EXCESSIVE ENTRY, INVESTMENT AND CONGESTION COST

IN THE ELECTRICITY INDUSTRY

by

ALEXANDER MCGAVOCK HILL

B.A., Claremont McKenna College, 2008

M.A. Boston University, 2011

M.A., University of Colorado, Boulder, 2014

A thesis submitted to the

Faculty of the Graduate School of the

University of Colorado in partial fulfillment

of the requirement for the degree of

Doctor of Philosophy

Department of Economics

2017

This thesis entitled:

Excessive Entry, Investment and Congestion Cost in the Electricity Industry

By Alex Hill

has been approved for the Department of Economics

Professor Enter Name, Chair

Professor Enter Name

4/12/2017
Date: ___________________________

The final copy of this thesis has been examined by the signatories, and we find that both the content and the form meet acceptable presentation standards of scholarly work in the above mentioned discipline.

Hill, Alexander McGavock (Ph.D., Economics)

Excessive Entry, Investment and Congestion Cost in the Electricity Industry

Thesis directed by Associate Professor Daniel Thomas Kaffine

Electricity is a large and crucial US industry, with annual sales and investment of $390 billion and $80 billion respectively. Two important topics in the industry, over the past 25 years, are the restructuring of state markets and variation electricity demand. Both changes occurred during the largest increase in electricity capacity in US history, the "Gas Boom," which increased generating capacity by over 33 percent. This thesis links these topics with the increase in gas capacity and estimates their effect on electricity industry investment and prices.

Chapter 1 estimates the effect of restructuring on US power plant investment, finding states that took steps toward retail competition built, on average, 593 MW more gas-fired CCGT power plants than if regulation had remained in place. Evidence of oversupply, bankruptcies, and falling capacity factors suggests this entry was excessive. Chapter 2 investigates the cause and cost of this excessive entry, finding it consistent with contagion, manager overconfidence, and falling long-term interest rates. The welfare cost is estimated to be $10.8 billion. Chapter 3 provides a first estimate of the impact of demand variance on electricity prices, finding an increase of 10 percent in demand variance leads to an increase of 9.7 percent in average state electricity prices. This impact was found to be significant across all three sectors of electricity use. These findings inform policymaker decisions in the future, suggesting states concern themselves with the lower and upper bound of capacity investment, while encouraging more measures to reduce daily and seasonal demand variation.

I dedicate this thesis to the two people that have contributed most to my starting and finishing a PhD in economics. Marcia McGavock Hill, my mother, is irrationally optimistic, a trait that has never allowed doubt to enter my mind. John Kenton Hill, my father, gave me the freedom and support to choose whichever career path I desired.

I would like to acknowledge the contributions made by the following people to this thesis: Dan Kaffine, who always gave positive feedback and never made me feel like I was wasting his time; Tyler Mangin, who I bounced good and bad ideas off of for years in the collection of these papers; Brian Cadena, whose insistence that I spend the summer after my second year focusing on research enabled me to finish in five years; and Kent Hill, my father, who convinced me that I was making a contribution in my research which gave me the fortitude to keep going.

CONTENTS

CHAPTER

TABLES

FIGURES

CHAPTER I

EXCESSIVE ENTRY AND INVESTMENT IN THE ELECTRICITY SECTOR

The restructuring of U.S. state electricity markets, beginning in the mid-1990s, provides an environment to empirically estimate the impact of free entry on investment in a large industrial sector. Leveraging the staggered restructuring by US states,, this chapter shows restructuring, after controlling for demand and supply factors, led to an average annual increase of 600 megawatts of state gas-fired power plant capacity over the period 2000-2006. This investment contributes to explaining the size and duration of the US Gas Boom, where over $240 billion was invested in gas-fired power plants. The restructuring increase in capacity can be considered excessive, or uneconomical, given it was not due to fundamentals, such as retail sales growth, and was later associated with owner bankruptcies and financial distress. (JEL L51, L94, L22, Q40)

1. Introduction

Free entry is a necessary, but not sufficient, condition for decentralized decision-making maximizing social welfare. When, for example, competitive pricing behavior is absent, free entry results in inadequate or excessive entry, relative to the social optimum.[1] Simultaneous firm entry can also lead to non-optimal market outcomes through coordination failure.[2] Empirical evidence of these effects is less clear, with the majority of papers finding the well-established result of S-shaped entry and exit in new product markets.[3] This chapter focuses on US electricity market restructuring to provide evidence that free entry led to overinvestment in electricity generating capacity.

---

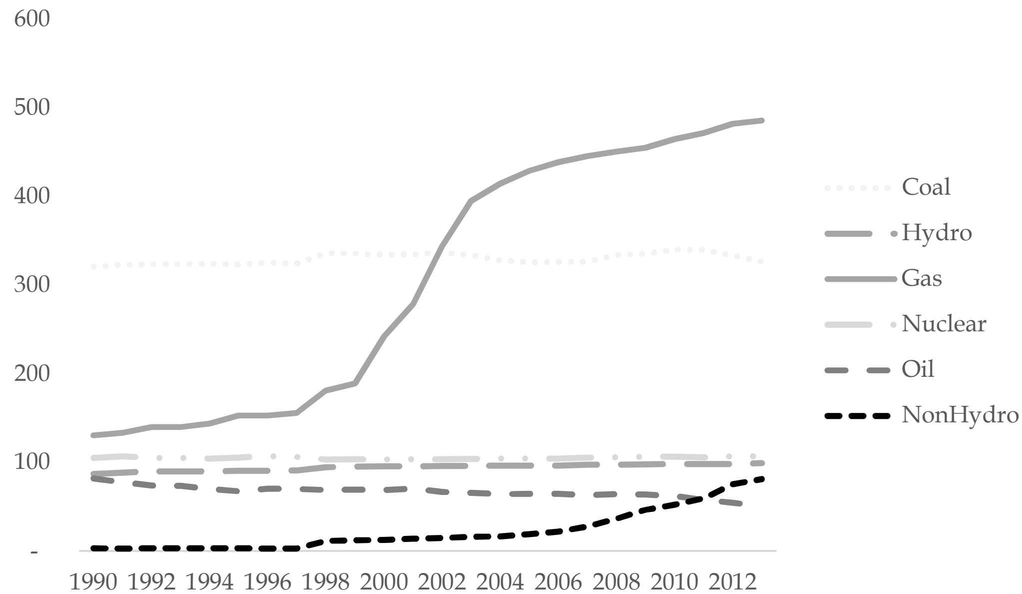
[1] See Spence (1976) and Dixit and Stiglitz (1977) for heterogeneous products and Weizsacker (1980), Perry (1984) and Mankiw and Whinston (1986) for homogeneous products.
[2] See Dixit and Shapiro (1986), Bolton and Farrell (1990) and Cabral (2004).
[3] See Jovanovic (1982), Jovanovic and MacDonald (1994), Klepper (1996), and Klepper and Simons (2000). Berry and Waldfogel (1999) provides empirical evidence using the radio broadcasting industry.

The electricity industry is a unique and important study of free entry for several reasons. First, it is a major US industrial sector accounting for over $390 billion in sales, $70 billion in investment and 39 percent of energy use (EIA 2016). Second, the electricity sector possesses two characteristics, mentioned in Berry and Waldfogel (1999), as required for free entry to lead to social inefficiency: the entrant's product can substitute for the incumbent's and average costs are decreasing in output. Average costs decline in both generation and transmission and the product is homogenous. Finally, free entry was the result of restructuring, as opposed to product innovation in other industries.

In addition to providing evidence of the impact of free entry on the electricity sector, this chapter contributes to the growing literature on the cost impact of electricity market restructuring. The improvement in the marginal cost of power plants following electricity market restructuring is a well-established result (Fabrizio et al 2007; Davis and Wolfram 2012; Cicalla 2015). However, the impact of competition on plant manager decision-making was expected to be a small part of the benefit of restructuring (Joskow 1997). The larger impact was predicted to be from long-term decision-making by firms on plant construction. This chapter is the first analysis of this longer-term impact.

State-level electricity market restructuring began in California and Texas in 1995. From 1995 to 2002, 22 states and DC began restructuring with the goal of increased generation and retail competition. During this time period, the US electricity industry experienced an investment boom. Due to changes in the structure of state electricity markets and relative fuel prices, the investment was primarily in gas-fired power plants. From 1999 to 2006, there was a net addition of over 250 gigawatts (GW) of gas-fired capacity to the U.S. electricity market (See Figure 1). The share of total US capacity using natural gas as a primary fuel increased from 23.7 to 41.5 percent and represented an investment of over $240 billion. Surprisingly, while there are many hypothesized reasons for teh dash for gas in the UK and the gas boom in the US, [4] careful empirical analysis has been lacking. In particular, very little has been written on the connection between the gas boom in the US and the restructuring of state electricity markets.

Figure 1: U.S Generating Capacity (MW)



Source: EIA (2016)

Prior to restructuring, industry regulation critics argued that cost-of-service regulation distorted firm incentives to profit maximize, both in the short-run and long-run. In the short-run, firms may not operate plants using the cost-minimizing amount of labor and maintenance. In the long-run, firms lack pressure to invest in the lowest cost form of generating capacity and may suffer from the Averch-Johnson effect, biasing investment decisions towards capital (Joskow 1997). Based on these incentive problems, states were encouraged to restructure their electricity markets by allowing retail competition. The decision to restructure electricity markets, by almost half of US states, created an opportunity to evaluate the impact of restructuring in both the short-run and

[4] See Winskel (2002) for a description of the dash for gas in the UK. See Kaplan (2010) and Macmillan (2013) for a review of the industry explanations for the gas boom in the US.

long-run. Topics previously investigated include the impact on efficiency of power plants operations (Bushnell and Wolfram 2005; Fabrizio, Rose, and Wolfram 2007; Davis and Wolfram 2012; Cicalla 2015), efficiency of wholesale electricity markets (Borenstein, Bushnell, and Wolak 2002; Bushnell, Mansur, and Saravia 2008), market power (Wolfram 1999; Joskow and Kahn 2002) and capital investments (Ishii and Yan 2007; Fowlie 2010).

This chapter makes three contributions to the economics literature. First, it provides an empirical example of an industry experiencing a free-entry "failure." Second, it is the first paper to empirically investigate the long-run gains of restructuring proposed in Joskow (1997), contradicting predictions of either investment efficiency (Joskow 1997) or underinvestment (Borenstein and Holland 2005). Third, it provides a robust analysis of the US gas boom from 2000-2006.

The staggered adoption of restructuring by 22 states and DC provides a setting which allows for identification of the causal impact of restructuring on gas-fired power plant investment in the US.[5] Comparisons are made throughout this chapter between states that restructured and those that did not. Observations are made at the state and year level, with the empirical model treating the state as the decision-level. While individual decisions are made by firms and utilities rather than state governments, the influence of state regulation, the level of the policy variable and the availability of information on a state-level makes this approach more accurate and feasible.

The hypothesis of this chapter is restructuring caused a surge in gas-fired power plant investment by encouraging entry by firms without complete information. However, once this surge occurred and margins began to shrink, investment decreased substantially. To identify this investment flow, the preferred specification includes an interaction of the restructuring variable with a binary variable indicating years since

[5] This chapter focuses on investment in gas-fired power plants because over 84 percent of investment in capacity during this period was in plants fueled by natural gas, as opposed to coal, nuclear, or renewables (Energy Information Administration (EIA), 2016).

restructuring. This specification fits the hypothesis of this chapter that the surge was temporary, instead of leading to permanent excessive entry of the type suggested in Mankiw and Whinston (1986).

In each specification, the variable of interest is the annual change in state combined cycle natural gas (CCGT) power plant capacity. This chapter shows, compared to the non-boom period, a restructured states built, on average, 593 MW more combined cycle gas capacity than if it was still regulated. This investment is approximately $600 million annually, per state, and over $95 billion total over the period. The result holds in a number of specifications that include differences in demand, prices, capacity needs and levelized costs, leading to the conclusion that there was excessive entry in restructured markets.[6] The implication is the nature of restructured markets lead to overinvestment in power plant capacity. The existence of large reserves, compared to non-restructured states, and a string of bankruptcies of firms that owned power plants in restructured markets, are further evidence of the excessive nature of the investment boom.

The structure of this chapter is organized as follows. Section 2 provides information on the electricity market setting of this chapter. Section 3 presents the methodology and model used to analyze the research question in this chapter and introduces the data. Section 4 presents the empirical results of the chapter. Section 5 shows the results of alternative specifications. Section 6 concludes.

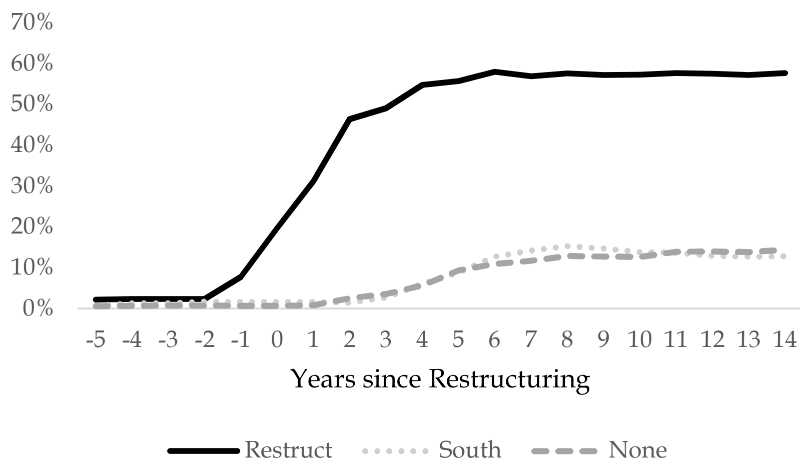
2. Background

Prior to the mid-1990s, US electricity generation was primarily supplied by regulated vertically-integrated public utilities. Prices were set by cost-of-service regulation, with utility rates determined by expenditure on fuel and operations and maintenance (O&M), prudent capital investment, and a reasonable rate of return on

[6] Excessive entry and overinvestment are defined in this chapter as the difference between the CCGT capacity increase of a state and what its synthetic control counterfactual would have built.

capital. The increase in prices challenged the system in the 1980s, as prices rose for consumers due to an excess of capacity over demand. Rising prices and two energy laws in 1978 (Public Utility Regulatory Policies Act) and 1992 (Energy Policy Act) led to the expansion of merchant gas-fired power plants built by independent power producers (IPP), which received further support from Federal Energy Regulatory Commission (FERC) orders 888 and 889 in 1996. While the impact of these orders was small (Figure 2), the combination of changing prices, policies and excess capacity created an environment where many states were prepared to restructure their electricity markets.

Figure 2. IPP Penetration by State Group[7]



Source: EIA Form 860 (2016)

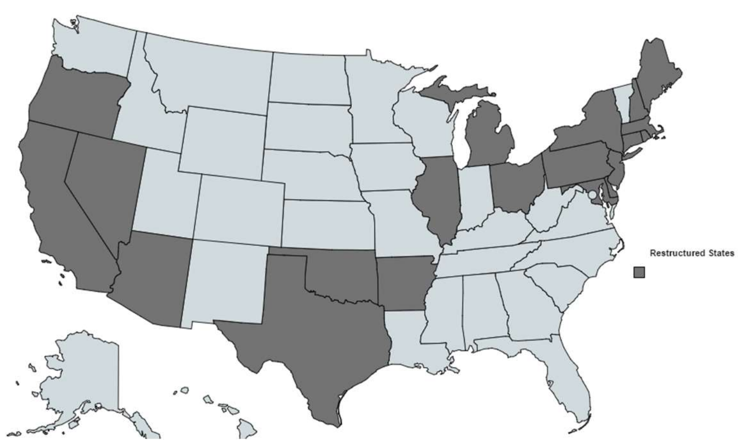
State-level electricity market restructuring split the industry into three components: generation, transmission and distribution, and retail. Transmission and distribution continued to operate as a regulated monopoly due to the efficiency of high voltage power lines and the undesirability of local utility power lines crisscrossing neighborhoods. Electricity generation and retail, however, were capable of operating in

[7] The running variable in this figure measures IPP penetration since the year a state started restructuring. For the South and None regions, 1996 is year 0, since this is when FERC began enforcing open access to transmission lines.

a competitive market (Joskow and Schmalensee 1983), particularly with new gas-fired plants not requiring the same economies of scale and build times as previous coal, gas and nuclear plants. Opening electricity markets to generation and retail competition was the final step in the attempt to provide power producers with an incentive to cost minimize, with previous attempts including state adoption of incentive regulation, where capped rates allowed utilities to earn rents by cost minimizing (Joskow 1997).

Restructuring varied by state, with all 50 states beginning the process and 22 states plus DC opening to retail competition (see Figure 3).[8] Early stages included hearings, retail pilot programs, and submissions by state regulatory commissions. States that continued restructuring passed legislation empowering regulators to enforce a timeline for the divestiture of assets and opening of retail markets to competition. [9] Restructuring concluded with asset divestiture by utilities and power marketers operating. A fully restructured market was intended to spur generation competition between utilities and IPPs and retail competition between utilities and power marketers.

Figure 3. Map of Restructured U.S. States



Source: EIA (2016)

[8] See Joskow (2008) for a list of rules for successful restructuring of electricity markets.
[9] Divestiture of assets was one of the crucial requirements noted in Joskow (2008) because of the ability of utilities, which owned the majority of generation facilities, to use their share of the market to raise prices.

The California Electricity Crisis of 2001 created concerns over the effective operation of restructured markets in several states, which led to a pause in restructuring in these states (EIA 2016). As noted in a number of papers on the crisis,[10] a combination of high demand due to unusual summer weather, reduced hydro generation due to low rainfall the previous winter, and strategic activity by firms created a series of blackouts and sent wholesale electricity prices soaring. This led several large California utilities to or near bankruptcy, as the utilities were not allowed to raise prices by the state regulatory commission but had to purchase electricity from the wholesale market, where prices were high. The crisis was seen as vindicating the concerns of restructuring opponents, who were focused on inherent volatility in restructured electricity markets and many states which were considering restructuring decided to continue regulation (EIA, 2016).

There were unique reasons behind each state's decision to restructure, but several trends are worth noting. First, high prices created consumer discontent, with a correlation between electricity prices and state decisions to restructure (Joskow 2008). Second, there were well-founded economic arguments for opening the generation and retail components of the electricity industry to competition. The most prominent are: 1) Generation competition incentivizes managers to choose cost-minimizing quantities of labor, fuel, materials, maintenance and capacity. 2) The Averch-Johnson effect distorts capacity choices towards more capital-intensive technologies. 3) Competitive retail and generation markets insulate electricity generation from politics. 4) Competition encourages the retirement of uneconomical plants. Finally, the late 1970s-1990s was a period of US and international market deregulation. The UK deregulated its electricity market in the early 1990s, while US deregulation of telecommunications and finance were taking place.[11] This environment encouraged restructuring of electricity markets.[12]

[10] See CBO (2001) and Joskow (2008) for more detailed descriptions.

[11] The Telecommunications Act of 1996 removed barriers to entry in telecommunications markets, while the Gramm-Leach-Bliley Act of 1999 repealed several regulations dating back to the 1930s.

[12] White (1996) and Joskow (1997) descirbe deregulation process and political influences.

Restructuring was one of several changes in U.S. electricity markets during this time period. A combination of technological improvement, supply choices on the state level, and demand-side changes occurred during this time period which had a substantial effect on gas-fired capacity. Table 1 shows the average annual net capacity addition, by state group, of both CCGT and CT plants during three periods: pre-boom (1990-1998), boom (1999-2006), and post-boom (2007-2013). While there is an identifiable CT boom, the CCGT capacity increase in restructured states from 1999-2006 is the clear outlier.

Table 1. CCGT and Gas CT Capacity Additions by State Group

	Restruct	None	South
CCGT			
1990-98	8,418	1,219	5,008
1999-06	129,381	16,332	45,128
2007-13	17,074	4,851	21,059
CT			
1990-98	19,349	8,367	12,861
1999-06	26,597	19,918	27,955
2007-13	10,358	3,876	4,201

Source: EIA (2016)

2.1 Explaining the Gas Boom

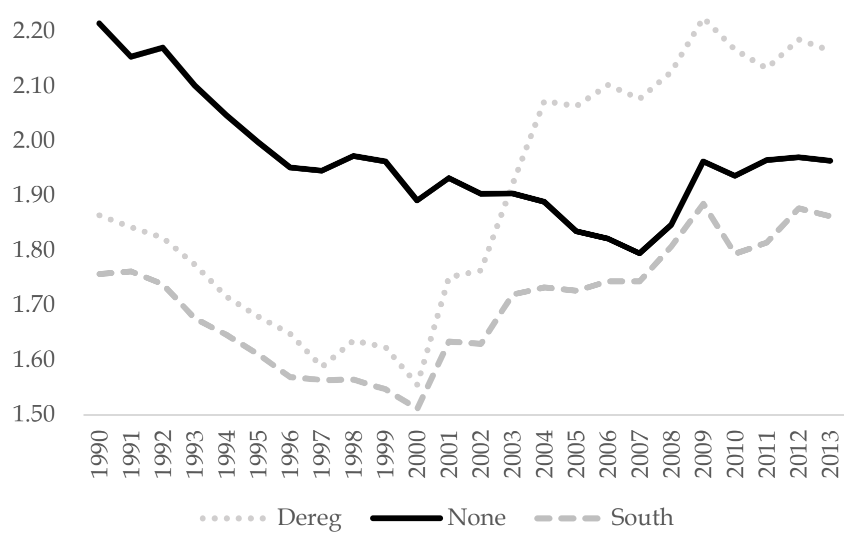
The shift towards gas-fired power plants is noted in both the economics and industrial literatures.[13] Five explanations are offered for the expansion of US gas-fired capacity. 1) The adoption of combined cycle technology and low gas prices made gas-fired plants cost effective, relative to coal and nuclear plants. 2) Deregulation of natural gas markets in the 1980s lifted restrictions on the use of natural gas in electricity generation. 3) Increase demand created a need for new investment. 4) Environmental concerns favor gas plants over coal plants. 5) Gas plants are built much quicker than coal plants. As noted later in this chapter, each of these explanations are not consistent with

[13] Kaplan (2010), Macmillan (2013), and Knittel et al (2016) summarize the gas boom.

the timing, size and cross-section of the boom, so this chapter provides a sixth explanation: restructuring lead to excessive entry in these markets and facilitated a boom.

It is important to note that, during the Gas Boom, there was a need for new investment in generation capacity. Figure 4 shows a measure of excess reserves for the three different groups of states. The measure of excess reserves in this chapter is the ratio of megawatt hours (MWh) of generation potential to electricity retail sales (ERS). [14] This ratio is split into three groups to gain perspective on the scale of the buildout in restructured states. The South is separated from the None group to provide more context.

Figure 4. Capacity to ERS Ratios by Group



Source: EIA (2016)

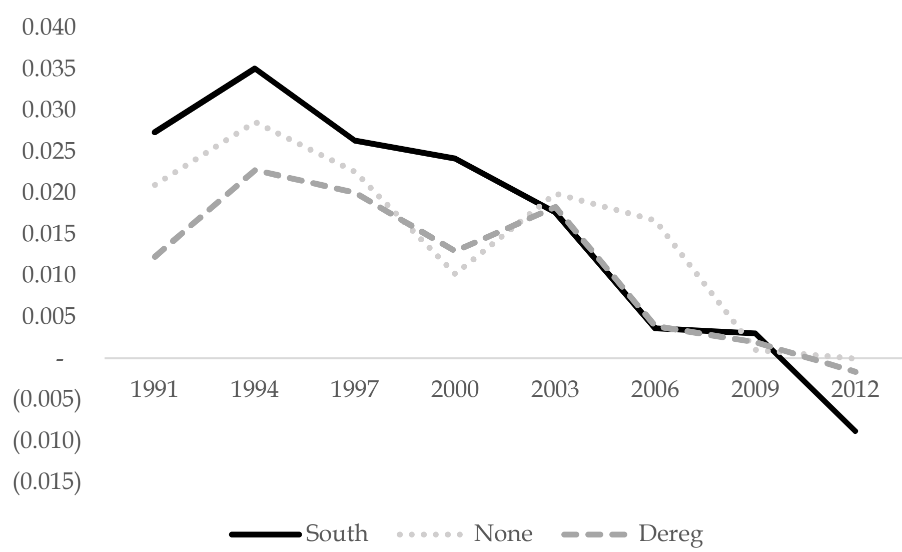
From 1990-1998, utilities across the US added little capacity due to idle plants built in the 1980s. While the optimal capacity to ERS ratio is unknown, these actions confirm that it was below that in the early 1990s.[15] Of the three state groups, the None group has

[14] Generation potential is calculated using EIA state capacity factors. Gas CT plants operate at a .3 capacity factor due to low usage. Biomass, coal, gas, geothermal, nuclear and oil plants do not have state specific capacity factors, so they are assumed to operate at the same capacity factor. Hydro, solar and wind have state specific capacity factors due to state weather patterns and resource reserves.

[15] Joskow (2008) notes the 1990s were a period of excess capacity and utilities were reducing capacity.

the largest ratio. This is mostly due to the inclusion of states like West Virginia and Wyoming, which are large electricity exporters. Starting in 1999, all three groups of states constructed new power plants, which is evidence that capacity to ERS ratios were too low. However, while the two non-restructured groups reached similar levels, there was a much larger increase in the restructured states. The ratio of capacity to ERS in these states reached historic highs (EIA, 2006). There is evidence that, prior to the Great Recession in 2008, state electricity markets attempted to reduce these high ratios, but, as Figure 5 shows, low or negative ERS growth caused capacity ratios to increase.

Figure 5. Average 3-year ERS Growth by Group



Source: EIA (2016)

Individual state ERS data are not available prior to 1990, but national-level data show a sharp slowdown in ERS growth starting in the early 1980s and continuing through the decade (EIA AEO, 1991). As the US economy grew at a faster rate in the 1990s, ERS growth increased unexpectedly, which accounts for the decline in individual state capacity to ERS ratios prior to 2000. Additionally, power plant investment involves significant lag times that range from 2-3 years for gas CT and CCGT plants, to 7 year for coal, and up to 15 years for nuclear and is heavily reliant on expectations about future demand (American Electric Power 2016; EIA 2016; Nuclear Energy Institute 2016). Plants that came online in 2000 were reacting to market conditions in 1997 and 1998.

The gas investment boom was a surprise to industry insiders. The EIA gathers information annually on the US energy sector from available data and industry experts and publishes expectations in their AEO reports. Of particular interest is their generating capacity projection over the next 20 years. These projections are a combination of (mostly) confirmed projects, along with speculation by the EIA about further investments needed to meet demand. The estimates are adjusted annually to reflect changes in either confirmed projects or updated estimates by EIA staff. As is evident in Table 2, the EIA forecasts for CCGT expansion through 2010 were increased continuously throughout the 1990s, as the construction boom was underestimated each year.

Table 2. Annual AEO Generation Capacity Projections for 2000, 2010 (GW) [16]

	Coal		CCGT		Gas CT	
AEO	2000	2010	2000	2010	2000	2010
1991	311	396	13	53	86	102
1992	317	395	15	55	88	109
1993	309	352	19	45	99	112
1994	299	326	27	58	70	87
1995	299	314	23	38	69	92
1996	301	313	23	45	88	135
1997	299	304	43	108	110	153
1998	297	305	41	107	140	191
1999	305	309	40	126	99	151
2000	302	302	51	93	93	154
2001	-	315	-	126	-	164
2002	-	306	-	140	-	129

Source: EIA AEO Archive (2016)

[16] How to read this table: Each row represents what the authors of the AEO, in that year (1991, for example), thought capacity of each of these fuels would be in 2000 and in 2010. Actual Gas CC GW capacity in 2010 was 239 GW. Actual Gas CT GW capacity in 2010 was 145 GW.

There was a change in US electricity markets in the late 1990s which caused the EIA to alter their projections. While the restructuring explanation is consistent with this trend, the hypotheses that previous restrictions, low natural gas prices and lack of available technology caused the boom are not. If these hypotheses were true, the boom would have started in the early 1990s and would have been a part of the EIA projections to 2010. Figures 5 and 6 suggest that the majority of US states had a need for capacity additions beginning in the late 1990s, but do not explain why the majority of plants were fueled by natural gas, the differences in capacity to ERS ratios among groups of states, or the shortfall in EIA expectations about capacity additions. The explanations to these anomalies are explored in the empirical analysis that follows.

3. Methodology and Data

The impact of electricity market restructuring on power plant investment is identified through the variation in investment decisions to build gas-fired plants in states over the period 1990 to 2013.[17] A reduced-form panel model with state and year fixed effects is used to measure the impact of restructuring on the annual change in state gas-fired capacity. This capacity measure is split by technology, as the motivation for building CCGT plants is assumed to be different from that for building gas CT plants. The choice in this chapter of using a panel model to explain generation investment decisions departs from the previous literature. Joskow and Mishkin (1977) and Ellis and Zimmerman (1983) used a conditional logit (CL) model for estimation. The choice of using a cost-based, discrete-choice model was due to the nature of the available data and the investment period. Both papers measured the decision of what type of power plant to build and possessed data on actual power plant construction.[18] While this method was appropriate for the time period and data the authors used, there are four problems that make a

[17] This time period was chosen due both to it containing the restructuring period and data availability
[18] Ellis and Zimmerman (1983) used a combination of actual and expected plant cost data.

reduced-form panel framework with a richer explanatory variable set preferable in describing the electricity industry of the last several decades.

First, the discrete-choice model assumes the dependent variable consists of only differences in fuel, with technology differences easily incorporated. This structure entails an analysis at the plant level and requires each plant to be the same size. [19] However, plants are not the same size and an analysis at the plant-level is subject to several problems explained further in the next section.

Second, the discrete-choice literature focused on cost as an explanatory factor for power plant decisions. During this period, the combination of rapidly expanding demand for electricity and cost-of-service utility regulation led cost-based models to largely ignore demand as a factor in power plant investment. This changed in the 1980s, with electricity industry models expanding beyond the cost-based approach (Peterson and Wilson 2011). The restructuring of electricity markets in the 1990s created further complexity in the building of these models. A subset of factors considered by planners, starting in the 1990s, included the cost of the plant, expectations about current and future electricity prices, future expectations of fuel and O&M costs, its ability to meet load requirements, the location of existing transmission lines with spare capacity, the impact of current regulation and possibility of future regulation, the probability of cost overruns, and the location of available land and water.[20]

[19] If the analysis is performed higher than the plant level, the problem is no longer discrete choice, as it entails the amount of investment in generating capacity rather than the decision to build or not.

[20] Consider the decision of where to site wind turbines. The factors involved in this decision include the wind level in the area, location of existing or planned transmission, distance to load centers, environmental impacts, state RPS, cost of materials and labor, availability of subsidies, expectations of future coal and gas prices, and a host of siting concerns explained in PSC (1999). A model based solely on levelized cost will capture some of these factors but not all of them and, as a result, offer poor predictions. For example, states such as New Jersey, California, Arizona, Nevada and Florida have made large investments in solar energy production facilities despite levelized costs that are significantly higher than traditional coal or natural gas resources (EIA, 2016). Their competition with expensive peaking plants and the need for renewable electricity retail sales to satisfy state RPS programs have spurred their construction, which would not have been predicted in a model based on cost.

Third, modeling decisions by firms as if they occur in independent environments ignores the influence of each state on the decision-making process. As a result, discrete-choice estimation may lead to biased coefficients and standard errors (Moulton 1990). While states only occasionally make capacity mix decisions directly, the influence of state regulation, topography, transmission and pipeline capacities, load and climate makes it a realistic unit of observation.

Fourth, the authors were reticent about employing expected cost data. Both made use of actual plant data and supplemented with best estimates of capital, fuel and O&M costs. This was largely due to the lack of available expected cost estimates on the detailed level necessary in these studies. These data are essential in a state-level study. Additionally, as Ellis and Zimmerman (1983) notes, using only actual data can lead to truncation bias by restricting the study to the least expensive planned plants.

This section will first motivate the investment problem at the heart of this chapter by constructing a structural framework to establish the relationship between restructuring and power plant investment. The result from this setup will be a framework that can be empirically tested using a reduced-form panel approach. Finally, this section concludes with summary statistics on power plant investment to motivate the results.

3.1 Build Model

Following the neoclassical model of investment and contributions from Bushnell and Ishii (2007), a firm enters the electricity market, or adds to its current position in that market, by constructing a power plant. The firm invests if the expected net present value (NPV) of profit earned from the operations of the power plant exceeds the investment cost. This decision includes both the impact of the power plant on profits, investment and competition today, but also in the future. Firm i chooses the level of its investment by maximizing the following:

$$(1) \qquad \max_{I_{i,t},A_{i,t}} E_t[\sum_{s=0}^{\infty} \delta^s \Pi_{i,t+s}(I_{i,t+s}, A_{i,t+s}, X_{i,t+s}, X_{-i,t+s}, \Omega_{t+s})]$$

$$\Pi_{i,t+s} = \pi_{i,t+s}\big(I_{i,t+s}, A_{i,t+s}, X_{i,t+s}, X_{-i,t+s}, \Omega_{t+s}\big) - (\psi(I_{i,t+s}))$$

where δ=discount factor, Π=net profit from the investment at time t, I=size of the power plant, A=prime mover,[21] X_i =generation portfolio of the firm, X_{-i}=generation portfolio of competitors, Ω=market conditions, π=variable profit and $\psi(I)$ is the investment cost function. The additional s subscript is included to show the effect on investment decisions on future profits, investment and competition. While the previous framework is closely linked with that in Bushnell (2007), the goal of this model is to create a framework for an empirical analysis. Keeping the insights from that paper of the inclusion of intertemporal and strategic decision-making a part of the model, the previous equation can be re-written in the following manner:

$$(2) \qquad \max_{I_{i,t},A_{i,t}} \sum_{t=1}^{T(A)}[\pi_{i,t}\big(I_{i,t}, X_{i,t}(I_{i,t}), X_{-i,t}(I_{i,t}), A_{i,t}, \Omega_t\big) - \psi(I_{i,t}(A_{i,t}))]$$

$$\psi(I_{it}) = f(A_{i,t}r_{i,t}, p_{i,t}^m I_{t+s}, R_t, Env_{i,t})$$

The first equation specifies the impact of current investments on the generation portfolios of both the firm and its competitors. As the time period moves forward and the generation mixes of all firms change, they are reacting to investment decisions made not only in that time period, but also in past periods and this specification reflects that. The second equation specifies investment as a function of the prime mover, interest rate (r), the price of construction materials and land (p^m), regulatory environment (R) and environmental regulation (Env). In order to provide a workable empirical model, the specific structure of variable profit is as follows:

[21] This is the electricity industry term for the turbine technology used in the power plant.

(3) $\pi_i = \sum_{d=1}^{D(t)} \sum_{h=1}^{24} (p_{dh}^E\, x_{idh}^D - c_i x_{id}^S\ )$

$p_{dh}^E = f(Inc, Weat_{dh}, h, ER)$

$x_{idh}^D = g(p_{dh}^E, Inc, Weat_{dh}, h, ER)$

$c_i = j[R, p_i^F(A_i, Res_i), OM_i(A_i, Env_i)]$

$x_{idh}^S = k(p_{dh}^E, c_i, I_i, X_{id}\ )$

Production constraint: $x_{id}^S \leq \mathrm{f}(I_i)$

Balancing constraint: $x_{id}^D = x_{id}^S$

Financing constraint: $\psi(I_i) \leq f(r, credit)$

where p_{dh}^E=hourly electricity price, x_{dh}^D= hourly market electricity demand, Inc=income of the region, which is assumed not to change on an hourly and daily basis. $Weat_{dh}$=hourly weather. This is typically captured by the temperature humidity and precipitation of the region in that particular hour and day. ER=economic makeup of the region the market is in, which does not vary by hour and day. R signifies whether the market is restructured or still regulated and is assumed not to vary by hour and day. c_i=marginal cost of producing electricity by the firm and is assumed not to vary within a year due to firm fuel and labor contracts. Marginal cost is determined by p_i^F=price of the plant's fuel source, which depends on prime mover choice and resource availability (*Res*), and OM_i=operations and maintenance cost of the plant, which is a function of prime mover choice and environmental regulation. $\mathrm{x}_{\mathrm{idh}}^{\mathrm{S}}$=electricity generated by the plant in each hour and day. The production of the plant depends not only on the variable profit of running the plant in each hour and day, but also on the size of the plant and the composition of the firm's generating portfolio.[22] The production constraint places an upper limit on the amount of electricity a firm can supply from the constructed plant. The balancing constraint is necessary as electricity storage is assumed to not be feasible.

[22] For example, a firm may wish to operate other plants in the market due to the cost of shutting down the plants that are currently operational.

Therefore, all electricity that is generated must be sold in that period. Additionally, firms face financial constraints based on interest rates and the credit rating of the company.

Expectations are an important consideration for utility planners. Plant profitability depends on the discounted stream of profits. Utility planners must forecast electricity prices and demand, entry of competitors and future plant construction and operation costs. Therefore, each component of this setup beyond t=1 is based on utility expectations, which are derived from past data and future projections. The following equation incorporates these expectations with the profit function above:

(4) $\max_{I_i, A_i} E_t \sum_{t=1}^{T(A)} \sum_{d=1}^{D(t)} \sum_{1}^{24} \delta^t \{p_{tdh}^E x_{itdh}^D - c_{it} x_{itd}^S - \psi(I_i)\}$

$p_{tdh}^E = f(x_{tdh}^D(Inc_t, Weat_{tdh}, h, ER_t), X_{itdh}(I_i), X_{-itdh}(I_i), R_{it})$

$x_{itdh}^D = g(p_{tdh}^E, Inc_t, Weat_{tdh}, h, ER_t)$

$c_{it} = j[R_{it}, p_{it}^F(A_i), OM_{it}(A_i, Env_{it})]$

$x_{itdh}^S = k(p_{tdh}^E, c_{it}, I_i, X_{itdh})$

$\psi(I_i) = (A_i, r_{it}, p_{it}^m, I_{t+s}, Env_{it}, R_{it})$

Production constraint: $x_{itdh}^S \leq f(I_i)$

Balancing constraint: $x_{itdh}^D = x_{itd}^S$

Financing constraint: $\psi(I_i) \leq f(r_t, credit_t)$

Taking partial derivatives $(\frac{\partial \Pi}{\partial I}, \frac{\partial \Pi}{\partial A})$ and simplifying:

(5) $I_i^{AF} = f[E_t(p_{tdh}^E, x_{itdh}^D, c_{it}, X_{itdh}, A_i, r_{it}, p_{it}^m, I_{t+s}, R_t)]$

The equation above states that the investment (MW) in a particular technology and fuel of power plant (coal, nuclear, CCGT, gas CT, renewable, oil) depends on the expectation of electricity prices, electricity demand, marginal cost of generating electricity, the generation composition of the firm, prime mover, interest rates, price of construction, future investment, and the state of regulation.

This is the individual firm framework that would be used as an estimation tool. However, there are several possible observation levels for this analysis. The micro-level consists of individual plants or firms, which may own several plants across the country. Intermediate-level observations include states or utility balancing areas, which can cross state boundaries. High-level observations consist of either North American Electric Reliability Corporation (NERC) regions, which encompass multiple states, or interconnects (East, West and Texas). A natural observation level for explaining power plant investment choices would consist of build decisions made by individual plant owners, as detailed above. These are the decision-makers that determine the characteristics of the investment and possess the most information about that investment.

However, the firm-level observation faces several problems. First, individual firm data on levelized costs, RPS, regulatory environment and demand forecasting are not available and the rise of single-plant firms makes assembling a panel challenging. Second, firms often own plants in different states, making data sets inconsistent with the introduction of restructuring. Third, there is a problem with identifying the correct electricity market a plant is built to serve, as electricity flow data from individual plants is not available. This problem is less severe the larger the region of analysis and explained further in the empirical model explanation. Based on the insights available through cross-sectional variation and the reduced impact of cross-border flows through the location of restructured states, the state is the most preferred level of observation of this group. The role of the state in determining regulation adds to the advantages of a state-level analysis.

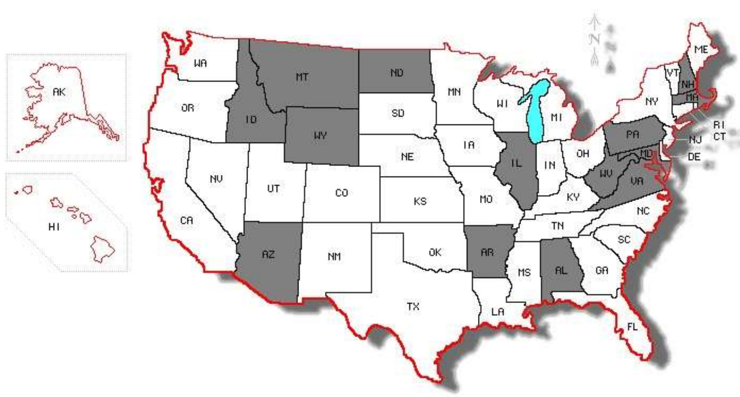
The framework above needs to be altered for the state level, as some of this information is not available to firms and not accessible at the state level. First, price expectations are based primarily on what utilities project about future entry (CAP) and demand (EIA 1997). Including these factors instead of prices increases model accuracy. Second, hourly demand is less important than its variability (VAR) and overall level (ERS). Therefore, hourly electricity demand expectations are replaced with expectations about future electricity retail sales and variability. Third, investment, operating and interest costs are included in firm levelized cost calculations.

$$(6) \qquad I_i^{AF} = f[E_t(ERS_{st}, VAR_{st}, X_{st}, CAP_{st}, R_t), LC_{st}^{AF}]$$

3.1.1 Cross-Border Flows

Figure 6 shows states with significant cross-border flows of electricity. Capacity invested in these states is often intended to serve other electricity markets, creating measurement error in the data. This problem is most severe on the plant or firm level, where electricity is impossible to trace. Using higher-level observations, such as the state, NERC region, and interconnect level, are an improvement but reduces the number of observations. There are 51 annual observations for states, 10 for NERC regions, and 3 for interconnects. As the number of observations falls, the power of the study is reduced, revealing a tradeoff between measurement accuracy and power of the test.

Figure 6. States with Significant Cross-border Flows, 2014[23]



Source: EIA (2016)

Measurement error in generation capacity due to cross-state electricity flows exists in several regions. While Texas is largely self-contained and all of the New England states

[23] Significant defined as difference between ERS and electricity production greater than 10 percent of state ERS.

except Vermont restructured, which lessens the impact of cross-state flows, borders like Illinois-Indiana and Virginia-North Carolina are a challenge for proper estimation. This requires identifying the state electricity market each plant was built to serve. While a significant amount of electricity in both regulated and non-regulated markets is sold in wholesale markets, the intended flow of electricity upon construction is what is important. There are two approaches used to estimate the intended state market for each plant. The first uses long-term power purchasing agreements (PPA), beginning upon construction of the plant, to infer which state the power plant was built to serve. The second uses ownership share data to assign capacity to each state. For example, investors in the large Palo Verde nuclear power plant in Arizona included several California utilities. A portion of the electricity produced at this plant is sent west to serve the California market. While these approaches do not capture every investment decision, their accuracy with large-scale plants ensures that the majority of cross-state flows are attributed to the intended state.

3.2 Empirical Model

The choice of size, prime mover, and fuel type of a power plant depends on the factors included in the investment equation:

(7) $$I_i^{AF} = f[E_t(ERS_{st}, VAR_{st}, X_{st}, CAP_{st}, R_t), LC_{st}^{AF}]$$ [24]

The goal of this analysis is to estimate the impact of restructuring on power plant investment. The change in CCGT and gas CT capacity are the dependent variables of interest instead of the change in all generating capacity, change in all gas capacity, or total CCGT and CT capacity. Gas capacity is the focus because there is little to explain about

[24] Levelized cost of electricity is an expectation, so including it in the expectation formula is redundant.

power plant investment during this period outside of gas-fired plants.[25] Capacity is split between CCGT and CT because this allows for better identification of the factors behind the investment decision in each prime mover. Finally, the change in the variable is analyzed, as opposed to the level, because the focus of this chapter is on additions made to utility capacity portfolios beginning in the 1990s. Analyzing the level includes past investment decisions, which are outside the scope of this chapter.

The definition of electricity market restructuring in the economics literature is focused on the opening of retail choice to consumers. However, the choice to build a power plant is influenced by generation competition, not retail competition. What allows this comparison to be made is the influence of retail competition on generation competition. As previously mentioned in Section 2, concerns about market power in generation led to state utility commissions requiring vertically-integrated utilities to sell significant assets to new and existing power producers. Figure 2 showed the effect of this policy, which increased IPP penetration to an average of over 50 percent in restructured states. Non-restructured states peaked at 15 percent, despite efforts by FERC to open transmission lines to new generation sources. Therefore, using retail and generation competition interchangeably in this analysis is feasible.

The primary equation for the panel model is derived from the theoretical analysis above, with separability and linearity assumed. State fixed effects are included in some specifications to control for possible time-invariant state differences which may bias the results. For example, a state's resource endowment will be captured in these fixed effects.[26] Year effects are also included to capture broad US trends, such as changes in technology and booms and busts in the business cycle.

---

[25] 84 percent of net capacity additions in the US from 1990-2013 were in gas-fired power plants. Leaving out renewable additions from 2007-2013, which are largely attributed to state RPS requirements, increases this amount to 95 percent (Powers and Yin 2010).

[26] An example of a factor accounted for in state fixed effects is California's preference for natural gas over coal due to environmental reasons.

$$(8) \qquad \Delta MW_{st}^{G} = aRestruct_{st} + \beta RestructxPeriod_{st} + \delta X_{st} + \theta_s + \omega_t + \varepsilon_{st}$$

In the equation above, s indexes states and t indexes years. *ΔMW* refers to the nameplate capacity of a plant in megawatts (MW) and is measured as the change in gas capacity in a state in a given year.[27] This equation is analyzed separately for CCGT capacity changes and gas CT capacity changes. *Restruct* is a binary variable detailing whether the state's electricity market was restructured. *RestructxPeriod* measures the impact of restructuring on power plant investment in groups of years after restructuring. This enables the model to capture short run dynamics of restructuring. Given the role of expectations in the impact of restructuring, the two restructuring variables have a lag to account for time to complete the power plant and begin when a state sends a strong signal it will restructure by passing legislation. *X* represents a set of controls in this regression, which include lagged capacity to ERS ratio (a ratio of available supply of electricity to demand (*CapERS*)), expected lagged electricity demand growth, expected lagged load variance, and levelized costs of competing fuels and technologies.

The coefficient of interest in this chapter is the interaction variable (*RestructxPeriod*). The choice to include a variable separating time periods following restructuring tests whether investment decisions in restructured versus non-restructured states differed based on the time period. This test the well-known theory of new-market investment following an S-shape, with rapid entry, a shakeout as firms leave, and an established equilibrium (Klepper and Miller 1995).

Firm expectations about future supply and both the level and variation of demand are important components of the investment decision. While firms assume that a decline in the CapERS ratio will be corrected in the future, as it encourages entry, their investment also serves to deter future entry by increasing the ratio. As a result, they are assumed to

[27] This functional form uses change instead of log due to the extreme swings in the data, with some states having no capacity prior to the gas boom.

take advantage of decreases in this ratio by investing, with expectations about the ratio in the future being less important due to their impact on other firms' investment. However, the firm does rely on expectations about future demand, as declining or slow growth in electricity demand reduces the profitability of plants. The level of growth is important to ensure higher prices and prevent idle plants, while the variance determines what the market for the plant will be. Each of these are lagged to correspond with when the decision to build the plant was made. In the dataset for this chapter, the observation is when the plant began operation, not when construction started.

A critical undertaking in this chapter was constructing cost data for each type of generating plant in each state and year, which did not exist in a complete panel. The structure of levelized costs can be modeled in the following way:

$$(9) \qquad LC_{st}^{AF} = f[CapCost_{st}^{A}, Life^{A}, DiscRte^{A}, E(p_{st}^{AF}, O\&M_{st}^{A})$$

where *CapCost* is the cost of plant by primemover, *Life* is the plant life by prime mover, and *DiscRte* is the discount rate applied to the plant by prime mover. Expectations only factor into levelized cost for fuel and O&M, as these are the costs of the plant impacted by future events. Plant life and discount rate vary little across time and state and only fuel prices are influenced by the choice of fuel. The other factors are distinguished by prime mover. Power plant investments are impacted by many factors, due to their size and length. Planning and construction alone take 7 years for coal plants and up to 15 for nuclear plants. Once online, a plant will operate, in the case of some coal and hydro plants, for as long as 70 years. Plant owners rely on expectations of future fuel prices and regulation and are wary of the near bankruptcy of nuclear power plant builders in the 1980s due to unexpected cost overruns (EIA 1993).

In a regulated state electricity market, utilities were confident in making these investments, as they could expect to be reimbursed for large capital outlays through cost-of-service regulation. There were no guarantees in restructured markets, which experienced a reverse Averch-Johnson effect. Capital intensive investments are a liability in markets where electricity prices could rapidly rise or fall, as they did in the mid-2000s. In these markets, expectations of fuel prices were key in new power plant investment.

3.2.1 Restructuring Exogeneity

A critical assumption in this model is the restructuring process had an exogenous influence on changes in state gas-fired capacity. If another factor was responsible for the gas buildout decision in restructured states, the empirical results of this chapter would be spurious. These concerns have undergone extensive vetting in the restructuring literature, as policy variables are often endogenous (Fabrizio et al 2007; Davis and Wolfram 2012; Cicalla 2015). There is evidence that the structure and expertise of state legislatures and high electricity prices were correlated with state decisions to restructure (White 1996; Ando and Palmer 1998; Damsgaard 2003; Ardoin and Grady 2006). There is no established link between legislature structure and preference for gas-fired capacity, so it is not considered a threat to identification. The link between electricity prices and restructuring is addressed in Fabrizio et al (2007), so it is worth noting how these authors, as well as Davis and Wolfram (2012), address endogeneity concerns.

The restructuring endogeneity concern is there is some factor influencing the decision to build gas-fired plants that changed during the restructuring time period and only for restructured states. Both Fabrizio et al and Davis and Wolfram note that any unobserved differences among states are most likely time-invariant and do not have a plausible connection to the dependent variable. This includes electricity prices, with a gap between restructured and other states existing for a long time period due to previous generation choices. The authors also used alternative methods to account for specific plant and geographic patterns relevant to their papers, but these were in support of their main specification. Instead, both papers rely on observation-level fixed effects to negate potential identification concerns.

While this chapter employs a similar strategy, the timing of restructuring is an important factor which reduces the chance of policy endogeneity. The time path for restructuring varies widely among states, with some states choosing to restructure early, some states employing a lengthy process, and others moving quickly to restructure. As a result, several states were still in the exploratory phase when the California Electricity Crisis occurred. This event is noted by several state public utility commissions (PUC) as part of their decision not to restructure. This exogenous event, due in large part to reduced rainfall in the Northwest and a summer spike in load, determined the restructuring path for multiple states. For those states that restructured early, due in part to high electricity prices, there is a question as to what made those states delay. Electricity prices were high for years. It is likely that the deregulatory environment during that time period in the US encouraged states to restructure their electricity markets, which has no plausible connection to states engaging in a large gas buildout.

Although the timing of restructuring is important in the exogeneity of restructuring, the primary evidence of this chapter is that states not only built gas, but built it in large quantities. The size of the buildout drastically increased the capacity ratios in restructured states and drove down average capacity factors for power plants. The only reason for these high ratios would have been if these states were large exporters of electricity. However, not only were these states not exporters prior to restructuring, but their low capacity factors following restructuring showed that they were not selling the additional electricity the new plants were capable of producing. Therefore, following Fabrizio et al (2007), Davis and Wolfram (2012) and Cicalla (2015), restructuring is assumed to be exogenous with the addition of year and state fixed effects.

3.3 Data and Summary Statistics

This chapter's approach to analyzing changes in electricity capacity requires a large amount of data on the individual state level from 1990 to 2013. All 50 states were included in this analysis.[28] While it would have been useful to incorporate data prior to 1990 in the estimation process, this is as far back as reliable state-level cost and demand projections are available. Data on plant capacity by state and year and ownership share are from EIA form 860. Additional nameplate capacity, demand and capacity factor information are from the EIA's AEO reports. FERC provides hourly load data.

There is no single source that provides estimates of levelized costs for each fuel and technology in each state from 1990-2013. Actual plant cost observations are available only in certain states, as not all fuels and technologies were constructed in each state every year. Using these estimates also creates the hypothetical bias mentioned in Ellis and Zimmerman (1983). Using estimates from several sources prevents a bias in the data coming from one source.[29] As a result, the data are a combination of both actual and estimated costs. Actual plant cost data are provided by Ventyx. The current and expected future prices of coal, natural gas, uranium, oil and biomass are provided by the AEO. Estimates of plant life are provided by the National Renewable Energy Laboratory (NREL), International Atomic Energy Agency (IAEA) and the National Association of Regulatory Utility Commissioners (NARUC). Actual and estimated O&M data are provided by Ventyx and the EIA. Capital cost data are from a variety of sources, including the AEO, the MIT interdisciplinary reports on natural gas, coal, nuclear, geothermal and solar, the University of Chicago report on nuclear power, Ventyx and NREL.

The restructuring variable uses detailed information on restructuring available from both the EIA and individual state regulatory commissions. These data list the major steps states initiated along the restructuring process. These include dated commissioned

[28] DC was excluded, as its small amount of capacity did not change during the time period.
[29] Utilities have frequently complained about the shortcomings in some of the EIA data, to the point where the EIA altered its methodology in 2010.

reports, pilot programs, legislative action, divestiture of assets, and the opening of markets to retail competition. A state is considered in this study to have restructured its electricity market after the state legislature passed a law directing the state regulatory commission to open the state's electricity market to retail competition. Assigning years to signal the official start of restructuring is challenging because of the role of expectations in these markets. This information was not hidden, so firms would be aware of the steps taken by states to restructure. Given the lag time in power plant construction, a firm seeking to enter a newly restructured market that is unable, or unwilling, to purchase a divested plant may invest prior to the official start date of the market. This would allow the plant to enter the market sooner than firms that began construction upon the date of retail competition. Therefore, this study assumes that a legislative order was considered sufficiently permanent to incentivize firms to begin construction of new power plants.

Throughout this chapter, a comparison is made between three groups of states: Restruct, South, and None. The Restruct group consists of states which started the process of retail competition through legislative or regulatory action. The South group consists of states located in the southeastern part of the United States, none of which restructured. The None group consists of the states that are not part of the South group that did not restructure. Separating these three groups provides for an intuitive understanding of the changes in this period and differences between the states.

Table 3 shows several important state variables by state group. The Restruct group has almost double the ERS, compared to the None group due to the inclusion of large population states like Texas and California. ERS growth in the South group was larger than the other two until 2008, which is an important component in understanding the increase in gas capacity in the South group. Finally the None group was heavily invested in coal and hydro and contained several exporting states, partly explaining its small investment in gas during the boom.

Table 3. State Group Characteristics

	# States	ERS	Coal %	Gas %	Nuclear %
Restruct					
1990	20	1,419,891	33.4%	25.1%	16.7%
2000	20	1,726,561	29.9%	36.2%	13.3%
2013	20	1,858,429	21.2%	48.3%	10.2%
None					
1990	21	607,544	65.5%	5.6%	5.7%
2000	21	770,348	59.0%	13.1%	4.7%
2013	21	864,852	45.7%	23.7%	3.6%
South					
1990	9	675,271	44.2%	14.7%	18.7%
2000	9	913,889	38.3%	25.8%	16.3%
2013	9	990,697	26.4%	46.5%	12.6%

Notes: ERS stands for Electricity Retail Sales. Coal%, Gas %, Nuclear % are each fuel's share of capacity for each region. Source: EIA (2016)

4. Empirical Results and Analysis

The results of this analysis are separated between CCGT and CT plants. The explanatory variables in each regression differ slightly, as the two technologies are built for different purposes. Their occasional overlap, particularly in the case of CCGT plants providing peak output, necessitates an analysis of both investments.

4.1 CCGT Impact

Table 4 shows the impact of electricity market restructuring on CCGT capacity. The variable being explained is the change in CCGT capacity in each state in a given year. The columns are differentiated by the number of years after restructuring. The primary explanatory variable, an interaction between restructuring and the post-restructuring time period, is lagged three years to account for plant construction. Therefore, column 1 shows the average change in CCGT capacity for restructured states four to seven years after restructuring compared to non-restructured states, column 2 shows this change for years four through eight and so on. Controls are included for available capacity to meet demand (CapRatio), expected electricity demand (ERSproj), expected load variance (Varratio) and the levelized cost ratio of CCGT to coal and state and year fixed effects. For alternate specification and robustness checks, see Section 5.

Table 4. Effect of Restructuring on CCGT Power Plant Investment

Period	(4)	(5)	(6)	(7)	(8)
Restruct	235.24**	116.83	14.37	37.52	49.25
	(118.12)	(71.29)	(91.72)	(96.85)	(99.22)
Restruct*Period	223.01	436.2**	593.71**	522.31**	485**
	(151.96)	(178.48)	(276.40)	(259.82)	(224.72)
CapRatio	-277.24***	-240.49***	-230.6**	-264.64***	-293.51***
	(87.75)	(82.86)	(87.89)	(87.07)	(91.51)
ERSproj	15.73**	15.18**	14.75**	13.77**	13.53**
	(6.39)	(6504.97)	(6.53)	(6.30)	(6.16)
Varratio	22***	21.54***	21.16***	20.73***	20.69***
	(7.43)	(7.24)	(7.11)	(6.87)	(6.73)
Gascoalratio	-4.85	-3.77	-5.87	-8.11	-9.52
	(13.45)	12.85	(13.69)	(13.80)	(13.34)
State, Year FE	Y	Y	Y	Y	Y
R^2	0.16	0.17	0.18	0.17	0.17

Notes: N=1050. Dependent variable is change in CCGT capacity (MW). *** significant at .01 level. ** significant at .05 level. * significant at .1 level

The peak impact of restructuring on CCGT capacity expansion occurs between years four and nine following restructuring (Period 6). Restructured states during this period add 593 MW additional capacity each year compared to non-restructured states throughout the sample period. As the number of years in the period increases, this effect is diluted but still significant 12 years after restructuring and nonexistent in a period of four years or less. RatioCapERS is significantly negative, suggesting that states with large amounts of capacity built very few CCGT plants. The signs and magnitudes of ERSproj and Varratio suggests that more CCGT plants were constructed when expected future demand and demand variance was greater.

There are three important conclusions to gather from the results in Table 4. The first is the magnitude of the CCGT capacity addition difference between restructured and non-restructured states. By 12 years after restructuring, a restructured state added over 4.1 GW of CCGT capacity more than a non-restructured state with similar supply and demand characteristics. Second, the expansion period was lengthy, with peak additions occurring nine years after restructuring and six years after plants began coming online. It was large enough to show a significant differential between restructured and non-restructured states 12 years after restructuring. Third, this relationship holds after controlling for a large number of factors, suggesting it is driven by restructuring and the nature of the market after free entry rather than other changes within the industry to which it was attributed.

4.2 CT Impact

Table 5 shows the effect of restructuring on CT plant investment. The variable explained in this table is the change in CT capacity in each state and year. Unlike CCGT, the effects are not separated by the period of time after restructuring, as there were no significant effects. Instead, each column is separated by the binary variable separating the time period. This setup not only shows what impacts CT construction, but also explains the negative effect of restructuring on CT plant investment. Other explanatory variables are included following the model setup at the beginning of this chapter, with a peaking ratio replacing the total capacity ratio due to its relevance to decision-making.

Table 5. Effect of Restructuring on CT Power Plant Investment

Period	(1)	(2)	(3)	(4)
Restruct	-101.95**	-123.96***	-110.34***	-126.47
	(42.18)	(36.34)	(46.96)	(82.64)
ERSproject	0.48	0.77	0.45	0.66
	(3.71)	(3.45)	(3.70)	(3.55)
Varratio	9.43	9.87	9.46	9.46
	(6.51)	(6.46)	(6.59)	(6.41)
PeakRatio	-2.32**	-2.51**	-2.39**	-2.16**
	(1.11)	(1.16)	(1.17)	(1.12)
Fuelratio	2.76	3.88	2.59	3.56
	(14.25)	(15.36)	(14.33)	15.26
Interval		-35.34	-30.74	-455.14***
		(28.38)	(27.94)	(92.99)
Restruct*Interval		189.47	14.66	45.17
		(224.14)	(105.79)	(112.97)
State FE		Y	Y	Y
R2	0.11	0.11	0.11	0.12

Notes: N=1100. Dependent variable is change in gas CT capacity (MW). *** significant at .01 level. ** significant at .05 level. * significant at .1 level

Column 1 shows the difference between CT construction in restructured and non-restructured states from 1990-2013. Restructured states built 100 MW less than non-restructured states each year, with the supply of peaking resources (represented by PeakRatio) having a significantly negative impact on CT construction. Given the expansion of CCGT plants detailed in the previous section, there is clearly a relationship between the impact of restructuring on excess CCGT construction and the dearth of CT construction. The nature of this relationship is in question, as it is feasible that either firms in restructured states focused more on CCGT and less on CT construction or the large CCGT expansion created excess capacity, which discouraged CT investment. This effect can be determined by the time period in which it occurred. If it is concentrated during the years of the CCGT expansion (2000-2006), then it was the preference for CCGT over CT that both explains some of the excessive build in CCGT and the lack of construction of CT plants in restructuring states, relative to non-restructured. However, if the effect is concentrated after 2006, the CCGT expansion was due to other factors and caused a later slowdown in CT plant construction.

Columns 2-4 identify the timing of this effect. Column 2 splits the effect of restructuring between pre-2000 and post-2000, with the significance of the restructuring variable suggesting the negative effect after 2000 is still strong. Column 3 increases the split to pre-2006 and post-2006, with similar results. Clearly, there is a negative impact of CCGT construction on CT plant construction post-2006. However, this does not leave out the possibility that CCGT construction was also substituting for CT plants during the boom. Column 4 separates the time periods into the boom period (2000-2006) and the non-boom (pre-2000 and post-2006). The lack of significance on the restructuring variable shows there is not strong evidence for substitution of CT for CCGT plants. Therefore, it is safe to conclude that CT plant construction declined in restructured states because of the expansion in CCGT power plants.

4.3 Analysis

The results in Tables 4 and 5 attribute a significant part of the gas boom to restructuring, with restructuring leading to an excess in CCGT capacity. Previous explanations of the gas boom focused on the adoption of new technology, the impact of low natural gas prices, deregulation of natural gas markets, increasing electricity demand, environmental concerns, and the fast construction time of gas plants. While many factors played a role in the gas boom, the timing and the geographic differences of the previous explanations are at odds with the nature of the boom. Each of these issues are addressed in the following paragraphs.

The introduction of both combined cycle and jet engine technologies into electricity markets had an impact on the cost of power plant construction. For CCGT plants, the levelized cost of capital declined due to the increased efficiency of the technology. [30] For gas CT plants, the introduction of jet engine technology not only reduced construction cost and increased efficiency, but also decreased the start-up time, which is important for plants designed for peak use. Both of these technologies lowered the levelized cost of electricity, making gas competitive with coal. However, this does not explain the timing of the gas boom, as these technologies were constructed in Japan, the UK, and the US beginning in the 1980s. Additionally, while this contributes to explaining why the majority of power plants built during the boom were fueled by natural gas, it does not explain why restructured states built CCGT plants than non-restructured states.

Similar to the adoption of new technology, the decline in long-run expectations of natural gas prices led to a reduction in the levelized cost of electricity from gas compared to other fuels. This decline was due largely to political factors in the Middle East and the introduction of new sources, both domestically and internationally (EIA, 1994). The combination of the technology change and the long-run expectations adjustment, as well as the flexibility of non-steam power plants, led to natural gas being preferred as a fuel in power plant construction over coal, nuclear and biomass. Therefore, this explains the choice of natural gas across the country, but not why it differed among states. It also conflicts with the timing of the boom, as these changes occurred in the early to mid-1990s.

In the mid-1990s, electricity demand growth began to pick up as a result of the increase in economic growth, as well as the further adoption of household electronics. The increase was largest in the South, both for these reasons and due to migration, but the Restruct group also experienced an uptick in electricity demand growth. Combined with the dearth of power plant construction during the 1990s, this created a need for new generating capacity just as the gas boom began. As a result, the timing of the increase in electricity demand fits the gas boom well, but does not explain the difference between investment in the Restruct group and the South group. The South group's electricity

[30] In the electricity industry, plant efficiency is measured by heat rates. This is the amount of BTUs of fuel energy necessary to create a kWh.

demand grew faster than the Restruct group but built fewer power plants in the 1990s than the Restruct group did. Electricity demand growth suggests those states should have built more plants than the Restruct group, while the opposite actually occurred.

Environmental pressure is a factor in the choice of natural gas plants, as they are cleaner burning than coal on a wide range of emissions. While there is no significant change in environmental protection during this time period, measured both in changes to the National Ambient Air Quality Standards (NAAQS) and regional agreements, it is still possible that utilities would be more wary of future regulation and curtail their choice of coal power plants in favor of gas. This could also explain regional variation in choice, with political differences existing between states that restructured, which tended to be left-leaning, and those that did not, which tended to be right-leaning. However, there was no regional variation in fuel choice, as almost no coal capacity was added during this period. Instead, there was variation in the size of investment, which can't be explained by differences in environmental attitudes.

The ability of gas-fired power plants to be built faster, and started quicker, is significant in the choice of fuel and technology. This is especially important in states that experience greater load variance and for those whose markets are liberalized. The average construction time for a gas CT plant is 2 years, while a gas CC plant can be built in up to 3. This compares favorably to coal, which can take 7 or more years, and nuclear, which began taking 15 to 20 years. This is an important factor for meeting immediate capacity needs but is crucial in restructured markets. Firms operating in liberalized electricity markets no longer receive guaranteed rates from the state regulatory commission. Instead, they are open to market fluctuations, which make it more difficult to pay off the large fixed costs associated with coal and nuclear power plants. The enhanced risk in these type of plants ensured that firms entering newly restructured markets would focus on gas plants. However, gas power plant construction occurred in both restructured and non-restructured markets and their ability to serve restructured markets does not explain why they were built in excess of demand.

The failure of these factors to explain the timing and scale of the gas boom suggests that alternative explanations are needed. This chapter has found separate explanations for CCGT and CT plant construction, with Table 4 showing restructuring played a large role in the construction of CCGT plants. The investment paths of the two plant types are explained in the following two sections by combining the explanations of previous analyses with new insights from this chapter.

4.3.1 Explaining the CCGT Boom

The South group of states was in need of capacity near the end of the 1990s due to a decade of low power plant investment and high ERS growth. CCGT plants were constructed in the South to meet the need for base load capacity and because the levelized cost of gas had fallen below coal. The leveling off that occurred in the South group's capacity to ERS ratio following 2006 was within the bounds of previously normal ratio levels, suggesting capacity had met its desired level. The increase in this ratio following 2008 is due to declining ERS, not changes in plant capacity.

The None group of states was not in need of capacity additions, in comparison to the Restruct and South groups, because of high capacity to ERS ratios entering the 1990s. As a result, there was significantly less construction in these areas. Power plants that were built during this time period were largely due to state RPS requirements or to meet outside state demand (Indiana, for example). The leveling off of these states at normal ratio levels suggests, as with the South group, that capacity had met its desired level.

The CCGT investment in the restructured states can't be explained by demand or supply factors. ERS growth did increase in the 1990s, but it was less than the South and significantly less than what would have been needed to purchase all of the new capacity coming online. Similarly, capacity ratios in restructured states were low in the late 1990s, but increased well past previously normal levels. Based on the results in Table 4, there is strong evidence of excessive entry and overinvestment in restructured markets.

A statistic used to measure excessive entry in electricity markets is the ratio of generating capacity to ERS. This annual ratio compares the megawatt hours of electricity that can be produced in a state if each plant operated at the planned capacity factor with the megawatt hours of electricity sold in each state. Figure 5 shows a decline in the ratio in both restructured and non-restructured markets from 1990-1998, followed by an increase until 2006, a slight decline from 2006-2008, and then a stagnant period until 2013. The difference in the ratios between the state groups beginning in 2000 provides further evidence that there was excessive entry into restructured markets.

The large number of bankruptcies, or near bankruptcies, and asset fire sales by both utilities and IPPs provides further evidence of excessive entry in these markets.[31] Descriptions of these firm difficulties regularly mention an oversupply of electricity. Additionally, the timing matches the description in Klepper and Miller (1995) of the shakeout, where less efficient firms leave and the market moves towards a long-run equilibrium. The decline in the capacity to ERS ratio post 2006, combined with the bankruptcies, suggests that the electricity market in deregulated states may have been starting to move to a long-run equilibrium before the combination of the 2008-2009 recession and the increase in energy efficiency programs led to a drop in ERS growth.

The factors mentioned above, combined with observations in papers like Borenstein and Bushnell (2016), which noted that large reserves required by ISOs do not fully explain the level of excess reserves in restructured markets, lead to the belief that restructuring electricity markets resulted in excessive entry and overinvestment. What is not clear, thus far, is the mechanism behind this excessive entry, which contradicts the

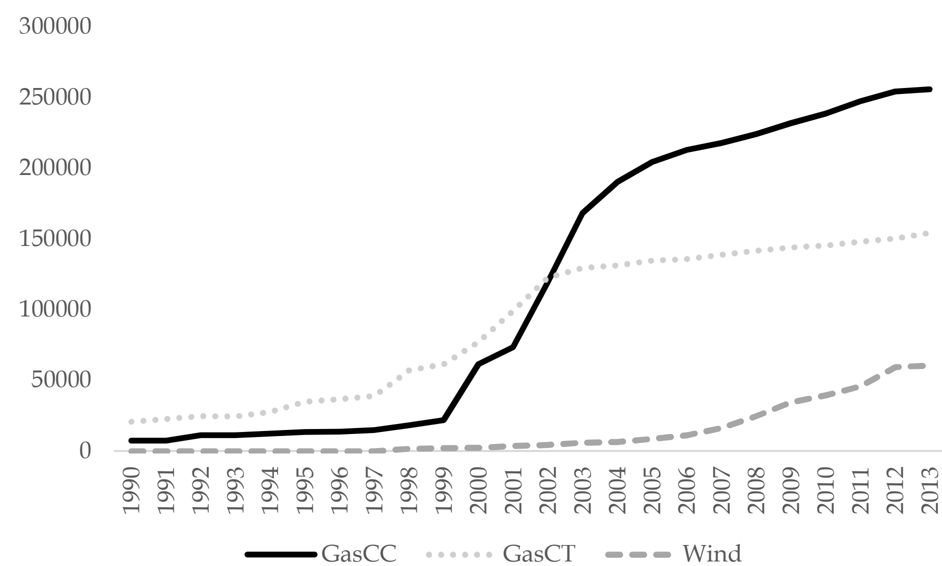
[31] Calpine, Mirant and NRG are a few examples of major bankruptcies occurring in the mid-2000s. Major utilities that were approaching bankruptcy include Pacific Gas & Electric and Southern California Edison. Dynegy, Williams, El Paso Energy and Duke are examples of large power suppliers that engaged in asset fire-sales and halted merchant energy trading. (Anderson and Erman 2006; Wharton 2006)

medium to long-run efficiency gains restructuring was supposed to provide (Joskow 1997). Chapter 2 of this thesis provides explanations for how restructuring could have led to overinvestment in newly liberalized markets.

4.3.2 Explaining the CT Boom

The increase in CCGT capacity was the most significant part of the US Gas Boom, but there was also a large increase in CT capacity around the same time period (as shown in Figure 7). Despite their similarities, the CT boom differed from the CCGT boom in several important aspects. First, the timing and spatial variation of the two buildouts are separate. Second, restructuring was not an important factor in the CT construction phase, as opposed to the impact on CCGT. Third, the investment decision in CT plants is different from that in CCGT plants. Each of these differences are expounded upon below.

Figure 7. Investment in Generating Capacity by Technology (MW)



Notes. Wind included for context on gas boom. Source: EIA (2016)

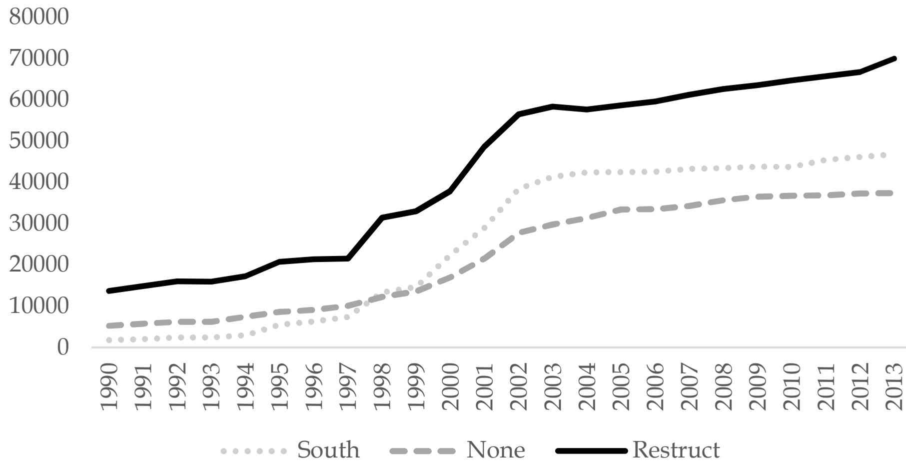
CT power plants are built to meet intermediate and, particularly, peak load demand. These are low capital, high marginal cost plants with the flexibility to start up quickly. Using the technology from jet turbines, CT plants were introduced in the US in the 1980s as an alternative to oil and gas steam turbine power plants. While the majority of these plants are able to operate using oil or natural gas, the cost difference prompts most utilities to choose natural gas as a fuel. The typical construction time for a CT plant is 2 years, which is quicker than CCGT (3), coal (6), and nuclear (15+). As Figure 7 shows, there was a large buildout of these plants from 1997-2002. Assuming a 2 year lag between the start of planning and construction and the beginning of operation, the relevant time period when analyzing CT investment is 1995-2000.

The timing and pattern of the CT buildout shown in Figure 6 contradicts the assumption that utilities add peaking capacity as needed. This strategy would consist of small additions throughout the years, which would appear as a more linear investment strategy. The actual buildout differs from this linear projection in several time periods, with less capacity added prior to 1997, more added between 1997 and 2002, and less after 2002. Of these three, the after 2002 period is most easily explained, as it was impacted by the large expansion from 1997-2002. With so many plants available from both the CT and CCGT buildouts, there was no need for further significant capacity additions.[32] Therefore, the focus of this section explains the CT investment pattern of 1990-2002.

Given the significant spatial variation between three groups of states (Restruct, None, South) in the CCGT, the CT buildout would be assumed to follow a similar pattern. Figure 8 shows this variation in CT plants. There are some differences between the groups of states, with the None buildout being smaller in magnitude and scale compared with the buildout by the other two groups. However, CT expansion is influenced by ERS growth in the period.

[32] This effect is apparent in the results in Table 5.

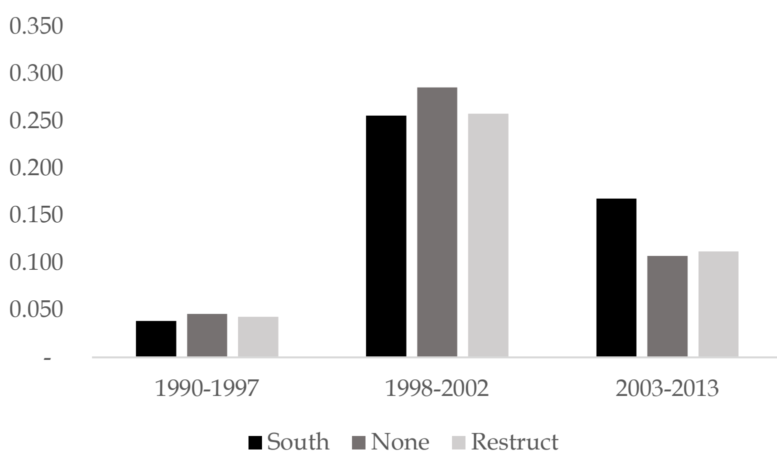
Figure 8. Gas CT Capacity Additions by Region (MW)



Source: EIA (2016)

Figure 9 shows the change in the ratio of Gas CT capacity to the change in ERS in three time periods. After this adjustment, there is no clear difference between the groups in the first two periods, with the South adding more CT plants in the third period as the only slight deviation. This result is in sharp contrast with the CCGT buildout, which was most heavily concentrated in the Restruct states. Combined with the results from Table 5, it is apparent the CT buildout was a national trend.

Figure 9. Gas CT Ratio by Time Period



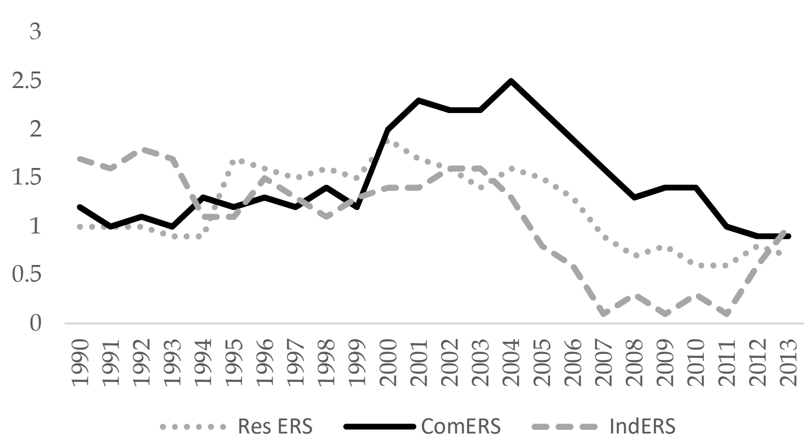
Notes: Ratio constructed comparing Gas CT and ERS additions. Source: EIA (2016)

Focusing on the national trend, there are three factors that influenced decisions to construct a large number of CT plants beginning in 1995: 1) a change in residential demand growth forecasts by utilities. 2) Falling relative natural gas price expectations and 3) the impact of the drop in ERS growth starting in the late 1970s. The combination of these factors influenced the timing and magnitude of the CT capacity buildout.

CT plants are built to meet peak load, which occurs during the summer in the late afternoon and early evening.[33] These are the hours of the day where residential demand peaks, as families return home and turn on the lights, TV and air conditioning. Therefore, investment in peaking plants by utilities is sensitive to changes in residential retail sales expectations (EIA 1995). On the other hand, increases in commercial and industrial ERS flatten the load curve, as they peak during the late morning and early afternoon.

Figure 10 shows two trends which impacted the investment behavior of utilities. First, the increase in residential ERS growth and decrease in industrial ERS growth expectations in 1995 would increase demand during peak hours, increasing the need for CT plants. Second, the increase in commercial and decline in residential ERS growth expectations would have flattened the load curve, reducing CT plant demand.

Figure 10. EIA 25 Year Projections of ERS Sector Growth



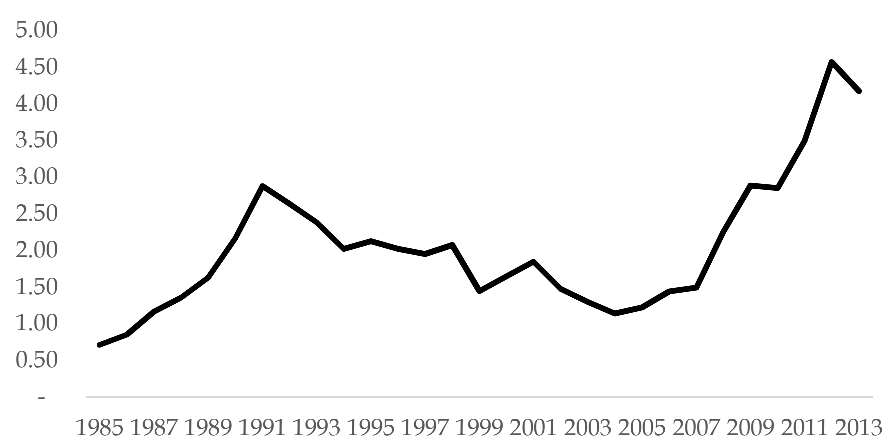
Source: EIA (2016)

[33] Winter energy demand for heating is met largely by natural gas and propane (EIA 2016).

The timing of both trends fits the CT investment pattern in the previous figures. Low residential and high industrial demand projections in the early 1990s coincided with only marginal investment in CT plants. The boom period of CT construction coincided with increasing residential demand projections and the decrease in industrial demand projections. The end of the CT boom coincided with the increase in commercial demand projections and the decline in residential demand projections.

A further factor influencing CT expansion was the change in relative price expectations between older oil and gas steam plants and newer combustion turbine plants. Figure 11 shows the change in the levelized fuel cost of oil and natural gas, beginning in 1985.[34] A high number implies utilities are more likely to switch from oil to gas, while a low number implies more oil and less gas. The acceleration of this ratio in the 1980s led to a value above 2 from 1988 to 1998, which influenced utilities to consider building CT plants. While the ratio is not the only factor in the decision to build Gas CT, it significantly influenced the decision-making of utilities (EIA AEO 1996).

Figure 11. Oil to Gas Levelized Fuel Cost Ratio



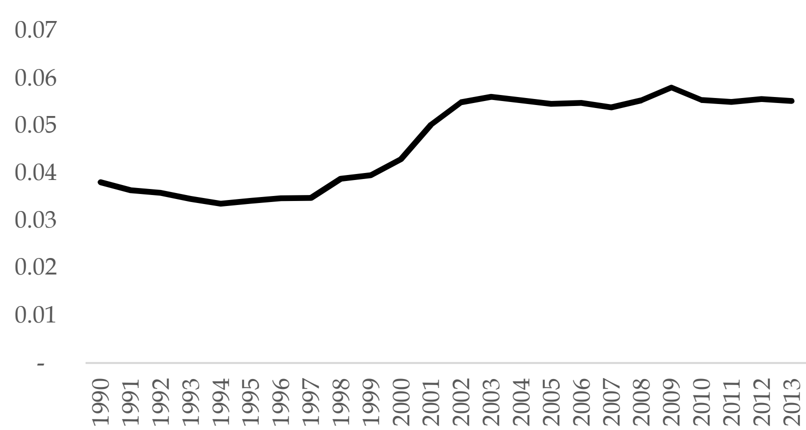
Notes: Switch from steam to CT assumed to occur in 1990. Source: EIA (2016)

[34] This is not a ratio of annual fuel prices, but rather includes projections on what the EIA's best estimate of future prices is. Each is converted from mmBtu to kWh to reflect different heat rates. CT plants are assumed to be available beginning in 1990. The ratio is mostly being driven, however, by changes in the fuel prices themselves.

Despite changes in ERS sector projections, an increase in ERS growth and an increase in the oil to gas fuel ratio during the 1990s, the lack of investment in the early 1990s and the size of the CT boom are not fully explained. The EIA AEO publications in the late 1980s and early 1990s mention repeatedly the hesitancy of utilities to invest in further capacity, even if only for peaking, due to the excess capacity they were left with after ERS growth changed in the late 1970s. This explains the lack of investment in the early 1990s and, due to both the increase in demand and the falling capacity ratios, explains why a surge in CT investment would occur in the mid-1990s.

The nature of other investment booms in electricity generation (coal and gas plants in the 1950s, nuclear in the 1970s, and Gas CC in the 2000s), the uncertainty of relying on ERS projections and the inevitable coordination failure lead to the belief that the investment pattern will often be more uneven than the optimal. However, looking at peaking ratios in Figure 12, it appears they reached their highest point in 2002 and have stayed relatively consistent since that time.[35] This suggests that the amount of CT capacity added in the late 1990s and early 2000s was close to the amount utilities required.

Figure 12. Estimated US Peaking Ratio



Notes: Assumes oil and gas steam units are used for peaking. Source: EIA (2016)

[35] The peaking ratio consists of the capacity of oil and Gas CT plants divided by ERS. While this is not a perfect peaking measure, it is a strong first approximation and explains the trend in peaking capacity.

The combination of changes in sector ERS growth projections and the relative fuel cost of natural gas, along with the impact of the 1980s buildout, explains the CT construction boom of the late 1990s and early 2000s. This boom was a national trend, with little difference between the South, None and Restruct groups of states. Following the end of the CT capacity expansion, the estimated US peaking ratio stayed constant, implying that, unlike the CCGT boom, the desired number of CT plants were built.

5. Alternative Specification Tests

Section 4 showed evidence of restructuring leading to a CCGT boom, which exceeded the capacity required to meet demand. To provide further support, alternative approaches are presented which eliminate potential confounding concerns. These include synthetic control, restricted data sets, and alternate LHS and RHS variables.

5.1 Synthetic Control

The staggered and incomplete adoption of restructuring by states created a reasonable counterfactual for the restructured states. However, synthetic control constructs the counterfactual in a more precise manner. Applying the method present in the literature to electricity markets (Abadie and Gardeazabal 2003; Abadie et al 2010; Buchmueller 2011; Bohn et al 2014), states are divided into restructured and donor groups. For each restructured state, the goal is to construct a counterfactual state from the donor group which shows the amount of CCGT capacity that would have been added in the restructured state if restructuring had not occurred. Following the notation in Abadie et al (2010),[36] the change in CCGT capacity can be written in the following way:

(10) $$Y_{it} = Y_{it}^{N} + \alpha_{it} D_{it}$$

[36] This explanation is meant to apply the synthetic control method to the specifics of electricity markets. Those interested in a more robust statistical explanation should consult Abadie et al (2010).

Y_{it} is the change in CCGT capacity for restructured states in each year, Y_{it}^N is the counterfactual change in CCGT capacity for each state and year, and a_{it} is the impact of restructuring on investment in CCGT capacity. Calculating a, which is the primary goal of this chapter, requires knowing Y_{it}-Y_{it}^N. Y_{it} is known, leaving only Y_{it}^N to be estimated. Estimating Y_{it}^N requires three inputs: 1) Important factors which influence changes in CCGT capacity; 2) The relative importance of each of those factors; 3) The weighting of each state in the counterfactual.

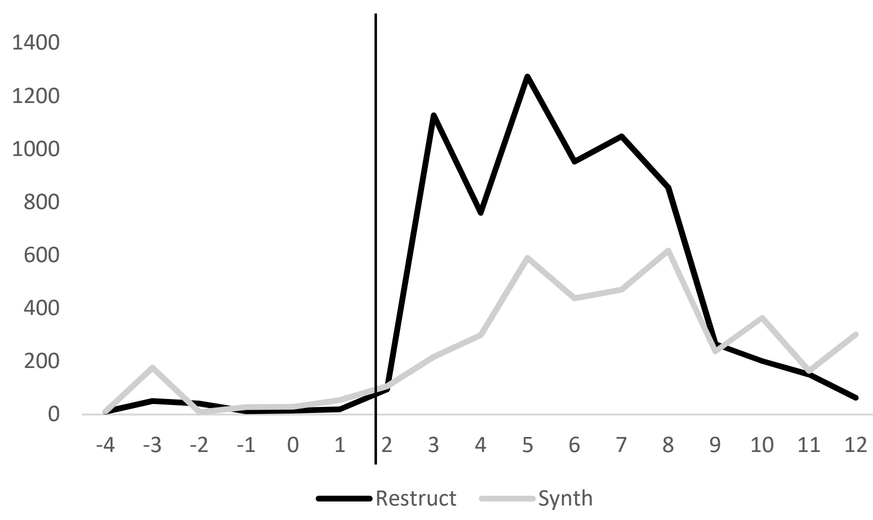
The important factors in changes in CCGT capacity, X, are used to match restructured states with the states which they most closely resemble prior to restructuring. These factors are selected both from insights in the industrial literature and from the model presented in section 3 and include ERS projections, capacity to ERS ratio, variance ratio and capacity mix ratios for relevant fuels like coal, gas, nuclear and hydro. Once collected, the importance of each of these factors, V, is estimated by minimizing the mean square prediction error of the change in CCGT capacity prior to restructuring. V is estimated in order to not impose the restrictive assumption that all factors equally affect the outcome variable. Finally, the weights of each state in the counterfactual are solved for by finding W^* that minimizes $(X_1-WX_0)'V(X_1-WX_0)$, where X_1 is the set of factors for the restructured states and X_0 the set of factors for the non-restructured states. As with the estimation of V, this step ensures the process does not rely on the assumption that all the chosen counterfactual states are of equal importance. Once the weights are calculated, a counterfactual state is formed for each restructured state and a_i is measured.

Estimating the synthetic control group requires choosing the correct Y. In previous studies, the outcome variable has been for one state or region. However, one of the strengths of this chapter is the adoption of restructuring by multiple states at different time periods. Therefore, a counterfactual must be constructed for several states, leaving two options: estimate the average of each individual state's synthetic control or the synthetic control of the average restructured state. In this case, these two are not identical, as restructuring occurred at different times for different states. For example, combining California, which restructured in 1995, with Maryland, which restructured in 1999, would provide a skewed synthetic control. Additionally, state electricity markets have differences in capacity choices and demand profiles, making an average state uninformative. Therefore, a synthetic counterfactual is estimated for each state, with the average effect then calculated and reported. The other restructured states were left out of the donor group for each restructured state, since they also received the treatment.

5.1.1 CCGT Synthetic Control

The factors included in estimating CCGT capacity were the same as those used in Section 4. However, in Section 4 the weights of all the non-restructured states were equal in the counterfactual. In this method, states received different weights based on factor proximity. Each restructured state was weighted differently to reflect differences in factor values and then averaged together to provide the results in Figure 13 and Table 6.

Figure 13. Average Restructured and Synthetic CCGT Capacity Change (MW)



Notes: Running variable: years since state restructured electricity market. Source: EIA (2016)

Table 6 compares the synthetic control results with several others in this chapter. The first column shows the amount of CCGT capacity (MW) added in restructured states compared to the synthetic control for the period defined by the variable. The second column show the same calculation for the years outside the period specified. The third column is the difference between the two, or the difference-in-difference. The main specification of Section 4 is shown in the DD result of the variable Boom. This is similar to what was estimated in the OLS specification presented earlier.

Table 6. Restructured and Synthetic Comparison Measures

Time Period	Difference	Other	DD
Period 4	642.8	-6.8	649.6
Period 5	629.6	-39.3	668.9
Period 6	564.2	-55.5	619.7
Period 7	487.7	-60.8	548.5
Boom	473.8	-54.7	528.5

Notes: Period= # of years of the buildout following the three year lag. Boom is 2000-2006. Source: EIA (2016)

The period results are included to accurately measure the restructuring period, as not all states started restructuring in 1997.[37] The first column period 4 result shows the average difference between capacity added by restructured states and capacity added by synthetic counterfactual states for four years after restructuring. Periods 5, 6 and 7 show this result for five, six and seven years after restructuring. The second column shows this difference for the years outside of the specified period. The results suggest, while the boom was largest in the early years following restructuring, it did not begin to decline substantially until seven years after restructuring.

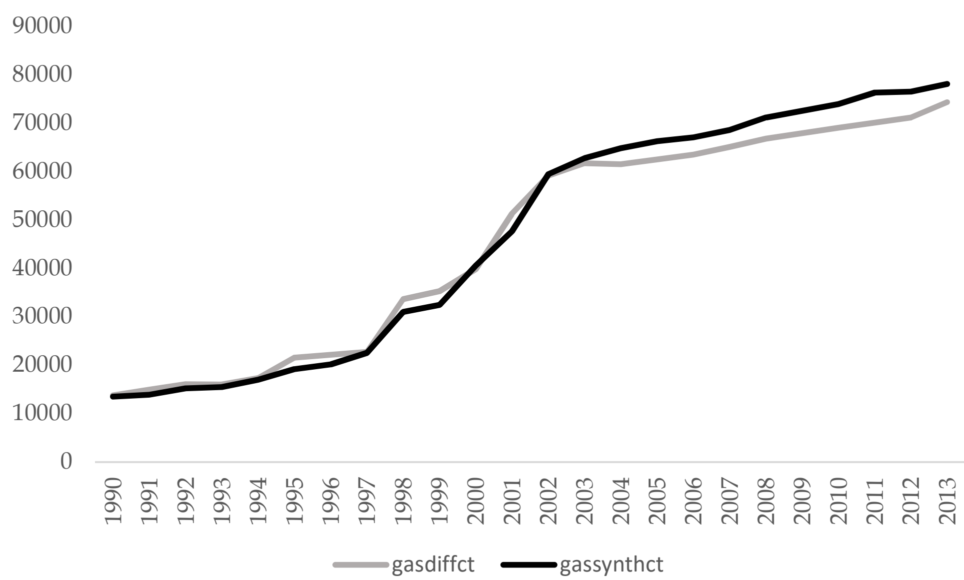
[37] This is consistent with the assumption that CCGT construction takes three years.

These results are consistent with the results of this chapter. First, restructuring led to significantly larger CCGT capacity additions. Second, the additions have the pattern of a boom, with a period of large capacity additions followed by a sharp decline. Third, the boom lasted for a long time, suggesting that the Contagion hypothesis is correct. Finally, the decline in years 8 through 12 can be seen as an attempt by electricity producers to try to balance supply and demand through a pause in capacity additions.

5.1.2 CT Synthetic Control

The factors used in the estimate of the change in CT capacity (Figure 14) in Section 5 are slightly different in this section due to a change in the estimation variable. Due to their decreased investment period and sensitivity to natural gas prices, capacity additions on the annual level are not a useful dependent variable in synthetic control. Total CT capacity is both more meaningful and easier to measure. Estimation includes all the previous variables as well as a measure of previous investment.

Figure 14. Average Restructured and Synthetic Total Gas CT Capacity (MW)



Notes: Running variable- years since state restructured. Source: EIA (2016)

Synthetic control matched restructured and synthetic states well prior to the start of the treatment. The close relationship between the two continues after restructuring, suggesting that restructuring did not change investment in gas CT plants. However, beginning 7 years after restructuring, the difference between the average restructured and synthetic state in Gas CT capacity widens by a small amount. This suggests that restructured states later in the study time period experienced reduced Gas CT investment, relative to what they should have without restructuring. The interpretation of the small gap during this time period is there was an effect in bidding markets of the large number of CCGT plants on CT plant construction.

5.1.3 A New Approach to Synthetic Control

To date, this study represents a new use of synthetic control, both in its use of multiple treatment states and its application to the electricity sector. Prior to this chapter, synthetic control was used for one treated group, like the state of California in Abadie et al 2010. As mentioned previously, this is a complication for multiple treatment groups, as combining the restructured states into one group and estimating the synthetic control is not a robust method of estimation. The strength of the approach taken here is it allows each state to have its own counterfactual. These counterfactuals are estimated using state-specific predictors, rather than a one size fits all approach. As a result, this chapter does not suffer from the estimator problems inherent in time-series analysis.

5.2 Alternative Specifications

The thesis of this chapter is not that restructuring led to a permanent increase in generating capacity, but a temporary surge which declined as time passed. The primary specification structure reflects this dynamic trend between periods. However, a simpler way of identifying this effect, which can be used to analyze changes in specification, is to separate the period 1990-2013 into the boom (2000-2006) and non-boom (all other years).

The interaction between the time period and restructuring allows the impact to be separated by restructuring and time period. In Table 7, the restructuring variable refers to restructured states in years outside of the boom, the boom variable refers to non-restructured states during the boom, and the interaction refers to restructured states during the boom. Each of the following specification are based around this structure.

Table 7. CCGT Buildout Alternative Specifications

	(1)	(2)	(3)	(4)	(5)	(6)
Restruct	332.66***	25.94	16.37	25.59	2.9	-20.88
	(111.96)	(71.98)	(72.55)	(75.33)	(111.08)	(91.42)
Boom		563.24***	200.27	215.76*	325.99***	274.84**
		(191.42)	(116.35)	(118.32)	(104.39)	(131.29)
RestrxBoom		587.59**	577.17**	581.04**	286.26**	551.94**
		(230.80)	(226.42)	(235.00)	(123.64)	(258.26)
CapRatio	-324.37***	-291.67***	-277.58***	-281.69***	-253.93***	-267.46***
	(83.68)	(82.96)	(79.74)	(81.29)	(123.64)	(87.87)
ERSproject	15.3**		13.96**	12.3**	15.93**	14.22**
	(6.55)		(5.95)	(5.67)	(6.80)	(6.39)
Varratio	21.87***	13.85**	21.5***	22.14***	23.27***	24.79***
	(7.61)	(5.55)	(7.03)	(7.23)	(7.40)	(8.06)
gascoalratio	-6.44		-4.33	-10.85	-11.88	-13.83
	(14.54)		(13.38)	(11.27)	(13.05)	(22.29)
FE	Y	Y	Y	Y	Y	Y
Obs	1050	1050	1050	1029	1050	1050
R^2	0.14	0.16	0.17	0.16	0.15	0.21

Dependent variable is change in CCGT capacity (MW) except in column 6, which is change in total gas capacity (MW). *** significant at .01 level. ** significant at .05 level. * significant at .1 level.

Column 1 shows the results of this study if the interaction with the boom is not included. This is measuring the impact of restructuring on CCGT investment across the entire time period. Not surprisingly, this is large in both magnitude and significance. However, this specification does not provide insight into investment path. Column 2 adds the interaction term, showing the significant impact of restructuring on CCGT investment during the boom period. However, demand projections and cost ratios are excluded from the analysis to show their importance in explaining the boom variable, which is the amount of CCGT capacity added in non-restructuring states during the boom. This effect is large in magnitude and significant, suggesting the model is not sufficient in explaining the increase in CCGT capacity in all of the states. Column 3 includes the missing two explanatory variables to show their effect on the boom variable, which is no longer significant. This is evidence that changing demand forecasts and the declining levelized cost of gas were significant in CCGT expansion outside of restructured states. Column 4 eliminates California from the sample to ensure that the largest state in the country is not significantly impacting the results. There is no evidence of any change from excluding California

Column 5 replaces the restructuring index used in this study with one from Fabrizio et al (2007). The two indices are similar, as described in Section 4 of this chapter, with only minor differences in the timing of restructuring for a few states. The table shows a restructuring effect which, while still significant, is much reduced and the boom variable has increased in magnitude and significance. This is not surprising, as the Fabrizio index showed delayed restructuring in several states, which would now be included in the boom group. Section 4 provides an explanation for the preference of the restructuring index used in this chapter over that used in Fabrizio et al.

Column 6 replaces the standard dependent variable, which is the change in CCGT capacity, with the change in total gas capacity. This includes any changes to CT and gas steam capacity. This specification tests two predictions of this chapter. First, if a boom occurred, it would be in CCGT plants, as their increased efficiency would be preferred by plant owners competing to provide intermediate and baseload power in restructured markets. If this is true, the impact of restructuring on the change in total gas capacity should be less significant than the impact on CCGT capacity. Second, this chapter predicts that restructuring leads to overbuilding of CCGT plants. This prediction should show a large impact on total gas capacity that overrides the investment in other gas technologies.

The results in column 6 confirm both of these predictions. The impact is still large and significant, but not as much as the impact on CCGT capacity. Additionally, the failure of restructuring to explain the increase in CT investment is apparent in this specification, as the boom variable is large and significant.

6. Conclusion

The consensus of the economics literature prior to restructuring is the US electricity industry would emerge more efficient, both in the short run and long run. Several papers have shown that the short-run cost impact of restructuring was positive, as it increased production and reduced O&M and fuel costs. Up to this point, however, there had been no study of the long-run cost implications of restructuring. It has now been 20 years since the first states began restructuring, which allows this chapter to analyze whether the long-run efficiency gains, through more effective plant investment, were present. The conclusion of this chapter is that, rather than increasing the efficiency of plant investment, restructuring caused a power plant construction boom that left states with bankrupt electricity firms and stranded power plants.

The natural gas power plant construction boom transformed the US electricity industry, which is in the process of switching from generation primarily from coal power plants to generation from cleaner-burning gas-fired plants. This switch, largely due to the low price of natural gas, would not have been possible without the investment boom, which left a number of stranded power plants that are still not approaching their efficient level of use. Explanations for this boom included changes in the levelized cost of gas-fired capacity, natural gas market deregulation, increasing electricity demand, environmental considerations, and the quick build times of gas plants. The timing and geographic nature of the gas boom suggested that these explanations were insufficient in explaining why so many gas plants were constructed during this time period.

This chapter shows that electricity market restructuring is an integral part of explaining the gas boom. While changes in technology and long-term price forecasts made natural gas more cost effective, compared to coal, in the 1990s, these changes occurred almost a decade before the boom began and do not explain why there was significantly more investment in restructured states. Compared to the non-boom period, restructured states built approximately 593 MW more CCGT power plants on average annually than non-restructured states, costing more than $95 billion in total. The boom in restructured states left them with excess capacity, which is still evident today.

The gas boom provides an informative lesson for policymakers as they consider restructuring other regulated markets. The success of markets in solving the coordination problem is heavily dependent on market information and economies of scale, each of which played a role in the inefficient transition to less-regulated electricity markets. While there is no reason to suspect that the inefficiency in investment by restructured markets is permanent, this chapter has shown the transition, when not handled correctly, can impose significant costs. Policymakers were concerned with this transition, but focused entirely on insufficient investment and changes in retail prices for consumers. It is possible to imagine ISOs imposing an upper bound on regional capacity, in addition to the lower bound they currently impose. The existence of a capacity upper bound during the transition would have prevented the excessive investment that still exists today.

CHAPTER 2

THE GAS BOOM AND DEREGULATION: THEORY AND WELFARE IMPACT

The Gas Boom in the U.S. electricity industry from 1999-2006 was strongly correlated with state choices to restructure their electricity markets. This relationship, supported by events in the industry during this time period, suggests that entry was excessive and investment non-optimal in these markets. Inefficient investment contradicts institutional assumptions about the effect of competition on firm investment behavior. This chapter analyzes the event in the context of four entry and exit models in the industrial organization literature: Business Stealing, War of Attrition, Contagion, and First Mover, finding the events of the period most consistent with the Contagion Model, coordination failure, and falling long-term interest rates. The $10.8 billion welfare loss from excessive entry was borne mostly by producers. (JEL L51, L94, L22, Q40)

1. Introduction

Beginning in 1999, the electricity industry underwent a transformative event dubbed "The Gas Boom." From 1999-2006, utilities and independent power producers invested over $240 billion in the US electricity industry, increasing capacity by over 250 gigawatts (GW). Of the 250 GW, as much as 90 GW was built excessively in states with restructured electricity markets, as shown in the first chapter of this dissertation. 21 years after the first state market was restructured, the legacy of the investment boom is still apparent, with most restructured states possessing larger than normal excess reserves. This chapter explains why restructuring led to over-entry and quantifies the welfare loss associated with the excessive investment.

Neoclassical investment models assume that firms construct capacity to maximize long-run economic profits. In the case of the electricity industry, a firm constructs a plant based on its projections of relevant factors like market demand and fuel prices. If the net present value of the plant is positive, a firm will build. The starting point for the analysis of the Gas Boom is the assumption that the plants were built based on profit maximizing decision-making. In this analysis, models are selected based on their characteristic match of the electricity industry. They also have the potential for non-optimal entry as a result of the findings in chapter 1. Two of the models in this chapter, Business-stealing (Mankiw and Whinston 1986) and First Mover (Spence 1977; Schmalensee 1981) assume firms operate under this assumption and their over-entry, or under-entry, can be seen as strategic. A third, War-of-attrition (Bulow and Klemperer 1999; Cabral 2004) also assumes firms are profit maximizers but suffer from a coordination failure which can lead to non-optimal entry. A fourth model, Contagion (Bikchandani et al 1992; Bannerjee 1992), differs from the other three in assuming firms act based on fads and herding behavior, rather than fundamentals. Contagion is closely associated with manager overconfidence, which assumes firm managers are acting according to a fundamental analysis, but all believe they are the low-cost producer in the area.

This chapter contributes to the literature in three different ways. First, it provides a theoretical underpinning for the Gas Boom and, in general, investment dynamics in the electricity industry. Second, it provides a prominent example of herding models in a large industry. Third, it quantifies the cost to the industry of the excessive entry which occurred following the opening of markets to competition. While it does not go so far as to say restructuring reduced total welfare, it provides an argument for the inclusion of entry effects on investment in any welfare analysis

The four models in this chapter were evaluated on whether the firm behavior they predicted was consistent with what was observed in the industry. In general, the failure of a large number of firms in 2005 and 2006, along with low plant usage and large-scale entry, contradicts any model which relies on firm profit maximization alone or coordination failure. When evaluated further, only the predictions of Contagion and manager overconfidence were consistent with the facts of the period. The Business-stealing model is static and shows excessive entry as an equilibrium outcome, both of which are inconsistent with the electricity industry in the time period in question. The First-mover model is dynamic and does not have excessive entry as an equilibrium outcome, but predicts the largest investments would have been at the start of restructuring, which is incorrect. The War-of-attrition model is consistent with the coordination failure present in electricity investment, but also predicts an equilibrium outcome not present during the period. The dynamic Contagion model does not predict equilibrium excessive entry, fits the investment time path and is consistent with industry publications during this period discussing imperfect information. Managerial overconfidence enhanced the contagion effect through the difference between the actual distribution of low-cost producers compared to what managers assumed.

This chapter concludes that there were three factors responsible for excessive entry in restructured electricity markets. First, imperfect information led firms to stray from a fundamental analysis and engaged in herd investment behavior, while also falling victim to managerial overconfidence. Second, due to the removal of utilities as planning authorities, firms suffered from coordination failure. Plants could be built simultaneously without knowledge of each other and which market they were intended to serve. Third, the decline of long-term interest rates, which began in the mid-1980s, pushed investors to search for yields. With the electricity industry being well-established, previously closed to almost all new entry and possessing high prices, investors took advantage of low borrowing costs and flooded in.

Three welfare effects are identified in this paper from excessive entry. The first is the impact on electricity prices paid by consumers from excess capacity, the second is the loss consumers bear from malinvestment and the third is the loss for producers. The literature finds no long-term effect on electricity prices from restructuring, eliminating the first effect from the analysis. The second effect is considered in the economics literature to be trivial and is also excluded. The third effect is large, totaling $10.8 billion.

The structure of this chapter is organized in the following manner. Section 2 introduces the models used in this analysis. Section 3 tests the predictions of these models. Section 4 sets up the welfare model and calculates the total welfare effect of excessive entry. Section 5 concludes.

2. Competition and Excessive Entry

When states restructured retail electricity markets, the transfer of transmission and distribution assets to non-profit ISOs and the forced divestiture of plants by vertically integrated utilities opened state electricity markets to generation competition. The combination of low state capacity ratios and the profits of merchant gas-fired power plants operating in the wholesale market encouraged entry into the newly competitive markets. Reserve margin requirements, typically set by ISOs around 15 percent of ERS (Joskow, 2008), reduced concerns of inadequate entry.[38] However, no consideration was given to the prospect of excessive entry. This section provides a framework for how electricity market restructuring could lead to excessive power plant investment.

Prior to introducing the framework, a clear definition of excessive entry in the electricity industry is necessary. Mankiw and Whinston (1986) defines excessive entry as an outcome where the equilibrium number of firms exceeds the socially optimal number. Cabral (2004) has a slightly different approach, stating that, if there is an entry tax that strictly increases social welfare, an industry has experienced excess entry. For the electricity industry, the corollary to the fixed cost of entry in the previous two papers is the capacity of the power plant a firm must construct in order to compete in the electricity

[38] There was concern, as markets began to restructure, of an underinvestment in peak resources, due to the difficulty of earning sufficient margins to cover fixed costs and the lack of real time pricing (Joskow, 2008).

market. Therefore, the number of firms is not as significant as the amount of capacity (MW) invested in each market. This chapter uses the difference between the actual CCGT capacity added compared to the synthetic control counterfactual as a measure of excessive entry. An additional statistic, the ratio of capacity to ERS (capacity ratio), is also useful for measuring industry entry and providing a historical comparison.

2.1 Four Models of Entry

Each of the four models in this section: business stealing, war-of-attrition, contagion, and first-mover advantage, analyze homogenous good market structures where excessive entry is a possible outcome.[39] The purpose of this section is not to prove outcomes theoretically, but rather to provide testable predictions to see which of these models is most consistent with entry and exit patterns following restructuring. The following assumptions are critical to the predictions that originate from the models.

Assumption 1: There are many entrants into each market and all sell a homogenous good.

Electricity is a homogenous good generated and sold by many types of firms, which include utilities, power marketers, government entities, and industrial firms.[40] Therefore, a large number of firms were able to enter the electricity market.

Assumption 2: Firms entering the market must construct a new CCGT plant.

While not all firms entering restructured markets built new power plants, as firms could purchase the divested assets of former vertically-integrated utilities, the purpose here is to explain the new construction of power plants, as opposed to the number of entering firms.[41] For those firms that chose to construct new plants, the overwhelming

[39] A large number of models were considered in this section, with selections based on the following criteria, which matched electricity markets in this time period: 1) Homogenous good. 2) Inclusion of economies of scale. 3) Imperfect information leading to the possibility of coordination failure.
[40] Information on participants in the generation and retail components of the electricity market is available from EIA form 861 and EIA form 860 respectively.
[41] See Ishii and Yan (2007) for an explanation of the build or buy decision of entrants.

choice for base-load market competitors was CCGT plants. While in a general build choice model, like those of Joskow and Mishkin (1977), firms choose the technology and fuel of the plant, the combination of reduced gas prices, short build times, and improved efficiency led new builders to CCGT.

Assumption 3: Incumbents have a different cost structure than entrants, with higher original investment costs and lower marginal costs.

The share of CCGT plants in restructured states prior to restructuring was very small. The majority of plants in operation during this time were nuclear, coal, and hydro. Each of these plant types have higher capital costs and lower fuel costs than CCGT plants.

Assumption 4: Entry suffers from a coordination failure.

Coordination failures are heterogeneous, ranging from external costs and benefits to Schelling's (1960) where-to-meet problems to multiple entrants in natural monopoly markets.[42] In the electricity industry, utilities filled the role of social planner, as they were aware of all plants being planned and built as well as what market they were serving. In restructured markets, this information is more difficult for entrants to obtain. Therefore, firms were most likely unaware of the intentions of other entrants upon investment.[43]

2.1.1 Business Stealing Entry

Consider a model of simultaneous entry with incomplete information.[44] The actors in this model consist of an incumbent firm with marginal cost c_1 and n identical firms with marginal cost c_2, with $c_1<c_2$. In order to compete in the market, entrants must invest a fixed amount (x_2), which is immediately sunk. The incumbent has already invested x_1,

---

[42] For more on coordination failures in the theoretical literature, see Dixit and Shapiro (1986), and Bolton and Farrell (1990). Cabral (2004) provides several industry examples.

[43] This assumption is consistent with Kydland and Prescott (1982), who state businesses may not be aware of entrants until the selling of goods commences.

[44] This model is most closely linked to Mankiw and Whinston (1986), but is similar in structure to a number of papers, such as Spence (1976), Dixit and Stiglitz (1977), Weizsacker (1980), and Perry (1984).

with $x_1>x_2$, and owes a portion of it (l). Prior to the start of the game, a regulator has fixed prices at a level where the variable profit of the incumbent is greater than x_1. Firms possess full information on past market prices and the cost structure of the incumbent, but are not aware of entry by other firms. Demand is inelastic.

The entry process is modeled as a two-stage game. In stage 1, the market for electricity is restructured, allowing for entry. This encourages m ($m \leq n$) firms to enter the market, each investing x. In stage 2, the m entrants and incumbent produce electricity as Cournot competitors. Variable profit for each firm is a function of the number of firms that enter the market and the marginal cost of the firm [$\pi = f(m,c_i)$].

Prediction 1. Under business-stealing entry,

a) *Output per firm falls as the number of firms increases (business stealing).*
b) *Entry reaches an equilibrium where variable profit equals fixed cost.*
c) *There is idle capacity in the industry following entry.*

Following the analysis in Mankiw and Whinston (1986) and the structure presented above, the business-stealing model, applied to the electricity industry, acts as follows. As a state begins the process of restructuring, firms enter the market due to their knowledge of the incumbent's cost structure and previous prices in the market.[45] These firms invest in a power plant with a fixed capacity and it is assumed the cost of investment is sunk.[46] Entering firms are assumed to build gas-fired power plants, which have a different cost structure than the incumbent.[47] As more firms enter, inelastic

[45] Largely this reflects expectations that regulators will fix prices high for a period to allow for utilities with stranded assets to recover their value (Borenstein and Bushnell, 2015).
[46] The firm can sell the plant, but may not be able to get market value for it. This scenario would occur if the firm is attempting to sell the plant during a fire sale, when assets are discounted. Evidence from a string of bankruptcies in 2005 and 2006 suggests that selling off generation assets could not save giants like Calpine (Anderson and Erman 2005).
[47] Marginal costs are higher for entrants than for incumbents due to the price differential between natural gas and coal or nuclear. Investment costs are lower, as large-scale steam plants are more expensive to build per MW of capacity than gas turbine and CCGT plants.

demand for electricity results in output and price per firm falling. This is the first prediction that will be tested in this chapter. Falling output and price reduces variable profit until it equals the cost of investment. At this point, firms stop entering the industry and an equilibrium is reached. This is the second prediction tested in this chapter. As shown in Mankiw and Whinston (1986), more firms have invested in investment costs than were necessary to serve the market. This inefficiency increases as the size of the investment cost increases and leaves firm capacity unused. This is a sign of excessive entry and is the third prediction to test.

2.1.2 War of Attrition Entry

Consider a model of simultaneous entry with incomplete information where investment takes multiple periods.[48] There is an incumbent earning positive profit (π_1) and n potential entrants, some ($m \leq n$) of which start investing a portion of the entry cost (x) in period t. The decision to invest in the market is based on the expected discounted profit of entering the market (π^e), which depends on firm assumptions about three factors: the path of wholesale electricity prices, number of firms entering the market and future natural gas prices. Firms are assumed to possess publicly available information for these factors and base their expectation on past prices and the number of firms observed investing in the market. Once an amount ($\frac{x}{3}$) is spent, it is considered sunk.[49]

A firm will continue investing in each period as long as the expected discounted profit of their project is positive. During the investment period, firms have complete information on the history of firm entry and pricing, but are unaware of the number of firms which invested in the market that period. After three periods of investment, they begin a Cournot competition game with the incumbent and other entrants. Variable profit

[48] This model is drawn primarily from Cabral (2004). Bulow and Klemperer (1999) summarizes the use of these models in the economics literature.
[49] Utility and EIA data suggest approximately 10 percent is spent in the first year, 50 percent in the second, and 40 percent in the third.

for each firm is a function of the number of firms that enter the market, the price of electricity, and the marginal cost of the firm [$\pi = f(m,p,c_i)$].

Prediction 2: Under war-of-attrition entry,

a) *As the market adjusts to capacity additions, new investment stops but capacity increases, as firms finish their investment.*
b) *Firms that completed their investment earn variable profit* $\geq x$
c) *Entry leads to excessive spending on investment, as firms invest simultaneously, unaware of other entry*

The structure and play of the game follow from Cabral (2004). The first stage consists of firms, following restructuring, investing in power plants. Construction of CCGT plants takes approximately three years, at which time firms are able to begin producing electricity. When making the investment, it is assumed that firms are unaware of potential competitors investing in the market in the same period. This follows from Assumption 4 of this chapter. In the following period, firms are able to observe new entrants that made an investment the prior year and update their expected profit function. If still positive, the firms make an investment in the following year and repeat the process. If the firm makes three investments, it then produces electricity in the following period and competes with the incumbent and other firms in a Cournot game.

Three testable predictions come out of this entry model. Firms in this model rely on their expectation of prices and entry when deciding whether or not to continue investing. As a result, a change in firm investment behavior should be visible when market expectations change, as they did in 2005. This is the first testable prediction. If firms complete the investment process and begin competing, it is assumed that firms unable to earn a profit based on market entry and prices will have left the market. Therefore, the remaining firms in the market should earn non-negative profits. This is the second testable prediction. However, there are still a large number of firms that entered and invested in power plants beyond the socially optimal number. This would be apparent in a large amount of capacity built beyond what is socially optimal and is the third testable prediction.

2.1.3 Contagion Entry and Exit

Consider a sequential entry model with incomplete information where firms make decisions based on the actions of other firms.[50] The players in this model are an incumbent and n potential entrants. The incumbent still owes a portion of its investment in the industry (l) and has marginal cost c_1. Potential entrants have the same marginal cost (c_2). Entrants know only their own cost structure. In order to compete in the market, the firm must make a one-time investment (x). Entry is based on expected future profits (π^e) and a surprise change in the number of firms (N^S_{t-1}).

11) $x_t = f(\pi^e, N^S_{t-1})$

$\pi^e = f(p_{t-1}, p_{t-2}, \ldots p_{t-T}, g_{t-1}, g_{t-2}, \ldots g_{t-T})$

$N^O_{t-1} - N^P_{t-1} = N^S_{t-1}$

Expected profits are a function of the firm's expectations about future electricity (p_t) and fuel (g_t) prices, which the firm estimates based on prices in previous periods. Firms are unaware of other entrants in the industry due to the coordination failure described in Assumption 4, so they look for other trends in the industry to help guide their expectations. They calculate a number of firms they believe should be operating in the industry based on their expectation of variable profits. The difference between the

[50] This model is based off what is used in Geroski and Mazzucato (2001) and Cabral (2004). Seminal papers in this literature include Bikchandani et al. (1992) and Bannerjee (1992). See Geroski and Mazzucato (2001) for a more complete review of the literature.

number of firms observed (N^{O}_{t-1}) and the number predicted (N^{P}_{t-1}) is the number of surprise firms in the market (N^{S}_{t-1}). This is either a response to uncertainty around future electricity and fuel prices or herd behavior and is called the contagion effect.

Firm decisions are made in the following sequence. In period 0, the incumbent sells electricity at a price greater than marginal cost and the regulator signals that, in the following period, the market will be open to competition. Out of n potential entrants, a portion, *m*, invest in period 0. In periods 1 through *T*, firms compete with the incumbent in a Cournot competition. New entrants make an investment to enter in the following period and, due to competition and new entry, prices fall. Falling prices or rising costs lead to negative profits for some firms, which exit the industry.

Prediction 3: Under contagion entry and exit,

a) *Firms continue to invest in markets where large investments have already been made*
b) *Exit occurs by firms previously producing in the market if variable profits fall significantly*
c) *Non-fundamentals decision-making leads to excessive entry*

At time 0, the incumbent is still operating under fixed pricing and earning a return based on cost-of-service regulation. Firms observe the high prices present in restructuring markets and assume they will continue based on the existence of stranded assets, the low capacity ratios in many restructured states, and rules put in place fixing prices to allow stranded assets to recover their value.[51] This encourages the entry of a large number of firms in period 1, which invest in CCGT plants and compete with the incumbent. In the following period, more firms enter as prices have remained high and, due to the coordination failure present in simultaneous entry situations, firms see a surprise number of firms entering the market, which encourages further entry. Firms investing in a market where large investments in CCGT plants have already taken place

[51] See Joskow (2008) for more on some of the strategies to save stranded assets.

is the first testable prediction. As CCGT capacity rises, firms compete with each other and price falls in restructured markets. Falling prices lead some firms to leave the industry through an asset fire-sale or declaring bankruptcy. This is the second testable prediction. As firms observe falling prices and other firms leaving the industry, investment declines. At this point, the industry is left with unused capacity due to excessive entry, which is the third testable prediction.

2.1.4 First-Mover Advantage

Consider a model of sequential entry with imperfect information where first-mover advantage affects investment.[52] There is a large incumbent firm with marginal cost c_1 that sells electricity at a price fixed by a regulator. The incumbent has a portion of its fixed cost that have yet to be paid off (lx). There are n potential entrants in the industry, all with the same cost structure and information (c_2). Firms are able to see market prices and investment in prior years before making an investment in each period. In order to compete in the market in period t, each firm must make an investment in period $t-1$. These investments, x_{it}, are individual to each firm, have increasing returns to scale in the production of electricity [$f''(x)>0$], and are a function of prices in previous periods and investment in the previous year [$x_{it} = f(p_{t-1}, p_{t-2}, \ldots, p_{t-T}, x_{t-1})$ where $f'(x_{t-1})<0$]. All investments are assumed to be completed before the start of the next period. Actual production of electricity (Z) is a function of variable profits and the size of investment made by the firm. [$Z=f(\pi_t, x_t)$].

In period 0, the incumbent sells electricity at a price greater than marginal cost. The regulator signals that, in the following period, the market will be open to competition. Out of n potential entrants, a portion, m, invest in period 0.[53] In periods 1

[52] See Spence (1977), Dixit (1979), Schmalensee (1981) and Hilke (1984) for prominent theoretical papers in this area. Berger and Dick (2007) provides a good review of both the theory and empirical examples in this literature.
[53] As noted in Gilbert and Vives (1986), some firms are quicker to act to new markets than others. This could be due to private information, low borrowing costs, or a difference in CEO risk assessment.

through N, controls on prices are lifted and the new entrants compete with the incumbent in a Cournot competition. A number of potential entrants ($< n\text{-}m$) invest each turn based on observed prices and the amount of the investment in the previous period and a number of participants ($\leq n$) leave the industry if variable profits are less than $\frac{x_{it}}{N}$.[54]

Prediction 4: Under a first-mover advantage framework,

a) *The first firms to invest in the new regulatory environment make large investments.*
b) *Investment from entrants exceeds the optimal amount due to excessive entry.*
c) *Entry is high after the introduction of competition, then drops swiftly.*

Entry in this model is ignited by the opening of the market to competition and the existence of variable profits in excess of investment costs.[55] In period 0, entering firms are encouraged to build large plants due to the combination of increasing returns to scale and the deterring effect of previous investment.[56] The existence of these large investments is the first testable prediction. Firms that build large gas-fired power plants in a given electricity market send a signal to future entrants that entry into this market may not be profitable, given the presence of a large power plant supplying electricity. Throughout the literature, the advantage is more significant if the firm is the first-mover and the investment large. The combination of increasing returns to scale and the opportunity to gain market share[57] by deterring future entry incentives firms to overinvest, leaving the market with capacity that outstretches demand. This is the second testable prediction.

In the following period, prices fall due to increased competition, driving variable profits below $\frac{x_{it}}{T}$. This leads to firms leaving the industry, as variable profit is not high

[54] It is assumed that firms can't continue to not pay off their investment, which is a fixed annual amount.
[55] Merchant gas-fired plants had entered the market before restructuring and made large profits selling electricity into the wholesale market. The reason for these profits was the high cost-of-service prices enforced by regulators to allow utilities to recoup large capital expenditures on coal and nuclear plants. Firms may have estimated variable profits based on the high prices that existed prior to restructuring.
[56] Spence (1977) notes this strategy relies on homogenous good markets with economies of scale.
[57] The actions of Enron suggest that gaining market share was a factor in the gas boom.

enough to pay creditors and the firms go in to bankruptcy. Additionally, investment falls in the first period in response to the large investments made prior to the first period. The decrease in investment accelerates in the second period, as firms respond to falling prices in addition to the large investments previously made. The quick drop in prices and investment following the first period investments is the third testable prediction.

2.2 Testing Predictions

Each of the predictions of these four models are able to be tested by data available from 1990-2013. Finding the model consistent with the events of this time period is important, as it serves to provide an explanation for how excessive entry can occur in homogenous-good industries with large fixed costs.

3. Entry Model Analysis

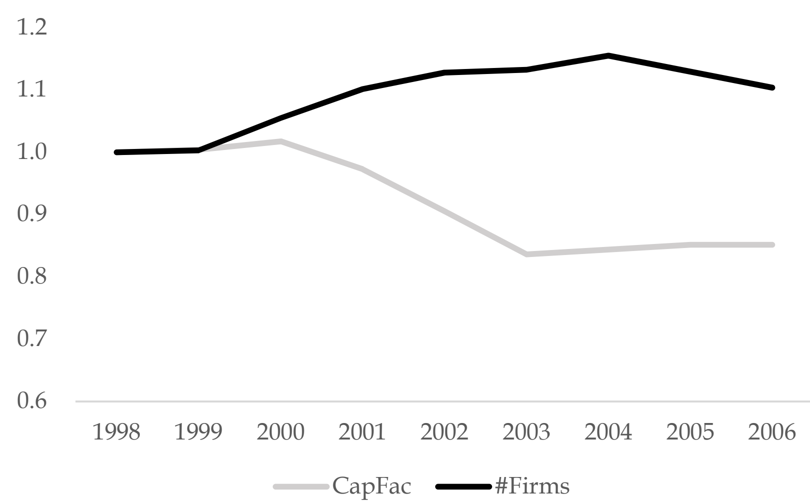
Chapter 1 presented significant evidence that restructuring led to overinvestment in CCGT capacity. However, this result is at odds with many economists' conceptions about firm decision-making and operations. For example, in the First Welfare Theorem, free entry is a prominent condition. Additionally, firms are thought to make investments in order to maximize profits. Yet not only did the restructured state capacity ratio exceed those for the two non-restructured groups, but in fact reached a 24 year high. How was it possible for electricity markets with free entry to provide the incentive for so many firms to invest beyond the optimal amount? Section 2 presented four models that provided a structure for analyzing this question. Each of those models generated testable predictions in order to determine which most closely is associated with the events of the electricity industry starting in the late 1990s. This section presents results to test these predictions and identify which model consistently explains the large investment boom in CCGT power plants from 2000-2006.

3.1 Model 1: Business-Stealing

The component of this model which drives its predictions is the existence of profits from imperfect competition, which encourages entry. As shown in Mankiw and Whinston (1986), excessive entry will occur and social welfare will be reduced due to too many firms paying the fixed cost of entry. In the electricity context, the corollary is plants being built and not used at full capacity. If this model is correct in explaining entry into restructured electricity markets, the following must be true: 1) Output per firm falls as the number of firms increases. 2) Entry reaches an equilibrium where variable profit approaches fixed cost. 3) There is idle capacity in the industry following entry.

Figure 15 shows the average capacity factor for all restructured states and the number of electricity producing firms operating each year. The table shows the drop in output per plant starting in 2001 and continuing until 2005. These are the peak years of the boom and the number of firms reflects the level of entry in this period. This result suggests the demand for electricity was insufficient to purchase the electricity capable of being produced by the firms in the industry, confirming a key aspect of business stealing.

Figure 15. Firm Entry and Capacity Factors in Restructured States



Source: EIA Form 860 (2016) and Author’s Calculations

The second prediction has a critical impact on the results. Firms are making the profit-maximizing decision to enter the industry. In order for this model to represent the electricity industry at this time, observed firms must earn a non-negative profit. However, this prediction is contrary to the evidence in this period. The string of high-profile bankruptcies in the electricity industry in the mid-2000s showed that not all firms which invested in this market were making profit. The number of entrants, therefore, was not only above the socially optimal level, but also above the level which could be profitably sustained by the industry.

The third prediction is consistent with the primary finding of this chapter, which is restructured states experienced excessive entry that led to idle capacity. The results in Table 1 clearly show this result. Capacity factors fell to a two decade low, as there was insufficient demand to use the created capacity, and have yet to recover.

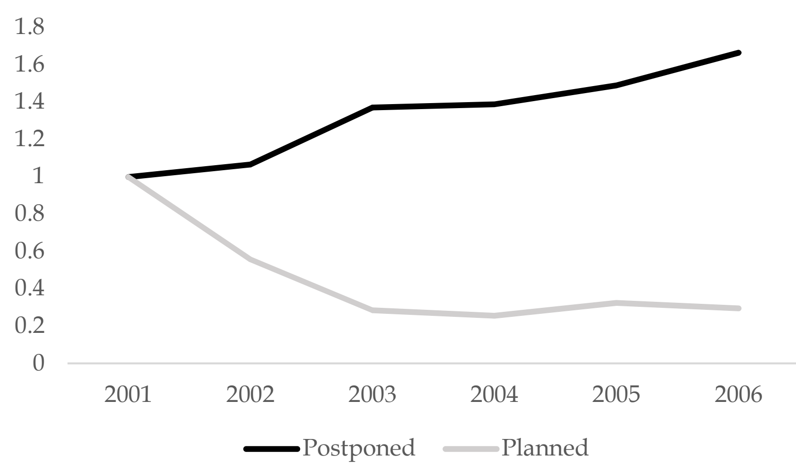
While two of the Business-Stealing model predictions are correct, the third shows the limits of this model in explaining the investment boom. Long-run profits in the industry were not sufficient to support the level of entry in this period. In particular, the Business-Stealing model fails to account for dynamics in entry and exit. The model makes no predictions regarding the path of investment, but rather describes the equilibrium and provides an explanation for how it is reached. In order to explain the timing of entrance and exit, a more robust model is needed.

3.2 Model 2: War-of-Attrition

There are two features of the War-of-Attrition model which drive its results. The first is the nature of entry, which is simultaneous. When firms enter, they do so without knowing who else has entered at that time. The second is the investment path, which requires firms to make new cash outlays for several periods prior to entering the market. These model specifics lead to the following three testable predictions: 1) As the market adjusts to new capacity coming online, new investment stops but ongoing investment will be completed. 2) Firms that have completed their investment will earn non-negative profit. 3) Entry leads to excessive investment spending, as firms invest simultaneously.

The first prediction of this model concerns investment patterns. If the electricity market follows a War-of-Attrition model, market operations lag investment decisions. Firms make decisions in each period whether or not to continue investing if there was a change in market conditions. Given the bankruptcies and market turbulence beginning in 2005, Figure 16 would be expected to show a sharp drop in planned investment and an increase in postponed investment following 2005. However, planned investment fell quickly after a peak in 2001 before stabilizing in 2003. Postponed (or cancelled) investments jumped in two periods, one after 2002 and then again following 2005. While the postponed portion followed the predictions of the model, the planned part differed substantially. Using this data, the model would have predicted the boom was coming to an end by 2003 instead of 2006. Therefore, this prediction contradicts the model.

Figure 16. Investment Planning in Restructured States



Source: EIA (2016) and Author's Calculations

The second and third predictions are similar to those made in the Business-stealing model. Firms that operate in the market should earn non-negative profit and the lumpiness and timing of firm investment guarantees that some investment spending will not be used. As in the Business-Stealing model, the third prediction is supported by this chapter, but the existence of large-scale bankruptcies invalidates the second.

The War-of-Attrition model is a dynamic model of entry and exit, where firms react to investment in prior periods and make a decision on the completion of their current project. It is able to explain how excessive entry could occur, but fails to explain the bankruptcies in the mid-2000s, rapidly declining capacity factors and the length of the boom. If this model is correct, firms would have reacted to declining capacity factors by suspending their projects. Instead, the boom continued through 2006. Similar to the Business-Stealing model, this model fails to account for the length of the boom.

3.3 Model 3: Contagion

The Contagion Entry and Exit model is the first encountered thus far with a sequential entry setup. Firms are able to observe entry in the previous period and update their decision framework. The key component of this model is the awareness of firms that they operate in an incomplete information framework. With the market newly restructured, firms must derive expectations about future prices and entry based on limited history. With uncertain fundamentals, firms are subject to fads and herd behavior. Seeing a surprise number of firms enter the electricity industry, in this case, increases the incentive to invest, as firms believe others may have better information. This model, based on surprise entry and exit, provides three testable predictions: 1) Firms continue to invest in markets where large investments have already been made. 2) Exit occurs by firms previously producing in the market if there are changes in price or cost fundamentals. 3) Outside-of-fundamentals decision-making leads to excessive entry.

One of the unique features of the boom is how long it lasted. Large-scale plant construction started in 1997 and continued until 2006. One of the failures of the previous two models was in explaining the length of the boom. Why, for example, would firms continue to build power plants after seeing other large plants already under construction or in operation? The answer, according to the Contagion model, is that firms saw this as a signal that there were opportunities for profit in the electricity market. As more plants began to go online in the early 2000s, this surprise number of firms entering the industry encouraged investment, as opposed to discouraging it. Therefore, the evidence from the early 2000s is consistent with this portion of the Contagion model.

Of the three models discussed thus far, the Contagion model is the only one to predict exit by loss-making firms. The reasoning is fairly clear. As firms began to rely on fads instead of fundamentals, the market became saturated with electricity producing firms. Each of these firms hoped to take part in what was clearly an industry on the rise. However, if they had stuck to a fundamental analysis, firms would have realized that there was insufficient demand to use all the capacity created. The result was shrinking margins and bankruptcies, which began in earnest during the rise in natural gas prices in the latter half of 2004. These firm failures are consistent with the Contagion model.

The Contagion model predicts more firms will enter than optimal. There is a diverse literature in economics on bubbles, herd behavior, and firm failure[58] which shows how straying from fundamental analysis can lead to more investment than is socially optimal. The findings of this chapter, that investment in restructured markets was more than optimal, are also consistent with the Contagion model.

This model is consistent with what occurred in restructured electricity markets starting in the late 1990s. In particular, it provides an explanation for the most perplexing issue in this chapter, that firms continued to invest even after observing the large power plants either already built or in progress.[59] Any basic fundamental analysis would have concluded that capacity factors would fall significantly, which is what occurred. Lower

[58] See, for example, Barbarino and Jovanovic (2007).

[59] Unlike models of simultaneous entry, the Contagion model does not rely solely on a coordination failure. If the boom were caused only by a coordination failure, it would not have stretched into 2006.

capacity factors would make it difficult for firms to pay the cost of capital construction. However, if firms did not follow a fundamental approach, entry would continue. Given the events in the telecommunications, technology and housing markets in this time period, this should not be surprising.

A contributing factor to this contagion effect is the role of manager overconfidence. If managers believe they are the low-cost producer, increasing supply and falling demand would not deter investment, as managers would be convinced they could compete in the market. If this were the case, excessive entry could occur, while also extending the boom and locating new plants nearby existing generation facilities. The existence of all three of these predictions, combined with insights contributed by industry insiders during the course of the formation of this chapter, lends support to this model of firm behavior.

3.4 Model 4: First-Mover

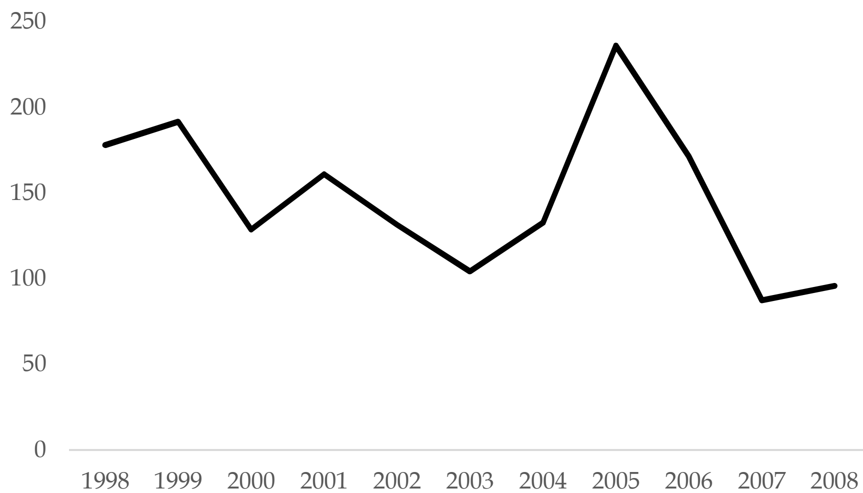
In the electricity sector, power plant investments are often lumpy due to economies of scale. As a result, firms could gain an advantage by entering a market early and make a large upfront investment. In order to prevent entry, a firm would overinvest to convince other firms that entering the market will only incur losses. Once other firms decide not to enter the market, the firm(s) that overinvested early would then have a degree of market power. If this model is consistent with the investment boom period, the following three predictions must be true: 1) Early entrants make large investments. 2) Investments from entrants exceed the optimal amount due to excess entry. 3) Following restructuring, there is a large amount of entry followed by a sharp drop in investment.

Figure 17 shows the average size of non-renewable plants started in each year from 1998-2008.[60] If the first prediction of the model is correct, the first plants in states that were restructuring their electricity markets should have been large to prevent entry. As the figure shows, the plants started in 1998 and 1999 were larger than the average through

[60] Renewable plants were excluded due to size and not being built as an investment deterrent.

the boom, but smaller than those that were started in 2005 and 2006. However, the small difference in average capacity size between the first two years and those that followed was unlikely to convince other firms not to invest. Additionally, this model does not explain the rise in investment size in the final years of the boom.

Figure 17. Average Built Power Plant Capacity in Restructured States (MW)



Source: EIA (2016)

While the first prediction of the first-mover model is not consistent with the nature of the investment boom, the second prediction is. The model predicts that firms will overinvest early, leading to excess capacity in an attempt to prevent future entry. This chapter has shown that excess capacity did develop, although the length of time in which it occurred suggests firms were motivated by factors other than investment deterrence. This is the essence of the third prediction, which forecasts excess capacity developed early in the boom, followed by severely diminished investment. While there were large net capacity increases in CCGT plants early in the boom, there were still significant investments later. This is inconsistent with the third prediction.

The First-mover model does not resemble the nature of the investment path during the boom. Despite the high-profile case of Enron attempting to use its market share in the

California market to exact rents, there is little evidence that plants were constructed to gain market share. As with the first two models, the failure of the First-mover model is its inability to explain why CCGT net capacity grew in the later years of the boom. By this time, any advantage of early entry in the market would have been exhausted.

3.6 Conclusions from Model Predictions

The size and length of the investment boom makes it difficult to explain the US Gas Boom. Had the boom been the same size but much shorter, the explanation would have been simple. Simultaneous entry leads to coordination failure, and once firms discovered the size of the investment, entry would have continued at a smaller, sustainable rate.[61] Economies of scale in power plant construction[62] and the importance of future price and demand expectations in entry decisions make the electricity industry particularly vulnerable to problems of coordination. For decades, vertically integrated public utilities acted as important collectors of private information and expertise, which were not easily duplicated by new entrants. Additionally, as noted in Camerer and Lovallo (2000), excessive entry in the early stages of a new industry can occur due to the overconfidence of inexperienced managers in their own abilities. While a coordination problem could theoretically lead to inefficient entry in either direction, the presence of ISO reserve margins ensured the only inefficient entry would be excessive.

The events that transpired in the electricity industry in the late 1990s and early 2000s support the hypothesis that a coordination problem existed. Industry professionals during the early 2000s noted that IPPs were unaware of the over-build in capacity and slowdown in ERS, which led them to react strongly to previous profit margins.[63] In the

---

[61] This explanation is not inconsistent with entry and exit findings in new product markets (Klepper and Miller 1995).

[62] While gas-fired power plants are more flexible in size than coal or nuclear plants, there is still efficiency in larger sized plants, particularly for CCGT (MIT, 2011).

[63] See Wharton (2006) for an overview of the industry during this time period and the entrance of new firms. The role of regulatory uncertainty was also a factor in an industry that had been predictable.

mid-2000s, large amounts of capacity came online, electricity demand declined and natural gas prices rose, causing profit margins to shrink. This lead to the bankruptcy of several large IPPs, large-scale fire sales of assets by power suppliers, and a significant decline in new power plant investment. In states that maintained regulated markets, these problems were not apparent, as large utilities were better situated to balance changes in demand with new supply coming online.[64]

The length of the boom, however, entails an explanation that goes beyond a coordination problem. Any firm operating under a fundamental approach would have stopped investing once they observed large entry. Instead, investment continued in these markets until a profit squeeze due to a natural gas price spike started a string of bankruptcies. As summarized in Table 6, this matches what the Contagion model predicts. Whether because of lack of faith in available market information or overconfidence in the decision-makers abilities and information (Camerer and Lovallo 1999), firms abandoned a fundamental analysis and invested based on fads. In this case, the fad was taking advantage of newly deregulated electricity markets. This shift in business strategy led to excess capacity in restructured states.

Table 1. Model Prediction Outcomes

	(1)	(2)	(3)
Business-Stealing	✓	×	✓
War-of-Attrition	×	×	✓
Contagion	✓	✓	✓
First-mover	×	✓	×

While the combination of a coordination failure and contagion effect were the most important factors in the gas boom, it is important to note the period in which this took place. The gas boom took place in the US at the end of the Dot Com bubble and

[64] While the experience of the 1980s shows that utilities are not immune from improper demand forecasting, their information on generation and transmission assets allows for a more efficient transition.

throughout the housing bubble. This is not a coincidence. The period 1998-2006 was one in which there was a change in the nature of global financial markets. Prior to 1998, large investments were made in the South American and East Asian economies. However, the experience of the Latin American Debt Crises, the Tequila Crisis in Mexico, the Asian Financial Crisis and the contagion which followed in Latin America, South America and Russia shifted the flow of global credit away from emerging markets and into the developing markets of Eastern Europe and the developed markets of Western Europe and the United States (Eichengreen, 2008). This flow was magnified by increased saving and reduced investment by the previously booming Asian economies, largely in response to perceived exchange rate vulnerability. The effect of these changes lowered real interest rates in the United States and was dubbed the "Asian Savings Glut" by Ben Bernanke (Bernanke 2005). The lowering of global real interest rates led to a flow of cheap credit to investment projects in many developed countries, including the US. At the same time, a combination of deregulation, a booming economy, and the introduction of new technology created a large number of investment opportunities in the United States. These opportunities were reflected in rapidly increasing asset prices in the United States in the late 1990s and early 2000s.

The link between long-run interest rates and investment is well-founded in the economics literature (Mundell 1963). As interest rates decline, more investment projects are profitable and an investment boom occurs. As Caballero et al (2008) notes, there is a well-founded relationship between lowering global interest rates and increased investment in U.S. assets. In the electricity industry, declining expectations about long-run interest rates lowers the cost of capital for building new power plants. This increased profitability encourages more entrants into the electricity industry, in particular those whose projects may be more risky and less profitable in the long run. This leaves the industry open to speculation and risk-taking, with excessive entry being one of the effects. However, if rates begin to rise, firms will begin to fall out of the industry.

The bubbles in telecommunications, technology and housing are well-studied, but the availability of cheap credit and deregulation were also factors in the rapid expansion of investment in the electricity industry (Financial Crisis Inquiry Commission 2011; Mian and Sufi 2014). Nominal 10-year sovereign yields for European and North American countries averaged nine percent in 1995 and four percent in 2003 (CEA, 2015). Industry insiders cite cheap credit as facilitating investment by IPPs (Wharton, 2006). The collapse of Calpine, a large builder and operator of gas-fired power plants, was partly attributed to excessive borrowing when credit was cheap (Tansey 2005).

However, low interest rates, by themselves, do not necessarily lead to excessive entry. The previously cited interest rate has averaged between two and three percent for the past five years without any substantial increase in electricity investment.[65] Utilities in regulated markets also faced similar interest rates in the late 1990s and early 2000s without responding by over-investing in power plants. While low interest rates contributed to the large increase in investment during this time period, they are more likely one of several factors than a primary cause, with restructuring providing an environment for low interest rates to spur investment.

4. Welfare Effects of Excessive Entry

When California and Texas started restructuring their electricity markets, a primary argument in favor of restructuring was the expected improvement in welfare. As outlined in Joskow (2008), cost of service regulation (COSR) reduced the incentive for utilities to be cost efficient, increasing the price of electricity for consumers and reducing consumer welfare. Moving away from COSR incentivized utilities to manage O&M and fuel costs in the short run, while in the long run, utilities were expected to choose the lowest cost form of generation. Joskow notes that the most significant impact is in the

[65] The one exception is in renewable generation, which is largely due to state RPS agreements (Powers and Yin 2010).

long run, as the largest portion of the levelized cost of electricity for a utility has traditionally been plant expenditure.[66] As a result, welfare improvements were expected to be substantial. What was not expected, however, was the over-build that occurred following restructuring. This section seeks to estimate the total welfare impact of the overbuild result found in this chapter.

Analysis of the introduction of competition into markets and its effect on welfare has a long tradition, with Schumpeter (1942) among the first seminal contributions. The term "creative destruction" was coined to describe the process of new products and firms replacing older ones. It was assumed that the gains from innovation and price competition outweighed the losses associated with firm failure and malinvestment, increasing total welfare. However, the electricity industry is substantially different from those dominated by new product formation, with the costs of malinvestment potentially very high and the impact of innovation less clear.

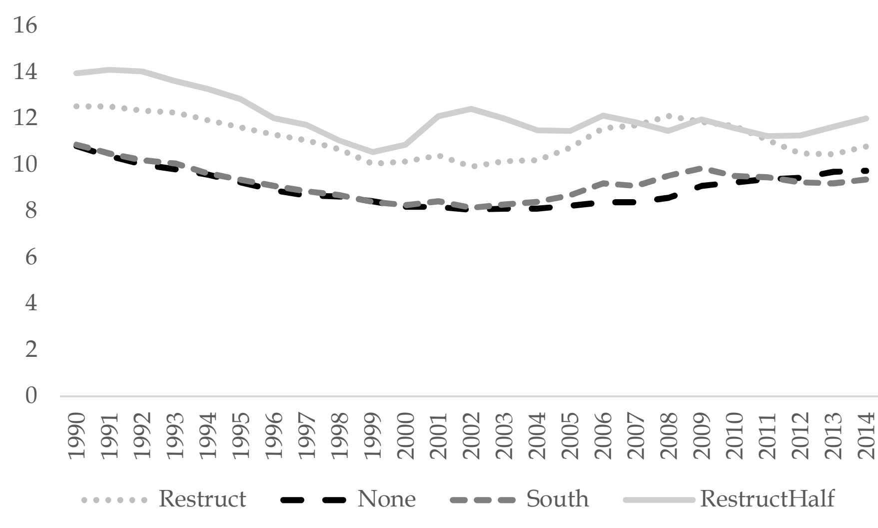
When electricity markets opened to retail competition, the number of firms producing electricity in a region increased for two reasons. First, utilities were required to divest large plant capacity to prevent market power formation. Second, this chapter has shown that the majority of regions in the US were capacity deficient at the start of the restructuring period, which encouraged entry. Part of this capacity deficiency was intentional, as utilities were deleveraging following the nuclear expansion of the 1970s and 1980s and ERS slowdown of the 1980s and early 1990s. However, the increase in ERS growth in the mid-1990s changed expectations about future growth in the industry, increasing the capacity deficiency in many markets and leading to entry.

The rise in the number of firms should have lowered electricity prices, yet evidence of this effect is lacking. A surprising, but well-known result, is there was not a major

[66] While the large-scale adoption of CCGT plants has reduced the capital portion of the cost of electricity, the scale of investment leaves Joskow's point still valid. Additional benefits discussed include insulating electricity generation from politics and increasing retirement of uneconomical plants (Joskow 2008).

change in average US electricity prices following restructuring (Borenstein and Bushnell 2016). Figure 17 shows trends in average electricity prices for four groups: states with restructured electricity markets as of 2015 (Restruct), states that started restructuring but since returned to regulation (RestructHalf), states in the Southeast that never restructured (South), and states not in the Southeast that never restructured (None). As illustrated, the Restruct groups entered the 1990s with higher electricity prices than the other two groups. However, after restructuring occurred in the late 1990s and early 2000s, there does not appear to be any convergence among the groups, with each group declining in the 1990s, due to low fuel prices, and rising in the 2000s, due to increasing natural gas prices. The lack of convergence in these groups indicates the benefits of restructuring were secondary to the importance of fuel availability and capacity choice differences between the groups.

Figure 18. Electricity Prices by State Group



Source: EIA (2016)

4.1 Welfare Problem

There are two markets, Restructured (*R*) and Synthetic (*S*), in which firms can choose to operate. Within these markets, there are two goods sold, electricity (x_1) and an alternative good (x_2). In order to produce either good, a minimum investment (X_1, X_2) is required. Any amount, (Y_1, Y_2), invested beyond the minimum requirement is assumed to go unused. Total investment by each firm, in each period, therefore, is $Q=X+Y$. It is assumed that firms are able to invest and operate within the same period. Demand for electricity in each region is represented by $D(x_1)$ and demand for the alternative good represented by $D(x_2)$, with $D'(x)>0$ and $D''(x)<0$. Each firm charges a uniform price for each good (p_1, p_2) and earns profit (π_1, π_2) from each good. Total welfare for each market can then be written as follows:

$$(12) \quad W^R = \int_0^{x_1^R} D^R(x_1) - p^R + \Pi_1^R(x_1)dx_1 + \int_0^{x_2^R} D^R(x_2) - p^R + \Pi_2^R(x_2)dx_2$$

$$W^S = \int_0^{x_1^S} D^S(x_1) - p^S + \Pi_1^S(x_1)dx_1 + \int_0^{x_2^S} D^S(x_2) - p^S + \Pi_2^S(x_2)dx_2$$

where $\Pi = \sum_{i=1}^{n} \pi_{it}$ for each good and market. Individual firm profit for each market can be written as follows:

$$(13) \quad \pi_i^R = p_{1i}^R x_{1i}^R - \sum_{a=1}^{2} c_{1ai}^R x_{1ai}^R - d_{1ai}^R X_{1ai}^R - d_{1ai}^R Y_{1ai}^R + p_{2i}^R x_{2i}^R - c_{2i}^R x_{2i}^R - d_{2i}^R X_{2i}^R - d_{2i}^R Y_{2i}^R$$

$$\pi_{it}^S = p_{1i}^S x_{1i}^S - \sum_{a=1}^{2} c_{1ai}^S x_{1ai}^S - d_{1ai}^S X_{1ai}^S - d_{1ai}^S Y_{1ai}^S + p_{2i}^S x_{2i}^S - c_{2i}^S x_{2i}^S - d_{2i}^S X_{2i}^S - d_{2i}^S Y_{2i}^S$$

i denotes individual firms, with i = 1,…..,n, t denotes the time period, which in this case is years t = 1,….,T, c the cost of producing a unit of x, and d is the cost of building a unit of X. a denotes which technology is being used to produce electricity (1=CCGT, 2=CT). Prior to reducing the above equations into a welfare function to estimate, the following assumptions are necessary:

Assumption 1: $D^R(x) = D^S(x) = D(x)$

Restructuring does not change the value of electricity and the alternative good.

Assumption 2: Y^S*=0*

Firms in synthetic states do not overbuild capacity in the electricity market. Said differently, there may be overbuilding in regulated markets, but that level serves as a baseline with which to compare the capacity built by restructured markets.

Assumption 3: $d_1^R Y_1^R = d_2^S X_2^S$

The amount ($) firms in restructured markets overinvest in electricity capacity is equal to the amount those firms in synthetic markets could invest in the alternative good. In other words, a full employment economy is assumed where capital is invested in some project, electricity or other.

Assumption 4: $X_1^R = X_1^S$

The amount of capacity needed to meet the demands of the restructured electricity market, before excess capacity is considered, is equal to the capacity needed to meet synthetic electricity demand.

Assumption 5: There are two plant technologies and one fuel type used for new plants.

Changes in the levelized cost of electricity production and length of build time led to over 95 percent of non-RPS plant construction from 1997-2013 to be gas-fired CCGT or CT plants. Therefore, it is safe to assume only CCGT and CT technologies are used for construction and only natural gas is used as a fuel.

Assumption 6: $x_{11} = f(X_{11}, X_{12})$ *and* $x_{21} = f(X_{21}, X_{22})$

Electricity demand can be met by either CCGT or CT plants. While it is more cost effective in the long-run to meet peak load with CT plants and baseload/intermediate load with CCGT plants, the lower marginal cost of CCGT plants allows them to bid lower than the CT plants. This leads to CCGT plants being used for peaking as well if they are sitting idle. While using CT plants for intermediate or baseload needs is less common due to higher marginal costs, it is possible they may be used during unexpected plant outages.

Once these assumptions are implemented in the model and the terms are rearranged (See Appendix 1 for details), the following welfare impacts from restructuring are identified:

$$(14) \quad \Delta W = \int_{x_2^R}^{x_2^S} p_2^S - D(x_2) + \sum_{i=1}^{n} c_{2i}^S x_{2i}^S - p_{2i}^S x_{2i}^S + \sum_{i=1}^{n} \sum_{a=1}^{2} \left(c_{1a}^S - c_{1a}^R\right) x_{1ai}^S + \sum_{i=1}^{n} \sum_{a=1}^{2} \left(d_{1ai}^R - d_{1ai}^S\right) X_{1a}^S + \int_{x_1^S}^{x_1^R} D(x_1) - \sum_{i=1}^{n} \sum_{a=1}^{2} c_{1ai}^R (x_{1a}^R - x_{1ai}^S) - d_{1a}^R (X_{1ai}^R - X_{1ai}^S)$$

Welfare Effect I: $\int_{x_2^R}^{x_2^S} p_2^S - D(x_2) + \sum_{i=1}^{n} c_{2i}^S x_{2i}^S - p_{2i}^S x_{2i}^S$

This is the total welfare loss from the marginal capital investment in restructured markets not occurring for the alternative good. Investment has already taken place in this market, so the additional investment raises quantity from x_2^R to x_2^S.

Welfare Effect II: $\sum_{i=1}^{n} \sum_{a=1}^{2} \left(c_{1ai}^S - c_{1ai}^R\right) x_{1ai}^S$

This is the change in welfare in the electricity market from improvements in power plant cost efficiency. Several papers in the literature estimated that restructuring reduced fuel and O&M costs for plants operating in restructured markets. Note that it is calculated over electricity demand in the synthetic market. The cost improvements that impacted the difference in electricity demand between restructured and synthetic markets is shown in welfare part V. The papers cited above did not distinguish between these effects.

Welfare Effect III: $\sum_{i=1}^{n}\sum_{a=1}^{2}\left(d_{1ai}^{R}-d_{1ai}^{S}\right)X_{1ai}^{S}$

This is the change in welfare from the restructuring impact on plant construction costs. No study as of yet has attempted to calculate this differential. As with the previous part, this is only calculated for the capacity built in the synthetic market.

Welfare Effect IV: $\int_{x_1^S}^{x_1^R} D(x_1) - \sum_{i=1}^{n}\sum_{a=1}^{2} c_{1ai}^{R}(x_{1ai}^{R}-x_{1ai}^{S}) - d_{1ai}^{R}(X_{1ai}^{R}-X_{1ai}^{S})$

This is the change in welfare from restructuring altering electricity demand due to a change in the price of electricity. The two components of this change are the variable and fixed cost components of electricity production. If restructuring lowered the price of electricity as it was intended to do, consumers would respond by purchasing more electricity, as identified in the difference between X^R and X^S.

4.2 Welfare Impact Estimation

Of the four welfare effects identified above, the focus of this chapter is on the impact of excessive entry on welfare, so only I and IV will be estimated. Effects II and III are of interest, but are not directly related to excessive entry.

4.2.1 Effect I

In order to calculate the total welfare change from additional firm investment in the alternative good market, it is worth noting that the economy is assumed to be at full employment. If this is not true, capital not invested in power plants may sit idle. Given that a large amount of the capital invested in the electricity sector originated from financial firms, this capital would otherwise have been invested in the next best alternative. The markets for this capital would most likely have already existed, with the additional amount contributing marginally to the existing market. With no market imperfections assumed outside of the excessive investment in the electricity sector, the increase in consumer welfare would be insignificant and no above-normal profits attainable. Therefore, the total welfare change is essentially equivalent to the net present value of the returns from the investment in the alternative good. This is simplified by summing the following equation over the number of state markets and years (See Appendix 2 for details):

$$(15) \quad \sum_{t=1}^{T}\sum_{s=1}^{S}(1+r)^{t} r d_{11st}^{R}(Q_{11st}^{R} - X_{11st}^{S}) + \sum_{t=1}^{T}\sum_{s=1}^{S}(1+r)^{t} r d_{12st}^{R}(Q_{12st}^{R} - X_{12st}^{S})$$

Construction cost of plants (d) is derived from a mixture of sources detailed in the Data section of this chapter, Q_{11}^{R} is the net CCGT capacity change in restructured states, Q_{12}^{R} is the net CT capacity change in restructured states, $X_{11}^{S}$is the net CCGT capacity change for the synthetic state and X_{12}^{S} is the net CT capacity change for the synthetic state. r is the assumed standard rate of return on capital investments of 10 percent. Since CCGT and CT are substitutes in production, excessive investment in CCGT resulted in insufficient investment in CT, as shown in Chapter 1 of this chapter. As a result, the two equations will have conflicting signs and magnitudes.

Given that power plants, once completed, are a durable investment, there are two underlying calculations. The first is the amount of plants that are still not fully used as of 2013.[67] The second is the cost of constructing plants before they were required. As a result, the NPV calculations in this section will be large and positive in the early years, when excess capacity was added, and negative in the later years, as that capacity is put to use.

Estimating the welfare impact of excessive CCGT and CT construction using synthetic control relies on the following assumptions:

Assumption 7: The synthetic control method properly estimates the CCGT and CT capacity that would have been added by states if they had not restructured.

Estimating the counterfactual is not a precise science, as only one outcome is observed. This chapter uses synthetic control, introduced in section 7, to approximate the

[67] Borenstein and Bushnell (2016) note that capacity factors are still low in many restructured states.

amount of CCGT and CT capacity added by restructured states if they did not restructure. The strength of this method is it identifies states that most closely resemble the restructured state and assembles a synthetic version of that state. The close match in CCGT and CT capacity additions in the period prior to restructuring provides confidence that this is an accurate counterfactual (See Chapter 1).

Assumption 8: CCGT and CT net capacity increases equal CCGT and CT capacity additions

The nature of the EIA Form 860 data entails using net capacity additions to substitute for actual new plant builds. While not precisely equal due to the existence of plant retirements, CCGT was newly introduced in the US starting in the late 1980s and very few CT plants were built until the 1990s, so plant retirements are not a concern.

Assumption 9: EIA construction cost estimates with regional cost adjustment approximate the actual build cost in the restructured states.

With comprehensive data on individual plant builds unavailable, this chapter uses standard estimates of construction costs. The EIA and other entities issue one cost estimate annually for the US, with regional adjustments to distinguish between less costly and more costly states. While this approach is not ideal, if a bias does exist, the direction is unclear. See Chapter 1 for a more thorough explanation of these data sources.

Assumption 10: The regulated level of capacity additions is in excess of the optimal amount.

The comparison in this analysis is between the actual added capacity and the regulated amount, not the optimal. As Joskow (2008) notes, the Averch-Johnson effect induces utilities to overinvest in capital, with the result being that utilities were compensated for more capacity than was required to adequately meet load. Therefore, the calculations in this section should be seen as a lower bound of the true welfare impact.

Table 8 shows the results from this calculation.[68] As of 2013, the net present cost of the buildout in restructured states was \$13.6 billion. This is largely due to CCGT plants that were built and are still not fully used. As the 1998-2013 summary row shows, the synthetic restructured states would have built over 59 GW less CCGT capacity without restructuring. The cost pattern over the years was as predicted, with losses incurred in the first seven years and the cost reduced in the final nine years, as there was a need for capacity to meet ERS growth. This amount was slightly reduced by the deficiency in CT plants predicted in the synthetic states. Moving past 2013, there are additional benefits to having these plants available which reduces the cost, with a small adjustment due to an increase in CT plant construction to meet the deficiency in the previous period. However, the benefit of having these plants operational is limited by low projected ERS growth in the US, due to the adoption of energy efficiency programs. Given EIA estimates of ERS growth, plant retirement and cost, excess capacity in states that restructured will not be eliminated until at least 2024.[69] Including these years, the net present cost of the buildout is \$10.8 billion.

[68] Further Assumptions: Discount rate=10 percent. No difference in technology between 1998 and 2013. Build time for CCGT was 3 years for 1993-2003, 4 years for 2004-2013. CT was 2 years for 1992-2013. Costs for three year CCGT build period were distributed 40%, 50% and 10%. Costs for CT build period were distributed 50% and 50%. Costs for four year period were distributed 5%, 25%, 55%, 15%. Future projections of construction are from EIA AEO 2015.

[69] One confounding fact of the post-buildout period is the construction of CCGT plants, despite the presence of excess capacity. This continues in the post-2014 period, with 15 GW of CCGT estimated to be added. Two reasons explain the continued construction. 1) Regions added capacity in a heterogeneous fashion, with some states adding more capacity than others. Therefore, there are balancing authorities in restructured areas which are in need of capacity. 2) Camerer and Lovallo (1999) suggests managers can suffer from overconfidence in their own abilities. Therefore, CCGT plants will continue to be added by firms that either are low-cost producers or believe they are.

Table 8. Alternative Market Producer Welfare Loss

Year	#states	CCAct	CCSyn	CTAct	CTSyn	PV($mil)
1997	2	0	0	-	286	(52)
1998	3	2,669	7	4,195	2,486	1,072
1999	13	2,371	724	1,187	701	506
2000	17	29,722	4,074	3,326	6,755	5,569
2001	21	7,186	3,833	11,463	7,038	1,533
2002	21	31,230	14,654	7,886	11,883	2,671
2003	21	36,517	16,818	2,565	3,268	3,502
2004	21	13,827	8,781	-	2,060	600
2005	21	7,589	11,407	1,042	1,434	(588)
2006	21	5,897	944	954	846	718
2007	21	2,077	6,098	1,605	1,499	(544)
2008	21	1,539	3,608	1,706	2,602	(380)
2009	21	3,110	10,704	1,139	1,366	(837)
2010	21	4,019	1,427	1,142	1,405	243
2011	21	4,729	6,539	1,038	2,367	(312)
2012	21	2,415	3,957	1,020	186	(78)
2013	21	1,488	3,684	3,231	1,582	(53)
1998-2013	21	156,384	97,257	43,498	47,765	13,570
2014-2024	21	14,951	74,080	6,859	2,597	(2,732)
Total		171,335	171,337	50,357	50,362	10,837

Notes: CC/CTAct is the amount of CCGT and CT capacity built (MW). CC/CTSyn is the counterfactual amount (MW). NPV is the net present value of the investment in millions of 2013 $.

The distributional burden of this buildout is very different from the large increase in nuclear capacity in the 1970s and early 1980s. In the prior nuclear construction period, states were tightly regulated under cost-of-service regulation, with the cost of investment errors often being passed on to the consumer in the form of higher prices.[70] In this case, the costs of the gas boom were on the firms, as consumers were insulated from investment errors through retail and generation competition. While the impact on firms would affect consumers through misallocation of resources, it did not have the same effect as in the nuclear case.

[70] As previously discussed in this chapter, the cost of nuclear power plants was underestimated due to additional safety measures and optimistic cost efficiency expectations.

While Joskow was certainly correct, that COSR incentivized firms to overinvest in capital-intensive methods of electricity generation, there are two other points worth considering. First, the cost of excess investment can be large, as evidenced by a number of events across the world over the last 20 years. Second, while COSR incentivizes firms to be capital-heavy, competitive electricity markets encourage firms to underinvest in capital-intensive plants. Uncertainty and expectations play a large role in the construction of power plants, which has traditionally been minimized by the ability of utilities to recoup the cost of their investment through rate increases. This is why new nuclear additions are not strongly considered without federal loan guarantees. However, making a large investment in a competitive electricity market is a risky venture and firms may opt for less capital-intensive projects even if they are projected to have a higher levelized cost. It is not surprising, therefore, that fast-building plants were chosen over coal and nuclear plants. While it appeared these were sound investments in the late 1990s, when gas prices were low, the high gas prices of the 2000s led to a string of bankruptcies.

4.2.2 Effect IV

The change in welfare from effect IV is dependent on the impact of restructuring on electricity prices and the elasticity of electricity demand. Theoretically, electricity prices could have increased or decreased as a result of restructuring. However, controlling for changes in fuel and construction materials prices is difficult, so analyses in this area of study must attempt to isolate only the effect on electricity prices of restructuring. This chapter identifies two decreasing and three increasing effects.

There are two channels for restructuring leading to falling electricity prices. The first is the reduction in construction and operation costs from competition that Fabrizio et al and others found. Given a downward-sloping demand curve and competitive electricity market that passes cost savings on to consumers in the form of lower prices, this would increase electricity demand from x_1^S to $x_1^{R_1}$. The magnitude of this effect is unknown, with the competitiveness of electricity markets often in question (Bushnell et al 2008). This effect is also outside the scope of this analysis, which is focused only on the effect of excessive entry on total welfare.

The second effect is the decrease in costs associated with excess capacity. With a larger number of more efficient CCGT plants in a market, the dispatch cost would be reduced, as balancing authorities would receive lower bids compared to less-efficient CT plants. The largest effect of lower CCGT bids would be in peak periods. Competitive electricity markets meet peak demand by enabling peaking plants to charge high prices during the few periods a year in which they operate to recover fixed costs. Peaking plants are able to charge these high prices because they lack competition in these rare periods. With the introduction of more plants, bids during this period would fall, lowering the price of electricity for consumers. In a real-time pricing market, this would influence consumers to increase electricity demand during the peak. However, the majority of markets during this time period did not have access to real time pricing, resulting in a decrease in average electricity prices. Given the assumption of a downward sloping demand curve for electricity, this will increase electricity demand from $x_1^{R_1}$ to x_1^R.

One channel for restructuring increasing electricity prices is the nature of price regulation prior to restructuring. Partly due to increasing nuclear costs and partly to tight price regulation, several utilities across the US received very low grades on their bonds. This suggests investors were wary of low profits in the industry. Following restructuring, firms were freed from price regulation that may have allowed for profits approaching the opportunity cost of capital. This would not have been the first time this occurred in US regulatory history. As noted in Winston (1998), railroads were losing money in the 1970s due to price regulation. Following the Staggers Act of 1980, railroad prices increased initially as firms were freed from being forced to set prices too low. This is attributed to the successful recovery of the industry and a similar effect may have been present in the restructured electricity industry.

A second channel, as noted in Su (2015), is the inclusion of search and switching costs. Traditionally, consumers had one provider for electricity, which eliminated any complications surrounding information gathering and switching by consumers. With more choices, costs can increase for consumers and lead to higher prices. This is particularly true for smaller customers.

A third channel for restructuring increasing electricity prices is the impact of market power on prices. Restructured markets gave firms the freedom to act strategically when providing electricity to the system. The combination of short-run inelastic demand and supply, along with the inability to store electricity, gives firms the ability to manipulate market prices by holding low-cost generation out of the market. The papers mentioned above are a small part of a large literature that has identified the existence of strategic behavior in markets. The increase in bids raises average electricity prices.

The magnitude of these five effects is unknown, as there is no literature present that attempts to estimate each. There is a large literature, however, on the estimation of electricity demand curve elasticity.[71] The studies vary from 0 to -1 in their estimates, with the majority clustered between -.1 and -.4 and the EIA adopting -.3 as an estimate in 2010. Therefore, it is not surprising that, given both negative and positive theoretical impacts and a relatively inelastic demand curve, a number of studies find little evidence for a significant fall in electricity prices following restructuring (Joskow 2008; Su 2015; Borenstein and Bushnell 2016).[72] Additionally, any change in price would be, at least, partly due to increased plant efficiency rather than excessive entry. Lacking evidence in the literature of lower prices due to excessive entry from restructuring, the assumption of this chapter is that excessive entry had a minimal impact on prices that is unable to be

---

[71] See Espey and Espey (2004), Paul, Myers and Palmer (2009) and Alberini, Gans and Velez-Lopez (2011) for summaries of elasticity estimates in this literature.

[72] Su (2015) finds short run impacts in the residential market, but these disappear in the longer-run. Commercial and industrial markets do not experience significant impacts. These are consistent with the findings of Apt (2005) and Fagan (2006), which found no difference in industrial electricity prices between restructured and non-restructured states.

distinguished from other factors like changing fuel prices, population growth, changes in economic activity, and others. If prices did not change as a result of excessive entry, the quantity of electricity demanded in restructured states would be no different than that demanded in the synthetic state. In this case, welfare effect V is insignificant and, therefore, not a factor in the welfare calculation.

4.3 Total Welfare Effect of Restructuring

The purpose of this section is to identify the total welfare impact of excessive entry due to restructuring. While there were other effects on welfare identified, these are outside the scope of this chapter, which is focused on the excessive entry caused by restructuring. Two separate effects were considered, with the impact on the alternative market being a negative effect and the impact on prices and quantity in the electricity market a positive effect. This chapter found welfare losses of $10.84 billion in the alternative market, with no evidence of a change in welfare in the electricity market due to excessive entry.

5. Conclusion

The US Gas Boom transformed the US electricity industry, providing the capacity necessary for the coal to gas switch which has had substantial environmental and cost benefits. However, it also left the industry over-capitalized and created large-scale panic in 2006. This finding was inconsistent with expectations prior to restructuring, when the consensus suggested restructuring would lead to a more efficient outcome, and this chapter sought to explain the results and provide an estimate for their impact. The conclusion of this chapter is the Gas Boom was caused by a combination of herding behavior and manager overconfidence, low interest rates, and coordination failure.

Of the four models tested, the Contagion model describing herding behavior, and a similar story of manager overconfidence, led to over-investment in the US electricity industry. Of particular importance is the length of the boom and the large number of bankruptcies in 2006, which suggest a non-equilibrium outcome and a phenomenon that existed for too long not to be noticed. Low long-term interest rates contributed to the Contagion effect, providing significant low-cost funding for new entrants beyond a sustainable level. Anecdotal evidence suggests that a coordination failure spurred early investment, with firms unable to determine the potential level of supply in these markets. These factors led to significant welfare losses from excessive entry for producers, which were most apparent in a series of major bankruptcies in 2006. The lack of evidence for increases in consumer welfare through lower electricity prices narrows the welfare focus to firms, with losses of approximately $11.2 billion.

Given the significance of this event for the electricity industry, additional work in other industries will be useful in crafting policy for the future. Of particular importance are those industries that are either newly deregulated or considering a change in regulation. In electricity, the majority of attention was placed on the lower bound of the investment function. This chapter shows that the entire investment path is worth considering, with the likelihood of over-investment being as or more likely than under-investment. While entry dynamics exist in all industries, further work is needed to prevent the investment problems the electricity industry faced and the effect this will have on the dynamism of the competitive process.

CHAPTER 3

THE IMPACT OF LOAD VARIANCE ON US ELECTRICITY PRICES

Reducing electricity demand variance through energy efficiency and real-time pricing has been a major US policy goal for over two decades. However, there has yet to be an estimate of the impact of reducing variance on electricity prices. Leveraging a unique electricity load dataset to estimate the impact of changes in load variance on electricity prices, a reduction in load variance of 10 percent leads to a fall in electricity prices of 9.7 percent. This effect is largest for residential and industrial users and shows there are potentially large gains in lowering electricity prices available through real-time pricing and other demand smoothing measures. (JEL Q41, Q49, L94)

1. Introduction

The problem of peak congestion is common to many industries. Inner-city roads are congested in both the early morning and early evening hours, airports see the most travelers in the early morning during the week and on the weekend travel days, and telecommunications companies often face a spike in internet usage during evening hours (Lawrence 2011). In the electricity sector, demand peaks in the late afternoon and early evening, as summer temperatures reach a daily high and people return home from work. Markets typically utilize price changes to reduce congestion, charging higher prices during peak hours and lower prices off-peak. However, demand-smoothing is reliant on the responsiveness of consumer demand to changes in prices, the ability to impose marginal cost pricing on consumers and the flexibility of supply. The electricity industry fails in its ability to price at marginal cost and have flexible supply, with the responsiveness of consumers to real-time pricing uncertain.[73] The inability to use prices

[73] Wolak (2011) finds large effects in Anaheim and DC. Allcott (2011) finds small effects in Chicago.

to smooth demand forces firms to maintain spare capacity to meet demand spikes and increases the marginal cost of electricity production, both of which raise retail prices for consumers. While this effect is most likely negative, the size of the cost and welfare losses from load variance are unknown. This chapter uses a unique dataset to provide a first estimate of the impact of load variance on electricity prices and consumer welfare.

Electricity is a large and important industry in the United States. In 2015, consumers spent over $390 billion dollars on electricity, $178 billion of which was spent by residential customers. Prices vary considerably across states and over time.[74] High electricity prices are a topic consumers are very sensitive to, with restructuring in the 1990s instigated primarily by high prices. A significant factor influencing electricity prices is demand, which increased by 38 percent in the US between 1990 and 2015. The national average hides significant variation among states, with electricity demand more than doubling in Nevada and North Dakota and decreasing in Washington. Changes in the structure of demand during this time period, through greater electronics and air conditioning use, occurred simultaneously with energy efficiency measures, both of which impacted electricity prices.

Changes in the variability of electricity demand affect policy outcomes in several ways. The impact on emissions was estimated in Holland and Mansur (2008), with the authors finding heterogeneous regional effects, depending on the carbon intensity of the marginal source of electricity. A second impact is on the electricity bills of consumers, which is the topic of this chapter. An important aspect of this impact is whether the relationship between load variance and electricity prices is linear. There is good reason to believe, with a convex market cost function, that the relationship is non-linear, which has important policy implications. A non-linear relationship implies that states with high

[74] In 2015, Hawaii was the highest priced state, at 26.17 cents/kilowatt hour (kwh), and Washington the lowest, at 7.40 cents/kwh. From 1990 to 2015, real US electricity prices fell and then rose, starting at 11.91 cents/kwh in 1990 and falling to 9.37 cents/kwh in 2000 before rising to 10.41 cents/kwh in 2015.

load variance will have a greater impact on electricity prices from reducing load variance than those states with low variance will. Therefore, this chapter shows not only the average effect, but also heterogeneous effects that are important as states consider new programs to reduce peak load.

This chapter identifies four channels through which load variance affects electricity prices. First, increasing variance in load increases the use of high fuel cost power plants, increasing the cost of providing electricity. Second, more frequent power plant starts and stops increase O&M costs. Third, higher load congestion raises the amount of transmission capacity needed. Fourth, additional capital is required to meet high peaks, an increasing share of which sits idle for long periods. Each of the first three channels are directly linked with increased electricity prices, but the fourth may not be, as optimal firm capacity choice in higher load variance markets may be less capital-intensive than in low load variance markets. The role of cost of production as an intermediary in these channels shows this is a two-stage problem. The first is between load variance and cost, with the ability of firms to pass on costs to consumers in the form of higher electricity prices the second stage.

A broad literature exists on the topic of peak congestion and pricing that spans across multiple industries. In the transportation literature, attempts at reducing the problem of peak congestion have been well documented, both in a theoretical sense and empirically (Arnott, De Palma, Lindsey 1993; Davis 2008). Electricity markets have drawn significant attention, with early papers recognizing the problems of capacity lumpiness due to economies of scale and the inability of consumers to respond to marginal cost pricing (Steiner 1957; Williamson 1966; Chao 1983). Later work measured the inefficiencies associated with inelastic demand and suggested optimal pricing and investment strategies (Joskow and Tirole 2007). The introduction of real-time pricing experiments led to simulations and empirical estimates of the impact of real time pricing on demand and, by extension, peak load, as well as welfare impacts (Borenstein and Holland 2005; Holland and Mansur 2006; Allcott 2011).

Research into the determinants of electricity prices falls into one of two categories. The first investigate the short-run fluctuations of electricity prices in wholesale markets (Barmack et al 2008), while the second attempt to measure prices without the influence of market power (Wolfram 1999; Joskow and Kahn 2000; Borenstein et al 2002). Holland and Mansur (2008) uses FERC Form 714 data to measure the impact of load variance on emissions. The authors' analysis at the NERC level found small effects from reducing load variance, largely due to regions using hydro as a peaking resource.

This chapter contributes to the literature in two ways. It provides the first estimate of the impact on electricity prices from changes in load variance. Given the substantial interest in real-time pricing (RTP) and energy efficiency in the literature, knowing the actual cost of load variance helps inform the welfare impact of these policies. It also uses sectoral data to estimate who is paying for higher load variance.

The unique contribution of this chapter is the construction of a state-level hourly demand dataset assembled from FERC Form 714. This chapter leverages variation in this data within states across time to identify the causal impact of load variance on electricity prices. I argue that, after controlling for average demand, plant heat rates and age, and state characteristics, the effect of load variance on electricity prices is exogenous. This chapter utilizes the panel dataset in OLS estimation at both the state and NERC region levels. I test for heterogeneous effects through several non-linear specifications. Given the convexity of market cost curves, the hypothesis of this chapter is the effect will be increasingly negative as load variance increases.

This chapter finds an increase of load variance by 10 percent leads to a rise in electricity prices of 9.7 percent. A 10 percent change is comparable to giving Michigan the load variance of Florida. This large effect is robust to a number of specifications and alternatively constructed variables. Residential and industrial consumers' electricity bills change the most from altering load variance, which is consistent with an effort to decrease load in peak periods and increase it in off-peak hours.

These results imply there are large gains for states that reduce emissions. Based on the findings in Section 5 of this chapter, these effects are not heterogeneous, so states with both low and high load variance can lower electricity bills by flattening their daily and seasonal load curves. As shown in Section 4, load variance has been declining in the US during the past 20 years, with the introduction of energy efficient appliances and lower cost peaking generation reducing the cost of peak periods. As states search for new ways to lower electricity congestion costs, improving RTP and storage are critical.

The structure of the remainder of the chapter is as follows. Section 2 provides background information on the electricity industry and load variance. Section 3 provides a causal link between load variance and electricity prices while also introducing the empirical strategy of this chapter. Section 4 explains the data used in this chapter and shows relevant summary statistics. Section 5 presents the empirical results of both the panel OLS and alternative specifications. Section 6 concludes.

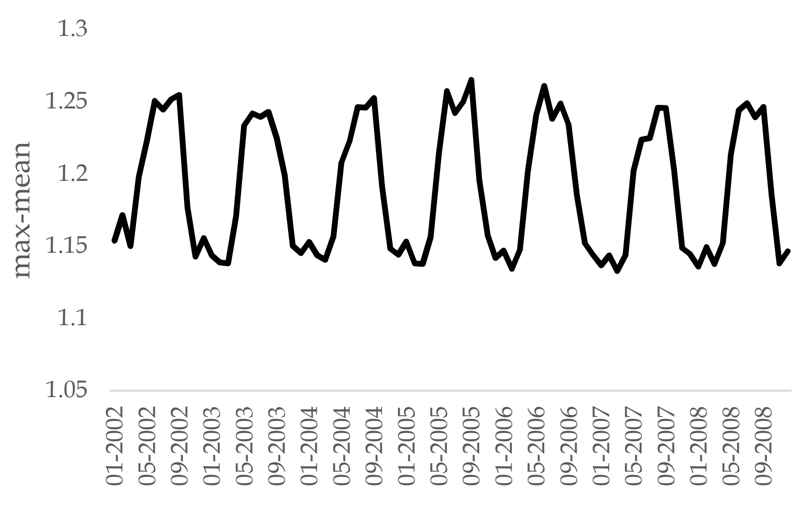
2. Background

With a few, small exceptions, US electricity demand must equal supply in each period. The lack of storability requires balancing authorities to ensure the system is neither overloaded, nor undersupplied. Throughout most of the 20th century, vertically-integrated utilities performed this function, as they owned the transmission and distribution networks. Within the last 15 years, many utilities have surrendered control of these networks to non-profit regional transmission organizations (RTOs) and independent system operators (ISOs). The amount of electricity on the grid at any particular moment is known as the load, which is similar to the electricity consumers wish to receive at that time (electricity demand). Throughout this chapter, the two are considered interchangeable, with the difference not impacting the estimation strategy.

Balancing authorities typically face two reliability problems. The first are supply failures, such as downed power lines and forced plant shut downs. These are often unpredictable and can lead to long periods before returning to full operation. The second is due to changes in the demand for electricity, such as rising or falling temperatures and increases in consumer electronic use. These changes are more predictable, leading to balancing authorities keeping excess resources to account for load variation.

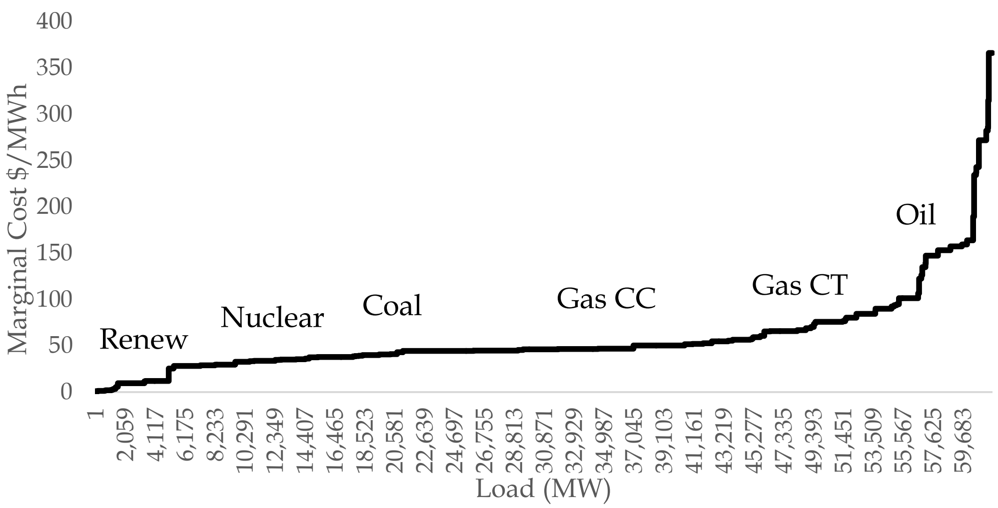
Figure 19 shows monthly load variance for Texas over a span of six years. The lows correspond with the winter months and the highs with summer months. Daily and seasonal load variation are due to a number of factors, such as weather, income, technology and working schedules (EIA, 2016). Temperature changes within a day cause load to peak in the late afternoon, particularly in the summer. Winter lows do not have the same impact on electricity demand, as a large amount of heating uses natural gas and propane. While temperature is a dominant factor in load variance, changes in income, electronics usage, and energy efficiency programs starting in the 1990s contributed to changes in load variance, both within a day and between seasons (EIA, 2004).

Figure 19. Texas Monthly Load Variance



To meet varying load requirements, balancing authorities utilize different types of power plants, typically known as the dispatch order (See Figure 20). The first plants used to meet the demand for electricity are low marginal cost plants, which typically have included wind, solar, hydro (Renew), coal and nuclear. Given the limited amount of wind and solar resources and the regional variation in hydro power plants, the majority of initial plants in the dispatch order are base load power plants, with high capital costs and low fuel costs. These plants are in operation all year, excluding scheduled maintenance.

Figure 20. Florida Balancing Authority Dispatch Order, October 2010



As the sun rises and the demand for electricity increases, the next group of plants in the dispatch order come online. These have traditionally been combined cycle gas turbine (CCGT) plants, although falling natural gas prices recently have resulted in competition between coal and CCGT plants in the dispatch order. These plants will run most of the day and, in some cases, overnight depending on the region and time of year. CCGT plants are typically built for intermediate and base load purposes due to their efficiency and, in warm and humid southern climates, will often run most of the year.

When temperatures reach a daily maximum and people return home from work, system load reaches its peak and the final part of the dispatch order engages. These are largely combustion turbine (CT) gas-fired power plants, although some areas use pumped hydro storage and older steam oil and gas plants. These plants have very high heat rates and fuel costs but low construction cost, so they are often the cost-minimizing investment choice for peaking purposes. They are also cheap and quick to ignite, making them ideal for a few hours of use in a day. These peak periods tend to be more pronounced in populated, temperature varying areas and have a seasonal component, so in winter months these plants are rarely used.

Electricity is not the only industry that faces peak congestion. Similar problems occur in transportation, telecommunications, entertainment, hospitality and retail. However, many of these industries are able to use changes in price to smooth out peaks in demand.[75] For example, hotels typically charge higher prices on the weekends and in peak seasons, water parks discount prices on the weekdays and some cities have toll lanes on their highways which activate during rush hour. In order to be able to smooth demand, however, firms must be able to charge customers marginal cost prices, supply must be flexible, and customers must respond to changes in price. The electricity industry faces problems with each of these factors which create greater congestion cost.

As shown in Figure 20, the marginal cost of electricity increases throughout the day, but in most cases, customer rates do not (Borenstein and Bushnell 2016). This is due to the lack of real-time pricing infrastructure, which would regularly update consumers on the cost of electricity throughout the day and bill them accordingly. While there are multiple reasons for the delay in implementation of real-time pricing,[76] its absence has led electricity-selling firms to use time-of-day pricing and tiered-pricing as approximations.[77] While these methods provide electricity consumers with more incentive to reduce consumption than average pricing, balancing authorities nevertheless

[75] It is worth noting that smoothing demand through peak pricing is not a first-best outcome, as there is a welfare loss associated with this. However, it is smaller than if there were no peak pricing.
[76] See Allcott (2011) for a description of some of the problems.
[77] See Borenstein (2012) for work on the impacts of these various programs.

face inelastic demand. The failure to fully adopt real-time pricing means customers are responding to incorrect price signals and demand can't be smoothed through higher prices. The electricity industry, therefore, utilizes power plants with low capacity factors that are built to operate, in some cases, only a few days out of the year.

The industry also must overcome inflexible investment size. Ideally, firms would be able to build plants of any size to fit the need in the market. The electricity industry, however, has always been characterized by economies of scale. Prior to the boom in CCGT and CT gas plant construction in the late 1990s, the majority of power plants built were steam-powered coal, nuclear, and gas, for which scale was crucial in cost minimization. Although the adoption of jet turbine technology for use in gas-fired plants has reduced the minimum efficient scale, it is still the case that investment in power plants is lumpy, leading to plants that are not the right size for the market.

Finally, there is mixed evidence of the impact of real-time pricing on electricity use. Allcott (2011) found small effects, while Wolak (2011) found larger effects. Initially, it appears the effects found in Wolak are consistent with intuition. If consumers face a high price for electricity, they should switch washing dishes and clothes and the usage of some electronics to later in the evening, when the peak begins to fall. However, enough rigidities exist in terms of work, dinner and TV scheduling that make it possible consumers may not switch enough to justify the cost of real-time pricing. Therefore, it is unclear that full time adoption of rea-time pricing would have significant effects.

Due to the factors listed above, US electricity markets consist of a large number of plants that have low capacity factors and high marginal costs. These plants are chosen by design, as they can only compete with more efficient plants when capacity factors are low. Given that there may be only a few days when these plants need to operate, only very high prices can incentivize firms to invest in peaking plants.[78] A more formal structure of this relationship is presented in the following section.

3. Methodology

Average state retail electricity prices are the focus of this chapter and vary based on market regulation. For a fully restructured market, the retail price is the outcome of competition between market retailers and depends on the wholesale market price. Regulators in non-restructured markets determine electricity prices based on cost-of-service, with firms allowed to recoup fuel and O&M costs, while also earning a return on capital. The purpose of this section is to uncover the influence pathways of load variance on electricity prices informing the reduced form estimation strategy of this chapter.

3.1 Restructured Price

Following Borenstein and Holland (2005), the retail sector is modeled as Bertrand competition and assumes no switching costs. Profits in this sector, after accounting for the cost of transmission and retail operating costs, are zero. More formally, the zero profit condition in the retail market can be written as:

$$(16) \qquad \sum_{t=1}^{T} (\bar{p}^E - w_t - c^{ret})\, D_t(\bar{p}_{SR}) = 0$$

$$\bar{p}^E = \frac{\sum_{t=1}^{T} (w_t + c^{ret}) D_t(\bar{p}^E)}{\sum_{t=1}^{T} D_t(\bar{p}^E)}$$

Retailers charge a price ($\bar{p}^E$) based on the cost of purchasing electricity from the wholesale market (w) and retail operations and distribution costs (c^{ret}), weighted by demand in that period. This price is often changed only a few times a year, with limited

[78] In years where the peak is not as high, these plants may not be able to compete due to high marginal costs. To combat this, ISOs were set up with reserve margin requirements to prevent supply shortages. Joskow (2008) discusses the concern that restructured markets would underprovide peaking resources due to the problem described. This does not reflect reality, however, as restructured markets have excess capacity (Borenstein and Bushnell 2016).

adoption of real-time pricing (RTP) and tiered rate structures. Therefore, $\bar{p}^E$ is not influenced by wholesale prices in the particular time period.

$$(17) \quad w_t = c_t(D_t) + mark$$
$$c^{ret} = tra(D_t) + \bar{o}$$

Wholesale electricity providers bid to provide power both in day-ahead and spot markets, where market operators use a uniform-price, sealed bid auction. With the exception of a small number of large capacity firms in particular market conditions, firms have an incentive to bid their marginal cost (Mansur 2008). Firms are then paid the market-clearing price, which is the cost of the marginal unit in that hour (c_t) plus any impact of firms using their market power to raise prices (mark). Retail and distribution costs consist of non-varying operations costs ($\bar{o}$) and transmission costs (tra), which are influenced by the market demand structure. The retail price consumers pay varies based on the wholesale price in these markets and the cost of transmission.

3.2 Regulated Price

All states without retail competition began the process of restructuring and stopped at various stages. Those states that advanced past initial reports by the public utility commission (PUC) and pilot programs but either did not open to retail competition or suspended it, often broke up vertically-integrated utilities. This process created broader wholesale markets and, in many cases, transferred transmission control to independent system operators (ISOs), leaving existing utilities with a smaller fraction of generation assets while still acting as single providers of electricity. Due to lack of retail competition, which would allow an unregulated utility to act as a monopoly in the retail market and have monopsony power in the wholesale market, PUCs continued to set the retail price for consumers in these states using cost-of-service regulation (COSR). Under

COSR, regulators set retail prices ($\bar{p}^E$) for utilities to generate a set rate of return (r) on capital (mK). Shown formally:

$$\text{(18)} \quad \sum_{t=1}^{T}(\bar{p}^E - c_t - c^{ret})\,\widetilde{D}_t(\bar{p}^E) + (\bar{p}^E - w_t - c^{ret})\widehat{D}_t(\bar{p}^E) = \sum_{t=1}^{T} r(\lambda_t)m_t K_t$$

$$D_t(\bar{p}^E) = \widetilde{D}_t(\bar{p}^E) + \widehat{D}_t(\bar{p}^E)$$

$$\lambda = \frac{\widehat{D}_t}{D_t}$$

$$\bar{p}^E = \frac{\sum_{t=1}^{T}\{r(\lambda_t)m_t K_t + c^{ret}[D_t(\bar{p}^E)] + c_t\widetilde{D}_t[(\bar{p}^E)] + w_t\widehat{D}_t[(\bar{p}^E)]\}}{\sum_{t=1}^{T} D_t(\bar{p}^E)}$$

Firms meet electricity demand both from their own plant's production ($\widetilde{D}_t$) and from wholesale market purchases ($\widehat{D}_t$). If the firm is a net seller of electricity ($\lambda \leq 0$), profit from these operations are accounted for by the regulator in the return on capital. This is consistent with standard practice of customer rebates when firms exceed their allowed rate of return. An alternative to this approach is to assume firms are fully integrated into the wholesale market, with a firewall between their retail and wholesale operations. However, this assumption is not consistent with transmission constraints and plant placement. Regulated prices rely on similar factors as restructured prices, with the main differences being the rate of return guarantees which are not present in restructured markets, and the opportunity for markups in restructured markets, which are not possible in regulated markets. From this point forward, both restructured and regulated markets are assumed to react to wholesale prices and the factors that determine them.

3.3 Cost Function and Demand

Firm profit is written as follows:

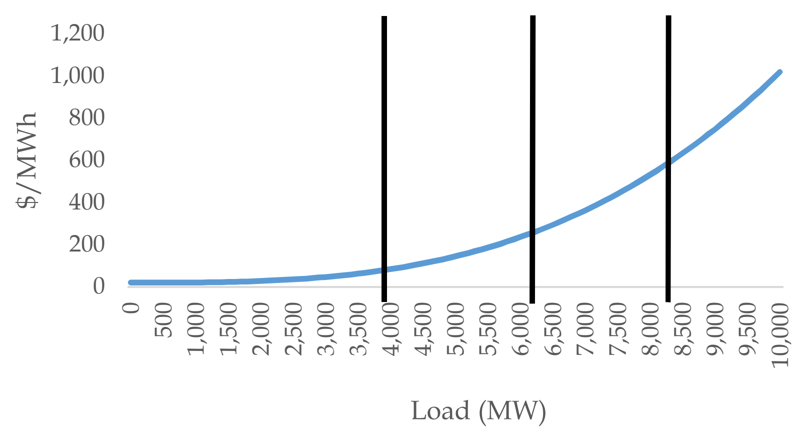
$$\text{(19)} \quad \pi_{it} = w_t D_{it}(\bar{p}^E) - c_{it}D_{it}(\bar{p}^E) - m_{it}K_{it}$$

$$c_t = f[fuel(D_{it}, Age_{it}, K_{it}(E(D_{it}), p_t^f), O\&M(w_t, D_{it}]$$

Firm marginal cost consists of fuel and O&M costs. Fuel cost is the product of the average heat rate of the firms' plants and the price of fuel (p_t^f) used by the firm. The fuel and heat rate of plants used by the firm varies based on the level of demand in the market. Heat rate, a measure of the efficiency of the plant in converting fuel to electricity, is inversely correlated with plant age (Age) and strongly influenced by plant choice (K_t).[79] O&M costs include labor, repairs and other operations costs and are determined by standard operating costs (w_t) and equipment repairs. The frequency of repairs increase with more plant starts and shut downs, which vary based on changes in demand.

Figure 21 illustrates how demand determines price within the wholesale market, with firms below the market-clearing price earning variable profit to provide a rate of return to capital owners. The cost curve varies based on the market, with the properties $c'(x)>0$ and $c''(x)>0$.[80] As demand increases throughout the day, wholesale prices rise, which is reflected in the average retail price.[81] By the peak period of the day, the most expensive plants come online, with all plants earning a price above marginal cost.

Figure 21. Hypothetical Wholesale Market



[79] Plant choice determines the prime mover of the turbine, which have different heat rates. Age impacts heat rate through the vintage of prime mover. Any reduction in heat rate is part of O&M cost.
[80] Balancing authority area dispatch curves resemble a function of this type, which are consistent with Mansur (2008), because peaking plants, which have low construction costs, can't use coal and uranium as fuel. Zero marginal cost renewable resources are intermittent, so are not available in large quantities.
[81] Beginning in 2014, the dramatic fall in natural gas prices has flattened this load curve considerably.

The impact of electricity demand on wholesale prices depends both on the average level of demand as well as the variance. When average demand changes, the amount of generating capacity needed changes along with the choice of plant type and fuel. For example, higher average demand leads firms to require more baseload power, which traditionally is met with coal and nuclear, although also increasingly from CCGT. Therefore, an increase in average demand would lead both to more plants being constructed as well as a preference for base load forms of power. When load variance changes, there is a shift in the type of plants and fuel required. Instead of requiring more baseload like an increase in average demand, increasing load variance will lead to more capacity required but much of less-capital intensive and flexible.

Electricity demand, which enters both fuel and O&M costs of the firm, is a product of the number of electricity consumers and the average consumer use intensity. The former is impacted by population, while the latter is influenced by income, weather, season, hour of day, price, environmental sensitivity and electricity sector ratio.[82] Average electricity demand changes based number of consumers and intensity, while variance is impacted only by changes in user intensity.

Incorporating these factors into the equations above and combining both fuel and O&M cost factors, average retail electricity prices can be written as:

$$(20) \quad \bar{p}^E = f(\overline{D_t}(\bar{p}^E), D_t^{Var}(\bar{p}^E), Age_t, K_t, p_t^f, O\&M, mark, restruct, \bar{o})$$

$$\overline{D_t} = f(Pop, Inc, W, Seas, \bar{p}_{t-1}^E, ES, Sector)$$

$$D_t^{Var} = f(Weath, Inc, Seas, H, Sector)$$

where $\overline{D}$ is average demand, D^{Var} is demand variance, Pop is population, Inc is income, W is weather, $Seas$ is season, ES is the environmental sensitivity of the market,

[82] This ratio compares the split of electricity retail sales (ERS) into residential, commercial and industrial. These sectors peak at different points of the day, so an electricity market with high residential demand may have greater demand in the evening than one with high industrial, which peaks mid-day.

H is hour and $Sector$ is a ratio of the residential and commercial electricity sectors to industrial. As shown above, retail prices are dependent on many factors, including the choice of generation capacity (K_t). These are not simultaneous choices, as several years separate the decision to build a new power plant and that plant's availability. Therefore, prices are a function of past factors which led to the generation portfolio that influences price in that time period. For this purpose, a build decision model is constructed and its results fed into the function for average retail price (see Appendix 3). Including the outcome from the build model, average retail prices are a function of the following:

$$(21) \quad \bar{p}^E = f(fuel, O\&M, tra, K, mark, restruct, \bar{o})$$
$$fuel = g(x^{Davg}, x^{Dvar}, Age_t, p_t^f)$$
$$O\&M = h(x^{Davg}, x^{Dvar}, Age_t, w)$$
$$tra = i(x^{Davg}, x^{Dvar})$$
$$K = E_t[j(x^{Davg}, x^{Dvar}, CAP_t, R_t, LC_{Af})]$$

3.4 Load Variance Channels

Load variance influences four of the factors in equation 21: 1) fuel prices, 2) O&M, 3) transmission and 4) capacity. 1) Varying demand increases the use of higher marginal cost plants. These high marginal cost peaking plants replace the medium cost plants which would have been previously operating. 2) Varying demand increases the number of plant starts and stops, leading to higher maintenance costs. 3) Varying demand requires enough transmission to meet demand at its peak, which is higher for a high variance market. Therefore, the high variance market would be expected to have to invest in more transmission capacity than a low variance market. 4), There is a cost of capital sitting idle without producing electricity in a high variance market.

Of these four channels, the first three are fairly straightforward, although the timing on the second and third channels is problematic.[83] Higher load variance should lead to greater fuel, O&M and transmission costs. The capacity channel is complicated in timing and in effect direction. The timing of capacity charges depends on how often variance deviates from market expectations, the length of rate cases and their effects. However, even if timing were not a concern, it is unclear whether the capacity channel raises or lowers electricity prices when demand variance increases.

The reason for this ambiguity is capacity in the electricity industry is different from other industries. The material used for road construction, for example, does not vary whether the road is being built for every day traffic or to meet peak demand. For electricity, however, firms optimize by comparing the marginal and capital costs of technologies and the cost differences in these technologies make it likely that the capacity cost of a low variance market may be higher than a high variance market. This is an optimal choice by firms to minimize the cost of idle capacity and leads to savings for electricity customers.[84] Assuming the firm made the correct profit maximizing choice, the overall effect after considering marginal and capacity cost must be to increase electricity prices, but the capacity effect specifically is ambiguous and most likely very small. This is explained further in the estimation section that follows.

The identification of these four channels shows that this is a two-stage problem. The first consists of increasing demand leading to higher costs for firms. The second is the ability of firms to pass through these costs to consumers. This second stage is influenced by the presence of market power in electricity supply. Following Mansur (2008) studying the PJM market, if there exist firms that hold a strategic number of assets

[83] Greater load variance in a particular month may lead to major O&M expenditures in later months and are, therefore, hard to identify. Similarly, the nature of transmission investment may not coincide perfectly with increases in load variance. Therefore, the effect of load variance is likely underestimated.
[84] For example, a low variance state that builds coal plants will have higher capital costs than a high variance state that builds peaking Gas CT plants. However, when fuel and O&M costs are included, the average cost of providing electricity in the low variance state is lower.

in an electricity market and are net sellers of electricity, these firms will be able to manipulate market prices by withholding supply. In this context, the increase in costs will be passed through in the form of higher electricity prices. The ability to markup and its correlation with other factors in the analysis is discussed further in the next section.

3.5 Estimation

This chapter uses a panel reduced-form estimation strategy with time and region fixed effects to identify the effect of load variance on average retail electricity prices. The outcome of the previous section showed average retail electricity prices depend on the following factors: average load, load variance, plant age, past expectations of factors that influence average demand, past levels of average demand, past expectations of factors that influence load variance, past levels of load variance, past expectations of capacity ratios, past expectations about regulatory status, past estimates of plant levelized costs, fuel prices, O&M costs, market markups, distributional and retail costs, income, weather, season, environmental sensitivity, sector ratio and hour. Given the use of year and state fixed effects in this estimation strategy, time-invariant state differences and national trends are excluded, leaving only the following factors: average load, load variance, plant age and heat rates, fuel prices and sector. Assuming separability and linearity leads to the following reduced-form specification, where "r" represents the region of analysis and "t" the unit of time aggregation. The primary estimation observation level is state by month, with other region choices shown as alternative specifications later in this chapter. The estimating equation is as follows:

$$ln(EP_{rt}) = \beta\, ln(Var_{rt}) + \gamma X_{rt} + \theta_r + \omega_t + \varepsilon_{rt} \tag{22}$$

EP stands for electricity prices, Var is a measurement of state load variance, X is a set of controls, θ and ω represent state and year fixed effects respectively and ε is the idiosyncratic error term.

The identification strategy of this chapter combines the factors identified in the previous section with the insights of Holland and Mansur (2008), hereafter referred to as HM. The variable to be explained in HM was emissions, compared to electricity prices in this chapter, with each facing its own challenges. HM used six different statistics of load variance (Var), with their primary focus being on the coefficient of variation (COV). While COV is used in this chapter as an alternative specification, there is a significant problem with using this statistic in this analysis. What impacts the excess reserves available to utilities, which is where the majority of the costs are, is not the variance of load compared to the mean, but the max load compared to the mean.[85] This is the standard measure used in the industry and is frequently cited in EIA documents pertaining to changes in load.

A log-log model was chosen for this analysis, which is consistent with HM. Not only does the data generating process favor this model, but there is a clear interpretation of the results. States have varied electricity prices, meaning a unit level increase which may be large for Iowa is not large for Hawaii, making an elasticity interpretation best.

A problem common to both HM and this chapter is the existence of cross-border electricity flows. If not properly accounted for, these flows create measurement error, which can bias the results. HM approached this problem by summing observations at the NERC level and including bordering state weather measures to control for load variance changes in neighboring states. The approach of this chapter is discussed further in the following subsection. Due to concerns with endogeneity, average load, heat rate, plant age, restructuring and sector were included, with fuel prices added for precision.

[85] The reason for this is balancing authorities must maintain enough capacity on hand to meet max demand. It is possible for a market to have a large gap between max and mean, yet have a lower COV than a situation with a lower gap. This would be true of sharply peaking days and confuses the analysis.

3.5.1 Cross Border Flows and Observation Level

Our approach differs from HM not only in the variable of interest (average electricity prices instead of emissions) but also in the time frame and choice of region observation. HM limit their sample to monthly observations from January 1997 to December 2000. Choosing such a restricted period, while it creates fewer data issues described in Section 4, ignores the longer-term changes in the different US regions. This chapter relies on differences in load variance that occurred in states across time to identify the impact on electricity prices and a four-year period is too small to show many of those changes. Additionally, their choice of region size, while reducing the cross-border flows problem, also reduces the amount of variation necessary to identify an effect.[86]

This analysis consists of a panel of time and region observations. For the time observation, data on load variance and electricity price can be gathered at the day, month or year level. The advantage of observations at the daily level is it allows for more observations and variation to measure the change. However, the downside is that many of the control variables are not available at that time level, making it more difficult to precisely estimate the effect. There may also be a lot of noise in the data at the daily level and aggregating up diminishes the problem. Additionally, the daily data is only available for some regions, as it is reliant on wholesale market data that is only collected at certain hubs. These locations mostly coincide with regions containing restructured electricity markets. Therefore, this chapter measures time at the month level.

The smallest possible region possible is the state, while data are also available at the ISO and NERC levels.[87] Electricity flows across balancing authority, state, regional and national borders in large amounts. As a result, there is no observation level which prevents measurement error. At the highest level, the interconnect, the country is split into three regions (Western, Eastern and Texas). There is little electricity trading that

[86] This was not a problem for HM because their unit of time observation was the day, not month.
[87] Electricity prices are not available at the balancing authority, county or city level. Even if they were, identifying the resource mix and load variance at these levels is not feasible.

occurs across these boundaries, but also little variation for identification. The next level down, the NERC level, consists of 10 regions and is the most common observation level (Holland and Mansur 2008), but there still exists significant trading over these boundaries. There is even less inter-ISO electricity trading, but ISOs lack complete coverage of the US and limit the time scope. Therefore, the next feasible level down from NERC is the state level, which, as Table 9 shows, has significant cross-border flows. Although, in most cases, it is the less-populated states with large cross-border flows, this nevertheless poses a problem for estimation at the state level. The states at the top of the table have significantly greater demand for electricity (ERS) than supply (Gen), while the states at the bottom export electricity to other states due to the surplus in supply.

Table 9. 2014 Cross-Border Flows of Select States (GWh)

State	ERS	Gen	Difference
Massachusetts	54,469	31,119	-23,351
Maryland	61,684	37,834	-23,850
Idaho	23,233	15,184	-8,049
Delaware	11,338	7,704	-3,635
Virginia	112,098	77,137	-34,961
Wyoming	17,134	49,696	32,562
West Virginia	32,696	81,060	48,363
Montana	14,102	30,258	16,155
North Dakota	18,240	36,463	18,223
New Hampshire	10,944	19,538	8,594

Source: EIA, 2017

The primary concern of cross-state border flows for this chapter is not omitted variable bias, as was present for HM, but measurement error. The inter-state activities of many balancing authorities make it difficult to precisely measure the amount of electricity being consumed in each state. Therefore, these uncertain cross-state flows create measurement error in the primary explanatory variable. HM has an additional problem because, unlike electricity prices, changes in demand in a region do not always result in emissions changes. For example, when a state has higher load variance in the summer, it may be more likely to import electricity from other states. This does not change emissions in the state, as the electricity is being produced in another state. The change in load variance does impact the electricity price in the state, however, as the electricity provider is paying for the electricity transfer across state lines. Therefore, load variance changes within a state are accounted for by changes in state emissions. As a result, this chapter does not employ the HM method of controlling for neighboring variance, but rather uses ownership and PPA agreements between utilities to identify the source of cross-state border flows and adjust for them.

3.5.2 Endogeneity and Fixed Effects

Threats to identification come in three forms. First, there are time-invariant state differences that are correlated with state load variance.[88] Second, there are time-varying factors that influence states and are correlated with state load variance.[89] Third, there are time-varying state specific factors that are correlated with state load variance. Including state and year fixed effects focuses this chapter on the third threat to identification.

As shown in Figure 19, the greatest variation in load variance occurs between states and seasons within states. The choice in this chapter of eliminating between state comparisons is due to the endogeneity concerns present without state fixed effects. Therefore, this analysis will be relying particularly on within state differences in seasons and across years. Of the number of factors that vary within states over time, average load, heat rate, plant age, restructuring and sector were present in the modeling process as threats to identification.

[88] Examples include climate and resources which both influence load variance and price through the availability of low marginal cost resources like hydro, wind and solar.

[89] Example: national energy efficiency programs aimed at reducing both load variance and average load, which impacts average prices.

Average load changes impact electricity prices, as they require utilities to construct new capacity, increasing costs for consumers. These changes are also correlated with load variance, as new population growth and other demand factors increase the daily average load requirements as well as their structure. Newer plants with lower heat rates will often lower electricity prices, due to their increased efficiency, and may also be correlated with states that have to meet increases in load variance, as these states are more likely to build new plants. States with a higher share of ERS from industrial users, compared to residential, will have lower demand variance, as this lowers demand at the peak and raises it during the other hours. These states will also have a different price structure, resulting in concerns about endogeneity. As cited earlier in this thesis, there is little evidence that restructuring is correlated with electricity prices, after controlling for state fixed effects, but the potential correlation with load variance warrants inclusion. There is no correlation between changes in state load variance and fuel prices, as fuel prices are strongly correlated with national trends. However, fuel prices are a significant component of electricity prices in many markets, so they are included for precision. Controlling for average load, plant heat rate and age, and state and year effects, the effect of load variance on electricity prices is taken as exogenous.

3.5.3 Lagged Effects

Changes in load variance affect electricity prices in the current and future periods. In the current period, the shift in demand increases fuel, O&M and transmission expenditures, as well as prompting new capacity expenditures. These changes also impact future capacity decisions, which are partly based on historical load variance. By not including lags, this chapter is focused only on the former and not the latter effects. There are several reasons for this. First, the effect is small due to the number of factors involved in firm expectations and the ambiguous cost impact from cost-minimizing firm capacity decisions described earlier in this section. Second, the timing of these lagged effects is unknown and heterogeneous, with the type of market, rules and frequency of rate cases creating differences among states. Finally, a significant part of these expenditures are reflected in current electricity prices through firm foresight. Given the lack of a correlation between past load variance and current price factors, not including lagged effects will not harm this estimation.

4. Data

The principle dataset used this chapter is the hourly load data by balancing authority available from FERC Form 714. This chapter combines the 714 data with information on balancing authorities to create a unique state-level monthly dataset stretching from 1993-2014. This dataset has not been widely used, with the few papers that have, like Holland and Mansur (2008), using only a subset of years and often combining at a higher level. This data has not been utilized due to gaps in reporting and the difficulty in attributing load by state.[90]

The initial approach of this chapter was to use utility reported lambdas to fill in the gaps. [91] However, comparisons between lambdas and actual demand data for balancing authorities that had both revealed very different variance patterns. Choosing to bypass lambdas, data sets were added outside FERC and several states dropped that had significant missing data.[92] Balancing authorities were grouped into state and NERC observations based on input from contacts at each of the NERC and balancing authority organizations as well as the EIA. Given that the focus is not total demand, but variance, I collected data from balancing authorities that were representative of load in the state.

---

[90] For example, Entergy data is often reported as a single entity, which covers parts of Louisiana, Arkansas, Texas and Mississippi.

[91] Lambdas are an estimate of the marginal cost of production in that hour. Therefore, they are a proxy measure for load. In this chapter, I have not used lambdas, as they are dependent on the resource mix available, and, therefore, may differ from load.

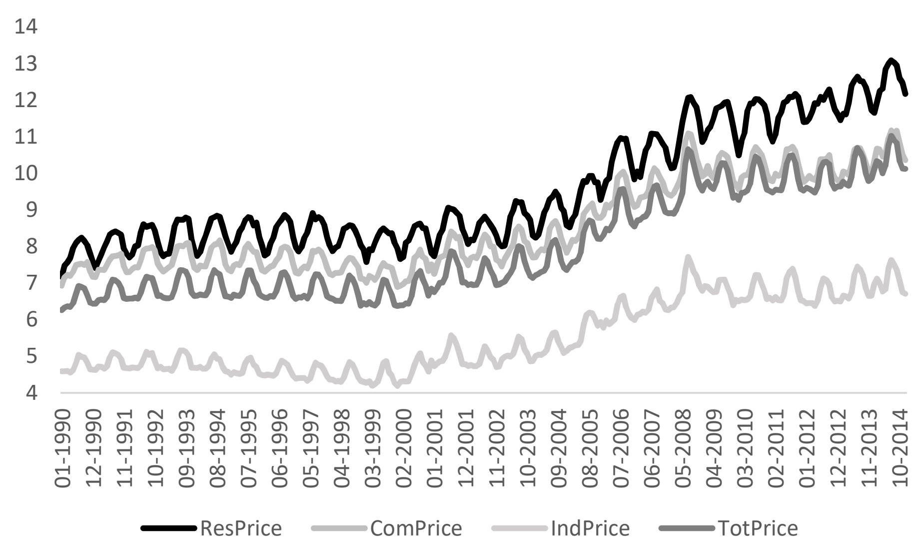
[92] States dropped: Kansas, Louisiana, North Dakota, Oklahoma, Wyoming

Data on electricity prices, generation mix and fuel costs are available from the EIA and Ventyx. Heat rates and plant age were gathered from EIA form 860 and combined at the state and NERC level. Data available at the state level was grouped at the NERC level by weighting each state observation according to its share of electricity use.

4.1.1 Electricity Prices

Figure 22 shows monthly US electricity prices by sector. There are several trends worth noting. First, there is strong seasonality in each of the sectors, with the peaks corresponding to the summer months and the valleys occurring in the winter. Second, industrial prices are well below the other two sectors, with residential consistently the highest. Third, electricity prices were relatively constant until the mid-2000s, when they increased substantially before stabilizing in 2009. Fourth, while commercial and industrial prices stayed level for the past 5 years, residential prices continued to rise.

Figure 22. US Electricity Prices by Sector



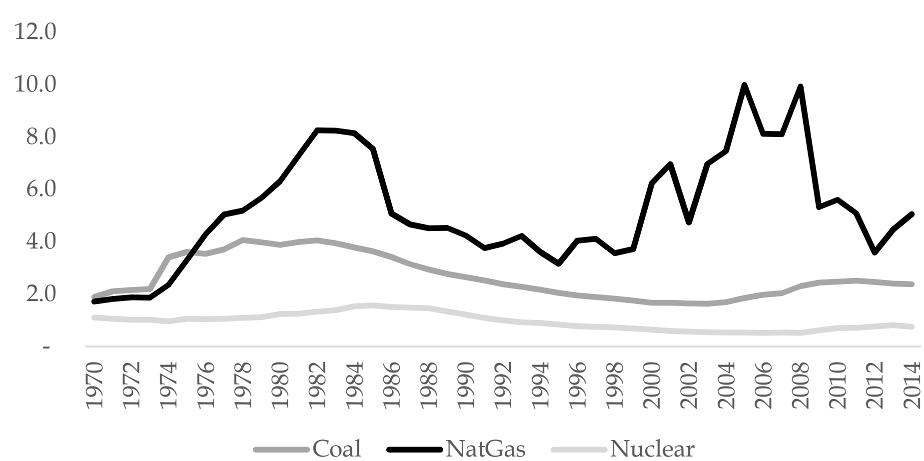
Notes: US electricity prices are measured in cents/kWh. Source: EIA 2017

Price seasonality occurs in each sector, but is most pronounced in the residential sector. The seasonal pattern is strongly correlated with increases in average demand as well as load variance, which increase in the summer months and decrease in the winter (See Figure 24). Taken together, these two statements suggest residential users pay a larger share of the increase in load variance than do commercial or industrial customers.

The gap between electricity prices in the three sectors is due primarily to two factors. First, industrial users are large and negotiate lower prices with utilities, with some also possessing their own power generating facilities (EIA 2004). Second, demand is highest when residential use increases. Industrial users both operate in the middle of the day, when demand is lower, and throughout the night, when demand is lowest.

The change in electricity prices in the mid-2000s was largely due to rising natural gas prices (EIA 2008). Prior to 2005, US fuel prices were low, with coal and nuclear maintaining a large share of generation. The natural gas construction boom shifted the portfolio mix of US utilities towards a greater reliance on gas, which experiences greater price volatility. Figure 23 shows the rapid increase in natural gas prices during the 2000s, which peaked in 2008 before dropping off significantly. This was a major factor in the electricity price increase in the mid-2000s, with natural gas prices continuing to be almost double the level of the late 1980s and 1990s even after the 2008 drop.

Figure 23. US Delivered Fuel Prices to the Electricity Sector (2014 $)



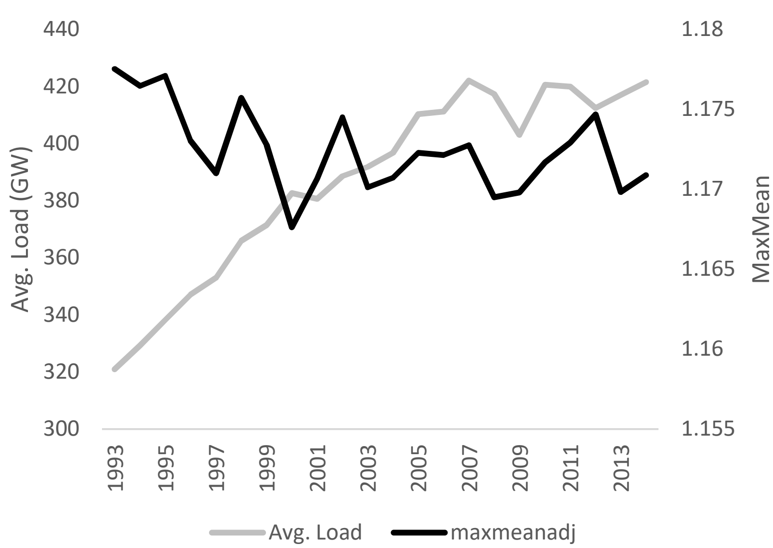
Notes: Natural gas prices are measured in $/mmBtu. Source: EIA, 2017

Increasing residential electricity prices after the natural gas price crash in 2008 are an anomaly when compared with the rest of the events in this period. Utility fuel costs were falling and commercial and industrial prices were relatively constant.

4.1.2 Electricity Demand

Figure 24 shows several important trends. First, US average load in 2014 was approximately equivalent to 2007. While in general load growth was slower in the 2000s than the 1990s, and slower in the 1990s than the 1980s and 1970s, the drop starting after 2007 is a significant break in trend. There are three explanations for this decline. The first is the substantial decrease in economic activity from 2008 to 2009 due to the recession. A smaller dip can be observed in 2001, which was another US recession. Electricity demand is linked to economic activity, with households and firms varying their electricity demand based on changes in their income.

Figure 24. U.S. Average Load and Max-Mean Ratio



Source: FERC, 2017 and Author's Calculations

This relationship provides support for the second explanation as well, which is slow GDP growth since the end of the recession has resulted in slow demand growth.

Average annual GDP growth since the end of the recession has been 2.1 percent, compared to 2.7 percent in the previous expansion from 2002-2007 and 3.8 percent from 1992-2000. As noted, falling electricity demand is a U.S. trend, with GDP growth and energy intensity per unit of GDP falling since the 1970s.

The third explanation is energy efficiency measures have had a significant impact on electricity demand (AEO 2015). The early improvements were largely in air conditioners and swamp coolers, both by replacing inefficient window units with central air as well as improving the efficiency of all units. Those improvements spread to other appliances, particularly electronics, which used significant amounts of electricity when they started to saturate U.S. households and businesses in the late 1990s. While not all these efficiency improvements occurred in the past decade, the improvements in the 2000s were obscured by rising electronics use and an increase in the average home size, which is correlated with increased electricity use.

US load variance has dropped slightly over the past 20 years. This is an initially surprising result, given the increased use of electronics and air conditioning over this time period, as well as the general warming of temperatures. For example, 2014 was the hottest year on record (NOAA, 2015), yet US load variance was lower than throughout the 1990s. While some of this is due to programs around the US aimed at reducing electricity usage in peak summer times, the majority is due to greater adoption of central air as well as increases in its efficiency. An air conditioner purchased in 2015 used between 30 to 50 percent less electricity than in 1980 (EIA 2015).

4.1.3 State-level Characteristics

Table 10 separates the states in this chapter into three groups based on electricity demand variance in 1993. The purpose of this section is two-fold. First, the table shows differences between the state groups, highlighting potential threats to exogeneity. Second, the statistics show factors that explain differences in variance between the states, providing an explanation for the results shown in the next section. The table highlights several interesting differences between groups.

Table 10. Summary Statistics by State Variance Group

	LowVar			MedVar			HighVar		
	1993	2014	%Δ	1993	2014	%Δ	1993	2014	%Δ
Restruct		51.3%			38.4%			20.0%	
ResPrice	12.60	12.02	2.1%	12.86	12.38	2.1%	15.35	15.52	2.3%
ComPrice	10.99	9.77	1.7%	11.64	10.51	1.8%	13.89	13.29	2.1%
IndPrice	7.20	6.78	2.0%	8.25	7.62	1.9%	10.64	10.55	2.2%
ElecPrice	10.01	9.55	2.1%	10.81	10.38	2.1%	13.41	13.46	2.3%
AvgCap	13.7	29.4	3.5%	16.0	34.3	3.5%	12.6	31.9	4.3%
PlantAge	23	21	-0.4%	23	22	-0.1%	21	19	-0.5%
PlantHR	10.9	10.2	-0.3%	10.7	10.1	-0.3%	10.6	9.3	-0.6%
GasPrice	$4.44	$5.42	3.2%	$4.10	$5.20	3.4%	$4.26	$5.47	3.4%
Cool Days	774	851	0.4%	992	1,026	0.2%	1,448	1,547	0.3%
PersInc	$33.1	$44.3	3.6%	$33.4	$45.2	3.7%	$36.5	$47.1	3.5%
Pop	5.04	5.60	0.5%	4.65	6.12	1.3%	6.01	7.78	1.2%
ERS	5,491	6,829	1.0%	6,364	8,623	1.4%	5,476	7,459	1.4%
Avg Load	6,270	7,800	1.0%	7,266	9,848	1.4%	6,251	8,514	1.4%

Notes: Restruct: percent of ERS in restructured markets. Electricity prices measured in cents/kWh. AvgCap: Average amount of generating capacity in each state. PlantHR: Average heat rate of plants per state (mmBtu content of kWh). Gas prices are $/mmBtu. Cool days: Average number of cooling degree days per state. PersInc: Average state personal income ($ thousand). Pop: Average state population (millions). ERS: Average state electricity retail sales (GWh). Average Load is in GW.

While electricity prices are similar between the low and medium variance group, the high variance group pays significantly higher prices, despite there being very little difference in the price of natural gas, average load, total electricity retail sales and personal income. Additionally, average plant age and heat rate are lower, implying the plants in these states are more efficient. The one significant difference is the number of cool days. On average, the high variance group has almost double the number of cool days compared with the low variance group. This means the high variance states have a much larger number of days when their residents would be running their air conditioners, increasing demand variance and, by extension, electricity bills.

The one other difference that stands out between the groups is the makeup of the plants in the high variance group. High variance states experienced a greater increase in capacity from 1993-2004, which lowered their age and heat rate. The increase in new capacity construction can be interpreted as a response to higher variance in load. However, despite the increase in efficiency, these states continued to have higher electricity prices than the other two groups.

5. Data Analysis

This section presents the results from two analyses: the impact of load variance on state electricity prices and the sectoral burden (residential, commercial, industrial) followed by a series of alternative specifications. The first results provide a first estimate of the magnitude of the problem of electricity congestion. The second set of results identifies the user burden and provides a comparison to a system where price can be used to shift demand.

5.1 Load Variance Effect on Electricity Prices

Table 11 shows the effect of changes in load variance on state electricity prices from 1993-2014. The variable of interest in this equation is total average state electricity prices (logged). The primary explanatory variable is the most commonly used industry measure of load variance, the max-mean ratio (logged).[93] Included controls are average plant heat rate (HR), average plant age (PlAge), average daily load (Load), the average weighted price of fuel (Fuel), the ratio of industrial to residential sales (IndRatio) and a measure of whether the market is restructured (restruct). Five different specifications are presented, with the preferred specification (5) including the entire set of controls, year and state fixed effects.

[93] Results using coefficient of variation and max-min ratio, two other popular measures of electricity demand variance, are shown in Section 6 of this chapter.

In the preferred specification (Column 5 of Table 11), a state with a 10 percent increase in load variance, on average, faced 9.7 percent higher electricity prices. This effect is both large and robust to a number of specifications and tests (see the following section for robustness checks). Given the downward trend in load variance during the time period of this study, these results should be interpreted as being driven by reductions in load variance and prices. This result is a factor in the slow rise in real electricity prices from 1993-2014.

Table 11. Impact of Load Variance on Electricity Prices

	(1)	(2)	(3)	(4)	(5)
lmaxmean	0.435***	0.495***	0.96***	0.958***	0.974***
	(0.067)	(0.050)	(0.033)	(0.142)	(0.142)
HR		-0.177***	-0.001	0.001	-0.004
		(0.004)	(0.003)	(0.030)	(0.030)
PlAge		-0.003***	0.004***	0.004	0.004
		(0.001)	(0.001)	(0.003)	(0.002)
Load		0.002***	-0.004***	-0.004	-0.004
		(0.001)	(0.001)	(0.003)	(0.003)
Fuel		0.032***	0.023***	0.023***	0.023***
		(0.001)	(0.001)	(0.006)	(0.006)
IndRatio		-0.138***	-.061***	-0.061***	-0.063***
		(0.006)	(0.004)	(0.020)	(0.019)
Restruct					-0.01
					(0.006)
State FE	N	N	N	Y	Y
Year FE	N	N	Y	Y	Y
R2	0.1	0.25	0.44	0.44	0.40

Notes: Dependent variable is total electricity price (logged). * significant at 10 percent level, ** significant at 5 percent level. *** significant at 1 percent level.

The results in Table 11 highlight several interesting facets of the problem of estimating the impact of load variance. First, while the direction of the effect was likely to be positive, the magnitude validates the importance of efforts to reduce load variance by states. Second, the inclusion of year fixed effects both corrects for a substantial downward bias in the estimate and significantly increases the amount of variation explained. Column 2, which excludes year fixed effects, shows a significantly lower impact, showing the downward bias effect of excluded endogenous national trends. Third, restructuring does not have a significant effect on load variance influencing electricity prices, as shown in the results in Table 11 and specifications interacting restructuring with load variance, which were insignificant.

Additionally, the two factors initially believed to be most influential in electricity prices, load variation and fuel costs, were consistently influential. Utilities under COS regulation are provided allowances for changes in fuel prices and wholesale markets are very sensitive to changes in fuel prices. As a result, electricity prices are strongly correlated with changes in fuel prices, particularly the price of natural gas.[94] Average load is not, which is not surprising, as states with a higher load do not necessarily require more expensive generation. If a state's electricity demand is twice as large as its neighbor, it can build double the number of coal plants without the plants being any more costly.

5.2 Sector Analysis

Electricity users are divided into three sectors: residential, commercial and industrial. Each sector is responsible for approximately one-third of electricity retail sales in the US, with significant variation between the sectors in time of use. Residential use is a primary driver of load variance within a day (EIA 2006). System load peaks in the morning and evening hours, as families get ready for work and school and then return back home. Each state, therefore, experiences a peak in load from 7 to 9 AM each morning and 4 to 8 PM each evening. Load reaches its lowest point overnight, as homes use little electricity. Commercial use is largely during business hours, while industrial use

[94] The increase in natural gas prices from 2005-2008 resulted in a substantial increase in wholesale electricity prices during this time period (EIA 2008).

continues throughout the day and evening. Large production companies will often run night shifts to fully utilize their capital investment.

Table 12 shows the impact of load variance on electricity prices by sector user. The first column is identical to column 4 in the previous table to compare with the sector results. Column two shows the impact of load variance on residential prices, column 3 shows the impact on commercial prices and column 4 the impact on industrial prices. Each of these is presented with the full set of controls and fixed effects, as this is the preferred specification of this chapter.

Table 12. Electricity Price Impact of Load Variance by Sector

	(1)	(2)	(3)	(4)
lmaxmean	0.958***	1.043***	0.707***	1.06***
	(0.142)	(0.169)	(0.166)	(0.158)
HR	0.001	0.002	0.003	-0.016
	(0.030)	(0.029)	(0.029)	(0.030)
PlAge	0.004	0.003	0.004	.009***
	(0.003)	(0.002)	(0.003)	(0.003)
Load	-0.004	-0.004	-0.005	-0.003
	(0.003)	(0.004)	(0.003)	(0.003)
Fuel	0.023***	0.02***	0.023***	0.027***
	(0.006)	(0.006)	(0.006)	(0.006)
IndRatio	-0.061***	.072***	0.020	-.055*
	(0.020)	(0.019)	(0.017)	(0.032)
State FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
R^2	0.44	0.35	0.34	0.38

Notes: Dependent variables are logged total price (1), residential price (2), commercial price (3), industrial price (4). * significant at 10 percent level, ** significant at 5 percent level. *** significant at 1 percent level.

The results from the first two columns are consistent with the data shown in Section 4. Residential prices experienced the most volatility and Table 12 provides confirmation, with residential prices increase 10.4 percent with every 10 percent increase in load variance. Likewise, commercial users are impacted less by load variance, with a

7.1 percent increase in prices with every 10 percent increase in load variance. These are consistent with a story of smaller users lacking market power and paying higher prices. However, the findings in column 4 do not follow this trend. Given the access of some industrial users to their own power supplies during peak times and their negotiated contracts, it was expected that industrial users would not be as greatly impacted by load variance. Column 4 suggests otherwise, with industrial prices increasing 10.6 percent for every 10 percent increase in load variance. This result may suffer from some selection bias, as not all industrial users pay for electricity, due to onsite generation, and the insignificant difference between the sectors is not sufficient to confirm that residential and industrial users pay more than commercial users.

As the popularity of demand reduction programs increases across the US, it is worth noting that the findings in Table 12 are consistent with this type of effort. The users who have the most flexibility in shifting their electricity use to off-peak overnight hours are residential and industrial. Residential users can run washers and dishes at night and program temperatures to be warmer during the day while industrial users can shift electricity-intensive production to night shifts. Commercial users are not as elastic. Their greatest electricity uses are lights and air conditioning during business hours, which precludes shifting electricity use to later in the evening.

5.3 Discussion

The findings in Tables 11 and 12 show that load variance significantly impacts electricity prices. As Figure 24 shows in Section 4, load variance has fallen across the US, which has been a factor, along with low fuel prices, of real electricity prices remaining relatively constant for 20 years. Therefore, not only are there potential environmental benefits to lowering load variance in some regions (Holland and Mansur 2008), but there is a more fundamental gain for consumers in the form of lower electricity prices. How can this be achieved?

Before discussing potential solutions, it would be naïve to assume that all states have the opportunity to achieve low load variance. When comparing load variance within a day and across seasons, the impact of climate is apparent. The states with lowest load variance, such as Maine, Ohio and South Dakota, do not face the same electricity demand as those with high variance, like Florida, Arizona and Nevada. There is no policy solution to give Florida the weather of Maine, so there are limits to how much load variance can be reduced.

With the inherent problems faced by the electricity industry, it is worth noting what an ideal system would look like. Assuming that congestion is the natural state of the system, due to the desire of households and firms to use power in peak periods, there are several changes necessary to reduce the problems previously identified.

First, full adoption of RTP will allow decision-makers to act on correct prices. In the current system, users face average prices that, in some cases, vary based on tiers. A person running their dishwasher at peak time pays the same price as someone that runs it in the middle of the night, even though the social marginal cost of these two electricity uses are very different. Users who face correct prices will be able to shift some electricity-using activities to non-peak periods, like the middle of the night. However, there are some activities, like clothes washing, TV viewing, computer use, and air conditioning that users will still engage in during peak use periods. By itself, therefore, RTP leaves significant cost problems with peak demand met by idle capacity.

A solution to idle capacity is storage of electricity. Electricity is stored to use plants that are lowest cost, running surpluses overnight and shortages during the day. With storability, the system would be able to shift away from idle, high cost oil and gas CT plants towards lower cost nuclear and coal plants. This would also allow for an expansion of intermittent, renewable sources like wind, which is expensive but often does not provide electricity during peak demand periods. Possible methods currently in use or being considered are pumped hydro, molten salt, battery and compressed air storage.

However, none of these are currently in large-scale usage, outside of molten salt in concentrated solar thermal (CST) plants (NREL 2017). Even if these technologies matured and provided economical storage, this is still an additional cost imposed.

Currently, there are three ways states can reduce their load variance cost. The first is to replace their remaining old steam oil plants with newer, more efficient gas CT plants. Second, they can encourage electricity providers to develop and implement creative methods of charging consumers real-time prices.[95] Third, states can continue to encourage energy efficiency measures. While the economics literature has found actual cost savings from energy efficiency measures smaller than those of engineering estimates, simple measures like changing to LED lightbulbs and installing insulation are still very cost efficient ways of reducing demand (Fowlie et al 2015). While none of these will lead to a truly flat electricity load profile, the success of the first and third solutions in the past 20 years at reducing load variance are evidence that reducing load variance is possible.[96]

5.4 NERC Level Analysis

The choice of unit of observation in this chapter involved a tradeoff between power and measurement proficiency. The previous analysis was at the state level, favoring the added variation of more observations over cross-border flow concerns. This section performs this analysis at the NERC level to test for measurement error in the previous specification. As with the state-level analysis, there are challenges to choosing NERC regions as the unit of observation. In 2005, NERC regions were reorganized which results in an unbalanced panel. For this analysis, areas that had been organized into a NERC region pre-2005 continued to be in that region through 2014.

---

[95] Papers on RTP, such as Alcott (2011), have outlined various ways utilities have experimented with methods to encourage people to react to marginal cost pricing, including blinking orbs and online alerts.
[96] A flat load profile is not ideal in this case, as consumers have a high value on electricity usage during prime periods (Alcott 2011).

Table 13 presents results for the same specifications as Table 11, with the only difference being the change in the unit of observation and removal of restructuring due to the level of observation. Comparing these results with those in Table 11, the findings are very similar. A 10 percent increase in load variance increases electricity prices by 9.59 percent, which is almost identical to the previous findings. There is also a significant gap present in these results between specifications with and without year fixed effects. This shows the significant downward bias present by not controlling for national trends during this time period. Overall, these results confirm what was found in the state analysis, that load variance in an electricity system imposes large costs.

Table 13. NERC Load Variance Impact on Electricity Prices

	(1)	(2)	(3)	(4)	(5)
lmaxmean	0.241	0.327***	1.0***	0.272	0.959***
	(0.157)	(0.101)	(0.055)	(0.259)	(0.256)
HR		-0.039***	0.032***	-0.079**	0.031***
		(0.004)	(0.003)	(0.031)	(0.008)
PlAge		-0.011***	0.008***	-0.008	0.009*
		(0.001)	(0.001)	(0.010)	(0.005)
Load		0.002*	-0.001***	0.004	0.001
		(0.001)	(0.001)	(0.004)	(0.001)
Fuel		0.062***	0.033***	0.057***	0.033***
		(0.001)	(0.001)	(0.004)	(0.003)
IndRatio		-0.089***	0.003	-0.059***	0.007
		(0.014)	(0.008)	(0.068)	(0.022)
State FE	N	N	N	Y	Y
Year FE	N	N	Y	N	Y
R2	0.07	0.16	0.65	0.25	0.62

Notes: Dependent variable is total electricity price (logged). * significant at 10 percent level, ** significant at 5 percent level. *** significant at 1 percent level.

As in the previous analysis, the impact of load variance on electricity prices is disaggregated by sector to uncover which sector is paying the most for load variance. These results are presented in Table 14, with Column 1 is identical to Column 5 in Table

5, Column 2 analyzes residential prices, Column 3 commercial, and Column 4 industrial. The findings are consistent with the state-level sector analysis. Residential and industrial users are impacted more by load variance than commercial users, although the gap in this case is smaller than the state-level findings.

Table 14. NERC Electricity Pirce Impact by Sector

	(1)	(2)	(3)	(4)
lmaxmean	0.959***	0.997***	0.856***	0.921***
	(0.256)	(0.291)	(0.266)	(0.214)
HR	0.031***	.03***	.028**	0.031***
	(0.008)	(0.008)	(0.010)	(0.010)
PlAge	0.009*	0.01	0.008	0.015***
	(0.005)	(0.006)	(0.005)	(0.003)
Load	0.001	0.002	-0.002	0.003
	(0.001)	(0.002)	(0.002)	(0.002)
Fuel	0.033***	0.027***	.034***	0.037***
	(0.003)	(0.003)	(0.003)	(0.004)
IndRatio	0.007	0.116***	0.096***	0.042*
	(0.022)	(0.029)	(0.027)	(0.022)
State FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
R2	0.62	0.58	0.65	0.59

Notes: Dependent variables are logged total price (1), residential price (2), commercial price (3), industrial price (4). * significant at 10 percent level, ** significant at 5 percent level. *** significant at 1 percent level.

Table 14 shows the results of alternative LHS and RHS variables, with all the controls and fixed effects in the preferred specifications above. Column 1 shows the main specification in levels instead of logs. An increase of 0.1 in load variance, on average, leads to a 0.6 cent/kwh rise in electricity prices. This is a large increase, the equivalent of making Michigan like Florida, so it is not surprising that it has a large effect. However, this effect is difficult to interpret, as it is small for states with high electricity prices, like Hawaii and California, and large for states with low electricity prices, such as Alabama.

Column 2 tests the fit of this relationship as well as a query from the introduction. It was proposed that states with high load variance may benefit more from a reduction in load variance than those states with more flat load profiles, due to dispatch curve convexity. If this were true, the relationship would appear quadratic, with the price impact increasing in load variance. However, as Column 2 shows, there is no clear relationship in the quadratic function, with higher orders yielding the same result.

Table 15. Alternative Specifications

	(1)	(2)	(3)	(4)
maxmean	6.621***	42.619		
	(1.071)	(30.594)		
maxmean2		-15.279		
		(12.999)		
COV				7.454***
				(1.287)
maxmin			1.186***	
			(0.380)	
HR	-0.035	-0.036	-0.023	-0.016
	(0.239)	(0.238)	(0.238)	(0.238)
PlAge	0.103**	0.103**	0.105**	0.104**
	(0.041)	(0.041)	(0.042)	(0.041)
Load	-0.048**	-0.044	(0.039)	(0.043)
	(0.036)	(0.036)	(0.034)	(0.034)
Fuel	0.529**	0.528**	0.538**	0.537**
	(0.223)	(0.223)	(0.224)	(0.224)
IndRatio	-0.192	-0.193	-0.312**	-0.335**
	(0.151)	(0.152)	(0.152)	(0.154)
State FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
R2	0.460	0.46	0.480	0.480

Notes: Dependent variable is total electricity prices (level) in each column. * significant at 10 percent level, ** significant at 5 percent level. *** significant at 1 percent level.

Columns 3 and 4 replace max-mean with two alternative variance statistics: coefficient of variation (COV) and max-min. Max-mean was chosen both because of its prevalence in the industry and potential misleading situations when using COV. The COV result is similar to the max-mean finding, with the small difference mostly due to slight differences in the spread of these variables. However, the max-min ratio is very different, reflecting the much greater standard deviation of the variable. COV had a sample standard deviation of 0.036, similar to the 0.042 of max-mean. The sample standard deviation of max-min is 0.148, so this finding should be interpreted differently. An increase of 0.1 units of the max-mean ratio is similar to an increase in the max-min ratio of approximately 0.4 units. Adjusting for this, the max-min finding is still smaller than the other two ratios but much closer in magnitude. Both confirm a large effect of load variance on electricity prices.

6. Conclusion

Congestion costs are common to many US industries where consumer demand for a product or service is not uniformly distributed across time. The solution, in most industries, is to use prices to smooth demand, with higher prices in the peak and lower prices off-peak. Electricity is one of the US industries with high congestion costs. Demand for electricity peaks both within days and across seasons. Lack of electricity storage at a large level and ability to charge people the marginal cost of production lead to significant costs. The increase in cost comes from plants sitting idle and the use of low efficiency, high fuel cost peaking generation. This chapter is the first to estimate these costs.

This chapter finds an increase in load variance of 10 percent leads to a rise in electricity prices of 9.7 percent. This large result is similar to making a state like Florida, with high load variance, more like Michigan, with lower than average load variance. The magnitude of this result is consistent through a series of alternative specifications and changes in the observation level. Furthermore, it holds when disaggregated by sector, with residential and industrial customers paying more than commercial users when load variance is higher. This is consistent with the elasticity of demand with respect to load variance being higher for residential and industrial than commercial users.

The time period of this analysis, 1993-2014, captured significant changes in load variance within states. There was a steady decline in load variance during the 22 years of this study, which coincides with improvements in appliance efficiency and the large increase in higher efficiency gas-fueled peaking generation. While these gains are not exhausted, future reductions in electricity industry congestion costs will rely on improvements in RTP infrastructure and ability to store large amounts of electricity.

Given the short-run nature of the analysis, this result can be considered a lower bound on the cost of load variance. A change in load variance in a year causes utility planners to alter their plans for future generation and capacity planning. Therefore, there are likely lagged effects not captured in this chapter. Future work is needed to be able to link firm decisions about future investments with changes in load variance in an environment of incomplete information as well as what has caused load variance and congestion costs to decline consistently in the US over the past two decades.

REFERENCES

Abadie, Alberto, Diamond, Alexis, and Hainmueller, Jens. 2010. "Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program." *Journal of the American Statistical Association*, 105: 493-505

Abadie, Alberto and Gardeazabal, Javier. "The Economic Costs of Conflict: A Case-Control Study for the Basque Country." *American Economic Review*, 93(1): 113-132

Alavalapati, Janaki, Carter, Douglas, Huang, Ming-Yuan, and Langholtz, Matthew. 2007. "Is the Choice of Renewable Portfolio Standards Random?" *Energy Policy*. 35(11): 5571-5575

Alberini, Anna, Gans, Will and Velez-Lopez, Daniel (2011). "Residential Consumption of Gas and Electricity in the U.S.: The Role of Prices and Income." *Energy Economics* 33: 870-881

Allcott, Hunt. 2011. "Rethinking Real-Time Electricity Pricing." *Resource and Energy Economics* 33 (4): 820-842

American Electric Power. 2016. https://www.aep.com/

Anderson, Leonard and Erman, Michael. "Calpine Files for Bankruptcy." *USA Today* 21 Dec. 2005: USAToday.com Web. 30 Mar. 2016

Ando, A and Palmer, Karen. 1998. Getting on the Map: The Political Economy of State Level Electricity Restructuring. Discussion Paper 98-19rev. Washington, DC: *Resources for the Future*.

Apt, Jay. "Competition Has Not Lowered U.S. Industrial Electricity Prices." *Electricity Journal* 18(2): 52-61

Ardoin, Phillip and Grady, Dennis. 2006. "The Politics of Electricity Restructuring Across the American States: Power Failure and Policy Failure." *State & Local Government Review* 38(3): 165-175

Arnott, Richard, De Palma, Andre and Lindsey, Robin. 1993. "A Structural Model of Peak-Period Congestion: A Traffic Bottleneck with Elastic Demand." *American Economic Review* 83 (1): 161-179

Atkinson, S. E. and Halvorsen, R. 1976. "Interfuel Substitution in Steam Electric Power Generation." *Journal of Political Economy* 84(5): 959-78

Bannerjee, A. 1992. "A Simple Model of Herd Behavior." *Quarterly Journal of Economics*, 107: 797-817

Barbarino, Alessandro and Jovanovic, Boyan. 2007. "Shakeouts and Market Crashes." *International Economic Review*, 48(2): 385-420

Barmack, Matthew, Kahn, Edward, Tierney, Susan and Goldman, Charles. 2008. "Econometric Models of Power Prices: An Approach to Market Monitoring in the Western U.S." *Utilities Policy* 16 (4): 307-320

Berger, Allen and Dick, Astrid. 2007. "Entry into Banking Markets and the Early-Mover Advantage." *Journal of Money, Credit and Banking*, 39(4): 775-807

Bernanke, Ben. (2005). "The Global Savings Glut and the U.S. Current Account Deficit." *The Federal Reserve Board*: 10 March 2005

Berry, Steven and Waldfogel, Joel. 1999. "Free Entry and Social Inefficiency in Radio Broadcasting." *RAND Journal of Economics*, 30(3): 397-420

Berzel, Yoram. 1964. "The Production Function and Technical Change in the Steam Power Industry." *Journal of Political Economy* 72(2): 133-150

Bikchandani, S, Hirschleifer, D, and Welch, I. 1992. "A Theory of Fads, Fashion, Custom and Cultural Change as Information Cascades." *Journal of Political Economy*, 100: 992-1026

Bohn, Sarah, Lofstrom, Magnus, and Raphael, Steven. 2014. "Did the 2007 Legal Arizona Workers Act Reduce the State's Unauthorized Immigrant Population?" *The Review of Economics and Statistics*, 96(2): 258-269

Bolton, Patrick and Farrell, Joseph. "Decentralization, Duplication and Delay." *Journal of Political Economy*, 98(4): 803-826

Borenstein, Severin and Bushnell, James. 2000. "Electricity Restructuring: Deregulation or Reregulation?" *Regulation* 23(2): 46-52

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

Borenstein, Severin and Bushnell, James. 2016. "The U.S. Electricity Industry after 20 Years of Restructuring." *Haas Institute of Energy working paper*

Borenstein, Severin, Bushnell, James and Wolak, Frank. 2002. "Measuring Market Innefficiencies in California's restructured Wholesale Electricity Market." *American Economic Review* 92 (5): 1376-1405

Borenstein, Severin and Holland, Stephen. 2005. "On the Efficiency of Competitive Electricity Markets with Time-Invariant Retail Prices." *RAND Journal of Economics* 36 (3): 469-493

Buchmueller, Thomas, DiNardo, John and Valleta, Robert. 2011. "The Effect of an Employer Health Insurance Mandate on Health Insurance Coverage and the Demand for Labor: Evidence from Hawaii." *American Economic Journal: Economic Policy*, 3(4): 25-51

Bulow, Jeremy and Klemperer, Paul. 1999. "The Generalized War of Attrition." *American Economic Review*, 89(1): 175-189

Bushnell, James and Ishii, Jun. 2007. "An Equilibrium Model of Investment in Restructured Electricity Markets." *UCEI CSEM Working Paper #164*

Bushnell, James, Mansur, Erin, and Saravia, Celeste. 2008. "Vertical Arrangements, Market Structure and Competition: An Analysis of Restructured Electricity Markets." *American Economic Review* 98(1): 237-266

Bushnell, James and Wolfram, Catherine. 2005. "Ownership Change, Incentives and Plant Efficiency: The Divestiture of U.S. Electric Generation Plants. *CSEM Working Paper 140*

Caballero, Ricardo, Farhi, Emmanuel and Gourinchas, Pierre-Olivier. 2008. 'An Equilibrium Model of "Global Imbalances" and Low Interest Rates.' *American Economic Review*, 98(1): 358-393

Cabral, Luis. 2004. "Simultaneous Entry and Welfare." *European Economic Review*, 48(5): 943-957

Camerer, Colin and Lovallo, Dan. 1999. "Overconfidence and Excess Entry: An Experimental Approach." *American Economic Review*, 89(1): 306-318

Chandler, Jess. 2009. "Trendy Solutions: Why do States Adopt Sustainable Energy Portfolio Standards?" *Energy Policy* 37(8): 3274-3281

Chao, Hung-po. 1983. "Peak Load Pricing and Capacity Planning with Demand and Supply Uncertainty." *Bell Journal of Economics* 14 (1): 179-190

Cicalla, Steve. 2015. "When Does Regulation Distort Costs? Lessons from Fuel Procurement in US Electricity Generation." *American Economic Review* 105(1): 411-444

Cohen, Bernard. 1990. "The Nuclear Energy Option." *Plenum Press* University of Pittsburgh

Congressional Budget Office. 2001. "Causes and Lessons of the California Electricity Crisis." *A CBO Paper*

Council of Economic Advisers. 2015. Long-Term Interest Rates: A Survey. *Executive Office of the President of the United States*, Washington D.C.

Damsgaard, Niclas. 2003. "Why Deregulation? Economics and Politics in Retail Electricity Markets." *Deregulation and Regulation*

Davis, Lucas. 2008. "The Effect of Driving Restrictions on Air Quality in Mexico City." *Journal of Political Economy* 116 (1): 38-81

Davis, Lucas and Wolfram, Catherine. 2012. "Deregulation, Consolidation, and Efficiency: Evidence from US Nuclear Power." *American Economic Journal: Applied Economics*, 4(4): 194-225

Dixit, Avinash and Stiglitz, Joseph. 1977. "Monopolistic Competition and Optimal Product Diversity." *American Economic Review*, vol. 67, pp. 297-308

Dixit, Avinash and Shapiro, Shapiro, Carl. Entry Dynamics with Mixed Strategies. In: Thomas III, L.G. (Ed.), The Economics of Strategic Planning: Essays in Honor of Joel Dean. *Lexington Books*, Lexington, MA, pp. 63-79

Eichengreen, Barry. 2008. "Globalizing Capital: A History of the International Monetary System." Princeton: *Princeton University Press, 2nd Edition*

Ellis, Randall and Zimmerman, Martin. 1983. "What Happened to Nuclear Power? A Discrete Choice Model of Technology Adoption." *Review of Economics and Statistics* 65(2): 234-242

EIA. 2017. Energy Information Administration. www.eia.doe.gov

EIA AEO. 1990-2015. https://www.eia.gov/outlooks/aeo/archive.cfm

Espey, James and Espey, Molly (2004). "Turning on the Lights: A Meta-Analysis of Residential Electricity Demand Elasticities." *Journal of Agricultural and Applied Economics* 36(1): 65-81

Fabrizio, Kira R., Nancy L. Rose, and Catherine D. Wolfram. 2007. "Do Markets Reduce Costs? Assessing the Impact of Regulatory Restructuring on US Electric Generation Efficiency." *American Economic Review* 97(4): 1250-1277

Fagan, Mark (2006). "Measuring and Explaining Electricity Price Changes in Restructured States." *Electricity Journal* 19(5): 35-42

Federal Energy Regulatory Commission (FERC). 2016. www.ferc.gov

Financial Crisis Inquiry Commission. 2011. "The Financial Crisis Inquiry Report." *US Government Publishing Office*, Washington D.C.

Fowlie, Meredith. 2010. "Emissions Trading, Electricity Restructuring, and Investment in Pollution Abatement." *American Economic Review* 100(3): 837-869

Fowlie, Meredith, Greenstone, Michael, and Wolfram, Catherine. 2015. "Do Energy Efficiency Investment Deliver? Evidence from the Weatherization Assistance Program." *NBER Working Paper No. 21331*

Geroski, P.A. and Mazzucato, M. 2001. "Modelling the Dynamics of Industry Populations." *International Journal of Industrial Organization*, 19(7):1003-1022

Gilbert, Richard and Vives, Xavier. 1986. "Entry Deterrence and the Free Rider Problem." *Review of Economic Studies*, 53(1): 71-83

Hilke, John. 1984. "Excess Capacity and Entry: Some Empirical Evidence." *Journal of Industrial Economics*, 33(2): 233-240

Hobbs, Benjamin. 2000. "Optimization Methods for Electric Utility Resource Planning." *European Journal of Operational Research* 83(1): 1-20

Holland, Stephen and Mansur, Erin. 2006. "The Short-Run Effects of Time-Varying Prices in Competitive Electricity Markets." *Energy Journal* 27 (4): 127-155

Holland, Stephen and Mansur, Erin. 2008. "Is Real-Time Pricing Green? The Environmental Impacts of Electricity Demand Variance." *Review of Economics and Statistics* 90 (3): 550-561

International Atomic Energy Agency. www.iaea.org

Ishii, Jun and Yan, Jingming. 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

Joskow, Paul. 1997. “Restructuring, Competition and Regulatory Reform in the U.S. Electricity Sector.” *Journal of Economic Perspectives*, 11(3): 119-138

Joskow, Paul. 2005. “Incentive Regulation in Theory and Practice: Electricity Distribution and Transmission Networks.” *AEI-Brookings Joint Center* Working Paper 05-18

Joskow, Paul. 2008. “Lessons Learned from Electricity Market Liberalization.” *Energy Journal* Special Issue: 9-42

Joskow, Paul. 2011. “Comparing the Costs of Intermittent and Dispatchable Electricity Generating Technologies.” *American Economic Review Papers and Proceedings* 100(3): 238-241

Joskow, Paul and Kahn, Edward. 2002. “A Quantitative Analysis of Pricing Behavior in California’s Wholesale Electricity Market during Summer 2000.” *Energy Journal* 23 (4): 1-35

Joskow, Paul and Mishkin, Frederic. 1977. “Electric Utility Fuel Choice Behavior in the United States.” *International Economic Review* 18(3): 719-736

Joskow, Paul and Schmalensee, Richard. “Markets for Power: An Analysis of Electric Utility Deregulation.” Cambridge: *MIT Press*, 1983

Joskow, Paul and Tirole, Jean. 2007. “Reliability and Competitive Electricity Markets.” *RAND Journal of Economics* 38 (1): 60-84

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

Jovanovic, Boyan and MacDonald, Glenn. 1994. “The Life Cycle of a Competitive Industry.” *Journal of Political Economy*, 102(2): 322-347

Kahn, Edward 1995. “Regulation by Simulation: The Role of Production Cost Models in Electricity Planning and Pricing.” *Operations Research* 43(3): 388-398

Kaplan, Stan. 2010. “Displacing Coal with Generation from Existing Natural Gas-Fired Power Plants.” *CRS Report for Congress*. January

Klepper, Steven. 1996. “Entry, Exit, Growth and Innovation over the Product Life Cycle.” *American Economic Review*, 86(3): 562-583

Klepper, Steven and Miller, John. 1995. “Entry, Exit and Shakeouts in the United States in New Manufactured Products.” *International Journal of Industrial Organization*, 13: 567-591

Klepper, Steven and Simons, Kenneth. 2000. “The Making of an Oligopoly: Firm Survival and Technological Change in the Evolution of the U.S. Tire Industry.” *Journal of Political Economy*, 108(4): 728-760

Knittel, Chris, Metaxoglou, Konstantinos, and Trindade, Andre. 2016. “Natural Gas Prices and Coal Displacement: Evidence from Electricity Markets, *NBER Working Papers 21627*, National Bureau of Economic Research, Inc.

Kydland, Finn and Prescott, Edward (1982). Time to Build and Aggregate Fluctuations. *Econometrica* 50(6): 1345-1370

Lawrence, Tom. “Evening Internet ‘Rush-Hour’ Affects Broadband Users.” *Independent*, 16 November, 2011. 17 April, 2016. http://www.independent.co.uk/life-style/gadgets-and-tech/news/evening-internet-rush-hour-affects-broadband-users-6262838.html

Lyon, Thomas and Yin, Haitao. 2010. Why Do States Adopt Renewable Portfolio Standards?: An Empirical Investigation.” *Energy Journal* 31(3): 133-158

Macmillan, Steven, Antonyuk, Alexander, and Schwind, Hannah. 2013. “Gas to Coal Competition in the U.S. Power Sector.” *International Energy Agency*

Mankiw, N.G. and Whinston, M.D. 1986. “Free Entry and Social Inefficiency.” *RAND Journal of Economics*, 17(1): 48-58

Mian, Atif and Sufi, Amir. 2014. “House of Debt.” *University of Chicago Press*, Chicago

Mundell, Robert. 1963. “Inflation and Real Interest.” *Journal of Political Economy*, 71(3): 280-283

MIT Interdisciplinary Study. 2003. “The Future of Nuclear Power.” *MIT Press*.

MIT Interdisciplinary Study. 2011. “The Future of Natural Gas.” *MIT Press*.

MIT Interdisciplinary Study. 2007. “The Future of Coal.” *MIT Press*.

MIT Interdisciplinary Study. 2006. “The Future of Geothermal.” *MIT Press*.

MIT Interdisciplinary Study. 2013. “The Future of Solar.” *MIT Press*.

Moulton, Brent. 1990. "An Illustration of a Pitfall in Estimating the Effects of Aggregate Variables on Micro Unit." *Review of Economics and Statistics,* 72(2): 334-38

National Association of Regulatory Utility Commissioners. 2015. www.naruc.org

NOAA. 2015. National Oceanic and Atmospheric Administration. www.noaa.gov

Nordin, J.A. 1947. "Note on a Light Plant's Cost Curves." *Econometrica* 15(3): 231-235

Nuclear Energy Institute. 2016. http://www.nei.org/

NREL. 2017. National Renewable Energy Laboratory. www.nrel.gov

Paul, Anthony, Myers, Erica and Palmer, Karen (2009). "A Partial Adjustment Model of U.S. Electricity Demand by Region, Season, and Sector." *RFF Discussion Paper No. 08-50*

Perry, M.K. 1984. "Scale Economies, Imperfect Competition, and Public Policy." *Journal of Industrial Economics*, 32: 313-330

Peterson, Paul and Wilson, Rachel. 2011. "A Brief Survey of State Integrated Resource Planning Rules and Requirements." *American Clean Skies Foundation*. April

Powers, Nicholas and Yin, Haitao. 2010. "Do State Renewable Portfolio Standards Promote In-state Renewable Generation?" *Energy Policy* 38(2): 1140-1149

Public Service Commission of Wisconsin (PSC). 2009. "Common Power Plant Siting Criteria." September

Schelling, Thomas. 1960. "The Strategy of Conflict." Cambridge: *Harvard University Press*

Schmalensee, Richard. 1981. "Economies of Scale and Barriers to Entry." *Journal of Political Economy*, 89(6): 1228-1238

Schumpeter, Joseph. 1942. "Capitalism, Socialism and Democracy." *Harper & Brothers*, New York

Spence, Michael. 1976. "Product Selection, Fixed Costs, and Monopolistic Competition," *Review of Economic Studies*, vol. 43(2): 217-235

Spence, Michael. 1977. "Entry, Capacity, Investment and Oligopolistic Pricing." *Bell Journal of Economics*, 8(2): 534-544

Steiner, Peter. 1957. “Peak Loads and Efficient Pricing.” *Quarterly Journal of Economics* 71 (4): 585-610

Stern D. I. 2012. “Interfuel substitution: A Meta-Analysis.” *Journal of Economic Surveys,* 26 (June): 307-331

Su, Xuejuan. 2015. “Have Customers Benefited from Electricity Retail Competition?.” *Journal of Regulatory Economics*, 47(2): 146-182

Tansey, Bernadette. “Calpine Runs out of Gas.” *San Francisco Gate* 30 Nov. 2005: sfgate.com Web. 30 Mar. 2016

University of Chicago Interdisciplinary Study. 2004. “The Economic Future of Nuclear Power.” *University of Chicago Press*

Ventyx. 2015. www.ventyx.com

Von Weizsacker, C.C. 1980. “A Welfare Analysis of Barriers to Entry.” *Bell Journal of Economics*, vol. 11: 399-420

Wharton (2006). “Continuing Turmoil in the Power Industry: What It Means for the Major Players.” *Knowledge@Wharton*, 22 Mar. 2006

White, Matthew. 1996. “Power Struggles: Explaining Deregulatory Reforms in Electricity Markets.” *Brookings Papers on Economic Activity-Microeconomics* 0(0): 201-250

Williamson, Oliver. 1966. “Peak-Load Pricing and Optimal Capacity under Indivisibility Constraints.” *American Economic Review* 56 (4): 810-827

Winskel, M. 2002. “When Systems are Overthrown: The 'Dash for Gas' in the British Electricity Supply Industry.” *Social Studies of Science*. 32 (August): 563-598

Wolak, Frank. 2011. “Do Residential Customers Respond to Hourly Prices? Evidence from a Dynamic Pricing Experiment.” *American Economic Review* 101 (3): 83-87

Wolfram, Catherine. 1999. “Measuring Duopoly Power in the British Electricity Spot Market.” *American Economic Review* 89 (4): 805-826

Wolfram, Catherine. 1998. “Strategic Bidding in a Multiunit Auction: An Empirical Analysis of Bids to Supply Electricity in England and Wales.” *RAND Journal of Economics* 29 (4): 703-725

APPENDIX

Appendix 1: Welfare Proof

$\Delta W = W^R - W^S$

$$= \int_0^{x_1^R} D^R(x_1) - p_1^R + \sum_{i=1}^{n}\sum_{a=1}^{2} p_{1i}^R x_{1i}^R - c_{1ai}^R x_{1ai}^R - d_{1ai}^R X_{1ai}^R - d_{1ai}^R Y_{1ai}^R$$

$$+ \int_0^{x_2^R} D^R(x_2) - p_2^R - \sum_{i=1}^{n} p_{2i}^R x_{2i}^R - c_{2i}^R x_{2i}^R - d_{2i}^R X_{2i}^R - d_{2i}^R Y_{2i}^R$$

$$- \int_0^{x_1^S} D^S(x_1) - p_1^S - \sum_{i=1}^{n}\sum_{a=1}^{2} p_{1i}^S x_{1i}^S - c_{1a}^S\, x_{1ai}^S - d_{1ai}^S X_{1a}^S \; - d_{1ai}^S Y_{1ai}^S - \int_0^{x_2^S} D^S(x_2)$$

$$- p_2^S - \sum_{i=1}^{n} p_{2i}^S x_{2i}^S - c_{2i}^S x_{2i}^S - d_{2i}^S X_{2i}^S - d_{2i}^S Y_{2i}^S$$

Imposing assumptions 1 through 4 simplifies the expression to:

$$\int_0^{x_1^R} D(x_1) - p_1^R + \sum_{i=1}^{n}\sum_{a=1}^{2} p_{1i}^R x_{1i}^R - c_{1ai}^R x_{1ai}^R - d_{1ai}^R X_{1ai}^R$$

$$- \int_0^{x_1^S} D(x_1) - p_1^S - \sum_{i=1}^{n}\sum_{a=1}^{2} p_{1i}^S x_{1i}^S - c_{1ai}^S x_{1ai}^S - d_{1ai}^S X_{1ai}^S - \int_{x_2^R}^{x_2^S} D(x_2) - p_2^S$$

$$- \sum_{i=1}^{n} p_{2i}^S x_{2i}^S - c_{2i}^S x_{2i}^S$$

Combining and re-arranging terms, the change in total welfare can be expressed as:

$$\int_{x_1^S}^{x_1^R} D(x_1) - \sum_{i=1}^{n}\sum_{a=1}^{2} c_{1ai}^R x_{1ai}^R + d_{1ai}^R X_{1ai}^R + \sum_{i=1}^{n}\sum_{a=1}^{2} c_{1ai}^S x_{1i}^S + d_{1ai}^S X_{1ai}^S$$

$$- \int_{x_2^R}^{x_2^S} D(x_2) - p_2^S - \sum_{i=1}^{n} p_{2i}^S x_{2i}^S - c_{2i}^S x_{2i}^S$$

Which can then be separated into the following five terms:

$$\int_{x_2^R}^{x_2^S} p_2^S - D(x_2) + \sum_{i=1}^{n} c_{2i}^S x_{2i}^S - p_{2i}^S x_{2i}^S + \sum_{i=1}^{n}\sum_{a=1}^{2}\left(c_{1ai}^S - c_{1a}^R\right)x_{1ai}^S + \sum_{i=1}^{n}\sum_{a=1}^{2}\left(d_{1a}^R - d_{1ai}^S\right)X_{1ai}^S$$
$$+ \int_{x_1^S}^{x_1^R} D(x_1) - \sum_{i=1}^{n}\sum_{a=1}^{2} c_{1a}^R\,(x_{1ai}^R - x_{1ai}^S) - d_{1ai}^R(X_{1a}^R - X_{1ai}^S)$$

Appendix 2: Welfare Aggregation

The producer surplus part of Welfare Effect I was found to be:

$$\sum_{i=1}^{n} c_{2i}^S x_{2i}^S - p_{2i}^S x_{2i}^S$$

This is the negative of the firm's variable profit from operating in the alternative goods market and can be re-written as:

$$-\sum_{i=1}^{n} r d_{2i}^S X_{2i}^S$$

by assuming a normal rate of return (r) on the capital investment ($d_2^S X_2^S$). Using assumption 3, the amount invested in the synthetic market ($d_2^S X_2^S$) is equal to the amount of excess investment in the electricity market ($\sum_{a=1}^{2} d_{1a}^R Y_{1a}^R$). Plugging this into the previous equation yields:

$$-\sum_{i=1}^{n}\sum_{a=1}^{2} r d_{1ai}^R Y_{1a}^R$$

Since Q=X+Y, Y_1^R can be replaced by $(Q_1^R - X_1^R)$ and, according to assumption 4, X_1^R=X_1^S, so the previous equation can be written as:

$$-\sum_{i=1}^{n}\sum_{a=1}^{2} r d_{1a}^R (Q_{1a}^R - X_{1a}^S)$$

This information is not available by firm, so it is aggregated and discounted at the state level over the study time period and the result is:

$$\sum_{t=1}^{T}\sum_{s=1}^{S}\sum_{a=1}^{2} (1+r)^t r d_{1ast}^R (Q_{1ast}^R - X_{1ast}^S)$$

Disaggregating this result by technology, the final equation to be estimated is:

$$\sum_{t=1}^{T}\sum_{s=1}^{S}(1+r)^{t} rd_{11st}^{R}(Q_{11st}^{R}-X_{11st}^{S}) + \sum_{t=1}^{T}\sum_{s=1}^{S}(1+r)^{t} rd_{12st}^{R}(Q_{12st}^{R}-X_{12st}^{S})$$

Appendix 3: Build Decision

Following the neoclassical model of investment and contributions from Bushnell and Ishii (2007), a firm enters the electricity market, or adds to its current position in that market, by constructing a power plant. The firm invests if the expected net present value (NPV) of the plant investment is positive. Firm i chooses the capacity (K)[97] and prime mover (A) [98] of the plant by maximizing:

$$\max_{K,A} E_t \sum_{t=1}^{T(A)} \delta_t[\pi_{it}(K_{it}, X_{it}(K_{it}), X_{-i}\ (K_{it}), A_{it}, \Omega_t, c_t) - \psi(K_{it})]$$

where t=time period, δ=discount factor, π=variable profit, X_i =generation portfolio of the firm, X_{-i}=generation portfolio of competitors, Ω=market conditions, c_t=marginal cost of the plant and ψ(K) is the investment cost function. Firms have the option of constructing a new plant or adding on to an existing one and must also prepare for plant repairs or reductions in usable capacity, so K will fluctuate with time. The amount of periods the plant will generate profit for depends on the type of plant chosen. For example, nuclear plants are expected to be in service longer than solar PV panels. Equation _ provides a more specific structure for the model:

---

[97] Plant capacity specifies how much electricity the plant could produce instantaneously and is typically measured in megawatts.

[98] This is the electricity industry term for the turbine technology used in the power plant.

$$\max_{K,A} \mathrm{E_t} \sum_{t=1}^{T(A)} \sum_{d=1}^{D(t)} \sum_{h=1}^{24} \delta^t \{ p_{tdh}^w \left[x_{tdh}^D, X_{itd}(K_{it}), X_{-it} \quad (K_{it}), R_{it} \right] x_{itdh}^D - c_{it} x_{itdh}^S - \psi_t(K_{Ait})$$

$$x_{tdh}^D = f(\bar{p}_{t-1}^E, Pop_t, Inc_t, W_{tdh}, Sec_t, \mathrm{h})$$

$$c_{it} = f(K_{it}, Age_{it}, A_{it}, p_t^f, w_t, x_{it}^S)$$

$$x_{itdh}^S = f(K_{it}, A_{it}, x_{tdh}^D)$$

$$\psi_t(K_{Ait}) = f(p_t^m, r_t, Res, K_{it})$$

Production constraint: $x_{itdh}^S \leq \mathrm{f}(K_i)$

Balancing constraint: $x_{itdh}^D = x_{itdh}^S$

Financing constraint: $\psi(K_{it}) \leq f(r_t, credit_t)$

where p^w=wholesale price of electricity, x^D= market electricity demand, x^S=electricity generated by the plant, p^E=retail electricity price, Pop=population of the region, Inc=income of the region, W=weather in the market, Sec=economic makeup of the region the market is in, h=hour, R=whether the market is restructured or still regulated, c=marginal cost of producing electricity by the firm, p^f=price of the plant's fuel source, Res=Resource availability in the region, w=operation and maintenance cost of the plant, p^m=construction material cost, and r=interest rate. Time is specified in greater detail in this equation because of its importance in distinguishing the effect of certain variables. For example, demand has both average and variance features, so it is useful to think of that variable at the hourly, daily and longer time period levels.

The production of the plant depends not only on the variable profit of running the plant in each hour and day, but also on the size of the plant and the composition of the firm's generating portfolio.[99] The production constraint places an upper limit on the amount of electricity a firm can supply from the constructed plant. The balancing constraint is necessary as electricity storage is assumed to not be feasible. Therefore, all

[99] For example, a firm may wish to operate other plants in the market due to the cost of shutting down the plants that are currently operational.

electricity that is generated must be sold in that period. Additionally, firms face financial constraints based on interest rates and the credit rating of the company.

Taking partial derivatives $(\frac{\partial \Pi}{\partial K}, \frac{\partial \Pi}{\partial A})$ and simplifying:

$$K_{iA} = f[E_t(x_{tdh}^D, p_t^f, w_t, X_{it+s}, X_{-it+s}, r_{it}, \mathrm{p}_{it}^m, R_t, Res)]$$

The equation above states that the investment (MW) in a particular technology and prime mover of the power plant[100] depends on the expectation of electricity demand, fuel prices, operations and maintenance costs, the generation composition of the firm and other firms both today and in the future, interest rates, price of construction, the state of regulation, and resource availability. With electricity demand varying by hour, day and longer time period, both average demand and variance are factors in firm decision-making.

Firms are interested in average demand for plants they plan to build that require a high capacity factor to be financially feasible. Their projections of future average demand depend both on their expectations about future growth in factors that influence demand ($DemFac$), as well as past demand, which provides insight on the importance of these factors and on broad trends in demand. Past demand is discounted by the firm, as more recent levels of demand are more likely to be useful in projecting future demand.

$$E_t(x_t^{Davg}) = f(E_t(g_{DemFac}), \delta^1 \overline{D}_{t-1}, \delta^2\, \overline{D}_{t-2}, \ldots\ldots, \delta^n \overline{D}_{t-n})$$

Firms are interested in changing load variance if they plan to build high marginal cost facilities that require periods of high congestion to generate a return for capital owners. As with average load, firms project variance by estimating the future growth of

[100] While technically fuel and prime mover are separate decisions, many times the fuel is chosen when the prime mover is chosen. For example, when a steam powered coal plant is built, it can't use uranium as fuel, even though nuclear also uses steam. For the purpose of this chapter, the following prime mover categories exist: coal steam, nuclear steam, combined cycle gas turbine, gas combustion turbine, renewable steam and oil combustion turbine.

factors that influence load variance ($VarFac$) and past measures of load variance, which are discounted as previously with average demand. Shown formally:

$$E_t(x_t^{Dvar}) = f(E_t(g_{VarFac}), \delta D_{t-1}^{Var}, \delta^2 D_{t-2}^{Var}, \ldots\ldots, \delta^n D_{t-n}^{Var})$$

The inclusion of demand growth factors allows equation _ to be narrowed to the expectation of four factors that firms consider when constructing new capacity: average demand and variance, levelized cost of the plant,[101] the existence of competition from the firm's own capacity or competitors'(CAP), and the state of regulation, whether it be environmental or market structure. Summarized:

$$K_A = f(E_t(x_t^{Davg}, x_t^{Dvar}, LC_{Aft}, CAP_t, R_t))$$

[101] This includes the construction cost of the plant, the fuel and operating costs of the plant, and the interest rate. Resource availability can also be included here, with unavailable resources having an infinite price.