C_k(q_k)]}_{\pi_k} \Big) \tag{3.1}$$

The right side of Equation 3.1 explicitly shows how common ownership can influence manager j's decisions. Investor i's profits are the sum of profits she receives from all output producing rival firms $k = 1 \ldots N$. The weight β_{ik} is the actual ownership stake that investor i has in firm k. Strategies that maximize the profit of firm j usually lower profits of other firms and total industry profits. Because these profits appear in the objective function of firm j through common owners, the manager's strategies may be less competitive. Equation 3.2 shows the first order necessary condition for profit maximization, which explicitly shows this incentive. Like in the typical, sole-ownership Cournot model, firm j trades off the profit from selling another unit with lowering the market price it gets for all units sold ($MR_j - MC_j$). But in the common ownership model, there is an extra effect of lowering the market price and revenue of all other rival firms $k \neq j$. For each rival firm, k, this effect is weighted by how much j's owners have in common with k's owners and how much the manager of firm j cares about those owners. This term, which is negative for downward sloping demand curves, amounts to a downward shift of the marginal revenue curve for firm j. Thus when the manager of firm j accounts for common ownership, the optimal quantity produced, q_j^* will be lower than under sole ownership, and the market price will be higher.

$$\underbrace{P(Q) + P'(Q)q_j}_{MR_j} + \underbrace{\sum_{k \neq j} \frac{\sum_i \gamma_{ij}\beta_{ik}}{\sum_i \gamma_{ij}\beta_{ij}} P'(Q)q_k}_{\text{Common Owner Effect}} = \underbrace{C_j'(q_j)}_{MC_j} \tag{3.2}$$

This Counot, homogenous goods model yields a connection between Modified HHI and the equilibrium price just as there is a connection between HHI and price in the typical, sole-ownership Cournot Model.[2] The share-weighted first order necessary conditions for profit maximization of each firm can be added up and rearranged to yield the relationship expressed in Equation 3.3. The constant price elasticity of demand times a market share weighted average of markups is equal to the MHHI. The MHHI can be broken up into

[2]The derivation is similar to the result in [**Cowling and Waterson, 1976**] that the share-weighted average margin times the demand elasticity is equal to the HHI

the sum of the usual HHI and a term called $MHHI\Delta$, the difference between MHHI and HHI. $MHHI\Delta$ captures the degree in which common ownership increases markups. If there is no common ownership, each investor i has financial interests in one firm j and no other firm $k \neq j$, then this part of the $MHHI\Delta$ term goes to zero.

$$\eta \sum_j s_j \frac{P - C_j'(q_j)}{P} = \underbrace{\sum_j \sum_k s_j s_k \frac{\sum_i \gamma_{ij}\beta_{ik}}{\sum_i \gamma_{ij}\beta_{ij}}}_{MHHI} = \underbrace{\sum_j s_j^2}_{HHI} + \underbrace{\sum_j \sum_{k \neq j} s_j s_k \frac{\sum_i \gamma_{ij}\beta_{ik}}{\sum_i \gamma_{ij}\beta_{ij}}}_{MHHI\Delta} \tag{3.3}$$

Because institutional investors own rival firms in the majority of industries, the control weights γ_{ij} are the key part of the model that differentiate the outcomes of this model from typical, sole-ownership models of imperfect competition. If $\gamma_{ij} > 0$ for only the investors that do not have ownership stakes in rival firms, $k \neq j$, then there is no additional markup from common ownership. If there is any weight on investors with common ownership, there will be higher prices. The key empirical questions are how to identify these weights and how are the weights determined in the real world.

Previous empirical research ([**Azar et al., 2017**] [**Azar et al., 2016**]) assumes the weights management puts on the interests of the firms' owners are precisely the voting shares of each owner, which is called proportional control. They argue this is a good measure of weight on management decisions because there is evidence that owners with larger voting shares have more ability to communicate with management. They document that large investors are able to arrange private meetings with management, design executive compensation, and they vote on the board of directors who can influence management directly.

[**O'Brien and Waehrer, 2017**] argue that proportional control may not be a reasonable assumption for the control weights. Many large institutional owners are only minority investors. The largest owner of a publicly traded firm typically owns just 3-7% of total equity. If the majority of equity is owned by smaller investors who own no rival firms, the manager who makes decisions according to proportional control does so to the detriment of the majority of her owners. They also argue it is difficult to incentivize a manager to weigh industry profits over firm profits. Typically executive compensation is a fixed salary plus stock options which should reward individual firm performance. However [**Kwon, 2016**] argues that as common ownership increases, individual firms' stocks are more correlated with industry profits. Regardless of the explicit mechanism to incentivize managers to account for common ownership, critics argue proportional control may not be the true reflection of owners' sway. Instead, the weights could be more abstract such as favoring "the squeaky wheel" and activist investors.

3.3. Electricity Market as a Setting

There are many characteristics of wholesale electricity markets that facilitate firms' abilities to exert market power. Firms' marginal costs are publicly available, they compete in markets every hour of every day, and their production decisions are made public soon after production. These ingredients are ideal for enforcing a cartel, and they are still helpful for tacitly exercising market power. Market regulators are constantly vigilant for market manipulation and allegations are made regularly. Notoriously, electricity generation companies exerted market power and exacerbated the California Energy Crisis at the beginning of restructuring[3] [**Borenstein, 2002**]. Because of the ability of firms to exert market power, the electricity market is a good candidate to test for evidence that common ownership further increases markups.

3.3.1. Wholesale Electricity Markets

Electricity is produced from a variety of technologies that convert primary energy into electrical energy. Fossil fuel generators, such as coal, natural gas, and oil, have the largest market share, producing over 70% of electricity in New England. Nuclear generators have a 13% market share. Finally, renewables, mostly hydroelectric dams, but increasingly wind and solar, make up the rest. Generated electricity is transmitted long distances over high voltage wires and then distributed to homes and businesses. While the transmission and distribution part of the industry is regulated at the state level, generation of electricity is a liberalized market in New England and much of the United States.

Independent System Operators[4] (ISOs) control the wires. ISOs are responsible for procuring power to meet demand, which is considered inelastic, for every hour of the day. They run auctions where firms bid supply curves for each generating unit. The ISO aggregates supply curves and gets energy from the lowest bidders up to the point that supply meets demand. Each unit that "clears" the market receives the price of the marginal bid. Typically, nuclear and coal generators, with low marginal costs, can clear the auction every hour and they supply power all year. Natural gas plants, with

[3]Prior to the late 1990s, electricity markets were regulated in a vertically-integrated regulated monopoly framework.

[4]Sometimes these are called Regional Transmission Organizations. The difference between the two classifications is not important for the analysis in this paper.

higher marginal costs and flexible production, run seasonally or during peak hours. Oil generators with high marginal costs come online only during peak hours.

Electricity markets are concentrated with the majority of electricity produced by a handful of firms. These firms typically own many generators of all different technology types. In addition to large firms, there are several small firms with just one or a few generators. Municipalities, which largely supply to their own jurisdiction, can also sell excess energy in the wholesale market. Finally, industrial firms with on-site generation, such as steel or lumber processing plants, can sell to the grid as well.

There are several large interstate ISOs in the United States. I focus on the New England ISO (ISONE) from 2006-2011 for several reasons. First, the same set of firms operated the same generators over the sample period, so that detecting effects of common ownership on price should come from changing ownership shares rather than other confounding factors. Second, there was one large ownership change when a firm that was owned completely by a private equity company was sold to a public company which had owners in common with the other energy firms in ISONE. This creates substantial variation in common ownership patterns, and one can identify an effect on price. Third, the data from the ISO is well maintained and readily available over many years. Finally, I lack data on renewable generation, but there was very little renewable generation capacity in ISONE during this period, meaning measurement errors from generation data should be small.

3.3.2. Data

Data on the electricity market comes from several sources. The Energy Information Administration (EIA), which is the statistical arm of the U.S. Department of Energy, is the main source of data on unit level generator characteristics. Almost every generator in the United States must frequently report data to the EIA, which then becomes available to the public. EIA-860 reports annual information on existing, retired, or planned generators with over 1 megawatt (MW)[5] of combined plant capacity. Each unit reports its capacity, generation technology, and ownership. EIA-923[6] collects monthly generation and heat rate (which is used to construct marginal costs) data from a broad survey of plants covering almost all in the sample. The particular data in this paper comes from SNL Energy, a subsidiary of S&P. SNL cleans the EIA's data and facilitates searching for generator characteristics. Hourly generation unit level production data comes from the Environmental Protection Agency's (EPA) Continuous Emission Monitoring Systems (CEMS). Finally, hourly market prices and quantities in the New England market come from the ISONE website.

Data on firm ownership comes from the Thomson-Reuters Spectrum dataset on 13F filings. The Securities and Exchange Commission (SEC) requires that all institutional investors managing more than $100 million file their portfolio holdings on a quarterly basis.[7] The dataset has some gaps when institutional investors go bankrupt or the stock tickers of firms in the portfolios change. In this paper, NextEra Energy changes its stock

[5] 1 MW can supply about 1000 homes. The average capacity of a generator is well over 1 MW
[6] Formerly EIA-906 and EIA-920
[7] This is a low threshold to cover institutional investors of interest. As of September 30, 2017 BlackRock managed $5.7 trillion, Vanguard managed $4.5 trillion, and State Street managed $2.8 trillion

ticker from "FPL" (NextEra was formerly Florida Power & Light) to "NEE" in the first quarter of 2011, meaning I am missing ownership data for this firm for 2011. I added in NextEra's stock holdings from the third quarter of 2011, which is available on the Nasdaq website through the WayBack Machine. I have extrapolated these holdings to the other three quarters of 2011.

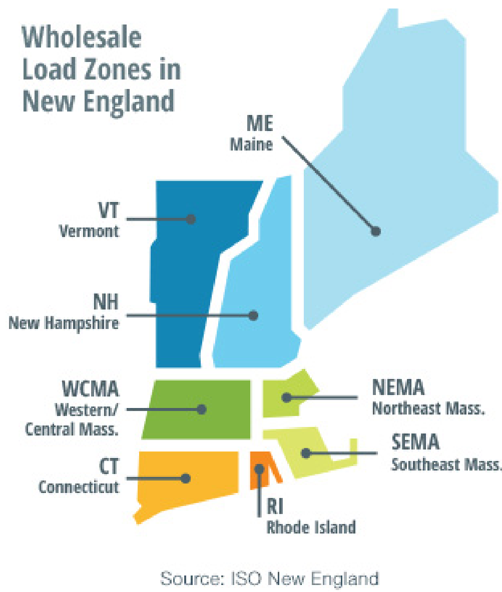
3.3.3. ISONE

ISONE covers the six states of New England: Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island, and Vermont. Figure 3.1 shows there are different load zones which can have different wholesale prices, but generally prices are similar across the ISO and electricity can be transported anywhere almost instantly. Throughout the paper, I consider the whole ISO as one large power pool and one large market. I use an average market price provided by ISONE and the total market quantities. In addition to the electricity generated and consumed in New England, power can be imported from and exported to the neighboring markets in New York (NYISO) and Canada. New England tends to import power from hydroelectric dams in Canada. While I lack data on imports, I construct a measure from an accounting identity[8].

There are dozens of firms in ISONE but I distinguish large firms as oligopolists and the remaining firms are the competitive fringe. I use a cutoff of 5% of total quarterly market share to divide the groups. This is in line with a Texas rule that "small fish swim free"

[8]The total load in the system, or total quantity demanded, equals the total supply. The supply comes from production within the system (fossil fuel generators from oligopolists and fringe players, nuclear and renewable generation) and imports from outside. $q^{demand} = q^{olig,fossil} + q^{fringe,fossil} + q^{nuclear} + q^{renew} + q^{imports}$. Then $q^{fringe} = q^{fringe,fossil} + q^{imports} = q^{demand} - q^{olig,fossil} - q^{nuclear} - q^{renew}$

Figure 3.1. ISONE Coverage Area



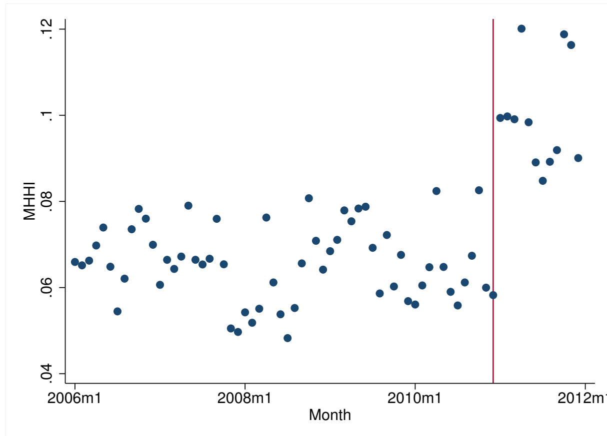
and firms wth less than 5% market share have a safe harbor from market manipulation charges. Over the period 2006-2011 there are 4 major firms that meet this threshold.[9]

The large firms are heterogenous in the generators they hold and their market share. Dominion is the largest firm with around 25% market share and it is publicly owned. NextEra (formerly Florida Power & Light) has around 12.5% market share and is also publicly owned. US Power Generating Company[10] also has around 12.5% market share. Importantly, it is owned by a private equity fund and assumed to have no owners in common with the other electricity generation firms. In 2010, the private equity firm declared bankruptcy for unknown reasons, but a bankruptcy court mandated the sale

[9] In some quarters, there are some firms that meet this threshold but I do not categorize as them oligopolists. In 2006 Q1, Eversource Energy has a 6% market share, but never gets above 5% again. In 2011 Q3, Capital Power Corp gets 5% market share. In 2011 Q1-Q4 Energy Capital Partners has an 8% market share.

[10] At the start of the sample, the firm is EBG Holdings LLC. It is bought by US Power which calls its New England subsidiary Boston Generating LLC

Figure 3.2. Average monthly MHHI



Notes: The figure shows the average monthly MHHI in ISONE. The red vertical line is between December 2010 and January 2011 when US Power Generation's assets were sold to Constellation which had owners in common with other firms in ISONE.

of the New England generation assets to pay off creditors. All assets were bought by Constellation Energy and operated in 2011. Constellation is publicly owned and has owners in common with the other firms, thus expanding the degree of common ownership in ISONE. Figure 3.2 shows how this acquisition dramatically increased the degree of common ownership. Entergy is the smallest oligopolist with just 10% market share. For most of the sample, Entergy only owns two large nuclear generators, which are largely nonstrategic. However, Entergy shares owners with the other publicly traded firms, so its presense incentivizes the other firms to raise prices if managers take common ownership into account. Entergy also bought 3 peaking plants from NextEra in 2011 Q4.

The Thomson-Reuters data gives the firms' institutional investors. Table 3.1 shows the top ten owners of each firm for 2011 Q3. At this point, Constellation bought US Power's generation assets. This quarter is a typical example of the ownership data. The largest shareholder owns about 5-7% of the firm. Firms may have owners in common with every other competing firm (State Street is at least the fourth largest owner in each of the firms), just a subset of competing firms (T. Rowe Price is the largest owner in two firms, but not a top 10 owner of the other firms), or may have an owner that has no other stakes in the industry (LSV Asset Management is only present in the top 10 owners of Constellation Energy). Each quarter, ownership stakes vary as institutional investors shift their ownership stakes. Prior to 2010, US Power (later bought by Constellation) is soley owned by itself. For each firm and quarter, I choose only owners with at least a 1% stake in the firm. I then reweigh the proportional control measures so they sum to 1. In previous literature using this dataset, [**Azar et al., 2016**] and [**Azar et al., 2017**] use data on voting shares to construct ownership control weights (γ_{ij}) and total ownership shares to construct ownership weights (β_{ij}). Because I am missing data on voting shares for NextEra for four quarters of data, I do not distinguish voting shares from regular shares so all $\gamma_{ij} = \beta_{ij}$, the total proportion of shares owned.

Table 3.1. Major owners of ISONE Energy Firms

Dominion Resources	
Capital Research Global Investors	5.74
State Street Corp	4.90
BlackRock Inc.	4.30
Vanguard Group, Inc.	4.00
Wellington Management Co, LLP	3.65
Barrow Hanley Mewhinney & Straus.	2.83
Northern Trust Corp	1.73
Franklin Resources Inc	1.36
BlackRock Advisors	1.18
Mellon Bank NA	1.04

NextEra Energy	
Wellington Management Co, LLP	7.88
State Street Corp	7.41
Vanguard Group, Inc.	5.96
FMR, LLC	4.90
Franklin Resources Inc	4.53
J.P. Morgan Chase & Co.	4.07
Barclays Bank PLC	3.68
Mellon Bank NA	2.92
Wells Fargo	2.65
Northern Trust Corp	2.07

Constellation Energy	
T. Rowe Price Associates, Inc.	7.34
State Street Corp	5.76
Vanguard Group, Inc.	5.21
BlackRock Inc.	4.41
Fidelity Management & Research	3.35
Pictet Asset Management Ltd.	2.62
J.P. Morgan Chase & Co.	1.87
Donald Smith & Co., Inc.	1.84
LSV Asset Management	1.44
Franklin Resources Inc.	1.36

Entergy	
T. Rowe Price Associates, Inc.	7.10
Franklin Resources Inc	6.52
BlackRock Inc.	6.17
State Street Corp	4.97
Evercore Trust Company	4.26
Vanguard Group, Inc.	4.17
Barrow Hanley Mewhinney & Straus.	3.56
Pzena Investment Management LLC	1.84
First Eagle Investment Management LLC	1.82
Amvescap PLC London	1.75

Notes: The table shows the top 10 owners of each firm for 2011 Q3. All data is from Thomson-Reuters except for NextEra which is missing from the data set. The top ten owners from NextEra are from an archived version of the Nasdaq website.

3.4. Reduced Form Approach

Equation 3.3 shows the connection between equilibrium price, $MHHI\Delta$, and HHI in the theoretical framework. It motivates finding the causal effects of $MHHI\Delta$ and HHI on price as in the reduced form Equation 3.4. p_t (\$/MWh) is the price of electricity for a given hour t. HHI_t is the standard Herfindahl-Hirschman Index of the sum of squared market shares of the oligopolists within the hour.[11] $MHHI_t\Delta$ is the difference between HHI and $MHHI$ for the given hour. While part of this term is constructed from ownership control parameters, γ_{ij} and β_{ij}, which only update quarterly, these parameters interact with "cross" market shares $s_j s_k$, which vary by the hour. When Equation 3.4 is interpreted causally, θ is the effect that increasing the degree of common ownership has on prices. The null hypotheses is that there is no effect: $\theta = 0$.

$$p_t = \theta \cdot MHHI\Delta_t + \eta \cdot HHI_t + controls + \varepsilon_t \tag{3.4}$$

I control for several variables that affect the prices observed in electricity markets. First, demand for electricity fluctuates throughout the day and seasonally. Within a day, demand is predictably low when most people are asleep and peaks in the evening when many people are home using electronics and appliances. Within a year, demand is higher in the summer and winter when there is greater need for heating and cooling. I add in hour-of-day, day-of-week, month-of-year, and weekend fixed effects to control for these systematic fluctuations. I also include dummies for each year to control for longer term changes in demand and competitiveness of imports, which are affected by long-run investments in capacity and transmission. Weather is an important predictor of energy use and I control for it with data on hourly temperatures and dew points. I add in the squares

[11] The market shares do not add up to 1. There are dozens of other small producers in the market (including unobserved importers of electricity), and I assume their squared market shares are 0. I do this because I do not have ownership data for each small producer and I do not observe the market shares of individual importers. Because other firms have very small market shares, their contribution to HHI should be small

of these variables for a more flexible effect, and I also interact temperature with the hour-of-day dummies. Electricity markets are supplied by a variety of generation technologies, but natural gas generators are typically the marginal, price-setting technology. I add in monthly spot market fuel prices for natural gas from the Henry Hub pipeline system, which is a common index used to quote natural gas prices.

Identifying the causal effects in a reduced form equation like in 3.4 is notoriously difficult. HHI_t and $MHHI\Delta_t$ are functions of market shares, which are endogenous. Additionally, the ownership shares of firms may be endogenous. For example, investors increase ownership of firms when they think prices will be higher. I ignore endogeneity problems in the main specification. In previous work, [**Azar et al., 2017**] find that OLS results, ignoring endogeneity problems, were close to the results of several other identification strategies with a causal interpretation.

3.4.1. Panel Regression

Table 3.2 reports estimates for different specifications of Equation 3.4. Specification (1) shows there is a negative raw correlation between $MHHI\Delta$ and price as well as HHI and price. Theory predicts that there should be a positive relationship between these variables. The negative correlation comes from fluctuating demand and capacity constraints. Demand in the electricity market is considered inelastic in the short run. Therefore, when demand goes up, prices go up. But, as demand grows, the oligopolists in the data hit capacity constraints and realize more inframarginal rents. Without the ability to flexibly respond to increased demand, the market share of the oligopolists decreases

Table 3.2. OLS Results for Equation 3.4

VARIABLES	(1) Main	(2) Load	(3) Peak	(4) Peak + Load
MHHI_Delta	-65.62***	85.34***	-72.19***	63.32***
	(22.93)	(17.51)	(21.01)	(13.49)
MHHI_Deltapeak			-169.6***	29.26
			(43.36)	(27.79)
HHI	-29.64**	81.53***	37.14**	51.61***
	(12.49)	(9.939)	(17.50)	(10.04)
HHIpeak			-505.0***	171.1***
			(106.7)	(32.99)
load		0.00774***		0.00598***
		(0.000485)		(0.000365)
loadpeak				0.00491***
				(0.000676)
Observations	52,387	52,387	52,387	52,387
R-squared	0.901	0.909	0.903	0.910

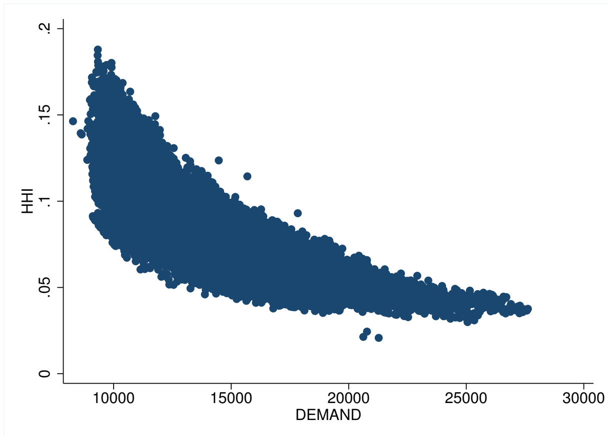
Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Notes: This table shows the OLS regression results for the specification given in Equation 3.4. Data is all hourly prices in ISONE from 2006-2011 with some hours missing at random. All specifications control for hour-of-day, day-of-week, month-of-year, and yearly fixed effects. There are also controls for hourly temperature, hourly dew point, and the monthly price of natural gas. Standard errors are clustered by the hour-of-day.

compared to lower demand hours. Therefore, HHI is actually lower in high demand, high price times. Figure 3.3 shows this relationship.

Specification (2) adds the market demand (called load in the industry) as a control. Typically, regressing price on quantity suffers from endogeneity problems. However, total hourly demand is inelastic; it is plausibly exogenous to price. Moreover, this control is important to the interpretation of the results. Because HHI and MHHI are lower when

Figure 3.3. HHI vs. Demand



demand is higher, θ and η should be the partial effects of an exogenous change in their respective variables conditional on a level of demand. Predictably, the coefficient on load is positive and significant. The coefficients on $MHHI\Delta$ and HHI become positive and significant when load is controlled for. The results in specification (2), the preferred specification, mean that an increase of $MHHI\Delta$ from 0 to 0.07 (going from no common ownership to the average level in the data) increases the price of electricity \$5.97. The average price of electricity in the data is \$57.76. Similarly an increase of HHI from 0 to 0.08 (going from perfect competition to the average level in the data) raises the price of electricity \$6.52.

Because peak hours are different than normal hours, specifications (3) and (4) explicitly control for $MHHI\Delta$ and HHI having different effects on price during peak hours. I define a peak hour as being above the 85th percentile of demand. In specification (3), there is still a negative correlation because of capacity constraints. Specification (4) controls for quantity demanded and the results are still positive.

Table 3.3. OLS Results for Equation 3.4 with $\log(p_t)$ Dependent Variable

VARIABLES	(1) Main	(2) Load	(3) Peak	(4) Peak + Load
MHHI_Delta	-2.140***	0.783***	-2.072***	0.765***
	(0.388)	(0.209)	(0.366)	(0.224)
MHHI_Deltapeak			-0.729	0.822*
			(0.461)	(0.414)
HHI	0.570*	2.388***	1.181***	2.211***
	(0.294)	(0.269)	(0.343)	(0.318)
HHIpeak			-4.542***	0.603
			(1.016)	(0.525)
logload		1.976***		1.876***
		(0.0702)		(0.0822)
logloadpeak				0.507***
				(0.0637)
Observations	52,387	52,387	52,387	52,387
R-squared	0.995	0.996	0.995	0.996

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 3.3 shows the analogous results as Table 3.2 but has a log-level specification. This is the specification chosen in [**Azar et al., 2017**]. Just as above, controlling for load matters. In Specification (2), the preferred specification, the coefficient on $MHHI\Delta$ means an increase of $MHHI\Delta$ from 0 to 0.07 means a 5.5% increase in price. The coefficient of 0.783 on $MHHI\Delta$ compares to a coefficient of 0.194 in [**Azar et al., 2017**]. However, their average level of $MHHI\Delta$ was larger, so they predict common ownership raises prices by 4%.

3.4.2. Robustness

Previous literature, [**Azar et al., 2016**] and [**Azar et al., 2017**], try several identification strategies to address the problem of reverse causality. Instead of prices rising because of increased common ownership, it could be that savvy investors anticipate rising industry prices and profits, so they increase their holdings across the industry. The literature tries to address this concern by regressing price on leads and lags of HHI and $MHHI\Delta$. They find that the coefficient on lags is significant and positive, meaning an increase of common ownership does indeed lead to high prices in the future. And the coefficient on leads is not significant meaning that it's difficult to say whether investors predict high prices in the future so they increase their holdings today. This exercise is difficult to replicate in this setting because of lack of data. In previous studies there was a large cross section of markets in the panel data, while I only study one market with 24 quarters of data. There is not enough data to find leads and lags of ownership changes.

The previous literature also studied an event where ownership was plausibly exogenous. They exploited the bankruptcy of Barclays during the Global Financial Crisis and the subsequent sale of its assets to BlackRock. This greatly increased BlackRock's ownership of airlines and increased MHHI. As long as airlines were not the main driver of the acquisition, then the ownership change is plausibly exogenous. Unfortunately, Barclays and BlackRock did not hold enough stakes of the energy firms in ISONE to create a distinguished shift in MHHI from the same acquisition. A similar event study using the sale of US Power Generation assets to Constellation Energy does increase MHHI significantly as Figure 3.2 shows. However, finding comparable hours prior to and after the "treatment" is difficult because of the seasonal and volatile nature of electricity markets. Also, the bankruptcy is less plausibly exogenous.

3.5. Structural Model Approach

There are many criticisms of using reduced form approaches to find the causal effect that increased common ownership has on price. In addition to the endogeneity of price and HHI and $MHHI\Delta$, [**O'Brien and Waehrer, 2017**] argue the specification in Equation 3.4 is not tied to the theoretical relationship in Equation 3.3, which also involves the price elasticity of demand and costs. Moreover, they argue that theoretical models can predict negative relationships between the definition of $MHHI$ and price. They argue for a structural approach to identifying the effects of common ownership on price. They outline an approach that estimates economic primitives like demand and cost as well the weights that firm management puts on its owners (γ_{ij}).

The electricity industry is an ideal setting to test for common ownership having an effect on price by directly estimating the weights firm management puts on its owners' profits. Unlike many other industries, there is good data on costs and demand is inelastic meaning only weights must be estimated. Previous literature [**Bushnell et al., 2008**] shows that prices and firm-level production can be simulated from imposing a Cournot competition game on the largest firms in the market, and the simulated results fit the true data well. I generalize the structural model of the electricity spot market by allowing for firms to care about their common owners. In an ideal framework, I would then estimate the γ weights directly from the data. Instead, I test two distinct candidate γs. The first assumes that firms maximize their own profits only; $\gamma_{ij} > 0$ for owners who only own one firm. The second model assumes proportional control and uses γ_{ij} that reflect the ownership shares of each owner in each firm. While this might not be the true γ, any amount of weight on common owners will result in higher prices. If that assumption helps better explain the true prices, then it is consistent with the argument that firms put at least some weight on common owners' interests.

3.5.1. Spot Market Model

While electricity generation firms actually compete by bidding supply functions on an hourly basis, I follow the framework of [**Bushnell et al., 2008**] and model the spot market for electricity as a capacity-constrained Cournot game.

Each oligopolist j, has a number, n_j, of generators with capacities described by $k_j \in \mathbb{R}^{n_j}$. Each generator has a constant marginal cost so that firm-level marginal costs are given by a corresponding vector $c_j \in \mathbb{R}^{n_j}$. Oligopolist j's generator capacities and marginal costs compose a firm-level marginal cost function, $MC_j(q_j; k_j, c_j)$, which is a weakly increasing step-function.

In addition to strategic suppliers of electricity, there is a competitive fringe that supplies electricity at marginal cost. The fringe's supply relation is linear in price $q_t^{fringe} = bP_t$

Each hour t, there is a perfectly inelastic and perfectly forecastable level of demand y_t. N oligopolists and a competitive fringe compete to supply this level of demand.

Oligopolists maximize the objective function given in Equation 3.1. That is they maximize the weighted profits of all their owners. In the special case where firms maximize their own profits, $\gamma_{ij} > 0$ for one and only one owner i who in turn only has an ownership stake in firm j. This is equivalent to maximizing firm level profits only.

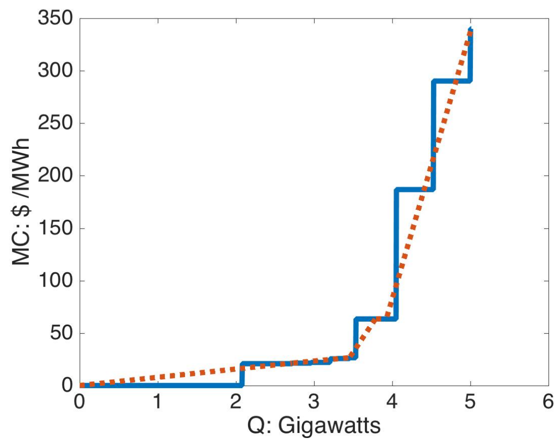
3.5.2. Estimation

As is common in the electricity industry, marginal costs of generators are data. For each unit, Equation 3.5 shows the two major parts of marginal cost (MC). The cost of fuel makes up the vast majority (usually more than 90%) of marginal cost. The heat rate (HR) is generator specific and measures how much fuel (mmBTU) it takes to generate 1 MWh of electricity. While heat rates can vary on a monthly basis, I am missing data on some months. I use the average heat rate across all heat rates I observe for each unit. Multiplying the heat rate by the price of the unit's fossil fuel ($/mmBTU) yields the fuel marginal cost of the unit. I have monthly coal, natural gas, and oil commodity data from the EIA. I use the quarterly average commodity price to compute the fuel portion of the marginal cost for each unit in each quarter. This is important as the price of natural gas halves in 2008 because of developments in extracting shale gas in the United States. In addition to fuel costs, there are variable operating and maintenance costs (VO&M) such as pumping and treating water for cooling. I use estimates of these costs from SNL, the data provider.

$$MC = HR \cdot P_{fuel} + VO\&M \tag{3.5}$$

I construct continuous, piece-wise linear approximations of firm-level marginal cost step-functions for each quarter. I use 5 segments for each firm and group generators by their quintiles. I then create 5 piecewise continuous segments to approximate the true curve. The approximation helps calculating the model in two ways. First, the model

Figure 3.4. Firm-Level Marginal Cost Function



Notes: The figure shows the true and approximated marginal cost functions for Dominion Energy in 2006 Q1. The blue step function is the true marginal cost with generator level costs constructed from Equation 3.5. The orange dotted curve is the approximation to the true curve with five piece-wise continuous segments.

will not give a unique price prediction when firms are capacity constrained, but price is not high enough to turn on the next marginal generator. Second, the model is faster to solve when all firms have the same number of generators. The 5 continuous segments are like firms having 5 generators with weakly upward sloping marginal costs. In cases where firms had less than 5 generators,[12] I split capacity over several identical generators.

While oligopolists make up the majority of generation in New England, there are dozens of other firms that supply electricity and imports into the market. I chose oligopolists as the firms that have more than 5% market share in each quarter over the period. The rest of the generators are part of the competitive fringe. While some of these firms are merchant power producers with a few generators, most are industrial firms like timber and paper plants. They typically generate electricity for their own use, but

[12] Entergy has just two nuclear generators for most of the sample

Table 3.4. Supply Relation Slope Estimate

	(1)
VARIABLES	Main
price	111.9***
	(1.583)
Observations	52,584

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Notes:

they can supply electricity to the grid. While I have these generators' marginal costs, I do not observe the true opportunity costs of selling electricity rather than self-supplying and producing manufactured goods. Therefore, I parameterize and estimate their supply relation as in Equation 3.6. I use the same controls as when estimating Equation 3.4.

$$q_t^{fringe} = bP_t + controls + \varepsilon_t \tag{3.6}$$

Of course, quantity and price are determined simultaneously, so Equation 3.6 suffers from an endogeneity problem likely to bias the estimate of b. Typically supply curve estimation requires an instrumental variable that shifts the demand curve. It must affect the market quantity, but not through price. Since demand is inelastic (at least in the short run), the actual hourly load is the demand shifter and a valid instrument. Table 3.4 shows results of estimating the supply relation slope over the full sample using two-stage least squares.

While demand, y_t is inelastic and given in the data, I construct an estimate of the intercept of the linear residual demand curve, $\hat{a}_t$. Equation 3.7 shows the intercept is the sum of actual oligopolist production in the data and the estimated fringe supply, the estimated slope of the demand curve multiplied by the actual market price. I use this intercept and estimated slope as the demand curve when calculating the results of the capacity constrained Cournot model. The intercept does not have a direct interpretation. However, if the model is correctly specified and the estimate of the slope is the true parameter, then the resulting endogenous variables, P_t and Q_t, should be close to the true outcomes in the data P_t^{data} and Q_t^{data}. The extent to which they are off could be because the demand curve is misspecified, the estimated slope of the demand curve is off, or the competitive assumption of the strategic setting is misspecified.

$$\hat{a}_t = \sum_j q_{jt}^{data} + \hat{b}P_t^{data} \tag{3.7}$$

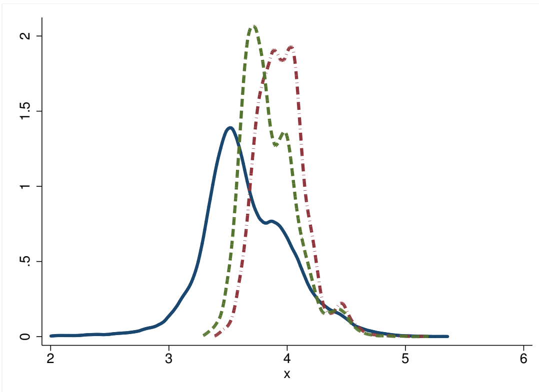
With data on marginal costs and an estimated residual demand function, I solve a Cournot model. I use the first-order conditions from the objective function in Equation 3.1. I represent the whole problem as a mixed-integer program with the first-order conditions of the firms and capacity constraints as constraints. The full set up is given in the Appendix.

3.5.3. Simulation Results

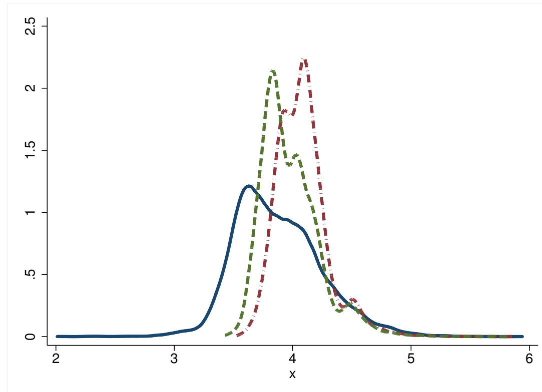
Figure 3.5 shows the distributions of log prices for the true data and the resulting log prices from the model. I divide the results into four groups based on the quartiles of market load. In general, the model cannot reflect the true spread of the data. On the left side, observed prices can be lower than anything the model predicts (including negative). The model will never predict a generator producing at lower than marginal cost, but in reality low prices exist from renewable generation and imports. On the right side, the model struggles to predict high prices. One reason is the linearity of the residual demand curve. At higher levels of demand (when there are higher prices) it should get more difficult to get more generation from the fringe (many are capacity constrained). I use a linear demand curve for computational reasons.

Model fit improves as market load increases. Table 3.5 shows the mean price in the data, the mean price assuming sole ownership, and the mean price assuming common ownership with proportional weights for each quartile of load. It also shows the mean-squared error (MSE) comparing the price resulting from a modeling assumption and the true data. In general, the sole owner assumption fits the data better, having the lower MSE. However, the fit for the common ownership assumption improves as the market load grows. In the first, second, and third quartiles, the structural model overpredicts the average prices for both assumptions. This is on average, and there are some hours where the model underpredicts prices for both assumptions. In the fourth quartile, the average price under the sole ownership assumption underpredicts the true average price and the average price under the common ownership assumption overpredicts the average price. [**Bushnell et al., 2008**] find that this structural model will overpredict the true price without data on contracts. Forward contracts, which are a major part of selling electricity, are pro-competitive and reduce market prices. In their analysis, adding contract data made the model have better fit by reducing the predicted prices. Because I am missing

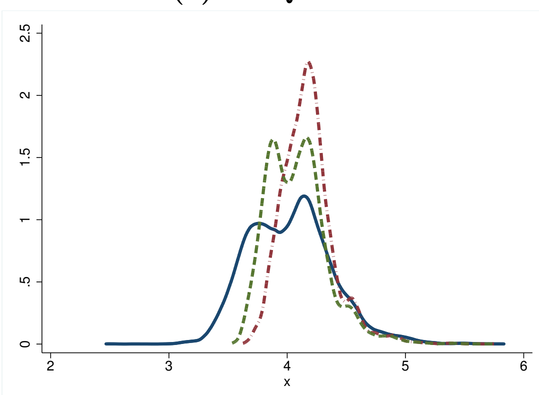
Figure 3.5. Distribution of Log Prices Data vs. Models



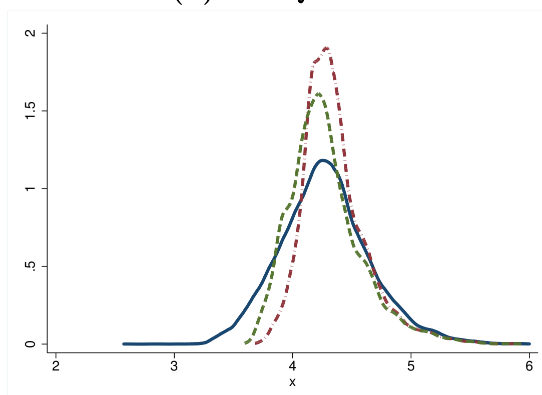
(a) 1st Quartile



(b) 2nd Quartile



(c) 3rd Quartile



(d) 4th Quartile

Notes: The figure shows the distribution of log prices for the true data and the results of the model. Each panel represents data for hourly loads broken up by quartile. The blue, solid lines are the distributions for the true data. The green, dashed lines are the distributions assuming sole ownership. The red, dash-dot lines are the distributions assuming common ownership

data on forward contracts, I expect predicted prices to be upward biased. If predicted prices were lower, given contract data, then the common ownership assumption may have better fit in the fourth quartile. This mean shift may extend so far to mean the common ownership assumption would better explain the data in all quarters.

Table 3.5. Structural Model Fit

	Q'tile 1	Q'tile 2	Q'tile 3	Q'tile 4
$\bar{p}_t^{data}$	41.37	51.78	59.81	78.10
$\bar{p}_t^{sole}$	48.15	55.27	61.45	76.51
$\bar{p}_t^{common}$	53.05	60.61	66.77	81.61
MSE Sole	90.74	68.57	55.26	52.27
MSE Common	188.52	148.83	118.11	84.98

Notes: The table shows the mean prices in the data and implied by the model under assumptions of sole or common ownership. The mean-squared errors are also reported. Each statistic is given within a quartile of total load demanded.

3.5.4. Monte Carlo Experiment

The structural model allows me to generate simulated electricity market outcomes for all types of behavioral assumptions. Therefore, I can examine if the reduced-form specification in Equation 3.4 can detect higher prices from common ownership in the ideal setting where I control firm behavior. If the reduced-form approach used above successfully identifies the effect of common ownership in the ideal setting, then there is hope the same method can be applied in other industries. The reduced form approach is easier to apply to a variety of industries in contrast to employing idiosyncratic structural models. However, failure to accurately detect the effect of common ownership on prices in an ideal setting would further validate concerns of using reduced-form approaches raised in [**O'Brien and Waehrer, 2017**].

For each hour in the data, I simulate equilibrium quantities produced and prices under two different assumptions. First, I assume that firms only maximize their own profits. Second, I assume they care about their common owners, according to proportional control. I then re-estimate the reduced-form model by OLS for each market with all of the controls mentioned before. Table 3.6 shows the results of the Monte Carlo experiment. In column 1, where the data was generated by firms who only maximize their own profits, the coefficient on MHHI is positive but not significant. The result is correctly inconclusive. The coefficient on HHI is positive and significant which leads to the correct conclusion that the ordinary form of market power (a few firms have significant control over the price) is a problem in this market even when firms do not acknowledge their common owners. In column 2, where the data was generated from firms maximizing the profits of their common owners, the coefficient on MHHI is positive and significant leading to the correct conclusion that common ownership does lead to higher prices in the setting. The coefficient on HHI is also positive and significant but about half the magnitude of the coefficient in column 1.

The results of the Monte Carlo experiment support using the reduced form approach commonly used in the literature. However, additional approaches such as a complementary structural model will add more credibility. As [**O'Brien and Waehrer, 2017**] point out, the reduced form specification may still falsely result in positive and significant effects of MHHI on price.

Table 3.6. Monte Carlo Results

VARIABLES	(1) Sole	(2) Common
MHHI_sole	18.61	
	(31.35)	
HHI_sole	480.7***	
	(66.02)	
load	0.00615***	0.00609***
	(0.000429)	(0.000344)
MHHI_common		469.1***
		(24.89)
HHI_common		280.2***
		(44.51)
Observations	52,578	52,578
R-squared	0.949	0.961

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Notes: The table shows the reduced form regression results of the Monte Carlo Experiment where the reduced-form model in Equation 3.4 is estimated using data from the structural model. Column 1 is from the structural model where firms only maximize their own profits. Column 2 is from the structural model where firms take common ownership into account

3.6. Counterfactual

Proponents of more stringent anti-trust enforcement argue that common ownership must be limited. Limiting common ownership is weakly better for consumers. If common ownership does indeed raise prices, then consumers will benefit from regulation. However, if firms only maximize their own profits, the regulation will have no effect on spot market prices. The cost of regulation is on the investor side. If common ownership raises prices and profits for rival firms, then returns for investors will decrease when common ownership is limited. Even if common ownership does not raise prices, limiting investor portfolios could increase risk as investors cannot diversify their holdings. Understanding these benefits and costs in any intervention to reduce common ownership is important for policy makers.

[**Posner et al., 2016**] argues for limiting institutional investors' stakes in firms with clearly defined markets. In their proposal, investors can invest freely in one firm in a market, but investment in any additional firm in the market is limited to 1% of total equity. They argue that the welfare for consumers with this regulation will outweigh any costs to investors.[13]

I calculate the welfare effects of this regulation using my structural model. In each quarter, I allow an investor's largest holding in an electricity firm to stay constant. If the investor has any other large stakes in a competing firm, I reduce the stakes to 1%. Holdings that are less than 1% remain the same. I then recalculate the model for each hour in the data and find the prices and quantities produced.

Table 3.7 shows the results of the exercise. Reducing common ownership reduced the average market price by $4.26. Consumer demand is inelastic in the short run. So in both scenarios they consume the same quantity of electricity, but they get reduced prices with the regulation. Because consumer demand is inelastic, measuring consumer surplus in either scenario is not meaningful. However, the savings in total electricity bills is the change in consumer surplus, and it is $3.294 billion dollars over the 6 year period. Oligopolists' production is elastic. With the regulation, they produce more and the market price is lower. Their producer surplus is their revenue less the variable cost of production, and, as expected, it goes down with the regulation: $663 million over the

[13] A major part of balancing welfare gains and losses is that a larger share of the population participates in a given market than owns equities.

Table 3.7. Limited Portfolio Comparison

	Proportional Control	Limited Portfolio	Change
$\bar{p}_t$ ($ / MWh)	65.51	61.25	4.26
Consumer's Exp ($ Bill.)	52.319	49.026	3.294
PS - olig ($ Bill.)	16.768	16.104	0.663
PS - fringe ($ Bill.)	14.803	13.121	1.682

Notes: The table summarizes the main results of the counterfactual

6 year period. The regulation also affects the fringe and reduces their producer surplus too. The total welfare of the market, consumer surplus less the total producer surplus (oligopolist and fringe), increases with the regulation.

Analyzing the regulation requires quantifying the loss to investors, not simply producer surplus. Mapping firm profits to equity valuations is outside the scope of this paper. Also, the ownership structure after the regulation was imposed rather than endogenously determined by investors. Investors might be able to increase firm profits by optimizing portfolios subject to the regulation. Also, there may be dynamic responses in long term strategies of electricity generation firms. They have the ability to exert market power through strategic investment or disinvestment in capacity which may change after the regulation.

3.7. Conclusion

This paper analyzes the effects of common ownership on price in the New England electricity market. In line with the previous literature, reduced form results find that increasing common ownership raises prices. However, a structural model taking into account firms' costs and incentives finds the data is more consistent with firms maximizing their own profits.

In an environment where firms in New England are proportionally controlled by their major shareholders, a proposed antitrust regulation to limit common ownership has a big impact. Consumers save more than 5% on their electricity bills and total surplus goes up.

CHAPTER 4

Capacity Markets and Outcomes in PJM

4.1. Introduction

Electricity generators earn the vast majority of their revenue from selling energy. However, in some electricity markets, generators can earn additional revenue from providing capacity, regardless of how much energy they produce. Capacity payments are becoming more widespread across the globe and an increasingly important source of revenue for generators. In the Pennsylvania-Jersey-Maryland Interconnection (PJM), the largest electricity market in the United States and one of the largest in the world, capacity charges to consumers are about 20% of the total electricity bill and have ranged from $5.2 billion in 2009 to $10.9 billion in 2017.[1] Despite the growing economic significance of capacity payments, the topic is still controversial among energy policymakers.

In markets that allow capacity payments, the stated goals of the payment are to ensure reliability and shape long-run investment. Sufficient capacity is important in electricity markets because total supply must meet demand at every moment to avoid power outages. Because consumers do not respond to price in real time, regulators prefer that suppliers have reserve excess capacity in case of an unexpected increase in demand. Capacity payments compensate suppliers for investing in this reserve and are like a form of insurance. While capacity payments could be set by a regulator to incentivize investment, many

[1] http://www.pjm.com/-/media/about-pjm/newsroom/annual-reports/2016-annual-report.ashx?la=en

electricity markets in the United States determine capacity payments in annual auctions. The goal of a capacity auction is to elicit the lowest capacity payment to procure the desired reserve margin of capacity. Critics of capacity payments argue that in standard economic models, energy sales alone provide enough incentive to invest in capacity to cover peak demand. “Energy-only” markets, where generators only earn revenue from energy sales, are still the predominant market structure in liberalized electricity markets across the globe. In general, they do not seem to suffer from significant reliability problems. Understanding the outcomes of both regulatory structures is important to minimize the cost of electricity for consumers.

The purpose of this paper is to analyze the PJM capacity market and its impact on the PJM energy market. PJM is an important market to study because it is a large market with over 50 million customers and it has one of the longest running capacity markets with payments starting in 2007. Analyzing capacity markets in the United States is difficult because there is no data on suppliers’ bids and which generators actually cleared the market. Instead of analyzing the capacity auction itself, I explore whether the resulting payments are shaping the long-run investment or disinvestment outcomes in PJM. I exploit a regulation that allows resulting capacity prices to be higher in certain areas of PJM. Using this regulation as a plausibly exogenous increase in capacity prices, I test to see if higher capacity price regions have more relative investment. I do not find convincing evidence to support this hypothesis.

This paper contributes to the small but growing literature on capacity payments in electricity markets. Early work discusses the theoretical need for capacity payments and the ideal design of a market: [**Joskow, 2007**], [**Cramton and Stoft, 2005**],[**Cramton**

and Stoft, 2006], [**Bushnell, 2005**]. The closest related paper to this one is [**Bushnell et al., 2017**], which compares the broad outcomes between markets with capacity payments, energy-only markets, and traditional regulated utility structures.

The contributions of this paper are twofold. First, I describe the goals of capacity payments and a typical capacity auction. I develop a simple framework to analyze the efficiency of the auction, and I find that outcomes are far from efficient. There is a possibility of finding a less costly mechanism to accomplish the same investment goals. Second, I analyze the effect of capacity prices on investment within a market. This is in contrast to [**Bushnell et al., 2017**], which compares outcomes across markets and regulatory regimes. Using within market variation avoids unobserved and meaningful differences between markets. I do not find convincing evidence that higher capacity prices cause more investment, but a richer analysis with better data is needed.

4.2. Capacity Payments

Electricity markets are vast networks of transmission lines that connect consumers to energy generators. In a typical market design, electricity generators compete in auctions to supply electricity for a given delivery hour. There are many different generation technologies with varying costs that can produce electricity. Nuclear generators have low marginal costs and typically supply electricity all year. Fossil fuel generators, which burn coal, gas, and oil, have higher marginal costs. Coal and gas generators can run all year, seasonally, or just during peak times. Oil generators generally have the highest marginal costs and run for just a few hours a year. Renewable generators, like solar and wind, have the lowest marginal costs and can undercut any fossil fuel generator, but only when the sun is shining or the wind is blowing. Hourly demand is inelastic, at least in the short run as consumers do not have the ability to know the current price and curtail consumption or they have contracted for a set price. Therefore, for each delivery hour, generators offer bids to supply electricity for a set amount of demand. The auctioneer orders bids from least to greatest until supply meets demand. The market price is determined by the last generator to clear the auction. As demand increases, higher marginal cost generators can clear the auction and bid higher prices.

Traditionally, generators earn revenue by selling electricity and ancillary services [2]. This is called an "energy-only" market. The total revenue earned by a generator must cover its total costs: variable costs, annual fixed costs,[3] and sunk costs of investment. Usually generators only supply electricity when the price is above marginal cost meaning variable costs are covered. Any additional revenue (when price is above marginal cost) must recoup annual fixed costs and sunk costs of investment. Low marginal cost suppliers frequently earn inframarginal rents, which help to cover their fixed and sunk costs. Low marginal cost suppliers typically have high fixed and sunk costs so earning inframarginal rents is important. High marginal cost suppliers typically have low fixed and sunk costs, so earning smaller inframarginal rents less often is enough to cover costs. In the long-run, the portfolio of generators adjusts as new plants enter, when they can earn enough to cover their costs, and as plants exit when they can't recoup their costs.

[2] Ancillary services help maintain the flow of electricity throughout the grid. The revenues are usually very small compared to energy sales, so I abstract away from these sales and focus on revenues from selling electricity only throughout the paper. In 2016 in PJM ancillary services averaged about \$1/ MWh while energy sales averaged \$30/MWh http://www.monitoringanalytics.com/reports/PJM_State_of_the_Market/2016/2016q1-som-pjm-sec10.pdf

[3] Salaries, rent on land, environmental permitting, etc.

Generally, competitive markets can attain long-run efficiency (least-cost outcomes) with suppliers earning revenues from the product or service market alone. In the energy-only market design described above, more efficient generation units have an incentive to enter and undercut the current portfolio of generators, lowering the cost of the overall system. Moreover, all demand will be covered by supply even though demand is inelastic. A shortage of supply would cause an enormous price spike because consumers are inelastic. The prospect of price spikes, even if for only a few hours a year, provides enough incentive for investment in new generators to cover all forecasted demand.

Many liberalized electricity markets around the world have an energy-only market design, but some policy makers have different objectives than attaining the least-cost generation portfolio. There are many other goals, but common objectives are to provide excess capacity to completely mitigate the risk of a shortage, implement price caps to limit the ability of suppliers to exert market power over inelastic consumers, or increase the amount of renewable energy sources in the market. While there are many ways to achieve these goals, subsidizing investment through capacity payments is becoming a common part of many markets.

The importance of reliability is one of the main concerns of regulators and a leading reason for capacity payments. Electricity markets must have supply exactly equal demand at every moment in time or else there will be power outages. Unlike in many other markets, consumers of electricity are largely inelastic so there will be shortages and power outages if there is not sufficient supply. Regulators consider power outages to be a significant risk. Electricity is a major input in most of the economy and its absence can be life-threatening in some cases. Therefore many policy-makers argue that supporting excess capacity, which may never operate in the energy market, with capacity payments is like paying insurance to avoid the damage of a power outage.

Regulatory price caps are another reason for capacity payments. Because consumers do not respond to price, there have been extraordinary price spikes in electricity markets. Regulators do not want consumers to pay extremely high prices for electricity when consumers would have curtailed usage if they had known the price. Therefore, many electricity markets have price caps to reduce the ability of suppliers to exert market power. While price caps can reduce market power, they also distort investment incentives. High prices signal scarcity and the need for investment. Without adequate returns on investment, generators may not build the additional capacity that would prevent future price spikes. Price caps also create problems for generators that already exist. [**Joskow, 2007**] shows that even the least-cost portfolio will *lose* money with price caps, and could cause unwanted exit of generators. This is called the "missing-money" problem in electricity markets. Capacity payments can incentivize building the least-cost portfolio while maintaining price caps.

The growing penetration of renewable energy generators in electricity markets is another reason for capacity payments. Renewables such as solar and wind have almost zero marginal cost but very high sunk costs of investment. Therefore, when renewables can supply the entire market, the market price is close to 0. Fossil fuel generators, which have higher marginal costs because of the cost of fuel, cannot compete with a large fleet of renewables. However, renewables are not as reliable as fossil fuels. If the wind is not blowing or the sun is not shining, then fossil fuels must cover the market. Therefore, a market with significant renewable penetration may also need the same fleet of fossil fuel generators that would exist, endogenously, without competition from renewables as backup. But this backup fossil fuel generator fleet would never be able to recoup all their fixed and sunk costs because they cannot competitively produce when the renewables are active. Capacity payments would help support investment in backup fossil fuels.

Critics of capacity payments argue that there are plenty of alternative solutions to the problems above that do not require additional non-energy market revenue for generators. First, demand for electricity is becoming more elastic. Sophisticated consumers of electricity, such as heavy industry, are increasingly entering in “demand-response” contracts where they are paid to reduce consumption in times of peak demand. This avoids price spikes and decreases the likelihood of power outages from a shortage. Also greater connectivity across grids can mitigate the risk of shortages, even when there is high renewable penetration. Demand is likely to be less correlated across larger areas because of differences in temperature, weather, sunlight, and economic needs. Therefore there is likely to be more excess capacity as the grid size grows. Additionally, weather patterns that affect renewables’ operations are likely to be less correlated. So if wind is not blowing on one side of the grid, another part could transport electricity there. In the sense that capacity payments are like paying for insurance against power outages, increasing the geographical footprint and connectivity across grids is like expanding the risk pool of insurance markets.

4.3. Capacity Markets

Regulators use capacity payments to support sufficient reserve capacity. Finding the least-cost way to compensate this capacity is a challenging problem. For example, regulators could pay new generators to be held in reserve, but the cost of new entry may be unobserved by the regulator. Also new generators may be more efficient in the energy market and displace older, less efficient generators. The regulator could compensate these older generators to remain in case of a shortage. Finding the least-cost capacity payment and which generators to pay is complicated by the private information of suppliers. Just like other economic problems of allocation and private information, capacity payments are usually determined in auctions called capacity markets.

Many real-world capacity auctions have a similar framework. First, regulators set a target level of capacity to procure in the auction. The target is usually based on expected peak demand for the delivery period (usually a day, month, or year) plus a reserve margin. This target is like the demand curve in the capacity market. It can be vertical like a completely inelastic demand curve, but some markets have a downward sloping demand. This reflects the regulator's lower willingness to pay for excessive levels of capacity past expected peak demand. It also limits the ability of suppliers to drive up capacity prices when there is a shortage of capacity.

Generators, or market participants, submit supply function bids for every unit of capacity they own. Because firms can own many generators (and electricity markets tend to be highly concentrated), this is done at the firm level. Supply function bids are cast privately and simultaneously. The regulator aggregates supply functions from least to greatest creating a supply curve for capacity. The price at which supply crosses demand becomes the daily capacity price and is paid to all units of capacity that "cleared" the auction. Capacity payments are made to winning generators whether they produce energy or not. However, many markets require production from these units in the event of an emergency or limit their offers in energy markets to reasonable levels. Losing generation units do not receive a capacity payment but can still participate in the energy market.

The typical capacity auction framework described above awards payments to many generators than just the desired reserve capacity. Often more is paid to the set of generators that are expected to produce and earn revenues in periods of peak demand than the set of generators that are expected to remain idle in reserve. Moreover, all generators are paid the same payment, regardless of their expected net revenues in the energy market. While some information rents are expected in auctions, there is a little known about how far typical capacity payments arising from these auctions are from a first-best, full information case. I develop a framework to study this question in a stylized setting and propose an alternative auction design to reduce payments while still supporting reserve capacity.

4.3.1. Model

I consider a setting with two electricity generation firms, A and B. Each firm owns a generator with one unit of capacity. Peak demand over the year is two units meaning the two units are sufficient to avoid power outages. Each firm's generator has a cost that is private information but known to be drawn independently from a uniform distribution: $C_A, C_B \sim U(0, \frac{1}{2})$. I abstract away from the details of the energy market except that the market provides enough revenue for the firms to cover their costs and there is no opportunity for entry of another generator. That is, a firm that builds a new generator will not be able to cover its investment cost.

The government wants the firms to build another one-unit capacity generator to keep in reserve. Each firm has an idiosyncratic fixed-cost of building a new generator independently drawn from a uniform distribution: $F_A, F_B \sim U[\frac{1}{2}, 1]$. The cost is private information for the firm.

Following the typical capacity auction framework, firms are allowed to submit bids for each generator they own: $b_A^1, b_A^2, b_B^1, b_B^2$. The government's target capacity is three units, so the three lowest bids clear the auction and the highest bid among those becomes the capacity price paid to the generators that clear. The new payment incentivizes the winning firm (the firm that cleared two generators) to build a reserve generator. The government caps the bids of each plant at the top of the support for its cost distribution. Legacy plants cannot bid more than $\frac{1}{2}$ and new plants cannot bid more than 1.[4]

The optimal strategy for each generator is to bid zero for its legacy plant $b_i^1 = 0$ and bid $b_i^2(F_i) = \frac{2}{3} + \frac{1}{3}F_i$. The full proof is in the appendix. The intuition is that bidding zero on the legacy plant guarantees a payment and bidding higher cannot change the payment. Indeed many generators in real capacity auctions bid all their capacity at zero. The bid for the potential new plant is the strategic bid. It is increasing in its cost because the bid must cover the cost of the plant. When the firm considers raising this bid higher, it trades off the benefit of a larger payment (which it will get two of) against the chance that the higher bid outbids its opponent and loses the opportunity to receive a second payment. Note that the minimum bid from the potential new plant is $\frac{2}{3}$ which already

[4]The results of the auction allow the the bid cap of legacy plants to be higher. In practice, a competitive check on these bids would be more bidders.

greatly exceeds many potential fixed costs. Also the optimal bid from the highest possible fixed cost, $F_i = 1$ just covers the cost of entry, but still secures a profit of 1 because the winner receives two payments.

The government must pay three generators the winning bid which is a total payment of $2+\min\{F_A, F_B\}$. This is, more than twice the cost of entry of the most expensive plant. Paying for inframarginal plants makes the outcome of the auction very expensive in this setting. Designers of this capacity market argue that paying inframarginal generators (generators that are not part of the reserve) is necessary because these generators still do not earn sufficient revenue in the energy market to cover their fixed costs and sunk costs of investment. In this setting, the costs, C_i, are the revenue deficiency of the legacy plants. The stylized model shows that the result of the auction is not changed by these inframarginal plants. The auctioneer learns nothing about the costs of these plants in the mechanism. A cheaper mechanism that retains these plants and incentivizes building a new plant would be to just give the legacy plants $\frac{1}{2}$ each and pay one firm 1 to build a new plant. This shows that there is room for an alternative mechanism than the current model to procure the same capacity and give less information rent,

4.3.2. Other Potential Problems

In addition to the high rents paid out of the capacity auction, there are several other potential ways that firms could manipulate the results of the capacity auction. The first is entry deterrence. Electricity markets are known to be highly concentrated and have high markups. Firms enjoying oligopoly profits in a market might forgo capacity auction profits to prevent entry of new generators which could undercut them in the energy market. They could collude to keep capacity auction prices low by bidding in their capacity at zero while making high profits in the energy market.

Capacity prices could also be inflated because of "strategic supply reduction." Because firms typically own many generation units, they have an incentive to manipulate their bids by strategically losing. For example, given an allocation of bids, a firm may have several units that "clear the auction." Pushing up the capacity price by hoping to be the marginal bidder may be unlikely. However, removing one clearing generator (by bidding at a unrealistically high price) will mean a "leftward" shift of the bid supply function from that point. The marginal generator is likely to "move up" and mean a higher clearing price. This is a profitable strategy as long as the gain in revenue from the remaining cleared generators is greater than the lost revenue from the one generator that fails to clear. [**Doraszelski Katja Seim Michael Sinkinson Peichun Wang et al., 2016**] analyzed this bidding behavior in a FCC procurement auction for TV licenses. In the electricity capacity market, many analysts suspect Exelon is doing this with their nuclear plants[5]. Their nuclear plants are not clearing the capacity auction, but they are are not retiring before the delivery year of the auction and are not stopping operation. Given their energy market behavior, they should bid in at zero to collect an extra payment. Instead, Exelon is bidding them high to not clear the market, pushing up the marginal generator's bid and earning higher payments for all their other plants.

4.4. The PJM Capacity Market

Analyzing real world outcomes of capacity auctions will help to understand their efficiency and their effectiveness in attracting sufficient investment. However, empirical

[5]`https://www.rtoinsider.com/exelon-pjm-capacity-mkt/`

analysis is difficult because of a lack of data. Capacity auction rules are constantly evolving as regulators adjust rules to better achieve targets. Because rules are constantly changing, it's difficult to compare auctions across years. Moreover, most auction data is kept confidential to prevent future collusion among bidders. Regulators do not reveal which generators "cleared" the auction and they do not release bids. Most auctions only release the clearing price and the aggregate capacity cleared. Further complicating empirical analysis of capacity auctions is that they happen infrequently and power plants are built for a variety of reasons besides capacity prices.

Instead of analyzing specific bidding behavior of capacity market participants, I describe the evolution of the PJM capacity market and analyze its outcomes. The PJM capacity market is one of the oldest in the United States and has had relatively consistent rules for over a decade. I document some of the problems regulators have encountered and the adjustments they have made. Then I show the evolution of capacity prices and analyze subsequent investment.

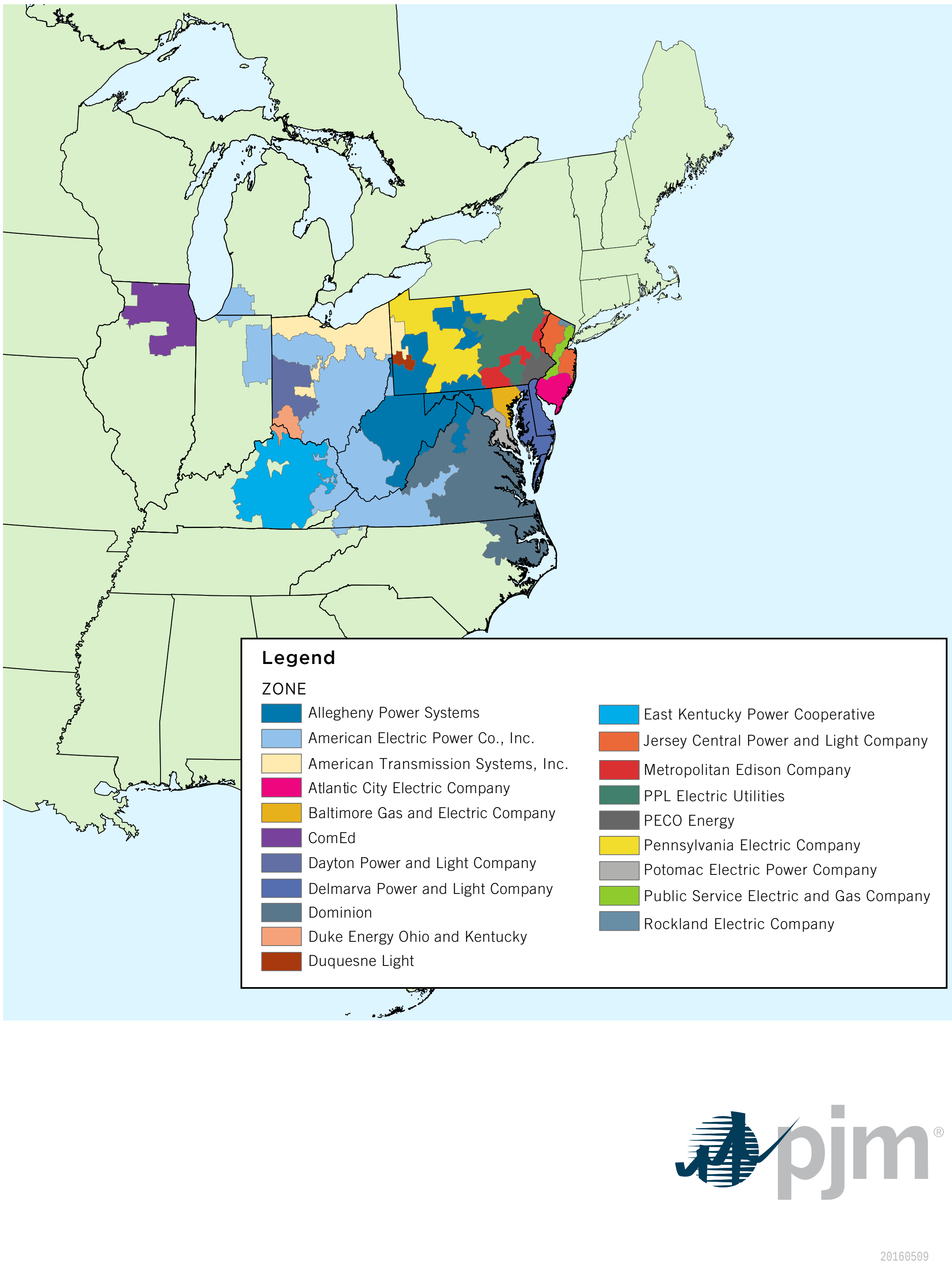
4.4.1. Energy Market History

The PJM Interconnection is one of the largest electricity markets in the world. Today it stretches from New Jersey's east coast to the western suburbs of Chicago, and it serves more than 65 million customers. In 2016, PJM customers spent about $20 billion on energy and $10 billion on capacity payments. PJM's current capacity market framework, starting in 2007, is one of the longest running in the world. Both the magnitude of the market and its longevity make the PJM capacity market worth studying.

The history of PJM is important to understand outcomes in its capacity market. In the United States, prior to the 1990s, electricity was supplied by regional, vertically-integrated regulated monopolies. A utility company was granted a service area and generated electricity, transmitted it, and delivered it to customers. Utilities were responsible for planning new generation resources and building new transmission wires. However, investments were approved of or vetoed by state regulators called Public Utility Commissions (PUCs). PUCs also set prices for utilities that allowed a reasonable return on assets. In the 1990s, there was a movement to liberalize electricity markets. Upstream generation could be a competitive market while downstream delivery was still considered a natural monopoly and regulated.

Like many electricity markets, PJM is a patchwork of historically independent utility companies. Figure 4.1 shows the current geographical footprint of PJM and each of its utility zones. This has been the footprint of PJM since 2012, but there is a history of constantly expanding territory. It was originally formed before the liberalization of electricity markets in 1927 when the Public Service Electric and Gas Company (a New Jersey utility), the Philadelphia Electric Company (PECO Energy), and the Pennsylvania Power & Light Company (PPL) connected their grids to form a larger power pool called the Pennsylvania-New Jersey Interconnection. Previously, these utilities generated their own electricity and distributed it over their own wires in their respective service territories. Interconnection increased gains from trade. In 1956, the Baltimore Gas and Electric Company and General Public Utilities joined the pool, and it was renamed the Pennsylvania-New Jersey-Maryland Interconnection or PJM.

Figure 4.1. Map of PJM and its Zones



After deregulation, PJM became a Regional Transmission Organization (RTO). The transmission wires within the market were still owned by their original utilities, but PJM controlled the flow of electricity across those wires and ran competitive auctions to supply electricity from all merchant generators. PJM is like a market place independent of any of the competing generation firms. PJM rapidly expanded across the Midwest in 2004 and 2005 all the way to Chicago, which is not connected to the rest of PJM directly, but through a market mechanism with bordering utilities called "The Path." Additional utilities in Virginia, West Virginia, and Kentucky were added during this time too. Finally in 2011 and 2012 American Transmission Systems and Duke Energy of Ohio and Kentucky were added making all of Ohio in PJM territory. While the footprint has been stable since 2012, generators in southern Illinois have expressed interest in joining PJM citing the higher capacity prices in PJM compared to their own RTO.

4.4.2. Capacity Market History

Capacity payments have existed in some form in PJM since the time of restructuring in the late 1990s. The determination of how to set payments and rules enforcing capacity obligations have evolved considerably over time. In a PJM document, [**Bowring, 2008**] provides a detailed history of early capacity payments in PJM and motivations for creating the Reliability Pricing Model (RPM), which is the foundation of the current PJM capacity market. This history is largely based on that document.

Explicit capacity payments were not part of the PJM rules when competitive energy markets opened on April 1, 1998. However, PJM required retailers of electricity to be able to procure enough dedicated generator capacity to cover their anticipated peak demand plus a reserve margin at any moment, which is called a capacity obligation. New retailers of electricity had to compete with incumbent utilities for retail sales. At the time, utilities still owned most of the generation capacity in PJM as new, entering merchant generators had not built much capacity. To cover their expected demand, the new retailers needed to contract with their competing utilities. Utilities, not wanting to yield retail sales to competitors, were unwilling to sell capacity at competitive prices. Acknowledging the market failure, on September 17, 1998, the Pennsylvania Public Utility Commission ordered that all electric utilities in its jurisdiction sell capacity at a price of $54.03 per MW-day, which was based on an estimate of the payment needed for generators to recoup their investment.

On January 1, 1999 PJM implemented an official capacity market. There were markets to procure capacity for a day, month, or many months. Retailers of electricity were required to procure enough units of capacity to cover their expected peak demand for each day plus a reserve margin imposed by PJM. They could procure this capacity in markets for daily, monthly, and multi-monthly commitments. Uncommitted generation resources could offer their capacity at any price in any of these markets. PJM aggregated generator bids from least to greatest up until the aggregate requirements by all retailers were fulfilled. The last generator to "clear" the auction set the capacity price for the day and every generator that cleared earned a payment. Generators that cleared the auction were required to submit bids in the energy market for that day and could be compelled to operate if PJM declared an emergency. While these generators were compelled to offer in the energy market, they could offer at any price. Thus generators could earn capacity payments but never actually produce in the energy market by bidding prices well above competitive levels. Generators that did not earn a capacity payment could earn revenue that day by actually producing energy. Absent an administratively declared emergency, excess capacity was the only check on energy prices and capacity prices.

The capacity market ran smoothly through 1999, averaging \$52.86 per MW-day over all capacity markets: daily, monthly and multi-monthly markets. In 2000 prices became more volatile and the daily price peaked on June 1, 2000 at \$350.43 per MW-day. PJM accused a seller of market manipulation throughout the first quarter of 2001 and changed the rules to mitigate market power when one seller becomes pivotal (when demand exceeds the sum total capacity of all other sellers). In these cases, penalties imposed on retailers for not procuring capacity were lowered. This was the outside option for retailers and therefore the price cap on the market.

This form of of PJM's capacity market ran through 2006. During this time, PJM expanded rapidly adding new zones. While the energy market had locational marginal pricing (LMP), the capacity market cleared for the entire footprint of PJM. PJM studied the investment incentives in this environment and found need for reform. They found that energy prices tended to be higher in constrained eastern zones which should signal a need for more investment. However, adding new zones with excess capacity put downward pressure on capacity market prices, attenuating the investment signals in the eastern zones. Their analysis on net revenue, revenue from energy sales and capacity payments less variable costs, showed that generators could not recoup investment costs at the current level. They thought that if there was just a capacity market in the east, demand for capacity would greatly exceed supply pushing up capacity prices and sending correct investment signals and allow sufficient return on investment.

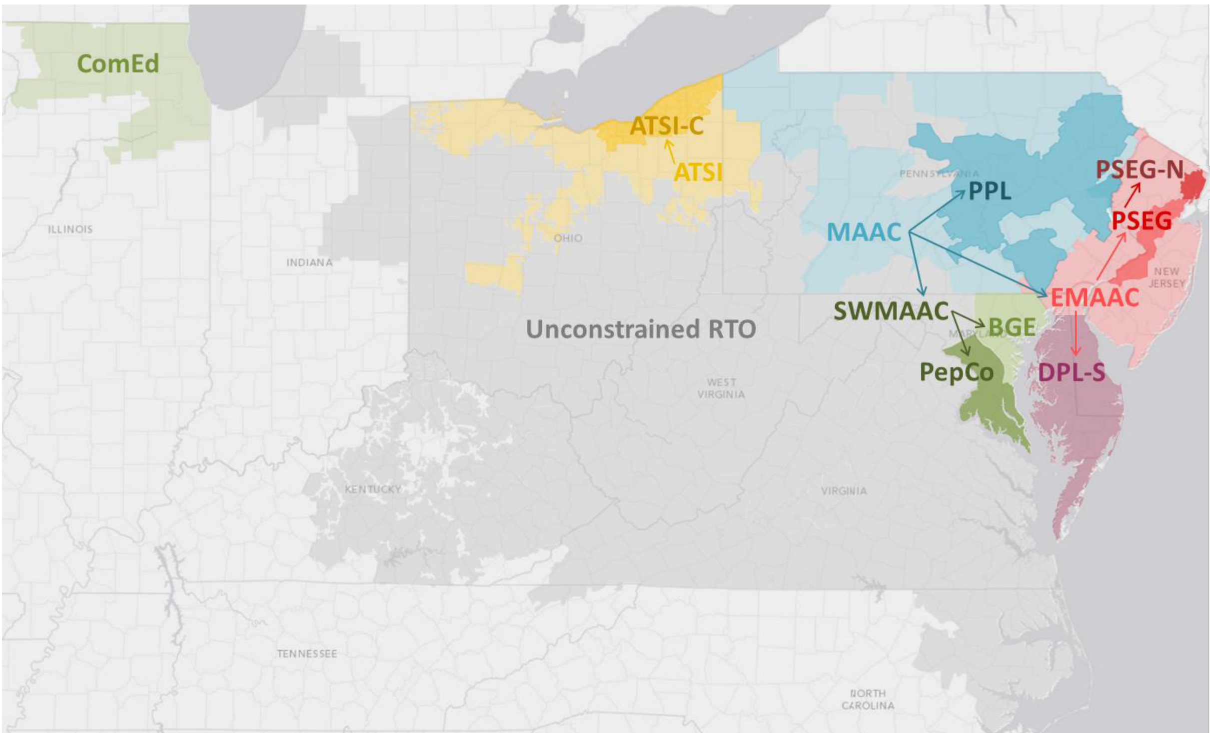
In the early 2000's PJM started planning the Reliability Pricing Model (RPM) for its capacity market, which is still the framework used today. There are several significant differences between PJM's early capacity market and the RPM. Importantly, the capacity payment resulting from the capacity auction in the RPM is set on an annual basis rather than daily, monthly, or multi-monthly. PJM states that forgoing the flexibility of a daily price to represent the "true" value of capacity is worth the guarantee of a stable payment to generators for a whole year. Moreover, the payment is guaranteed three years in advance. The capacity auction, called the Base Residual Auction (BRA) is done three years prior to the delivery year. And each year, up to and including the delivery year, there are incremental auctions where generators can buy or sell their payment. Because the auction happens well in advance, suppliers of capacity can bid into the auction without owning any capacity. Thus a potential investor can guarantee a payment for a generator that has not been built yet.

The RPM can also be location based rather than across all of PJM. Each year, PJM studies the entire footprint of its territory and the supply needs of each of its Locational Delivery Areas (LDAs), shown in Figure 4.2. LDAs are largely based on the utility transmission zones in Figure 4.1, but are more coarse. Transporting electricity within an LDA is relatively seamless (there is little congestion), while transportation across LDAs may be more difficult. For each LDA, PJM defines a reliability-based requirement for capacity based on its transmission capabilities. It defines a zone or set of zones as constrained when the requirement for capacity in the zone can only be met by capacity resources in the defined area, taken out of the merit order. That is, there are much cheaper outside generating units available that cannot reach the zone because of transmission constraints. PJM holds separate capacity auctions for these zones in the delivery year, and all other zones participate in an auction with the rest of the "unconstrained RTO."

Figure 4.2 shows the possible constrained LDAs in PJM. In general, a few areas of PJM are considered constrained for each delivery year, but the number of constrained regions vary. For example, delivery year 2011/ 2012 had just one full PJM price. Also, certain LDAs can clear with one price or even be divided into sub-LDAs. These are indicated by arrows in Figure 4.2. For example, the MAAC LDA also has sub-LDAs PPL, EMAAC, and SWMAAC. In turn, the SWMAAC has sub-LDAs PepCo and BGE. PJM may just have a RTO price and MAAC price where the MAAC price applies to all its sub-LDAs. A typical capacity auction may have an RTO price, MAAC price, and SWMAAC price meaning that every LDA outside of MAAC gets the RTO price, every region in MAAC except for SWMAAC gets the MAAC price and SWMAAC gets its own price.

The final distinguishing feature of the RPM is the downward sloping capacity demand curve. Instead of inelastically demanding a set level of required capacity, PJM uses cost estimates to construct a downward sloping capacity demand curve. Figure 4.3 shows the demand curve used for the first RPM auction in the 2007/2008 delivery year. The first segment of the curve is at a price that is 1.5 times the estimated cost of new entry and the quantity is based on the level of peak demand. The next segment is at the cost of new entry and the quantity is based on some fraction of the desired reserve capacity. The next segment is 0.2 times the cost of new entry and is based on some part of the reserve. This pattern continues. The estimated cost of new entry is thus an important parameter set by PJM. In general, PJM lowers this cost when energy prices are high so that revenues in the capacity market are lower. PJM reasons that high energy prices give an incentive to

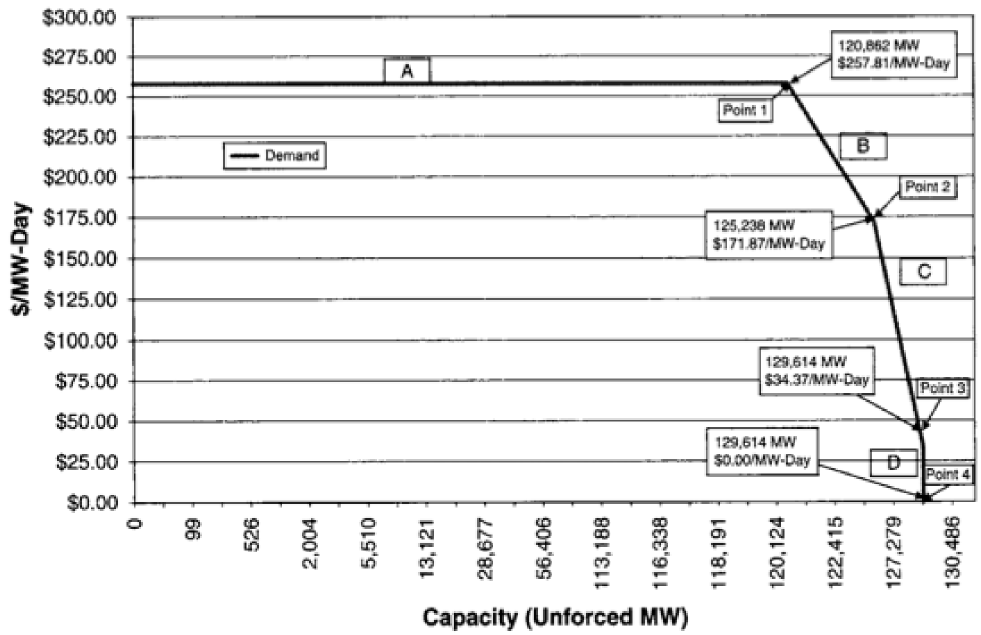
Figure 4.2. Locational Delivery Zones



invest in new capacity (and the net cost of entry lowers), so capacity market prices need not be as high.

The RPM also includes several regulations to address market power concerns. First, in the capacity market there are generation unit-specific offer caps. These only apply when there is enough incumbent generation to serve the entire target level of capacity. When new entrants are required to meet the target capacity, there are no constraints. The reasoning is that competitive forces will drive down prices when there is opportunity to invest. Finally, every generator that offers capacity in the capacity market must offer energy in the energy market. Generators pay fines if their withholding increases the market price by 5%.

Figure 4.3. PJM Aggregate Capacity Demand Curve for Delivery Year 2007/2008



In general, the RPM has run smoothly since its inception in the delivery year 2007/2008. There was a significant change in delivery year 2013/2014 as additional Ohio territories were added to the market. The most significant change came after the Polar Vortex in 2014, which caused many generators to fail.[6] PJM made changes to auction to account for "capacity performance." At first, a certain portion of retailers' required capacity had to meet certain capacity performance standards. By the 2020/2021 delivery year, all capacity in the auction had to meet capacity performance standards. The additional standards preclude certain generators from bidding while the overall levels of desired capacity remained constant.

[6]22% of capacity was unavailable because of weather https://www.greentechmedia.com/articles/read/effects-of-pjms-first-capacity-performance-auction-on-distributed-resources#gs.j6q599w

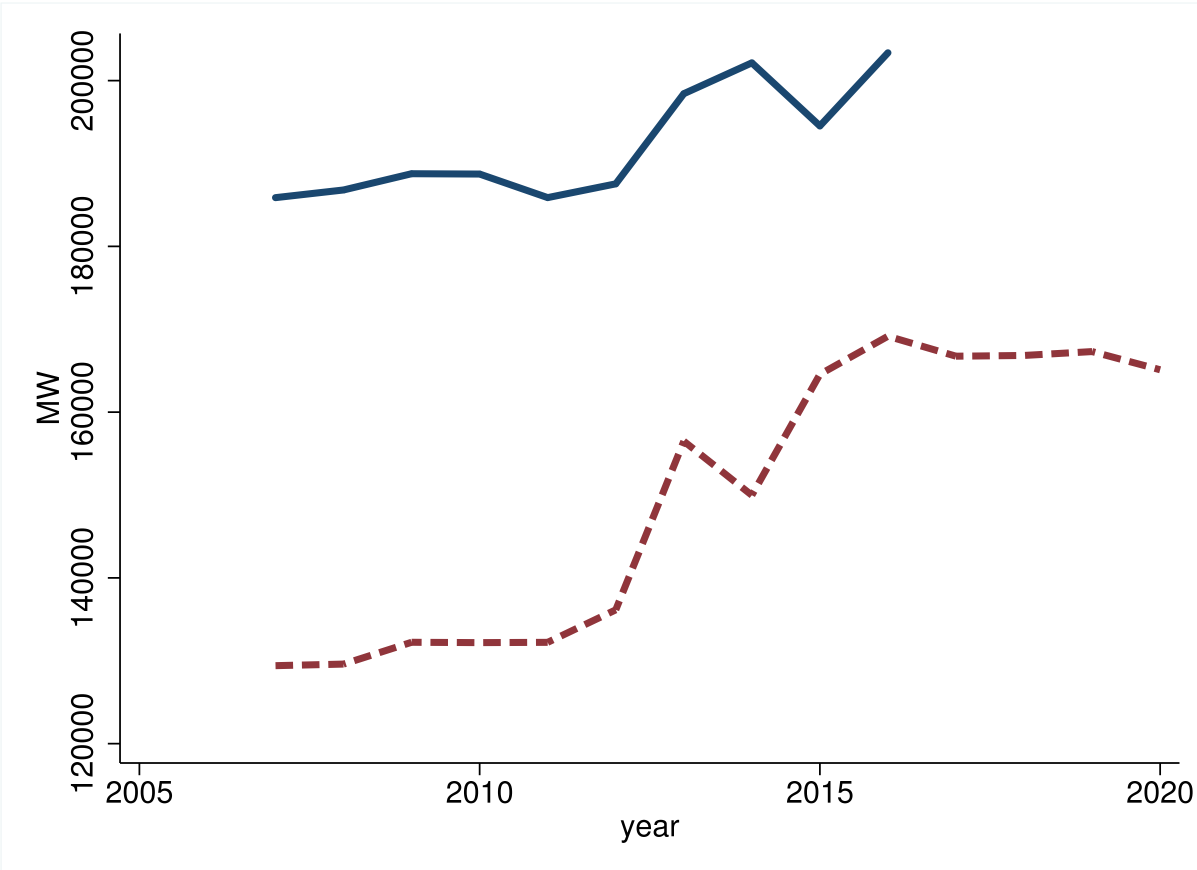
4.5. PJM Capacity Market Outcomes

Analyzing the capacity market is difficult because bid data and capacity payment winners are kept confidential. However, more is known about the outcomes of the capacity market. Targeted levels of capacity and clearing prices by zone are made public by PJM. Also, the Energy Information Administration (EIA), keeps data on every electricity generator in the country. There is data on capacity, ownership, territory, and operations. With this data, there may be ways to identify the effects of capacity prices on investment.

Demand for capacity is largely determined by PJM. As described above, PJM creates a downward sloping demand for capacity in its BRA. The quantity schedules are determined by estimates of peak demand plus a reserve margin. The price schedules are determined by the estimated cost of new entry, net expected energy market revenue for a new entrant. So while the actual amount of capacity cleared in the auction is determined endogenously, PJM designs its demand curves to achieve close to a targeted amount. Figure 4.4 shows the total cleared capacity in the auction, by delivery year, and the actual total capacity in PJM for each delivery year. The targeted level extends into future years because auctions occur three years in advance of the delivery year. The difference between the two curves is realized excess capacity. The difference is large for a few reasons. First there can be operational generators that will not clear the capacity market. Second, the total capacity includes renewable generators which may not be eligible for a capacity payment because they are not reliable sources of generation. The target and actual capacity jump up in 2013 because some utilities in Ohio and Kentucky joined PJM in 2012.

Each year, there can be many different capacity prices depending on how many LDAs PJM defines as constrained. Figure 4.5 shows some of the resulting capacity prices. The

Figure 4.4. PJM Capacity Procurement



Notes: The figure shows the resulting procurement of capacity in MW. The solid blue curve shows the actual level of capacity in PJM. The dashed red curve shows the target level of capacity to procure by delivery year.

lower, solid blue curve is the "unconstrained RTO" price. This is bulk of the market and in 2011 was the whole market. The dashed red line shows the maximum capacity price in the year. Capacity auctions for constrained LDAs always result in higher prices than the RTO-wide price. Therefore the two curves give the range of possible capacity prices for a given delivery year. In general, equilibrium prices are volatile with no clear upward or downward trend. While the offer curve, the supply relation of the market, is

Figure 4.5. PJM Capacity Prices

Notes: The figure shows the resulting capacity prices from the BRA. The solid blue curve shows the unconstrained RTO prices. The dashed red curve shows the maximum price from all the constrained LDAs in the delivery year.

kept confidential, we can infer that it must also be volatile from year to year because the targeted demand (which is not very elastic) is relatively stable.

Identifying the causal effect of capacity prices on investment in capacity is challenging because both are set simultaneously. This is a classic identification problem in econometrics: separately identifying supply and demand. In addition to well specified functions for each (controlling for all shifters of both curves), identification requires an instrument for each curve. For example, an instrument that shifts the supply of capacity without shifting the demand of capacity is needed to identify the supply function of capacity. An example could be a tax break on investment in power plants. This will shift the supply function of capacity. It is a valid instrument if it does not affect the demand for electricity (and therefore the demand for capacity).

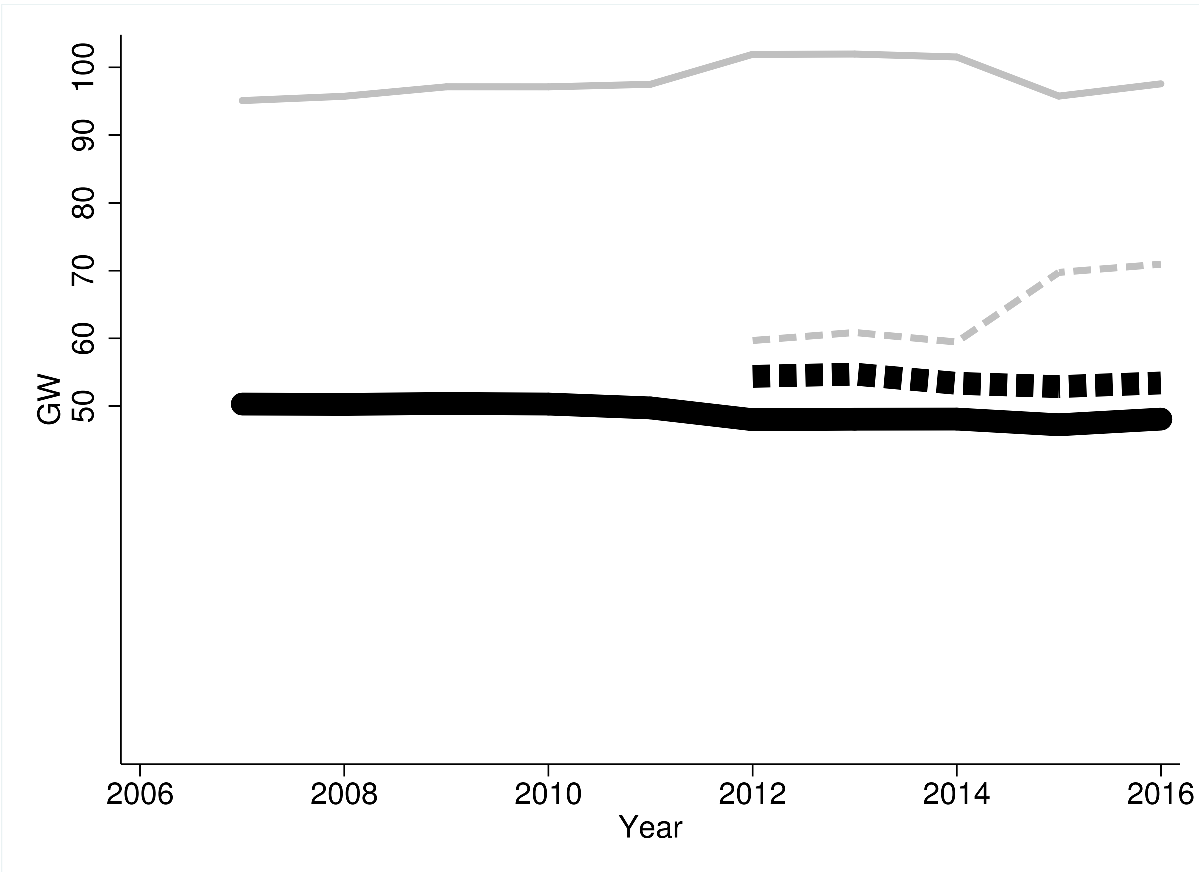
There is simply not enough data to credibly separately identify supply and demand curves for capacity. Instead, I try to find a causal effect of higher capacity prices on investment by exploiting PJM's ability to allow certain zones to have higher prices. I separate PJM into a treatment group, where zones are consistently designated as constrained, and a control group, where zones have never been constrained. I will conclude that expected higher capacity prices stimulate more investment if there is more relative investment in the treatment group than in the control group.

There is no perfect treatment or control group. All zones were considered unconstrained in 2011 and only one zone was considered constrained in 2010. Also, additional zones were added to PJM in 2012. Therefore I define the control group as any zone that has existed since the first delivery year in 2007 and has never been designated as a constrained zone. There are five zones meeting these criteria: American Electric Power Company, Commonwealth Edison. Dominion, Duquesne Light Company, and Dayton Power and Light. I define the treatment group as any zone that has existed since 2007 and was designated as a constrained zone in every year except for 2010 and 2011. There are seven zones meeting this criteria: Atlantic City Electric Company, Baltimore Gas and Electric, Delmarva Power and Light, Jersey City Power and Light, PECO, Potomac Electric Power, and Public Service Electric and Gas.

Ideally, the treatment and control groups would be almost identical except one was randomly assigned higher capacity prices. In reality, each group is different. First, the sizes are different. Figure 4.6 shows their relative sizes by capacity. The control group has much more capacity and its target level is higher than the treatment group. So even if there was more building in the control group compared to the treatment group, we might expect that from the size alone. Second, they have different excess capacity. Figure 4.6 also shows that the treatment group has less capacity than the target set by PJM while the control group has plenty of excess capacity. This is expected as PJM considers the treatment group in need of more investment. The treatment and control group can have even more differences. Energy needs within each group can be different. The treatment group may have more predictable demand and not need as much excess capacity. Also the technology mix could be different. The control group's excess capacity could all be from renewables that run intermittently, meaning its reliable capacity could be just as tight as the treatment group.

Regardless of the many unobserved factors that could be driving investment or disinvestment, the treatment consistently has higher capacity clearing prices compared to the control group. Figure 4.7 shows the net capacity additions, in megawatts (MW), of the treatment and control groups by year. Overall, both groups lost capacity over the period. The treatment group lost 3608 MW while the control group lost 1619 MW. For almost every year the control group had higher net capacity additions than the treatment group. This was in spite of the larger size and excess capacity of the control group. While

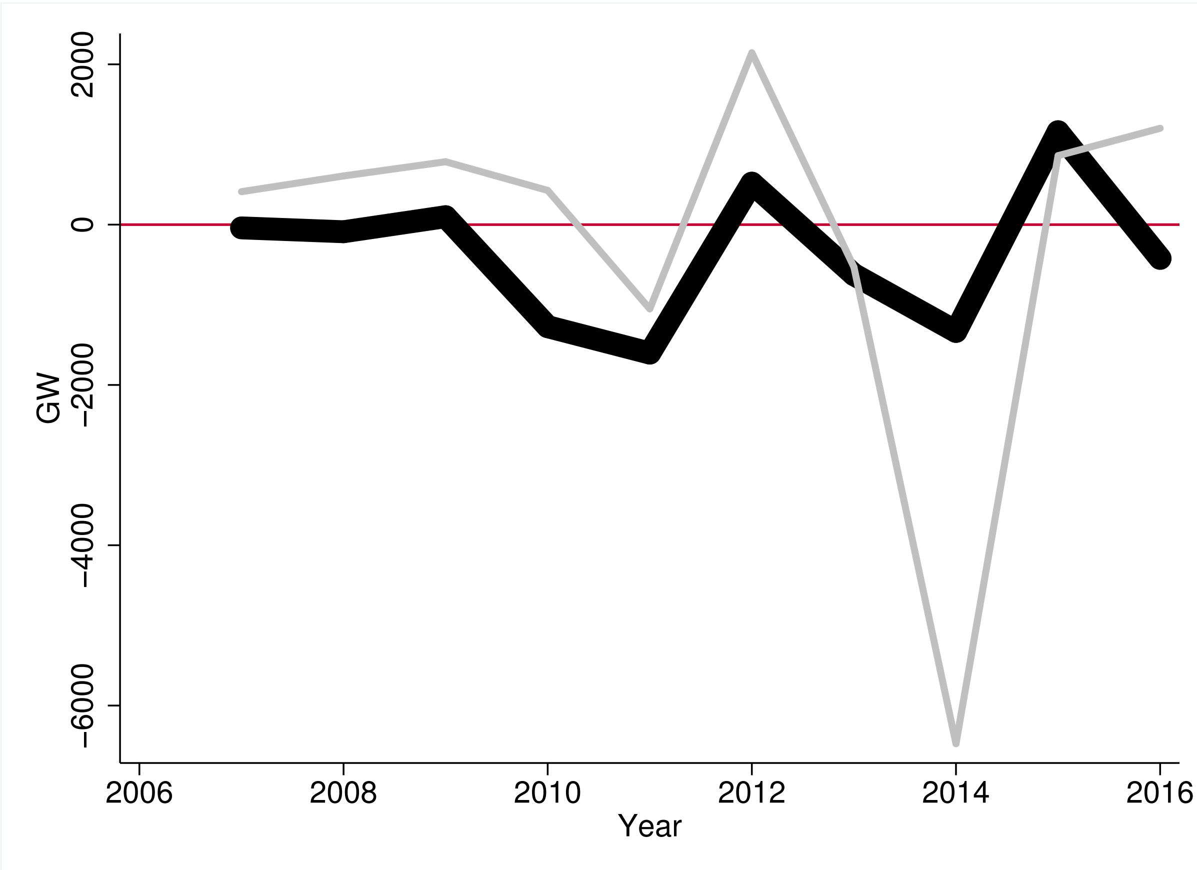
Figure 4.6. PJM Capacity Procurement: Treatment vs. Control Zones



Notes: The figure shows the resulting procurement of capacity in Gigawatts (GW) for two groups of zones. The thinner curves are from the control group while the thicker curves are for the treatment group. In each group, the solid line is the actual capacity in the group, and the dashed line represents the PJM's target capacity.

there can be many other confounding explanations for the different investment and disinvestment levels in the two groups, there does not seem to be any evidence that higher capacity prices are driving higher relative investment.

Figure 4.7. Net Capacity Additions: Treatment vs. Control



Notes: The figure shows the net capacity additions, total capacity built - total capacity retired, for each group of zones. The thicker curve is the treatment group and the thinner curve is control group.

4.6. Conclusion

This paper describes the justifications for capacity payments and mechanics of capacity markets. I analyze the expected payments of a stylized capacity auction modeled after the common set up in capacity auctions around the world. The results of the model show that capacity auctions are an expensive way to procure reserve capacity and it is likely that less expensive mechanisms to procure the same level of capacity exist. I then analyze the investment outcomes in the PJM market and cannot conclude that higher capacity prices are causing higher investment, but more detailed analysis is needed.

Capacity payments and capacity auctions are becoming a major part of electricity markets, and they are increasingly a large part of consumers' bills. More research is needed to analyze and improve the auction environment. Also there needs to more research on the efficacy of capacity payments to drive desired investment.

References

[Abito, 2017] Abito, J. M. (2017). Agency Problems and Environmental Regulation: Evidence from Electric Utilities under Rate of Return Regulation. *Working Paper*.

[Arellano and Serra, 2007] Arellano, M. S. and Serra, P. (2007). A model of market power in electricity industries subject to peak load pricing. *Energy Policy*, 35(10):5130–5135.

[Azar, 2017] Azar, J. (2017). Portfolio Diversification, Market Power, and the Theory of the Firm. *SSRN Electronic Journal*.

[Azar et al., 2016] Azar, J., Raina, S., and Schmalz, M. C. (2016). Ultimate Ownership and Bank Competition. *Working Paper*.

[Azar et al., 2017] Azar, J., Schmalz, M. C., and Tecu, I. (2017). Anti-Competitive Effects of Common Ownership. *SSRN Electronic Journal*.

[Borenstein, 2002] Borenstein, S. (2002). The Trouble With Electricity Markets: Understanding California's Restructuring Disaster. *Journal of Economic Perspectives*, 16(1):191–211.

[Borenstein and Bushnell, 2014] Borenstein, S. and Bushnell, J. (2014). The US Electricity Industry After 20 Years of Restructuring. *Annual Review of Economics*, 7(1):150504160918000.

[Bowring, 2008] Bowring, J. E. (2008). The Evolution of PJM's Capacity Market. In *Competitive Electricity Markets: Design, Implementation, Performance*, chapter 10, pages 363–386.

[Brehm, 2017] Brehm, P. (2017). Natural Gas Prices, Electric Generation Investment, and Greenhouse Gas Emissions. *Working Paper*.

[Bresnahan and Salop, 1986] Bresnahan, T. F. and Salop, S. C. (1986). Quantifying the competitive effects of production joint ventures. *International Journal of Industrial Organization*, 4(2):155–175.

[Bushnell, 2005] Bushnell, J. (2005). Electricity resource adequacy: Matching policies and goals. *Electricity Journal*, 18(8):11–21.

[Bushnell et al., 2017] Bushnell, J., Flagg, M., and Mansur, E. (2017). Capacity Markets at a Crossroads. *Energy Institute at HAAS Working Paper*.

[Bushnell and Ishii, 2007] Bushnell, J. and Ishii, J. (2007). An Equilibrium Model of Investment in Restructured Energy Markets. *Center for the Study of Energy Markets*.

[Bushnell et al., 2008] Bushnell, J. B., Mansur, E. T., and Saravia, C. (2008). Vertical arrangements, market structure, and competition: An analysis of restructured US electricity markets. *American Economic Review*, 98(1):237–266.

[Bushnell and Wolfram, 2005] Bushnell, J. B. and Wolfram, C. (2005). Ownership Change , Incentives and Plant Efficiency : The Divestiture of U . S . Electric Generation Plants. *University of California Energy Institute*.

[Cicala, 2015] Cicala, S. (2015). When does regulation distort costs? lessons from fuel procurement in US electricity generation.

[CollardWexler, 2013] CollardWexler, A. (2013). Demand Fluctuations in the Ready-Mix Concrete Industry. *Econometrica*, 81(3):1003–1037.

[Cowling and Waterson, 1976] Cowling, K. and Waterson, M. (1976). Price-Cost Margins and Market Structure. *Economica*, 43:267–274.

[Cramton and Stoft, 2005] Cramton, P. and Stoft, S. (2005). A capacity market that makes sense. *Electricity Journal*, 18(7):43–54.

[Cramton and Stoft, 2006] Cramton, P. and Stoft, S. (2006). The Convergence of Market Designs for Adequate Generating Capacity. *Report for California Electricity Oversight Board*, (April):1–72.

[Davis, 2017] Davis, L. W. (2017). Evidence of a decline in electricity use by U.S. households. *Economics Bulletin*, 37(2):1098–1105.

[Davis and Wolfram, 2012] Davis, L. W. and Wolfram, C. (2012). Deregulation, consolidation, and efficiency: Evidence from US nuclear power. *American Economic Journal: Applied Economics*, 4(4):194–225.

[Doraszelski Katja Seim Michael Sinkinson Peichun Wang et al., 2016] Doraszelski Katja Seim Michael Sinkinson Peichun Wang, U., Hortaçsu, A., Levy, J., Lewis, G., Kwerel, E., and Vohra, R. (2016). Ownership Concentration and Strategic Supply Reduction *. Technical report.

[Fowlie et al., 2014] Fowlie, M., Reguant, M., and Ryan, S. P. (2014). Market-Based Emissions Regulation and Industry Dynamics. *Journal of Political Economy*, pages 1–53.

[Hortaçsu and Puller, 2008] Hortaçsu, A. and Puller, S. L. (2008). Understanding strategic bidding in multi-unit auctions: A case study of the Texas electricity spot market. *RAND Journal of Economics*, 39(1):86–114.

[Hortascu et al., 2017] Hortascu, A., Luco, F., Puller, S., and Zhu, D. (2017). Does Strategic Ability Affect Efficiency? Evidence from Electricity Markets. *Working Paper*.

[Ishii and Yan, 2007] Ishii, J. and Yan, J. (2007). Does divestiture crowd out new investment? the "make or buy" decision in the U.S. electricity generation industry. *RAND Journal of Economics*, 38(1):185–213.

[Ito, 2014] Ito, K. (2014). Do consumers respond to marginal or average price? Evidence from nonlinear electricity pricing. *American Economic Review*, 104(2):537–563.

[Ito and Reguant, 2016] Ito, K. and Reguant, M. (2016). Sequential markets, market power, and arbitrage. *American Economic Review*, 106(7):1921–1957.

[Joskow, 2007] Joskow, P. (2007). Capacity Payments in Imperfect Electricity Markets: Need and Design. *MIT Memo*.

[Kwon, 2016] Kwon, H. J. (2016). Executive Compensation under Common Ownership. *Working Paper*.

[Lim and Yurukoglu, 2016] Lim, C. and Yurukoglu, A. (2016). Dynamic Natural Monopoly Regulation: Time Inconsistency, Moral Hazard, and Political Environments. *Working Paper*.

[O'Brien and Salop, 2000] O'Brien, D. P. and Salop, S. C. (2000). COMPETITIVE EFFECTS OF PARTIAL OWNERSHIP: FINANCIAL INTEREST AND CORPORATE. *Source: Antitrust Law Journal*, 67(3):559–614.

[O'Brien and Waehrer, 2017] O'Brien, D. P. and Waehrer, K. (2017). The Competitive Effects of Common Ownership: We Know Less than We Think. *SSRN Electronic Journal*.

[Posner et al., 2016] Posner, E. A., Scott Morton, F. M., and Weyl, E. G. (2016). A Proposal to Limit the Anti-Competitive Power of Institutional Investors. *SSRN Electronic Journal*.

[Ryan, 2012] Ryan, S. P. (2012). The Costs of Environmental Regulation in a Concentrated Industry. *Econometrica*, 80(3):1019–1061.

[Rydqvist et al., 2014] Rydqvist, K., Spizman, J., and Strebulaev, I. (2014). Government policy and ownership of equity securities. *Journal of Financial Economics*, 111(1):70–85.

[von der Fehr and Harbord, 1997] von der Fehr, N.-H. M. and Harbord, D. (1997). Capacity Investment and Competition in Decentralised Electricity Markets. *Memorandum No. 27, Department of Economics, University of Oslo*.

APPENDIX A

Appendix

A.1. Capacity Constrained Cournot

I used a mixed-inter solver to find the solution to the capacity-constrained Cournot equilibrium in the hourly spot market of each ISO. The framework was first calculated as a complementarity problem in [**Bushnell et al., 2008**]. I follow the framework used in [**Ito and Reguant, 2016**] which expressed the problem in an equivalent mixed-integer representation.

Given estimates of $\hat{b}$, the slope of the competitive fringe's supply curve and $\hat{a}$ the corresponding intercept, firms with market power face a residual demand curve of $Q(p_t) = \hat{a} - \hat{b}p_t$. Firms with market power have piece-wise continuous marginal cost curves made of five segments, $h = 1, 2, 3, 4, 5$ (In Chapter 2, the game is solved when marginal cost curves have four segments). Each segment is linear: $C_{jh}(q_{jh}) = c_{jh} + m_{jh}q_{jh}$ and each segment has maximum capacity $\bar{q}_{jh}$. Marginal costs are constructed so that the firm level marginal cost curve is continuous: $c_{jh} + m_{jh}\bar{q}_{jh} = c_{jh+1}$.

The formulation here is generalized to account for common ownership. β_{ij} is the ownership share that investor i has in firm j. γ_{ij} is the weight that the manager of firm j puts on the profits of investor i. In Chapter 2 where there is no accounting for common ownership and in Chapter 3 where firms sometimes do not account for common ownership, γ_{ij} and β_{ij} are equal to one for only one investor with no interests in competing firms. This reduces back to standard Cournot first-order necessary conditions for profit maximization.

Define $\underline{u}$ and $\bar{u}$ as vectors of dummies of length N x 5 that specify whether a given segment in the marginal cost curve is used at all ($q_{jh} > 0$) and whether it is used at full capacity ($q_{jh} = \bar{q}_{jh}$), respectively. Define $\psi_{jh} \geq 0$ as the shadow value when $\bar{u}_{jh}$ is binding.The equilibrium solves for vectors $\underline{u}$, $\bar{u}$, ψ, and q. The conditions to solve for equilibrium are

$$
\begin{array}{rlr}
\text{[FOC 1]} & \displaystyle\sum_{i \in I} \gamma_{ij}\Big(\beta_{ij}(P - c_{jh} - m_{jh}q_{jh}) - \sum_{k \in J} \beta_{ik} q_k / b\Big) - \psi_{jh} \leq 0 & \forall j, h = 1, 2, 3, 4, 5 \\
\text{[FOC 2]} & \displaystyle\sum_{i \in I} \gamma_{ij}\Big(\beta_{ij}(P - c_{jh} - m_{jh}q_{jh}) - \sum_{k \in J} \beta_{ik} q_k / b\Big) - \psi_{jh} \geq M\underline{u}_{jh} - M & \forall j, h = 1, 2, 3, 4, 5 \\
\text{[Complementarity]} & \psi_{jh} - M\bar{u}_{jh} \leq 0 & \forall j, h = 1, 2, 3, 4, 5 \\
\text{[Definition } \underline{u}\text{]} & q_{jh} - \bar{q}_j \underline{u}_{jh} \leq 0 & \forall j, h = 1, 2, 3, 4, 5 \\
\text{[Definition } \bar{u}\text{]} & \bar{q}_j \bar{u}_{jh} - q_{jh} \leq 0 & \forall j, h = 1, 2, 3, 4, 5 \\
\text{[Sorting 1]} & \bar{u}_{jh} - \underline{u}_{jh} \leq 0 & \forall j, h = 1, 2, 3, 4, 5 \\
\text{[Sorting 2]} & \bar{u}_{jh} - \underline{u}_{jh-1} \leq 0 & \forall j, h = 2, 3, 4, 5 \\
\text{[Sorting 3]} & \bar{u}_{jh} - \bar{u}_{jh-1} \leq 0 & \forall j, h = 2, 3, 4, 5
\end{array}
$$

where $P = \frac{\hat{a}_t - \sum_{j,h} q_{jh}}{\hat{b}}$ and $M = 10^5$, a sufficiently large value.

A.2. Proof of Auction Result

Proof of Model Result.

Firms A and B own legacy plants which have costs that are private information but are known to be distributed $C_A, C_B \sim U[0, \frac{1}{2}]$. Revenues in the energy market for each firm for each firm are at least $\frac{1}{2}$ even if the other opponent builds a new plant (which could cut into firm revenue), meaning firms A and B will keep their legacy plants regardless the outcome of the capacity auction.

Each firm has an idiosyncratic cost of building a new plant $F_A, F_B \sim U[\frac{1}{2}, 1]$ which are statistically independent and private information. Building a new plant is unprofitable without a payment to recoup the cost draw.

The government wants to incentivize building a new plant with a capacity payment determined by the auction in the text. Also, the government caps legacy bids at 1/2 knowing the costs of legacy plants are less than building a new plant.

Claim: The optimal strategy is symmetric and each firm bids $b_i^1 = 0$ and $b_i^2(F_i) = \frac{2}{3} + \frac{1}{3}F_i$

Proof:

Search for solutions where the opponent bids an affine function of its fixed cost. Also assume that $b_i^1 < b_i^2$, so that each firm's first bid cannot set the price, therefore bidding 0 does not change expected profits. Consider the problem from the perspective of Firm A. Given that $b_i^2(F_i) = \lambda + \mu F_A$, Firm A maximizes its expected payoff.

$$\max_{b_A^2}(2b_A^2 - F_A)P(b_A^2 < b_B^2) + E(b_B^2 | b_A^2 > b_B^2)P(b_A^2 > b_B^2)$$
$$= \max_{b_A^2}(2b_A^2 - F_A)P\Big(\frac{b_A^2 - \lambda}{\mu} < F_B\Big) + \Big(\lambda + \mu E\Big(F_B | \frac{b_A^2 - \lambda}{\mu} > F_B\Big)\Big)P\Big(\frac{b_A^2 - \lambda}{\mu} > F_B\Big)$$
$$= \max_{b_A^2}(2b_A^2 - F_A)\Big(1 - \frac{b_A^2 - \lambda}{\mu}\Big)2 + \Big(\lambda + \mu\Big(\frac{1}{2} + \frac{1}{2}\Big(\frac{b_A^2 - \lambda}{\mu} - \frac{1}{2}\Big)\Big)\Big)\Big(\frac{b_A^2 - \lambda}{\mu} - \frac{1}{2}\Big)2$$

The first-order necessary condition reduces to

$$b_A^2 = \frac{2}{3}(\mu + \lambda) + \frac{1}{3}F_A$$

Assuming that bid functions will be symmetric, I set constants equal to one another

$$\mu = \frac{1}{3}$$

$$\lambda = \frac{2}{3}(\mu + \lambda) = \frac{2}{3}(\frac{1}{3} + \lambda)$$
$$\Rightarrow \lambda = \frac{2}{3}$$

Resolving the problem for the optimal b_A^2 when $b_A^1 = b_B^1 = 0$ and $b_B^2(F_B) = \frac{2}{3} + \frac{1}{3}F_B$ results in the affine bid function $b_A^2(F_A) = \frac{2}{3} + \frac{1}{3}F_A$.

The last thing to check is that $b_i^1 = 0$ is optimal. Given that $b_B^1 = 0$ and $b_B^2(F_B) = \frac{2}{3} + \frac{1}{3}F_B$, then any bid below $\frac{1}{2}$ (can even be up to $\frac{2}{3}$) will not set the price and will still clear the auction, so there is no profitable deviation of $b_A^1 > 0$.

The above setting depends on the bid cap of legacy plants. While unit specific bid caps are a part of the PJM capacity market, a more general solution where bids are unconstrained may be interesting to regulators. The above strategies are not optimal with unconstrained bids. From Firm A's perspective, if Firm B is bidding the above strategy, then bidding b_A^1 some very large number will dominate. It will clear the auction because it is the third bid and its payment will be the bid. Any bid higher than 2 will earn more than Firm B's strategy which has maximum bid of 1 and could at most earn two payments of 1.

Another interesting scenario would be if all bids were capped at 1. Note that the conditional expected payoff of the bid (given Firm B does the above strategy) is $F_A^2 - 2F_A + \frac{23}{12}$. This payoff is greater than or equal to 1 for $F_A < 1 - \frac{1}{2\sqrt{3}} \approx .711$. So a better response to Firm B is symmetric for low costs of Firm A. But bid $b_A = 1$ for higher realizations of cost. Given Firm B's strategy, this guarantees a payoff of 1 because A's first plant will clear and win the bid of 1. Firm A's second potential plant will not clear and not incur the cost. There would be a completely different equilibrium strategy in this case.