The U.S. Electricity Industry after 20 Years of Restructuring

Severin Borenstein and James Bushnell
September 2014
The U.S. Electricity Industry after 20 Years of Restructuring

By Severin Borenstein and James Bushnell*

During the 1990s, the traditional U.S. model of a vertically-integrated, cost-of-service regulated utility was disrupted by a variety of state restructuring plans, which allowed some combination of sales by unregulated merchant generators to retailers and unregulated retailing of electricity to end-use customers. We review the original economic arguments for electricity restructuring, the potential winners and losers from such regulatory reform, and what has actually happened in the subsequent years. We argue that the greatest motivation for restructuring was rent shifting not efficiency improvements, and that this explanation makes sense of the cycles of enthusiasm for electricity reform. While electricity restructuring has brought significant efficiency improvements in generation, it has generally been viewed as a disappointment, because the outsized promises made by some advocates were based on politically unsustainable rent transfers. In reality, the electricity rate impacts of restructuring have been more driven by exogenous factors, like generation technologies and natural gas prices, than by the endogenous effects of restructuring itself. We then apply the same premise to the current political momentum behind distributed generation (primarily rooftop solar PV) which remains costly from a societal viewpoint, but privately economic due to the rent transfers it enables.

* When citing this paper, please use the following: Borenstein, S, Bushnell, JB. The U.S. Electricity Industry after 20 Years of Restructuring. Annu. Rev. Econ. 7: Submitted. Doi: 10.1146/annurev-economics-080614-115630. We are very grateful for the excellent research assistance of Walter Graf in the development of this paper. Borenstein: Haas School of Business, UC Berkeley, and NBER.
I. Introduction

In the mid-1990s, the great majority of electricity customers in the U.S. were served by an investor-owned, vertically integrated utility (IOU), with the remainder receiving power from a publicly-owned municipal utility. Investor-owned electric utilities were closely regulated by state-level public service commissions under cost-of-service principles. In the seven years between 1995 and 2002 a wave of major restructuring transformed the industry. These changes followed closely on the heels of what was seen as the successful economic deregulation of many other industries, including airlines, railroads, telecommunications, gasoline retailing, and the production of oil and natural gas.

At the time, it was widely expected that this transformation would eventually transition the entire industry to a much less regulated and more market-based structure. Yet in the years following 2002 – after the California electricity crisis of 2000-01 – the movement for electricity deregulation encountered a significant backlash. While there was some rhetorical focus on “rolling back” deregulation during this period, it is more accurate to portray actions since 2005 as a cessation of any further deregulatory moves. For the last decade, the policy focus for the electricity industry has turned elsewhere and the loud debates from the early 2000s over the merits of restructuring have been reduced to a background murmur. However, many of the same incentive issues that created the movement toward restructuring 20 years ago remain at work today. They are instead manifested in the increasing focus of the industry on “distributed” generation, which envisions a future in which utilities play a greatly reduced role.

The central premise of this paper is that the political economy in the industry over the last two decades can be viewed as part of a longer cycle of consumer and producer tensions. The cycle is broadly defined by fluctuations in the relationship between the average cost and marginal cost of producing and delivering electricity.
Average cost is the basis for price setting under regulation, while marginal cost is the basis for pricing in a competitive market. That relationship is in turn influenced by many factors. Some of these – such as productivity, level of investment, and the choice of type of investment – are influenced by the transitional incentive problems attributed to rate-of-return regulation. Others are influenced by factors largely beyond the control of state utility commissioners. Two critical exogenous trends during this period have been technology innovations adapted from other sectors (such as aircraft design in the case of gas turbines and semiconductors in the case of solar power) and trends in the prices of natural gas, which is generally the fuel setting marginal costs in most electric systems.

Thus while the restructuring era dawned with great hope that regulatory innovations and the incentives provided by competition would dramatically improve efficiency and greatly lower consumer costs, that hope was largely illusory. Subsequent studies of retail rates in restructured and non-restructured states have generally overlooked the key point that exogenous shocks to the industry often dominate the incremental benefits that restructuring can provide. There is clear evidence that competition has improved efficiency at power plants and improved the coordination of operations across a formerly balkanized power grid. But the impact of gas price movements and new technologies have had a far larger impact.

In section 2 we review the expectations the drove the push for electricity restructuring in the 1990s and how those beliefs shaped the market-based models for electricity markets in each vertical component of the industry: generation, transmission, distribution and retailing. In section 3, we examine the evidence on what effect restructuring has actually had, as well as the most common confusions that confound electricity restructuring with input cost and other changes. Section 4 looks ahead to the most pressing challenge the industry will face in the coming years, the increasing role of renewable and intermittent energy sources, both from utility-scale generation plants and from much smaller scale distributed generation at households and commercial customers. We conclude in section 5.
II. Theory and Implementation of Electricity Restructuring

One of the challenges for an analysis of electricity restructuring is that there are several competing definitions of what restructuring or deregulation actually is. Outside the United States, a key step in electricity restructuring was the divestiture of the government-owned assets that had comprised a nationalized power sector. In the United States, government ownership was never the dominant form of organization and the exceptions in the U.S. – federally marketed hydro-electric power and municipally-owned generation and distribution companies – have remained largely unchanged during the restructuring era. Technically, wholesale electricity markets are still regulated by the Federal Energy Regulatory Commission (FERC) under the authority granted by the 1938 Federal Power Act. The wave of state-level restructuring did not change this fact, although FERC has applied its authority flexibly by allowing states and regions to set “market-based” rates. Such authority can be revoked, however, so it is inaccurate to label even wholesale markets in fully restructured regions as “deregulated.”

In a market-based system for electricity provision, the industry is generally considered as four separate activities: generation of electricity, transmission over high-voltage lines and long-distances, voltage step-down (to the 110V common in the U.S. or 220V used in Europe and elsewhere) and local distribution to end users, and retailing. The last activity includes procurement of power under long-term contracts, rate setting, billing, and collection. The U.S. restructuring process was focused on generation, transmission and retailing, all of which were dominated by investor-owned regulated utilities rather than public ownership. The local distribution lines continued to be considered a natural monopoly that would be subject to either regulation or municipal ownership.

Changes to generation, transmission and retailing were pursued with varying levels of commitment in different parts of the country. Independent oversight and control of the transmission networks was viewed by many as the backbone of restructuring, because transmission was critical to generators accessing a com-
petitive market into which they could sell and to retailers accessing competitive sellers from which they could buy. Restructuring of generation resembled most closely the deregulation that had taken place in other industries, with free entry of merchant electricity plants that would live or die by their cost of production and the price they could get for their output. Finally, retail restructuring, in the limited areas it has taken hold in the U.S., has allowed non-utility companies to become the wholesale procurement entities for retail customers, offering customers alternative retail pricing structures, though across a rather limited spectrum as we discuss below.

In theory, at least, the three aspects of restructuring were closely intertwined. Without independent oversight of transmission, a merchant electricity plant would be at the mercy of the local transmission owner, which could extract large shares of the quasi-rents available once the plant was built, thereby discouraging entry of competitive generation. Even with transmission access, a merchant generator would be in a very weak position if there were only one retail electricity provider to which it could sell its output. A monopoly retail provider (a distribution utility) could still engage in competitive procurement, but that creates a narrower spectrum for competitive generation and it means that the monopoly retailer is the single determinant of the range of products that might be procured for retail. For instance, the monopoly retailer might not pursue low-carbon sources even if there are many retail customers who would be willing to pay a premium for greener energy. Thus, retail competition potentially makes competitive generation more viable. Likewise, competitive generation is central to the retailer being able to offer better procurement options, different generation sources, or alternative billing mechanisms, which the retailer would likely want to balance with the wholesale contracts it has with producers.

In practice, while pursuit of restructuring in the three activities has been regionally correlated, many areas have developed generation restructuring without retail competition. And independent transmission operators have taken over large
swaths of the U.S. grid in which both generation and retail competition varies greatly.

A. Transmission Access Reforms

Transmission restructuring proceeded along two paths, a regulatory path that attempted to impose rules upon vertically integrated utilities that would promote third party access to their networks, and an institutional path that encouraged the creation of Independent System Operators (ISOs) and later Regional Transmission Organizations (RTOs).\(^1\) FERC attempted through a series of orders during the 1990s and 2000s to force the creation of more transparent online market places for available transmission capacity and to require vertically integrated utilities to provide transmission service to third-party independent power producers (IPPs, also known as merchant generators). These efforts have achieved at best mixed success.

The more successful path to nondiscriminatory grid access appears to have been the creation of the RTO/ISO. These entities are organized as user-supported non-profit companies and operate essentially as regulated entities overseen by FERC. In the U.S. these transmission companies do not own the transmission assets in their jurisdiction, but rather they control access to those assets by virtue of approving, and in some cases setting, the production schedules of the power plants within their regions, as well as operating real-time balancing markets that adjust supply as needed to maintain network reliability. In each case, the decisions made by the ISOs with regards to generation operations are dominated by a mandate to respect the constraints of the transmission network and other reliability considerations. Unlike the vertically integrated network entities, ISOs have no generation assets or retail consumers, and are therefore credibly impartial as to specific market outcomes as long those outcomes do not threaten reliability.\(^2\)

\(^1\)Both types of organizations are tasked by FERC to coordinate investment and operations of regional power grids in a non-discriminatory transparent manner.

\(^2\)Indeed, RTO/ISOs have at times been criticized as being too exclusively focused on reliability and
Initially the RTO/ISO model was largely restricted to markets undertaking the full suite of restructuring steps described in this section. The full and unfettered access of disparate power producers to the available population of electricity customers dictated an institutional structure that would eliminate concerns over vertical barriers. Conversely, jurisdictions that wanted no part of retail competition were equally suspicious of the RTO/ISO structure as an initial step down the slippery slope to full restructuring. Thus, many municipal utilities and some of the largest and strongest integrated utilities, as well as the Federal Marketing Agencies, kept their transmission systems organized along traditional structures in which they directly controlled access and real-time use.

This changed in the latter half of the 2000s. As we discuss below, the pressures to restructure other aspects of utility operations receded in many regions, so joining an RTO/ISO market no longer implied the inevitable dissolution of the traditional utility franchise. At the same time, the benefits of better coordination of operations and lowering of transactions costs within ISOs appear to have been substantial. Figure 1 illustrates the geographic reach of North American ISOs and RTOs as of 2012. Currently, RTOs such as the Midcontinent Independent System Operator, Southwest Power Pool, and PJM each contain several states that never seriously considered restructuring their generation or retail sectors.

The creation and expansion of the RTO/ISO model may be the single most unambiguous success of the restructuring era in the United States. The U.S. has historically suffered from a utility system that was highly balkanized relative to most other countries. The evidence suggests that the lack of coordination across utility control areas impeded Pareto-improving trades worth billions of dollars. Although the early momentum for aggregating utility control areas into

---

3See Joskow (2006), Wolak (2011a), and Mansur and White (2012).

4See White (1992), Joskow (1997), Kleit and Reitzes (2008), and Mansur and White (2012).
more regionally managed RTOs was provided by it being seen as a necessary step toward the ultimate goal of deregulating generation and retail, the expansion of the RTO structure has come to be viewed as a valuable legacy of this period even for states that never showed serious interest in these other aspects of restructuring.

B. Restructuring of Generation Ownership

The second dimension of restructuring impacted the ownership status and remuneration of generation assets. Large amounts of generation capacity were converted from utility status to independent power producer (non-utility or “merchant”) status. Effectively, these assets transitioned from a rate-of-return regulation model in which they were compensated based upon average production cost to a market-based pricing model under which these assets earned a market price for the output they were able to produce.

To the extent one considers the electric sector to be “deregulated,” it is due to this fundamental shift in the paradigm for compensating owners of generation. In addition to the divestiture of much of the existing generation fleet previously
owned by IOUs, an equally dramatic change impacted the investment in new generation. The construction of generation assets was no longer coupled with a guarantee to recover a positive return on those capital costs. In 1997 only 1.6% of U.S. electricity was produced by generation owned by firms classified as Independent Power Producers (IPPs). That figure rose to 25% by 2002 and was just under 35% in 2012. The share of nuclear generation owned by IPPs rose from zero in 1997 to almost 50% in 2012.

Figure 2 displays the diversity of ownership patterns across the U.S. as of 2012 and illustrates the strong regional pattern of generation restructuring. The Southeast, with its large and regionally powerful IOUs, and much of the Pacific Northwest, with its dominance of federally operated generation and municipal utilities, have largely resisted changes in generation ownership. Importantly these regions also enjoyed amongst the lowest average retail rates in the country in 1997. The Northeast and Illinois have almost fully transitioned to a non-utility form of ownership, while Texas, California and Montana have also seen large shares of IPPs.

As we will discuss below, we consider this dimension of restructuring to be the most economically meaningful in its consequence. This is mainly because the
majority of costs still reside in the generation sector and the fact that the most potential variation in costs and prices resides in this sector.

Political attitudes toward the effects of restructuring during the last 20 years have also been dominated by outcomes in the generation sector. These attitudes can largely be captured through the lens of comparing average to marginal costs.

In the early 1990s, just prior to the initial years of restructuring, much of the country experienced large generation reserve margins (see Figure 3). Until the last few years (with the rise of intermittent renewable generation), this statistic was a very good proxy for measuring the efficient deployment of capital. Larger reserve margins generally imply installed capacity (and capital) that is underutilized. Lower utilization implies higher average costs as the capital expenditures are spread amongst a smaller consumer base. Lower utilization rates also often implied that generation with relatively low marginal cost was often available, and marginal, thereby contributing to relatively low regional wholesale prices. Historically low natural gas prices during the 1990s also greatly contributed to low regional wholesale prices.

The industry during the late 1990s was therefore experiencing a period of unusually low marginal costs and unusually high average costs. This is the fundamental source of the pressure for restructuring. While, as discussed above, much of the rhetoric at the time focused on retail deregulation, this needs to be seen from the perspective of customers (often large industrial customers) who saw great opportunity in being able to gain “direct access” to the wholesale market.

Of course, what appeared as a great opportunity for customers conversely created a real threat to utilities who were the residual claimants on generation assets for which the market value in a competitive wholesale market would have been well below the depreciated capital value that remained on the utilities’ books at the time of restructuring. This fact was quickly internalized by equity markets.

\[\text{5In Borenstein and Bushnell (2000), we pointed out this tension between efficient economic decision making and incentives for rent shifting.}\]
Share prices of the largest utilities in California, Pennsylvania, and New England all experienced sharp declines during the mid-1990s. The concern among holders of utility stocks soon gave way to a period of reflection and negotiation over an acceptable transition from an average-cost to market-based pricing paradigm. The political and regulatory process was forced to confront the uncomfortable fact that much of the consumer appeal of restructuring was rooted not in cost savings and productivity gains, but rather in an opportunity to shift responsibility for paying the sunk costs of what were considered uneconomic “stranded assets.” This meant that immediate consumer savings were largely dependent upon an equivalent reduction in returns for utility shareholders. This is an important theme we will return to when we examine the current rhetoric about the “utility of the future.”

In the end, utilities in all restructuring states persuaded regulators that they should not simply have the cost recovery rug pulled out from under them. Since the assets then viewed as excessive white elephants had for the most part been
built with the approval, and in some cases mandate, of regulatory commissions, it was generally concluded that to force restructuring without compensation for stranded assets would violate the “regulatory compact,” or implicit guarantee that utilities could recover the costs of capital investments made with the approval of regulators. Most state restructuring schemes included a plan for 100% recovery by utilities of any stranded investment and the others aimed at nearly 100% recovery.

The mechanism for recovering stranded cost was to allow a transition period in which portions of utility retail prices would be frozen at what were then considered to be above-market rates during a transition period. Utilities would therefore be allowed to apply these excess retail margins to pay down the stranded costs on their divested and retained generation assets. The exact design of this mechanism varied by state, and it produced tragic consequences for California where the excess retail margins suddenly turned negative and caused the 2000-01 California electricity crisis. In order to avoid conflict between the goals of fostering retail competition and recovery of stranded costs, these competition transition charges were generally applied as surcharges to the bills of distribution companies who maintained a monopoly franchise over the wires components of the business. Therefore, somewhat ironically, while the customer impetus that started electricity restructuring was a desire to avoid paying for average costs during a period when marginal costs were lower, the transition charges largely guaranteed that utilities recovered something close to those costs anyway.

In addition to generation assets operated by utilities, stranded assets in several states included uneconomic long-term contracts with IPPs that were mostly mandated by PUCs under the Public Utilities Regulatory Power Act of 1978. See White (1996)

Through a combination of real scarcity and generator market power (caused in part by high local natural gas prices that followed a pipeline explosion), California wholesale electricity prices skyrocketed in summer 2000 and remained extremely high into May of 2001. Under the competition transition plan, the two largest utilities in the state were not allowed to raise retail rates to reflect the high wholesale prices. One of them, Pacific Gas & Electric, was forced into bankruptcy and the other, Southern California Edision nearly followed. For detailed discussions of these events, see Joskow (2001), Blumstein, Friedman and Green (2002), Borenstein (2002), Wolak (2003b), and Bushnell (2004).
C. Restructuring and Reform of Retail Services

The aspect of restructuring to receive the most rhetorical attention and market hype was the relaxing of the utility monopoly franchise over retailing. Phrases evoking liberty and freedom, such as “customer choice” and “freedom to choose” were rhetorical staples of the restructuring process. There was also much hope that electricity retail competition might spur innovation in retail services in the way that it had for telecommunications. Exactly how this was supposed to be achieved was never clear.\(^8\) Electricity service has proven to be less amenable to the sorts of usage and complementary product innovation that wired telecom service experienced in the 1980s and 1990s. Perhaps this isn’t surprising given that the product is so narrow – just the electricity, not any devices that use it – and so homogeneous. In order to use the grid, electricity must meet exact specifications that make one provider’s product indistinguishable from another’s. The place where innovation did seem valuable and likely to occur with retail choice was in financial arrangements: price schedules, payment plans, and options to bundle purchases with complementary products.

More concretely, retail deregulation involved the opening of customer access to new “energy-only” retail providers who produced or acquired wholesale power for sale to end users. The incumbent utility (and the grid operator) maintained a franchise over distribution and transmission related functions. In many cases the incumbent utility was allowed to continue to offer a default “bundled” retail rate for customers who did not switch retailers.\(^9\) Customers who did switch received a bill for “energy-only” service from the third-party retailer they chose, and a separate charge, intended to recover transmission and distribution system investments made by the incumbent utility.

The extent to which this transformation has materialized has varied greatly

\(^8\)See Joskow’s (2000) discussion of the potential for new product innovation under electricity restructuring.

\(^9\)The bundled rate combined energy with the incumbent utility’s transmission, distribution and retailing charge. This was sometimes called the “default provider” or “provider of last resort” (or POLR) rate. In some states, the default provider franchise is selected through auctions overseen by local regulators.
Figure 4. Share of Retail Sales from Retail Power Marketers

around the U.S. Figure 4 illustrates the fraction of total sales in each state from entities with an ownership classification of 'retail power marketer.'\[^{10}\] Texas has far outstripped the rest of the country on the retail competition front, with the only other significant activity clustered in the Northeast.

Retail Price Reform. — To understand the potential for efficiency improvements in pricing electricity, it helps to review the inefficiency concerns raised by the typical 1990s electricity retail tariff. Throughout most of the history of electric utilities, retail pricing policy has been driven more by equity than efficiency considerations. Because customers had little alternative to the monopoly utility provider, and the utility was focused on meeting a rate of return standard not maximizing profits, there was little initiative to improve the efficiency of pricing. However, with greater competition and demand elasticity – from non-utility energy sources and retail suppliers, and more recently from improved opportunities to generate electricity behind the meter – the pressure to align prices with marginal costs has grown.

\[^{10}\]These data are compiled from EIA form 861.
Efficient retail prices should reflect the short-run marginal cost in every hourly (or even shorter) time period at every location on the grid. At the beginning of restructuring, nearly all residential, commercial and industrial customers faced prices that did not vary hour to hour. Furthermore, utilities recovered nearly all of their costs through volumetric charges, including the substantial share of costs that are fixed with respect to a customer’s marginal consumption. For most residential customers, the rate was a simple constant price per kilowatt-hour (kWh) consumed, regardless of when the energy was used, set to cover all the utility’s costs, variable and fixed.\textsuperscript{11}

Setting price equal to short-run time-specific and location-specific marginal cost leads to efficient consumption given the level of investment, but only under a very narrow set of conditions does it exactly cover total costs.\textsuperscript{12} In reality, there are almost certainly some costs that scale less than proportionally with the total quantity sold, so efficient marginal pricing would result in a revenue shortfall.

A fixed charge can be used to capture the additional needed revenue. A fixed charge (per month, for instance) is particularly efficient in residential electricity markets because the elasticity of connecting to the grid with respect to the monthly fixed charge is likely near zero over a wide range of charges. Thus, the deadweight loss that could result if some customers chose to consume zero because the fixed charge exceeds their consumer surplus is likely to be small.

For basically the same reason, however, the distributional consequences of a fixed charge are of great concern. Moving from a flat volumetric rate and no fixed charge to a lower flat rate and a fixed charge is very regressive. Borenstein (2011) shows that such a revenue-neutral change to a higher fixed charge and lower volumetric rate would raise the average bills of low-income customers by 69% to 92% of the fixed charge across the three large investor-owned utilities.

\textsuperscript{11}Borenstein and Holland (2007) show that the break-even flat price could be higher or lower than the second-best optimal flat rate, depending on whether peak or off-peak price elasticities are higher.

\textsuperscript{12}Under constant returns to scale, optimal pricing covers costs if capacity is also set optimally. If capacity is greater than the optimum level, optimal pricing will generate less revenue than is needed to cover total costs.
in California. Equity notions often suggest that the fairest allocation of such a revenue requirement would be in proportion to quantity consumed.\textsuperscript{13} That approach, however, steers back towards average cost pricing and the inefficiencies that it is known to produce.

The problem of average cost pricing is exacerbated in the electricity industry by the nature of the contract between the retail provider and the customer. In nearly all cases, the customer has an option, but not an obligation, to purchase any quantity at the announced price, known in the industry as a “requirements contract.” This in itself wouldn’t be a destabilizing force if price adjusted quickly,\textsuperscript{14} but with long lags between cost changes and price adjustment, this creates an opportunity for buyers to switch between alternative suppliers inefficiently. This is the same phenomenon as described earlier for the state decision to deregulate, but manifest in contemporaneous customer choice among competing sources. The combination of requirements contracts and average-cost retail pricing will create increasing problems if distributed generation (“behind the meter”) continues to expand, as we discuss below.

Thus, as restructuring began 20 years ago, retail pricing deviated considerably from the ideal efficient structure. It seemed at least possible that competitive pressure on the existing structure would lead to substantial changes in pricing, and the potential for differentiation among the products retailers sold. The technological and market configuration, however, turned out to leave much less space for pricing innovation than was suggested at the time.

The principle technological constraint was metering: in the 1990s, virtually all residential customers, and most commercial and industrial customers, had meters that recorded only the aggregate amount of electricity that had flowed through them. They did not have the capability to collect information on when the electricity was consumed. This meant that time-varying pricing wasn’t feasible with-

\textsuperscript{13}Or, if data were available, in proportion to consumer surplus gained by each customer.

\textsuperscript{14}In a sense, sellers in any commodity market operate under requirements contracts, at least over a large range of purchase, but they can and do change prices rapidly as market conditions change.
out a significant investment in metering. Nor could a retailer necessarily overcome this constraint just by metering its own customers, because the arrangements for billing and payments among retailers and the utility providing distribution services were generally not set up to accommodate time-varying pricing. Instead, in most cases a retailer was deemed responsible for providing power to its customers – either generating it, signing long-term contracts, or buying on the spot market – based on a standard assumed “load shape” (a time-varying pattern of consumption) that was applied to all customers within broad location, customer type, and sometimes size, classes. The assumed load shape was independent of the prices the customer faced, so the retailer had no incentive to charge time-varying prices. With the expansion of smart meters in the late 2000s, the groundwork is now being laid for broader use of time-varying pricing, but the vast majority of residential customers with non-utility retail providers still see no time-variation in the prices they pay. Commercial and industrial customers have experienced a much greater shift towards time-of-use pricing, which entails two or three different pre-set prices that apply at different times of the week. TOU pricing, however, is known to capture a small share of the hourly variation in wholesale electricity prices.\textsuperscript{15}

A second way in which retailers might have offered greater differentiation was in reliability, but this too was undermined by the structure of the retail markets that were established. Because the grid operator must always balance supply and demand to avoid service disruptions, the grid operators in these markets procured enough reserves to make sure that the full expected demand could be met. If one retailer did not procure sufficient supplies to meet its retail demand obligation, the result was not reduced supply to the customers of that retailer – as would occur with nearly any other product. Instead, the grid operator drew on its reserves to make sure all demand was met. The cost of those reserves was spread over all retail quantities whether or not the provider to a particular customer ever caused

\textsuperscript{15}See Borenstein (2005).
the grid operator to need those reserves. Reliability was assured by the grid operator and charged to every kWh supplied, so there could be no differentiation on reliability. Alternative arrangements – in which the customer either lost power when its supplier had procured insufficient quantities (which posed technological challenges along the same lines as real-time metering) or the retailer or customer was charged a very high fee for being short – would have created a significant cost for insufficient supply and likely led to greater product differentiation along these lines, but these weren’t widely adopted. The lack of retailer responsibility for reliability also undermines the incentive to implement price-responsive demand, which could be a valuable tool for a retailer in balancing its supply and demand while keeping costs down.

Reliability differentiation also could extend to exit by the retailer. If a retailer exits the market, what cost is borne by its customers? If customers can easily switch to another supplier at a pre-determined rate, then a similar moral hazard problem arises in which a retailer can procure short-term power at spot prices when that price is low, but exit if the spot price rises, leaving the customer to switch to some default rate. If that default rate is a price that reflects average procurement costs over a longer period, then once again the variation in average versus marginal price drives behavior in the market. Enron and some other retail providers in California took this path when prices in the California wholesale market spiked in 2000. In Texas, which has the most extensive retail residential competition (see figure 4), rules have been adjusted so that customers of a retail provider that exits are moved, by default, to a tariff that reflects the contemporaneous marginal cost of procuring power.

III. Electricity Market Performance Since Restructuring Began

The most consequential economic changes took place on the wholesale production and marketing sectors of the industry. We therefore begin by discussing the evolution of the industry since 1997 at the wholesale level. As discussed above,
formal centralized markets only formed in the parts of the country that embraced the RTO/ISO structure, which were also the areas with the highest prices and for which the average cost exceeded marginal cost by the largest amount.

A. Wholesale Markets

The regions with RTO/ISOs are also the markets for which the best data on wholesale prices are available. Figure 5 summarizes annual average prices from two data sources. For 1998 through 2001 we use data from Bushnell, Mansur and Saravia (2008), which are drawn from ISO websites. For 2001 on, we report data from the Intercontinental Exchange (ICE) for trading hubs in Southern California (SP15), western Pennsylvania (PJM) Massachusetts (ISO-NE) and the Pacific Northwest (Mid C) hubs. The dashed line in figure 5 summarizes the U.S. average city gate natural gas price, taken from the Energy Information Administration.

Since 1998, two facts are worth noting. First, although somewhat muted by the annual aggregation in the data, the California market stands out as suffering
from sustained extremely high price levels during the period from 2000-2001. Both academic research and subsequent regulatory findings have determined that this market suffered from a lack of competition made acute by a combination of tightening capacity and a near total absence of forward contracting.\(^\text{16}\) Second, in the other markets wholesale power market prices are dominated by natural gas prices, though somewhat less so in the Pacific Northwest. This is consistent with the general fact that natural gas fired generation units are the marginal source of power in most markets during most times, but the Pacific Northwest is influenced more by the availability of hydro-electric power.

Because gas generation comprises a minority share in most electricity markets, under average-cost based regulation it did not dominate rate making. Prices for deregulated generation, however, are driven by the marginal producer, which is much more commonly gas generation. Thus to a degree that was not appreciated at the time, power market reformers hitched their rate determination carts to the natural gas market horse. As natural gas prices nearly tripled during the first half of the 2000s, the impact on retail rates and the rents created for infra-marginal generation were far greater than they would have been under regulation.

During 2006 and 2008 the U.S. natural gas price peaked above $11/MMBTU. The higher gas prices drove up generation costs and power market prices. By this time, the relationship between marginal and average costs of power production had again reversed so that marginal cost-based market prices were higher than the average costs of operating and producing from a mixed generation portfolio. Many of the nuclear and coal-fired power plants in restructured states, which had been considered “stranded” assets in the late 1990s, were by 2007 tremendously profitable due to their low operating costs and the relatively high market prices they earned for their output.

The combination of higher prices and healthy profits earned by power produc-

ers in restructured states contributed to a strong dissatisfaction with restructuring in several states.\textsuperscript{17} This mood of ex-post regulatory regret in restructured states peaked in 2007-2008. States such as Illinois, Maryland and Maine initiated proceedings that were characterized as rolling back deregulation.\textsuperscript{18} After 2009, however, with plummeting natural gas prices and increasing reserve margins, momentum for significant changes declined.

\textbf{B. Restructuring and Plant Operations}

One aspect of restructuring that has been studied at a micro level has been its impact on the performance and efficiency of power plants. Overall the results point to a positive influence of restructuring on the operations of plants. Unfortunately, while cost data on regulated plants are extensive, there is much less data available on the costs of non-utility generation. Thus, studies of plant-level impacts of restructuring have either focused on its impact on regulated plants or were limited to a focus on the few performance variables that continue to be reported for deregulated plants. Fabrizio, Rose, and Wolfram (2007) compared the performance of regulated plants in states that pursued restructuring (by the Energy Information Administration’s definition, which we discuss further below) against regulated plants in states that did not initiate restructuring, and against publicly owned plants in both types of states. They find modest efficiency gains in the restructured states, much of these focused on employment and labor productivity. There is some evidence that the efficiency of fuel consumption, the largest single variable expense in power plants, can be influenced by incentives and skill,\textsuperscript{19} but to date the evidence on fuel efficiency at restructured plants has been inconclusive.

The most dramatic documented impact of restructuring on power plant operations has been on the performance of nuclear plants, shown by Davis and Wolfram

\textsuperscript{17}See Johnston (2007).
\textsuperscript{18}See Sharp (2007) and Behr (2009).
\textsuperscript{19}See Bushnell and Wolfram (2009)
(2012). Almost half of the nuclear generation plants in the U.S. were divested to non-utility producers since 1998.\textsuperscript{20} Davis and Wolfram show that industrywide U.S. nuclear power plants have greatly increased capacity factors since 1998, but relative to their regulated counter-parts, output at the restructured plants increased 10 percent between 1998 and 2010. They estimate this additional output has a market value of $2.5 billion dollars annually.\textsuperscript{21}

C. Restructuring and Retail Prices

It is useful to begin a review of retail prices under deregulation by examining conditions in 2007, when dissatisfaction with restructuring peaked. In 2007, the New York Times ran a series of articles highlighting the fact that rates had risen faster in restructured states than in regulated ones. The articles cited studies that relied upon average retail price data reported to the Energy Information Administration and essentially performed a difference in difference comparison between restructured and non-restructured States.\textsuperscript{22}

A central challenge in studies like this is to identify what constitutes “restructured” in order to assign a state to one category or the other. Many papers have relied upon the Energy Information Administration’s definition, which in turn is focused on the status of retail competition. An alternative measure of deregulation is based upon the fraction of energy generated in a state that is produced by Independent Power Producers (IPP). Figure 2 illustrates these values for 2012, but we can apply the full panel of values to capture the underlying points of transition in each state.

As one examines recent data on retail rates, it is clear that many of the conditions of 2007 have since dramatically reversed. Table 1 summarizes the average retail rates in states considered “deregulated” according to two alternative mea-

\textsuperscript{20}Since 1998, no new nuclear plants have come online.
\textsuperscript{21}Hausman (2014) concludes that the gains in utilization were not accompanied by degradation of safety among deregulated plants.
\textsuperscript{22}See Showalter (2007) and Tierney (2007).
sures against those that are still regulated.\textsuperscript{23} The first measure is the one used in a study by Showalter (2007) for Power in the Public Interest (PPI) that is cited in the NY Times article. This definition excludes from the restructured category states such as Illinois and Pennsylvania which by 2012 have almost all of their energy provided from non-utility sources. As an alternative measure, we assign states to the restructured category if they had more than 40\% of their energy provided by non-utility sources in 2012.\textsuperscript{24}

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|c|c|c|c|}
\hline
\textbf{Definition} & \textbf{Status} & \textbf{Average Retail Price} & \textbf{Percent Change} \\
\hline
 &  & 1997 & 2007 & 2012 & 97 to 07 & 07 to 12 & 98 to 12 \\
\hline
PPI Definition & Regulated & 5.89 & 7.44 & 8.72 & 0.21 & 0.15 & 0.32 \\
& Restructured & 8.96 & 12.53 & 12.35 & 0.29 & -0.01 & 0.27 \\
\hline
At least 40\% & Regulated & 5.67 & 7.23 & 8.57 & 0.22 & 0.16 & 0.34 \\
IPP in 2012 & Restructured & 8.83 & 11.99 & 11.95 & 0.26 & 0.00 & 0.26 \\
\hline
\end{tabular}
\caption{Summary of Retail Price Changes}
\end{table}

From Table 1 one can see that at this level of analysis the definition of restructured makes only a small difference. The time period examined, however, makes an enormous difference as rates in restructured states increased at a pace nearly 50\% higher than those in regulated states between 1997 and 2007 but have actually declined slightly since 2007. Average rates in regulated states have continued to increase since 2007 at a slightly lower pace than between 1998-2007. Overall there is almost no difference in the change in average rates for the two groups over the full sample from 1998-2012.

Figure 6 illustrates the annual levels of rates in restructured and regulated states using our generation-based definition, along with the national average city

\textsuperscript{23}Retail price data come from the EIA form 861, which report sales and revenues by utility. We examine the average rate across all major rate categories, including residential, industrial and commercial. Several previous studies, including Showalter (2007) and Apt (2005) have focused on rates paid by industrial customers.

\textsuperscript{24}The NY Times article lists the restructured states as CA, CT, DC, DE, MA, MD, ME, MI, MT, NH, NJ, NY, RI, and TX. Our generation-based definition puts CA, CT, DE, IL, MA, MD, ME, MT, NH, NJ, OH, PA, NY, RI, TX, and VT into the restructured category.
gate natural gas price. Restructured states experienced higher rates during the 1990s, a major factor in their election to adopt restructuring. The gap between regulated and restructured states narrows around 1998, reflecting the impact of legislation that required immediate rate reductions accompany restructuring in several states. Since that time, rates in restructured states more closely follow the trajectory of gas prices up during the early 2000s and back down since then.

To further test this relationship between natural gas prices, restructuring and electricity rates we estimate the following regression on state level annual changes in electricity prices and city-gate natural gas prices.

\[ \Delta \text{Elec}_{s,t} = \alpha + \beta_1 \text{FractionIPP}_{s,t} + \beta_2 \Delta \text{NGas}_{s,t} + \beta_3 \text{FractionIPP}_{s,t} \times \Delta \text{NG}_{s,t}, \]

where \( \Delta \text{Elec}_{s,t} = \ln(\text{Rate}_{s,t}) - \ln(\text{Rate}_{s,t-1}) \) and \( \Delta \text{NGas}_{s,t} = \ln(\text{NG\_CityGate}_{s,t}) - \ln(\text{NG\_CityGate}_{s,t-1}) \) are the annual changes in log state average electricity rates,
and log state average city-gate natural gas prices, respectively. We estimate for 1998 (the change from 1997) to 2012. Table 2 presents the summary statistics for these variables in the years 1997 and 2012. We estimate (1) clustering standard errors at the state level.

| Table 2—Summary Statistics of Retail Electric and Natural Gas Prices |
|-----------------------------|-----|-----|-----|-----|
| Variable | Mean | S.D. | Min | Max |
| Price    | 6.72 | 2.03 | 3.87 | 11.66 |
| Fraction IPP | 0.03 | 0.07 | 0.00 | 0.46 |
| Nat. Gas | 3.54 | 0.64 | 2.12 | 5.18 |
| Data for 2012 |
| Price    | 9.70 | 2.30 | 6.90 | 15.54 |
| Fraction IPP | 0.35 | 0.33 | 0.00 | 0.99 |
| Nat. Gas | 4.90 | 0.97 | 3.46 | 7.73 |

The results of regression (1) are reported in table 3. As table 1 suggests, restructuring, which we are representing with fraction of energy generation from non-utility sources in that year (Fraction IPP), has no statistically discernible effect on average changes in rates over the 1997 to 2012 period. The point estimate implies that a state with 100% merchant generation has a 0.6% higher average annual rate increase, but one cannot reject no effect at conventional significance levels. Changes in local natural gas prices, however, do influence rates. The second column of table 3 suggests that a 1% increase in natural gas prices implies a 5% increase in electricity prices on its own. The third column in the table yields greater clarity on the mechanism. When the change in natural gas price is interacted with the Fraction IPP, the results suggest that the effect of natural gas is much greater in restructured states as the earlier discussion would suggest.

Both time series are from the Energy Information Administration. Electricity rates are the “Total Electric Industry” average price across all customer classes, per state, as reported at http://www.eia.gov/electricity/data/state/ and derived from EIA form 861 data. Natural Gas prices are available at http://www.eia.gov/dnav/ng/ng_pri_sum_dcu_nus_a.htm, and derived from EIA form 857 data.
The influence of natural gas price on retail rates is estimated to be nearly twice as large in a state with all merchant generation than in a state with none. The effect of natural gas prices in a state with no merchant generation is not statistically significant, while the interacted effect with Fraction_IPP is highly significant.

Table 3—Analysis of Retail Price Changes

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pct IPP</td>
<td>0.006</td>
<td>0.007</td>
<td>0.006</td>
</tr>
<tr>
<td></td>
<td>(0.005)</td>
<td>(0.005)</td>
<td>(0.005)</td>
</tr>
<tr>
<td>Pct Change in Nat. Gas</td>
<td>0.051</td>
<td>0.023</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.016)</td>
<td>(0.016)</td>
<td></td>
</tr>
<tr>
<td>ΔNGasxPctIPP</td>
<td>NA</td>
<td>0.018</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.005)</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>720</td>
<td>720</td>
<td>720</td>
</tr>
</tbody>
</table>

*Dependent variable is change in log annual state-level average electricity rates. Standard Errors are clustered by state.*

We do not intend this to be an exhaustive analysis of the drivers of retail prices. However these data are strongly supportive of the argument that, apart from the California electricity crisis, any harm that electricity restructuring has done to consumers was a side-effect of changes in the price of natural gas. In restructured markets, natural gas generation determines market prices and therefore the remuneration for all the non-utility assets. The more non-utility assets featured in a state’s generation mix, the more exposed that state is to the natural gas market.

Simply put, restructuring in the U.S. was in hindsight very poorly timed. Assets that were viewed as stranded in 1998 were sold as white elephants at prices far below what they would have fetched in 2007. Conversely large customers who gazed longingly at reported wholesale prices in the $20/MWh range in 1996 switched to market pricing just in time to ride (or, more accurately, be hit by)

Other such as Apt (2005) and Taber, Chapman and Mount (2006) have performed more extensive exercises, but only utilizing data during the early years of restructuring.
the wave of natural gas price increases up to its peak in 2007. This timing is not coincidental, however, because the very factors that contributed to the low valuation of utility assets in the late 1990s (low wholesale prices) were the ones that made the prospect of restructuring so appealing to customers and policy makers.

D. The Evolution of Retail Price Structures

Unfortunately, data on retail price structures are much less available than data on retail price levels. Nonetheless, it is clear that there has been gradual movement towards time-varying pricing, primarily for commercial and industrial customers. In the last decade – partially in response to funding from the 2009 American Recovery and Reinvestment Act (ARRA) – many utilities have rolled out so-called “smart meters” to even residential customers. Estimates vary, but by 2013 it is likely that more than 40% of all customers in the U.S. had smart meters.27

These meters record total electricity consumption in hourly or shorter periods, and can facilitate much wider use of pricing that changes frequently to reflect real-time supply/demand balance, known as dynamic pricing.28 So far, such granular and timely pricing has appeared for only a narrow slice of large industrial and commercial customers, but with smart meters now in place, most of the financial cost of dynamic pricing down to even residential customers has been sunk. Still, there remains substantial resistance to dynamic pricing among residential consumers and groups that represent them. Data from a 2012 EIA survey of utilities suggest that only a few percent of customers are on tariffs that have any dynamic pricing component.29

Of course, the efficiency gain from dynamic pricing depends on the ability and

27 See FERC (2013).
28 The meters also communicate information to the utility without onsite visit by a meter reader. Savings on meter reading labor have been the largest benefits projected by utility installing smart meters.
29 See FERC (2013) and EIA Form 861.
willingness of customers to respond to those prices. Opponents have generally argued that households won’t pay the attention necessary to adjust thermostat settings, washer/dryer use, and other electricity-consuming activities in response to dynamic prices. Simple calculations, such as in Borenstein (2013), show that the financial gain from paying attention to such price fluctuations has been modest. Still, increased penetration of intermittent resources is likely to increase at least wholesale price volatility and raise the social return to such attention, while automation is likely to continue lowering the cost of the necessary attention.

A very large literature has now developed using randomized control trials, randomized encouragement designs and quasi-experiments to analyze just how much consumers do respond to dynamic pricing. The evidence is fairly consistent that even without automation, customers respond significantly on average, though with a fairly small elasticity, generally estimated to be in the -0.1 to -0.2 range.30 The research suggests that the larger elasticities result from interventions that include technology to convey information, such as emails, text messages, and in-home electricity usage displays.

The literature on elasticity with automated demand response is much thinner; it is pretty much non-existent in economics outlets. But programmable controllable thermostats – which can permit a person to automate response to a price or other warning signal or allow an authorized third party to do so – have been in use for more than a decade. Industry publications suggest these technologies greatly increase potential demand response.31

IV. The Next 20 Years

After a tumultuous period from 1996 to 2005, the regulatory/legal status of electricity restructuring – in generation, transmission, distribution and retailing – has changed little in the last decade. In recent years, however, the continuing

30 See Jessoe and Rapson (2014), Ito (2014), and Wolak (2011b).
31 See Faruqui and George (2002).
evolution in technology and in environmental concerns has disrupted the industry in new ways. These changes are ongoing and are likely to continue for many years.

The greatest change occurring in electricity markets today – and likely going forward for many years – is the increased recognition of environmental costs of electricity generation, most notably (but not exclusively) greenhouse gas emissions. Environmental issues have played a significant role in electricity for decades, but most of the emphasis in past years was on limiting the local air and water pollution from traditional generation sources. Of course, appropriate pricing of the environmental externalities – either through a tax or a cap-and-trade program – would be the simplest and most efficient way to incorporate these environmental costs. Currently, all U.S. utilities either pay zero for their greenhouse gas emissions or pay prices well below the most common estimates of the social cost of those emissions. In that situation, raising marginal retail price above the utility’s private marginal cost can be efficient, of course, and it can at the same time reduce the need for fixed charges discussed earlier.

In the last decade, with growing concern about climate change and with improving technology, environmental stakeholders have turned more and more to goals for increasing generation from renewable sources. While hydro-electric and nuclear generation are by far the largest low-carbon sources in the U.S., wind and solar are growing rapidly, as shown in figure 7.

The growth of wind and solar generation sources raises two issues that are now coming to dominate policy discussions among utilities and policy makers: (1) economic and technical management of intermittent-production resources for which costs are largely sunk before production begins and (2) policy towards distributed generation resources that are on the property of the end user (so-called “behind the meter” generation). The latter is primarily an issue with rooftop solar PV today, but could expand to batteries and other generation or

---

32 “Appropriate” is a key word here. Simply setting a tax or a quantity cap addresses the issue efficiently only if the tax or quantity limit is set correctly. This is an obvious point, but one that seems to be missed or ignored by many policymakers.
storage devices in the future.

A. Management of Intermittent Generation Resources

Numerous regulatory and legislative initiatives, including President Obama’s Clean Power Plan proposed in 2014, are pressuring electricity providers to reduce the greenhouse gas footprint of the power they supply. Many options exist for reducing GHG emissions from electricity, but among the most prevalent today are greater use of wind and solar power. Economic and technical integration of these intermittent renewable generation resources is likely to be one of the principal challenges facing the electricity industry in the next few decades.

The technical challenge stems primarily from the fact that production from these resource occurs intermittently and largely outside the control of the owner – when the wind blows or the sun shines.\textsuperscript{33} Because the physics requires that

\textsuperscript{33}In reality, these resources can be adjusted downward, just not upward if wind or sun aren’t present. Both wind and solar PV are potentially curtailable, but require communication between the grid operator and the resource. Wind turbine blades can be positioned so as not to catch the wind and stop turning. Solar PV curtailment requires a smart inverter that can be told to disconnect the PV system from the
quantities supplied and demanded in an electrical grid must balance at all times for the system to be stable – and because storage is still quite expensive – the intermittency of wind and solar implies that either other flexible supply resources must be available to offset these fluctuations or demand must change in response. Both solutions are technically feasible, though supply-side responses have been the focus of more discussion.

Intermittency problems occur on both short and long timescales. Large fluctuations in electrical generation can occur second by second from solar PV, and minute by minute from wind. On a longer scale, both wind and solar can exhibit many hours of higher or lower production than was forecast even a day in advance. Short scale intermittency is generally localized and idiosyncratic, so a diversity of locations may substantially mitigate the problem, though studies suggest that some additional balancing resources or demand responsiveness will still be necessary at high penetration.

Longer-scale intermittency is likely to be a more formidable problem if wind and solar become a large share of generation capacity. Absent inexpensive electricity storage, days or weeks without much sunshine or wind would create energy supply fluctuations that would be very costly for demand to follow. If the existence of those days requires full or nearly-full capacity coverage from conventional fossil resources, then the full cost of supplying power with high renewables penetration grows significantly.

Further complicating the technical challenge, conventional fossil generation is constrained in how quickly it can “ramp” output up and down to offset large changes in output from renewable resources. In general, the most flexible conventional generation is from gas-fired peaker plants, which are also the least efficient and most expensive. Larger combined-cycle gas turbine plants are somewhat less flexible, but lower cost, and coal and nuclear plants are the least flexible.

A well-know concern is illustrated in what has become known as the “duck grid. The inverters currently on nearly all residential and small commercial systems do not have this
Figure 8. Projected Hourly California Demand and Solar Production (on a sunny, cool March weekday) with Increased Solar Penetration

The duck chart shown in figure 8. The duck chart presents the forecast total demand and net demand for the California electricity grid on a sunny spring day with high penetration of solar PV. The lowest line shows the net demand after subtracting solar PV generation from total electricity consumption with solar penetration projected for 2020. Even if solar generation were perfectly forecastable, the rapid drop in net demand as the sun rises and increase in net demand as the sun sets would be difficult to meet with the current mix of gas-fired generation in California.

The most cost-effective solution proposed by a recent study would be to run more gas-fired plants in the middle of the day and curtail production from solar PV. In other words, the least costly engineering solution at this point may be to forego electricity that has zero marginal cost. It seems quite possible that if retail prices at these times were set at or near zero to reflect this situation,
consumers would find innovative ways to use nearly costless electricity, but that requires adoption of high-frequency, time-varying pricing. While such pricing is completely feasible with current smart-meter technology, it has not been widely adopted, as mentioned earlier. In this way, technical challenges to integration overlap a great deal with economic policies.

Further economic challenges arise with the addition of subsidized renewable resources, because they change the economic returns to conventional generation. The most notable change is that because solar and wind generation have near-zero marginal cost they are generally used virtually all the time they are available. This pushes out the supply curve and lowers the market clearing price for electricity, reducing profits for all conventional generation in the market. In the longer run, this worsens the economics of conventional generation and can lead to exit. All of that would be a description of an efficiently operating competitive market if no generation sources were subsidized, all sources paid their full social marginal cost, and electricity prices reflected the social value of marginal production at every point in time. However, renewable generation costs are artificially low due to investment and production subsidies, while conventional generation does not pay for its negative pollution externalities. And wholesale prices do not reflect the value of marginal power at a specific point in time or space; instead, the system operator separately arranges for electricity needed to maintain voltage in specific areas, to offset fluctuation of intermittent resources and for other operational constraints, and to respond to un-forecasted demand volatility. One of the common ways to assure needed capacity does not exit is through capacity payments, which generally pay companies to have generation available regardless of the electricity it is called upon to generate.\textsuperscript{38}

\textsuperscript{38}See Joskow (2008) for a broad overview of the role of capacity payments.
B. Policy towards distributed generation

Cost reductions in solar PV technologies have also changed the economics of self-generation by end-use customers, known as distributed generation. In California, Hawaii and other sunny locations with high electricity prices, falling PV system costs have combined with substantial federal and state subsidies to make installing solar PV a money saver for some customers. The result has been a booming market in behind-the-meter solar PV. In the U.S., distributed solar PV capacity installation has increased from 400 MW in 2009 to about 1900 MW in 2013, with about half of new installations occurring in California.\(^\text{39}\)

This trend has led some observers and utility executives to predict a “death spiral” in which a significant number of customers self-generate much of their electricity, forcing the utility to raise rates for the electricity they still sell in order to cover fixed investments, in turn making solar PV economic for a larger set of customers who then reduce their purchases, leading to a greater revenue shortfall and another rate increase, and restarting the cycle. Ultimately, some argue (some gleefully), the monopolist utility disappears. This scenario has triggered widespread debate – both positive and normative – about the future and viability of the utility. The regulator in New York state has even proposed a complete redesign of utility systems that is focused on customers also being generators.\(^\text{40}\)

The social welfare gain from increasing reliance on distributed PV generation, however, is still far from clear. Even the most optimistic cost scenarios suggest that the full social levelized cost of electricity from residential solar PV is likely at least $0.20/kWh in relatively sunny areas, more than double the full cost of gas-fired generation including a greenhouse gas cost of $40/ton.\(^\text{41}\) Distributed PV generation is eligible for the same tax benefits as large scale solar, a 30%
tax credit through the end of 2016 and accelerated depreciation. Borenstein
and Notsund (2014) estimate that the accelerated depreciation amounts to an
additional effective subsidy of about 15%.42

Distributed PV generation also benefits from being compensated at retail prices
for the power it produces. Under “net metering,” which has been adopted in
most of the U.S., customers are credited for all power produced from their PV
system by deducting the quantity from the customer’s consumption.43 In reality,
calculations by Darghouth, Barbose and Wiser (2013) suggest that less than half
of the power produced by a typical household PV system is consumed onsite –
actually reducing the customer’s retail demand – but net metering treats all power
as demand reduction, thereby crediting it at the retail rate the customer would
have paid.44 If the full benefits that DG solar PV power brings to the market are
less than the marginal rate the customer pays, then net metering policies lead to
over-compensation of DG solar production. A simple calculation suggests this is
very much the case, but the full system benefits are a matter of some dispute.45
What is clear is that retail electricity rates are set in ways that are not closely tied
to long-run marginal cost, so incentivizing DG solar through net metering will
conflate solar policy with rate design policy and will have unpredictable effects
on the incentive to install residential solar.

Probably the clearest illustration of perverse incentives from net metering policy
is in California, where more than half of U.S. residential PV has been installed and
where the gap between marginal retail rates and marginal cost may be highest.
Most California utilities use increasing-block residential electricity pricing, mean-

---

42 Actually, the accelerated depreciation benefit is available only if the system is owned by a company,
not an individual. This has been a significant factor behind the rapid growth of third-party owned resi-
dential systems in which the third-party owner leases the system to the homeowner or, more commonly,
sells the electricity from the system to the homeowner. Third-party owners of these systems point out
that this model also greatly lowers, or eliminates, the up-front payment the homeowner would otherwise
have to make.

43 See http://dsireusa.org/solar/solarpolicyguide/?id=17 for timely information on U.S. state net me-
tering policies.

44 This is for a system that generates electricity equal to half of the household’s demand. The figure
could be even lower for a system that is larger relative to household demand.

45 See Borenstein (2012) and Cohen and Callaway (2013)
ing that the marginal price a customer faces increases in steps as the customer’s consumption increases during the billing period. The two largest California utilities, each of which has an average residential retail price around $0.18/kWh, have four blocks in their residential tariffs with prices from about $0.12/kWh up to more than $0.35/kWh on the highest block. Borenstein and Notsund (2014) report that a greatly disproportionate share of California households installing PV from 2007 to 2013 had consumption levels that reached into the two highest price tiers. They also find that over time installations have been calibrated more precisely to eliminate consumption on the highest-price tiers, but not to crowd out the lower-price consumption. Borenstein and Notsund (2014) estimate that the average bill savings from installing DG solar for customers of these utilities was about 20% greater due to increasing block pricing than it would have been if the utility charged a flat rate equal to their average residential price per kWh. They estimate that the bill savings were more than double what they would have been if the utilities had charged $0.10/kWh, a rough approximation of social marginal avoided cost.\footnote{The best estimates of long-run marginal cost from gas-fired generation is about $0.06/kWh as mentioned earlier, but DG solar PV consumed onsite also avoids the 7%-9% of electricity that is dissipated through line losses as the power flows from generation through transmission and distribution lines to the end user. See Borenstein (2008). Accounting for line losses, the electricity delivered for consumption from conventional generation has a marginal cost closer to $0.065/kWh. The timing of power from solar PV also boosts its value, or the cost of alternative sources. Solar PV generation produces more at peak times, so it is replacing power at times when marginal electricity costs are higher. Borenstein (2008) estimates that in real-world grid operation this increases the cost of the alternative power source by an average of 20%, bringing marginal cost of alternative generation to around $0.078/kWh. Inclusion of the cost of GHG emissions raises the cost of alternative generation by $0.015-$0.02 per kWh at a GHG price of $40/ton, bringing the alternative marginal cost to about $0.10.}

Talk of a “death spiral” and questions of the viability of utilities, however, raises a question that extends far beyond these issues of implicit and explicit subsidies and the value of incremental DG solar generation. Can DG really function without the grid? Without low-cost electricity storage, and tolerance of less reliable electricity at some times (\textit{e.g.}, a week without sunshine), it seems unlikely that most customers will be ready to operate off the grid anytime soon. If the grid is needed, how should it be paid for? The model to date has been based on volu-
metric average cost pricing. Distributed generation at this point looks very much like the push for restructuring discussed in section 2: a comparison of average cost to marginal cost that ignores that the difference is not a real savings, but rather cost shifting. To the extent that a DG solar household has costs greater than or equal to the social marginal cost of grid-supplied electricity, the private savings are offset, or more than offset, by a revenue shortfall at the utility. That shortfall must then be made up by utility shareholders or, more likely, remaining rate payers. In fact, the notion of a death spiral – with rising retail rates as consumption declines – necessarily implies that price is above marginal cost, and an excessive incentive to install DG.

V. Summary

The changes in the electricity industry over the last two decades have been dramatic, but many were not the changes that were anticipated at the beginning of the industry’s grand experiment with market-based pricing of generation and retail services. While the revenues for much of the nation’s conventional and nuclear generation sources are now based upon market prices rather than production costs, retail pricing for the vast majority of residential customers remains dominated by state regulatory processes.

In the mid-1990s, the strong momentum for restructuring was driven by a large gap between market-based prices – which were based upon marginal cost in competitive markets – and regulated rates – which were based upon average production costs. During this period of relatively large capacity margins and low natural gas prices, market-based pricing appealed to customers and terrified utility shareholders whose assets would become stranded absent other compensation. However, despite the allure of market-based pricing, the reality of the regulatory process, and of case law, dictated that utilities be allowed to recover the bulk of what appeared at the time to be stranded costs.

The great irony of this period is that a half decade after transition arrangements
largely compensated utilities for the losses incurred in selling or transferring these assets, the market value of those same assets had fully recovered. By the mid-2000s the relationship between average and marginal cost had largely reversed, and many states expressed a great deal of regret about the decision to restructure. However, since the formerly regulated generation assets were now largely held by private, deregulated firms, there was no clear path to dramatically “re-regulate” the industry without paying full market value for those assets. Looked at this way, one can view the disappointment with restructuring as being driven by magnificently poor market timing. Utilities sold off their assets at the nadir of their value, and as natural gas prices climbed throughout the 2000s, those assets became quite valuable under market-based pricing.

Since 2009, this story has largely reversed yet again. Natural gas prices have declined sharply, nearly to the levels seen at the dawn of the restructuring movement. The attention of policymakers has now been consumed by environmental priorities, particularly the implications of coal generation decline and renewable generation growth for costs and greenhouse gas emissions. A surge of subsidized renewable generation, combined with low natural gas prices, has driven wholesale prices steadily lower. As one would expect, in the short run this has benefited consumers in market-based states disproportionately more than those in regulated states.

Going forward, the role of intermittent renewable generation at both the wholesale and distributed level is likely to continue to dominate the economics and policy of the industry. The low wholesale prices that have resulted from expansion of subsidized renewables are not sufficient to cover the total cost of renewable or conventional sources, so the prominence of extra-market sources of revenue – such as above-market contracts and capacity payments – is likely to continue to grow. This will mean that even in the “market” states, the true cost of supply will increasingly diverge from the underlying price of the fundamental commodity, electrical energy.
At the retail level, distributed energy threatens to unravel the economics of retail distribution supply. Again the juxtaposition of average and marginal cost is a driving force here, although the differences are exacerbated by inefficient rate-making and political economy. Current rate-making practices encourage individuals to install distributed generation, such as solar PV, that is privately economic because rates, which include the fixed costs of transmission and distribution, exceed the marginal cost of generated energy by a comfortable margin. The next natural step in the rate-making process will be a move to two-part tariffs that include monthly charges decoupled from the volume of electricity consumed. There is speculation that the cost of storage technologies, perhaps deployed in a joint application such as electric vehicles, could decline enough that households might bypass the grid completely.47 Such an outcome would be a giant leap forward in technology, but it could be a step backward in economics if such decisions would again be motivated by an ability to shift sunk costs – this time of grid assets – onto other customers or utility shareholders. Policymakers again have a chance to make economically rational decisions based on true incremental costs. We can only hope that this time they will grab that opportunity.

VI. References


47Lacey, 2014


Kleit AN, Reitzes JD. 2008. The Effectiveness of FERC’s Transmission Policy: Is


