Carnegie Mellon University
Carnegie Institute of Technology

Thesis

Submitted in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

Title: Meeting Electric Peak on the Demand Side: Wholesale and Retail Market Impacts of Real-Time Pricing and Peak Load Management Policy

Presented By: Kathleen Spees

Accepted by the Department of: Engineering and Public Policy

Professor Lester Lave, Advisor | Date

Professor M. Granger Morgan, Department Head | Date

Approved by the College Council

Dean | Date

Meeting Electric Peak on the Demand Side

Wholesale and Retail Market Impacts of Real-Time Pricing and Peak Load Management Policy

A Dissertation Submitted in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

Kathleen Spees
September 2008

Advisor
Lester Lave

Thesis Committee
Jay Apt
Seth Blumsack
Keith Florig
Lester Lave
Rahul Tongia

Carnegie Mellon University
Carnegie Mellon Electricity Industry Center
Department of Engineering and Public Policy
Pittsburgh, Pennsylvania 15213 USA

kathleen.spees@gmail.com
kspees@cmu.edu
412.445.2694

Front Matter

Acknowledgements

My graduate research was funded in part by the National Science Foundation Graduate Research Fellowship Program, and in part by the Alfred P. Sloan Foundation and the Electric Power Research Institute through grants to the Carnegie Mellon Electricity Industry Center. I further acknowledge Jeanne and Dick Berdik and Bev and Steve Elliott of the Achievement Rewards for College Scientists Foundation of Pittsburgh for their financial support.

I would like to thank Lester Lave for advising my doctoral research and acknowledge him as the co-author of the papers representing the work in this thesis. I would like to thank Jay Apt, Seth Blumsack, Keith Florig, and Rahul Tongia for their assistance in completing this work as members of my thesis committee and for all of the professional guidance that each has provided.

For helping me to acquire valuable data and industry insights, I would like to thank Tom Hurkmans of BottomLine Resource Technologies, and Ruth DeLost, Ribeka Kovach, Karen Murray, Morgan O'Brien, and Bill Pfrommer of Duquesne Light.

I would like to thank Mike Berkenpas, Lee-Ken Choo, Steve Fienberg, Lis Gilmore, Lee Gresham, Marija Ilic, Warren Katzenstein, Ryan Kurlinski, Drew Miller, Sompop Pattanariyankool, Costa Samaras, Fallaw Sowell, Chris Weber, and Le Xie, for the thought-provoking discussions, ideas, sources, and data that they have shared.

I would like to thank Barbara Bugosh, Vicki Finney, Dolores Matthews, Patricia Porter, and Patti Steranchak for all of the many times that each has assisted and advised me.

Finally, I would like to thank my parents Marsha and Andy Spees, my grandmother Elizabeth Weymouth, my dear friend Jessica Fauver, and my boyfriend Drew Miller for all of their love and support.

Abstract

Traditionally, the participation of customers in the electric market has been weak or non-existent. Almost all customers have paid a flat rate for power without variations based on the time of their consumption, so these customers have had no incentive to reduce their usage during times of capacity shortage and very high wholesale prices. Perhaps even more importantly, customers have not participated in forward decisions about whether it would be better to build additional capacity at very high cost or to commit to peak load reductions during a few peak hours each year. In this thesis I present the status of efforts to incorporate customer decisions into the electric market place and calculate the possible system benefits.

In Part I I discuss recent activities relating to demand response and demand-side management. Although interest in demand response is growing among policy-makers and industry participants, the process of making this possible will be a complicated navigation among the incentives of involved parties and the jurisdictions of state and federal regulators. One of the key problems in developing a coordinated policy is that the wholesale markets covering generation and transmission are under the jurisdiction of the federal government represented by the Federal Energy Regulatory Commission while electric distribution and retail markets are under the jurisdiction of the state, represented by state public utility commissions (PUC).

In Part II I investigate the value to the system of reducing peak demand and compare this value to the current costs of peak load reductions. Peak load reductions are currently being achieved at \$21/kW·y, or less than one fourth of the \$94/kW·y it costs to build new capacity. Similarly, energy efficiency is being achieved at \$29/MWh, or roughly one third of the \$92/MWh retail price for electricity. At current rates, peak load could be cost-effectively reduced by some 17%, although I expect that at greater levels of peak reductions the marginal cost of achieving more reductions will increase, it is clear that significant peak load reductions can be achieved cost-effectively.

I further investigate the value to the system of shifting the burden of uncertainty in peak load on to customers and the utilities acting on their behalves who have the most ability to determine

what peak load will be. The traditional means of accounting for uncertainty in peak load has been to build enough excess capacity that the chance of shortages is low. I calculate that a right-sizing peak capacity to the best estimate of peak load would reduce the amount cost of supplying capacity by 8.5% below the current level.

In Part III I investigate the short-run economic impacts of a policy change from flat-rate retail electric pricing to real-time prices (RTP) or time-of-use (TOU) prices. If retail prices reflected hourly wholesale market prices, customers would shift consumption away from peak hours and installed capacity could drop. I use hourly price and load data from Pennsylvania-New Jersey-Maryland Regional Transmission Organization (RTO) to estimate consumer and producer savings from a change toward RTP or TOU. Surprisingly, neither RTP nor TOU has much effect on average price under plausible short-term consumer responses. Consumer plus producer surplus rises 2.8%-4.4% with RTP and 0.6%-1.0% with TOU. Peak capacity savings are seven times larger with RTP. Peak load drops by 10.4%-17.7% with RTP and only 1.1%-2.4% with TOU. Half of all possible customer savings from load shifting are obtained by shifting only 1.7% of all MWh to another time of day, indicating that only the largest customers need be responsive to get the majority of the short-run savings.

Placing customers on an RTP can benefit them through lower average rates for energy and capacity, but the advanced metering infrastructure (AMI) required to make RTP and customer response possible is a large investment. In Part IV I determine how many customers can be cost-effectively placed on RTP from the perspective of a PUC. I calculate that for wide scale implementation of AMI, all customers above 2.5 kW in coincident peak load (about 40% of all customers, representing all industrial, all commercial, and large residential customers) could be cost-effectively placed on RTP if there are no benefits to the AMI other than demand response from RTP. For the customers below size 0.31-0.73 kW (the smallest 10%-20% of customers, representing small residential loads), installing an AMI is not cost effective even under the most favorable assumptions about other AMI benefits and highly responsive customers. For intermediate-size customers the investment would be justified in some cases but not others.

Table of Contents

List of Figures

Body Figures

Appendix Figures

List of Tables

Body Tables

Appendix Tables

Part I Introduction to Electric Demand-Side Issues

Part I of this thesis introduces the state of affairs in policy and research on demand response and demand-side management[1].

Chapter 1 The Importance of Price Responsive Demand

Historically, end users have had few or no opportunities to interact with electric market. Most, although not all, customers have paid a flat price for electricity no matter when they use it, even during times of peak electric use when wholesale prices skyrocket and capacity shortages threaten system stability. Even beyond these short-term considerations, customers have not been engaged in the process of deciding how much generation capacity to build now in order to supply their future peak electric use.

Hourly, daily, and seasonal fluctuations in consumer demand require additional generating capacity, particularly peaking plants that were needed only a few hours per year. If these fluctuations are treated as facts of life where load must be served at lowest cost, then the traditional utility would build baseload plants, usually coal, with high capital costs and low operating costs to run most of the time; they would also build peaking plants, usually gas plants, with low capital costs and high operating costs to run only a few hours per year. Therefore peaking plants add cost in two ways: first, their operating costs are much higher than average, meaning that the marginal cost of supplying electric energy is very large during peak events; second, even though the capital costs of these peakers are low on a per kW basis, their very low capacity factors result in very high capital costs per unit of peaking MWh produced..

Under regulation, the cost of peakers was spread over all kilowatt-hours generated, adding little to the average cost of producing power, and therefore to customer price. Even though the capital

[1]The substance of Part I was published under Spees and Lave [1].

costs of these peakers represent only a small addition to the average cost of power, they do represent a large total amount.

To avoid some of these costs, some utilities, under the oversight of their state Public Utility Commissions (PUC), have implemented load management programs, with examples of radio-controlled end-use devices going back to the 1930s [2]. These load management programs would alleviate both capital and operating costs associated with peaking plants.

A more advanced load management system would benefit the system by allowing customers respond to real-time system conditions and real time prices[2] (RTP) as Fred Schweppe envisioned decades ago [3-5]. However, despite decades of advances in technology and more recent developments in the industry structure, Schweppe's vision of a dynamic demand-side electric marketplace has as yet failed to materialize.

Industry restructuring has breathed new life into demand response and generated a wide range of demonstration projects and pilot programs [6]. Many market operators in the United States have developed initiatives to invite demand into the marketplace, but enrollments have been small and sluggish. Market operators publish lists of private parties who provide demand response services, but only a few end users currently employ these services [7, 8]. I explore here the obstacles that public regulators and private ventures face in developing the load-side resource and also the possible benefits to be had.

Market restructuring turned the issue of high peak demand into a major problem. On the demand side, the systems operators run markets that represent customers who are presumed not to want to alter their electricity use, no matter how high the price. Thus customers face a fixed retail

---

[2] I will use RTP to refer to any combination of day-ahead and balancing market prices and distinguish among these only where relevant. I do note however that I am usually referring to a price that is essentially represented by the day-ahead market price as discussed in Chapter 12.

price, e.g., \$0.10/kWh[3], even when the wholesale price hits its maximum of \$1/kWh. A customer has no reason not to use an electric dryer at 5 PM on the hottest day in August because she always pays the same \$0.10/kWh. If the customer faced the wholesale market price of \$1/kWh, she would demand much less electricity at that price. Once consumers have the technology to respond to day-ahead or balancing market prices, they would be able to reduce consumption during these hours and mitigate the high price extremes that we see in current wholesale energy markets.

On the supply side, the Independent System Operators (ISOs) and Regional Transmission Operators (RTOs)[4] determine the price in an auction market with all successful generators paid the locational market clearing price (capped at \$1000/MWh[5] in most RTOs). All generators receive this price, from a baseload nuclear plant generating power at a marginal cost of \$20/MWh to an expensive light oil generator at \$240/MWh (which operates only a few hours per year) [11]. In a competitive wholesale market, baseload plants can earn high profits during the high demand periods in a competitive market, but, if the market clearing price reflects the marginal cost of the most expensive peaker running as intended, the highest cost peaking unit only receives its marginal costs and cannot cover its fixed costs.

Recognizing the problem of peakers being unable to cover their fixed costs through energy markets alone, some market operators have introduced installed or forward locational capacity markets, with nettlesome early results and rapidly evolving market rules. Another way to

---

[3] Retail prices are higher than wholesale prices because the retailer ads an additional amount for billing and local distribution.

[4] I will use RTO to refer to both ISOs and RTOs. When referring to a single state or multi-state entity, the acronym ISO or RTO will be used as appropriate.

[5] ERCOT is one exception with the recent increase of its price cap to \$2,250/MWh [9]. California ISO is another exception with a \$400/MWh soft cap on energy and ancillary service bids [10]. Generators may bid above a soft price cap and will be paid as bid; other generators will receive payment only as high as the cap. The neighboring Western Electricity Coordinating Council (WECC) has the same price caps although WECC is not a market operator.

recover peaker costs proposed by economists such as Bill Hogan is to remove all price caps and allow very high prices during peak hours of capacity shortage [12].

Pennsylvania-New Jersey-Maryland (PJM) RTO, ISO New England (ISO-NE), and New York ISO (NYISO) have created capacity markets to pay for fixed costs. Although these market structures might provide a price incentive for suppliers to build new peaking capacity, the structures will ultimately be economically efficient only if the price signal also reflects the true demands of customers. Even if a reasonable wholesale market structure for incenting peaking capacity investments were to materialize through the RTOs' rapidly evolving market structures, customers' demands would not be accurately represented on the demand side. Current capacity market "demand curves" are not gathered based on information from customers or their representative load-serving entities (LSE) at all, but rather are developed by the RTO staff. The curves are developing using engineering estimates of the cost of peaking capacity, a target level of capacity based on traditional resource planning methods, and an essentially arbitrary shape and slope for the downward-sloping "demand curves" [13, 14].

The essence of this problem from a policy standpoint is that the regulation of wholesale markets for generation and transmission is under the jurisdiction of the federal government through the Federal Energy Regulatory Commission (FERC), while the regulation of electric distribution and retail rates is under the jurisdiction of state lawmakers represented by state PUCs. This means that an appropriate wholesale market structure would not translate into an appropriate retail market structure without additional legislative action in each state. The level of complication and muddling through that will be required in order to create efficient markets for peaking capacity is daunting at a time when many states are doing their best to grapple with the effects of rapid fuel price increases and adjust to the consequences of a first wave of retail restructuring [15].

State regulators will have to navigate through the traditional commitment to finding the lowest rate possible for utility customers, the physical reality of one set of distribution lines, the promises of innovative retail structures with electric choice, and the lack of control over wholesale structures. The trick for state regulators will be to allow retailers enough latitude and flexibility to provide retail customers the array of retail agreements to best meet true demands and represent the fast wholesale market structure changes without giving retailers the ability to exploit the natural monopoly created by the physical distribution and metering system.

This thicket of policy issues and business incentives will be difficult to straighten out, but I do not see the problems as intractable. I lay out these problems and opportunities for utilities, state regulators, and federal regulators more fully in Part I and place them within the historical context of related regulatory efforts, beginning with a discussion of traditional demand-side management.

Chapter 2 Conservation Initiatives and Effectiveness

Electricity conservation policies since 1975 have been expensive but cost-effective. A recent Resources for the Future (RFF) retrospective estimated expenditure and savings numbers from large federal energy efficiency efforts with results shown in Table 2.1 [16]. Voluntary programs appear to have energy savings on the same scale as some mandatory programs with small federal government costs, but voluntary program results are uncertain and difficult to verify. Mandatory residential appliance standards and utility demand side management (DSM) programs both show benefits at more than twice the cost even without considering environmental costs.

Table 2.1. Slice-of-time program costs and benefits for the year 2000, 2007$[6] [16].

Program			Energy[7] Savings Quads/Year	Program Costs		Cost-Effectiveness \$/MWh[8]	Retail Price[9] \$/MWh	Benefit-Cost Ratio[10]
				\$Billion/ Year	Costs Reported			
Mandatory	Appliance Standards	Residential	0.77	\$2.89	Consumer, manufacturer	\$43.78	\$99.22	2.27
Mandatory	Appliance Standards	Commercial	0.43	--	--	--	--	--
Mandatory	Utility DSM		0.62	\$2.05	Utility	\$38.57	\$82.00	2.13
Voluntary	Energy Star[11]		less than 0.93	\$0.06	Government	--	--	--
Voluntary	1605b Registry		less than 0.41	\$0.0004	Government	--	--	--
Voluntary	DOE Climate Challenge		less than 0.81	--	--	--	--	--

6 These dollar values are updated to 2007$ using Bureau of Labor Statistics inflation data [17]. Costs and benefits result from all programs or standards up until the year 2000, the numbers have been annualized so that the costs and benefits can be viewed over a one year slice of time.

7 Energy savings are reported in quadrillion BTUs of source energy.

8 Cost effectiveness numbers are reported assuming that all energy is converted to electric energy. A conversion factor of 11660 $BTU_{source}/kWh_{electric}$ corresponding to a conversion efficiency of about 29% was used.

9 Year 2000 residential and average retail prices are reported for comparison with residential appliance standards and utility DSM programs respectively, although the prices are updated to 2007$ as are the rest of the numbers [18].

10 Benefit:Cost ratio compares benefits accrued to the end user to costs reported. Environmental benefits and costs to unlisted parties are not considered.

11 Energy Star cost and savings numbers are reported for the year 2001; all other program numbers are reported for the year 2000.

2.1 Efficiency Standards

Conservation activists insist that appliance efficiency regulations are needed because consumers notice an increase in purchase price but give less attention to the lower electricity payments over time. Regulations initiated in California and other states were later adopted at the federal level. Standards covering devices from washing machines to exit signs to mobile homes have had a large impact on end user efficiency. Federal appliance efficiency standards began in earnest with the sweeping 1987 National Appliance Energy Conservation Act and have been supplemented and updated frequently [19].

Table 2.1 shows that the year 2000 residential savings from appliance efficiency standards are estimated to be $43.78/MWh, less than half the retail residential electricity price of $99.22/MWh [16, 18] [12].

Building efficiency codes have developed similarly, with an indispensable role played by professional societies. In 1977 the American Society of Heating, Refrigerating, and Air-Conditioning Engineers (ASHRAE)[13] and the Council of American Building Officials (CABO)[14] developed initial versions of their energy codes for commercial and residential buildings respectively. Every state had instituted a building energy code based on one of these standards before the 1992[15] Energy Policy Act mandated them [21]. Given the high level of technical

[12] Original RFF and EIA numbers were both converted to 2007 dollars. The EIA number refers only to retail residential sales; commercial and industrial rates were $76.7/MWh and $47.9/MWh respectively.

[13] The ASHRAE 90.1 series with its periodic updates has been widely adopted for commercial and high rise residential facilities. Forty states have adopted a version of this code [20].

[14] The CABO developed the original Model Energy Code (MEC) for residences, which is now the International Energy Conservation Code (IECC). Some version of this code is enacted in 40 states. The list of noncompliant states is not identical between commercial and residential sectors [20].

[15] Federal standards were again updated in the 2005 Energy Policy Act to reflect the most recent versions of these codes, ASHRAE 90.1-1999 and IECC.

complexity and domain expertise necessary to develop and maintain these standards, the roles of ASHRAE and CABO have been essential.

2.2 Demand Side Management

In the mid 1970s, California and Wisconsin ordered utilities to work with customers to increase energy efficiency. Congress picked up DSM in the 1978 National Energy Conservation Policy Act [22]. Utilities were expected to treat peak demand reduction as an alternative to capacity growth from an integrated resources planning (IRP) perspective. During the next decade the meaning of DSM evolved to incorporate efficiency as well as load profile management. Since utilities were compensated for their DSM programs and reported energy savings without a detailed audit, some analysts were skeptical of the reported savings, but Parfomak and Lave used ex post econometric analysis to verify that 99.4% of the reported savings were statistically observed [23, 24].

Effective DSM programs are expensive and labor intensive; if the administrator of a DSM program wants to ensure that certain measures are taken and implemented according to plan then the surest method is to purchase and install more efficient equipment at many locations. Utility DSM programs grew both in size of expenditure and size of electric energy savings from their conception until their peak expenditure in 1993, partially shown in Figure 2.1 [25][16]. The RFF 2000 cost estimate for avoided energy from DSM programs is $38.57/MWh which shows slightly better performance than appliance efficiency standards with benefit-cost ratios of 2.13 and 2.27 respectively[17] [16]. Although DSM has cost customers less than it has saved them, the programs do require large expenditures to

---

[16] Spending trends are from the EIA as estimated from numbers reported by utilities.

[17] See Table 1.

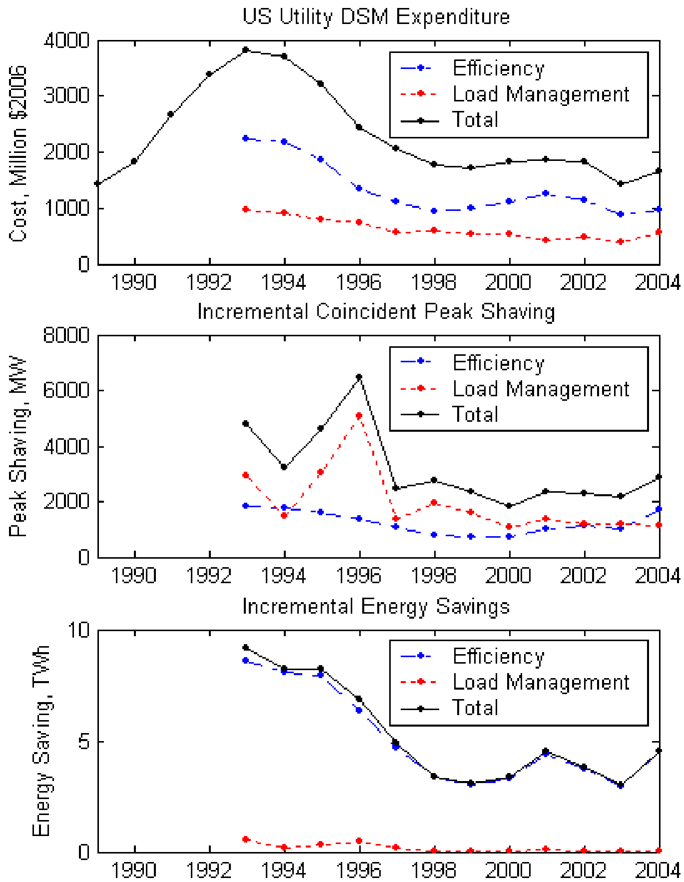


Figure 2.1. DSM expenditure[18] and savings 1989-2004[19] [25, 26].

With industry restructuring, DSM expenditure declined dramatically as shown in Figure 2.1. Restructuring focused on lowering price and there was less ability to hide the program expenditures from customers. Incremental[20] energy and peak savings from efficiency efforts have generated net benefits. From Figure 2.1 it appears that load management expenditure had almost no payoff in energy savings and a volatile relationship with peak shaving. Peak shaving spiked just as much of the industry was preparing for restructuring, even though load

[18] Total utility DSM expenditure includes indirect costs as well as efficiency and load management costs. Indirect costs represent between 8.5% and 17.7% of total expenditure in this time period.

[19] Only total expenditure data were available prior to 1993 because the EIA did not collect the more complete data before that year.

[20] "Incremental" savings are attributable to expenditures in the current year, not from previous years. The EIA also reports "annual" numbers that represent current year savings from all previous investments.

management investments were on a steady decline. This might indicate that utilities were increasingly accountable for coincident peak load. Some federal and state efforts have tried to stem the decline in efficiency investments with public benefit funds such as the Low Income Weatherization Assistance Program, which may account for the increased expenditures on efficiency after 1998.

2.3 Energy Services Companies

The energy services sector was created by DSM programs. Some utilities created subsidiaries for the DSM programs while others contracted with independent companies. In 2000 90% of all energy services company (ESCO) revenues were earned by subsidiaries of an energy company as shown in Figure 2.2 [27]. Although independent ESCOs are numerous, they are not nearly as large as their subsidiary competitors[21].

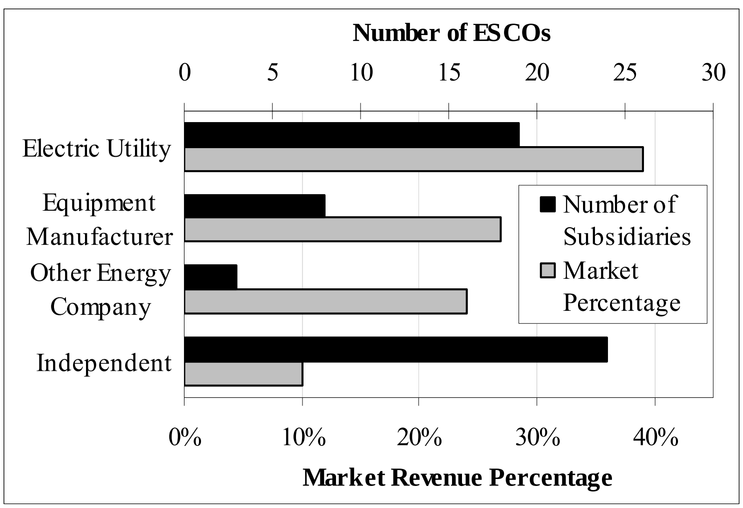


Figure 2.2. Year 2000 market percentage based on 54 ESCOs by parent company type [27].

[21] "Equipment Manufacturers" generally make building equipment and controls; "Other Energy Companies" can be fuel producers, pipeline owners, etc.

Initial ESCO industry revenues were from performance based contracts[22], but have shifted toward packages of services including procurement and risk management. Throughout the entire restructuring period of the late 1990s, ESCOs have continued to grow; market revenues first hit $2 billion in 2000 [27].

Chapter 3 State of Demand Response Technology and Policy

Customers could save money from demand response and load shifting by using less expensive energy. System benefits from economic load response should be larger than a responding end user's benefits per unit, especially in the long run, since they include congestion relief, improved reliability, and a lower capacity requirement.

Day-ahead prices have been used in nearly all related programs and demonstrations to date, allowing the end user time to plan and respond even without having to invest in automated enabling technology. Even though the day-ahead price is a strong predictor of the balancing price under most normal conditions, it cannot communicate unforeseen system conditions such as unplanned outages or other emergencies. System benefits from immediate load curtailment and load shedding in contingency situations can only be garnered from active load management or immediate prices, for example PJM's five-minute balancing locational marginal prices (LMPs) [28]. Immediate response requires automated enabling technology that acts on behalf of the end user in response to an electronic price broadcast. Providing customers with information on both balancing and day-ahead prices would allow both planning and real-time response as long as the retail agreement reflected both numbers.

[22] A performance based contract is an arrangement in which an energy services company will install efficiency upgrades for a client. The client and the energy company then share the savings accrued from the lower energy bills.

3.1 Real-Time Pricing

Electricity retailers, whether traditional utilities or alternative generation suppliers under retail choice, buy electricity from the wholesale market and sell it to the end user. Most of the roughly 70 utilities that offer RTPs in the US developed optional programs in the mid-1990s in order to retain large industrial customers under the threat of retail competition or relocation. Other primary motivations were to lower peak consumption, to encourage overall load growth, and to comply with a mandate. These non-exclusive motivations are shown in Figure 3.1 [29]. These utilities tend to offer implicit hedges to protect valuable customers from price spikes.

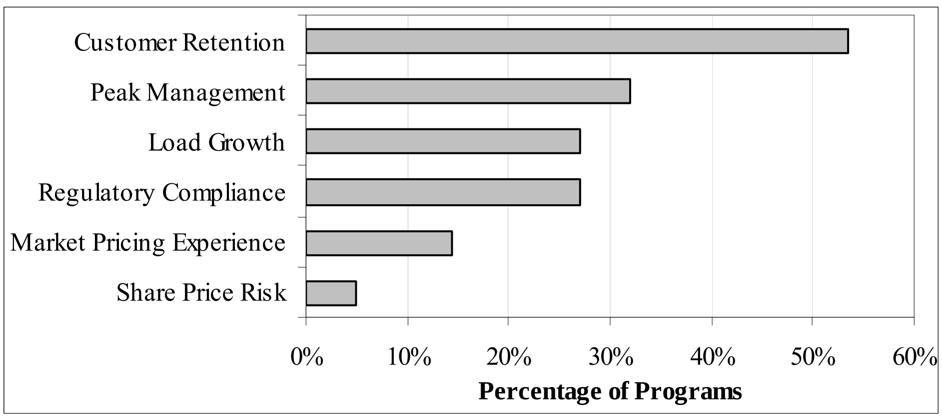


Figure 3.1. Utility reported motivation for offering RTPs to customers.

When some utilities offered all their large customers the option of RTP, they did so knowing that some would pay lower average prices without making any changes. Because some utilities never expected customers to respond to the RTP, it is not surprising that only 35% of them offered technical assistance for RTP response, and only 49% provided customers a way to monitor usage regularly[23]. What *is* surprising is that these utilities have reported 12-33% reduction in participants' coincident peak load even under these circumstances [29].

[23] "Regular" means real-time energy consumption or consumption information from the previous day.

3.2 Economic Load Response

Even if a current consumer paying a fixed tariff learned the five-minute LMP values by looking at the PJM website, the price would be irrelevant since the consumer would face a fixed price. Although operational demand response programs have yet to demonstrate large enrollment, most market operators in the United States offer some combination of economic load response, emergency response, and ancillary service programs as shown in Table 3.1.

Table 3.1. Market operator demand response programs.

Market Operator[24]	Demand Response Programs [30]			
	Economic[25]	**Contingency**	**Ancillary Services[26]**	**Size [31]**
CAISO	None	Voluntary load reduction[27], investor-owned utility curtailment	Non-spinning reserve, replacement reserve, supplemental energy[28]	500 MW in VLRP, Up to 800 MW shaved in 2005
ERCOT [32]	None	Included in ancillary services	All ancillary services	2.5% of total load is registered
ISO-NE	Day-ahead, real-time	Emergency	Investigating stage for operating reserves	Up to 5% of peak demand in emergency
MISO	None	Emergency	None	--
NYISO	Day-ahead	Emergency	Installed capacity or special case	2,300 customers, $75 million in capacity revenues
PJM [33, 34]	Day-ahead, real-time	Emergency	Limited ancillary services including spinning reserve [35]	6,000 commercial and industrial customers, more than 45,000 small customers [34]
SPP	None	None	None	--

---

[24] All of the members of the ISO/RTO Council that are in the United States are examined here. California ISO (CAISO), Electric Reliability Council of Texas (ERCOT), ISO New England (ISO-NE), Midwest ISO (MISO), New York ISO (NYISO), Pennsylvania-New Jersey-Maryland (PJM) RTO, and Southwest Power Pool (SPP).

[25] Market operators that do not offer economic response programs state that they allow price response via a third party intermediary, but do not support such response with electronic price broadcasts.

[26] All ancillary services here require that load have automated response to identical signals given to generators and demonstrate their ability to respond.

[27] CAISO used to offer more programs but has eliminated them as investor-owned utility distribution companies (UDCs) have increased their own load curtailment programs.

[28] Supplemental energy is a near real-time response.

The economic load response programs within ISO-NE, NYISO, and PJM are similar. If a customer is not large enough to interact directly with the wholesale market, it must participate in demand response programs via a licensed curtailment service provider (CSP). Minimum individual or aggregated curtailment is 100 kW in PJM and ISO-NE. At low prices, load usually has the option to respond to day-ahead prices but will be compensated for curtailment only when prices are above $75/MWh in PJM or $100/MWh in ISO-NE. Reporting and metering requirements are extensive; curtailments are verified based on a weather-adjusted customer baseline usage.

Double-counting is implicit in the early versions of these programs because load not only has the choice not to pay for the power, but also receives a payment. The customers that do not participate benefit from lower electricity prices. Curtailment payments do not reflect the systems benefit of response; they were set at an arbitrary level to jumpstart enrollment.

Even though PJM, ISO-NE, and NYISO compute day-ahead and balancing LMPs for every bus in the system, only a subset of these are posted online in real time [28]. All demand response programs are settled at the aggregate zone level. This averaging prevents localized congestion from being reflected and alleviated through demand response. The internet-based communication system used in ISO-NE to transmit the real-time zonal prices might be the most advanced system in operation. Responders in New England can receive up to $2800 in reimbursement for compatible communications devices [36].

Back-up generation can be employed in these programs with proper permitting, but not if the same resource receives capacity payments as then it is treated as a capacity resource rather than as a demand resource.   Table 3.2 shows the sizes of several contingency and market-based demand response programs, many of which are not operated by ISOs or RTOs. Actual curtailments are much higher in contingency programs than they are in economic response programs, possibly because involvement is often binding. Back-up generation serves as a significant but not overwhelming proportion of curtailed load [6].

Table 3.2. Response rates and back-up generation proportion of several demand response programs [6].

Program Type	Number of Programs	Average Curtailable Load, MW	Average Load Curtailed, MW	Percent[29] of Enrolled Load	
				Actually Responded	Back-Up Generation
Contingency	8	158	84	64%	31%
Market	10	204	21	17%	12%

3.3 Load in Ancillary Service

Using load as an ancillary resource is an old idea that has been developed for specific applications from voltage support, to spinning reserve, to stochastic frequency control [37, 38]. National laboratory projects have also demonstrated the technical feasibility of using municipal pumped water and residential air conditioners to provide spinning reserve. Incorporating load as a regulation and reserve resource might become even more important if wind resources grow into a significant generation asset [39-41].

Many enacted projects fall under the category of demand response in ancillary services. Most common among these are emergency load curtailment programs instituted by investor owned utilities [6]. Market operators also employ load shedding under stress; in PJM an emergency responder collects either \$500/MWh or the zonal LMP, whichever is higher.

[29] Percentages are an average of percentages for individual programs, not a percentage of totals from all programs.

Market operators ERCOT, CAISO, and more recently PJM have instituted programs allowing load resources to bid and receive payment for the provision of ancillary services [32, 35]. Load receives the same control signal given to generators for spinning reserve and regulation response. A licensed provider must demonstrate both ability to respond and the level of response before the market operator will recognize bids. These programs have been developed and implemented quickly considering that the Federal Energy Regulatory Commission (FERC) and the North American Electric Reliability Corporation (NERC) regulations only began allowing for ancillary services on the demand side beginning in 2002 [40]. Including loads as a resource became possible when national standards moved away from proscriptive standards of how ancillary services should be provided and toward performance-based standards. Regional reliability bodies and market operators can still decide whether to allow demand-side provision of ancillary services.

So far, ERCOT appears to lead the market operators in providing technical and market tools for the private sector to use in integrating load into ancillary provision. In its Load Acting as a Resource program, ERCOT will employ load for any ancillary service as long as it is enabled with the stipulated communications and control devices [32].

Chapter 4 Magnitude of Electric Energy Savings

Comparing the magnitude of possible savings from efficiency with that from demand response is important for guiding public and private investments. The comparison is nuanced because energy savings are most important in evaluating efficiency investments while peak load reduction is most important in evaluating demand response. There are cases where incentives toward load flattening would be at odds with those toward energy efficiency; for example if cheap off-peak power were used to charge batteries which were then discharged on-peak there would be an efficiency loss via the battery. In other cases, the incentives are in line; for example very high peak prices would affect air conditioning load more than other loads and may incent the investment in a more efficient model. Potential savings from both types of DSM investments will be informed by exploring retrospective and prospective estimates.

4.1 Energy Efficiency Savings

An energy efficiency savings projection relies on the combination of an economic model and a policy scenario. A 1999 study that analyzed environmental energy policies over the entire United States projected electric savings of 5% in a moderate and 11% in an advanced policy scenario[30] [42]. A set of nine prospective efficiency savings estimates from seven studies is featured in Figure 4.1 [43]. The national study and five state or regional studies show variability stemming from policy assumptions, locational differences, and fundamental uncertainty. The "economic" savings are those that can be achieved at less than the retail price of electricity; the "achievable" savings are lower than the "economic" savings because some upgrades that would make sense economically could not be implemented for practical reasons.

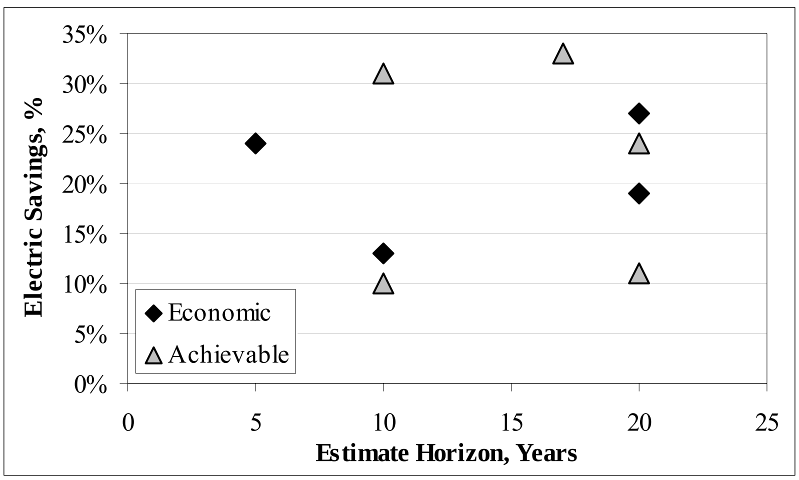


Figure 4.1. Economically feasible and practically achievable electric savings[31] [43].

These nine studies project a range of 10-33% in potential energy efficiency gains from aggressive policy changes. Policy strategies included in these studies reflect efforts similar to traditional demand side management tools and have time horizons from 5 and 20 years.

[30] These numbers represent yearly savings after a 10 year time horizon based on the Energy Information Administration (EIA) base case projection.

[31] Source reports additional numbers that are not reported here. Those numbers are higher possible gains from technically feasible but economically infeasible options; only "economic" or "achievable" results are examined here.

4.2 Elasticity of Demand

Many analyses and experiments have been undertaken in order to examine price responsiveness as well as the responsiveness to shifting demand to a lower cost hour[32] [44]. Some experiments are more relevant to demand response because they examine responsiveness to day-ahead hourly prices or with enabling technology [45][33]. Results are highly variable, partly because responsiveness behavior is complex and highly dependent on the details of the experiment including how prices are communicated. For example, if customers are recruited into a program by being assured that they would not have to pay a higher bill than if they had not participated in the experiment, their incentives are eroded. Similarly, if they know the program will last for only a year or two, they have little incentive to replace appliances or make a capital expenditure that would pay off under a long-term program.

Price responsiveness is much greater when customers have an incentive to react by purchasing more efficient appliances and equipment; in the short run end users can reduce usage only by forgoing consumption. A 1984[34] review of 34 short run and long run estimates found median elasticities of -0.20 and -0.90 respectively, implying that a 10% price increase would reduce consumption by 2% in the short-run and 9% in the long-run [44]. Over the long run these same customers can make additional choices about buying efficient appliances and equipment. Figure 4.2 shows the difference between short-run and long-run responsiveness.

---

[32] Price elasticity of substitution is a measure of load shifting in this context, generally measured between on peak and off peak hours. There is no standard definition of peak hours.

[33] Appendix C of this source contains a review of elasticity studies and their relationship to demand response. Elasticity numbers reported here are obtained from this source.

[34] The short run numbers were recently updated in another review of 36 estimates with a median of -0.28 [44].

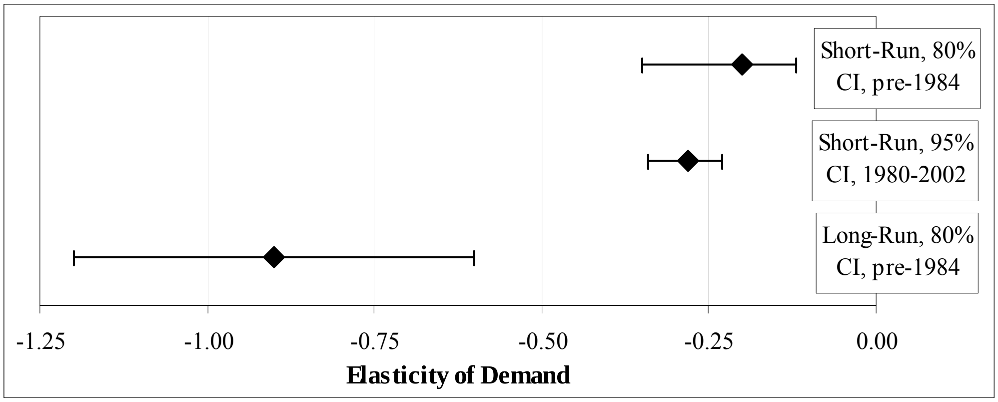


Figure 4.2. Short-run and long run residential elasticity median and confidence intervals (CI) [44].

A recent Department of Energy review published price elasticities of substitution under TOU, critical peak pricing (CPP), and day-ahead RTP situations [45]. Figure 4.3 shows averages and ranges reported from four of these studies in residential and commercial and industrial (C&I) sectors. The range of elasticities of substitution was 0.02 to 0.27.

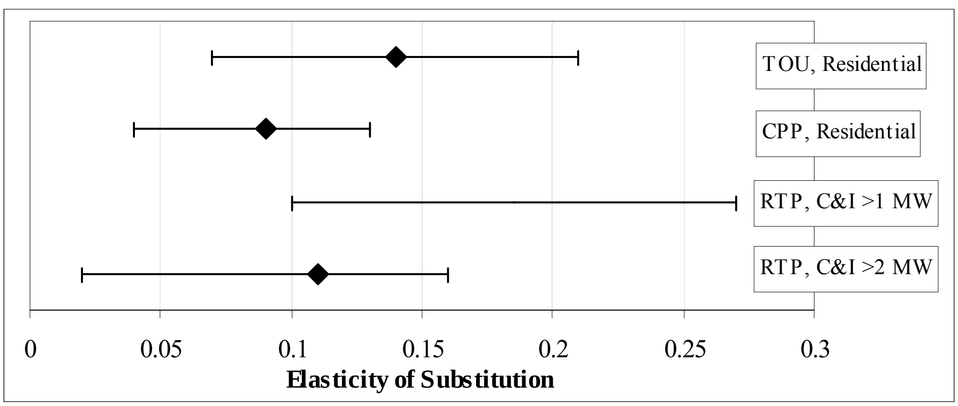


Figure 4.3. Elasticity of substitution average and range [45].

In the future, short-run price elasticity and elasticity of substitution will depend on the sophistication of enabling technologies. Modern electronics allow customers to respond to each price change without further thought or effort by having an "energy manager" run electric hot water heaters, dishwashers, pool pumps, and air conditioners during less expensive hours.

4.3 Demand Response Savings

Projecting the savings in a switch from an average-price system to a real time price (RTP) system is complicated by the uncertainty in how customers will respond. Borenstein has projected that if all customers faced the RTP, equilibrium[35] customer dollar savings would range from 2.0% to 13.7% depending on the responsiveness of demand [46][36]. Table 4.1 shows the theoretical projected savings when different fractions of load face the RTP and the demand elasticity is -0.1. Coincident peak load reductions are large, implying that RTPs would indeed be an effective means of addressing peak demand problems.

Table 4.1. Equilibrium savings in switching from average price to RTP, elasticity -0.1 [46].

Participating Load	Customer Bills, $	Energy Consumption, MWh	Peak Power, MW
33.3%	3.51%	-0.53%	14.0%
66.7%	5.25%	-0.92%	20.3%
99.9%	6.52%	-1.23%	24.5%

Overall energy consumption actually increases under this model because customers can increase usage when prices are low. An increase in energy consumption or profile-dependent pollution[37] under RTP is a real concern. One effect that this model does not address is that responsive customers who have greater control over *when* they use electricity would also have greater control over *how much* electricity they use. For example the Carrier ComfortChoice thermostats that have been used to demonstrate spinning reserve from load reductions also allow customers to specify timed usage [40]. A homeowner can leave the air conditioner off all day while she is at work and have it turn on in time to return to a cool house; she can also control the device over the web if she forgets to turn it off before a vacation. A converse effect is that if customers get a

---

[35] This simulation used California market loads and a mix of three generator types.

[36] The elasticity reflecting these estimates ranges from -0.025 to -0.5. It also reflects scenarios in which demand is more, less, or equally responsive during peak times.

[37] For example if peak generation is natural gas and baseload is coal, a flat load profile would increase emissions of sulfur, particulates, mercury, and other pollutants which are much higher from coal generation than from gas generation [47].

lower average rate for power they may end up wanting to use more of it; for example by using more electricity to pre-cool a space at a time when electric rates are lower.

One question to ask is whether most of the savings from RTPs could be gained from applying the much simpler time of use (TOU) rates. Borenstein has projected that when switching from flat-rate tariffs, total economic surplus[38] increases with TOU rates are only 8% to 29% of the surplus increases with RTPs as shown in Figure 4.4 [48]. The surplus increase is expressed as a percentage of customer baseline expenses. The three TOU rate schedules represent progressively more detailed price granularity. This indicates that if end users really can be responsive in real time, then the savings from the most accurate price signals are substantially greater than those from TOU.

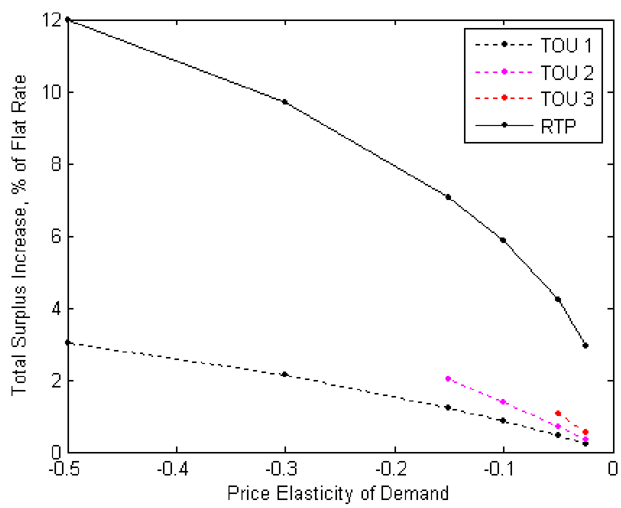


Figure 4.4. Total surplus increase using RTP or TOU pricing, as a percent of flat rate bills[38] [48].

Chapter 5 Barriers to Electricity Market Efficiency

A frustration to policy makers is the continued inaction to reap the savings when an investment in energy efficiency would have a high return. Some failures to invest in efficiency appear

---

[38] The sum of consumer and producer surplus is the total surplus. Under all TOU and RTP pricing structures examined, the total surplus increased compared to the total surplus under flat rate pricing. To scale the total surplus increase numbers, they are expressed as a percentage of the entire consumer electric bills under the flat rate scenario.

irrational from the engineering economic analysis but make sense when hidden costs are included. Other parts are viewed as market failures. Either way, advocates cite barriers to realizing efficiency investments as reasons to enact correcting policies.

Most of the recognized barriers in adopting energy efficiency technology will also inhibit the adoption of demand response technology and strategies; some of the same impediments have already been noted [49].

Few customers can, or have the time to, calculate the return to energy efficiency investments. A more subtle barrier to implementing efficiency programs might be a limited range of features in the efficient models [50]. The EnergyStar program informs consumers about which appliances are efficient with an accessible labeling system at very low cost to the manufacturer or federal government, although the resulting benefits of EnergyStar are difficult to quantify [16].

High-level macroeconomic models that attempt to evaluate economically efficient outcomes are not detailed enough to capture hidden costs at the technology level where they occur. Accounting for the engineering economics of current physical capital and investment costs is becoming a more important part of policy modeling. A proxy for hidden costs is included in the National Energy Modeling System by introducing technology adoption rates and hurdle costs. Models that incorporate these hidden costs explicitly tend to have outcomes with lower energy efficiency [51].

Lack of consumer knowledge about energy efficiency and related costs can result in a less energy-efficient choice. End users may not be able to afford the more efficient appliance or might be financing the purchase with a credit card. Many efficiency investments that are attractive at social rates of return of 2-5% are unattractive at credit card interest rates of 18% or more. Some other situations lead to suboptimal efficiency investments. When different budgets are used for technology investments and for energy costs, the incentive is to decrease up-front costs even at the expense of long-term gains. This situation is acute in a landlord-tenant situation where a landlord buys the least expensive, inefficient air conditioning or water heating equipment but the tenant will have to pay the electric bills [52]. A similar situation can occur even within one firm with a putative common bottom line; the purchasing department might try to minimize the cost of procuring lighting fixtures without considering the long-term electric costs that will be paid by facilities management. Still another situation arises when firms have capital budgets with hard limits; such firms may refuse to buy efficient products regardless of payback. At any rate, once technology is installed, the energy efficiency decision is unlikely to be undone until the end of equipment lifetime; the only changes that can be made until the equipment is replaced are laborious behavioral and usage changes.

Chapter 6 Outlook on Demand-Side Activities

6.1 National Standards

In general, choices of demand response technology, communication, and contractual structures need not be decided by FERC or NERC. The role of regulators and standards bodies is to open markets to competition and participation for all generators and loads. Stipulating that large customers must face RTPs is a prerequisite to making demand response possible without subsidy [53]. Although FERC standards make it possible for demand to have equal opportunity for energy and ancillary market participation, policy movement toward more responsive demand requires legislative action at the state level and is not under FERC jurisdiction.

Already, some states have mandated or allowed utilities to offer RTPs to their largest customers; other utilities create some incentive for load flattening using TOU rates or a customer demand charge [29, 54, 55]. The typical demand charge for large customers records and charges based on the peak kW usage in a month. The problem with this type of demand charge is that it is the same price applies whether that customer's peak demand occurred during the peak time of day, incurring great cost on the system, or whether that customer's peak demand occurred away from peak hours, incurring little cost on the system.

6.2 Market Operator Responsibility

Market operators are rapidly updating the structures of their demand response programs and especially their capacity markets, with a trend toward forward locational capacity markets [13].

The market operators appear to be looking toward a future in which they hope to be able to integrate responsive demand into the markets, and are attempting to structure the markets in a way that will allow for this type of integration, but each has a long way to go. As an example of the support that market operators are willing to provide for possible load-side integration, PJM has structured its energy market to allow LSEs the opportunity to bid in a real demand curve rather than a simple quantity at any price [56]. Creating the structure that would allow for wide-scale energy market participation from demand resources is decidedly forward-looking since in their highest activity year reported in 2006, the total participation in energy market demand response was less than 250 GWh, or about 0.03% of the total energy market [57, 58].

Ultimately, market operators will have to find structures in which load following, wind balancing, and all other ancillary services can be provided by load resources, either directly to the market in a few cases of very large customers, or more likely through a load aggregator such as the new company EnerNOC [59]. This means at the very least that market operators need to be able to broadcast price signals and allow load resources to prove that they can meet the same requirements for reliability and response to control signals.

In more ambitious scenarios, the market operators would have to develop support for load-side resources by changing the market structures fundamentally. One of the biggest challenge will be to develop forward capacity markets that have true demand curves rather than the artificial demand curves that have been used to date [13]. In an improved capacity market, customers would truly be a party to the decision of how much capacity they want built and are willing to pay for. Obviously, as in all cases, the market operator alone cannot accomplish the integration of customers into the markets because the market operators answer to the FERC and the local retailers interacting with customers answer to state regulators.

6.3 State Legislation and Public Utility Commissions

State efforts toward retail competition could have at least a small impact on demand response. A form of time of use pricing happens in deregulated retail markets when a broker buys electricity in the wholesale market for customers. The price that the broker can offer depends on the time profile of customer usage. The broker can show a customer how much the total electricity bill will decline by shifting some demand to off-peak hours. Similarly, the broker can contact customers to tell them when the current wholesale price is very high or very low.

These retail competition efforts are not likely to go very far without state-mandated hourly metering however, because no competitive retailer would be able to assess or prove what the customer was actually using without making the investment in such a meter on a customer-by-customer basis. For this reason the public utility commission (PUC) approval and even the mandate of hourly meters is essential for demand response for all but the largest customers, even where competitive retailers are allowed to offer arbitrary rate structures.

In 2006, the penetration of advanced metering was at 5.9% according to a FERC study, although an unknown but probably large fraction of utilities reporting advanced metering would not be able to implement demand response programs other than RTP or TOU rates without upgrades[39] [60]. Growth in the use and planned use of advanced metering infrastructure (AMI) has more than quadrupled in between 2005 and 2008, driven by state legislation and PUC approval of these expenditures [61].

Small customers may not offer enough system benefit to warrant the expense of time of use or real-time metering as I discuss in Part IV. Using the observed variability of wholesale prices, the expense of a smart meter, the consumption level of a consumer, and the likely response to higher prices, it is possible to estimate whether installing a smart meter will save money for the customer. Aggregators have already organized customers into large loads to realize savings [62]. Eventually aggregators will organize even residential customers if the state regulations allow it and there is profit in it. Requiring large loads to face RTPs does not mean that they cannot get a flat rate contract; a broker would be willing to offer any contract that the customer wants, at a suitable price. Similar implied hedges have already been observed in RTP tariffs [29].

---

[39] For example, a system installed 10 years ago may have been able to conduct automated meter reading but have no ability for two-way communication.

6.4 Distribution Companies

Traditional distribution companies must propose rate structures and have them approved by the PUC. The PUC in turn must act on behalf of the state legislature and in an effort to fulfill its duty to the public. This means that a distribution company does not have the incentive or even the opportunity to move toward RTP or demand response without a state legislative mandate or the commitment of the PUC. Even so, given the trend in states' regulations toward AMI, peak reductions, and demand response, it is likely that many distribution companies are preparing for the eventuality of a mandate [60, 61]. This urgency is compounded by the threat presented by retail competition at the retail level and the possibility that competitors could offer innovative retail rates even if the traditional utilities do not [15].

A big challenge that traditional utilities will have to face in the years to come is in convincing the PUC to allow them enough latitude on offering choices among rate structures. If traditional utilities are forced to offer RTPs alone or flat rates alone or are forced to offer both with no premium on the flat rates to account for risk and unequal load profiles, then the utility could lose customers to competitive suppliers or the new RTP rates could inspire complaints. In any outcome where traditional utilities are able to compete with alternative suppliers, these traditional utilities must be able to offer their customers choices without cross-subsidizing between flat-rate and RTP rate customers.

6.5 Alternative Retailers and Energy Services Companies

A study of 1379 recent energy services company (ESCO) projects shows that these companies are cost-effectively upgrading the electric efficiency for their clients [27]. When ESCOs have upgraded lighting equipment, they have delivered median energy savings of 47%[40] below original consumption from lighting equipment. When ESCOs have performed services beyond

[40] The 50% confidence interval is 37% to 56%.

lighting, they have delivered median savings of 23%[41] from the entire electric bill. Figure 6.1 shows the percentage of these projects that have made improvements of various types. Traditionally inefficient systems such as lighting and heating ventilating and air conditioning (HVAC) are often addressed, but a significant portion of projects involve "other" services as well. These other services can be backup fuel choices, training, or rate analysis.

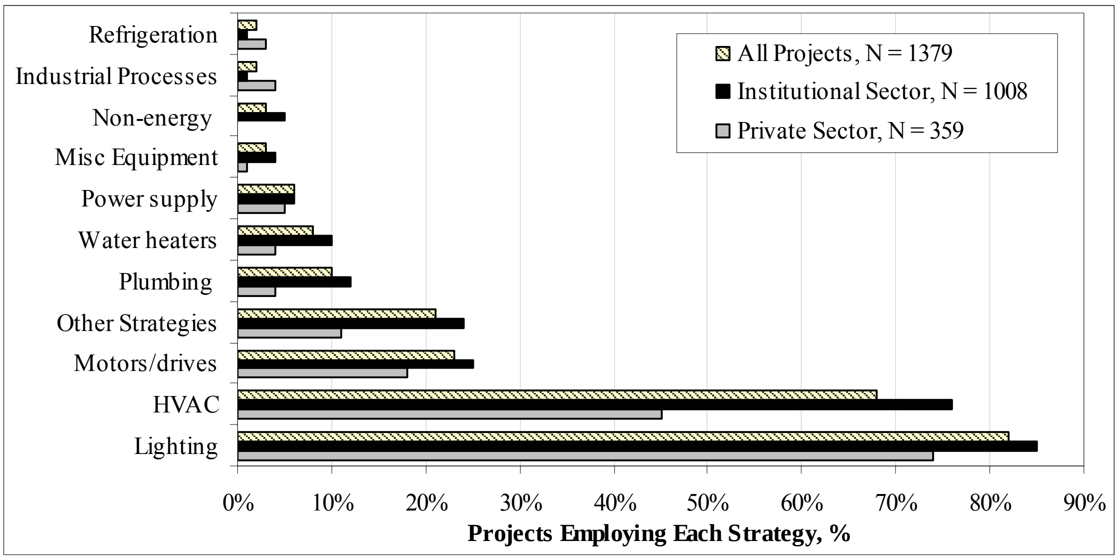


Figure 6.1. Percent of ESCO projects that employed various cost saving strategies [27].

Energy services are a growing market that would find more opportunities for growth if many more customers are subjected to RTPs. Although most of the customer base for ESCOs is in publicly funded facilities, 26% of revenues are from the private sector, especially office space and industrial facilities [27]. Demand response can be added to the portfolio of packaged services that ESCOs offer. Some market operators appear to value ESCOs as intermediaries between the load and the marketplace, but not all market operators offer demand response programs [7, 8]. Market rule changes and additional communication services might be necessary for ESCOs to offer demand response and these needs should be communicated to the market operators.

---

41 The 50% confidence interval is 17% to 32%.

While traditional ESCOs could only add value for customers by reducing consumption and in some cases by reducing monthly peak load, there are new opportunities for ESCOs in demand response. Before customers or ESCOs have any reason to make efforts at load profile management, state regulators must place those customers on interval meters and allow for either traditional utilities or competitive retailers to offer these rates to customers. Further, the introduction of retail competition in some states will allow ESCOs and competitive retail providers the possibility of offering entirely different retail rates and agreements to customers [15].

6.6 Customers

When loads are subjected to RTPs, customers will react to the prices and may invest in automated demand response with the help of a load aggregator. Internalizing the externalities from limiting emissions of pollutants and greenhouse gases will increase the average cost of power; decreasing the relative cost of peaking power with a coal-gas mix and increasing the cost of peaking power in regions where hydro is used for peak power [47]. However, no reactions can occur unless customers know the price in real time.

Initial load response will reflect the easiest and cheapest ways of reducing expense. Figure 6.2 shows the response strategies used by Niagara Mohawk Power Company's large customers under mandatory RTP billing [63]. Among firms that reported shifting load, 47% said they would shift to the next day, 18% to the following day, and only 35% to another time of day. Evidently time of day is more important than actual day in consuming electricity for these customers, possibly because of scheduled shifts and operations. Large customers might find it too expensive and disruptive to flatten their load profiles, even if they are willing to make some changes.

Among firms that reported forgoing load, 65% said it had minimal or no impact on facility operation, 20% reported significant inconvenience or discomfort, 9% had to adjust business operations, and 6% reported not knowing. If many firms can respond to high prices without impacting their missions, then some of the benefits of demand response can be achieved without significant drawback.

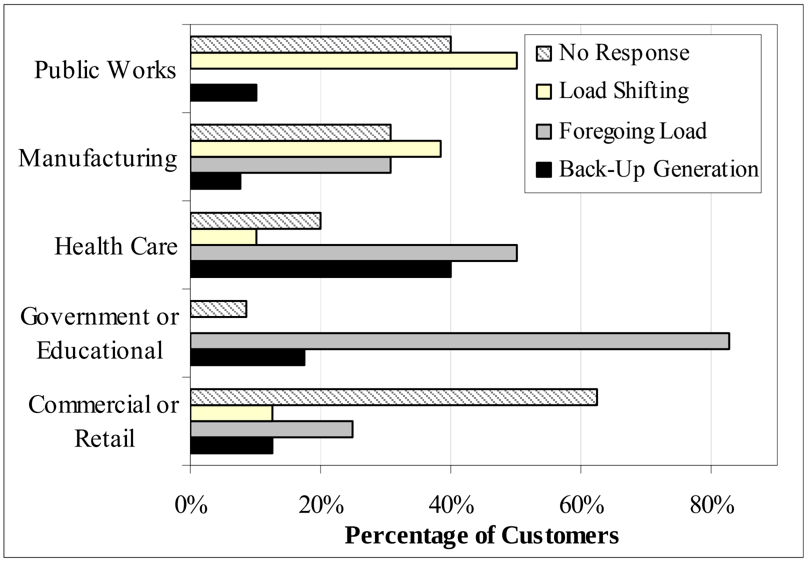


Figure 6.2. Price response strategies employed by various load segments[42] [63].

Although regulators might be hesitant to impose RTPs for fear of end user pushback, in Mohawk only 15% of customers were dissatisfied with a switch to RTPs from TOU even though 54% reported that they did not respond in real time [64]. Some customers, especially governmental and educational facilities, report that they have responded to system emergencies not because prices were high but rather because it was a civic duty [49]. The only customer who would protest the RTP would be one who refused to change her usage and who used more power during the peak hours and so was free-riding on customers who used more power during the off-peak hours.

[42] Percentages do not sum to 100% because response categories are non-exclusive. Sectors had response numbers N = 10, 26, 10, 23, and 8 from top to bottom.

Chapter 7 Challenges and Opportunities

One way to lower average demand is to have consumers understand the implications of their purchases of appliances and other devices that use energy. In many cases, consumers purchase inefficient air conditioners, hot water heaters, and other devices, although paying a bit more for an efficient appliance would save money over time. Government programs attempt to deal with the situation by requiring appliances to have prominent efficiency labels and by setting minimum standards. While much has been accomplished here, much remains to be done in situations where the person paying the electricity bill does not select the appliance or the person making the purchase does not have the money to buy the more efficient appliance.

Another important way to achieve savings is to allow end users to stop buying additional kWh when the RTP exceeds the price they are willing to pay. Just as consumers have learned to respond to the volatile prices of gasoline, fruits and vegetables, and other commodities, so they can learn to respond to electricity prices. The largest difference is that customers purchase electricity every hour of the year and therefore need automated devices to react to changing prices without spending all their time looking up prices and making decisions.

Customer response has been a neglected way of solving electricity industry problems. Historically, providers have focused on supply, assuming that consumers are unwilling or unable to modify their consumption. Contrary to these expectations, customers respond to higher prices that they expect to continue by purchasing more efficient appliances and other efficiency measures. When there are power shortages, customers have shown that they will respond to a peak price signal to reduce demand. Large industrial and commercial customers currently respond to time of use and real time pricing. With the addition of an electronic energy manager, small consumers could respond in real time to price fluctuations or to the more manageable critical peak pricing (CPP) [60]. This customer response has the potential to lower costs by eliminating the most expensive peaking generators, as well providing ancillary serves on the demand side and virtually eliminating blackouts.

While some policy makers and utilities fear that consumers will protest RTPs, experience has found few unhappy customers. Even if they do not change their usage patterns, most customers would find no change in their total bills, since they already pay the average of all high and low price hours. Those customers who *do* choose to react to the high priced hours would lower their own bills, and even lower the bills of unresponsive customers because peak prices would fall. All of the customers on RTP would have to pay an additional charge to cover the cost of the AMI. For some of these customers, the cost of the AMI would outweigh the customer's savings from response, as discussed in detail in Part IV.

A service provider or market operator already has sufficient information to inform individual consumers as to the real-time LMP of electricity. The principal barrier to RTP is the installation of an hourly meter. Even if hourly meters are not cost-effective to change out for very small customers, as the current stock of energy meters are replaced, they should be replaced with real-time meters. Smart meters do not necessarily have to be monitored in real-time, they only need to record hourly consumption data. Additional communication expense is incurred if an LSE is to monitor real-time usage and provide the customer with this information. Some retailers already find it worth their while to install communications with their meters so that they do not have to pay the labor costs of meter readers [65]. Customers must decide for themselves whether to invest in automated devices or ESCO services that would allow them to react to the RTP.

Demand response will become more important as electricity prices rise due to fuel price increases, the need for new generation and distribution, and some of the price increases that have come from unfreezing prices after deregulation. Investment in expensive new capacity can be obviated by demand response and market clearing prices can be lowered. As wind power realizes large scale deployment, the ability of load to shift power use to coordinate with availability will become more valuable and even essential. We will have limited ability to incorporate wind generation into our portfolio mix at large scale without compensating for intermittency on the demand side. Further, compensating in the traditional method with gas causes those units to run less efficiently and emit more air quality pollutants [66].

Customer ability to respond and adapt to these additional costs and system pressures will be greater with more accurate price signals and greater load response. Demand response capability can be part of an overall package of services and greater controls offered by ESCOs. The adaptability that ESCOs have exhibited through deregulation will be invaluable when taking on the additional challenge of making demand response available to consumers.

Part II Meeting Peak Capacity on the Demand Side

In Part II I will present the problem of meeting peak capacity as a two-stage problem within a framework of uncertainty in load growth. The crux of the peak load problem is obvious when describing the traditional situation:

Stage 1 – The amount of peaking generation capacity that will be built and available can be determined with high certainty, but the quantity that will be available must be set on a three-year time horizon because generation suppliers must have enough time to build new capacity. Once the three-year mark has passed, generation capacity is fixed and no more can be built.

Stage 2 – The amount of peak load that will actually be used is determined in real time with high predictive uncertainty and no response to real-time system conditions.

In the traditional model there is no way of ensuring reliability of service except by building a lot of capacity that will rarely be used. This process of building large amounts of excess capacity has been very costly. A more reasonable approach to reliability and meeting peak load would be to involve consumers in the decision-making. If customers were faced directly with a choice between paying the full costs of building new capacity that would run only a few hours per year and the alternative of forgoing a fraction of their consumption for those hours, perhaps most customers would cut their peak use.

Even if customers had to pay some amount of money for devices to control the timing of their consumption, they would make those investments as long as the cost of reducing peak load was less than their savings. In an entirely efficient market, the marginal cost of building one more kW of peak generation would equal the marginal cost of reducing peak load one kW further.

One problem is the magnitude of peak load relative to average load. Another problem is the uncertainty around the predicted peak load. As I will show, the uncertainty bounds around the prediction of peak load are large and getting larger in ISO-NE. This means that it is becoming more and more costly to meet the North American Electric Reliability Council's (NERC's) reliability constraint that the reserve margin be large enough that the loss of load expectation (LOLE) is only one day in ten years. The costs of this reliability are high and getting higher. If the burden of reliability were left to customers rather than mandated system-wide, perhaps some would prefer to pay less and bear the risk of shortages themselves. Perhaps customers would prefer to risk a load curtailment two days in ten years, or five days, or one day per year. I suspect that most customers are getting more reliability than they would want to pay for given the choice.

Indeed, customers, along with aggregators and LSEs acting on their behalf, are the only parties that can do anything to mitigate or reduce the magnitude and uncertainty around peak load. There are many ways that end users and their surrogates might cost-effectively reduce the magnitude or uncertainty of peak load including time-varying rates, direct load control, interruptible rates, or permanent load shifting. I calculate in Part II the value of these reductions in the magnitude and uncertainty in peak load in terms of the avoided costs of building new generation.

Chapter 8 Uncertainty in Peak Load, ISO-NE

In Chapter 8 I use hourly load data for the ISO-NE system (previously the New England Power Pool), over the years 1980-2007 to develop a model representing expected peak load [67]. The model accounts for growth over time in both expected peak load and also in the uncertainty around that expectation.

8.1 Increasing Intensity of Peak Load Problem in ISO-NE

In 2006, the highest peak load year on record in ISO-NE, 15% of all capacity ran 0.90% of the time or less, 25% of all capacity ran 2.92% of the time or less [67]. The number is even more astounding when considering that the 9.9% reserve margin that existed at peak load in 2006 was not included in the calculation[43] [68]. This statistic is bad and getting worse over time, with a comparison of 1980 and 2006 in Figure 8.1 showing how large the fraction of capacity serving just peak load has gotten over the years; similar histograms for all years 1980-2007 are shown in Appendix A.1. In each histogram, peak load hours are highlighted. The width of the colored bands indicates the number of kW that can be considered peak hours; the bands' widths show the quantity of capacity that must exist just to serve demand during peak hours. The darker colored bands indicate that a smaller number of hours are considered. The red band indicates the amount of capacity existing to serve just the top 30 hours (corresponding to a capacity factor of 0.34%); the combination of all red and yellow bands indicates the amount of capacity that exists to serve the top 500 hours (corresponding to a capacity factor of 5.7%).

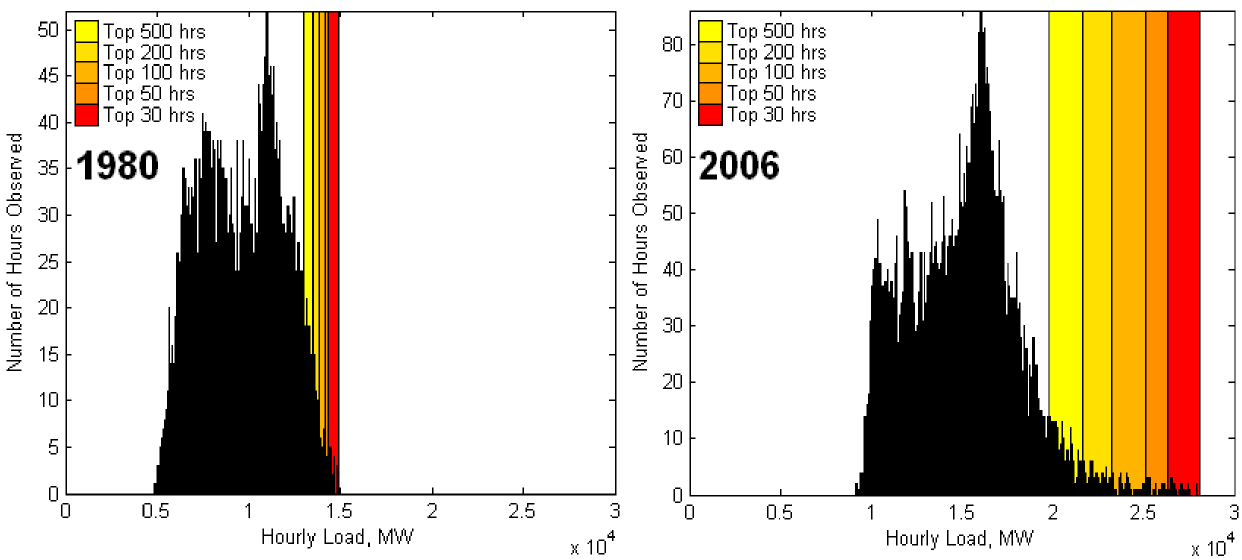


Figure 8.1. Histogram of ISO-NE hourly loads in 1980 and 2006 with peak hours highlighted [67].

The overall assessment is that an increasingly large amount of capacity must be available in ISO-NE in order to run a smaller number of hours per year. This amounts to large capital investments that must exist to ensure reliability but that society puts to very little productive use. If expected to make returns only from the energy market, these peakers with 5.7% capacity factors or less

---

[43] This low reserve margin is made up for by very large reserve margins in the Canadian portions of the Northeast Power Coordinating Council (NPCC) of which ISO-NE is a subregion [68].

would never make a return. Instead they can only be incented if capacity market prices are high enough to cover the entire capital costs of a plant that is not running.

8.2 Model of Peak Load Magnitude and Uncertainty

In this section I develop a model to use later in Chapter 10; this model for peak load accounts for an increase in magnitude and uncertainty over time. I treat peak load as a distribution around the median prediction with the shape of a generalized extreme value distribution $G(L_{peak}|k,\sigma(t),\mu(t))$ in which the location μ and spread σ parameters, analogous to the mean and standard deviation in a normal distribution, are increasing linearly over time t, but in which the shape parameter k, indicating skew, is constant over time.

Treating peak load as an extreme value is both theoretically appropriate because of the properties of the generalized extreme value (GEV) distribution and empirically supported by examining how well the model matches observed data. The GEV is theoretically appropriate because it is the theoretical limit of the maximum value of a sample drawn from any underlying distribution with finite variance [69-71]. The way sample maximum converges to a GEV is analogous to the way that a sample mean converges to a normal distribution by the central limit theorem.

The treatment of the distribution of peak load as a GEV is empirically supported by examining how well the distribution matches observations. Figure 8.2 shows the median prediction as well as the 80% and 95% confidence intervals for fitting a GEV over time to the weekly peak load observations on the left and the annual peak load observations on the right[44]. Figure 8.3 shows the annual and weekly fits along with observed data after the time trend has been removed by translating the raw data into z-values according to Equation (1).

$$(1) \qquad z_i = \frac{L_i - \mu(t_i)}{\sigma(t_i)}$$

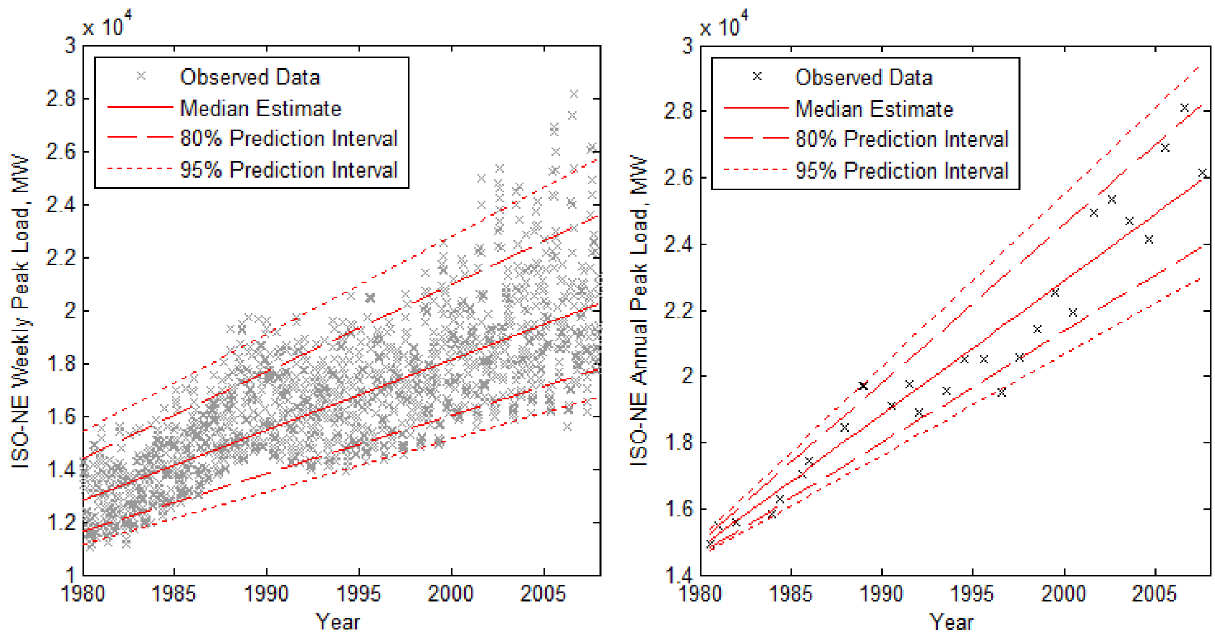


Figure 8.2. Peak load models over time, weekly peak loads (left) and annual peak loads (right)..

---

[44] I fit the five parameters (one representing *k*, and two each representing σ and μ as linear over time) using a maximum likelihood approach. The final distributions have negative values for *k* indicating that they are Type III or Weibull type distributions, indicating that the underlying distribution of loads has a finite tail. This makes sense because peak load is bounded on the high end by the system capacity limit.

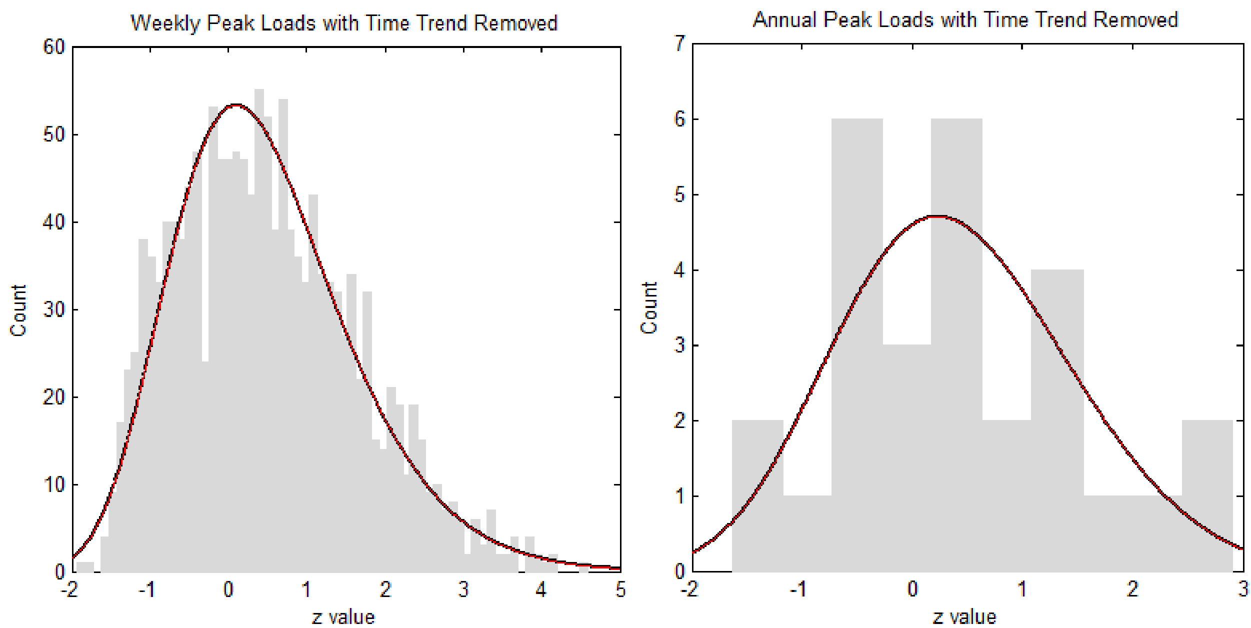


Figure 8.3. Weekly (left) and annual (right) peak load models shown with observed data with time trend removed.

Chapter 9 Value of Peaking Power

In this chapter I present an estimate of the annualized marginal cost of building new peaking capacity and compare this number with current costs of peak load reductions.

9.1 Engineering Economic Value of Capacity

The cost of new generation capacity has increased dramatically in recent years, with natural gas capacity cost having increased by 86% between 2000 and 2007[45] [72]. These dramatic increases have made peak load reductions ever more important; eliminating or delaying the need to build new generation capacity is worth more and more money.

I use the capital cost of a simple cycle gas turbine as the basis of the cost for peaking capacity. I estimate with the recent increases of capacity cost that the price of a simple cycle turbine is

[45] I calculated the 86% number from the two publicly quoted numbers on natural gas: that the capital cost of natural gas has increased 3% between 2007 to 2008, and increased 92% between 2000 and 2008 [72].

$728/kW overnight[46] or $81/kW·y annually[47] [72-74]. The amount of capacity needed to reliably serve the system is greater than the amount of end-use load delivered because of system losses and the necessary reserve margin. I use the same 8% transmission losses that ISO-NE uses in its forecasting processes [75]. When relevant, I assume a required reserve margin[48] of 15%, which was the requirement for the ISO-NE 2008/2009 capacity market auction [76].

Based on these values of T&D losses and required reserve margin, I estimate that $81/kW·y in peak capacity cost translates into a value of $89/kW·y for peak load reductions if T&D losses are considered but the margin for reliability is not. If both reliability and T&D losses are considered, then the cost is $94/kW. I highlight the distinctions among the three numbers to emphasize that a kW of reduction in peak load is worth significantly more than a kW of additional new capacity.

9.2 Current Costs of Load Reductions

The EIA-861 database contains historic data on utility demand-side management programs [77]. Several hundred utilities reported costs related to load-management and energy efficiency as well as total coincident peak load saved for residential, commercial, and industrial customers in 2006. Summary numbers are reported[49] in Table 9.1 through Table 9.4. These tables display average and peak load reductions from load management and energy efficiency programs. I show

---

46 I use Integrated Environmental Control Model (IECM) estimates from year 2000 for a natural gas combined cycle (NGCC) plant and then inflate the cost by 86% according to the Power Capital Cost Index (PCCI) to year 2007 values [72, 73]. I then estimate the cost of a simple cycle turbine by applying the ratio of costs between simple and combined cycle plants costs used in the National Energy Modeling System (NEMS) [74].

47 I inflate the overnight cost over a construction time of three years and then annualize the cost of capital over a 30 year plant life with an 8% cost of capital.

48 Reserve margin is expressed as a percent of peak load, distinct from the alternative measure of capacity margin, which is expressed as a percent of installed capacity.

49 I interpret peak reductions as annualized $/kW·y numbers, although the original data were reported as total dollars spent in load management programs and total "annual" (as opposed to "incremental") "actual" (as opposed to "potential") peak reductions. I similarly treat the energy efficiency numbers as annualized values. This is because although reductions are reported in terms of "incremental" new savings this year and "annual" savings from legacy investments, costs are reported as one number for energy efficiency and one number from load management without differentiating between operating and capital costs. I similarly interpret the energy efficiency numbers as annualized values, making $/MWh the appropriate reporting unit.

median of utilities' reported values as well as the range of the middle 50%[50]. Numbers are updated from 2006 to 2007 for inflation [78].

One important thing to keep in mind while interpreting these numbers is the concept of low-hanging fruit. If a utility's customers have been very inefficient with their electric use and the state has never before implemented a DSM program, then the costs of saving the first few MWh will be low since some very easy and effective changes like subsidizing compact fluorescent light bulbs will be made first. The low range of costs reported here are probably of this sort. After a DSM program has existed for many years and constitutes a large percentage of what would have been used, the marginal cost of more savings will continually increase as savings require more complicated customer interactions and more expensive equipment. Costs of load management will similarly increase with a greater scale of reductions.

[50] I report these values rather than mean and extreme value numbers because I believe the database has a large number of respondents that misreport their numbers resulting in poor data quality and interpretability outside the middle 50% of those reporting.

Table 9.1. Utilities' reported costs of coincident peak load reductions[51] from load management [77].

	25th Percentile, 2007$/kW·y	Median, 2007$/kW·y	75th Percentile, 2007$/kW·y	Number of Utilities Reporting
Residential	$8.37	$25.88	$60.04	119
Commercial	$6.43	$22.18	$60.17	82
Industrial	$3.92	$18.86	$58.72	80
Overall	$5.55	$21.34	$59.78	157

Table 9.2. Utilities' reported costs of coincident peak load reductions from energy efficiency [77].

	25th Percentile, 2007$/kW·y	Median, 2007$/kW·y	75th Percentile, 2007$/kW·y	Number of Utilities Reporting
Residential	$20.89	$50.80	$134	122
Commercial	$33.00	$91.36	$189	106
Industrial	$42.62	$75.96	$161	51
Overall	$24.70	$73.34	$160	141

Table 9.3. Utilities' reported costs of electric energy reductions from load management [77].

	25th Percentile, 2007$/MWh	Median, 2007$/MWh	75th Percentile, 2007$/MWh	Number of Utilities Reporting
Residential	$75.97	$1,348	$5,228	47
Commercial	$102	$1,028	$5,857	51
Industrial	$52.27	$479	$2,448	29
Overall	$74.87	$501	$2,038	79

Table 9.4. Utilities' reported costs of average load reductions from energy efficiency [77].

	25th Percentile, 2007$/MWh	Median, 2007$/MWh	75th Percentile, 2007$/MWh	Number of Utilities Reporting
Residential	$12.29	$29.14	$84.29	182
Commercial	$12.05	$26.18	$74.52	140
Industrial	$11.34	$24.26	$48.87	79
Overall	$12.51	$29.40	$94.22	187

51

As expected, load management programs are more cost-effective for reducing peak load reductions than energy efficiency programs; similarly, energy efficiency programs are more cost-effective for reducing energy consumption than load management programs. The tables show that currently coincident peak load reductions are costing \$21.34/kW·y[52] in the median or \$5.55-\$59.78/kW·y in the middle 50% of programs; energy efficiency is costing \$29.40/MWh in the median or \$12.51-\$94.22/MWh in the middle 50% of programs. Clearly, achieving peak load reductions and energy efficiency are a much cheaper means of satisfying demand than are supplying more capacity and more electric energy, until scaled up to some percent of load where the marginal costs of achieving more reductions are much higher and the marginal benefits much lower.

Peak load reductions are currently being achieved at \$21/kW·y, or less than one fourth of the \$94/kW·y[53] it costs to build new capacity. Similarly, energy efficiency is being achieved at \$29/MWh, or roughly one third of the \$92/MWh[54] retail price for electricity [78, 79]. These costs indicate that current markets and regulations do not make sufficient use of demand-side means of meeting either peak load or energy efficiency.

Reducing load is cheaper than building new capacity or providing more power right now, but I do not expect this to be the case forever. As successively more investments are made in peak load and average load reductions, the low-hanging fruit will have been picked and the marginal cost of achieving the next kW or MWh reduction will increase. When the marginal cost of peak reductions and efficiency equal the marginal costs of providing more capacity and more energy, then the market will have reached an efficient state. This possible end state will not exist without market structures and state regulations that create appropriate incentives for generators, utilities, and customers.

---

[52] Reported costs include incentive payments to end users for curtailment along with all other costs.

[53] See Section 9.1.

[54] This is the national average retail price for power in 2006, updated for inflation to 2007\$.

Also noteworthy, although not unexpected, is that in most cases peak reductions and energy efficiency are being achieved at a lower cost with larger customers. Peak reductions and energy efficiency are being achieved most inexpensively with industrial customers, followed by commercial and finally residential customers. With large industrial customers, the administrators of a DSM program could examine a large quantity of energy use all under one roof, rather than incurring the costs of interacting with many small residential customers in order to have affected the same total load. I examine the effects of customer size in scaling up one type of peak reduction method, real-time pricing, in Part IV.

Chapter 10 Meeting Peak Capacity on the Demand Side

In a perfectly efficient market, or under a perfect integrated resources planner, the marginal cost of curtailing load at times of peak demand should equal the marginal cost of new capacity. In this Chapter I assume that the cost of a peaking generator determines the marginal cost of supply for capacity. This means that peak load curtailment should be ramped up until the marginal cost of achieving more peak reductions is the same as the marginal cost of new supply.

In achieving the first few kW of peak load curtailments, the value per MWh curtailed is very large because there are few hours with very high demand. As curtailment progresses however, more and more days and hours of load will have to be curtailed in order to make further peak load reductions. I use the known target \$/kW·y for efficient peak reductions to determine a perfect integrated resources planner's willingness-to-pay for peak load reductions on a \$/MWh basis.

Similarly, I develop a perfect integrated resources planner's willingness-to-pay customers for reducing the uncertainty in peak load. If customers were willing to accept a relaxed reliability constraint, then the quantity of excess capacity that must be available at peak load in order could be reduced with capacity savings to the system. On the customer's end, a relaxation in the reliability constraint could mean that the customer's utility more carefully manages the size of peak load with targeted curtailments that cause minimal end user inconvenience; alternately, relaxing the reliability constraint could mean that the customer does not enter into a curtailment program, but rather gets hit by rolling blackouts more often.

10.1 Value of Load Shifting and Curtailment

Figure 10.1 shows the fraction of total capacity that was used to deliver a given fraction of total electric energy in ISO-NE over the year 2007 [67]. It shows that 20% of peak capacity was used to deliver only 0.34% of all MWh, and 30% of peak capacity was used to deliver only 1.63% of all MWh, even though 2007 was not the highest peak load year on record. These low capacity factors for peaking generators indicate very high capacity costs associated with producing peaking MWh.

Figure 10.1 must be interpreted with some nuance, as it is created using only information about observed load, not information about the actual installed capacity. The curve does not account for the roughly 16% reserve margin, based on projected capacity[55] and observed load [67, 68]. Recall from Section 9.1 that about 8% of the required reserve margin is necessary to cover T&D losses, which would not affect the shape of Figure 10.1 if it were included. The remaining roughly 8% is required for reliability, I discuss the impact of the reliability requirement in Section 10.2, but note that if that number were included in Figure 10.1, the situation would look much more stark, pushing the entire curve some 8% to the right.

---

[55] Peak capacity is projected by the ISO-NE subregion of NERC using their estimates of capacity that will be available from generation within ISO-NE as well as that available for import via tie-lines from Canada or other outside locations [68].

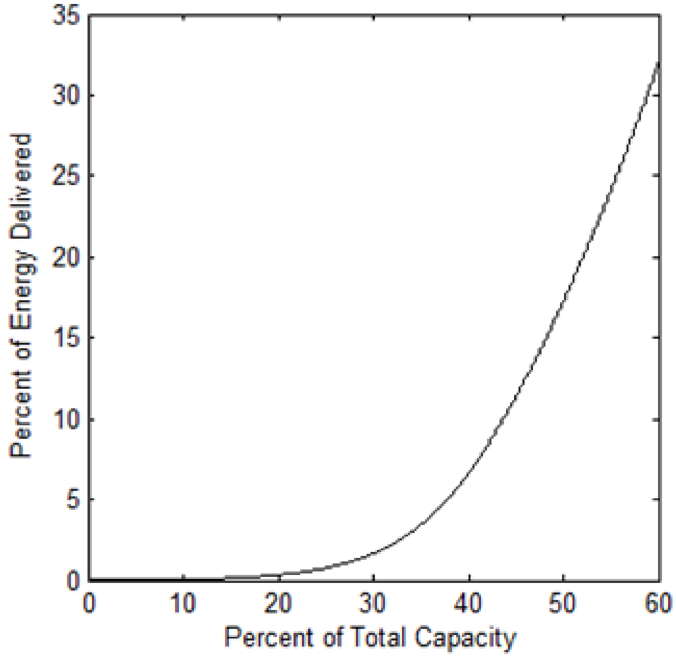


Figure 10.1. Fraction of capacity used to deliver a fraction of electric energy in ISO-NE over 2007 [67].

If peak load reductions are valued at \$93.72/kW·y[56], then the capital cost associated with a generator that runs exactly 1 hour per year is \$93,720/MWh. If the generator runs all 8760 hours of the year, then the per-unit capital cost is \$10.70/MWh[57]. Apportioning capacity costs this way, I can calculate the capital cost incurred by each MWh consumed depending on when it was consumed.

An integrated system planner could use this method to determine a willingness to pay for load curtailment off of peak hours in order to avoid the alternative of building new capacity. Figure 10.2 shows the marginal amount per MWh that an integrated system planner should have been

---

[56] Based on a capacity cost of \$81/kW·y and a 15% reserve margin as in Section 9.1.

[57] More realistically, a baseload generator that will operate nearly all the time would be a capital-intensive coal plant with a lower operating cost than the peaking gas generator. A baseload coal generator would have an overnight cost of about \$2040/kW, or \$228.31/kW·y [72, 73]. The per-unit capital cost of this generator would be \$26.06/MWh at 100% capacity factor or \$32.58/MWh at 80% capacity factor. Therefore, this approach is most valid for peaking hours with low enough capacity factors that capital costs dominate the overall cost of supply and thus simple cycle gas turbines with their high operating costs are still cheaper than other technologies. However, I consider an end state where no peaking gas generators at all are ever used to be an unrealistic end state and therefore do not consider that possibility here.

willing to pay to shift or curtail peak consumption off of peak times. That is, if peak load reductions were already achieved up to a given percent on the x-axis, the downward sloping curve represents the value of curtailing one additional MWh off of peak load. The left-hand side of the figure shows the full range of values, and the right-hand side shows a blown-up version for detail. In order to compare the value of curtailment with the costs of curtailment, utilities' reported costs for load management from Table 9.3.

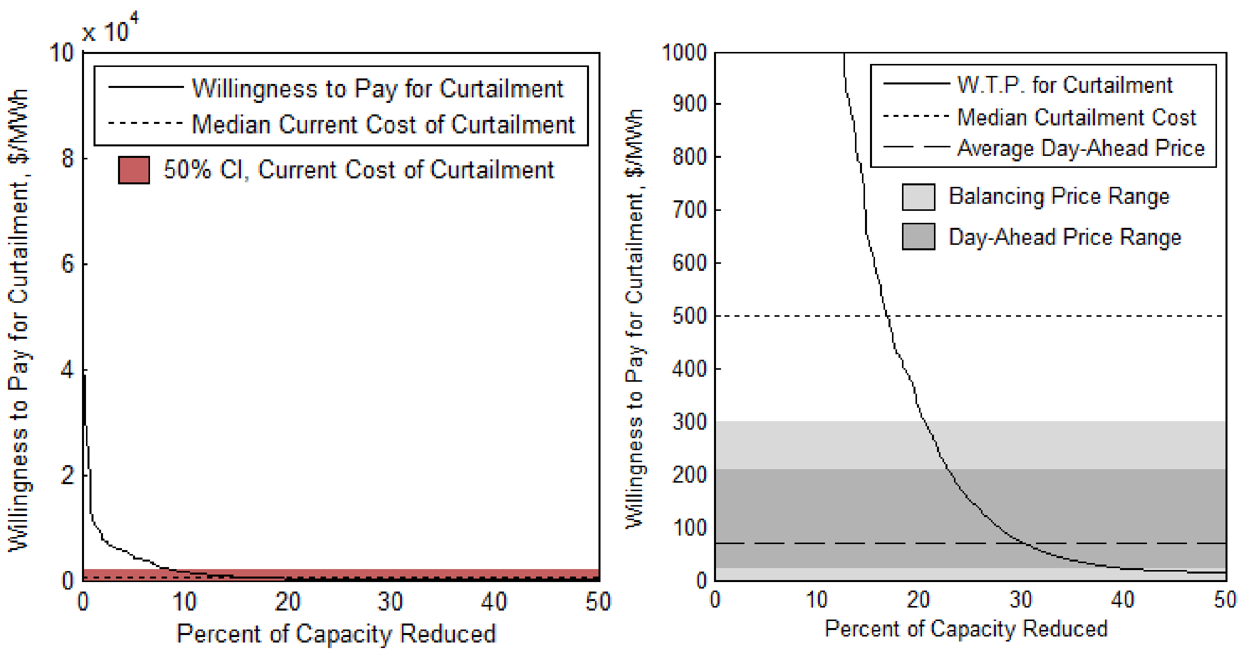


Figure 10.2. Downward sloping curve is willingness to-pay for curtailment per MWh in order to avoid paying for capacity as a function of how much capacity is reduced. Left: the entire range[58] of willingness-to-pay values; range of utilities' reported costs of load curtailment programs are shown from Table 9.3. Right: vertical axis blown up for detail; range of wholesale prices[59] from the day-ahead and balancing markets are shown for scale [67].

From Figure 10.2 it is clear that the value to the system for curtailing the first few MWh off peak hours is enormous, being some two orders of magnitude larger than the average price for power.

[58] The very first MWh reduced is worth $93,720/MWh as discussed previously; the large magnitude of value for the first few MWh curtailed dwarfs the rest of the figure.

[59] The ISO-NE Hub price is the number shown here. The Hub price is the simple (not load-weighted) average of prices in the ISO-NE market each hour.

The capacity costs associated with providing peak power dwarf even the highest peak prices observed in the hourly energy market. When looking at the curtailment costs incurred in current peak load reduction programs, with the 50% confidence interval shown in red on the left and the median shown in both graphs, it is clear that the costs of curtailing peak load are currently much less than the cost of supplying more peak capacity.

Table 10.1 displays the cross-over points from Figure 10.2, indicating what fraction of peak MW and peak MWh could be cost-effectively curtailed if the current $/MWh costs of peak load reductions from Table 9.3 persisted. The table indicates that at the current median cost of curtailment of $501/MWh, 16.9% of peak MW should be curtailed while only having to curtail or shift 0.2% of all electric energy consumed. Eliminating the need to build such a large amount of capacity while inconveniencing customers only a few hours per year would clearly result in large systems benefits.

It is more realistic to think that the cost of peak reductions would become more expensive as more peak load is curtailed and most of the low-hanging fruit is gone. In that case we might expect that the cost of peak curtailments would rise to the current 75th percentile of reported utility program costs of more than $2000/MWh for peak reductions. Even at this higher price, 8.8% of peak MW could be cost-effectively reduced, representing the curtailment of only 0.03% of annual MWh consumption.

Table 10.1. Fraction of peak load to curtail at current peak reduction costs.

	Current Curtailment Costs, $/MWh	Peak Load to Curtail, %MW	Peak Energy to Curtail, %MWh
25th Percentile	$74.87	29.9%	1.60%
Median	$501.35	16.9%	0.20%
75th Percentile	$2,038.99	8.8%	0.03%

Under any realistic assumption about the cost of curtailing progressively more MWh from peak load, including the additional costs associated with imperfect implementation of such a program, significant reductions in peak load could be cost-effectively implemented with small inconvenience to the end user.

10.2 Moving the Risk of Peak Load Uncertainty onto Customers

One of the key problems in meeting peak demand has been the traditional practice of placing the entire burden of uncertainty on the supply side. The result is that systems planners require that additional capacity be built to maintain a low likelihood of capacity shortage. Holding customers and the utilities that represent them harmless in the case of unexpectedly high peak load is puzzling since these demand-side parties are in the best position to do something to reduce that uncertainty and decide what peak demand will be. Further, if customers had a real choice between paying more to overbuild capacity and paying less while enduring a larger chance of curtailment a few days per year, I believe customers would prefer to pay less.

In Section 10.1, I asked how much an integrated system planner ought to be willing to pay customers to reduce peak demand; in Section 10.2 I ask how much a system planner ought to be willing to pay customers to reduce the uncertainty in coincident peak load, or at least bear the consequences of the uncertainty. This does not mean that customers would necessarily have to have smaller deviations in unconstrained peak demand, but would have to bear the risk of the uncertainty themselves by enduring blackouts or curtailment in the case of high demand.

Each NERC region, sub-region, or RTO acting in conjunction with one or more NERC regions establishes its own reliability criterion, nearly all[60] of which are some variation on the criterion that the target frequency for outages due to capacity shortages should be once in 10 years [80-82]. In different regions the loss of load expectation (LOLE) is more specifically described as 1 day in 10 years, 0.1 days per year, or 1 event in 10 years, and is calculated in different ways[61]

[60] Notably different is the Maritimes sub-region of the Northeast Power Coordinating Council, which does not use a probabilistic assessment of peak load as the basis for the reliability margin. The reserve margin of 20% there is based an N-1 criterion, indicating that the reserve margin must be greater than the largest one power source in the system.

[61] In some cases the model predicts daily peak loads with uncertainty for every day and the probability that the reliability margin will be exceeded is summed over all days with a limit of 0.1 days per year. In other cases, the model predicts only average and peak load each year, with the uncertainty around peak load being more analogous to the probability distribution that I use.

depending on how projected load is modeled [81]. For my purposes, I treat the criterion to mean that there should be no more than a 10% chance that peak demand will exceed available capacity in any given year.

I calculate the system cost of building to meet the reliability constraint at the current level, and the cost incurred if the reliability constraint were relaxed to 15% chance of peak shortage, or 20%, or 50%. No matter how the criterion is evaluated, a resource adequacy planner could use this same approach to determine the cost imposed on the system from various levels of reliability and the value to the system if the risk of high peak loads could be shifted to customers.

In Figure 10.3 and Table 10.2 I display the results from setting the reliability constraint at various levels. Using the projected distribution around peak load developed in Section 8.2, I determine the total capacity necessary to serve load in the median projection. I refer to this median projection as the uncurtailed demand. I calculate the quantity of capacity required in order to ensure meeting peak demand with a given certainty; in each case I use ISO-NE assumption that T&D losses are always 8% of generation[62] [76].

The left-hand side of Figure 10.3 shows the reserve margin implied by a certain level of reliability. Relaxing the reliability constraint reduces the required reserve margin, even moving to a negative value with low levels of reliability. Traditionally, a low or negative reserve margin would indicate that customers would experience rolling blackouts at times of peak demand, but this need not be the case in the future. If customers were able to enroll in peak curtailment programs, their non-essential loads could be shifted or shut off during times of peak demand, without suffering from a low quality of service.

The right-hand side of Figure 10.3 shows the cost of meeting peak demand with a given level of certainty, using the calculated reserve margin and \$81/kW·y for supplying peaking capacity. If the reliability constraint were any more stringent, the cost of meeting it would increase steeply; if

[62] Without the inclusion of T&D losses, the reserve margin would be zero at 50% chance of exceeding peak load.

the reliability constraint were more relaxed, the cost of meeting it would drop, but at a slower pace.

The numbers from Figure 10.3 are displayed for sample reliability levels in Table 10.2, with the current reliability level of 10% highlighted for reference. The reserve margin I have calculated here is larger than the 15% that ISO-NE required in their 2008/2009 capacity market auction, indicating that their model predicts a tighter spread around their peak load projections, although their required reserve margin has been higher in other years [76].

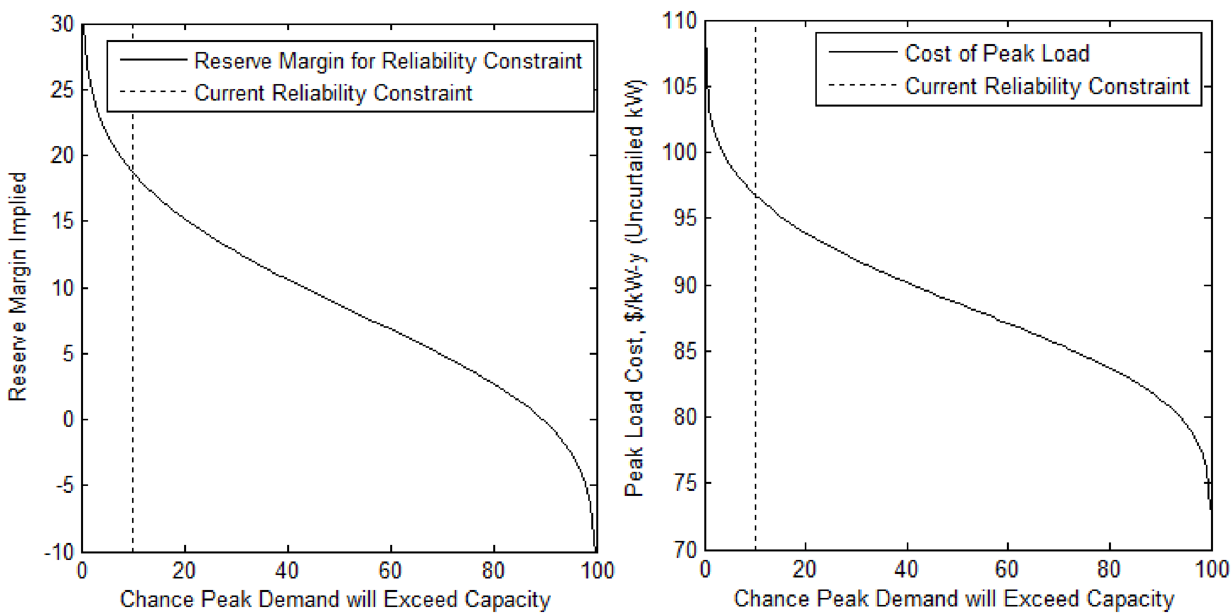


Figure 10.3. Impact of reliability constraint on necessary reserve margin (left) and cost of meeting peak load (right).

Table 10.2. Impact of reliability constraint on necessary reserve margin and cost of meeting peak load..

Chance that Peak Demand Will Exceed Capacity	Reserve Margin Needed to Meet Reliability Constraint	Marginal Capacity Cost per Unit of Unconstrained Peak Demand, $/kW·y
1%	26.8%	$103.32
5%	21.6%	$99.14
10%	18.8%	$96.78
20%	15.2%	$93.90
30%	12.7%	$91.86
40%	10.6%	$90.15
50%	8.7%	$88.59
60%	6.8%	$87.07
70%	4.9%	$85.49
80%	2.7%	$83.71
90%	-0.2%	$81.36

The results displayed in Figure 10.3 and Table 10.2 indicate that dropping the reliability criterion from 10% to 50% would drop the cost of supplying peak demand from $97/kW· y to $89/kW·y. Meaning that right-sizing peak capacity to the best estimate of peak load would reduce the amount cost of supplying capacity by 8.5% below the current cost of overbuilding capacity for reliability.

Part III Short-Run Impacts of Market-Wide Response

Part II of the thesis will investigate the impacts of demand response using a model of PJM's wholesale electricity market[63].

Chapter 11 Introduction and Literature Review

The electricity industry uses much of its generation and transmission capacity only a small fraction of the time. Over the calendar year 2006, 15% of the generation capacity in the Pennsylvania-New Jersey-Maryland (PJM) territory ran less than 1.1% of the time (96 hours or less), and 20% of capacity ran less than 2.3% of the time (202 hours or less) [84][64]. The result is tens of billions of dollars[65] invested in peaking generation that has low capital cost, but high generation cost and life cycle social cost.

The excessive peaking capacity has two causes. The first is technical: there must be enough system capacity to satisfy demand at all times or there will be a blackout. The second is regulatory: most customers pay a constant flat price for power rather than responding to the changing hourly price of the wholesale market. Flat-rate customers have no incentive to shift consumption away from times of peak demand.

Some electricity customers face "time of use" (TOU) pricing that charges them a higher price during on-peak hours, with the fixed on-peak and off-peak rates calculated as the delivered cost averaged over a year. A few customers face "real time pricing" (RTP) where the hourly wholesale generation price determines the retail price. The TOU price gives better information

---

[63] The substance of Part III was published under Spees and Lave [83].

[64] This is based on the entire PJM hourly load profile in 2006 [84]. The system had 17.5% excess available generation capacity. We do not include generation excess at coincident peak load in this calculation because some generation excess is necessary for reliability purposes.

[65] At \$700/kW, a reasonable simple-cycle natural gas generator cost from 2006, this 15% of PJM's generation capacity is worth \$15 billion. At \$2000/kW, a reasonable price for a coal generator, 15% of PJM's capacity is worth \$43 billion.

and incentives than a single fixed tariff, but does not account for the times when wholesale prices spike because of high demand or equipment problems. Some view a TOU rate as a good compromise that frees customers from having to be informed about constantly changing prices and adjusting their consumption accordingly.

Few end users have any opportunity to react to real-time market conditions or to the location-specific costs of generation and transmission. A PJM survey of load-serving entities (LSE) reported that only 5.4% of end user MW are on rates directly or indirectly related to the real-time or day-ahead locational marginal price (LMP) [54, 55][66]. Companies currently offering RTP rates usually have a variety of partial-hedging options as well [29]. Some additional customers are enrolled in direct load control, interruptible contracts, or other subsidy programs that offer curtailment incentives during the top few load hours per year. A Federal Energy Regulatory Committee (FERC) report estimates that 4% of peak MW in Reliability*First* Corporation (RFC) territory[67] could potentially have been curtailed via either RTP rates or non-price response programs, but the maximum response in 2005 was only 0.7% of MW [60]. Actual reductions are usually much smaller than program enrollments, partly because reduction is often voluntary [6].

I view the current flat tariff as both inefficient and inequitable. It is inefficient because it raises system costs and requires much more capital equipment to deliver the same quantity of power. It is inequitable because flat and counter-cyclical customers subsidize customers with high coincident peak demand.

I present a short-run analysis of a change to a more responsive demand-side market. In Chapter 13, I use one year of PJM data to build a supply model that implicitly accounts for dispatch constraints and varying conditions observed over a year. I use this model in three different simulations to estimate the impacts of responsive load. The first in Chapter 14 is an assumed

[66] Estimate is from 3716 MW on locational marginal price (LMP) based rates and 69,064 MW represented in survey responses. Both distribution utilities and competitive suppliers are represented as survey respondents and the rates charged by both types are reported here.

[67] The RFC territory does not match up exactly with PJM territory.

load-shifting scenario that finds the effects of small changes in load profile on overall price. The load-shifting simulation does not consider customer time preference, but does show how quickly savings could be achieved. The final two simulations in Chapter 15 are more realistic; they use hourly demand curves to predict short-run impacts from change toward TOU or RTP from flat-rate pricing.

Borenstein's long-run RTP analysis predicts more than double the peak load savings I predict in my short-run analysis, see Chapter 15 [46]. His conclusion results from using a long-run supply curve in estimating hourly equilibrium conditions. Because Borenstein includes capital costs in his supply curves, he predicts hourly prices up to $90,772/MWh; this implies that 22% of the annual bill is accounted for from the top hour. Those high wholesale prices would only be possible if market rules change dramatically since hourly energy prices are hard-capped at $1000/MWh[68] in all but one United States market and determined based on short-run conditions with a fixed generation portfolio [11]. The high prices that Borenstein uses represent a case in which the RTP reflects the results of both an energy market and a capacity market in which the entire cost of capacity is applied to the consumption of just one hour. Further, I believe that Borenstein's exercise is intended to be primarily illustrative on peak load reductions since his resulting load duration curves are abruptly leveled off on the high end. The short-run analysis presented here reflects current PJM conditions using observed market data.

Holland and Mansur predict less than half the short-term peak load savings that I predict from RTP, see Section Chapter 15 [47, 85]. The modest impact is due to their method of using one constant stacked marginal cost curve to represent supply over the entire year[69]. I use observed

[68] California ISO is the exception with a $400/MWh soft cap on energy and ancillary service bids [10]. Generators may bid above a soft price cap and will be paid as bid; other generators will receive payment only as high as the cap. The neighboring Western Electricity Coordinating Council (WECC) has the same price caps although WECC is not a market operator.

[69] Their stacked marginal cost curve is based on generator heat rates, fuel prices, emissions prices, and other publicly available data for the time frame in question.

market prices to account for transmission and other constraints[70] while they assume constraint-free economic dispatch of system generators to estimate marginal cost. Holland and Mansur attempt to correct for one of these constraints, generator availability, by discounting the capacity of each generator by an expected "outage" factor, but the method cannot capture the observed phenomenon of very high prices at moderate demand levels. Based on the empirical analysis in Appendix C, I find that a constraint-free stacked marginal cost curve underestimates price by $15.88/MWh on average[71], and, more importantly, it also underestimates the *slope* of the real supply curve. The supply curve slope determines the impact that a small change in load has on price, meaning that ignoring transmission and dispatch constraints can lead to qualitatively wrong policy conclusions for RTP, see Appendix C.

Power engineers account for real-time transmission constraints by solving the security-constrained direct-current optimal power flow (DCOPF) problem in example cases. This approach is similar to how PJM sets market prices. Wang, Redondo, and Galiana used a DCOPF-based model to examine demand-side participation in wholesale energy and ancillary services markets [86]. Their results indicate that demand participation erodes generator market power. However, results from test systems with a few buses do not translate directly into implications for the PJM system with roughly 7800 pricing points. Fitting supply curves to daily market data accounts for the effects these constraints have on the resulting energy market price.

Chapter 12 Data

The data that I use for this analysis are system-wide hourly PJM market clearing results. I examine aggregate load and PJM average prices[72] in the day-ahead and real-time (balancing)

---

[70] Examples of other constraints include limits on run times, ramp rates, reserve margins, local reactive power generation, scheduled maintenance etc.

[71] The estimate uses PJM generator bid data over a calendar year from June 2005 through May 2006. The bid data are publicly available after a six-month delay. Full details of this calculation are available in our working paper [83].

[72] The PJM price is a load-weighted average of all system LMPs.

markets over 2006 [84]. Day-ahead demand bids L_{DA} from LSEs are charged at the day-ahead price P_{DA}, the real-time increment or decrement L_{RT}-L_{DA} is charged or credited at the real-time price P_{RT}. . Overall revenue and price are calculated with Equations (2) and (3).

(2) $$R = L_{DA} \cdot P_{DA} + \left(L_{RT} - L_{DA}\right) \cdot P_{RT}$$

(3) $$P_0 = \frac{\sum_{hours} R}{\sum_{hours} L_{RT}}$$

Overall realized price and the real-time demand for each hour are the most accurate data for evaluating demand response. In the implementation of RTP rates, customers should have access to both day-ahead and real-time market prices. I assume that nearly all power continues to be purchased in the day-ahead market; both markets are counted as RTP.

Chapter 13 Market Model

I construct a short-term equilibrium model accounting for producer, consumer, and LSE participation. Results from the full model for RTP and TOU pricing are in Chapter 15. The load-shifting scenario in Chapter 14 uses only the supply-side model developed in this section.

13.1 Short-Term Equilibrium Model

My base case model treats the retail and wholesale markets separately as shown in Figure 13.1. In the retail market, I assume that all consumers currently pay a flat rate P_0 for all their power, making the supply curve appear completely elastic to consumers. In the wholesale market, the market operator treats hourly demand L_0 as completely unresponsive to price. While each hour has wholesale price P_W above or below retail price, the profits and losses are temporarily absorbed by the LSE and sum to zero over the year. This disconnect between wholesale and retail is a good characterization of current conditions since few customers face RTP [44, 54, 55, 60].

Under TOU the retail price takes on a value of p_{on} during on-peak hours and p_{off} during off-peak hours. In PJM off-peak hours are weeknights 11 PM to 7 AM and all day on weekends and the six NERC holidays [87]. On and off-peak prices are set so that LSE profit sums to zero over on-peak hours and off-peak hours separately.

When modeling RTP, I set the retail price equal to the wholesale price, eliminating the disconnect between the wholesale and retail markets (neglecting distribution costs).

13.2 Demand Side

I assume that each hour has a unique demand curve with constant own-price elasticity as shown in Equation (4) where the hourly offset parameters β are determined by base case price and hourly load [46, 85]. The assumed constant elasticity of demand E is assumed to be zero in the base case.

$$P_D(L) = \beta \cdot L^{1/E}$$
$$\beta = \frac{P_0}{L_0^{1/E}} \tag{4}$$

The left side of (4) is replaced with the retail price $P_D(L)$ that applies in the flat (5), TOU (6), or RTP (7) cases.

$$P_D(L) = p_0 \tag{5}$$

$$P_D(L) = P_{TOU} = \begin{cases} p_{on} \\ p_{off} \end{cases} \tag{6}$$

$$P_D(L) = P_S(L) \tag{7}$$

In order to model demand using the most realistic elasticities, I use estimates from the literature. A 1984[73] review of 34 studies found short run and long run price elasticities to be approximately -0.20 and -0.90 respectively, implying that a 10% price increase would reduce consumption by 2% in the short-run and 9% in the long-run [44]. Most of these estimates were made based on a change from one flat rate for power to another, not responses to hourly changing prices, and so the short run number only hints at the appropriate number for my purposes.

More telling is that after 5 years of experience with default RTP for customers larger than 2 MW, Niagara-Mohawk Power Corporation has observed an average demand elasticity of substitution of -0.11 [88, 89]. A Department of Energy study reviewed price elasticities of substitution under TOU, critical peak pricing (CPP), and day-ahead RTP situations [90]. The range of elasticities of substitution was 0.02 to 0.27.

The level of responsiveness that would be observed under RTP is uncertain and could depend on a variety of factors including customer class, weather, and enabling technology. Regardless, there have been enough empirical estimates to place the plausible short-run elasticities of demand between 0 and -0.4 under RTP conditions. I examine this full range. I will not specify exactly how an aggregate elasticity is achieved, for example having all customers on RTP with an elasticity of -0.1 would be approximately the same as having only half of all MW on RTP with an elasticity -0.2.

13.3 Wholesale Supply Side

At one extreme, I might hypothesize that the wholesale supply-side relationship between price and load is the same over an entire year. At the other extreme, I might hypothesize that the relationship is unique to each day. The market clearing price at a specified load level may differ from one day to another because some generating units or transmission lines are not available, fuel prices have changed, or weather is impeding supply. Fitting unique parameters for each day

[73] The short run numbers were recently updated in another review of 36 estimates with a median of -0.28.

would give a better fit than insisting that one set of parameters must fit the entire year. However, the former is not a parsimonious model and says nothing about what parameter values should be used in future days.

The wholesale price of electricity for each hour in a day follows a predictable pattern of being low in the early morning and at night with one or two peaks during the day. I fit the price and load data for each day with a third-degree polynomial. To investigate the similarity of the polynomial parameters across days, I employ dummy variables, taking on values of 0 or 1.

Equation (7) models price as a function of load represented by an intercept, load, load squared, and load cubed. The equation uses dummy variables δ_1 and δ_0 to allow for the possibility that the coefficient of load and the intercept might vary each day. I also examined the possibility that the coefficients of the squared and cubed terms take on unique values each day but determined that the additional dummy variables improved explanatory power very little. I explored a range of specifications and selected (7) as a model with good explanatory power, only half the number of parameters as employing the additional two dummy variables, and as a good fit to the plotted data. For detailed results from trying a range of models, see Appendix A.

$$(8) \qquad P_S(L) = a \cdot L^3 + b \cdot L^2 + \sum_{t=1}^{n} \{\delta_1 \cdot c_t \cdot L + \delta_0 \cdot d_t\}$$

The adjusted R^2 is 0.949, the F-statistic of 223 is highly significant[74], and the estimated parameters *a* and *b* are highly significant[75] all with p-values $\ll 0.001$. This model accurately represents observed data and is therefore an accurate base case, but I must add a note of caution to the reader when interpreting the simulation cases. I cannot presume to know all of the structural changes that would be part of integrating the expectation of responsive load into the

[74] Model significance test has $F(731,8028) = 223$ with p-value $\ll 0.001$.

[75] Studentized t-tests have $t_a(8028) = 10.9$ and $t_b(8028) = 33.0$ with p-values $\ll 0.001$ in each case.

wholesale market[76]. Therefore, the larger the difference between the base case and the simulation cases, the less confidence I have in the results.

13.4 Economic Result Definitions

Changes in consumer surplus ΔCS and producer surplus ΔPS between flat rate and RTP conditions are calculated in Equations (9) and (10) and shown graphically in Figure 13.1. Producer surplus is easier to calculate by integrating over load than over price. Change in consumer surplus in Equation (9) can be calculated in the TOU case by replacing P^* with the retail TOU price p_{off} or p_{on}. Change in producer surplus calculated in Equation (10) is the same formula under a change toward TOU or RTP because the wholesale electric price determines the producer surplus.

$$(9) \qquad \Delta CS = \sum_{hours} \int_{P^*}^{P_0} L(P_D)\partial P = \sum_{hours} \int_{P^*}^{P_0} \left(\frac{P_D}{\beta}\right)^E \partial P = \sum_{hours} \left(\frac{1}{E+1}\right)\left(\frac{P_D}{\beta}\right)^{E+1} \Bigg|_{P^*}^{P_0}$$

$$\Delta PS = \sum_{hours} \int_{P_S(L_0)}^{P^*} L(P_S)\partial P = \sum_{hours} \left\{ P^* L^* - P_0 L_0 - \int_{L_0}^{L^*} P_S(L)\partial L \right\}$$

$$(10) \qquad \Delta PS = \sum_{hours} \left\{ P^* L^* - P_0 L_0 - \int_{L_0}^{L^*} \left(aL^3 + bL^2 + cL + d\right)\partial L \right\}$$

$$\Delta PS = \sum_{hours} \left\{ P^* L^* - P_0 L_0 - \left[\left(\frac{a}{4}L^4 + \frac{b}{3}L^3 + \frac{c}{2}L^2 + dL\right)\right]_{L_0}^{L^*} \right\}$$

With flat-rate or TOU pricing there is deadweight loss in both high-priced hours and low-priced hours. Because the RTP case has no deadweight loss, we calculate the deadweight loss in the flat rate and TOU cases based on the surplus changes in Equation (11). Both deadweight loss and LSE profit Π are shown in Figure 13.1for a sample high-priced hour. In particular, note that

[76] For example, we speculate that the market operator would use some different scheduling rules when dispatching generators if a downward-sloping demand curve were part of the day-ahead dispatch problem.

the increase in total surplus increase after accounting for the customer, the LSE, and the supplier is exactly the same size as the deadweight loss decrease between any one rate and any other rate. For this reason all information about deadweight loss is exactly contained in the total surplus results, which are the numbers I will report in the results.

$$
\begin{aligned}
DW_{flat} &= \Delta\Pi_{flat}^{RTP} + \Delta CS_{flat}^{RTP} + \Delta PS_{flat}^{RTP} = \Delta CS_{flat}^{RTP} + \Delta PS_{flat}^{RTP} \\
DW_{TOU} &= DW_{flat} - \Delta DW_{flat}^{TOU} = DW_{flat} - \left(\Delta CS_{flat}^{TOU} + \Delta PS_{flat}^{TOU}\right)
\end{aligned}
\tag{11}
$$

The model is shown graphically in Figure 13.1for a sample high-priced hour. The base-case retail market is represented by demand curve $P_D(L)$ and completely elastic supply P_0; the wholesale market is represented by supply curve $P_S(L)$ and completely inelastic demand L_0. The base case model has two different resulting prices P_0 and P_W that apply in the retail and wholesale markets respectively, but resulting load has to be the same in both. The arrow shows how load drops under RTP when the integrated market is represented by supply and demand curves $P_D(L)$ and $P_S(L)$.

The shaded areas in Figure 13.1show a change from flat-rate to RTP for a high-price hour. The farthest left plot shows the LSE's hourly deficit under flat-rate pricing due to buying power from the wholesale market at P_W and retailing it to consumers at P_0; under RTP the LSE's hourly deficit is eliminated, increasing its welfare by $(P_W\text{-}P_0)\cdot L_0$. The increase in the LSE's welfare in moving from a flat-rate to RTP comes from three places as shown in the remaining three shaded areas from left to right: 1) hourly consumer surplus drop in moving from flat-rate to RTP due to having to pay the higher price as in (9); 2) hourly producer surplus drop in moving from flat-rate to RTP due to the drop in price and quantity as in (10) ; and 3) hourly deadweight loss as in (11). Note that in a corollary low-price hour, load would increase under RTP but the mathematical definitions would hold. Although the LSE may have a positive or negative profit in any one hour with TOU or flat rate, it has zero profit over the year under any of these pricing scenarios.

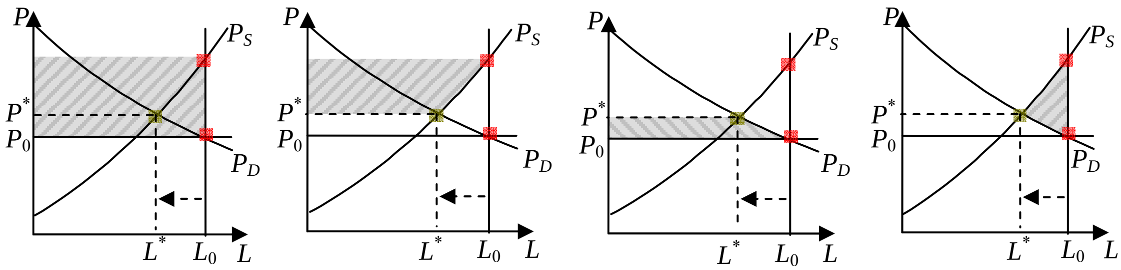


Figure 13.1. From left: retailer loss, producer surplus drop, consumer surplus drop, and deadweight loss.

Chapter 14 Load Shifting

Assume that customers can be induced to shift their demand to be more level over the day. Although the resulting load profiles may not be realistic, I use this simulation to show how much shifting is necessary to flatten load and how quickly savings can be achieved.

14.1 Method

I scale possible consumer savings from demand response by incrementally shifting load to achieve a totally flat daily load profile without changing total consumption. Although this method does not consider real-world preference effects, it does set an upper bound on customer savings. The simulation allows load shifting to any other time of day but does not allow shifting from one day to another.

For a particular day, I simulate shifting an increment of demand from the highest load hour to the lowest load hour. I continue shifting demand increments so that there is one wholesale price for the hours of greatest use and another (lower) wholesale price for the hours of least use. The maximum fraction f that is curtailed off the peak load hours is the same for all days. I stop shifting load when the quantity and wholesale price are the same for the high and low-priced hours. The simulation reaches maximum shifting with 5.3% of all MWh shifted away from peak hours and $f = 0.158$ (or 15.8% of MW) at which point the load profile is flat over each day, but not between days.

Figure 14.1 illustrates the effects of shifting on load and price profiles of one week beginning Monday, June 19, 2006. This week originally exhibited moderately high load and price. Results are shown when 3% of all yearly MWh are shifted and after the maximum shifting of 5.3% of all yearly MWh. This method does not change total daily consumption in MWh, but the extremes of usage and price variation are reduced.

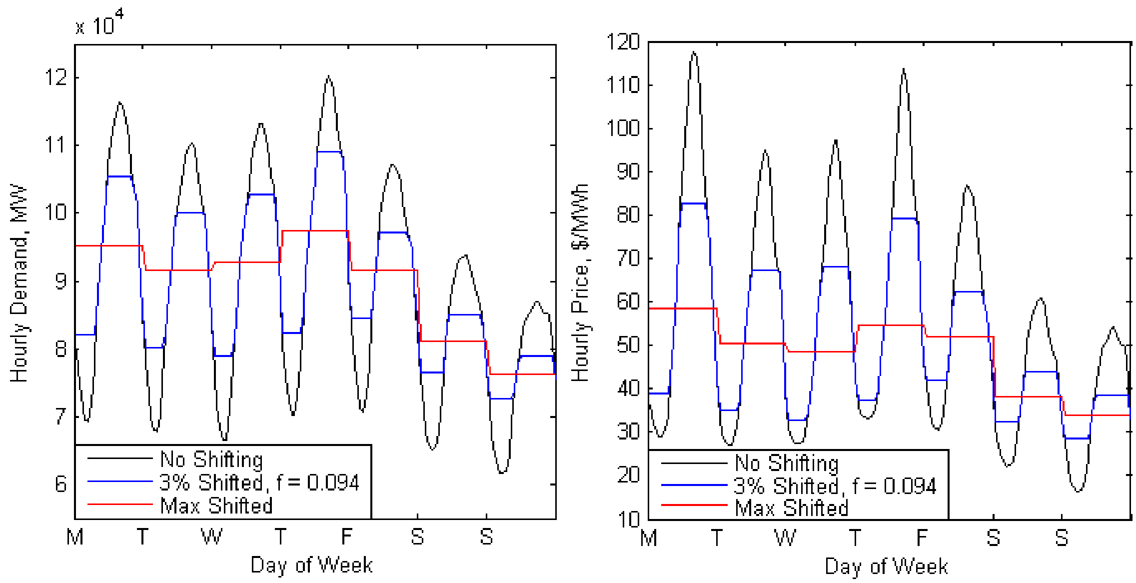


Figure 14.1. Load and price profiles for a July week; base case, 3% shifting (*f* = 0.093), and max shifting.

14.2 Results

I remind the reader that the load shifts are imposed, rather than resulting from consumer preferences and so no conclusions can be drawn about consumers being better or worse off.

Customer expenditure savings from load shifting are shown in Figure 14.2. Savings are also split out by the amounts received by shifters and those received by free riders that do nothing. The left-hand plot in Figure 14.2 displays decreasing marginal savings with more shifting; when the daily load is leveled, there are no further savings. The right-hand plot of Figure 14.2 shows that shifters' percentage savings drop with increased shifting. This is because the price differential over a given day can be large under current conditions but approaches zero in the limit; small marginal savings steadily reduce average calculated savings. Total customer savings increase with the amount of shifting with an ultimate limit of 10.7% of the annual electric bill.

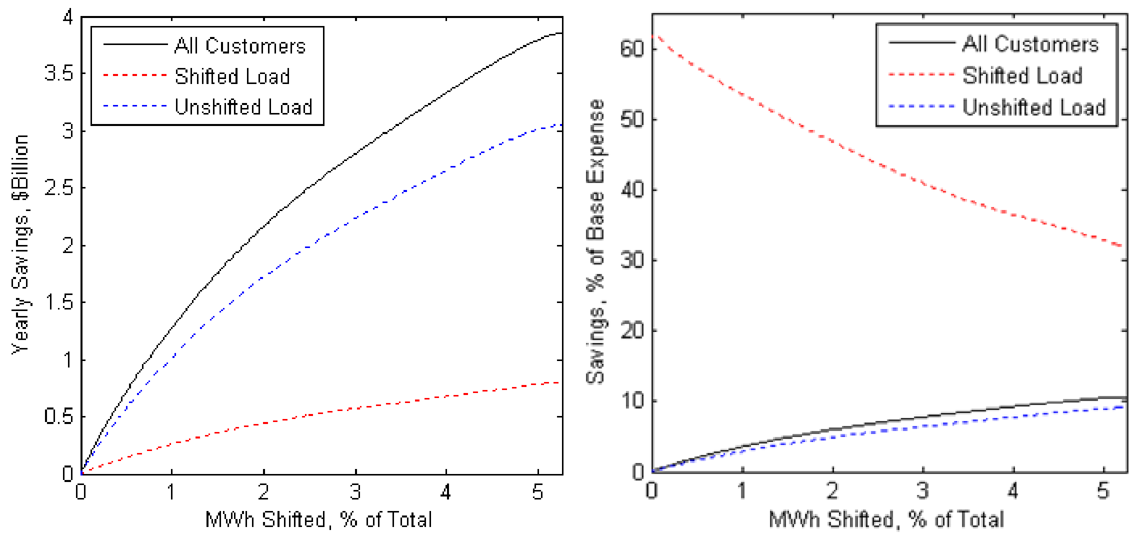


Figure 14.2. Savings to shifters, free riders, and total in dollars (left) and as a percentage of bill (right).

Load shifting reduces peak load dramatically as shown in Table 14.1, obviating the need for costly investment in generation and transmission.

Table 14.1. Peak load and overall cost savings with daily shifting.

Shifted Load, %	Peak Load, GW	Peak Load Saved	Total Expense, $Billion[77]	Average Cost, $/MWh	Customer Bill Savings
0%	145	0.0%	$36.17	$51.96	0.0%
1%	138	4.8%	$34.90	$50.13	3.5%
2%	134	7.3%	$34.03	$48.88	5.9%
3%	131	9.3%	$33.37	$47.94	7.7%
4%	128	11.6%	$32.84	$47.17	9.2%
5%	122	15.8%	$32.38	$46.51	10.5%
5.3%	122	15.8%	$32.32	$46.43	10.7%

[77] This is the total customer bill to all customers, shown in order to scale the magnitude of savings in relation to the total market size.

Table 14.2 shows how quickly customer savings are reached by load shifting. Half of all the possible savings from load shifting are achieved by shifting only 1.69% of all energy. This indicates that a small amount of demand response is all that is needed to get most of the benefits.

Table 14.2. Load shifting necessary to achieve a portion of limiting savings with daily shifting.

% of Savings in Limit	% Load Shifted	Maximum Hourly % Curtailed
25%	0.70%	3.9%
50%	1.69%	6.6%
75%	3.15%	9.6%
90%	4.26%	12.4%
95%	4.66%	14.0%
99%	5.06%	16.5%

Chapter 15 Time of Use and Real Time Pricing

I turn from calculating the savings from assuming that load can be shifted to an analysis of how much consumers would shift load in response to price differentials between high and low demand hours. I use a simulation to determine the magnitude of effects from a change to RTP or TOU.

15.1 Sample Price and Load Profiles

The new price and load under RTP and TOU conditions are calculated as in Chapter 13. Figure 15.1 shows load and wholesale price profiles P_S over a week in the base case, under TOU, and under RTP conditions with elasticity -0.2. Under RTP, the price that consumers face is the same as the one paid to generators in the wholesale market, $P_D = P_S$ as in (6). Under flat or TOU rates, the wholesale price P_S can be higher or lower than the retail prices. For reference the flat and TOU retail rates p_0 and p_{TOU} are shown in dashed lines for the flat-rate and TOU cases respectively (without accounting for the distribution charge). The June week shown originally had moderately high load and wholesale price, so the RTP case shows steep drops in price and load during peak hours.

The left-hand graph in Figure 15.1 shows that RTP reduces peak loads much more than TOU pricing, which is only slightly better than flat rate pricing. The right-hand graph shows wholesale prices reflecting the marginal generation cost as solid lines; retail tariffs are in dashed lines. Under RTP the wholesale and retail prices are the same solid line. Wholesale price peaks are moderated much more under RTP than under TOU pricing. A TOU rate actually exacerbates wholesale price peaks on weekends because end users see the off-peak price all day.

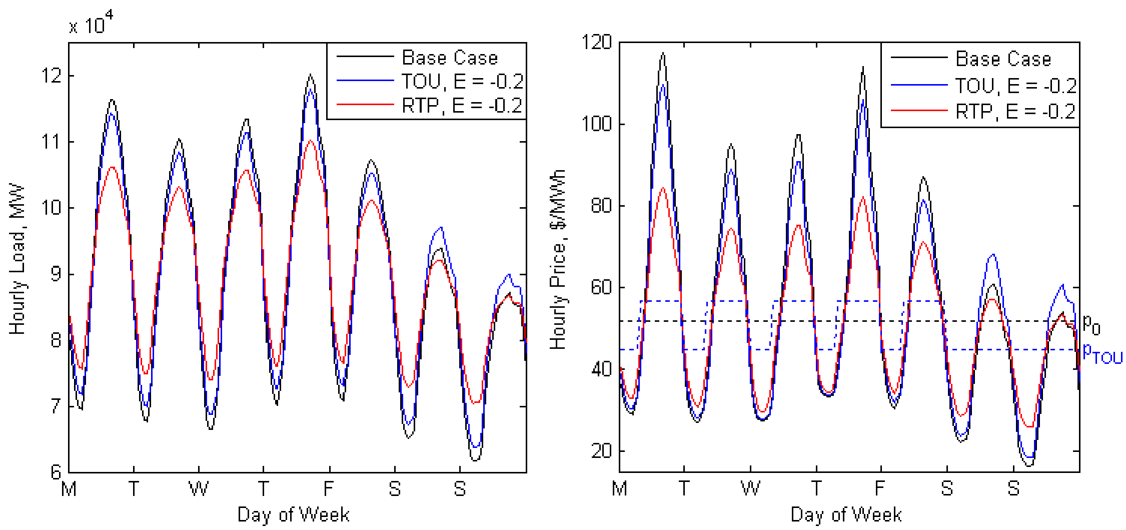


Figure 15.1. Load and price profiles with elasticity -0.2 for a July week with flat-rate, TOU, and RTP.

15.2 Economic Impacts

Market outcomes depend on the assumed demand elasticity[78]. Table 15.1 and Table 15.2 summarize impacts on consumption, expense, average price, and peak load with TOU and RTP rates respectively. The impacts from TOU pricing are a fraction of those from RTP. Impacts from TOU in peak load shaved, consumption increase, and consumer expense saved are never more than 14.4%, 22.3%, and 21.9% respectively of the impacts from changing to RTP at any elasticity.

---

[78] Customers are more responsive when elasticity is more negative; responsiveness increases as one moves to the left in these plots.

Impacts on consumer expense and consumption increase are small under either rate structure change. The most striking result in these tables is that with RTP, peak load reductions are large even with highly (but not completely) inelastic demand. I estimate a 10.4% reduction in peak demand at elasticity $E = -0.1$, a huge reduction at a modest assumed responsiveness. Holland and Mansur's prediction with all customers on RTP at this same elasticity is less than half mine at 3.91%, while Borenstein's estimate is more than twice the size at 24.5% [79,80] [46, 85].

Table 15.1. Load increase, peak shaving, and price savings with TOU pricing.

Elasticity of Demand	Peak Load, GW	Peak Load Saved	Total Energy, TWh	Consumption Increase	Total Expense, $Billion	Consumer Expense Saved	Average Price, $/MWh
0	145	0.0%	696	0.0%	$36.17	0.0%	$51.96
-0.05	144	0.6%	697	0.1%	$36.04	0.4%	$51.72
-0.1	143	1.1%	697	0.2%	$35.95	0.6%	$51.54
-0.15	143	1.5%	698	0.3%	$35.91	0.7%	$51.44
-0.2	142	1.9%	699	0.4%	$35.90	0.8%	$51.38
-0.25	142	2.2%	699	0.4%	$35.90	0.7%	$51.35
-0.3	141	2.4%	700	0.5%	$35.91	0.7%	$51.34
-0.35	141	2.6%	700	0.5%	$35.93	0.7%	$51.34
-0.4	141	2.8%	700	0.6%	$35.95	0.6%	$51.34

79 Holland and Mansur also predict a 5.88% peak load reduction at E = -0.2, where I predict a 15.1% savings. Borenstein also predicts 35.2% peak load reduction at E = -0.3 where I predict a 17.7% savings. The modest impacts predicted by Holland and Mansur are largely dictated by their method of using a stacked bid curve, see Appendix C. Borenstein's large projected peak reduction has to be understood knowing that his supply curve comprised of three generator types results in a load duration curve that is completely chopped off on the high end; he does not argue that this is a realistic resulting load duration curve.

80 Holland and Mansur examine the PJM electricity market as I do. They used observed load data as I do and their own price estimates based on generating units' operating performance from eGrid [85, 91]. Borenstein uses observed hourly load data from California along with sample costs from a mix of three generation technologies [46].

Table 15.2. Load increase, peak shaving, and price savings with RTP.

Elasticity of Demand	Peak Load, GW	Peak Load Saved	Total Energy, TWh	Consumption Increase	Total Expense, $Billion	Consumer Expense Saved	Average Price, $/MWh
0	145	0.0%	696	0.0%	$36.17	0.0%	$51.96
-0.05	137	5.7%	699	0.4%	$35.52	1.8%	$50.82
-0.1	130	10.4%	702	0.8%	$35.11	2.9%	$50.02
-0.15	126	13.3%	705	1.2%	$34.94	3.4%	$49.59
-0.2	123	15.1%	707	1.6%	$34.90	3.5%	$49.35
-0.25	121	16.6%	709	1.9%	$34.93	3.4%	$49.23
-0.3	119	17.7%	711	2.2%	$34.99	3.3%	$49.18
-0.35	118	18.7%	713	2.4%	$35.07	3.0%	$49.18
-0.4	117	19.5%	715	2.7%	$35.16	2.8%	$49.20

On-peak, off-peak, and average wholesale prices are shown in the left-hand side of Figure 15.2 for TOU pricing and in the right-hand side for RTP. Prices drop more with RTP; they are about 4% lower. Both schemes moderate on-peak and off-peak prices on average.

Table **15.3** shows the same on- and off-peak prices as in Figure 15.2 at sample customer elasticities as well as showing results for the most extreme prices. A regulator looking only at prices might be deceived by the apparently small difference between RTP and TOU on average prices, the actual impacts on price and total customer bill are more meaningful on a percentage basis as shown in Table 15.4. We must also look at the impacts on peak load reduction and equity among customers to get a full picture of the factors important to policy makers.

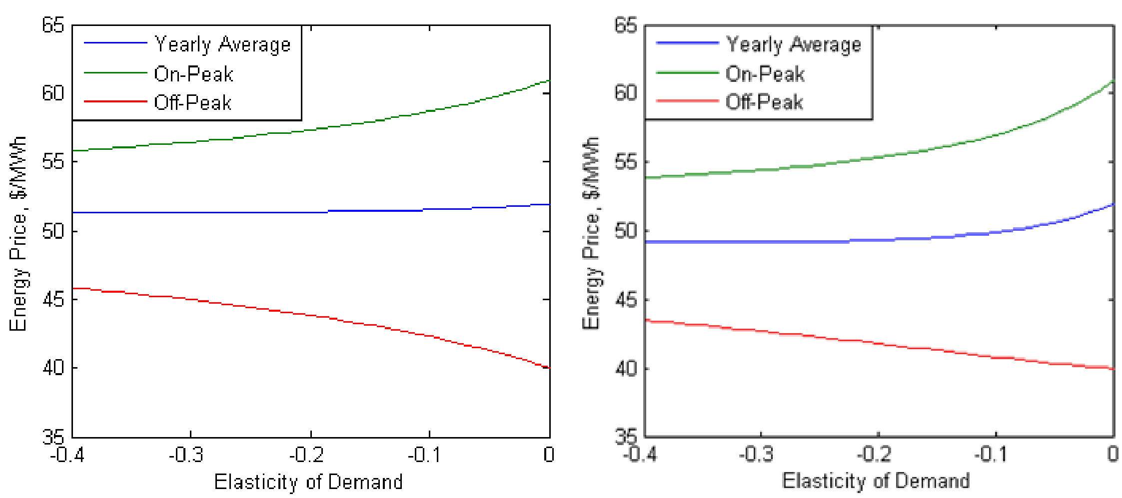


Figure 15.2. On-peak, off-peak, and average prices under the TOU scenario (left) and RTP scenario (right).

Table 15.3. Yearly prices with a change to TOU or RTP.

Elasticity of Demand	On-Peak Price, $/MWh		Off-Peak Price, $/MWh		Highest Price, $/MWh		Lowest Price, $/MWh	
	TOU	RTP	TOU	RTP	TOU	RTP	TOU	RTP
0	$60.92	$60.92	$40.01	$40.01	$292.03	$292.03	$5.23	$5.23
-0.05	$59.87	$58.86	$41.03	$40.28	$286.12	$234.09	$5.89	$10.13
-0.1	$58.86	$57.23	$42.08	$40.72	$280.43	$189.03	$6.58	$14.42
-0.15	$58.08	$56.17	$42.96	$41.20	$276.02	$160.92	$7.13	$17.76
-0.2	$57.46	$55.43	$43.70	$41.69	$272.50	$141.83	$7.60	$20.50
-0.25	$56.95	$54.89	$44.33	$42.16	$269.63	$128.14	$7.99	$22.80
-0.3	$56.53	$54.49	$44.87	$42.61	$267.23	$117.90	$8.33	$24.78
-0.35	$56.18	$54.17	$45.34	$43.04	$265.20	$109.99	$8.62	$25.96
-0.4	$55.87	$53.92	$45.75	$43.43	$263.46	$103.72	$8.88	$26.86

Consumers elect to buy more energy under RTP or TOU conditions as shown in Figure 15.3. Note that TOU and RTP result in prices above the flat-rate price for some hours and below it for others. The result is a drop in the quantity demanded during the high price period and an increase during the low price period. Since there are net customer savings, there is a small net increase in the quantity demanded. Marginal impacts diminish with more responsive load. Customer expenditure on electricity decreases steeply if elasticity is low in magnitude as shown in Figure 15.4. With inelastic demand most of the changes in consumption patterns are small reductions at peak prices. With greater elasticity, dollar savings drop as the effect of greater consumption dominates the overall expense.

These RTP results are explained by the large positive skew in electricity prices and the increasing steepness of supply curves at high load. Large price reductions from small amounts of curtailment at high prices dominate results at elasticities near zero. With increasing responsiveness, the load profile becomes flatter and flatter but overall consumption increases. Under these conditions, the effect of the consumption increase dominates other results. Results with TOU pricing have similar characteristics but only a fraction of the magnitude.

These load flattening and overall consumption increases indicate environmental concerns for two reasons. First, the greater consumption means greater generation and more of the carbon dioxide and air quality emissions associated with that generation. Second, PJM's resource mix is such that peak power is supplied primarily with natural gas and baseload power is supplied primarily with coal. Therefore, a flatter load profile would be supplied by a greater proportion of baseload and shouldering coal plants, resulting in greater emissions of carbon dioxide and air quality pollutants.

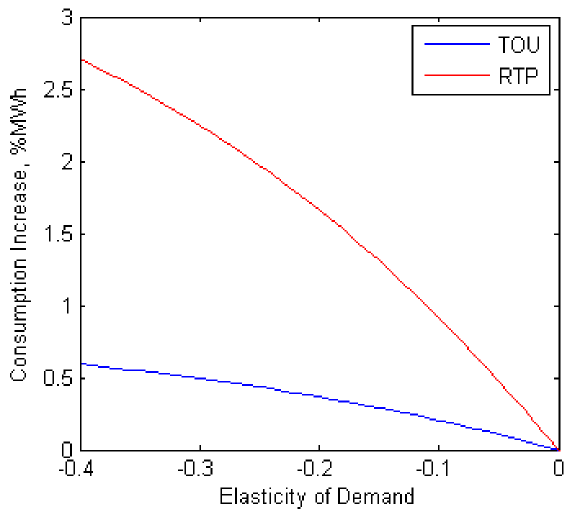


Figure 15.3. Consumption increase, TOU and RTP.

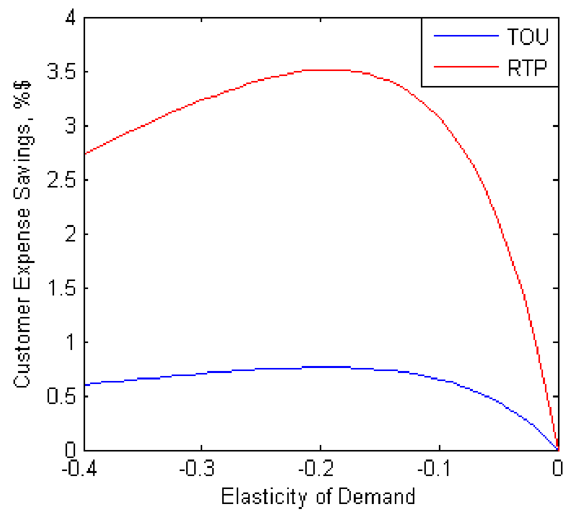


Figure 15.4. Customer bill savings, TOU and RTP.

Because consumers are buying more energy with less total expenditure, the overall impact on consumers is more easily understood by looking at a customer who refuses to change behavior as others do under TOU or RTP. In Figure 15.5, savings are shown for a single customer who has elasticity zero, while the aggregate system has an elasticity shown on the x-axis. I show savings for three types of customers:

Flat – Customer uses a constant level of power during all hours of the year.

Typical – Customer load profile is proportional to the original system load profile.

50% More Extreme – During each hour, the customer demands the typical customer's load *plus* an additional 50% of the difference between the typical customer's load for that hour and the minimum load for the day.

An unchanging typical customer saves less per unit than a responsive customer, but slightly more overall because she does not increase consumption[81]. More interesting is that a flat customer would save 7.0% of her annual electric bill even if *no one* responded to price. She would save the amount that currently goes to subsidize the excesses of more peaky customers. This savings highlights the issue of equity that I raised earlier: under flat rates, moderate and counter-cyclical customers subsidize the consumption of customers with high coincident peak loads.

The more extreme customer loses money under RTP if no one responds, but will have net savings if the aggregate elasticity is even slightly responsive, $E \leq -0.04$.

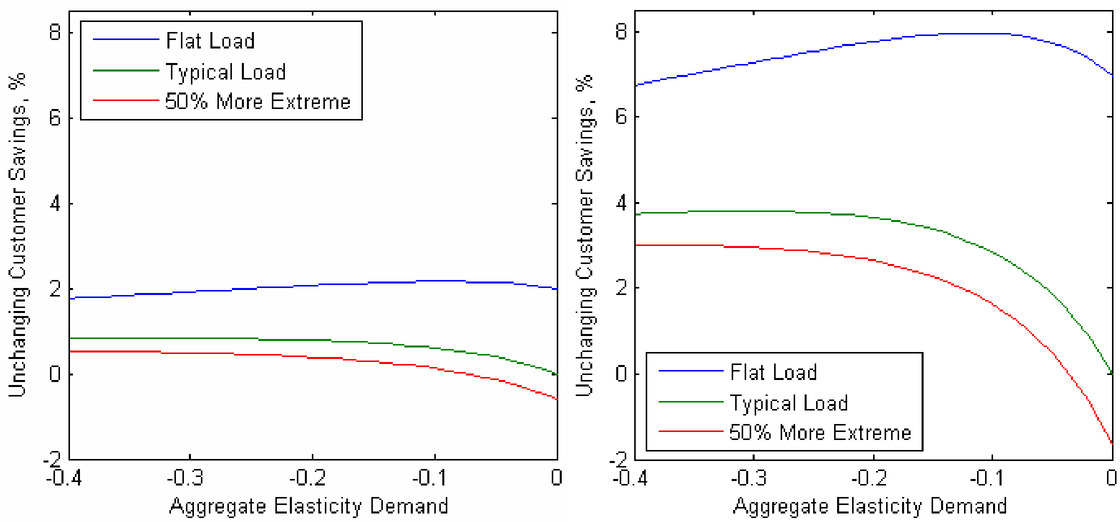


Figure 15.5. Expense savings to an unresponsive customer if others respond, TOU (left) and RTP (right).

Peak load reductions are extreme with a small amount of responsiveness but marginal savings taper with greater responsiveness as shown in Figure 15.6. Discontinuities in Figure 15.6 are caused by a change in the day upon which peak load is observed.

---

[81] At E = -0.2, the typical responsive customer saves 5.0% per unit and 3.5% overall; the typical unresponsive customer saves 3.6% although her quantity consumed is constant.

The large peak load reductions under RTP may have huge implications for total system cost if such reductions persisted over the long run. Peak load determines the total capacity investment necessary for the system to operate reliably. Although no short-run savings will be made on peak capacity that has already been built, there will be long-run savings via unneeded capacity investment as generators have to be replaced or load increases over time.

At elasticity -0.2, peak load drops by 15.1% with RTP. At that level, an overnight capacity value of $700/kW or $2000/kW, corresponding roughly with the overnight capital costs of gas and coal generation from 2006, translates into a dollar savings of $15- $43 billion from a change to RTP if the peak reductions persist in the long run. A change to TOU pricing would reduce $2.0 to $5.6 billion in overnight capacity investments under the same conditions. Note that these capacity savings numbers are ballpark numbers that apply only if our short-run peak savings estimates persist in the long run.

If state regulators and utilities begin to treat RTP as an alternative to investments in new generating capacity, then they will have to compare the costs of investing in new capacity against the costs of implementing RTP. At $15 billion in avoided capacity costs, an integrated resources planner would be willing to spend $294 for each of the 51 million people (note that the population quoted here is much larger than the actual number of customers and therefore the actual number of meters) in PJM territory to implement RTP [92]. Compared to the hardware and installation costs of $243-$311 per unit[82] for the advanced metering infrastructure required to implement RTP, these capacity savings justify RTP rates starting with the largest and most responsive customers [60].

Although I fully investigate these issues is Part IV, I conjecture here that only large customers need to face RTP to achieve most of these savings. From the experience in Niagara Mohawk, the "18% [of customers] with elasticities greater than -0.1 provide 85% of the aggregate price

[82] Original numbers were updated to year 2006$ for inflation [78].

response" [88]. If only a fraction of customers need smart meters, then the cost of implementing RTP would be much smaller than the social benefit, with all customers receiving some benefits via lower average price.

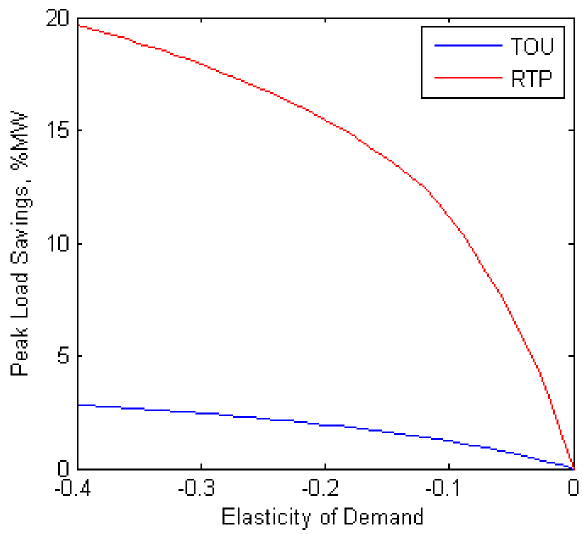


Figure 15.6. Peak load reductions, TOU and RTP.

Figure 15.7 and Table 15.4 show surplus increases with a time-varying rate. Neither consumer nor producer surplus changes monotonically with elasticity. Producer surplus drops slightly with peak price reductions but then increases with overall consumption. Producer surplus is equal to revenue minus operating costs and so indicates profitability if capital costs are not considered. Because we see almost no change in producer surplus, these results indicate that producers will not see the large reduction in profits that they might have feared from RTP. There is no change in consumer surplus for an elasticity of zero, but for an elasticity of -0.2, consumer surplus increases 0.7% for TOU pricing and 3.2% for RTP. We find that TOU pricing has only 20.3%-

21.8% the impact in increasing total surplus that RTP would have[83]. No matter what the assumed elasticity, consumer surplus increases with RTP or TOU[84].

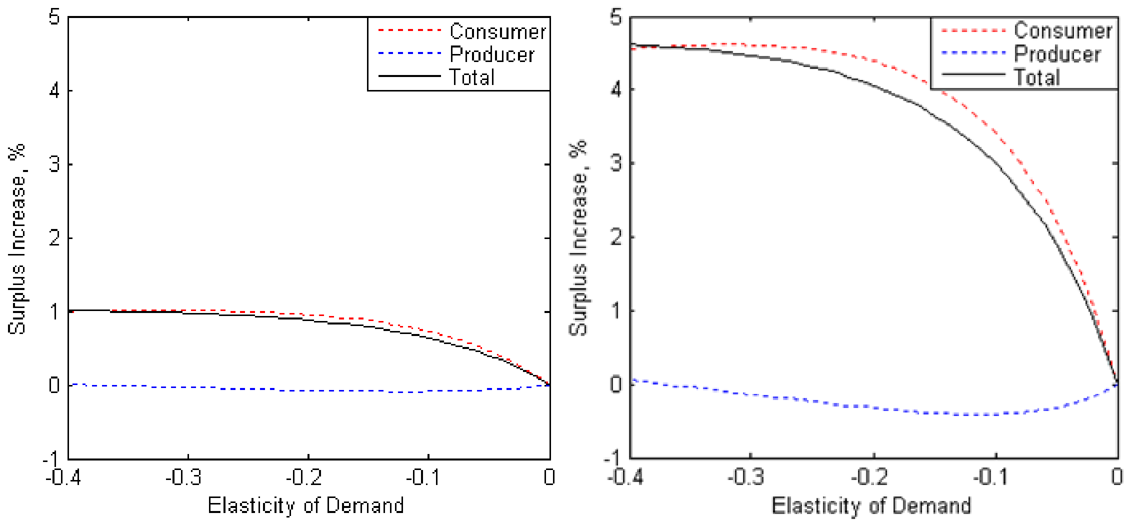


Figure 15.7. Surplus increases with TOU (left) and RTP (right) as a percent of baseline expense.

[83] Although the magnitude of our surplus estimates are much smaller than Borenstein's and much larger than Holland and Mansur's, the ratio of surplus increase between TOU and RTP are remarkably close given the different definitions of TOU used in each case. Borenstein predicted that TOU would have 8-25% the effect of RTP on surplus; Holland and Mansur predicted 15% [46, 85].

[84] The reason for the lack of monotonicity in consumer surplus can be understood by seeing what happens to the area representing ΔCS in Figure 13.1 with extremely steep, moderate, and extremely flat demand curves. A similar figure should be drawn and examined for the case in which load and price increase with RTP.

Table 15.4. Economic outcomes with RTP as a percentage of baseline expenditure.

Elasticity of Demand	Flat-Rate Deadweight Loss	TOU Rate Deadweight Loss	Surplus Increase with TOU			Surplus Increase with RTP		
			Consumer	Producer	Total	Consumer	Producer	Total
0	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
-0.05	1.6%	1.3%	0.4%	-0.1%	0.3%	1.9%	-0.3%	1.6%
-0.1	2.8%	2.2%	0.7%	-0.1%	0.6%	3.2%	-0.4%	2.8%
-0.15	3.5%	2.8%	0.8%	-0.1%	0.8%	3.9%	-0.4%	3.5%
-0.2	4.0%	3.1%	0.9%	-0.1%	0.9%	4.3%	-0.3%	4.0%
-0.25	4.3%	3.3%	1.0%	-0.1%	0.9%	4.5%	-0.3%	4.3%
-0.3	4.4%	3.5%	1.0%	0.0%	1.0%	4.6%	-0.2%	4.4%
-0.35	4.5%	3.6%	1.0%	0.0%	1.0%	4.6%	-0.1%	4.5%
-0.4	4.6%	3.6%	1.0%	0.0%	1.0%	4.6%	0.0%	4.6%

Before looking at these results, a regulator might be concerned about charging RTP for customers who have no ability to respond. It would seem unfair to charge customers high RTPs if they could not react. These results indicate that even if customers had no means of knowing or responding to the RTP, the adverse effect of extremely high prices would not cause any problems on average over the year. Flat and countercyclical customers would benefit by not having to subsidize the excesses of others. Even customers with high coincident peak load would not have a large change in average price and could actually save money from other customers' responses. These results indicate that regulators need not worry about the effect of RTP on poor or unresponsive consumers since they will be better off under RTP even if they did not respond.

Part IV How far can Demand Response Go?

Based on the results from Part III, it appears that the installation of an advanced metering infrastructure (AMI) is a cost-effective alternative to new investments in peaking capacity from an integrated resources planning (IRP) perspective. Further, RTP benefits customers by lowering their overall electric bills. It is not yet clear however, how far it makes sense to take those results. Because I expect most of the response and most of the peak load reductions from RTP to come from a small number of customers, perhaps almost all of the benefits from peak load reductions can be achieved by placing only a small number of customers on AMI at a much reduced cost.

In this section of the thesis I attempt to discover how many customers can be cost-effectively placed on RTP the market model for predicting the short-run impacts of RTP developed in Part III, again using PJM data. I use known load and price information covering a real set of utility customers to determine how many customers should be placed on RTP and how large the peak reductions from those customers are likely to be.

Chapter 16 The Costs of AMI and the Demand Response Gap

In order to determine the per-customer costs of an AMI, I examine a historic range of costs, as well as examining the non-demand response benefits of these systems. I again use PJM data as the basis for simulations, but I used the more recent year 2007 data.

16.1 Costs of Advanced Metering Infrastructure

In a recent report on advanced metering infrastructure (AMI), the Federal Energy Regulatory Commission (FERC) presented the hardware and total capital cost information in Table 16.1, updated here for inflation [60, 78]. I have annualized these numbers over 20 years using an 8% cost of capital. The total capital cost number includes all infrastructure, communications, and installation costs that were included in the AMI, as well as the hardware costs that are also reported. Based on these numbers, I expect installing an AMI to cost $15-$27/meter annually in hardware costs or $26-$32 annually in total capital costs.

Table 16.1. Hardware costs and total capital costs for AMI systems [60, 78].

Utility[85]	Year	Meters, Millions	Total, Million 2007$	Total Capital, Million 2007$	Annualized Hardware per Meter, 2007$	Annualized Capital per Meter, 2007$
DLC	1996	0.6	$159	--	$26.79	--
Virginia Power	1997	0.5	$114	--	$25.81	--
PREPA (Puerto Rico)	1998	1.3	$332	--	$25.99	--
Enel (Italy)	2000	30	$6,456	--	$21.92	--
JEA[86]	2001	0.7	--	$352	--	$25.55
PPL	2002	1.3	$259	$370	$20.29	$28.98
Bangor Hydro	2004	0.1	$17	$33	$15.29	$30.58
TXU	2005	0.3	$40	$81	$16.40	$32.54
PG&E [87]	2005	9.8	$1,536	$2,828	$15.96	$29.39
SDG&E [87]	2006	2.3	$411	$679	$18.16	$30.06

Note that the costs reported in Table 16.1 do not represent a uniform type of AMI. The definition of AMI used by the FERC report is as follows.

> *Advanced metering is a metering system that records customer consumption [and possibly other parameters] hourly or more frequently and that provides for daily or more frequent transmittal of measurements over a communication network to a central collection point [60].*

[85] Utility full names are Duquesne Light Company (DLC), (Dominion) Virginia Power, Puerto Rico Electric Power Authority (PREPA), Ente Nazionale per l'Energia Elettrica (Enel), Jacksonville Electric Authority (JEA), Pennsylvania Power and Light (PPL), Bangor Hydro Electric Company, Texas Utilities (TXU) now the Energy Future Holdings Corporation, Pacific Gas and Electric Company (PG&E), and San Diego Gas and Electric (SDG&E).

[86] Electric and water meters.

[87] Electric and gas meters.

This indicates that any of these systems would be capable of implementing an RTP rate, but these systems could have very different other sets of other capabilities. For instance, these may or may not allow for bi-directional communication or remote connect and disconnect. None of the costs reported here include the additional costs that would be incurred by supplying end users with price displays or equipment for automated price response.

16.2 The Demand Response Gap

The primary intent of Part IV is to examine the benefits from peak load reductions and reduced customer bills in comparison with the costs of implementing an AMI that would make RTP possible. Examining the full costs of an AMI leaves out a good bit of the picture however, since AMI systems have other benefits aside from the possibility of achieving peak load reductions. For example, one of the other reasons a company might roll out an AMI would be if it had very large labor costs from meter reading or service calls for initiating and ending service.

For these reasons, I present in Figure 16.1 results from three AMI business case filings with the California Public Utility Commission (CPUC), showing these utilities' estimates of the full costs and benefits from their proposed AMI installations [78, 93-96]. San Diego Gas and Electric (SDG&E), Southern California Edison (SCE), and Pacific Gas and Electric (PG&E) each presented a different set of costs and benefits to the CPUC. Issues particular to each utility and the legacy system that the AMI is to replace have had a large impact on the numbers reported. Perhaps most relevant is to note that in two of the three cases the utility is installing a system for gas as well as electric meters. In each case I updated present value numbers for inflation and annualized over 20 years at the cost of capital used in the original source as shown in Table 16.2 along with other parameters from the original analysis.

I have shown separately the costs and benefits in terms of capital and operations and maintenance (O&M), and shown the benefits from demand response separately. In each case the demand response benefits are in two components: capacity savings from peak load reductions and energy savings from load shifting to consumption of less expensive electricity.

In none of these three cases is the installation of an AMI justified based solely on the O&M and capital cost savings to the utility; in each case the AMI benefits outweigh the implementation costs only if the utility expects to achieve demand response benefits as well. The additional cost that is not justified by O&M and capital savings is referred to in these studies as the demand response (DR) gap; I will also use this term. Based on these three case studies I expect that the DR gap might be \$4.92-\$10.34/meter/year, but caution that these numbers could be low since the three utilities examined here may have had more non-DR benefits to gain from an AMI than a typical utility and therefore have been more inclined to be first movers in this area.

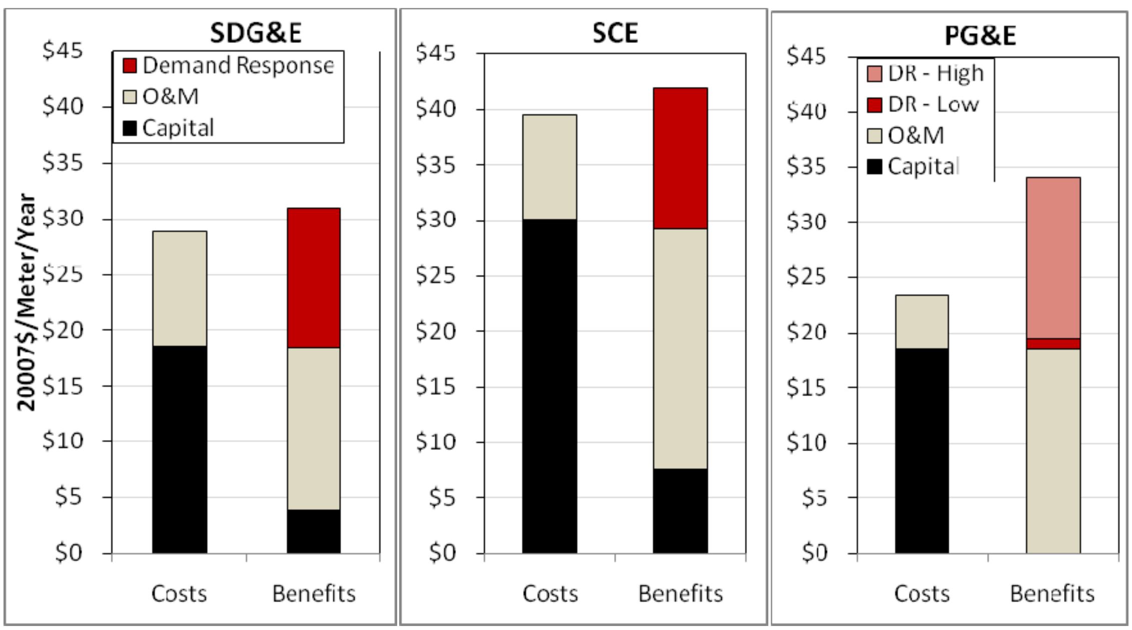


Figure 16.1. Estimated costs and benefits of AMI for all customers in three California utilities [78, 93-96].

Table 16.2. Benefit-cost analysis parameters and number of meters for three AMI business cases [93-96].

Utility	Weighted Cost of Capital	Millions of Meters		Capacity Cost, Nominal $/kW·y	Year of Nominal Dollar
		Electric	Gas		
SDG&E	8.18%	1.4	0.9	$85	2005
SCE	10%[88]	4.5	0	--	2007
PG&E	7.6%	5	4.1	$85	2004

Chapter 17 The Cost of Building New Peaking Capacity[89]

The cost of new generation capacity has increased dramatically in recent years, with natural gas capacity cost having increased by 86% between 2000 and 2007[90] [72]. These dramatic increases have made peak load reductions ever more important; eliminating or delaying the need to build new generation capacity is worth more and more money.

I use the capital cost of a simple cycle gas turbine as the basis of the cost for peaking capacity. I estimate with the recent increases of capacity cost that the price of a simple cycle turbine is $728/kW overnight[91] or $81/kW·y annually[92] [72-74]. The amount of capacity needed to reliably serve the system is greater than the amount of end-use load delivered because of system

[88] This study used this assumed interest rate, but it does not represent the weighted cost of capital as it does in the cases of the other two studies.

[89] The content of this Chapter is largely replicated from that in Section 9.1.

[90] I calculated the 86% number from the two publicly quoted numbers on natural gas: that the capital cost of natural gas has increased 3% between 2007 to 2008, and increased 92% between 2000 and 2008 [72].

[91] I use Integrated Environmental Control Model (IECM) estimates from year 2000 for a natural gas combined cycle (NGCC) plant and then inflate the cost by 86% according to the Power Capital Cost Index (PCCI) to year 2007 values [72, 73]. I then estimate the cost of a simple cycle turbine by applying the ratio of costs between simple and combined cycle plants costs used in the National Energy Modeling System (NEMS) [74].

[92] I inflate the overnight cost over a construction time of three years and then annualize the cost of capital over a 30 year plant life with an 8% cost of capital.

losses and the necessary margin for reliability. I assume a reserve margin[93] of 15% based on the requirements of the Reliability*First* Corporation (RFC), the NERC reliability region that primarily overlaps with the Pennsylvania-New Jersey-Maryland (PJM) Regional Transmission Organization (RTO) [82]. Of this about half is needed for reliability and about half is needed to cover transmission and distribution (T&D) losses; the national average for T&D losses 7.1% [97].

Based on this required reserve margin, I estimate that \$81/kW·y in peak capacity costs translates into a value of \$94/kW·y for peak load reductions. I highlight the distinction between the two numbers to emphasize that a kW of reduction in peak load is worth significantly more than a kW of additional new capacity.

I make one other note about the numbers presented in Figure 16.1 and Table 16.2. In both studies where the annualized value of capacity is noted, the number is \$85/kW·y; a number that is somewhat lower than my expectation. I believe that the higher, updated number more accurately represents the cost of building new capacity and the value of reducing peak load, but recognize that others' assessments may differ. For this reason I show results for my estimate in Chapter 19 as well as providing a full sensitivity analysis in Appendix D.

Chapter 18 Method and Data

The method and data that I use to predict the impacts of a policy change from flat-rate pricing to RTP in PJM are based largely on the method and data used in Part III. I have updated numbers and data to represent the calendar year 2007 and altered the method in order to address a new set of questions. Namely, what is the effect of moving to RTP by placing the largest customers on

---

[93] Reserve margin is expressed as a percent of peak load, distinct from the alternative measure of capacity margin, which is expressed as a percent of installed capacity.

the RTP tariff first and progressively adding smaller customers; at what point do the costs of an AMI outweigh the benefits from RTP?

18.1 Economic Model

I calculate the market equilibrium results with RTP as in Part III but alter the fraction of customers on RTP. The supply side model is unaltered as in Equation (12).

$$(12) \qquad P_S(L) = a \cdot L^3 + b \cdot L^2 + \sum_{t=1}^{n} \{\delta_1 \cdot c_t \cdot L + \delta_0 \cdot d_t\}$$

The demand side model is now considered in two parts. Total load L consists of load L_{RTP} from customers on RTP and load $\underline{L_{flat}}$ from customers on flat rates as in (13). The fraction of load L on RTP is *f*. Load from flat-rate customers does not change with RTP as in (14).

$$(13) \qquad L = L_{RTP} + L_{flat}$$

$$(14) \qquad L_{flat} = (1 - f) \cdot L_0$$

Load from RTP customers *does* change as a function of price according to the constant elasticity demand function in (15), where P_{D0} is the flat-rate price of power in 2007[94].

$$(15) \qquad \begin{aligned} P_D(L_{RTP}) &= \beta \cdot (L_{RTP})^{1/E} \\ \beta &= \frac{P_{D0}}{(f \cdot L_0)^{1/E}} \end{aligned}$$

[94] The number P_{DO} is the sum of the average wholesale price as calculated in (3) and the distribution charge *C* assessed by the local utility.

Equilibrium conditions when only a fraction f of customers are on RTP is in (16), where C is the offset for distribution charges. Price offset between wholesale and retail rates is 3.21cents/kWh in 2006 (cost of delivery-only service); when updated to year 2007 for inflation that results in 3.30 cents/kWh [78, 79].

$$(16) \qquad P_S(L_{RTP} + L_{flat}) + C = P_D(L_{RTP})$$

Results can be calculated after determining any fraction $0 \leq f \leq 1$of load on RTP and any assumed aggregate elasticity of those customers E.

18.2 Utility Customer Data

Not all customers are the same. There is a small number of large commercial and industrial customers and a large number of small residential customers. The makeup and relative sizes of small and large customers determines how many customers have to be on RTP in order to place a certain fraction f of the load under RTP. In general the entire load profile of each customer under flat rates is relevant in predicting what that customer's load profile will be under RTP. However, this hourly load profile information generally will not be available prior to the installation of an AMI.

An appropriate proxy for my purposes is the use of coincident peak load for each customer. Figure 18.1 shows the distribution of customer sizes as measured by coincident peak load for a sample utility with the customers sorted from largest to smallest [98].

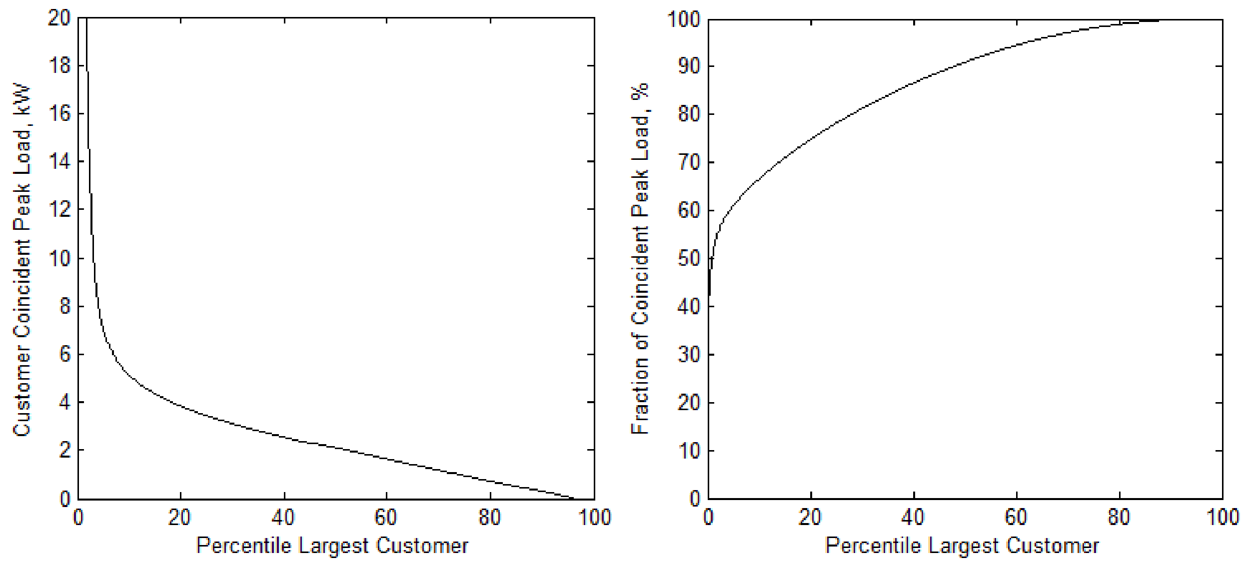


Figure 18.1. Customers sorted from largest to smallest; coincident peak load (left) and fraction of peak load accounted for by the largest customers (right) [98].

I use the coincident peak load contribution of each customer to scale its size, but treat each customer as if it had the same load profile. Coincident peak load is the most relevant piece of information for my purposes because peak load reductions that are possible is dependent upon how much peak load is placed under RTP. Note that the more detailed the information available on each customer or group of customers, the better the installation of an AMI could be targeted to the largest customers and those with the highest coincident peak loads.

Chapter 19 Results

The results of moving customers onto RTP starting with the largest and moving to progressively smaller customers are presented in this section. I use the model and data from Chapter 18 to determine a new predicted load and price profile over the year and then determine the total peak load reduction and total energy bill savings that customers would enjoy from RTP. Although the true economic benefits from this type of policy change are measured by the total consumer and producer surplus change, the real world decisions about whether an AMI should be installed for RTP are made by public utility commissions looking at predictions on customer bills.

For these reasons I calculate the same two customer benefits that were calculated and used in the three CPUC cost-benefit analyses presented in Figure 16.1. As in those cases, the demand response benefits are calculated in two parts: a peak load reduction and a reduction in electric energy bills. These benefits show much money the PUC should be willing to approve in AMI investments in order to place these customers on RTP. As already noted, these "benefits" are not complete from the economist's point of view. Benefits as calculated here represent the perspective of the PUC which is acting on behalf of the customer based on total energy bill, rather than based on consumer surplus or total societal benefits.

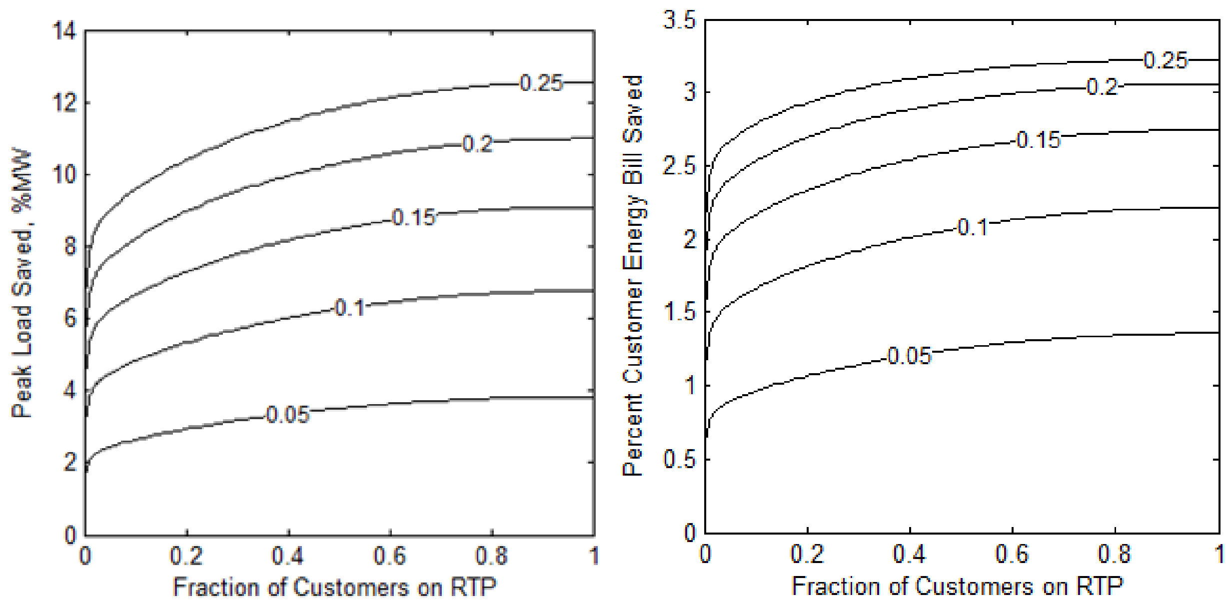


Figure 19.1. Peak load (left) and customer bill (right) savings from placing the largest to the smallest on RTP; various customer elasticities are shown in absolute values in the isocurves.

Once determined on a per-customer, per-year basis, the resulting benefits from placing an increasing fraction of customers on RTP starting from the largest customers are shown in Figure 19.2 through Figure 19.4. Figure 19.2 shows the annual capacity benefits per customer when calculated based on the peak capacity reduction, Figure 19.3 shows the annual electric energy bill reductions, and Figure 19.4 displays the sum of both types of demand response benefits. See Appendix D for a sensitivity analysis of the results from Figure 19.2 and Figure 19.4 depending on the assumed value of capacity.

In each case, the benefits are plotted along with the annualized per-customer costs of installing an AMI. Hardware and total capital costs are plotted based on the numbers from Table 16.1. As discussed in Section 16.2 however, the AMI may have other benefits that are not related to demand response as shown in Figure 16.1. The costs of the AMI minus these other possible benefits is also plotted and labeled as the DR gap in the following figures.

Figure 19.1 shows the savings from placing a certain fraction of customers on RTP with a given elasticity in terms of both peak load reductions and customer energy bill savings. The customer bill savings are expressed as a percent of the total customer bill, including all of those that are on flat and RTP rates, not just those moved to RTP rates. Benefits from both types of customers are relevant because all customers will benefit from lower prices, not just those on RTP. As expected, the more customers are on RTP and the more responsive they are to price, the greater the customer benefits. The most interesting finding from Figure 19.1 is that most of the total possible benefits are achieved with the first few customers.

I note here a few complexities in the real-world implementation of a move toward RTP that are not captured in this analysis, but that should be examined in a similar benefit-cost analysis before moving ahead in a real utility. First, co-located customers may have to be treated in groups so that the communication infrastructure can be implemented effectively; rather than moving from the biggest customer to the next biggest customer for roll-out, the average customer size in a neighborhood would determine the order of rollout. For example, a neighborhood with many pools would be enrolled early.

Second, the initial rate plans of various customer classes would be known and taken into account in more detail than is possible here. Most utilities have a mix of rate classes for their customers, and in PJM, some 5.4% of end user MW in PJM are on rates related the RTP [54, 55][95]. I have not accounted for this complexity, but rather performed the calculations assuming that all customers start on flat rates.

In determining how many customers can be cost-effectively placed on RTP at a given cost per customer, the PUC might look at the benefits in two different ways. In the first, the total benefits achieved from placing customers on RTP is calculated and then averaged over the number of customers that on the new rate. This average cost-benefit calculation is of the type presented in Figure 16.1 and displayed on the left-hand side in Figure 19.2 through Figure 19.4. In this line of thinking, if the total RTP benefits are greater than the total AMI system costs, then the investment in the AMI is justified.

Another way of assessing the possible benefits from RTP, is to recognize that a small number of customers represent a disproportionately large fraction of the load and potential for demand response. In this case, it is relevant to consider the marginal benefit of adding one more customer to RTP. That is, that the incremental capacity and energy benefits achieved from adding the next smallest customer are evaluated and used to determine whether it is cost-effective to place that next customer on RTP. Capacity and energy benefits calculated in this way are shown on the right-hand side in Figure 19.2 through Figure 19.4.

The three figures show, in order, capacity benefits from demand response, average electric bill savings from demand response, and the sum of those two demand response benefits. In each case, the isocurves are labeled with the absolute magnitude of the assumed elasticity of demand. The gray-colored bands represent the range of costs of an AMI depending on how these costs are measured. The range of system hardware costs and total capital costs are the range of values from actual projects as reported in Table 16.1. The DR gap is the range of demand response

[95] Estimate is from 3716 MW on locational marginal price (LMP) based rates and 69,064 MW represented in survey responses. Both distribution utilities and competitive suppliers are represented as survey respondents and the rates charged by both types are reported here.

benefits that would have had to be justified to the PUC in the three utilities examined in Figure 16.1.

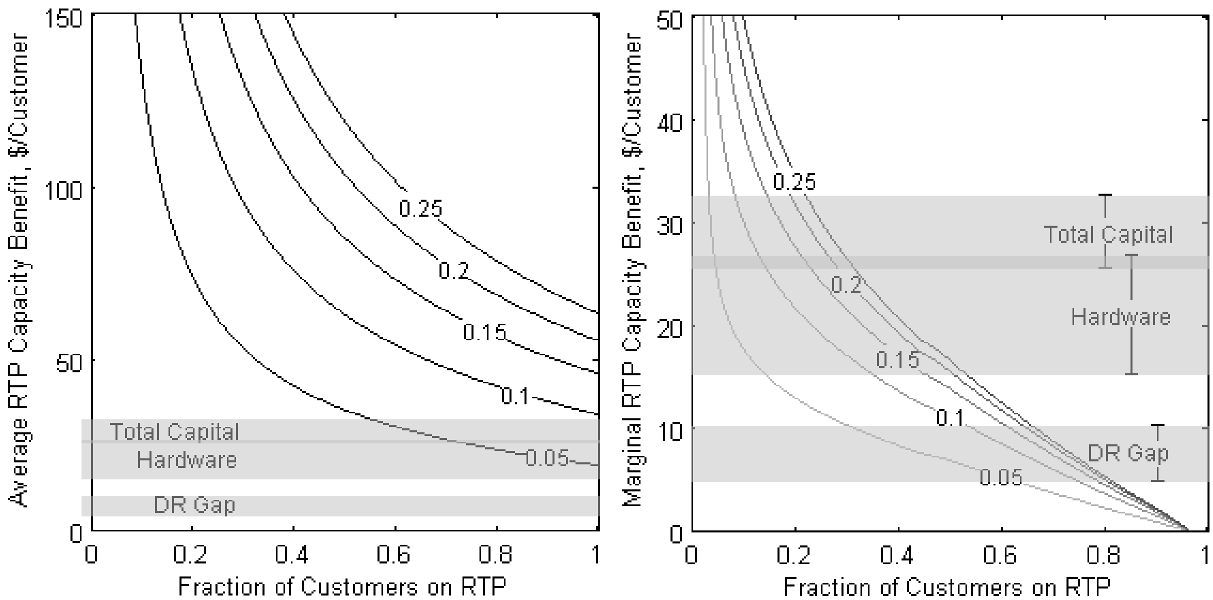


Figure 19.2. Average (left) and marginal (right) capacity benefits from placing a fraction of customers on RTP; various customer elasticities are shown in absolute values in the isocurves.

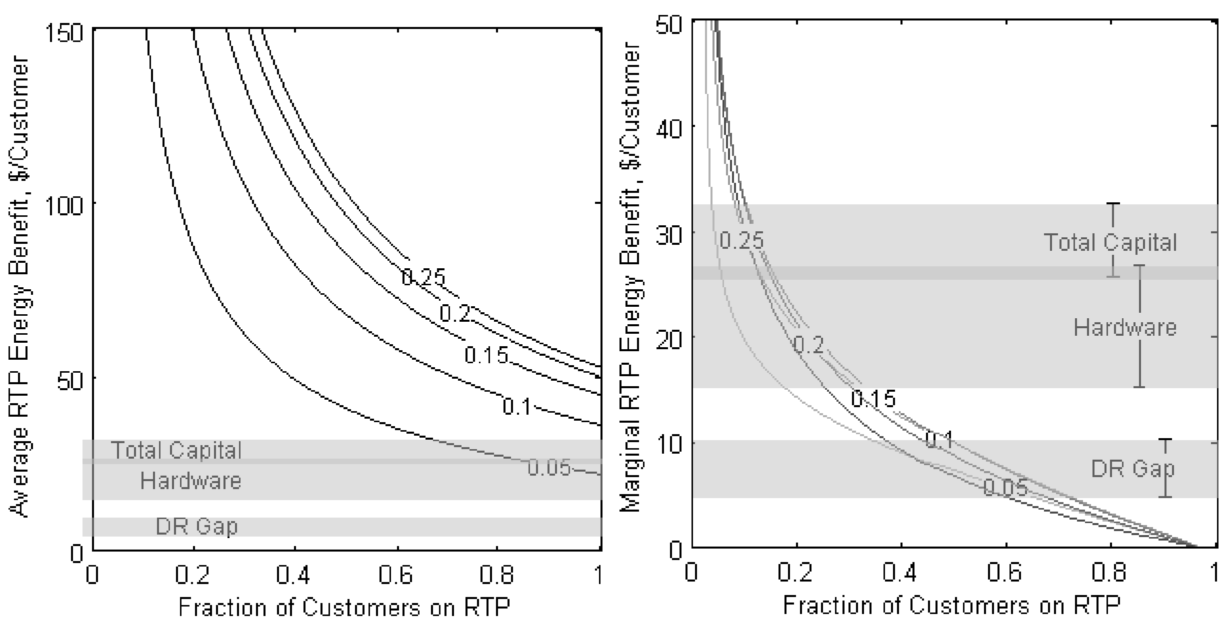


Figure 19.3. Average (left) and marginal (right) electric bill benefits from placing a fraction of customers on RTP; various customer elasticities are shown in absolute values in the isocurves.

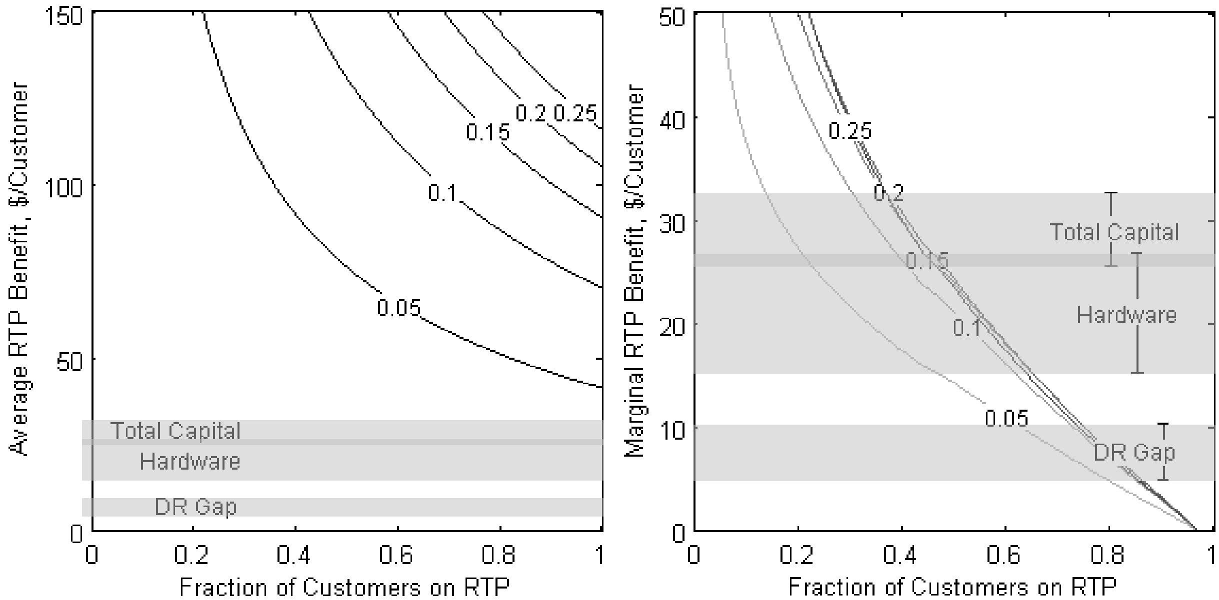


Figure 19.4. Average (left) and marginal (right) total capacity and energy benefits from placing customers on RTP; various customer elasticities are shown in absolute values in the isocurves.

The results displayed in the previous figures and summarized in Figure 19.4 show that an AMI system is justified on average for all customers even if the PUC expects only the very smallest levels of responsiveness.

Looking at the marginal benefits of demand response shows a different picture. If the incremental benefit from placing each additional customer on RTP is weighed against the incremental cost of placing another customer under an AMI, then it becomes clear that there are very large benefits from placing the largest customers on RTP and vanishingly small benefits from placing the smallest few customers on RTP. The figures indicate that it is worthwhile to place the very largest 20% of customers on RTP even if they are not expected to respond much to price and the expense of placing these customers on RTP is very high. This means that an AMI for RTP is justified for industrial, commercial, and a fraction of the largest residential customers.

For the next largest portion of residential customers, the costs do not outweigh the benefits in all cases. If the total capital costs of AMI must be paid via demand response benefits, then 20% to 60% of customers should be placed on RTP depending on the cost of the AMI system and the expected responsiveness of the customers. The tradeoff here is further complicated by the fact that these customers can be made more responsive if additional investment is made in enabling technology. Therefore, responsiveness can also be treated as something that can be influenced by more advanced technologies such as thermostat controls.

In the case where the AMI system has additional benefits to the local distribution utility and only the DR gap must be justified in order to make the investment worthwhile, many more customers are cost-effective for RTP. Some 60% to 85% of customers are appropriate for RTP in that case. In no cases should the utility make an investment to place the very smallest 15% of customers on RTP.

A more detailed display of the demand response benefits from RTP is shown in Table 19.1 and Table 19.2 for cases where the customers are not very responsive with $E = -0.05$ and where the customers are quite responsive with $E = -0.25$. If we have an estimate of the cost of an AMI system, then we can use the results in the following tables to determine what fraction of customers can be cost effectively placed on an AMI under RTP and the minimum size customer appropriate for that investment.

Table 19.1. Demand response benefits from RTP if customers are not very responsive, *E* = -0.05.

Percent of Customers on RTP	Min Customer Size, Coincident Peak kW	Average Benefits, $/Cust/y			Marginal Benefits, $/Cust/y		
		Capacity	Energy	Total	Capacity	Energy	Total
0.5%	85.7	$3,249	$3,940	$7,189	$351	$407	$758
1.0%	36.4	$844	$1,013	$1,856	$134	$154	$288
2.5%	13.0	$501	$599	$1,099	$45.84	$52.06	$97.90
5.0%	7.04	$207	$246	$453	$24.74	$27.89	$52.63
10%	5.16	$121	$144	$265	$17.73	$19.79	$37.52
20%	3.85	$70.57	$83.09	$154	$12.98	$14.27	$27.25
30%	3.13	$51.46	$60.28	$112	$10.36	$11.26	$21.61
40%	2.55	$41.13	$47.98	$89.11	$8.37	$9.00	$17.37
50%	2.13	$35.66	$41.47	$77.13	$6.90	$7.36	$14.26
60%	1.65	$30.67	$35.57	$66.24	$5.31	$5.63	$10.93
70%	1.18	$26.91	$31.14	$58.04	$3.78	$3.99	$7.77
80%	0.73	$23.90	$27.62	$51.52	$2.32	$2.44	$4.77
90%	0.31	$21.42	$24.73	$46.14	$0.98	$1.02	$2.00
100%	0.00	$19.30	$22.28	$41.57	$0.00	$0.00	$0.00

Table 19.2. Demand response benefits from RTP if customers are quite responsive, *E* = -0.25.

Percent of Customers on RTP	Min Customer Size, Coincident Peak kW	Average Benefits, $/Cust/y			Marginal Benefits, $/Cust/y		
		Capacity	Energy	Total	Capacity	Energy	Total
0.5%	85.7	$12,868	$13,540	$26,408	$1,151	$922	$2,074
1.0%	36.4	$3,202	$3,202	$6,405	$420	$313	$732
2.5%	13.0	$1,876	$1,848	$3,725	$138	$95.10	$233
5.0%	7.04	$760	$731	$1,490	$72.20	$47.32	$120
10%	5.16	$439	$414	$853	$49.83	$30.32	$80.15
20%	3.85	$249	$228	$476	$34.54	$18.66	$53.20
30%	3.13	$178	$159	$337	$26.44	$12.90	$39.34
40%	2.55	$140	$123	$263	$20.66	$9.21	$29.87
50%	2.13	$120	$104	$224	$16.58	$6.82	$23.40
60%	1.65	$102	$87.41	$190	$12.48	$4.79	$17.28
70%	1.18	$89.22	$75.34	$165	$8.75	$3.18	$11.93
80%	0.73	$78.85	$66.14	$145	$5.32	$1.86	$7.19
90%	0.31	$70.45	$58.86	$129	$2.22	$0.76	$2.98
100%	0.00	$63.44	$52.95	$116	$0.00	$0.00	$0.00

The tables show that to put the very largest few percent of customers on RTP it would be worth investing several multiples of the cost of a large-scale AMI. For the largest 0.5% of customers with coincident peak loads at 85 kW and higher, it would be worth making significant investments in enabling technology that would be cost-effective even at a cost of some $760 per customer per year if they are not expected to be very responsive, and perhaps $2100 per customer per year if they are expected to be quite responsive.

For wide scale implementation of AMI, all customers above 2.5 kW (about 40% of all customers) could be cost-effectively placed on RTP if there are no benefits to the AMI other than demand response from RTP. For the smallest 10%-20% of customers of size 0.31-0.73 kW in coincident peak load, installing an AMI is not cost effective even under the most favorable of assumptions about other AMI benefits and highly responsive customers.

Part V Conclusions and Recommendations

The traditional assumption that end users cannot vary their consumption as prices change has led to large, unnecessary investments in peaking plants. In 2006, 15% of the generation capacity in PJM territory ran less than 1.1% of the time (96 hours or less), and 20% of capacity ran less than 2.3% of the time (202 hours or less) [84][96]. These under-utilized peak generation investments are a luxury that neither providers nor customers want to pay for.

Peak load reductions are currently being achieved at less than one fourth of the cost of building new capacity at a cost of \$93.72/kW·y and energy efficiency is being achieved at roughly one third the cost of providing more power at \$91.50/MWh.

The good news is that the peak load problem can be mitigated by moving some flat rate customers onto RTP tariffs. Even with little price responsiveness, surprisingly large peak load reductions can be achieved; at elasticities -0.1 and -0.2, 10.4% and 15.1% respectively can be shaved off of coincident peak consumption. Most other quantities of interest such as generator profitability, overall consumption, and average end user expense will not be affected greatly by a change toward RTP. However, policy makers will be disappointed with the short-term reduction in overall bills. A move toward RTP should be driven by concerns about meeting peak load at the lowest cost, enhancing system reliability, and creating equity among end users.

Under current conditions counter-cyclical end users subsidize the high coincident peak loads of others. When problematic, high-peak customers are confronted with higher bills, they will want to make small but important changes. If a peaky customer does not want to alter her consumption habits, then she will face the full price of her own load profile rather than having it subsidized by the rest of the system. Just as consumers have learned to respond to the volatile

[96] This is based on the entire PJM hourly load profile in 2006 [84]. Even at peak load, the system had 17.5% excess available generation capacity. We do not include generation excess at coincident peak load in this calculation because some generation excess is necessary for reliability purposes.

prices of gasoline, fruits, vegetables, and other commodities, so they can learn to respond to electricity prices. The largest difference is that customers purchase electricity every hour of the year and therefore some customers will want automated devices to react to changing prices. Further, for the customers that place a high value on stability in price, retailers could provide any combination of hedges or flat rates; these rates would charge a premium above the RTP rate reflecting the higher cost of service.

Because only modest aggregate price elasticities are necessary for large peak capacity savings, most of the benefits can be achieved by shifting only large, responsive customers to RTP. Further, 50% of all possible customer expense savings from load shifting could be achieved by shifting only 1.7% of all MWh to another time of day. Large, responsive users are the customers who would benefit the most by installing the equipment necessary for automated response to RTP. With RTP, each customer is free to react in the ways that best serve her interest.

Even though RTP is beneficial to the system, for some small customers the expense of installing an AMI is greater than the possible benefits. Customers above 2.5 kW (about 40% of all customers, representing all industrial, all commercial, and large residential customers) could be cost-effectively placed on RTP even under assumptions of high AMI costs and low responsiveness to price. For the smallest 10%-20% of customers of size 0.31-0.73 kW in coincident peak load (representing the smallest residential loads), installing an AMI is not cost effective even under the most favorable of assumptions about other AMI benefits and highly responsive customers.

End Matter

References

[1] Spees, Kathleen, and Lester Lave. "Demand Response and Electricity Market Efficiency." *The Electricity Journal.* Volume 20, Issue 3, April 2007.

[2] Goh, Chong Hock K., and Jay Apt. "Consumer Strategies for Controlling Electric Water Heaters under Dynamic Pricing." Carnegie Mellon Electricity Industry Center. 2004. Available: http://wpweb2.tepper.cmu.edu/ceic/PDFS/CEIC_04_02.pdf

[3] Schweppe, Fred C. "Power Systems 2000: Hierarchal Control Strategies." *IEEE Spectrum,* 1978.

[4] Schweppe, Fred C, Richard D Tabors, James L Kirtley Jr, Hugh R Outhred, Frederick H Pickel, and Alan J Cox. "Homeostatic Utility Control." *IEEE Transactions on Power Apparatus and Systems,* Vol. PAS-99, No. 3 May/June 1980.

[5] Schweppe, FC, Tabors, RD, and Kirtley, JL. "Homeostatic Control for Electric Power Usage." *IEEE Spectrum*. Vol. 19, No. 7. 1982.

[6] Heffner, Grayson C. "Configuring Load as a Resource for Competitive Electricity Markets - Review of Demand Response Programs in the US and Around the World." Lawrence Berkeley National Laboratory. November 2002. Available: http://repositories.cdlib.org/cgi/viewcontent.cgi?article=1747&context=lbnl

[7] *Curtailment Service Providers.* PJM. Available: http://www.pjm.com/services/demand-response/curtailment.html

[8] "Demand Response Service Provider Directory." ISO-NE, June 2006. Available: http://www.iso-ne.com/genrtion_resrcs/dr/providers/drp_directory_060106.pdf

[9] "Board Approves Market Changes to Impose Price Cap." ERCOT Press Release. June 18, 2008. Available: http://www.ercot.com/news/press_releases/2008/nr06-18-08b

[10] *FERC Order on Section 206 Investigation into WECC-Wide Price Cap and the CAISO Ancillary Service Capacity Bid Cap*. February 13,2006. Available: http://www.caiso.com/179e/179ead8256e10.pdf

[11] *RTO-ISO Handbook, Revised - February 28, 2006.* Federal Energy Regulatory Commission. Available: http://www.ferc.gov/industries/electric/indus-act/rto/handbook.asp#skipnavsub

[12] Hogan, William W. On an 'Energy Only' Electricity Market Design for Resource Adequacy. Working Paper. Available: http://stoft.com/metaPage/lib/Hogan-2005-08-energy-only-adequacy.pdf

[13] "Capacity Markets." Federal Energy Regulatory Commission. August 3, 2007. Available: http://www.ferc.gov/market-oversight/reports-analyses/overview/capacity-markets-report.pdf

[14] *PJM Manual 18: PJM Capacity Market, Revision 2.* PJM Forward Market Operations. 2008. Available: http://www.pjm.com/contributions/pjm-manuals/pdf/m18.pdf

[15] *Status of Electric Industry Restructuring Activity.* February 2003. Energy Information Administration. Available: http://www.eia.doe.gov/cneaf/electricity/chg_str/restructure.pdf

[16] Gillingham, Kenneth, Richard G Newell, and Karen Palmer. *Retrospective Examination of Demand-Side Energy Efficiency Policies.* Resources for the Future. Revised September 2004. Available: http://www.rff.org/Documents/RFF-DP-04-19REV.pdf

[17] Ruff, Larry E. "Economic Principles of Demand Response in Electricity." Prepared for Edison Electric Institute. October 2002. Available: http://www.eei.org/industry_issues/retail_services_and_delivery/wise_energy_use/demand_response/economicprinciples.pdf

[18] *Average Retail Price of Electricity to Ultimate Customers by End-Use Sector.* Energy Information Administration. November 2005. Available: http://www.eia.doe.gov/cneaf/electricity/epa/epat7p4.html

[19] Geller, Howard, Toru Kubo, and Steven Nadel. *Overall Savings from Federal Appliance and Equipment Efficiency Standards.* American Council for an Energy Efficient Economy. February 2001. Available: http://www.standardsasap.org/stndsvgs.pdf

[20] VanGeem, Martha G. Energy Codes and Standards. *Whole Building Design Guide.* Available: http://www.wbdg.org/design/energycodes.php

[21] Hildenbrand, Jim L. *Design and Evaluation Criteria for Energy Conservation in New Buildings.* Available: http://nvl.nist.gov/pub/nistpubs/sp958-lide/260-265.pdf

[22] Eto, Joseph. *The Past, Present, and Future of US Utility Demand-Side Management Programs.* Lawrence Berkeley National Laboratory. December 1996. Available: http://eetd.lbl.gov/ea/emp/reports/39931.pdf

[23] Parfomak, Paul, and Lester Lave. "How Many Kilowatts are in a Negawatt? Verifying Ex Post Estimates of Utility Conservation Impacts at the Regional Level." *The Energy Journal*, Vol 17, No.4. October 1996.

[24] Parfomak, Paul. *Electric Utility Conservation Programs: Empirical Studies of Impacts and Cost-Effectiveness.* Doctoral Thesis, Carnegie Mellon University. 1996.

[25] Joskow, Paul L. "Incentive Regulation in Theory and Practice: Electricity Distribution and Transmission Networks" Prepared for the National Bureau of Economic Research Conference on Economic Regulation, September 9-10, 2005. Available: http://www.electricitypolicy.org.uk/events/autumn05/joskow.pdf

[26] "Table 8.13: Electric Utility Demand-Side Management Programs, 1989-2003". *Annual Energy Review.* Energy Information Administration. Available: http://www.eia.doe.gov/emeu/aer/txt/ptb0813.html

[27] Goldman, Charles, Nicole Hopper, Julie Osbern, and Terry Singer. *Review of US ESCO Industry Market Trends: An Empirical Analysis of Project Data*. Lawrence Berkeley National Laboratory. January 2005. Available: http://repositories.cdlib.org/lbnl/LBNL-52320/

[28] Liao, Huaiwei, Jay Apt, Sarosh Talukdar. "Phase Transitions in the Probability of Cascading Failures." Carnegie Electricity Industry Center, 2004. Available: http://wpweb2.tepper.cmu.edu/ceic/papers/ceic-04-08.asp

[29] Barbose, Galen, Charles Goldman, and Bernie Neenan. *A Survey of Utility Experience with Real Time Pricing.* Lawrence Berkeley National Lab. December 2004. Available: http://www.osti.gov/bridge/servlets/purl/836966-SZe2FO/native/836966.pdf

[30] *RTO-ISO Handbook, Revised - February 28, 2006.* Federal Energy Regulatory Commission. Available: http://www.ferc.gov/industries/electric/indus-act/rto/handbook.asp#skipnavsub

[31] *The Value of Independent Regional Grid Operators.* ISO/RTO Council. November 2005. Available: http://www.nyiso.com/public/webdocs/newsroom/press_releases/2005/isortowhitepaper_final11112005.pdf

[32] Talukdar, Sarosh N, Jay Apt, Marija Ilic, Lester B Lave, and M Granger Morgan. "Cascading Failures: Survival vs Prevention." *The Electricity Journal.* November 2003, Vol. 16, Issue 9. Available: http://wpweb2.tepper.cmu.edu/ceic/papers/ceic-03-12.asp

[33] Apt, Jay. "Competition Has Not Lowered US Industrial Electricity Prices." *The Electricity Journal*, Elsevier Inc. 2005. Available: http://wpweb2.tepper.cmu.edu/ceic/papers/ceic-05-01.asp

[34] *PJM Economic Load Response Program.* Available: http://www.pjm.com/services/demand-response/downloads/documentation/20020315-tariff-revisns.pdf

[35] *Load Response in the Ancillary Service Markets.* Available: http://www.pjm.com/services/training/downloads/load-response-in-ancillary-service.pdf

[36] *ISO New England - Communications Hardware Reimbursement Requirements*. ISO New England. 6/10/2003. Available: http://www.iso-ne.com/genrtion_resrcs/dr/ibcs/Communications_Hardware_Reimbursement_Requirements.pdf

[37] Hiskins, IA, and B Gong. "Voltage Stability Enhancement via Model Predictive Control of Load." *Bulk Power System Dynamics and Control VI*, August 22-27, 2004, Cortina d'Ampzzo, Italy. Available: http://www.pserc.org/ecow/get/publicatio/2004public/hiskens_irep.pdf

[38] Black, Jason, and Marija Ilic. "Demand-Based Frequency Control for Distributed Generation." IEEE 2002. Available: http://ieeexplore.ieee.org/iel5/8076/22348/01043270.pdf?arnumber=1043270

[39] Kirby, Brendan J, and John D Kueck. *Spinning Reserve from Pump Load: A Technical Findings Report to the California Department of Water Resources.* November 2003. Oak Ridge National Laboratory. Available: http://certs.lbl.gov/pdf/tm-2003-99.pdf

[40] Kirby, BJ. *Spinning Reserve From Responsive Loads.* March 2003. Oak Ridge National Laboratory. Available: http://certs.lbl.gov/pdf/spinning-reserves.pdf

[41] Parsons, BK, Y Wan, and B Kirby. *Wind Farm Power Fluctuations, Ancillary Services, and System Operating Impact Analysis Activities in the United States.* National Renewable Energy Laboratory. Available: http://www.nrel.gov/docs/fy01osti/30547.pdf

[42] *Scenarios for a Clean Energy Future.* Interlaboratory Working Group on Energy-Efficient and Clean Energy Technologies. November 2000. Available: http://www.nrel.gov/docs/fy01osti/29379.pdf

[43] Nadel, Steven, Anna Shipley, and R Neal Elliott. *The Technical, Economic and Achievable Potential for Energy-Efficiency in the U.S. - A Meta-Analysis of Recent Studies*. American Council for an Energy Efficient Economy. 2004. Available: http://www.aceee.org/conf/04ss/rnemeta.pdf

[44] King, Chris S, and Sanjoy Chatterjee. Predicting California Demand Response: How do Customers React to Hourly Prices? Public Utilities Fortnightly, July 1, 2003. Available: http://www.americanenergyinstitutes.org/research/CaDemandResponse.pdf

[45] Benefits of Demand Response in Electricity Markets and Recommendations for Achieving Them: A Report to the United States Congress Pursuant to Section 1252 of the Energy Policy Act of 2005. US Department of Energy, February 2006. Available: http://www.electricity.doe.gov/documents/congress_1252d.pdf

[46] Borenstein, Severin. "The Long-Run Efficiency of Real-Time Electricity Pricing." *The Energy Journal.* 2005; 26; 3. pp 93. Available: http://repositories.cdlib.org/cgi/viewcontent.cgi?article=1036&context=ucei/csem

[47] Holland, Stephen P. and Erin T. Mansur. *Is Real-Time Pricing Green? The Environmental Impacts of Electricity Demand Variance.* Available: http://repositories.cdlib.org/cgi/viewcontent.cgi?article=1039&context=ucei/csem

[48] Borenstein, Severin. "The Long-Run Efficiency of Real-Time Electricity Pricing." *The Energy Journal.* 2005; 26; 3. pp 93. Available: http://www.ucei.berkeley.edu/PDF/csemwp133r.pdf

[49] Goldman, C, N Hopper, O Sezgen, M Moezzi, and R Bharvirkar. *Customer Response to Day-Ahead Wholesale Market Electricity Prices: Case Study of RTP Program Experience in New York.* Lawrence Berkeley National Laboratory. June 2004. Available: http://drrc.lbl.gov/pubs/NMPC_LBNL_54761.pdf

[50] Mark D Levine, Mark D, Jonathan G Koomey, James E McMahon, and Alan H Sanstad. "Energy Efficiency Policy and Market Failures." Annual Review of Energy and the Environment. 1995. Available: http://arjournals.annualreviews.org/doi/abs/10.1146/annurev.eg.20.110195.002535

[51] Worrell, Ernst, Lynn Price, and Michael Ruth. "Policy Modeling for Energy Efficiency Improvement in US Industry." *Annual Review of Energy and Environment.* 2001. 26:117-43.

[52] Golove, William H, and Joseph H Eto. *Market Barriers to Energy Efficiency: A Critical Reappraisal of the Rationale for Public Policies to Promote Energy Efficiency.* Lawrence Berkeley National Laboratory. March 1996. Available: http://eetd.lbl.gov/ea/EMS/reports/38059.pdf

[53] Cramton, Peter. "Electricity Market Design: The Good, the Bad, and the Ugly." *Proceedings of the 36th Hawaii International Conference on System Sciences*. 2003 IEEE. Available: http://ieeexplore.ieee.org/xpl/abs_free.jsp?arNumber=1173866

[54] *Assessment of PJM Load Response Programs.* PJM Market Monitoring Unit. Report to the Federal Energy Regulatory Commission, Docket No. ER02-1326-006. August 29,2006. Available: http://www.pjm.com/markets/market-monitor/downloads/mmu-reports/dsr-report-2005-august-29-%202006.pdf

[55] *2005 Price Responsive Load Survey Results.* Available: http://www.pjm.com/committees/working-groups/dsrwg/downloads/20060615-05-price-responsive-load-survey.pdf

[56] *PJM Manual 11: Scheduling Operations, Revision 35.* Pennsylvania-New Jersey-Maryland Regional Transmission Organization. June 13, 2008. Available: http://www.pjm.com/contributions/pjm-manuals/pdf/m11.pdf

[57] Assessment of PJM Load Response Programs. PJM report to the FERC. Docket ER02-1326-006. August 29, 2006. Available: http://www.pjm.com/markets/market-monitor/downloads/mmu-reports/dsr-report-2005-august-29-%202006.pdf

[58] *Energy Market Demand Response.* PJM. 2008. Available: http://www.pjm.com/markets/demand-response/energy-market.html

[59] "Network Operations Center Tour." EnerNOC. Available: http://www.enernoc.com/noc_tour.html

[60] *Assessment of Demand Response and Advanced Metering*. Federal Energy Regulatory Commission, Staff Report. Docket Number: AD-06-2-00. August 2006. Available: http://www.ferc.gov/legal/staff-reports/demand-response.pdf

[61] Assessment of Demand Response and Advanced Metering. Federal Energy Regulatory Commission. September 2007. Available: http://www.ferc.gov/legal/staff-reports/09-07-demand-response.pdf

[62] *PJM Economic Load Response Program.* Available: http://www.pjm.com/services/demand-response/downloads/documentation/20020315-tariff-revisns.pdf

[63] Goldman, C, N Hopper, R Bharvirkar, B Neenan, R Boisvert, P Cappers, D Pratt, and K Butkins. *Customer Strategies for Responding to Day-Ahead Market Hourly Electricity Pricing*. Lawrence Berkeley National Laboratory. August 2005. Available: http://eetd.lbl.gov/ea/EMS/reports/57128.pdf

[64] Neenan, Bernie, Donna Pratt, Peter Cappers, and Richard Boisvert. *Does Real-Time Pricing Deliver Demand Response? A Case Study of Niagra Mohawk's Large Customer RTP Tariff.* Lawrence Berkeley National Laboratory. August 2004. Available: http://eetd.lbl.gov/ea/EMS/reports/54974.pdf

[65] Black, Jason W, and Marija Ilic. "Survey of Technologies and Cost Estimates for Residential Electricity Services." IEEE 2001. Available: http://ieeexplore.ieee.org/iel5/7659/20922/00970022.pdf?isnumber=&arnumber=970022

[66] Katzenstein, Warren, and Jay Apt. "Air Emissions Due To Wind and Solar Power - DRAFT." 2008. Carnegie Mellon Electricity Industry Center.

[67] "Hourly Historical Data." ISO-NE. 2008. Available: http://www.iso-ne.com/markets/hstdata/hourly/index.html

[68] "Historic Capacity and Demand." North American Electric Reliability Council. 2007. Available: http://www.nerc.com/page.php?cid=4|38|41

[69] de Haan, Laurens, and Ana Ferreira. *Extreme Value Theory: An Introduction*. 2006. Springer Science+Business Media, LLC.

[70] "Generalized Extreme Value Distribution". MATLAB Documentation. Available: http://www.mathworks.com/access/helpdesk/help/toolbox/stats/index.html?/access/helpdesk/help/toolbox/stats/f4218.html&http://www.mathworks.com/cgi-bin/texis/webinator/search/

[71] "Extreme Value Distribution." Wolfram MathWorld. Available: http://mathworld.wolfram.com/ExtremeValueDistribution.html

[72] "Construction Costs for new Power Plants Continue to Escalate: IHS CERA Power Capital Cost Index." Cambridge Energy Research Associates, Inc. May 27, 2008. Available: http://www.cera.com/aspx/cda/public1/news/pressReleases/pressReleaseDetails.aspx?CID=9505

[73] *Integrated Environmental Control Model.* Developed by Carnegie Mellon University for the National Energy Technology Laboratory. IECM-cs 5.2.2 (c) 2007. Available: http://www.iecm-online.com/

[74] *Assumptions to the Annual Energy Outlook 2007.* Energy Information Administration. 2007. Available: http://www.eia.doe.gov/oiaf/aeo/assumption/index.html

[75] *2008-2017 Forecast Report of Capacity, Energy, Loads, and Transmission.* ISO New England, System Planning. April 2008. Available: http://www.iso-ne.com/trans/celt/report/2008/2008_celt_report_final.pdf

[76] "Installed Capacity Requirement & Reserve Margin Values for the Power Year 2008/2009." ISO-New England. 2008. Available: http://www.iso-ne.com/markets/othrmkts_data/inst_cap/icap/reserve_margins_2008-2009.pdf

[77] *Annual Electric Power Industry Database.* Form EIA-861. Energy Information Administration. Available: http://www.eia.doe.gov/cneaf/electricity/page/eia861.html

[78] *Inflation Calculator.* Bureau of Labor Statistics. Available: http://www.bls.gov/

[79] *Average Retail Price of Electricity to Ultimate Customers by End-Use Sector.* Energy Information Administration. October 2007. Available: http://www.eia.doe.gov/cneaf/electricity/epa/epat7p4.html

[80] *Resource and Transmission Adequacy Recommendations.* North American ELectric Reliability Council. June 15, 2004. Available: ftp://ftp.nerc.com/pub/sys/all_updl/pc/rtatf/RTATF_ReportBOTapprvd_061504.pdf

[81] "Inventory of NERC Regional & Sub-regional Resource Adequacy Assessment Criteria." North American Electric Reliability Council. 2008. Available: http://www.nerc.com/docs/pc/ris/NERC_Res_Adeq_Criteria_Inventory_080130.xls

[82] *2007 Long-Term Reliability Assessment.* North American Electric Reliability Corporation. October 2007. Available: http://www.nerc.com/files/LTRA2007.pdf

[83] Spees, Kathleen, and Lester Lave. "Impacts of Responsive Load in PJM: Load Shifting and Real Time Pricing." *The Energy Journal,* Vol. 29, No. 2. 2008. Available: http://wpweb2.tepper.cmu.edu/ceic/papers/ceic-07-02.asp

[84] PJM Market Data. Available: http://www.pjm.com/markets/market-monitor/data.html

[85] Holland, Stephen P, and Erin T Mansur. "The Short-Run Effects of Time-Varying Prices in Competitive Electricity Markets." *The Energy Journal*, Vol. 27, No. 4.

[86] Wang, Jing, Nuria Encinas Redondo, and Francisco D Galiana. "Demand-Side Reserve Offers in Joint Energy/Reserve Electricity Markets." *IEEE Transactions on Power Systems*, Vol. 18, No. 4, November 2003. Available: http://ieeexplore.ieee.org/iel5/59/27910/01245550.pdf?arnumber=1245550

[87] *PJM Manual 06: Financial Transmission Rights, Revision 8.* Effective Date: March 08, 2006. Available: http://www.pjm.com/contributions/pjm-manuals/pdf/m06v08.pdf

[88] Boisvert, Richard N, Peter Cappers, Charles Goldman, Bernie Neenan, and Nicole Hopper (2007). "Customer Response to RTP in Competitive Markets: A Study of Niagara Mohawk's Standard Offer Tariff," *The Energy Journal,* Vol. 28, No. 1.

[89] Neenan, Bernie, Donna Pratt, Peter Cappers, and Richard Boisvert. *Does Real-Time Pricing Deliver Demand Response? A Case Study of Niagara Mohawk's Large Customer RTP Tariff.* Lawrence Berkeley National Laboratory. August 2004. Available: http://eetd.lbl.gov/ea/EMS/reports/54974.pdf

[90] *Benefits of Demand Response in Electricity Markets and Recommendations for Achieving Them: A Report to the United States Congress Pursuant to Section 1252 of the Energy Policy Act of 2005.* US Department of Energy, February 2006. Available: http://www.oe.energy.gov/DocumentsandMedia/congress_1252d.pdf

[91] eGrid. Environmental Protection Agency. eGRID2006 v 2.1. Data on 2004. Available: http://www.epa.gov/cleanenergy/energy-resources/egrid/index.html

[92] *PJM Overview*. Available: http://www.pjm.com/about/overview.html

[93] Agerter, Linda L., and Peter Ouborg. "Updated Preliminary AMI Business Case Analysis of Pacific Gas and Electric Company (U 39-E)." Pacific Gas and Electric Company before the Public Utilities Commission of the State of California. March 15, 2005. Available: http://www.energetics.com/madri/toolbox/pdfs/business_cases/final_businesscase.pdf

[94] "Testimony Supporting Application for Approval of Advanced Metering Infrastructure Pre-Deployment Activities and Cost Recovery Mechanism: Volume 3 - AMI Preliminary Cost Benefit Analysis". Southern California Edison. December 21, 2006. Available: http://www.sce.com/NR/rdonlyres/B8B338D4-A893-4269-98F0-FDF296628170/0/Vol3_Testimony_AMIPhaseIIApplication.pdf

[95] "Preliminary Statement: II. Balancing Accounts Advanced Metering Infrastructure Blaancing Account (AMIBA)." San Diego Gas & Electric Company. Decision No. 07-04-043. Filed with CPUC May 14, 2007. Available: http://www.sdge.com/tm2/pdf/ELEC_ELEC-PRELIM_AMIBA.pdf

[96] "Application of San Diego Gas & Electric Company (U-902-E) for Adoption of an Advanced Metering Infrasstructure Deployment Scenario and Associated Cost Recovery and Rate Design." Prepared Direct Testimony, March 15, 2005. Available: http://www.energetics.com/madri/toolbox/business.html

[97] "Electricity Overview, 1949-2007." *Annual Energy Review.* June 23, 2008. Available: http://www.eia.doe.gov/aer/elect.html

[98] Proprietary Utility Customer Data. Obtained 2008.

[99] *PJM Manual 06: Financial Transmission Rights, Revision 8.* Effective Date: March 08, 2006. Available: http://www.pjm.com/contributions/pjm-manuals/pdf/m06v08.pdf

[100] Navidi, William. *Statistics for Engineers and Scientists.* McGraw-Hill. 2006.

[101] Krapels, Edward N., and Paul Flemming. *Impacts of the PJM RTO Expansion. November 2005.* Available: http://www.pjm.com/documents/downloads/reports/20051101-impact-pjm-expansion.pdf

[102] Kirschen, Daniel S . "Demand-Side View of Electricity Markets." *IEEE Transactions on Power Systems,* Vol. 18, No. 2, May 2003. Available: http://ieeexplore.ieee.org/iel5/59/26968/01198281.pdf?arnumber=1198281

[103] *Benefits of Demand Response in Electricity Markets and Recommendations for Achieving Them: A Report to the United States Congress Pursuant to Section 1252 of the Energy Policy Act of 2005*. US Department of Energy, February 2006. Available: http://www.oe.energy.gov/DocumentsandMedia/congress_1252d.pdf

[104] "Natural Gas Summary." *Natural Gas Navigator*. Energy Information Administration. Available: http://tonto.eia.doe.gov/dnav/ng/ng_sum_lsum_dcu_nus_a.htm

Appendix A Peak Load Problem over Time in ISO-NE

This appendix shows additional information about the hourly load data from ISO-NE used in Part II and the model used to represent load over time.

A.1 Hourly Data for All Years

Figure A.1. Histograms of hourly load profiles from all years 1980-2007 with peak hours highlighted [67].
Shows histograms of the hourly load data for ISO-NE for each year 1980-2007 [67]. The histograms display the information in the same way that the 1980 and 2006 data are displayed in Figure 8.1. The following figure shows more fully the point illustrated and discussed in Section 8.1 that a larger fraction of capacity in ISO-NE must be available in order to support peak demand over fewer and fewer hours per year.

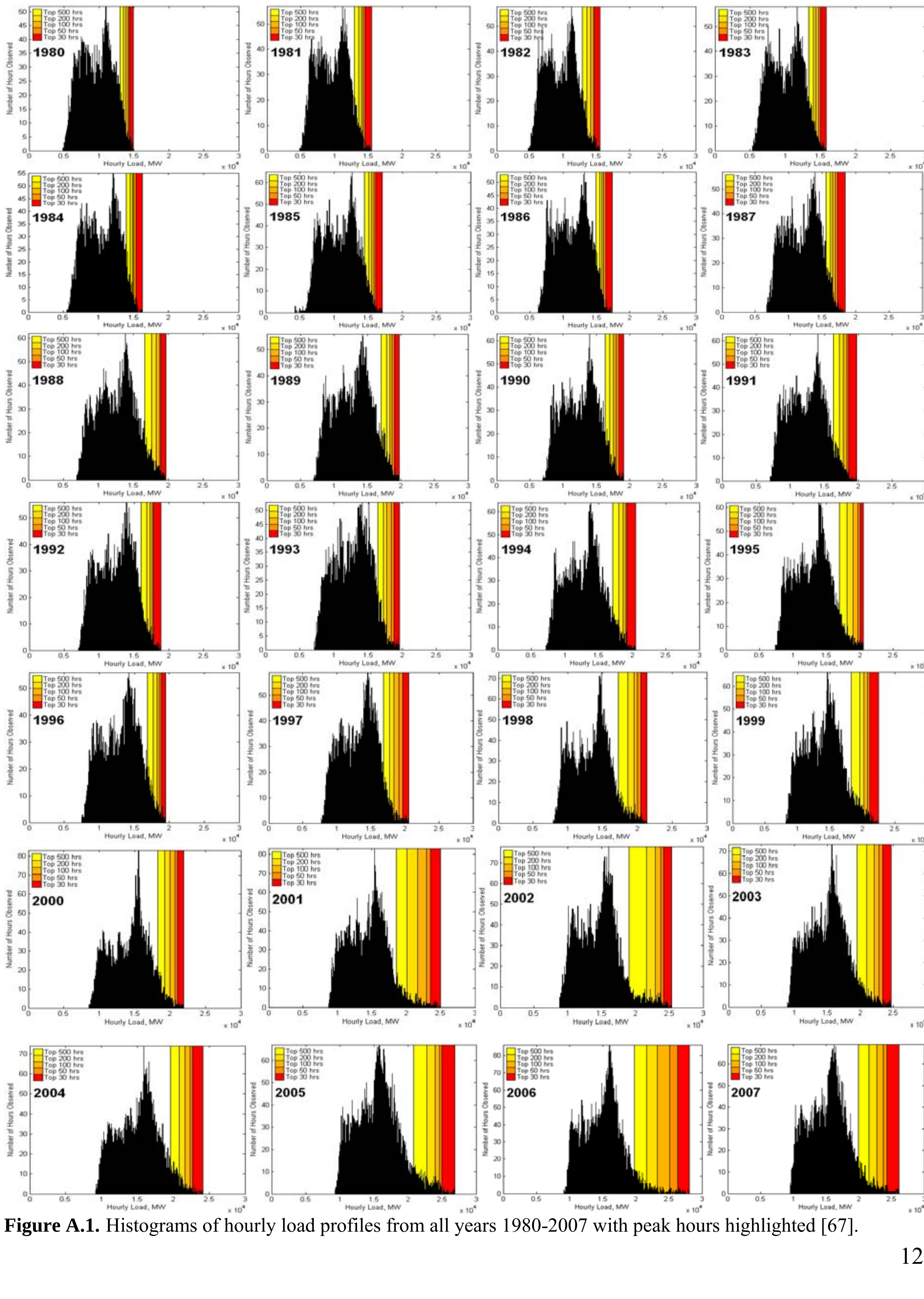


Figure A.1. Histograms of hourly load profiles from all years 1980-2007 with peak hours highlighted [67].

A.2 Additional Peak Load Uncertainty Model Information

Figure A.2 and Figure A.3 display the results when a generalized extreme value distribution is fitted to the weekly observations, with the distribution fitted separately for each year of data. The figures show trends over time for shape, spread, and location parameters along with the 95% confidence intervals for the parameter estimates from each year. Also shown are linear trends on the parameter best estimates with 95% confidence bounds, although the uncertainty in the annual estimates is not accounted for.

The spread parameter σ and the location parameter μ both increase over time; these increases appear to be appropriately represented with a linear trend. The shape parameter *k* shows that it may also have a trend over time, but a constant value over time is well within the predictive uncertainty. I choose to treat the parameter *k* as a constant for simplicity, although the model appears to represent the weekly peak data equally well whether k is treated as constant or with a linear trend.

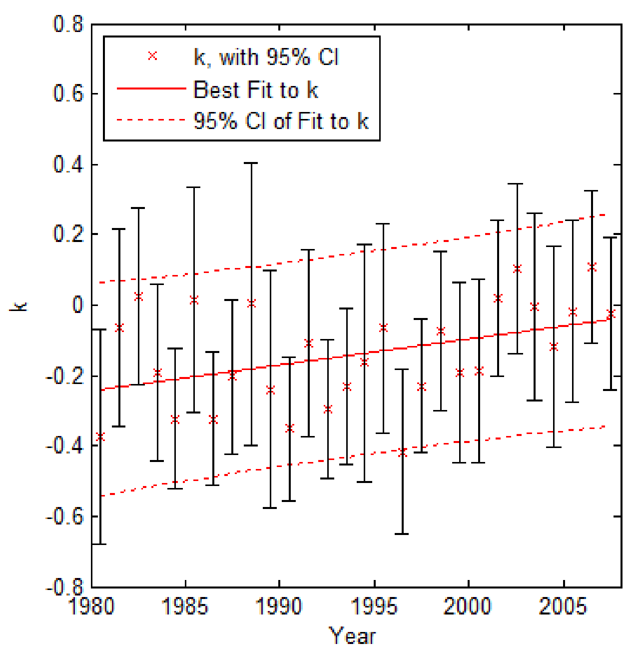


Figure A.2. Generalized extreme value shape parameter *k*, fitted to weekly peak loads for each year.

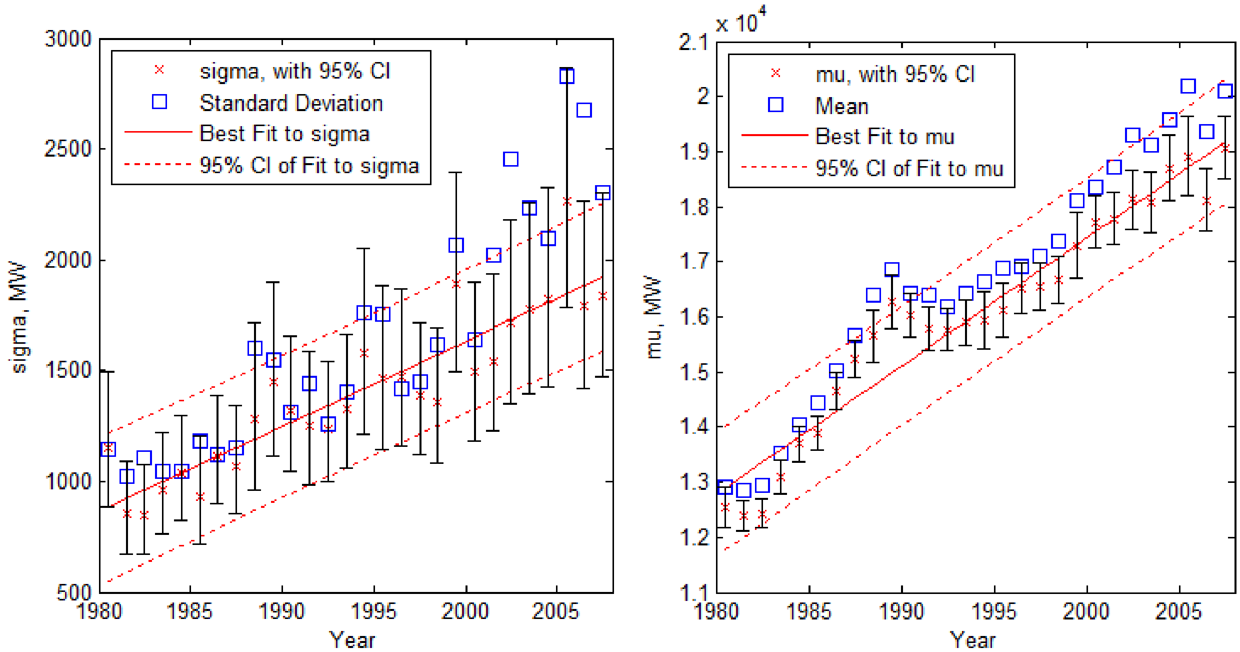


Figure A.3. Generalized extreme value spread (left) and location (right) parameters, fitted to weekly peak loads.

Appendix B Detail on Building a General Supply-Side Model

B.1 Possible Model Structures

I have examined several possible models for predicting price from load using variations on the third-degree polynomial in Equation (17).

$$(17) \qquad P_S(L) = \sum_{t=1}^{n} \left\{ \delta_3 \cdot a_t \cdot L^3 + \delta_2 \cdot b_t \cdot L^2 + \delta_1 \cdot c_t \cdot L + \delta_0 \cdot d_t \right\}$$

Each term in (17) is multiplied by a dummy variable δ_i that has possible values one and zero. These dummy variables act as on-off switches for the term parameter based on the time period t. For example, if I want to assume that each day has a unique third degree supply curve, then the number of time periods is n = 365 and each term will have n = 365 different parameters a_t, b_t, c_t, and d_t, one for each day. The dummy variables δ_0, δ_1, δ_2, and δ_3 ensure that only the parameters appropriate for the time period in question are considered; all others are zeroed out. The resulting values of $P_S(L)$ are then the same as they would have been had I fit 365 different third degree polynomials to the data; overall goodness of fit statistics for that model would have $4 \cdot n$ = 1460 parameters.

Another advantage of the dummy variable approach is that I am able to selectively drop dummy variables from the model to simplify it. For example, my conclusion from these regressions is that the daily fits are the same shape in the second and third degree terms as long as a linear offset is applied to each day as in (8). The simplified model includes only the zero and first degree dummy variables δ_0 and δ_1, and has $2 \cdot n+2$ = 732 p.

I have examined models with eight different definitions of time period *t* and consequent values for *n*. Table B.1 is a summary description of each of the eight time period definitions and the number of model parameters resulting from including a given number of dummy variables from 1 to 4. In each case except for the yearly model there are 15 ways to combine dummy variables.

Table B.1. Description of the eight time period definitions examined.

Time Period	*n*	Number of Parameters by Included Dummy Variables				Description
		1 δ_0	2 δ_0, δ_1	3 $\delta_0, \delta_1, \delta_2$	4 $\delta_0, \delta_1, \delta_2, \delta_3$	
Year	1	NA	2	3	4	One curve. Dropping a dummy means dropping the entire term.
Month of Year	12	15	26	37	48	One curve for each month.
Week of Year	53	56	108	160	212	Week is Mon-Sun. Data begin and end with Wed.
Day of Year	365	368	732	1096	1460	One curve for each day.
Week or Weekend of Year	105	108	212	316	420	One curve for each week Mon-Fri; one curve for each weekend Sat-Sun.
Week or Weekend, Holidays as Weekend	105	108	212	316	420	Append 6 NERC holidays[97] to closest weekend, all happen to fall on Mon or Fri.
Day of Week	7	10	16	22	28	One curve for each day of week.
Hour of Day	24	27	50	73	96	One curve for each hour of day.

B.2 Statistical Significance and Goodness of Fit

For each of the 109 models I have evaluated the goodness of fit statistics. By examining adjusted R^2 values indicating explanatory power, I have concluded that the best way to drop dummy variables is starting with the highest order term and working downward. That is to drop δ_3, then δ_3 and δ_2, then δ_3, δ_2, and δ_1. This ordering is consistent for almost all model types[98]. Table B.2 displays these adjusted R^2 values for the 31 models consistent with this drop ordering. Models are listed in order of decreasing explanatory power; the ordering of models by

---

[97] North American Electric Reliability Corporation (NERC) holidays are considered off-peak hours in PJM [99]

[98] In the day of week model, keeping higher order terms is preferred. Hour of day and month of year models also prefer a higher order term when only two dummy variables are included. These models are poor representations based on the adjusted R^2 values, and so I do not consider these issues further.

explanatory power is identical no matter how many dummy variables are included. Even after dropping two dummy variables, the daily model has more explanatory power than any other model with all four dummy variables.

Table B.2. Model adjusted R^2 values.

Model Sorted in Order of Descending Adjusted R^2	Dummy Variables Included			
	1 δ_0	2 δ_0, δ_1	3 $\delta_0, \delta_1, \delta_2$	4 $\delta_0, \delta_1, \delta_2, \delta_3$
Day of Year	0.9096	0.9488	0.9630	0.9661
Week/WeekendorHoliday	0.8866	0.9124	0.9223	0.9241
Week/Weekend	0.8859	0.9118	0.9221	0.9240
Week of Year	0.8725	0.8961	0.9061	0.9079
Month of Year	0.8521	0.8774	0.8853	0.8887
Hour of Day	0.7990	0.8151	0.8208	0.8225
Day of Week	0.7942	0.8001	0.8085	0.8088
Year	--	0.6925	0.7453	0.7805

The same ordering for dropping dummy variables is dictated by the F-statistic for overall model significance as shown in Table B.3. Because of the large DOF, the p-values associated with these F-statistics are vanishingly small and therefore uninformative.

Table B.3. Overall model F-statistics.

Model Sorted in Order of Descending Adjusted R^2	Number of Dummy Variables Included			
	1 δ_0	2 δ_0, δ_1	3 $\delta_0, \delta_1, \delta_2$	4 $\delta_0, \delta_1, \delta_2, \delta_3$
Day of Year	241	223	210	172
Week/WeekendorHoliday	641	433	331	256
Week/Weekend	637	430	330	255
Week of Year	1091	707	533	410
Month of Year	3607	2509	1879	1490
Hour of Day	1340	789	558	428
Day of Week	3758	2338	1762	1374
Year	--	19732	12815	10387

Model ordering is largely dictated by the number of parameters, the only exception being the month of year and versus hour of day models. Because the theoretical import of the model as decreases as the number of parameters increases, it may be a good idea to accept a model with less explanatory power to obtain a more parsimonious model. Table B.4 shows the explanatory power as calculated by the adjusted R^2 value lost by dropping to the next best model.

Temporal resolution always improves the explanatory power of the model, but the meaning of this observation is clouded by the fact that the higher temporal resolution models use more parameters. The largest drop in explanatory power occurs when moving from sequential time-series to non-sequential bunches of data. That means that Mondays have no interesting common characteristics, but that hours within one day or one week *do* have common characteristics. I conclude from this observation that system conditions change slowly over time and that grouping consecutive hours is a good way to capture these effects.

Table B.4. Adjusted R^2 lost by dropping to next best model.

Model Sorted in Order of Descending Adjusted R^2	Number of Dummy Variables Included			
	1 δ_0	2 δ_0, δ_1	3 $\delta_0, \delta_1, \delta_2$	4 $\delta_0, \delta_1, \delta_2, \delta_3$
Day of Year	0.0230	0.0364	0.0407	0.0420
Week/WeekendorHoliday	0.0007	0.0006	0.0002	0.0001
Week/Weekend	0.0134	0.0157	0.0160	0.0161
Week of Year	0.0204	0.0187	0.0208	0.0192
Month of Year	0.0531	0.0623	0.0645	0.0662
Hour of Day	0.0048	0.0150	0.0123	0.0137
Day of Week	--	0.1076	0.0632	0.0283
Year	--	--	--	--

In deciding how many dummy variables to drop, it is useful to examine the explanatory power lost in dropping the least important dummy variable as shown in Table B.5. I look primarily at the models with time-sequential data groupings. Dropping the δ_3 variable drops explanatory power a miniscule amount. Dropping δ_2 is only slightly worse. Based on this assessment, we conclude that including only linear offsets is a powerful way to represent price and load data. By dropping the number of dummy variables to two, the number of parameters in the model is roughly halved.

Table B.5. Adjusted R^2 lost by dropping one dummy variable.

Model Sorted in Order of Descending Adjusted R^2	Adjusted R^2 Loss from Dropping One Dummy			
	δ_0 to Year	δ_0,δ_1 to δ_0	$\delta_0,\delta_1,\delta_2$ to δ_0,δ_1	$\delta_0,\delta_1,\delta_2,\delta_3$ to $\delta_0,\delta_1,\delta_2$
Day of Year	0.1291	0.0392	0.0142	0.0031
Week/WeekendorHoliday	0.1061	0.0258	0.0099	0.0018
Week/Weekend	0.1054	0.0259	0.0103	0.0019
Week of Year	0.0920	0.0236	0.0100	0.0018
Month of Year	0.0716	0.0253	0.0079	0.0034
Hour of Day	0.0185	0.0161	0.0057	0.0017
Day of Week	0.0137	0.0059	0.0084	0.0003
Year	--	--	0.0528	0.0352

For further insight in determining how many dummy variables to drop, I have calculated an F-statistic for model improvement with and without each dummy variable according to (18) from [100]. The variable k represents the number of parameters a_t through d_t; the variable *SSE* represents the sum of squared error between the real data and model prediction; *N* is the number of data. Subscripts *full* and *reduced* refer to the models with and without the dummy variable respectively.

$$\frac{\left(SSE_{reduced} - SSE_{full}\right)/\left(k_{full} - k_{reduced}\right)}{SSE_{full}/\left(N - k_{full}\right)} \tag{18}$$

Calculated F-statistics are in Table B.6; associated p-values are again vanishingly small. This indicates that keeping additional dummy variables would be justified, although the higher order dummy variables are less important.

Table B.6. F-Statistic for testing the hypothesis that a model is no better than the next best model.

Model Sorted in Order of Descending Adjusted R^2	F-Statistic for Dropping to the Next Best Model			
	δ_0 to Year	δ_0,δ_1 to δ_0	$\delta_0,\delta_1,\delta_2$ to δ_0,δ_1	$\delta_0,\delta_1,\delta_2,\delta_3$ to $\delta_0,\delta_1,\delta_2$
Day of Year	35	19	10	3
Week/WeekendorHoliday[99]	80	25	11	3
Week/Weekend	79	25	12	3
Week of Year	123	39	19	4
Month of Year[100]	386	120	97	26
Hour of Day	36	23	24	4
Day of Week	13	20	65	113
Year	--	--	1814	1409

Each parameter in each model has a t-statistic and an associated p-value measuring its significance in improving the model. Some of the models I examined have upwards of one thousand parameters, so I have grouped the parameters d_t through a_t corresponding to term order zero through three respectively. Table B.7 and Table B.8 show mean and median t-test p-values in the daily and week or weekend-holiday model respectively.

When the dummy variable associated with each parameter is included, the number of parameters is large and the mean and median p-values are displayed without shading. Shaded p-values indicate that the associated dummy variable has been dropped and there is just one parameter of that order that applies to the entire model. In those cases mean and median are the same by definition and so only one is displayed. Bolded results indicate significance at the $p<0.05$ level.

[99] The next best model for the F-statistic is "Week of Year". The "Week/WeekendorHoliday" model has the same number of parameters and a smaller *SSE* than the "Week/Weekend" model, rendering the F-statistic meaningless for that pair.

[100] The next best model for the F-statistic is "Year". The "Month of Year" model has fewer parameters and a smaller *SSE* than the "Hour of Day" model, rendering the F-statistic meaningless for that pair.

Table B.7. Daily model t-test p-values by parameter order and dummy variables included.

	Number of Dummy Variables Included			
Median p-Values	1 δ_0	2 δ_0, δ_1	3 δ_0, δ_1,δ_2	4 δ_0, δ_1,δ_2,δ_3
a_t	**0.000**	**0.000**	**0.000**	0.257
b_t	**0.000**	**0.000**	**0.000**	0.131
c_t	**0.000**	**0.008**	**0.018**	0.058
d_t	**0.001**	**0.000**	**0.000**	**0.002**
Mean p-Values	1	2	3	4
a_t				0.318
b_t			**0.046**	0.254
c_t		**0.000**	**0.000**	**0.000**
d_t	0.139	0.111	0.106	0.134

Table B.8. Week or weekend-holiday t-test p-values by parameter and dummy variables included.

	Number of Dummy Variables Included			
Median p-Values	1 δ_0	2 δ_0, δ_1	3 δ_0, δ_1,δ_2	4 δ_0, δ_1,δ_2,δ_3
a_t	**0.000**	**0.000**	**0.000**	0.277
b_t	**0.000**	**0.000**	**0.001**	0.144
c_t	**0.000**	**0.000**	**0.001**	**0.033**
d_t	**0.000**	**0.000**	**0.001**	**0.003**
Mean p-Values	1	2	3	4
a_t				0.311
b_t			0.067	0.234
c_t		**0.000**	**0.000**	**0.000**
d_t	0.147	0.101	0.087	0.136

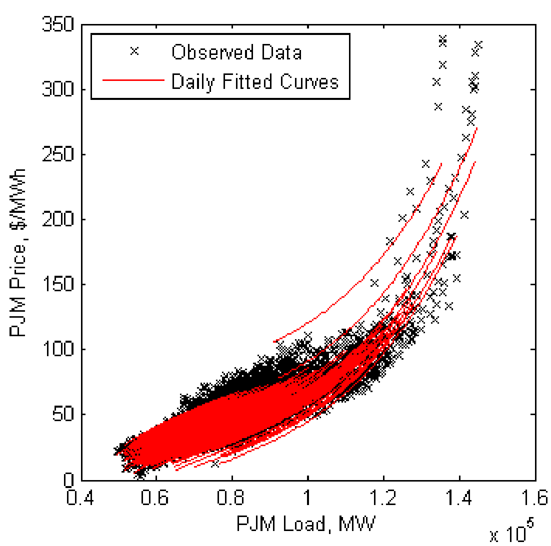
From Table B.7 and Table B.8 it is clear that if a dummy variable is dropped, then including a single parameter for that order term is always a statistically significant improvement to the overall model. The y-intercept, first order, and second order dummy variable parameters are statistically significant in the median but not always in the mean. The median number is a more useful measure because these distributions have strong positive skews. The third order dummy variable parameters do not show statistical significance. Examination of these t-test results justifies dropping one dummy variable and keeping the remaining three.

B.3 Visual Examination of Model Characteristics

Adjusted R^2 results suggest moving ahead with the daily model and only two dummy variables. The t-test results indicate that including three dummy variables is justified. F-statistics suggest reducing the time resolution.

Figure B.1 through Figure B.4 display predictions from models including 1 through 4 dummy variables respectively. Original data are plotted in black in the background; curves representing each time-period *t* are plotted in red in the foreground. Prices for each time period *t* are plotted over the range of loads observed in that time period. Left-hand plots represent the daily models, right-hand plots represent the week/weekend-holiday models.

From Figure B.1 it is clear that more than one dummy variable must be included in order to get a decent representation of the overall data characteristics. The weekly/weekend-holiday models in Figure B.2 through Figure B.4 do appear to represent general characteristics of the data but do poorly in the extremes. Especially obvious is the inability of the weekly/weekend-holiday models to capture the excessively high prices that are an important part of a demand-response analysis. The daily models are able to capture these high-price characteristics by including two to four dummy variables.



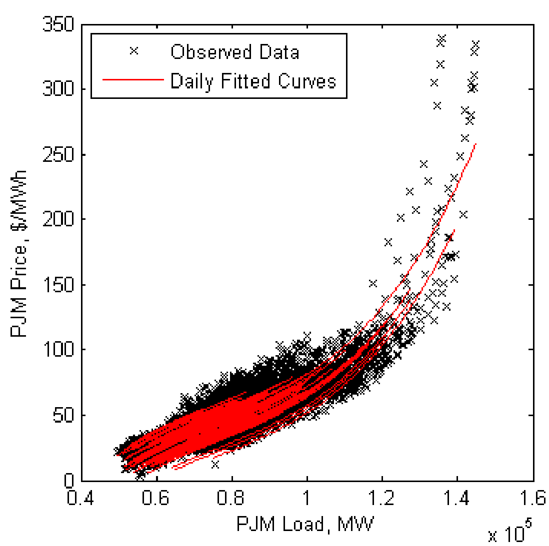


Figure B.1. Data plotted with model curves using 1 dummy variable, daily (left) and week/weekend (right).

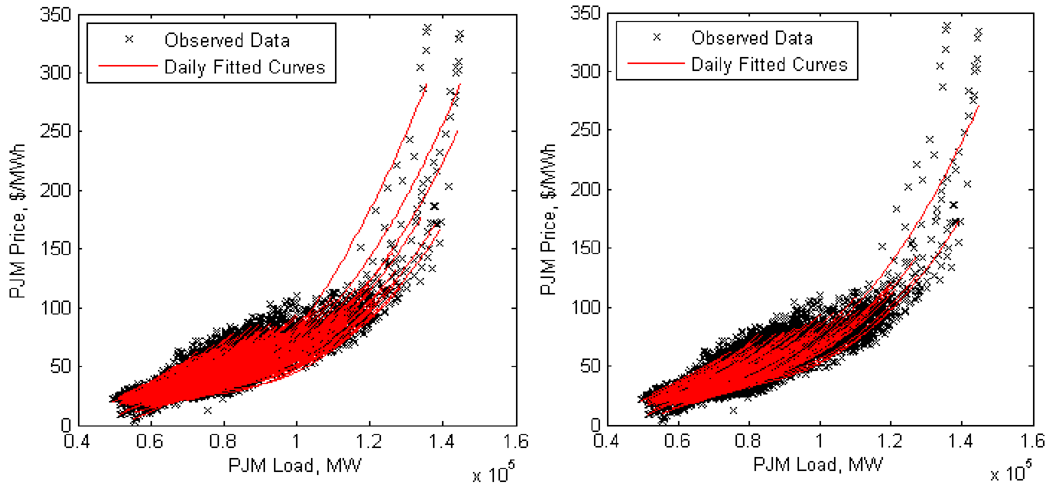


Figure B.2. Data plotted with model curves using 2 dummy variables, daily (left) and week/weekend (right).

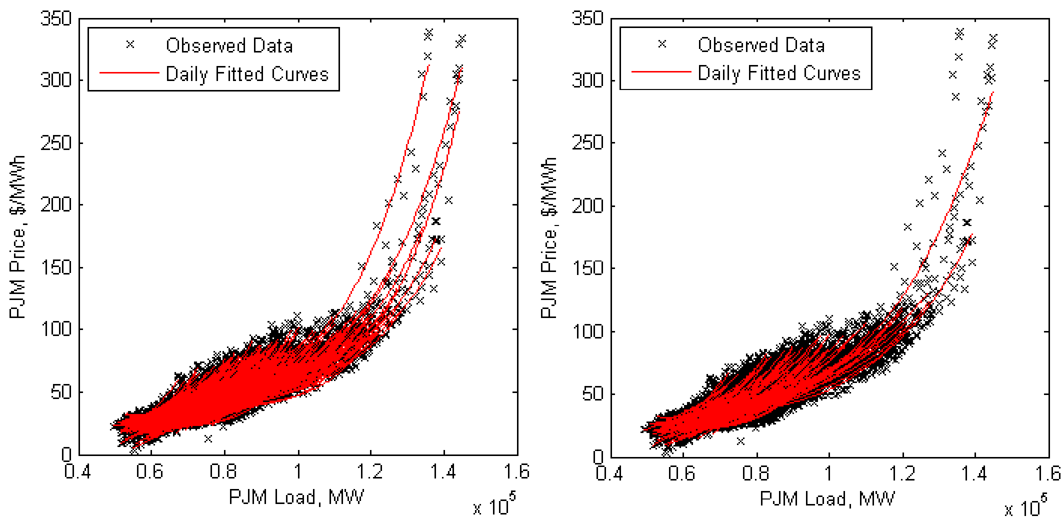


Figure B.3. Data plotted with model curves using 3 dummy variables, daily (left) and week/weekend (right).

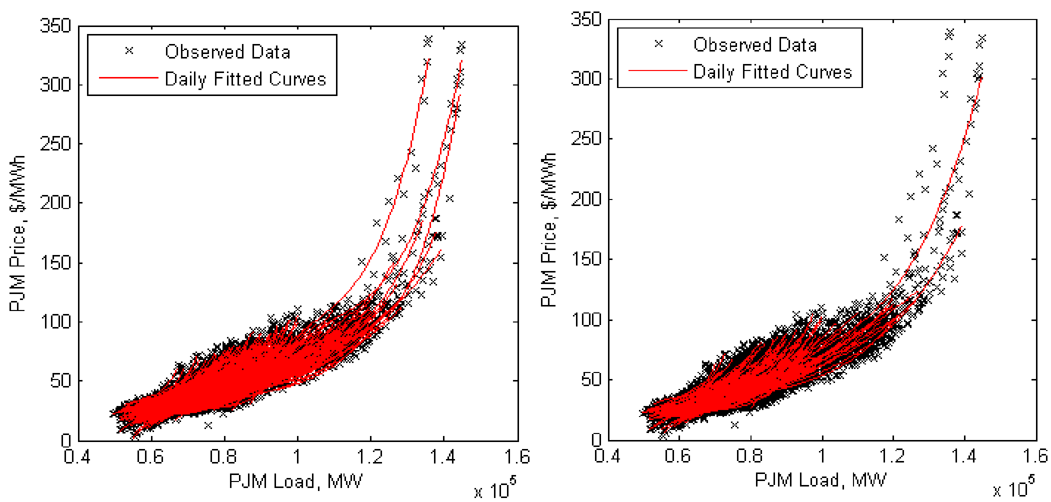


Figure B.4. Data plotted with model curves using 4 dummy variables, daily (left) and week/weekend (right).

Each dummy variable adds some small predictive ability to the daily model, especially in the high and low price extremes. Another important issue is whether the model can make small extrapolations outside observed daily loads. In addressing that issue we have looked at data and prediction plots for all 365 days for each set of dummy variables. I plotted along with those curves the most extreme daily demand curves[101] to determine the largest amount of extrapolation required. When I include all 4 dummy variables the price predictions can go off-course with extrapolation, but with fewer dummy variables the extrapolative ability improves. Every single day appears to have acceptable extrapolative ability when using only two dummy variables.

Based on these observations, I conclude that the best overall supply-side model for analyzing RTP effects is the third-degree polynomial model with linear daily offsets as in (8).

[101] From (4) with elasticity -0.4 and minimum or maximum daily load.

Appendix C Split between Supply Curve and Stacked Bid Curves

I have claimed that using generator marginal cost curves to approximate supply curves underestimates both price and slope. In order to support that claim, I have used PJM data on generator bids into the market to construct day-ahead hourly bid curves [84]. Most generators supply one bid curve into the market that will apply for the entire 24 hours, but others self-schedule their generation amounting to an hourly zero-price offset. A small number of hourly increments or decrements are bid at a non-zero price. I have constructed bid curves for every hour of the year from June 1, 2005 to May 31, 2006 by accounting for each of these bid types.

In this Appendix I examine an earlier time period than in the rest of this paper because the generator bid data are released only after a six-month delay and are unavailable at this time[102].

Aggregate bid curves vary little over the course of one day; the maximum total available load offset in the time frame we observed was 5.6% between daily maximum and daily minimum. On the left-hand side of Figure C.1 I have plotted the bid curves for noon of every day on the left. The bid curves have the hockey-stick shape typical of system marginal cost curves. As noted earlier it is common in literature to find that the bid curve is assumed to be the true supply curve [85, 102, 103]. When this assumption is made the conclusion is that small changes in load have almost no effect on price except at high loads above the "elbow".

On the right-hand side of Figure C.1 real market clearing results are shown along with the same noon bid curves. In this second plot, I have shown only the section of bid curve corresponding with the actual daily load range. If the bid curve were a good approximation of the system supply curve, then real market results would be close to the bid curve range. This graph makes it clear that bid curves are a poor approximation of overall supply. This is because the real-time

[102] Dominion merged with PJM on May 1, 2005 and increased system peak load increased by 18.6% [101]. The data start date allows us to study one contiguous years' worth of data without a territory expansion; the end date allows access to the generator bid data on a six-month delay.

constraints on generator dispatch including unit commitment, transmission constraints, and operating reserves are ignored. Real market prices are much higher than would be predicted by these bid curves.

Figure C.2 shows the same data along with daily fitted supply curves from (8); this model has overall adjusted $R^2 = 0.942$ and an overall model F-statistic of 194. By comparing Figure C.2with Figure C.1, it is clear that after accounting for real-time system constraints, the supply curves have a much steeper slope than the bid curves even at moderate and low load. This implies that small changes in load can have large impacts on price that would not be predicted by examining bid curves. Supply curve slope is the most influential factor in determining the impact of a small change in load on price.

These data covering the summer of 2005 have some qualitative differences from the 2006 data examined elsewhere. High price extremes were greater in 2006 because load extremes were also higher. We also see that high prices were observed even on days when load was moderate or low. This is because electric generators faced high natural gas prices in the fall of 2005 [104]. Natural gas generators are more versatile in load-following and are scheduled during a few hours of day even when overall demand is not high.

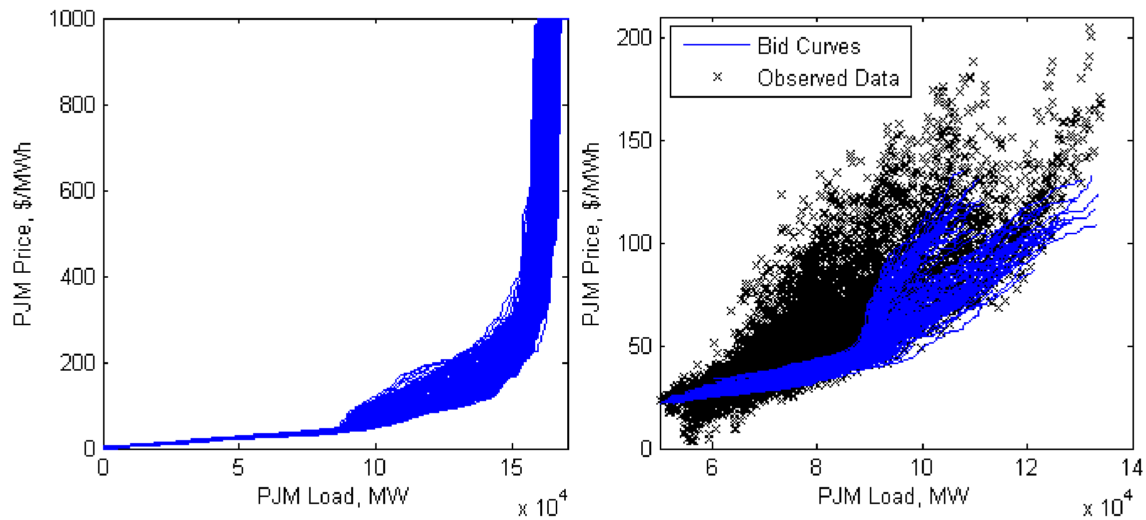


Figure C.1. Daily bid curves at noon (left); noon bid curves with observed data (right).

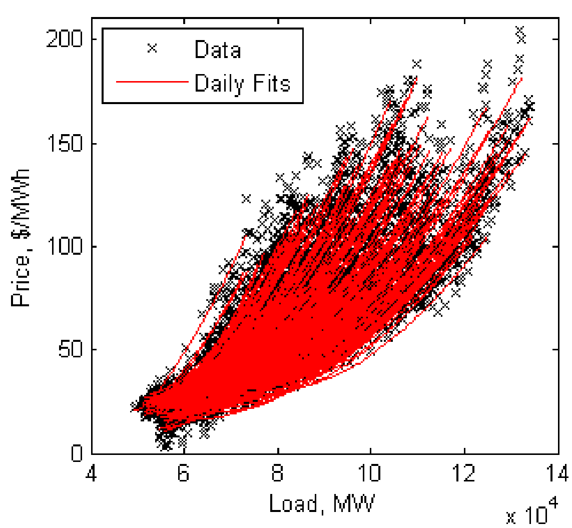


Figure C.2. Data plotted with daily fits from (8).

Many analyses of demand response have assumed away the effects of system constraints outright and taken the implications of a constraint-free stacked bid curve to their logical conclusions [46, 85, 102, 103]. By examining aggregate system results shown here, the magnitude of the discrepancy between the bid curves and actual system results becomes clear.

With a fitted supply model and the observed load, I can predict what price would have been in any hour. Table C.1 compares the average, minimum, and maximum prices predicted by a bid-curve model, a supply-curve model, and the actual observed prices. Based on this comparison it is clear that the bid-curve model predicts prices that are much too low, although not as low as they ought to be in low-load hours. The comparison implies a $15.88/MWh average premium for system constraints.

Table C.1. Prices observed versus predicted by bid curve and supply curve models.

	Bid Curve	Supply Curve	Observed
Minimum Price	$21.87	$10.66	$3.34
Maximum Price	$138.73	$181.64	$204.46
Median Price	$36.97	$54.17	$53.10
Average Price	$48.57	$69.44	$69.44

The comparison of prediction and observed price duration curves in Figure C.3 again show that bid curve price predictions are almost always too low. The duration curve predicted by the supply curves is indistinguishable from the curve actually observed.

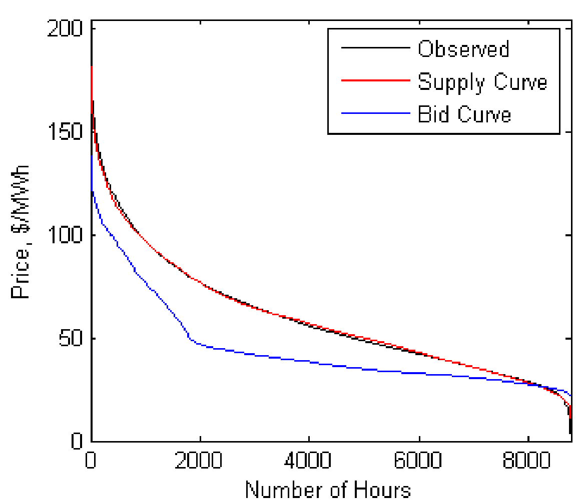


Figure C.3. Predicted and observed price duration curves.

By plotting observed prices against predicted prices in Figure C.4, a richer comparison of model quality can be made. If either model were perfect, then the scatterplot of predicted price and observed price would fall along the identity line. In order to show how close each model comes to the identity line I have plotted that along with the line outputted from a linear least-squares regression. Bid curve predicted prices are systematically lower than real prices, and are never observed in the high price region. The odd-looking heteroscedasticity in the left graph can be understood by comparing it with Figure C.4, but the general conclusion that the bid curve does not accurately represent the characteristics of observed prices and has poor predictability with adjusted $R^2 = 0.673$. The supply curve model best fit line is indistinguishable from the identity line and error appears to be evenly distributed up and down except at the most extreme prices.

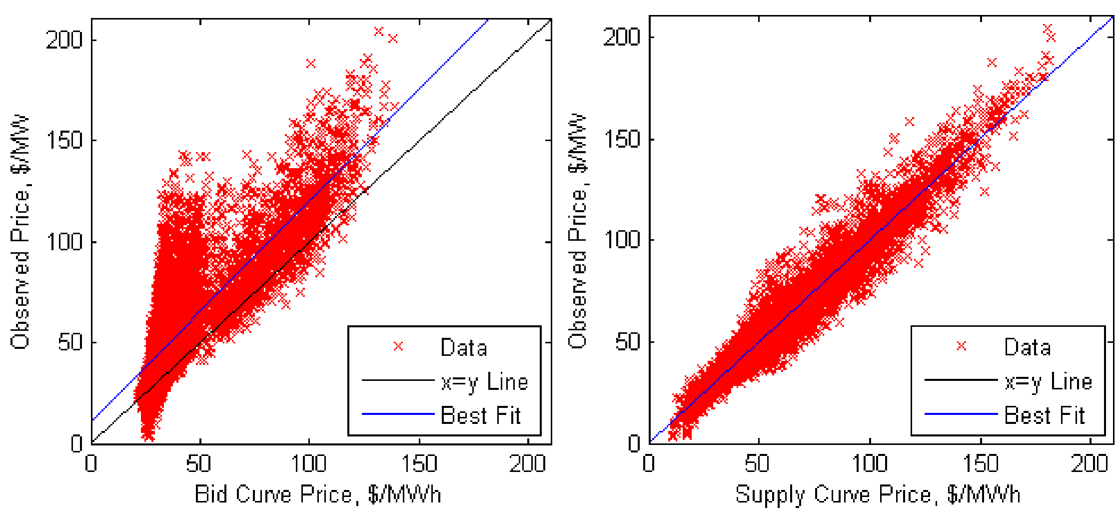


Figure C.4. Observed price versus prices predicted by bid curves (left) and supply curves (right).

Appendix D Sensitivity of Smallest RTP Customer to Peak kW Value

I present here a summary of the results from Chapter 19 with varying capacity costs.

D.1 Results with $55/kW·y

The capacity benefit and total benefit results from Chapter 19 are displayed here if the value of peak reductions is assumed to be $55/kW·y.

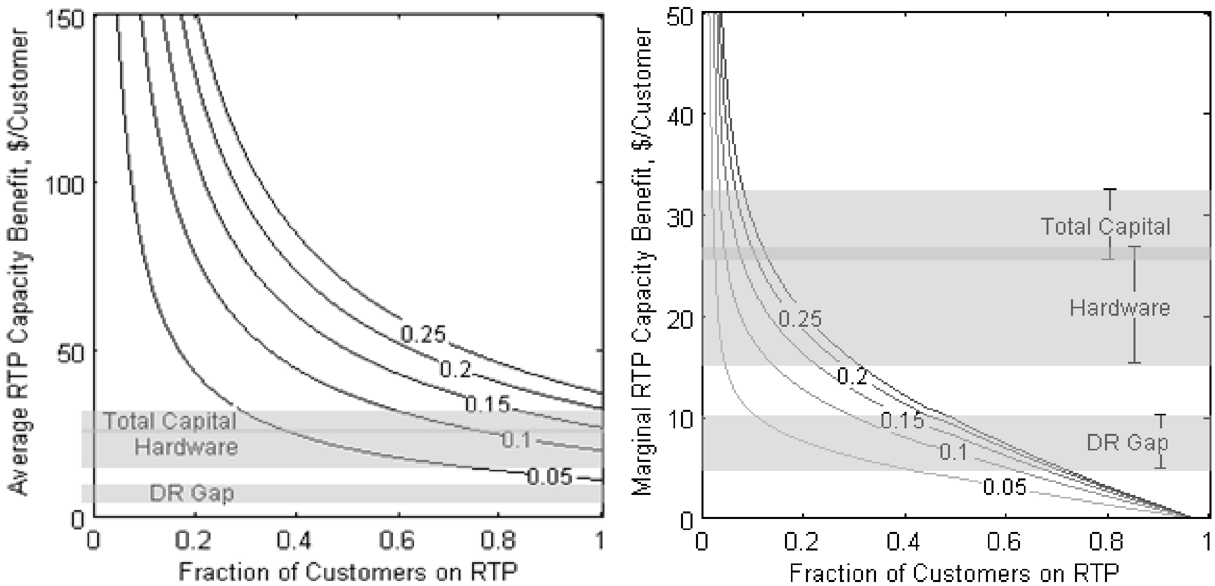


Figure D.1. Average (left) and marginal (right) capacity benefits, peak reduction value $55/kW·y.

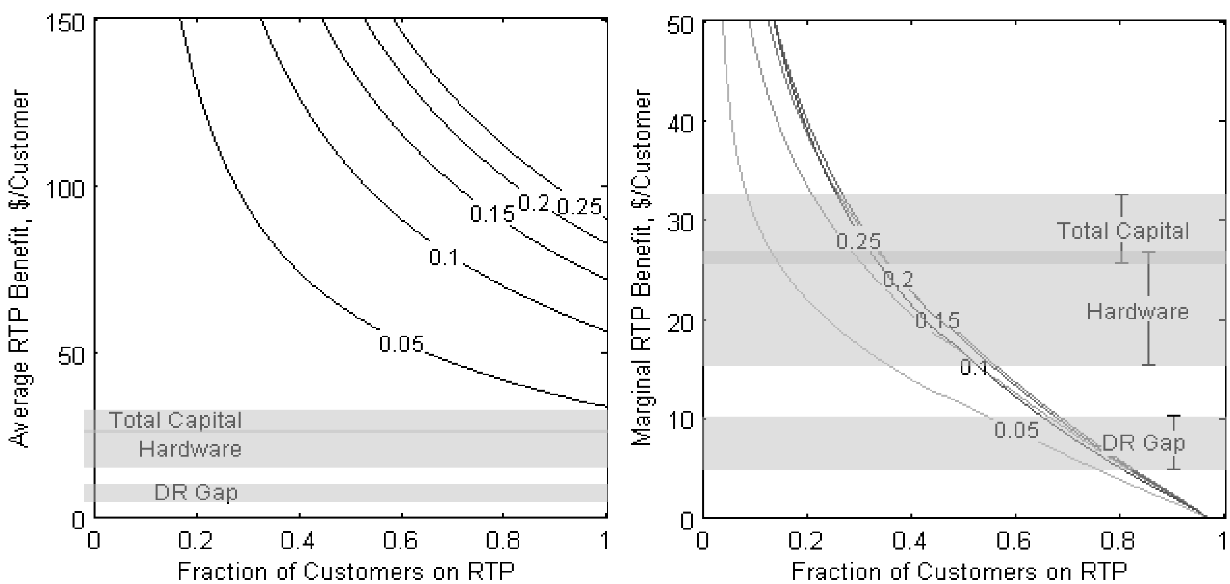


Figure D.2. Average (left) and marginal (right) total capacity and energy benefits, peak reduction value $55/kW·y.

D.2 Results with \$70/kW·y

The capacity benefit and total benefit results from Chapter 19 are displayed here if the value of peak reductions is assumed to be \$70/kW·y.

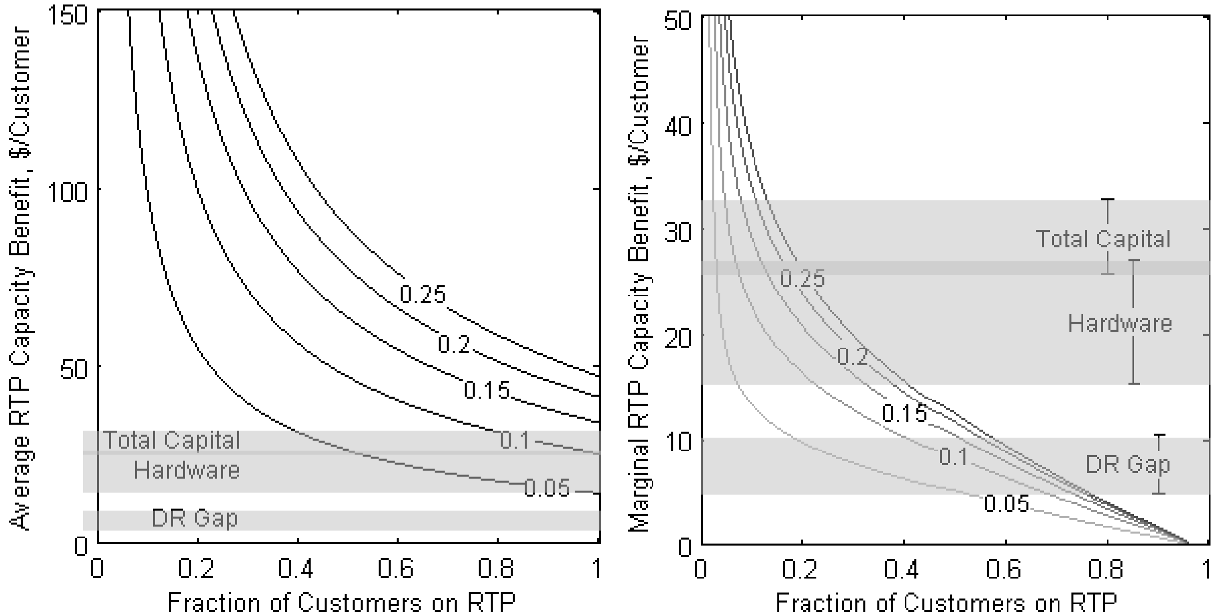


Figure D.3. Average (left) and marginal (right) capacity benefits, peak reduction value \$70/kW·y.

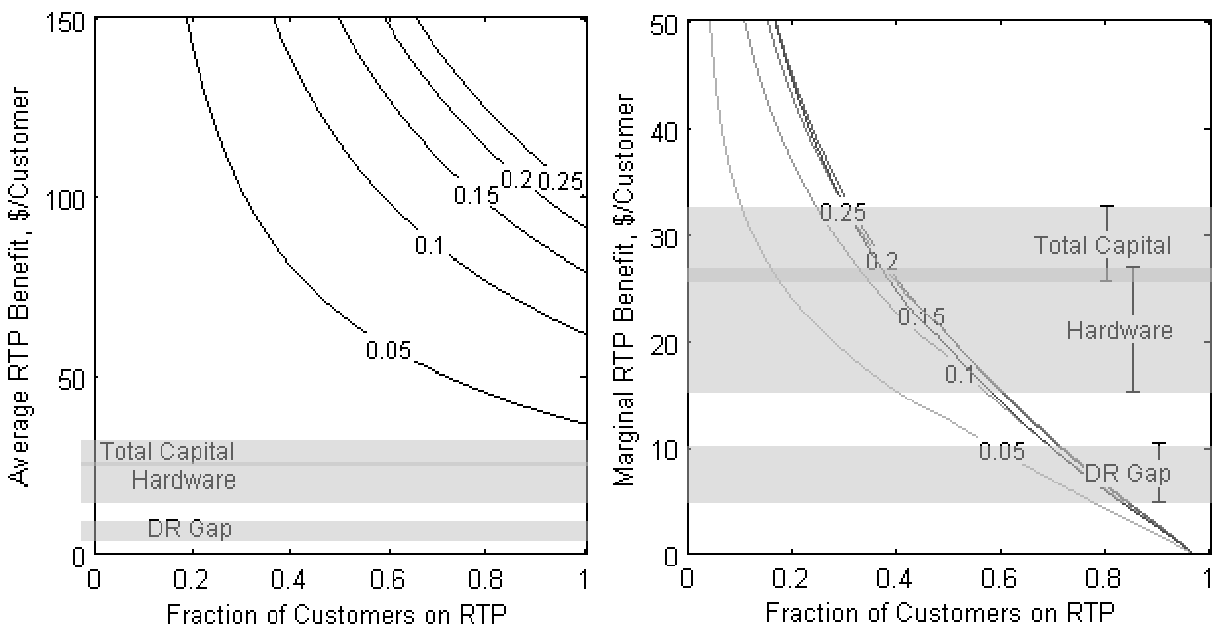


Figure D.4. Average (left) and marginal (right) total capacity and energy benefits, peak reduction value \$70/kW·y.

D.3 Results with $85/kW·y

The capacity benefit and total benefit results from Chapter 19 are displayed here if the value of peak reductions is assumed to be $85/kW·y.

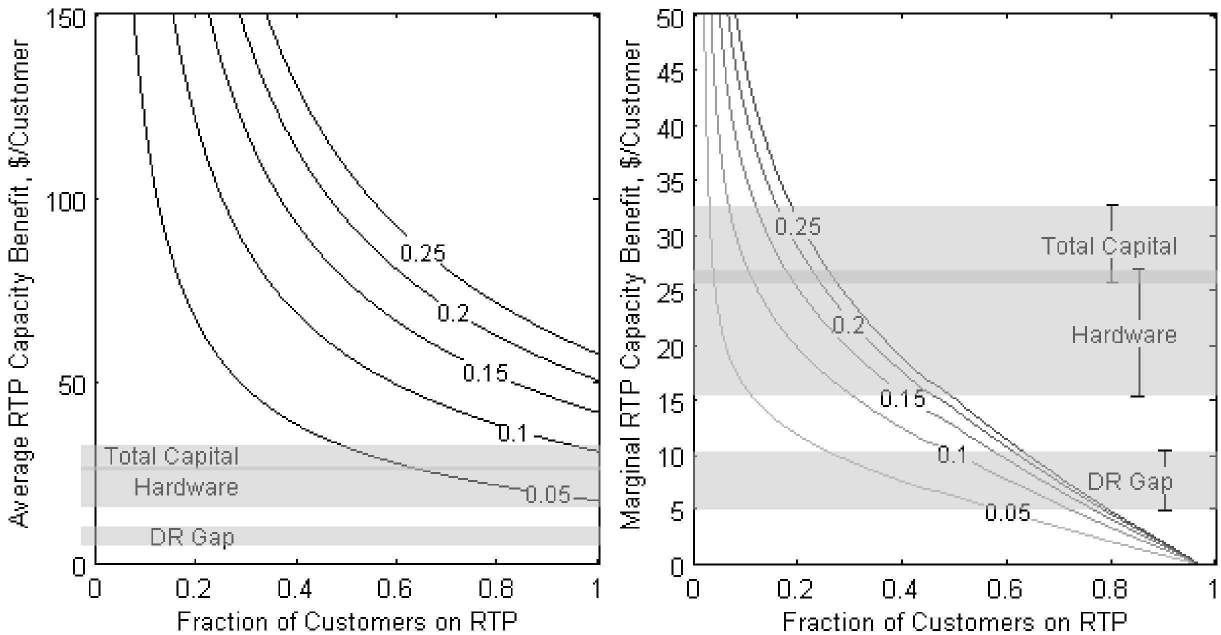


Figure D.5. Average (left) and marginal (right) capacity benefits, peak reduction value $85/kW·y.

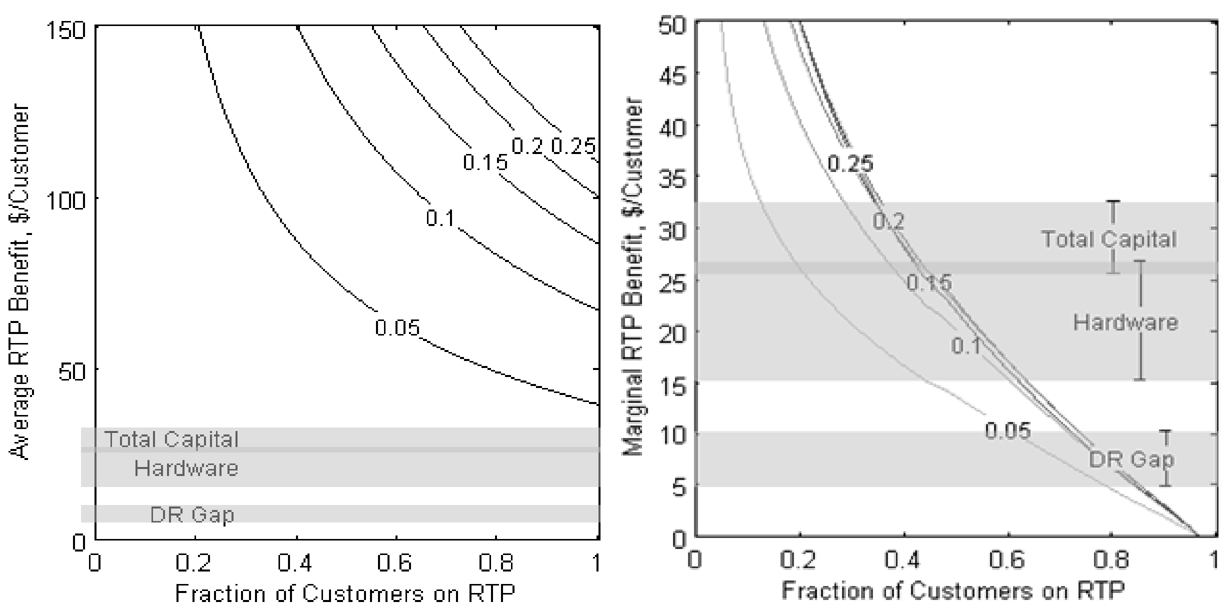


Figure D.6. Average (left) and marginal (right) total capacity and energy benefits, peak reduction value $85/kW·y.

D.4 Results with \$100/kW·y

The capacity benefit and total benefit results from Chapter 19 are displayed here if the value of peak reductions is assumed to be \$100/kW·y.

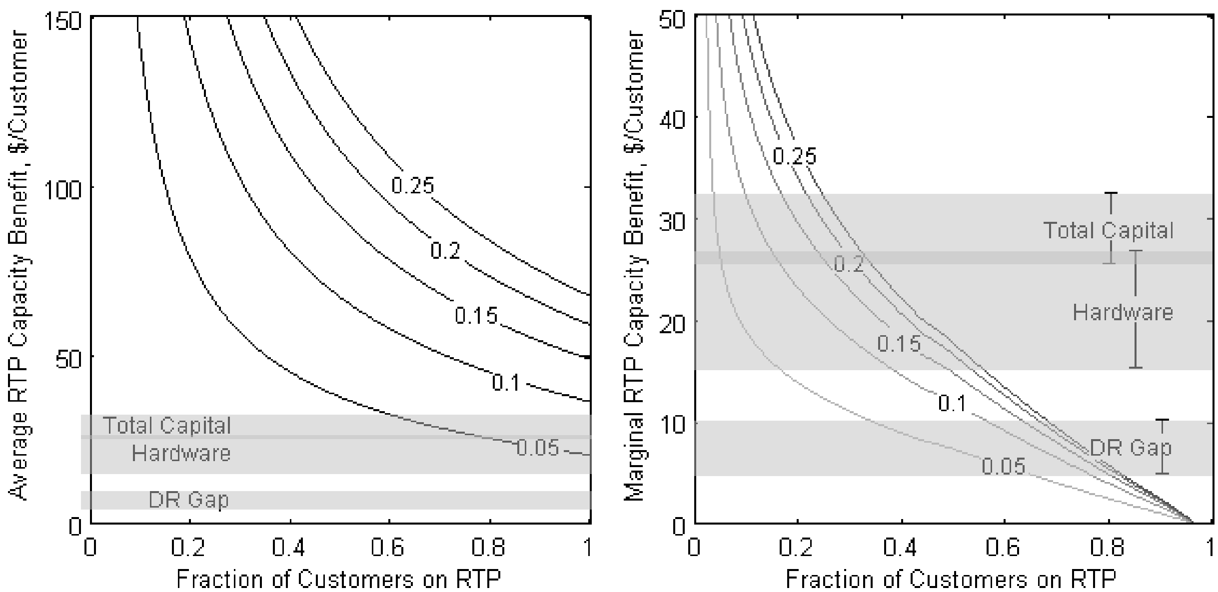


Figure D.7. Average (left) and marginal (right) capacity benefits, peak reduction value \$100/kW·y.

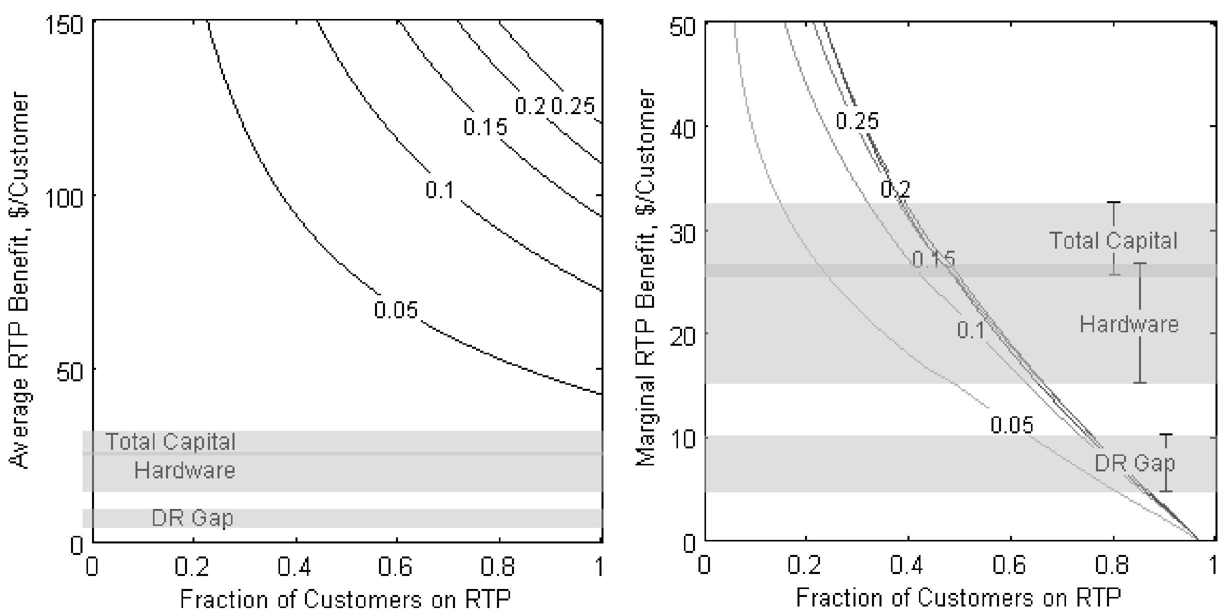


Figure D.8. Average (left) and marginal (right) total capacity and energy benefits, peak reduction value \$100/kW·y.

D.5 Results with $115/kW·y

The capacity benefit and total benefit results from Chapter 19 are displayed here if the value of peak reductions is assumed to be $115/kW·y.

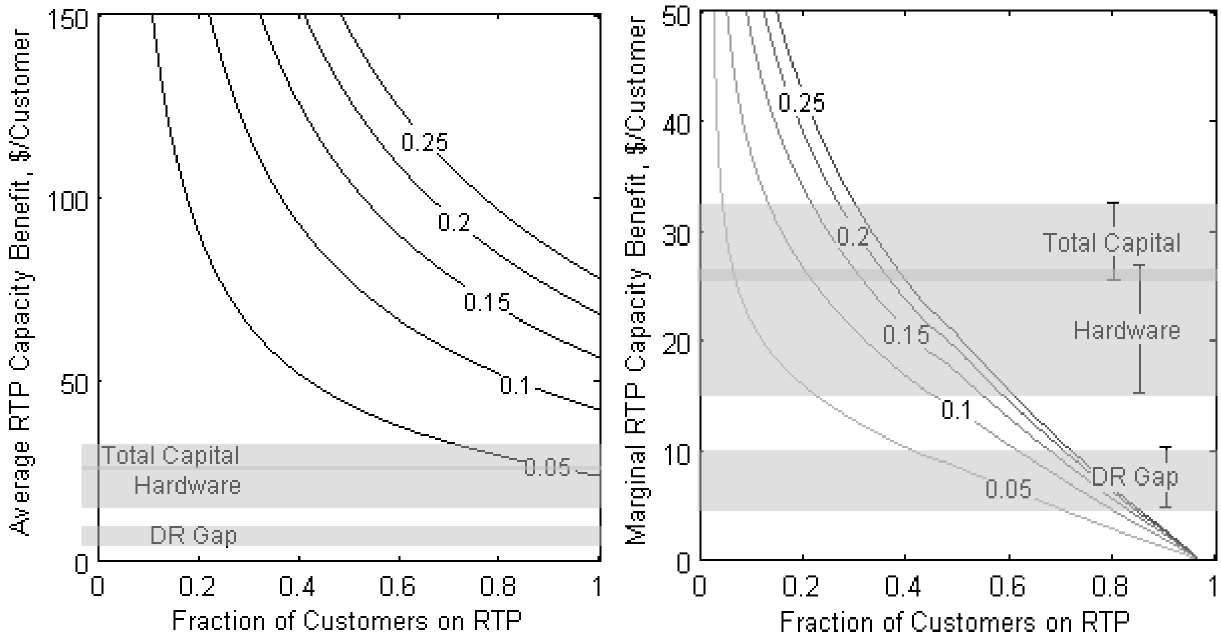


Figure D.9. Average (left) and marginal (right) capacity benefits, peak reduction value $115/kW·y.

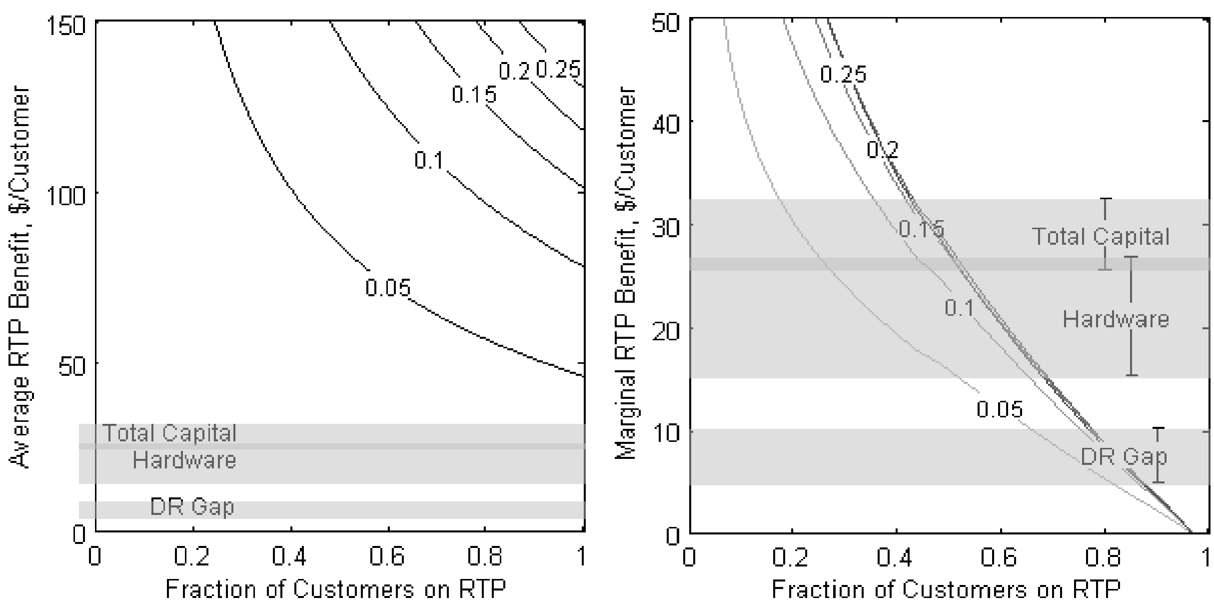


Figure D.10. Average (left) and marginal (right) total capacity and energy benefits, peak reduction $115/kW·y.