RAP Energy solutions for a changing world

Demand Response as a Power System Resource

Program Designs, Performance, and Lessons Learned in the United States

Authors

Doug Hurley

Paul Peterson

Melissa Whited

Synapse Energy Economics, Inc.

May 2013

The Regulatory Assistance Project (RAP)

The Regulatory Assistance Project (RAP) is a global, non-profit team of experts focusing on the long-term economic and environmental sustainability of the power and natural gas sectors. RAP has deep expertise in regulatory and market policies that promote economic efficiency, protect the environment, ensure system reliability, and allocate system benefits and costs fairly among all consumers.

RAP works extensively in the European Union, the US, China, and India. We have assisted governments in more than 25 nations and 50 states and provinces. In Europe, RAP maintains offices in Brussels and Berlin, with a team of more than 10 professional experts in power systems, regulation, and environmental policy. For additional information, visit the RAP website www.raponline.org.

Electronic copies of this paper and other RAP publications can be found on our website at **www.raponline.org**.

To be added to our distribution list, please send relevant contact information to **info@raponline.org**.

RAP Foreword

The nations of Europe, and the European Union as a whole are in the midst of a transition across the power sector that, at the highest level, aims to deliver the benefits of market competition and transnational market integration to customers across all of the Member States – and to do so while meeting 21st century power reliability standards, increasing the share of variable renewable generation, decreasing the climate and environmental impacts of power generation, and managing total power costs for the benefit of families and businesses that must compete in world markets. This is, all in all, a tall order.

European policymakers, power providers, and utilities are making significant progress towards these challenging goals, but challenges are also evolving and growing. Greater reliance on competitive markets exposes customers and investors to more volatile prices, and greater reliance on renewable generation will put pressure on grid transfer capabilities and will make it harder for system operators to align generation and customer demand levels in real time. For these and other reasons, it is increasingly apparent that investments on the supply-side alone will be insufficient to optimize the costs and environmental footprint of European power systems in the coming decade. *It is crucial now to examine options on the demand side of the power system,* and to design market rules and public policies that will enlist customers and their agents as power sector resource providers.

Demand-response (DR) resources can provide numerous benefits to power systems, particularly those seeking to integrate a large fraction of renewable generation, but DR is a challenging new area for most utilities and grid managers. As EU policymakers, system operators, and power providers examine options to tap DR resources, it is very useful to build on lessons from those power markets and system operators that have had some years of experience in this arena. Due to the expressed interest of European policymakers, including the electric power team at ACER, Europe's Agency for the Cooperation of Energy Regulators, the Regulatory Assistance Project commissioned this paper to provide a detailed and highly expert review of DR tools and results in several of the leading markets in North America, where DR has become a significant and valued resource. RAP itself has had substantial experience with these policies, having launched and led several of the national and regional initiatives that advanced DR as a resource in US power, reserves, and ancillary services markets.

This paper was completed by a team of highly-regarded experts from Synapse Energy Economics, who have been deeply involved in many of those measures. The scoping and drafting was prepared in close collaboration with Meg Gottstein and Mike Hogan, senior RAP policy advisors on market design, along with the invaluable project management assistance of RAP Associate Sarah Keay-Bright. We are grateful for the opportunity to provide this information in the European context, and stand ready to assist Governments, system operators, and other stakeholders on DR policies and programme designs as DR resources rise in importance in Europe's power systems.

Richard Cowart
Director of European Programmes
Regulatory Assistance Project
Brussels, May 2013

Synapse Foreword

Demand Response as a Power System Resource was prepared by Synapse Energy Economics, Inc. for the Regulatory Assistance Project. The report focuses on the ways that demand response resources effectively participate in and improve the performance of coordinated electric systems in the United States. The report reviews the many types of services that demand response can provide and the early history of demand response programs in the United States. The bulk of our research examined the specific applications of demand response in several US regions. This report includes numerous examples of demand response successfully providing reliable system services at competitive prices, and ends with lessons learned and key challenges for the near future.

Instead of attempting to translate the American experience into numerous European structures, we have tried to use well-defined, simple terms to describe how a variety of demand response resources provide reliability, energy, and ancillary services. Demand response resources have varied capabilities and services that they can provide, just as supply resources such as central station power plants do. We have minimized the use of acronyms in an industry that flourishes with them because the challenges facing the electric industry are complex and technical enough; the language itself should not be an additional barrier to communication and problem-solving.

We appreciate the information, suggestions, and insights provided by numerous members of the electric industry, colleagues, and friends that assisted us in writing this report. We are particularly grateful for comments from staff members at the Lawrence Berkeley National Laboratory and the Regulatory Assistance Project who reviewed early drafts of the report.

Doug Hurley, Melissa Whited, and Paul Peterson
Synapse Energy Economics

Table Of Contents

List of Figures

List of Tables

List of Acronyms and Abbreviations

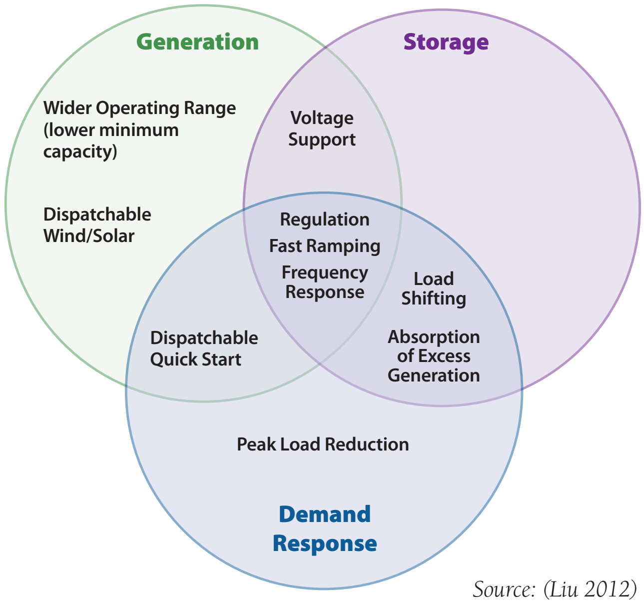
ACEEE	American Council for an Energy-Efficient Economy
BPA	Bonneville Power Administration
BRA	Base Residual Auction
CAISO	California Independent System Operator
CSP	Curtailment Service Provider
DADRP	Day-Ahead Demand Response Program
DALRP	Day-Ahead Load Response Program
DOE	Department of Energy
DR	Demand Response
DRR	Demand Response Resources
DSASP	Demand Side Ancillary Services Program
DSM	Demand Side Management
DSR	Demand Side Resources
EDR	Emergency Demand Response
EDRP	Emergency Demand Response Program
ERCOT	Electric Reliability Council of Texas
EILS	Emergency Interruptible Load Service
EPA	Environmental Protection Agency
EPACT	Energy Policy Act of 2005
ERS	Emergency Response Service
ETS	Electric Thermal Storage
FCA	Forward Capacity Auction
FCM	Forward Capacity Market
FERC	Federal Energy Regulatory Commission
HVAC	Heating/Ventilation/Air Conditioning
ICAP	Installed Capacity
ISO	Independent System Operator
ISO-NE	Independent System Operator New England
LaaR	Load Acting as a Resource
LMP	Locational Marginal Price
LMR	Load Modifying Resources
LOLE	Loss of Load Expectation
LOLH	Loss of Load Hours
LOLP	Loss of Load Probability
LR	Load Resource
LSE	Load Serving Entity
MISO	Midwest Independent System Operator
MW	Megawatt
MWh	Megawatt hour
NARUC	National Association of Regulatory Utility Commissioners
NAESB	North American Energy Standards Board
NERC	North American Electric Reliability Corporation
NYISO	New York Independent System Operator
PJM	PJM Interconnection (formerly Pennsylvania-New Jersey-Maryland Interconnection)
PRD	Price Responsive Demand
PRP	Price Response Program
RFP	Request for Proposal
RPM	Reliability Pricing Model
RTDR	Real Time Demand Response
RTEG	Real Time Emergency Generation
RTO	Regional Transmission Organization
SCR	Special Case Resources
SPP	Southwest Power Pool
SRMCP	Synchronized Reserve Market Clearing Price
T&D	Transmission and Distribution
TDU	Transmission and Distribution Utility
TPRD	Transitional Price Responsive Demand
TVA	Tennessee Valley Authority
UCAP	Unforced Capacity
VOLL	Value of Lost Load

Executive Summary

Demand response refers to the intentional modification of electricity usage by end-use customers during system imbalances or in response to market prices. While initially developed to help support electric system reliability during peak load hours, demand response resources currently provide an array of additional services that help support electric system reliability in many regions of the United States. These same resources also promote overall economic efficiency, particularly in regions that have wholesale electricity markets. Recent technical innovations have made it possible to expand the services offered by demand response and offer the potential for further improvements in the efficient, reliable delivery of electricity to end-use customers. This report reviews the performance of demand response resources in the United States; the program and market designs that support these resources; and the challenges that must be addressed in order to improve the ability of demand response to supply valuable grid services in the future.

Figure ES-1

Services Provided by Various Resources



Source: (Liu 2012)

A. Services Provided by Demand Response in the United States

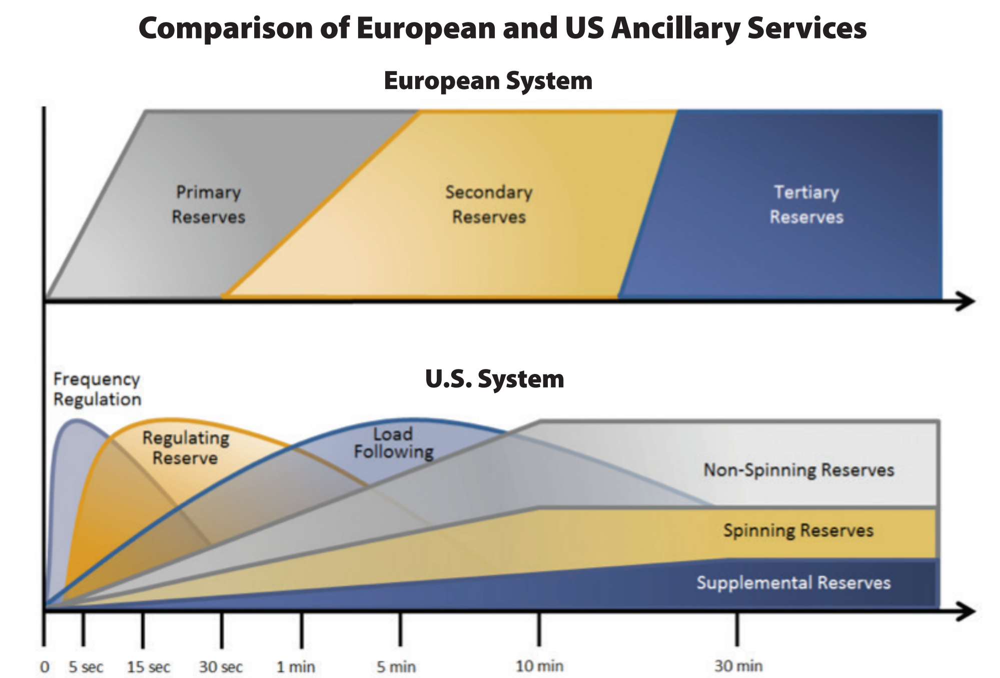
In Section 2 we provide an overview of the many diverse types of demand response resources and define the categories of demand response that we reviewed for this paper. We primarily focus on resources that are capable of being dispatched by system operators through direct controls and economic incentives (or a combination). In Section 3 we review historical demand response participation and how the programs developed by vertically integrated utilities became the precursors to today's demand response programs.

Sections 4 through 6 provide descriptions of the key services that demand response provides to the US electric system. These services range from ensuring resource adequacy and providing ancillary services to reducing high energy prices through participation in energy markets. An overview of these capabilities is shown in the illustration below.

In particular:

- We look at demand response resources that support resource adequacy through wholesale market designs in PJM, New York, and New England. These market designs treat demand response resources as rough equivalents to traditional generation resources to ensure sufficient capacity during peak-load hours. We also briefly look at non-centrally dispatched demand response resources that provide a similar capacity/resource adequacy service in California. These demand response resources are often called emergency resources, but we prefer the terms "resource adequacy" or "capacity" resources to reflect their expanded use in situations other than system emergency events.
- We look at demand response resources that provide energy reductions in the day-ahead or real-time wholesale energy markets and review the market designs currently in place. We also discuss the numerous new market designs going through the development, review, approval, and

Figure ES-2



Sources: Adapted from Ruggero Schleicher-Tappeser diagram based on (NERC 2011) and (Amprion 2013).

implementation process to comply with FERC Order No. 745.

- We look at ancillary services such as spinning and non-spinning reserves, and regulation and load-following services. A comparison of ancillary services in the United States and Europe is shown in the illustration above while a more detailed description of each service is provided in Section 6 and the appendix. Texas has proven demand response programs that act as spinning and non-spinning reserves to provide resources during system emergencies. Other regions are exploring innovative ways that demand response can provide a variety of ancillary services.

B. Lessons Learned from Existing Programs

In Section 7 we discuss lessons learned from demand response in the United States. These lessons include the following:

1. Demand response can provide capacity/resource adequacy services for system peak load days that are equivalent, and perhaps superior, to services provided by traditional resources. PJM and New England have existing market designs that have successfully incorporated significant amounts of demand response into their capacity markets by enabling demand response to compete with generation on a relatively level playing field. The Midwest and New York also have robust demand response participation in their capacity programs. In these designs, demand response resources have demonstrated reliable performance and provided substantial contributions to system resource adequacy goals. The table below demonstrates this by comparing the megawatts of demand response delivered relative to the quantity that had an obligation to be available, as well as the performance of traditional resources.
2. Demand response can provide energy services that primarily enhance efficient price formation in wholesale energy markets, but also enhance reliable operation of the system. Market designs are still being developed pursuant to Commission Order No. 745 to ensure that pricing and verification mechanisms are optimal. We focus largely on PJM and New England program and market designs for demand response resources that can provide energy services. The figure

Table ES-3

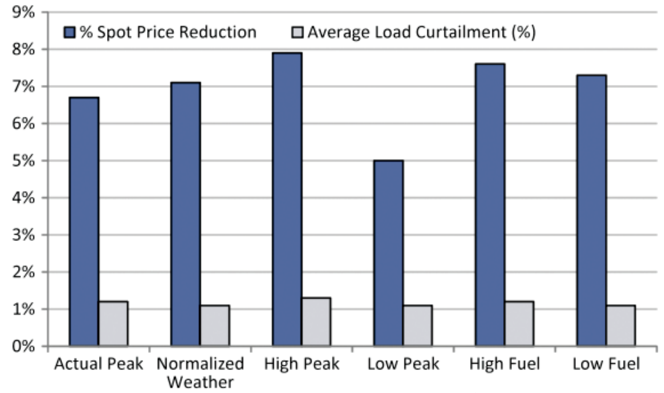
New England Capacity Resource Performance

	Real Time Demand Response		Real Time Emergency Generation	
Load Zone	MW	Performance	MW	Performance
Maine	278	100%	25	100%
New Hampshire	45	93%	33	98%
Vermont	33	100%	13	98%
Connecticut	261	72%	254	86%
Rhode Island	40	90%	56	88%
Southeastern Mass.	136	78%	37	86%
Western/Central Mass.	132	97%	577	94%
Northeastern Mass.	198	80%	78	89%
Total New England	1,124	86%	553	90%
Generation Fleet Average EFORd	94.5%			
Quick Start Generation	Assumed 80% during planning			

Source: (Scibelli 2012)

Figure ES-4

Percent Load Curtailment and Simulated Corresponding Price Reduction in Mid-Atlantic States



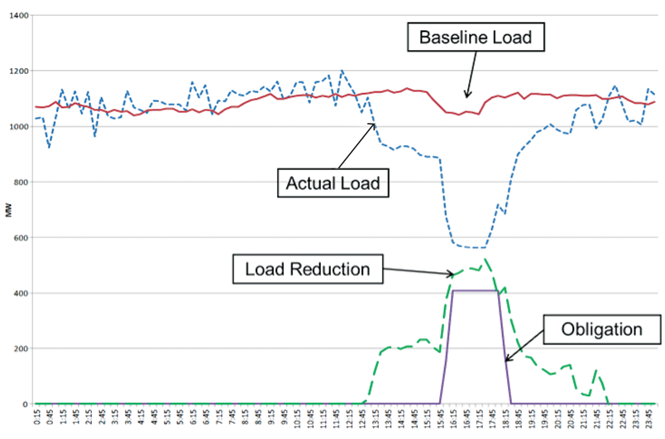
Source: (Brattle Group 2007)

above shows the potential reductions to PJM hourly clearing prices from demand response, based on a study by the Brattle Group.

3. Demand response can provide ancillary services that include various reserve services, dynamic system regulation, and load-following capabilities that can deliver value to the grid during any hour of the year. The Texas ten-minute reserves programs for demand response resources have demonstrated reliable performance over numerous system operator dispatch events (see Figure ES-5 below). Current program and market designs for demand response participation provide excellent examples of this type of demand response service. Additionally, pilot programs in several regions and new technology adaptations are suggesting

Figure ES-5

Demand Response Deployment of Emergency Reserves in Texas on August 4, 2011



Source: (ERCOT 2012)

a more robust role for demand response resources in providing balancing services to correct system imbalances on a minute-by-minute or even second-by-second basis.

C. New Opportunities and Challenges for Demand Response

Opportunities and challenges for demand response are also explored in Section 7. Future expansion of demand response participation in wholesale markets includes several promising areas:

- Coordination of demand response resources by system operators to more accurately match resource needs with system conditions. Rather than relying on individual suppliers or distribution utilities to activate demand response resources, regional system operators could provide more reliable and efficient implementation that utilizes demand response resources across multiple utility areas.
- Deployment of technologies that enhance the ability of system operators to integrate new forms of demand response into normal system operations during any hour, rather than just peak demand periods. These new resources include price-responsive demand that is enabled by advanced meters and demand response resources with a storage component such as water pumping and space heating that can increase demand during periods of excess generation. Such forms of demand response can reduce costs and enhance system efficiency during any hour of the year.
- Enhanced or new market designs that provide compensation for a variety of demand response (and other) resources to meet future system needs. These could be changes to existing capacity programs or separate programs for specific demand response resources that supply regulation or balancing services.

Key challenges include:

- Establishing baselines to accurately measure demand response contributions to ensure appropriate performance and compensation.
- Resolving dispatch issues between traditional distribution utilities and regional system operators or balancing authorities. As the proportion of variable generation resources increases, demand response can play a critical role by providing more flexible options to system operators to ensure reliability and improve overall system efficiency.

In order to achieve that greater flexibility, system operators must be able to directly dispatch demand response resources located across multiple utility distribution areas.

- Integrating sufficient metering and communications equipment to provide accurate and timely information about overall electric system performance as well as specific demand response resource operations, while ensuring requirements are not cost prohibitive for smaller resources.
- Modifying current market and operational rules to remove numerous market barriers to demand response participation, such as minimum size requirements and prohibitions regarding demand response aggregator participation.
- Addressing the lack of demand response participation in certain regions that is often related to traditional utility incentive structures that do not reward utilities for incorporating demand response. In some regions, third-party aggregators are obstructed from enrolling demand response resources due to utility opposition and regulator concerns about consumer impacts and benefits.

D. Summary and Conclusions

Demand response provides a variety of valuable services to the US electric grid and has the potential to enhance system efficiency to an even greater extent in the future. To that end, we offer the following conclusions and recommendations for facilitating optimal demand response participation going forward:

Provide greater revenue certainty. The growth of demand response has been strongest where a steady monthly payment exists, and where multiple streams of revenue are present to support different types of loads and different types of customers. Trying to rely on unpredictable and infrequent high-priced events is a business model that is too risky to incentivize significant demand response participation. Mechanisms to provide greater revenue security could include targeted instruments such as procurement of specific critical services, or less-targeted instruments such as forward capacity markets.

Remove restrictions on demand response participation. Regions that do not limit demand response resources and allow demand response to provide multiple types of services (energy, reserves, regulation) have demonstrated greater participation by demand response resources.

Ensure demand response providers face adequate incentives. Independent demand response aggregators have a greater financial incentive to sign up as many customers with load reduction capabilities as possible. Utility providers can provide reliable demand response, but often have conflicting financial incentives. Shifting traditional utility compensation to a more nuanced compensation system will help align incentives towards a more efficient overall use of resources.

The residential market remains largely untapped for now. Few demand response providers have even approached the residential market to date due to the amount and variety of load available from large customers. However, cost-effective technology to provide small amounts of demand response from a very large number of residential customers is not far away, and may lead to widespread implementation by the end of the decade.

Demand response, with sufficient compensation, can help integrate variable resources. The ability of storage-type demand response that can ramp in both directions to both reduce load and also absorb excess generation is a new and developing area of demand response that has proven reliable for balancing services (regulation and load-following) at a small scale. However, it remains to be seen whether the revenue generated through such services is sufficient to sustain demand response providers. A re-determination of the value of balancing services under resource mixes with greater quantities of variable resources (such as wind and solar) may need to occur. In addition, compensation that rewards the speed and accuracy of the response will help incentivize demand resources to participate in this market.

Regulatory support is necessary to level the playing field. Finally, regulatory support at both the state and federal level has been critical. In order for demand response to flourish, the necessary policy and regulatory framework first had to be established to govern the treatment of demand response and enable it to be compensated in a manner comparable to generation resources. The consistent attention to these issues by the Federal Energy Regulatory Commission has proven essential to the success of demand response resources to date.

1. Introduction

Demand response (DR) encompasses numerous types of load-modifying resources that provide a variety of electric system functions. Over the last few decades, the original utility programs that were developed primarily to provide load reductions during system emergencies have evolved into more sophisticated programs capable of providing a range of targeted services. Demand response has transitioned from simply a means for shaving peak demand into a valuable tool enabling grid operators to manage the challenges of the modern grid. Demand response program structures have likewise expanded from simple incentives that enable a utility to temporarily interrupt consumption to more sophisticated market arrangements, including three-year forward commitments to provide a guaranteed level of energy reduction based on a central operator dispatch signal, and balancing services that may employ storage to better integrate renewable resources.[1]

Significant effort has been invested in current demand response programs and market constructs to ensure the development of appropriate incentives, regulations, and technologies. There have been bumps in the road; some programs have worked better than others. Despite being originally viewed solely as a seasonal peak load reducer, demand response has demonstrated that it can provide cost-effective, year-round reliability services, daily energy services, and ancillary services that include reserves, load-following, and regulation. Programs and markets are likewise evolving to incorporate these diverse functions and encourage greater demand response participation.

This paper provides a cursory review of earlier demand response programs, a detailed look at recent programs, and summarizes the lessons learned, the most promising future applications, and key challenges facing demand response in the United States. The report is organized in sections:

Section 2 covers background issues related to demand response resources. This includes defining the different types of demand response resources and the services they can provide; the benefits that demand response can provide to bulk power system operations; and the specific characteristics of demand response resources reviewed in this report.

Section 3 covers the history of demand response initiatives in the context of the transition from vertically integrated utilities to more competitive regional markets operated by system operators who provide coordinated dispatch and exercise day-to-day operational control.

Sections 4 through 6 review current programs developed and implemented by numerous entities across the United States. We review program designs, the level of participation by demand response resources, and the performance of demand response resources during specific events. The reviews are organized into three broad categories: capacity/resource adequacy, energy markets, and ancillary services.

Section 7 describes the lessons learned, near-term opportunities for demand response, and key challenges to expanded participation by demand response resources.

The Terminology Appendix provides definitions of terms used in this report, and an acronym list is provided prior to the Executive Summary.

1 In this report we use "balancing services" to mean regulation and load-following services that are employed under normal system conditions to correct supply and demand imbalances within seconds or minutes.

2. Background

A. What is Demand Response?

Electricity demand varies significantly by time of day as well as by season. Historically, the balancing of electricity supply and demand was performed only by increasing or decreasing the electrical output of power plants, but this often requires large investments in capital-intensive facilities that are infrequently used, or the dispatch of increasingly inefficient (and therefore expensive) generators.

Demand response was originally developed by electric utilities in order to increase flexibility on the demand side by temporarily shifting or reducing peak energy demand, thereby avoiding costly energy procurements and capacity investments for a small number of hours of need. With the shift toward competitive electricity markets, demand response has become an important tool used by many utilities and system operators in the United States to enhance grid reliability and market outcomes.

The Federal Energy Regulatory Commission (FERC) currently defines demand response as:

> *Changes in electric usage by demand-side resources from their normal consumption patterns in response to changes in the price of electricity over time, or to incentive payments designed to induce lower electricity use at times of high wholesale market prices or when system reliability is jeopardized.* (FERC 2012)

As variable resources such as wind and solar provide an increasing proportion of electricity to the grid, new forms of demand response are being developed with capabilities that surpass traditional peak load-reducing demand response.[2] This next generation of demand response is automated and often linked to some form of energy storage in order to quickly respond to changes in system frequency or to increase demand during periods of oversupply, which improves the utilization of many renewable resources, as well as traditional thermal units.

In recognition of this expanded role, FERC's 2010 *National Action Plan for Demand Response* emphasizes that demand response includes actions that can change any part of a customer's load profile, not just the period of peak demand. This definition specifically includes "the smart integration of changeable consumption with variable generation," such as through energy storage (using devices such as electric vehicle batteries and thermal storage), and the associated provision of ancillary services such as regulation and reserves (FERC 2010).

The benefits offered by demand response are numerous, but they fall into three general categories: economic efficiency, system reliability, and environmental benefits. The economic benefits consist primarily of lower wholesale market prices due to demand response's ability to displace the most expensive peak generation resources, as well as the deferment or avoidance of more costly new capacity construction by flattening the demand curve.[3] The flexibility of demand is key to ensuring wholesale market efficiency; enhancing the elasticity of demand more accurately reflects consumers' willingness to pay and mitigates the ability for suppliers to exercise market power. Additional economic benefits may include

2 "Variable" as used in this paper refers to any source of electricity production where the availability to produce electricity is largely beyond the direct control of the operator. It can be simply variable—changing production independently of changes in demand—or variable and uncertain. Another term for this latter category is "intermittent."

3 A 2007 report found that in the PJM regional electricity market, a three percent load reduction in the 100 highest peak hours corresponds to a price decrease of six to twelve percent. (Brattle Group 2007)

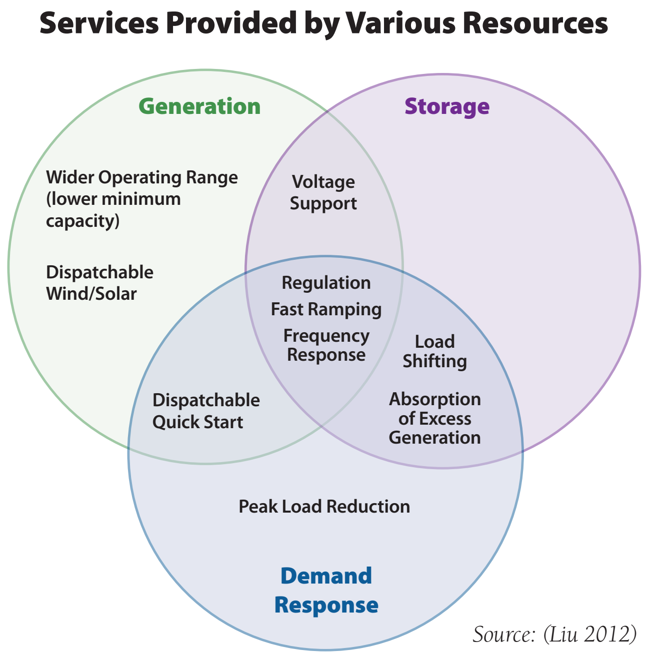
participant financial benefits and risk management.[4]

Demand resources may also be called upon by system operators to maintain the reliability of the electric system in the event of an emergency and avoid brownouts or blackouts. In addition to reducing capacity constraints, some demand resources can be used to provide ancillary services such as reserves or balancing by quickly increasing or decreasing demand. System stability is thus improved through better aligning the movement of generation supply and electricity demand. The services provided by various resources – from generation to storage to demand response – are depicted in Figure 1. The figure demonstrates the degree to which demand response can provide both unique services (i.e., peak load reduction), as well as some of the services offered by traditional generation and storage resources.

The environmental benefits of demand response resources will vary from region to region. Environmental benefits are mostly associated with the displacement of marginal fossil fuel resources. Each region of the country has a different generation resource mix with various percentages of fossil resources (coal, oil, and natural gas), nuclear, hydroelectricity, renewables, etc. During different seasons of the year and depending on the time of day, the displacement of marginal emissions due to the dispatch of demand response resources can vary substantially.[5]

Additionally, environmental benefits may result from demand response's ability to facilitate the integration of renewable resources. The flexibility of demand response allows the electric system to accommodate higher penetrations of variable resources such as wind and solar, whose energy output can fluctuate quickly and lead to excess generation supply on the system. Demand resources that include a form of energy storage are particularly well-equipped to facilitate the management of periods of over-

Figure 1



Source: (Liu 2012)

supply through the provision of load-following or regulation services that can both increase or decrease demand.

Demand response's load modifying capability thus enables more efficient use of current electricity generation resources, while yielding economic, reliability, and environmental benefits. Yet demand response is not a homogenous resource; it is provided by a highly diverse set of actors in numerous different ways, and with varying capabilities. This diversity precludes any simple characterization of demand response types and also contributes to the flexibility of demand response to meet multiple system needs. An overview of the various forms of demand response is given in the following section.

4 Participant financial benefits consist of the bill savings and incentive payments that customers receive in return for curtailing, shifting, or otherwise modifying their load. Risk management benefits are related to the ability of demand response investments to diversify generation portfolios and avoid large capital investments in new power plants that may experience shocks in terms of fuel costs, construction costs, or future environmental regulations (Binz, et al. 2012).Through its ability to increase the elasticity of demand, demand response also deters generators from exercising market power (US Department of Energy 2006).

5 We note, however, that the use of distributed back-up generation from fossil fuels (e.g., diesel fuel or natural gas) as "demand response" can reduce demand response's environmental benefits. For this reason, some regions have restricted the use of fossil-fuel powered back-up generators that may qualify to provide demand response.

B. Demand Response Provision and Classification

Demand response can be provided by all categories of customers (industrial, commercial, and residential) employing many different technologies or strategies to achieve shifts in demand. Common examples include:

- Reducing or interrupting consumption temporarily with no change in consumption in other periods
- Shifting consumption to other time periods
- Temporarily utilizing onsite generation in place of energy from the grid[6]

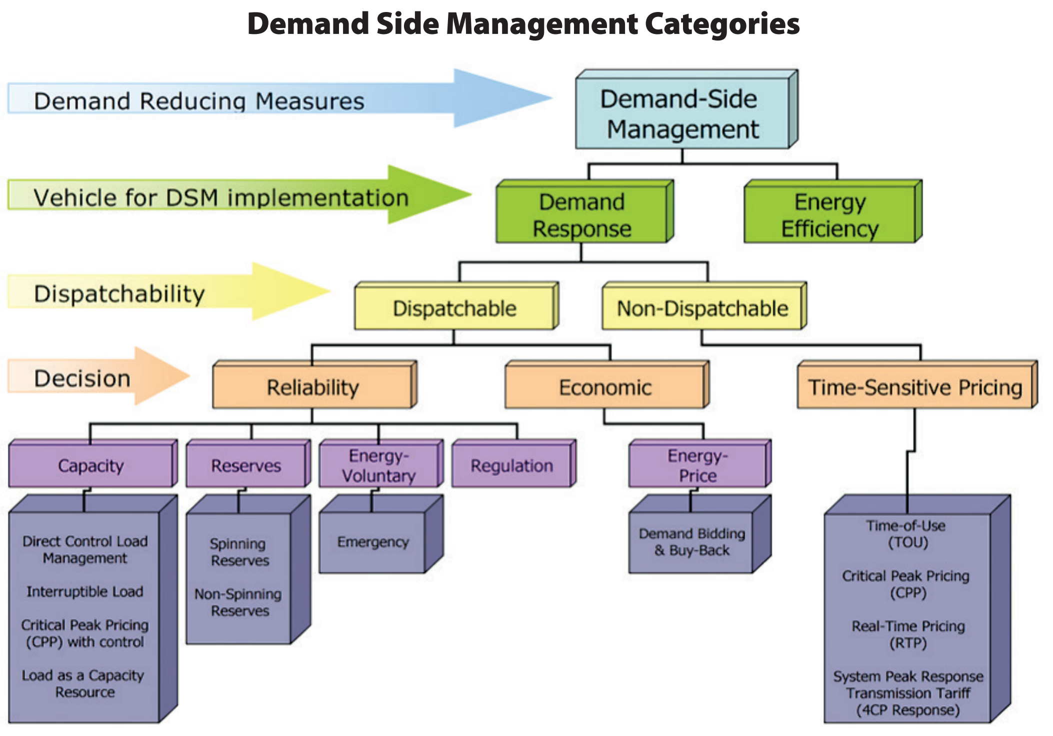
In addition, demand response can provide frequency regulation and load-following services. During periods of excess energy production, demand response resources that have an element of storage may increase the energy used for heating or pumping water, charging batteries, compressing air, or freezing ice for cold storage. The rate at which these activities occur can be automatically adjusted to align consumption with generation output.

Demand response resources interface with wholesale markets in two distinct ways: either as resources that are dispatched by a system operator, or as non-dispatchable resources that may elect (voluntarily) to adjust their consumption based on price signals. Dispatchable resources typically bid directly into wholesale markets or enter into contracts to receive payments for demand reduction, whether in response to a reliability event or high market prices.

In contrast, non-dispatchable resources generally participate in price-based demand response programs such as real-time pricing, critical peak pricing, and time-of-use tariffs. These price-based programs provide users with ongoing price signals to encourage lower energy consumption during periods of high electricity prices, but are generally not considered "firm" resources, as they are not dispatchable and grid operators do not know the degree to which customers will respond. These non-dispatchable price-based programs may become more prevalent for residential and small commercial customers as "smart meters" are deployed. Advanced meters also create opportunities for aggregating residential and small commercial customers in ways to provide dispatchable services to system operators, which may be compensated through either contract prices or market prices. The various types of demand side resources available, spanning both demand response and energy efficiency, are depicted in Figure 2, below.

For this report, we focus primarily on dispatchable demand response resources that may be dispatched

Figure 2



Source: (NERC 2011)

6 As noted previously, the environmental benefits of demand response may be reduced by fossil-fueled back-up generation. In this report we refer to demand response generally, irrespective of the means employed to shift or reduce load, in part due to the lack of data regarding the amount of demand response provided by back-up generation.

for the variety of purposes described above, including: ensuring resource adequacy (capacity), providing other reliability functions such as reserves and balancing services, and responding in energy markets to high market prices. In general, this report examines the ways that demand response can provide specific services at the control and discretion of system operators to improve the overall performance and stability of electric power systems.

Demand response participates in all regions of the United States. The participation occurs across a full continuum of structures from integrated, centrally managed, mandatory wholesale markets at one extreme to vertically integrated utility areas that have voluntary balancing services at the other extreme.

One key distinction along this continuum of market participation is the ability of fully-integrated demand response to set the market clearing price. Any reduction in load will reduce the overall cost of serving electricity during that timeframe, but fully integrated demand response in areas with wholesale markets can have a much larger price impact due to its ability to reduce the market clearing price for all market participants. There is further discussion of this distinction in section 6.D.

In regions where demand response does not participate in wholesale markets (whether or not these markets exist), demand response may be carried out by distribution utilities to avoid high peak energy costs, ensure system reliability, or provide balancing services.

Infrastructure Requirements

Irrespective of the market structure in which dispatchable demand response operates, to receive compensation, these resources must comply with dispatch signals from the system operator and changes in demand must be measured and verified. Measurement and verification typically requires a certain level of metering accuracy and telemetry infrastructure investment.[7] In particular, programs that offer incentives for participation (and/or penalties for non-compliance) must calculate the baseline load and measure the change from this baseline that occurs during a demand response event in order to calculate the total change in demand.

Meter requirements for dispatchable demand response may include 15-minute or five-minute interval meters, while telemetry requirements vary. Telemetry is generally added to ensure stable operation of the network, and, depending on the size of the demand resource and type of service offered, telemetry requirements may range from after-the-fact metering to four-second real time telemetry equipment to enable system operators to monitor loads and ensure that the contracted change in demand is met (Isser 2008). In order to set the real-time market clearing price, demand resources are typically required to have sufficient telemetry and capability to receive a system operator dispatch. However, for small resources, advanced metering and telemetry requirements can be prohibitively expensive and may not be necessary (Pfeifenberger and Hajos 2011).[8]

Telemetry is required for regulation by all system operators, but not all system operators require it for spinning reserves. In some regions, data granularity is two seconds for telemetry, but it can be batch sent once every minute for demand response. This requirement is easier to accommodate, particularly for demand response aggregators. As metering accuracy increases, so do costs. Some system operators also require that data are reported from each individual resource, while others only require that data be available and verifiable in aggregate form (MacDonald, et al. 2012).

A key element in enabling demand response to provide these various services has been the development of appropriate market regulations. Through a series of recent orders, FERC has established wholesale market rules that facilitate demand response resources' ability to participate actively in markets to provide energy, capacity, or ancillary services in a manner similar to generation resources. These regulations and case studies of demand response utilization will be explored in greater detail in sections 4 through 6.

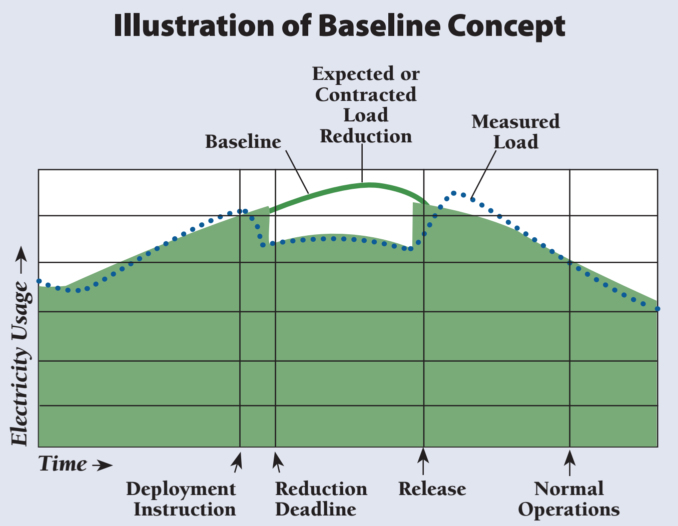
7 A notable early failure of telemetry was with electric water heater control programs that relied on a timing clock attached to the individual water heater. The timing device that switched the electric water heater on and off was affected by local power outages and the timers could only be reset by the local utility. The result was that many of the electric water heater timers were not properly synchronized and some of performed opposite to the program design.

8 For example, the Midwest Independent System Operator (MISO) initially required real-time telemetry for demand response resource participation in all types of ancillary services, but later found this to be unnecessary for provision of reliable spinning and non-spinning reserves (Pfeifenberger and Hajos 2011).

Customer Baselines

The amount of electric service provided by generation – whether central station power plants or distributed – can be measured through metering the actual energy output from the plant over a certain time interval. In contrast, demand reduction at a customer facility is a change in usage pattern and cannot be as easily calculated. Underlying the performance of demand response is the concept of a "baseline" – the amount of energy the customer would have consumed absent a dispatch signal from the system operator.

Figure 3



Source: (NERC 2011)

Figure 3 illustrates the basic idea behind baselines. On a day without a dispatch signal, the customer would have used electricity over time along the green line labeled "baseline." Given the dispatch signal, the customer will reduce their load during a few hours, by the amount of the gap between the baseline and the dotted line labeled "Measured Load." As indicated above, knowledge of load over time requires time-interval meters and a method for recording their output for reporting to the grid operator or other authority.

The trouble, of course, is that few customers have a load profile that is so regular that their baseline is simple to calculate. While retail outlets may have a predictable daily and weekly usage pattern, their load often shifts significantly with holidays. Primary and secondary schools have a drastically different usage during summer and school holidays. Factories often undergo shutdowns for routine maintenance, and order requests from their customers may fluctuate over time. As such, different regions have employed differing methods for calculating baselines. Numerous studies have analyzed diverse approaches to ensuring that customer baselines are accurate and remain accurate over time, under various dispatch scenarios.

An example will be helpful. At one time, the Independent System Operator of New England had a baseline methodology that assumed that all customers would respond on only a very small number of days. As such, the method for setting baselines required that a new customer report interval load data to the system operator for ten days before being declared fully ready to respond. Once this initial period ended, the baseline was always set by the most recent ten days of load, which was accurate enough to capture weekly load patterns, and seasonal differences. A small number of customers, however, found a loophole. These customers turned off on-site distributed generation units for maintenance during the initial reference period. During this time, the facilities were pulling the entire electric load needed to operate from the grid, and their baselines reflected this level of usage. Once the initial ten-day period ended, they restarted the onsite generation, and the load pulled from the grid dropped significantly, every day. They offered and cleared their resources as demand response every day, and the baseline was not reset for many months. They were paid for apparent reduction of load that was not a change from their normal usage. Although they appear to have been acting within the rules of the program, all outside entities aware of the situation - including the FERC Office of Enforcement - view this activity as a violation of the spirit of demand response. The baseline methodology for New England has since been changed such that no baseline can remain at a set level for more than seven consecutive business days.

Any method for calculating baselines must meet two basic criteria. First, it must be one that is trusted by the grid operator to be reliable. The entity that dispatches the demand response resources must be confident that they are receiving the change in load that they expect and that is being reported to them. This is true both for the reliable operation of the electric system and for the accurate settlement of payments for the electric service provided, whether it be energy, capacity, reserves, or other service. Second, it must be a method with which the customer and/or the DR provider can comply, both financially and feasibly. Reporting interval meter data to the regional operator days after a demand response event is reasonably inexpensive and easy, and is often sufficiently accurate and timely. On the other hand, sub-second, real-time metering and communication equipment may be feasible but not financially viable, especially for large aggregations of smaller customers.

Without feasible, trustworthy baselines, demand response will not succeed.

3. History

A. Demand Response as a Reliability Resource Prior to Restructuring

Demand response in the United States originated in the 1970s, in part due to the spread of central air conditioning, which resulted in declining load factors and needle peaks during hot summer days.[9] The advent of "integrated resource planning" in the late 1970s and 1980s drew attention to the high system costs of meeting these peak loads and encouraged utilities to look for load management alternatives (Cappers, Goldman, and Kathan 2009).[10] Rate design (particularly time-of-use pricing) and incentive programs became standard demand response programs at many regulated utilities.

Incentive programs such as direct load control and interruptible/curtailable programs allow a utility to curtail a portion of a customer's load in exchange for a monetary incentive, such as a credit on the customer's monthly bill or a lower overall electricity rate. Direct load control programs involve the installation of control technologies on a customer's appliance – typically an air conditioner, water heater, or pool pump – and are primarily offered to residential consumers. Interruptible/curtailable programs target large industrial and commercial consumers who have the capability of completely shutting down their operations or reducing their demand by a predetermined amount upon notice by the utility or system operator. In exchange, these large commercial and industrial customers often receive lower electricity rates.

Although interruptible/curtailable rate programs were popular during the 1980s and 1990s, in many cases customers were rarely called upon to reduce their load. For example, Southern California Edison had the largest interruptible load program in California, but did not invoke a single interruption for fourteen years until June 2000, despite customers receiving bill reductions of approximately 15 percent for participating (Marnay, et al. 2001). In Vermont, many ski areas received reduced electric rates in exchange for the ability to interrupt snow-making equipment during times of peak winter loads. A major issue arose with one ski area when an unusual weekend peak load coincided with a holiday weekend and the ski area was unwilling to shut down its snow-making equipment during one of its busiest ski weekends. The low utilization of these programs resulted in unofficial economic incentives for large customers and a limited ability or willingness to curtail load by participants when called more frequently, as described more below (Fryer, et al. 2002).

Until the late 1990s, the US electric industry consisted primarily of vertically integrated utilities that managed their own generation and distribution assets. The demand response programs of the 1970s through much of the 1990s were largely conducted by such utilities in a structured, regulated environment, and therefore consumers were not exposed to real-time wholesale price signals, nor were consumers compensated for the full system value of their demand reduction. This began to change in the 1990s as the US electric industry initiated the restructuring process.

B. Early Wholesale Market Programs

The US electric industry began to shift toward greater competition in the 1990s following the Energy Policy Act of 1992 that allowed independent power generators to participate in wholesale markets and FERC Order 888 that mandated open access to transmission systems. Wholesale markets themselves had undergone a transition from cost-of-service principles to greater competition in the late 1980s when FERC began to grant

9 Some northern states also experienced sharp peaks during winter cold snaps, particularly in regions with high penetrations of electric space heat.

10 Integrated resource planning refers to the evaluation of demand and supply resources by public utilities and state regulatory commissions to cost-effectively provide electricity service. Integrated resource planning differs from earlier planning techniques in that it also considers environmental factors, demand-side alternatives, and risks posed by different investment portfolios (Hirst and Goldman 1991).

wholesale power producers the ability to sell at "market-based rates" based on the dynamics of supply and demand (Joskow 2001).[11]

As the process of restructuring gathered steam in the late 1990s, many states elected to experiment with competitive markets, transforming their vertically-integrated utilities into stand-alone generation companies, regulated distribution companies, and regional grid operators. The regional grid operators are referred to as either independent system operators (ISOs) or regional transmission organizations (RTOs), and exist in many regions of the United States and Canada, as shown in Figure 4.

Regional system operators are responsible for managing the wholesale power markets (including real-time and day-ahead energy markets, ancillary services markets, and, in some cases, capacity markets). These wholesale power markets represent approximately two-thirds of US electricity demand (Market Committee of the ISO/RTO Council 2007).

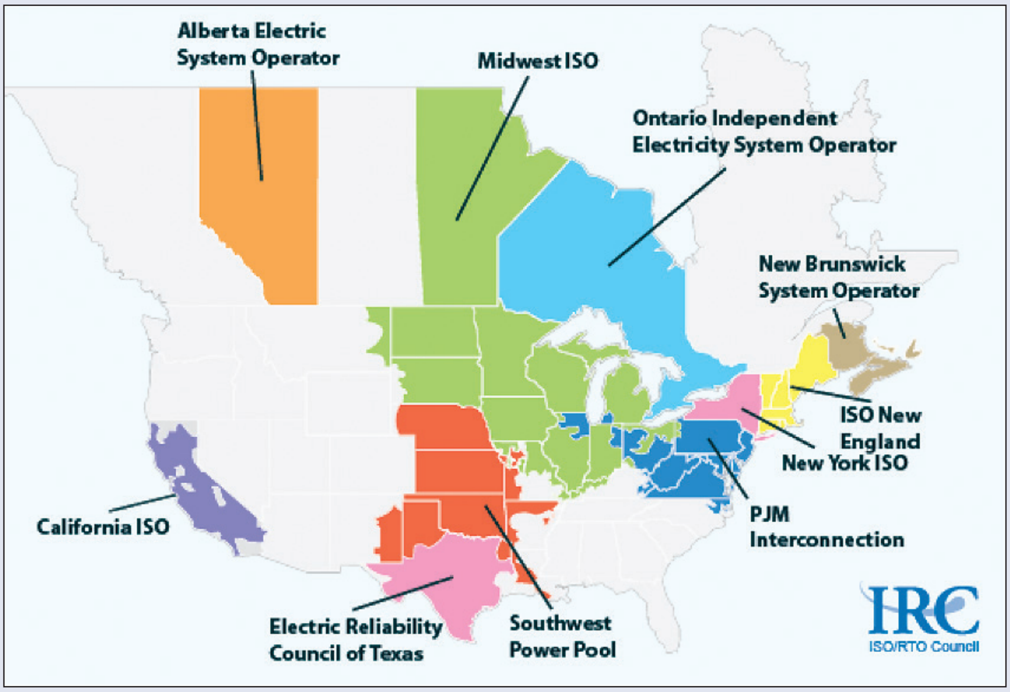
ISO-New England (ISO-NE) operates the power market in the Northeast; New York ISO (NYISO) manages the markets in New York State; PJM Interconnection (PJM) is responsible for the markets that cover the Mid-Atlantic states and parts of Ohio, Indiana, and northern Illinois; the Midwest ISO (MISO) operates the power markets that span much of the Midwest and northern Great Plains; the Southwest Power Pool (SPP) operates an energy imbalance service market in the Central Plains states; California ISO (CAISO) covers the majority of California; and the Texas market is managed by the Electric Reliability Council of Texas (ERCOT). ERCOT is the only ISO region that is not subject to FERC regulation. ISOs and RTOs also oversee and operate the high-voltage bulk power system, coordinate electricity generation, and conduct long-term regional planning (US Department of Energy 2006).

Other regions – including the Southeast, Southwest, Inter-mountain West, and Pacific Northwest – chose to retain the traditional vertically-integrated utility model. Balancing authorities operate in these areas to maintain the minute-to-minute balance between electricity supply and demand within their borders. Many utilities and/or balancing authorities in these regions operate demand response programs, such as the Bonneville Power Administration in the Pacific Northwest and the Tennessee Valley Authority in the Southeast.

While the process of restructuring increased competition and established regional wholesale markets, it also shifted responsibility for maintaining grid reliability away from utilities to system operators, thus reducing incentives for traditional utility-run demand side management programs. As utilities divested their generation assets, many no longer saw value in maintaining such programs to ensure reliable and efficient grid operations (Electric Energy Market Competition Task Force 2007). Demand side management spending peaked in 1993 at approximately $2.7 billion nationwide. By 2003, this

Figure 4

Regional System Operators in the United States and Canada

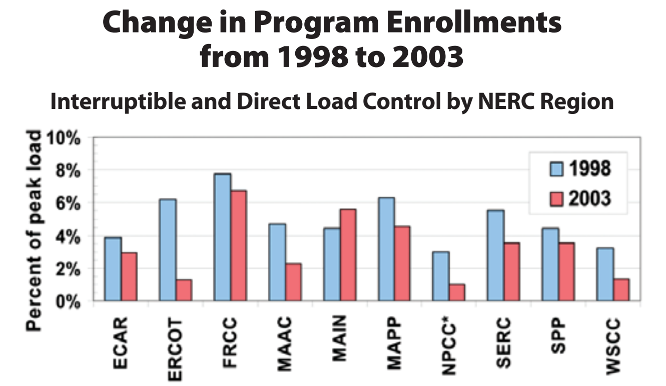


Source: (ISO/RTO Council 2011)

11 FERC granted power producers the right to sell at market rates only if they could show that they lacked market power and that the prices reflected actual market dynamics of supply and demand.

value had fallen by more than half to $1.3 billion. The change in program enrollment from 1998 to 2003 for each reliability region is shown in Figure 5, illustrating that program enrollments dropped precipitously along with funding, although it is worth noting that enrollments declined in virtually every region, even those where no divestment or restructuring occurred.[12]

Figure 5



Note: Data from NERC 1998 and 2003 summer assessments. NPCC data is for for 1998 and 2002.

Source: (Goldman 2005)

Soon it became evident that competitive wholesale markets, as initially implemented, would not automatically produce all the benefits initially expected, and that additional structures would be needed to manage energy price volatility, ensure system reliability, and guard against market power. An example of the difficulty involved in the restructuring process was that encountered during the Western States Power Crisis of 2000-2001. During this period, California's total wholesale power costs more than tripled and wholesale market prices soared to almost $400 per megawatt hour (Weare 2003, Joskow 2001). Traditional demand response programs were subjected to new and unexpected stresses.

During this time, California customers on interruptible/curtailable rate programs, many of whom had never previously been called upon to curtail, were suddenly subject to frequent interruptions – 23 times during the last eight months of 2000 alone. Due to the high number of interruptions, many customers began to leave the program or chose to ignore curtailment instructions and suffer penalties (Goldman, Eto, and Barbose 2002).

Following the Western States Power Crisis and subsequent sharp electric rate increases in some regions,[13] policymakers concluded that investments in demand response were necessary to ensure the efficient functioning of the wholesale markets and reliability of the grid (Cappers, Goldman, and Kathan 2009).

At the same time, the potential for demand response to bring benefits to electricity customers was expanded by restructuring's greater emphasis on wholesale markets. Previously, vertically integrated utilities had carried out demand response programs to prevent blackouts or control costs during peak periods, but with the expansion of wholesale markets, the scale of impact was greatly broadened. Instead of waiting for a utility to curtail load during a few hours a year, demand response providers were empowered to participate on an ongoing basis in the market to reduce volatility, improve the elasticity of demand, and potentially reduce the market clearing price for energy purchases for a much larger number of customers across entire regions.

Moreover, the pool of potential participants widened considerably due to the restructured wholesale markets' creation of opportunities for entrepreneurs to find innovative means to supply demand response (York and Kushler 2005). First, however, the necessary policy and regulatory framework had to be established to govern the treatment of demand response and enable demand response to be compensated in a manner comparable to generation resources. In the following section, we describe these developments, which were initiated in large part by FERC proactively responding to the identification of market barriers to demand response.

12 Reliability regions were created under the North American Electric Reliability Corporation (NERC) after the Northeast Blackout of 1965. Note that the reliability regions do not correspond directly to the ISO/RTO areas. In addition, some of these reliability regions have merged since 2003.

13 In Maryland, Baltimore Gas & Electric's retail rates increased 72 percent in 2006, leading to the dismissal of the state's five public utility commissioners (Pfeifenberger, Basheda, and Schumacher 2007). Other examples of price increases following restructuring include a 50 percent increase in Connecticut and a 59 percent increase in Delaware (Pfeifenberger, Basheda, and Schumacher 2007). In several situations, the rate increases reflected up to ten years of suppressed rates due to rate caps negotiated at the start of restructuring. In California, total wholesale power costs rose from $7.4 billion in 1999 to $27 billion per year from 2000 through 2001 (Weare 2003).

C. Recent Developments to Enable Demand Response in US Markets

The transition to competitive wholesale markets entailed significant market design efforts, which initially focused on supporting the participation of traditional generation resources, rather than demand-side resources. This supply-centric focus created numerous barriers to demand response participation. These included restrictive rules that increased the cost of participation; limitations on the entities allowed to bid in demand response; a failure to provide compensation for certain services; and, in some cases, the outright prohibition of demand response's participation in the market (FERC Staff 2009). As a result, modifications by FERC and state regulators have been necessary in order to provide legacy utility demand response programs as well as new providers of demand response with the opportunity to fully participate in the organized wholesale markets (Cappers, Goldman, and Kathan 2009).

Federal support for demand response was underscored with the Energy Policy Act of 2005, which stated that the official policy of the United States was to encourage demand response, facilitate the deployment of enabling technology and devices, and to eliminate unnecessary barriers to demand response's participation in energy, capacity, and ancillary service markets. The Act also declared, "the benefits of such demand response that accrue to those not deploying such technology and devices, but who are part of the same regional electricity entity, shall be recognized," implying that an accurate assessment of the benefits of demand response must take into account impacts on all regional customers.

In order to implement the policy goals articulated in the Energy Policy Act of 2005, modifications to the operation of wholesale markets have been required to set demand response on an equal footing with generation resources. Over the past few years, FERC has taken multiple steps to remove remaining barriers to demand response. These include:

- Order No. 890, issued in February 2007. This order modified the Open Access Transmission Tariff to allow non-generation resources such as demand response to provide certain ancillary services (e.g., regulation and frequency response, spinning reserves, and supplemental reserves services) on a comparable basis to services provided by generation resources. The order also directed transmission providers to treat demand response comparably to traditional resources in the transmission planning process.
- Order No. 719, adopted in October 2008. Known as the Wholesale Competition Final Rule, this order established regulations to improve the competitiveness of wholesale electric markets through demand response. The rule required system operators to study, and if necessary, reform market rules "to ensure that the market price for energy reflects the value of energy during an operating reserve shortage," in order to encourage the entry of new resources, including demand response.
  Order No. 719 also directed system operators to accept bids from demand response resources in the provision of certain ancillary services on a basis comparable to other resources, and permits aggregators of retail customers to bid demand response on behalf of retail customers directly into the wholesale market, unless prohibited by law or regulation (FERC Staff 2009).
- Order No. 745, issued in March 2011. This rule addresses the compensation for demand response in wholesale energy markets, requiring that demand response be compensated the full market price of energy,[14] when it is determined that the resource is capable of balancing supply and demand and is cost-effective. Further, the order specifies that costs are to be allocated among customers who benefit from the lower market price of energy resulting from demand response.
  The basis for this compensation requirement is to provide comparable compensation to both generation and demand response providers, based on the premise that they provide comparable services to the grid operator. In addition, compensation based on the full market price of energy is designed to facilitate the recovery of demand response technology investment costs, thereby encouraging greater participation of demand response in wholesale markets.
- Order No. 755, issued in October 2011. This order pertains to compensation of resources providing regulation service. FERC found that resources providing such services differ in their ramping

14 The FERC Order refers to the full market price of energy as the "locational marginal price" (LMP) to reflect the fact that in FERC regulated markets the energy price includes a locational signal (to varying degrees).

ability and the accuracy of their response, yet compensation by system operators did not account for such differences. Thus FERC found rates unjust and required system operators to base payment in part on the performance of each resource.

The long-term impacts of FERC's more recent orders are yet to be determined. In several cases, system operators have not yet finalized their compliance filings detailing their market rule changes. In other cases, the compliance filings have been approved, but the implementation of market designs to accommodate changes are not scheduled to occur for several years. For example, ISO-NE recently made its third compliance filing relating to Order 755 requirements to make regulation market changes to provide for two-part bidding, uniform pricing, and two-part payments to providers of regulation service. These changes will allow new technologies (other than traditional generation resources) that can provide regulation services to compete in ISO-NE's regulation market. However, the earliest implementation of the changes would be for the 2015-2016 power year.

In regard to Order No. 745, each ISO or RTO has submitted compliance filings to FERC including tariff revisions to implement the order's requirements. PJM implemented the changes in April 2012 and observed increased demand response participation in the summer of 2012 (discussed in Section 5). ISO-NE has commenced a multi-year transition period to fully implement the rules. The other major system operators, MISO, NYISO, SPP, and CAISO, are in the process of finalizing their Order No. 745 filings. ERCOT is not subject to FERC regulation of its markets and has elected to determine compensation for ERCOT demand response programs separately from the energy market.

D. Demand Response Today

The amount of demand response available to system operators has begun to rebound in many regions since the low levels reached at the beginning of restructuring.

Table 1

Demand Response Available at US ISOs and RTOs[15]

	2009 (MW)	Percent of 2009 Peak Demand[9]	2010 except as noted (MW)	Percent of 2010 Peak Demand[9]
California ISO	3,267[1]	7.1%	2,135[1]	4.5%
Electric Reliability Council of Texas	1,309[2]	2.1%	1,484[3]	2.3%
ISO New England, Inc.	2,183[2]	8.7%	2,116[4]	7.8%
Midwest Independent Transmission System Operator	5,300[2]	5.5%	8,663[5]	8.0%
New York Independent System Operator	3,291[2]	10.7%	2,498[6]	7.5%
PJM Interconnection, LLC	10,454[2]	7.2%	13,306[7]	10.5%
Southwest Power Pool, Inc.	1,385[2]	3.5%	1,500[8]	3.3%
Total RTO/ISO	27,189	6.1%	31,702	7.0%

Sources:
[1]California ISO 2010 Annual Report on Market Issues and Performance
[2]2010 FERC Survey
[3]ERCOT Quick Facts (June 2011)
[4]2010 Annual Markets Report, ISO New England Inc.
[5]2010 State of the Market report, Potomac Economics (Midwest ISO)
[6]2010 State of the Market report, Potomac Economics (New York ISO)
[7]PJM Load Response Activity Report, July 2011, "delivery year 2011-2012 active participants"
[8]Informational Status Report Concerning Incorporation of Demand Response In SPP Markets and Planning (September 2, 2011)
[9]Estimated based on peak demand data from the following: California ISO 2010 Annual Report on Market Issues and Performance, 2010 State of the Market Report for the ERCOT Wholesale Electricity Markets, 2010 Assessment of the Electricity Market in New England, 2010 State of the Market Report for the MISO Electricity Markets, New York ISO 2010 State of the Market Report, 2009 State of the Market Report for PJM and 2011 Quarterly State of the Market Report for PJM: January through June, and the Southwest Power Pool 2010 State of the Market.

Source: (FERC 2011)

In 2010, the potential resource contribution of demand response to system operators in the United States totaled 31,702 MW. As a percentage of peak demand, these resources provided between 2 and 10 percent of each region's peak demand, as shown in Table 1, above (FERC 2011).

Currently there are numerous ways in which dispatchable demand response can operate. In regions with organized wholesale markets, demand response resources can typically bid directly into the market or be dispatched in response to market signals. However, the degree to which demand response is integrated into the wholesale market varies, with some regions allowing demand response to set the market clearing price, while other regions restrict demand response's ability to influence market prices. Finally, across the United States,

15 The decline in demand response enrollment for CAISO is due to the way that DR capacity is assessed and reported, from planning estimates to *ex post* estimates. The decline for NYISO is not consistent with NYISO annual reports and is likely due to differences in definitions and the way data were reported between the FERC survey and NYISO's annual report.

and particularly in areas without wholesale markets, utilities may maintain their own demand response programs such as direct load control for water heaters and air conditioning units.

The remainder of this report focuses primarily on demand response programs in regions with wholesale markets, but also includes results from a pilot program in the Pacific Northwest that operates outside of any wholesale market. In general, regional dispatch of demand response through system operators provides a more flexible and sophisticated means of addressing a variety of system needs. Therefore, the majority of the case studies in this report relate the US experience to date with centrally-dispatched demand response.

Fully-Integrated Market-Based Demand Response

In some organized wholesale markets under FERC regulation, demand response has been fully integrated into the various electricity markets. Fully-integrated "market-based" demand response implies that demand response can set the market clearing price, rather than merely reacting to the clearing price. These resources are also dispatched by the system operator. Today, market-based demand response performs the following roles:

- **Energy Resource:** Demand response that participates in the energy market is dispatched for economic reasons. Demand response providers may bid their demand reduction directly into either the day-ahead market or the real-time market. If the bid is less than the market clearing price, the resource is dispatched by the system operator and receives the energy market price as payment.
- **Capacity Resource:** Demand response can play a key role in ensuring resource adequacy. In regions that have capacity markets, demand response providers may bid in a set amount of load that can be curtailed during a capacity shortfall. These providers must have the capability of curtailing load on short notice (usually within 30 minutes to two hours). The number of times that the provider may be called upon to provide a demand reduction varies with the specific market product, e.g., ranging from a maximum of ten times a year to an unlimited number of interruptions. Providers of this type of demand response receive capacity payments.
- **Ancillary Services Resource:** Demand response resources that have the ability to curtail on very short notice (30 minutes or less) may participate in ancillary services markets. Examples of ancillary services include frequency regulation, spinning reserves, and non-spinning reserves. A brief description of these services is provided in Table 2, above. Additional descriptions of ancillary services are provided in Section 7 and the appendix.

Table 2

Ancillary Services Provided by Demand Response

Service	Description
Frequency Regulation	Increase or decrease load in response to a real-time signal, generally within a few seconds
Spinning Reserves	Load reductions synchronized and responsive within the first few minutes of an emergency event
Non-Spinning Reserves	Demand resources available within 10 minutes

(Adapted from Levy, Kiliccote, and Goldman, 2011)

Market-Reactive Demand Response Programs

Demand response may also participate in the wholesale market in a more limited manner through programs that are reactive to market signals, but not fully-integrated. These programs typically offer services similar to fully-integrated market-based demand response and are also dispatched by the system operator, but resources in these programs are unable to set the market clearing price. Rather, these resources are dispatched in response to a market signal such as high energy prices, and they do not influence the price signal set by the intersection of traditional supply and demand.

An example of a market-reactive program is ISO-New England's day-ahead load response program. Once the day-ahead energy market clearing price has been set through the participation of traditional generation supply and demand, demand response resources that bid less than the clearing price are scheduled for dispatch in real-time. The clearing price in the day-ahead energy market is not affected by the demand response resources; only the real-time energy market will see an impact. This implies that reactive demand-response programs are unable to provide a market-wide price reduction effect, since they are prevented from impacting the clearing price set by traditional generation resources in the day-ahead market. This lessens the price impact from the participation of demand response.

In other cases, demand response programs are even more detached from the wholesale markets, with payments determined ahead of time by the system operator or utility. For example, demand response providers may agree to provide a capacity service by curtailing load when called by the system operator during system emergencies, and may be compensated based on a pre-determined price per kilowatt hour.

Non-Market Local Demand Response Programs

Where demand response is unable to participate in wholesale markets (whether or not these markets exist), demand response may be dispatched by local utilities or regional operators based on utility or balancing authority needs. Examples include CAISO, where distribution utilities inform the system operator of expected curtailments a day in advance but these resources are not bid into the wholesale market, and the Bonneville Power Administration in the Pacific Northwest that dispatches balancing resources – including demand response – as needed.

Despite the lack of central coordination through a system operator and the exclusion of these resources from market participation, these programs have numerous benefits to the electric system and utilities, particularly through the enhancement of reliability and the procurement of the most cost-effective resources. However, the quantity of resources that participate and the potential economic benefits may be more limited than where demand response participates in wholesale markets through dispatch by system operators.

4. Demand Response for Resource Adequacy

Resource adequacy is one element of overall system reliability, and refers to the procurement of sufficient resources to meet annual peak demand. To better understand demand response services (actual and potential), it is useful to review the actual performance of demand response in specific programs across various regions. Although a capacity market is not required to achieve demand response participation, we have found that a mechanism providing a steady monthly payment in exchange for a guaranteed response has been successful in the United States. Capacity markets are one such mechanism, but not the only one. Because they have been successful at attracting demand response resources, we start our discussion with them, but also discuss other mechanisms that use demand response to address resource adequacy.[16]

A. Forward Capacity Markets

Regional system operators are responsible for procuring sufficient capacity to ensure system reliability, and in some cases do so several years into the future by holding forward-looking capacity auctions. In these auctions, the system operator solicits bids to meet estimated future peak demand, and then provides market-based compensation to resources who agree to be available during the expected peak hours. Resources that are cleared in the auction receive a stream of payments at the auction price, typically in dollars per megawatt-day or dollars per kilowatt-month (Gottstein and Schwartz 2010).[17]

Independent System Operator of New England

The Independent System Operator of New England operates the electric grid for six states in the northeast corner of the country: Maine, New Hampshire, Vermont, Massachusetts, Connecticut, and Rhode Island. The region experienced its all-time peak load of just over 28,000 MW on August 2, 2006.

For a number of years before December 2006, the ISO-NE operated a monthly residual capacity market simply called the Installed Capacity Market. All load serving entities were required to demonstrate that they had contracted for sufficient capacity to serve their share of the region's peak load plus annual reserves. Any shortage would be purchased in the monthly residual Installed Capacity Market. However, prices in the Installed Capacity Market were extremely low for many years, below $1/kW-month, and more than 5,000 MW of generation resources sought retirement, citing economic woes.

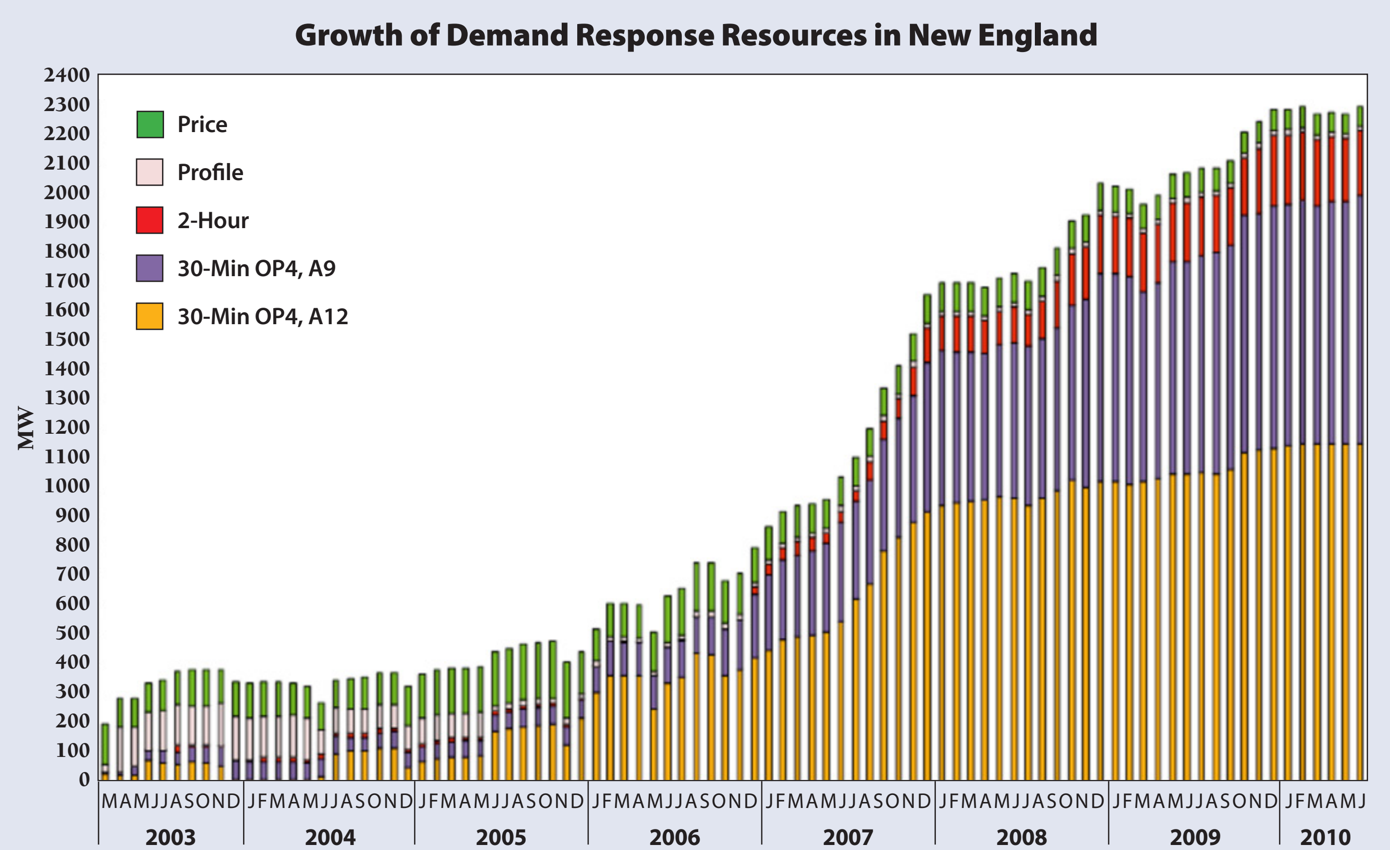
Upon orders from FERC, the ISO-NE spent many years designing the nation's first Forward Capacity Market. The Forward Capacity Market required all resources seeking to provide capacity to qualify for and participate in an auction that would occur three years in advance of the required delivery date. The auction price is good for one year of services, although new resources may select the first year clearing price to remain in effect for up to five years. The three-year-in-advance feature is intended to allow new generation resources sufficient time to finish development and become operational. Demand response resources can qualify to participate in the Forward Capacity Market, but were never officially a part of its predecessor, the Installed Capacity Market.

The growth of demand response in New England correlates directly with market opportunities.

16 Other elements of reliability include responding to system interruptions, maintaining a balanced system, and ensuring a constant voltage. These elements of reliability are typically met with various types of ancillary services.

17 For more information, the reader is directed to the 2010 report titled *The Role of Forward Capacity Markets in Increasing Demand-Side and Other Low-Carbon Resources: Experience and Prospects* by Gottstein and Schwartz available at www.raponline.org/docs/RAP_Gottstein_Schwartz_RoleofFCM_ExperienceandProspects2_2010_05_04.pdf.

Figure 6



Source: (ISO New England Demand Response Working Group, June 2012)

Figure 6 shows the growth of demand response in New England from 2003 - 2010. The "Price" category is, as the name implies, a voluntary response based upon a prediction of energy market prices exceeding $100/MWh. We would loosely categorize this program as demand response in the energy market, and it was a precursor to the ISO-NE's Day Ahead Load Response Program, which later evolved, through FERC Order 745, into the Transitional Price Responsive Demand program and finally into the fully-integrated Price Responsive Demand program. We discuss energy market participation more fully in Section 6.

The other categories on this chart represent demand response with a mandatory response obligation and a monthly reserve payment. In other words, demand response as a capacity resource, even though they were not specifically allowed to participate in the Installed Capacity Market. There were a small number of resources in the Profile and 2-Hour programs. These two programs paid customers willing to respond either within two hours of a dispatch signal, or on a custom response profile as approved on a case-by-case basis. The bulk of the demand response customers were in the 30-minute programs that were called during specific actions taken when the New England system was operating with a shortage of real-time reserves. This condition allowed the ISO-NE to use actions under its Operating Procedure 4, and demand response could be dispatched at Action 9 (reduction in demand) and Action 12 (backup generation subject to environmental restrictions, called only in conjunction with voltage reduction on the system).

Figure 6 shows that from 2003 through most of 2006, these programs sustained little or no growth in participation. Throughout the winter of 2005 and into 2006, ISO-NE and their stakeholder group – the New England Power Pool – negotiated and settled upon a design for the Forward Capacity Market. This design included a transitional Forward Capacity Market phase that started in December 2006 and lasted through May 2010, and then prices set by an auction with an administrative floor price. The price paid to all capacity resources was substantially higher than prior levels of less

Table 3

New England Capacity Prices Applicable to All Capacity Resources Dec 2006 – May 2013

Capacity Market	Months	Capacity Price ($/kW-month)
Monthly Installed Capacity	Pre- Dec 2006	Residual auction whose price varied monthly. Generally <$1.00
Transitional Forward Capacity Market	Dec 2006 – May 2008	$3.05
Transitional Forward Capacity Market	June 2008 – May 2009	$3.75
Transitional Forward Capacity Market	June 2009 – May 2010	$4.10
Forward Capacity Market	June 2010 – May 2011 (first Forward Capacity Market period)	$4.50
Forward Capacity Market	June 2011 – May 2012	$3.60
Forward Capacity Market	June 2012 – May 2013	$2.95
Monthly Weighted Average	Dec 2006 – May 2013	$3.61

Source: Compiled from Forward Capacity Market Settlement Agreement and Forward Capacity Auction Results filings

than $1/kW-month, as shown in the table above.

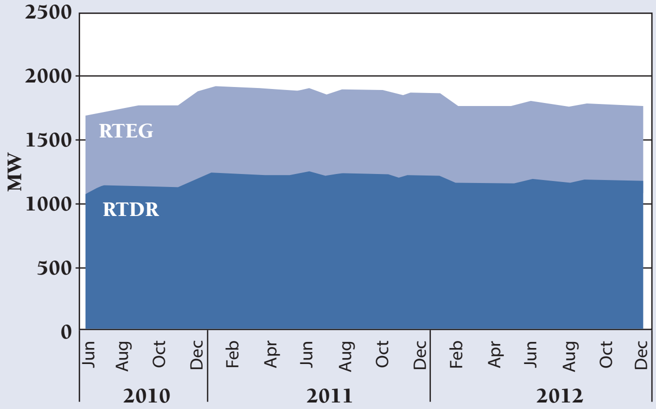
The prices in Table 3 and the amount of demand response provided in Figure 6 show that demand response providers increased their participation in response to a clear signal of a multi-year, reasonably steady monthly price for providing capacity.

With the adoption of the Forward Capacity Market, demand response was allowed to directly participate in the market and the number of capacity programs was reduced to only two: Real Time Demand Response and Real Time Emergency Generation. Real Time Demand Response refers to a reduction in energy usage at an end-use customer facility, while Real Time Emergency Generation refers to an on-site generator behind the customer meter that has environmental permits limiting its operation to "emergency" hours when the system operator calls upon them in order to prevent the load shedding. Both types receive the monthly capacity payment and in return must respond within 30 minutes when dispatched by the ISO-NE control room operators. In the first three commitment periods of the Forward Capacity Market, the price has dropped slowly, and the growth of demand response participation has plateaued.

Figure 7 shows that in New England, nearly 2,000 MW of customer demand are willing to accept an obligation to respond to a dispatch signal from the ISO in return for a monthly reserve payment that has averaged slightly more than $3.50/kW-month over the most recent six-year period. Their expectation from recent events is that they will be dispatched rarely. Real

Figure 7

ISO-NE Demand Response



Source: ISO New England Demand Resources Working Group. December 2012.

Time Emergency Generation resources were dispatched region-wide on only two days in August 2006 for a total of 13 hours, and three days in 2010 for a total of 11 hours. Real Time Emergency Generation resources in Maine were dispatched during a gas pipeline interruption in December 2007 for 23 hours. They were dispatched only in Boston for eight hours in May 2008. Real Time Demand Response resources were dispatched only slightly more frequently, for 15 hours over three summer days in 2006, 16 hours over two days in 2007, eight hours in one day in 2008, only 3.5 hours in the summer of 2010, eight hours over two days in 2012, and not at all during the summer of 2009 (ISO-NE 2006-2012).

Table 4

Real Time Demand Response Performance
June 24, 2010

Load Zone	Dispatched (MW)	Performance (MW)	Percent
Connecticut	226.8	170	75%
West Central Massachusetts	79.6	79	99%
Northeast Massachusetts	70.7	46	65%
Southeast Massachusetts	45.2	30	66%
Rhode Island	27.8	27	97%
Vermont	23.7	29	122%
New Hampshire	29.1	33	113%
Maine	166.2	239	144%
Total	669.0	653	98%

Source: (Taniwha 2010)

Table 5

Real Time Demand Response Performance – Period of 100% Dispatch
July 22, 2011

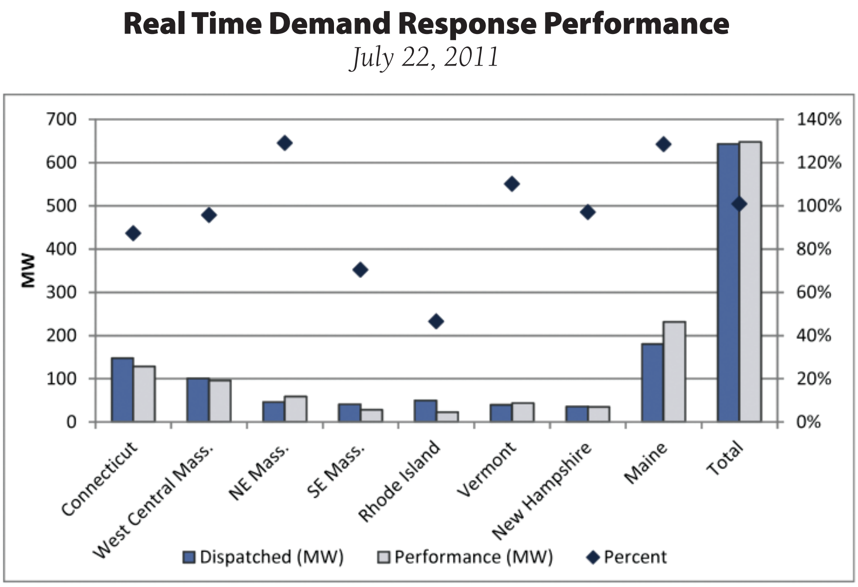
Load Zone	Dispatched (MW)	Performance (MW)	Percent
Connecticut	148.2	129.1	87%
West Central Massachusetts	100.5	96.1	96%
Northeast Massachusetts	46	59.3	129%
Southeast Massachusetts	41	28.8	70%
Rhode Island	50.2	23.3	46%
Vermont	40.2	44.2	110%
New Hampshire	36	34.9	97%
Maine	180.8	232.1	128%
Total	642.9	647.8	101%

Source: (Smith 2011)

In the past six years, neither Real Time Demand Response nor Real Time Emergency Generation resources have been dispatched to meet their capacity requirement for more than 24 hours in any year.

In actual events, the overall performance of demand response resources in the New England capacity market relative to their obligation is quite high. Several generation reductions and outages over the course of the morning and early afternoon on June 24, 2010 led to a capacity deficiency event. Of the 669 MW of demand response called, 653 MW responded, as shown in Table 4. On July 22, 2011, 643 MW of demand response were dispatched in New England, and 648 MW responded (Table 5 and Figure 8). Performance in each dispatch zone varied, however, from 46 percent in Rhode Island to 129 percent in Northeastern Massachusetts (Smith 2011). It is important to note that over-performance can be just as problematic as under-performance for system operators who are tasked with balancing load and supply.

Figure 8



Source: (Smith 2011)

PJM

PJM is the regional system operator for a region that covers 14 states in the Mid-Atlantic and Midwest regions of the United States. PJM manages a peak load of over 160,000 MW. As an RTO, PJM implements the open-access transmission tariff, designs and operates the wholesale markets, and oversees regional system planning. PJM also functions as the balancing authority for the entire market area as a single balancing area. PJM was an early leader in developing opportunities for demand response resources and continues to look for new and innovative demand response applications. In 2002, PJM requested FERC approval of initial demand response program designs for demand response participation in resource adequacy and energy markets.

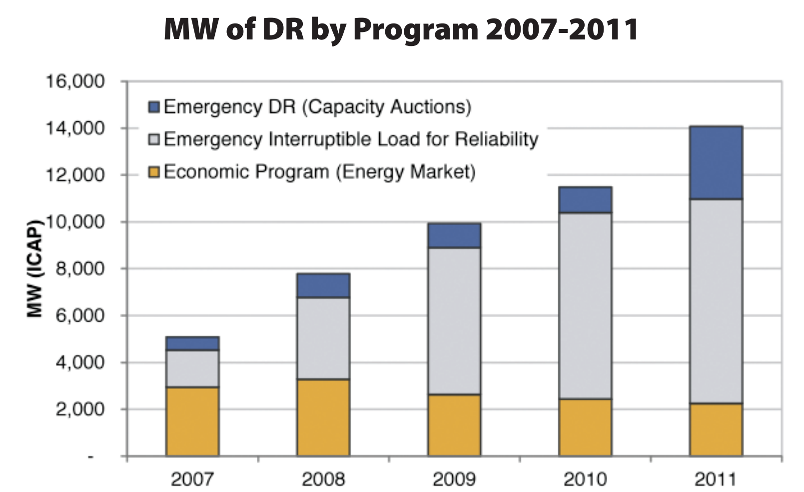
PJM provides detailed annual reports from 2007 through 2012 that cover both Emergency demand response resources (providing capacity, or resource adequacy, service) and Economic demand response resources (energy market services). The Economic program is discussed in a later section of this report. In this section we look at the PJM Emergency capacity program.

Since 2007, PJM has operated a forward capacity market called the “Reliability Pricing Model,” which replaced an earlier capacity market design from 1999. Through the forward capacity market, PJM procures capacity resources for resource adequacy three years in advance of a future delivery year by way of an auction that uses a demand-curve mechanism. In most years, the annual auction procures the full amount of capacity needed three years into the future, or an excess of that quantity if offer prices are low. If offer prices exceed a pre-determined cap, the annual auction will procure less than the needed amount, with the remainder acquired in any of three supplemental auctions prior to the delivery year (Gottstein and Schwartz 2010).

Although there was some minimal participation of Emergency demand response resources in PJM’s capacity market in 2004 through 2006, significant participation did not begin until implementation of the forward capacity market in 2007. In 2008, Emergency demand response resources began to dominate in terms of megawatts and revenues as compared with demand response in the Economic program (the energy market). That domination in terms of megawatts has continued every year, although the total compensation fluctuates based on annual clearing prices in each forward capacity market auction.

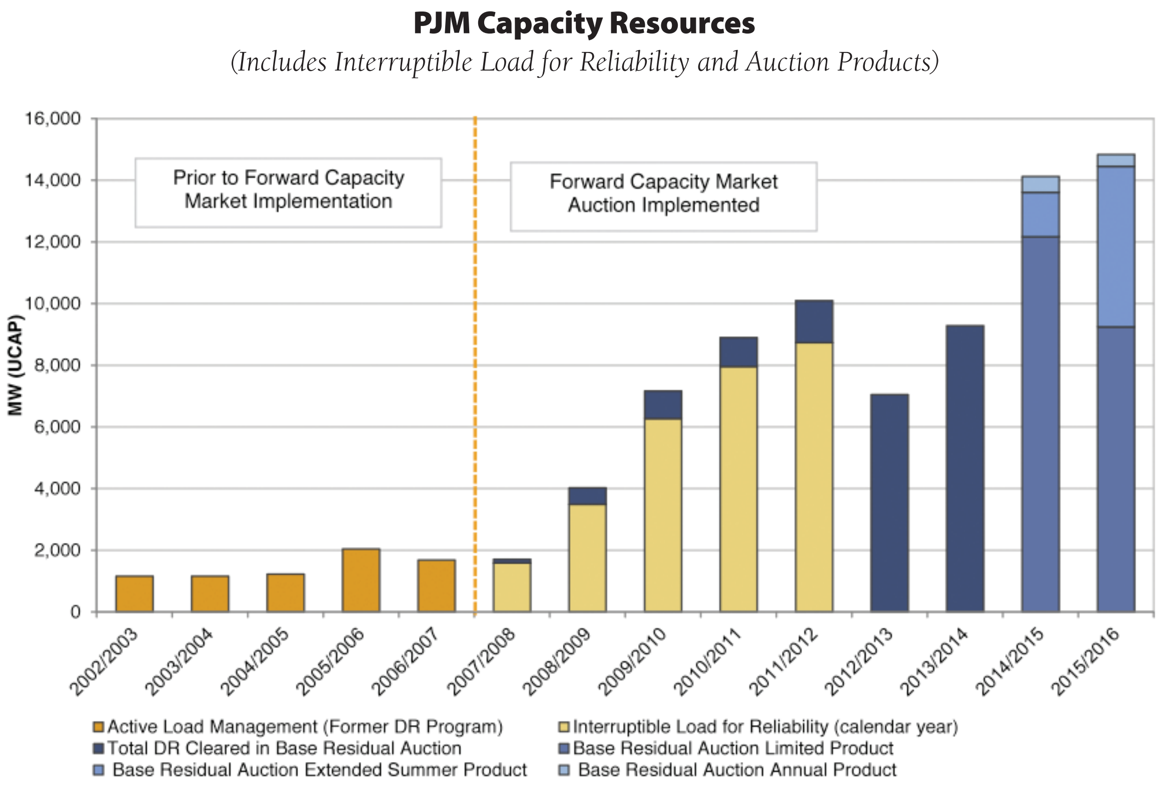
Figure 9 shows the amount of demand response capacity resources (Emergency Interruptible Load for Reliability and Emergency demand response from the capacity auctions) compared with the amount of Economic Program (energy market) demand response from 2007 to 2011. After 2007, Emergency Program

Figure 9



Source: (PJM Load Response Activity Reports 2007-2011)

Figure 10



Sources: (PJM Load Response Activity Reports 2007-2011); (2007-2012 State of the Market Reports); (PJM 2014/2015 Base Residual Auction Results)

demand response increasingly exceeds the amount of demand response offered in the energy market.

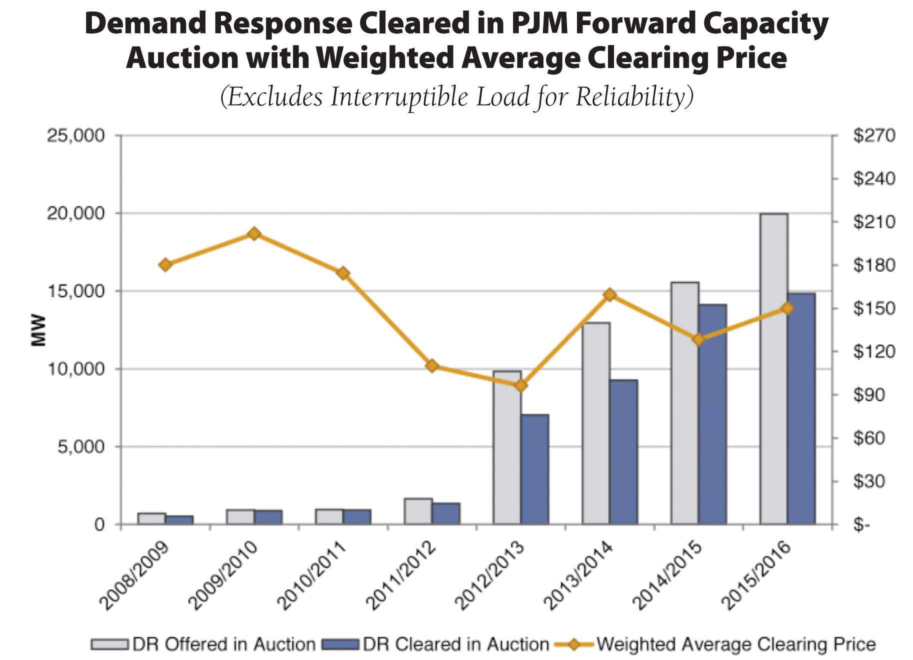
For the first few years of forward capacity market auctions, there were two demand response products: demand response that offered into the three-year forward auction (Emergency DR), and demand response that offered to participate a year or less in advance (Emergency Interruptible Load for Reliability) that did not participate in the three-year forward auctions. Initially, the Interruptible Load for Reliability resources were mostly legacy programs that had been operated by distribution companies and early demand response providers.[18] PJM reduced the total quantity of resources purchased in each three-year forward capacity auction by approximately 2,000 MW to account for these legacy Interruptible Load for Reliability resources that were likely to offer closer to the delivery year (outside of the capacity auction). PJM agreed to provide the set-aside for the first few years of forward capacity market auctions based on claims from legacy Interruptible Load for Reliability resource providers that their customers who signed up to provide demand response could not make a binding commitment three years in advance.

However, as demand response participation increased, particularly in the Interruptible Load for Reliability program, PJM became concerned that the integrity of the annual three-year forward auction was being undermined by the increasingly large quantity of Interruptible Load for Reliability resources that were outside the single clearing price auction. Figure 10 provides a breakdown between Emergency demand response and Interruptible Load for Reliability for the first five years of auctions (delivery years 2007/2008 through 2011/2012). For the 2012/2013 delivery year (auction held in 2009), the legacy Interruptible Load for Reliability program was ended and all demand response resources were required to participate in the three-year forward capacity auction.

Figure 11 shows the dramatic increase in demand response resources offered and cleared in the forward capacity auction for the 2012/2013 delivery year following the conclusion of the Interruptible Load for Reliability program.

18 In PJM, aggregated demand response resources had to be offered by a curtailment service provider (CSP) registered with PJM and the distribution utility.

Figure 11



Source: (PJM 2014/2015 Base Residual Auction Results)

Extended summer DR resources agree to be interrupted on any day from 10:00 am to 10:00 pm (June through October and the following May) for no more than ten hours per event and for an unlimited number of events in the season. Resources in this category have a higher value to PJM operators because they are available seven days a week over a wider range of hours, a two-month longer season than the limited summer product, and have no limit on the number of events for which they can be called.

Annual DR resources have the most value to the system operator. They agree to be interrupted for up to ten hours during any day in the extended summer months and up to ten hours during any day in the winter months.[19] There is no limit to the number of events for which these demand response resources can be interrupted.

In 2010, PJM began discussions about further modifications to demand response participation in annual forward capacity auctions. One of PJM's chief concerns was developing a way to differentiate between demand response resources that could only participate in the limited number of program hours and for no more than ten events a year versus demand response resources that could participate in a broader range of hours, over a longer season, and with more interruptions. The change was intended to address concerns about an over-reliance on demand response resources (with limited dispatch) in particular load zones. These program changes were implemented for the 2011 auction (2014-2015 delivery year).

Instead of a single demand response category, the limited summer demand response already in place, the 2011 forward capacity auction added two more demand response resource categories: extended summer, and annual. These three categories are shown in Figure 10 for the demand response resources that cleared in the last two auctions for the 2014-2015 and 2015-2016 delivery years.

Limited summer DR resources agree to be interrupted on weekdays from 12:00 pm to 8:00 pm (June through September) for no more than six hours each event and no more than ten events in the season. These demand response resources have the lowest value to PJM because of the limitations on their dispatch, but they are the most closely aligned with expected summer peak load hours.

Demand response resources may submit offers that cover just one category (Limited, Extended, or Annual) or offers that cover more than one category. PJM accepts offers based on pre-determined maximum quantities of Limited DR resources for each load zone. All demand resource offers and supply offers then compete to clear in the annual capacity auction. Price-separation between the different demand response resource categories may occur on a system-wide or zonal basis in the annual forward capacity auction.

The differentiation of demand response resources into three categories (Limited, Extended, and Annual) is a logical program design improvement that recognizes the different services that are needed and that demand response can provide. The Annual DR category is for demand response resources that can perform a reserve service (or provide a reserve product) in addition to the resource adequacy (summer peak load) service that is provided by the Limited DR resource. The Extended DR resource is providing a less robust reserve service than the Annual DR resource while also providing the resource

19 Extended summer hours occur during the months of June through October, plus May during the hours of 10:00 am -10:00 pm. Winter hours are November through April during the hours of 6:00 am to 9:00 pm.

adequacy service of the Limited DR resource.

This evolution of demand response resources from a single service to varied services is something that we observe in New England and Texas as well, even though these two regions have very different program designs. Demand response resources that provide only a limited, summer peak service are less valuable to system operators than demand response resources that can provide services in many sets of hours, and over a longer portion of the year. We also see this trend towards more versatile demand response resources in pilot projects and program designs in MISO, CAISO, and the Pacific Northwest. For example, demand response in New England can be called in any hour of the year, even though these events most often occur during summer peak load hours.

Demand Response Performance

Demand response resources' performance is tested through either curtailment events or, for resources that are not called upon during a delivery year, through formal tests to assess performance capability.[20] In the 2011/2012 delivery year, PJM performed testing in 17 zones of 8,860 MW of demand response.[21] Results of the test indicated over-compliance of 660 MW, or a performance level of 107 percent, as shown in Table 6.

Table 6

Load Management Commitments, Compliance, and Test Performance

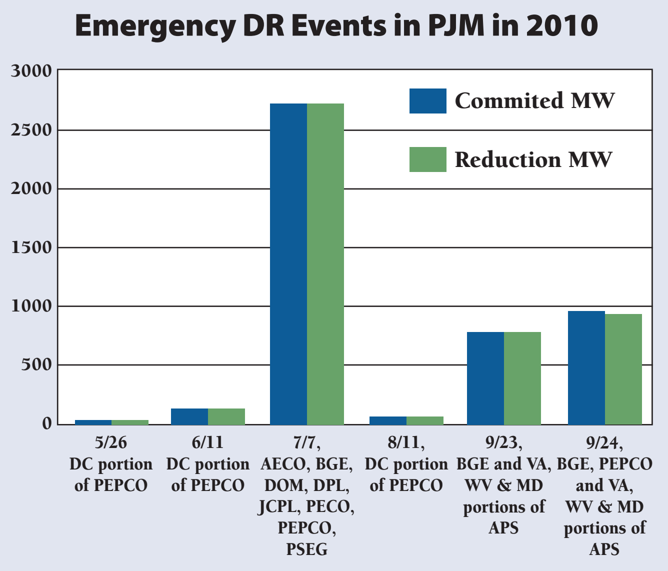
	Test Results			
Zone	Committed MW	Reduction MW	Over/Under Performance MW	Performance Percentage
AECO	90	90	0	**100%**
AEP	1991	2148	157	**108%**
APS	908	943	35	**104%**
ATSI	1107	1220	113	**110%**
COMED	1633	1729	96	**106%**
DAY	219	243	25	**111%**
DOM	1025	1088	63	**106%**
DPL	49	49	0	**100%**
DUQ	5.9	7.5	1.6	**127%**
JCPL	27	27	0	**100%**
METED	3.8	5.2	1.4	**137%**
PECO	1.4	1.2	-0.2	**86%**
PENELEC	393	433	40	**110%**
PEPCO	268	260	-9	**97%**
PPL	734	837	103	**114%**
PSEG	398	436	38	**110%**
RECO	6.4	4.6	-1.8	**72%**
Total	**8860**	**9521**	**660**	**107%**

Source: (PJM Load Management Performance Report 2011/2012)

2010 Events

In a PJM report, the performance of demand response was examined for several dispatch events in 2010. As demonstrated in Figure 12, demand response resources performed very close to their dispatch targets in each event in 2010.

Figure 12



Source: (PJM Load Management Performance Report 2011/2012)

PJM also provided a summary of the demand response performance by zone for these 2010 events as shown in Table 7 on the following page. The range of performance for the 17 events is 87-106%, with 14 events between 94-105%.

20 When we discuss the performance of demand response resources we are referring to the quantity of service (MW or MWh) delivered versus the quantity that had an obligation to be available.

21 Megawatts measured as installed capacity (ICAP).

Table 7

2010 PJM DR Performance Statistics

Event Date	Committed MW	Reduction MW	Over/Under Performance MW	Performance Percentage	Zone	Lead Time
5/26/2010	34	32	-2	94%	PEPCO	Long
6/11/2010	137	132	-6	96%	PEPCO	Long
7/7/2010	70	72	2	103%	AECO	Long
7/7/2010	410	422	12	103%	BGE	Long
7/7/2010	975	935	-39	96%	DOM	Long
7/7/2010	138	144	6	104%	DPL	Long
7/7/2010	155	155	-1	100%	JCPL	Long
7/7/2010	419	438	19	105%	PECO	Long
7/7/2010	179	171	-9	95%	PEPCO	Long
7/7/2010	380	384	4	101%	PSEG	Long
8/11/2010	60	53	-7	89%	PEPCO	Long
9/23/2010	1.5	1.3	-0.2	87%	APS	Short
9/23/2010	378	366	-12	97%	APS	Long
9/23/2010	410	433	23	106%	BGE	Long
9/24/2010	378	355	-23	94%	APS	Long
9/24/2010	410	430	20	105%	BGE	Long
9/24/2010	179	173	-7	96%	PEPCO	Long

Source: (PJM 2010)

July 2012

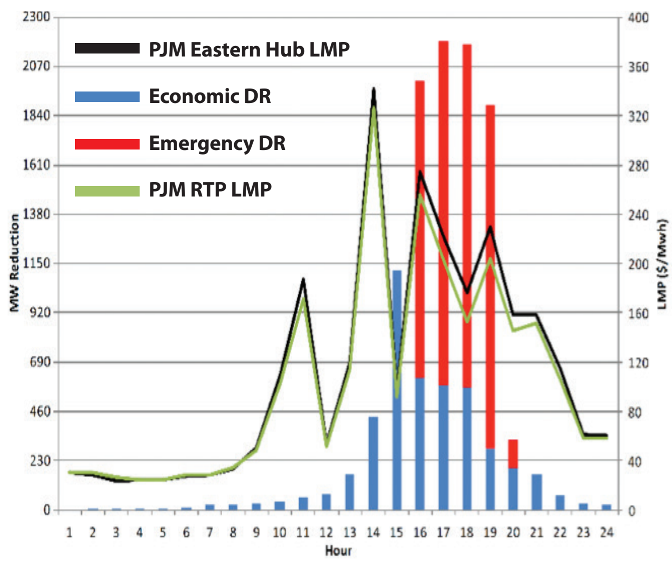
On July 17 and 18, PJM dispatched demand response resources due to hot weather and the potential for loads to exceed the resources available. Figure 13 and Figure 14 show the amount of DR that responded for those two days, along with the corresponding energy market price.

Note that after the first rise in energy prices, there was an initial sharp decrease in the clearing price in hour 15 after 1100 MW of Economic DR was included in the energy market bid-stack. By hour 16, when PJM activated the Emergency DR resources for reliability concerns, energy prices had rebounded a bit, but then decreased for the next few hours as these additional demand response resources displaced more expensive generation resources. The jagged pattern of prices throughout the day also indicate that other factors are likely at play, perhaps including failed start of additional supply resources or unit trips. The exact interactions between supply and demand offers can be ascertained after the event through an analysis of the specific bids and PJM's dispatch decisions.

The pattern observed a day later, on July 18, is more straightforward. Energy market prices were generally rising quickly during the morning ramp-up period, but then dropped precipitously as demand response resources

Figure 13

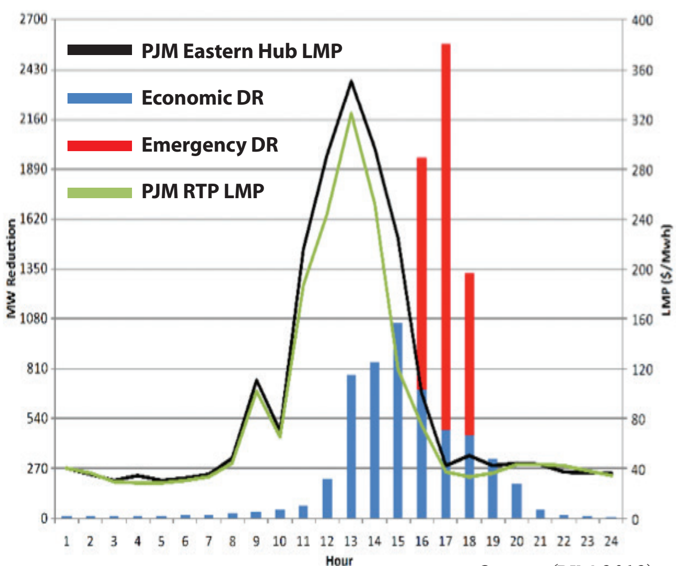
Estimated PJM Demand Response on July 17, 2012



Source: (PJM 2012)

Figure 14

Estimated PJM Demand Response on July 18, 2012



Source: (PJM 2012)

began to respond. The match between the timing of the drop in energy prices and the dispatch of emergency DR may have occurred because many demand response resource owners anticipated the hot weather and began to reduce load in advance of the official start time of the event.

B. Other Capacity Markets

In addition to the three-year forward capacity markets developed by PJM and ISO-NE to maintain resource adequacy, several other regions have implemented capacity market-type requirements.

New York Independent System Operator

NYISO administers its Installed Capacity Market to ensure resource adequacy for its one-state territory with an all-time peak load just over 33,000 MW. A demand response resource is eligible as a capacity resource if it is registered as a Special Case Resource.

The Special Case Resource program is a true capacity product. Market participants – often demand response providers who aggregate many retail customers – receive a monthly payment in response for a mandatory obligation to reduce load when dispatched by NYISO. Special Case Resource events occur when NYISO runs short of operating reserves.

NYISO also runs a voluntary demand response reliability program called the Emergency Demand Response Program that is called during the same hours. Response is voluntary, but those who do respond are paid the higher of the real-time energy market price or $500/MWh. Similar to the reliability programs in other regions, these events are rare, and have occurred for not more than 35 hours in any year since 2001, and only once for more than 24 hours.

In 2003, the Emergency Demand Response and Special Case Resource programs were made mutually exclusive, and the energy payment for energy reductions during a demand response event was incorporated into the Special Case Resource program. Since these modifications, enrollment in the Special Case Resource program has grown rapidly (as shown in Figure 15). As of July 31, 2011 there were 2,173 MW of demand response in the Special Case Resource and Emergency Demand Response programs, from 5,807 customer locations (NYISO 2012c). These resources were called on only two event days in the summer of 2011, and the 1,505 MW with capacity market obligations performed at 90.1 percent and 92.8 percent of their capacity value on these two days (NYISO 2012c, 20). This represents 6.4 percent of that region's peak load in that year. Figure 15 shows the historical growth of participation in the New York Emergency Demand Response and Special Case Resource programs.

Table 8

Total Number of Hours of Special Case Resource/Emergency Demand Response Events in NYISO, by Year

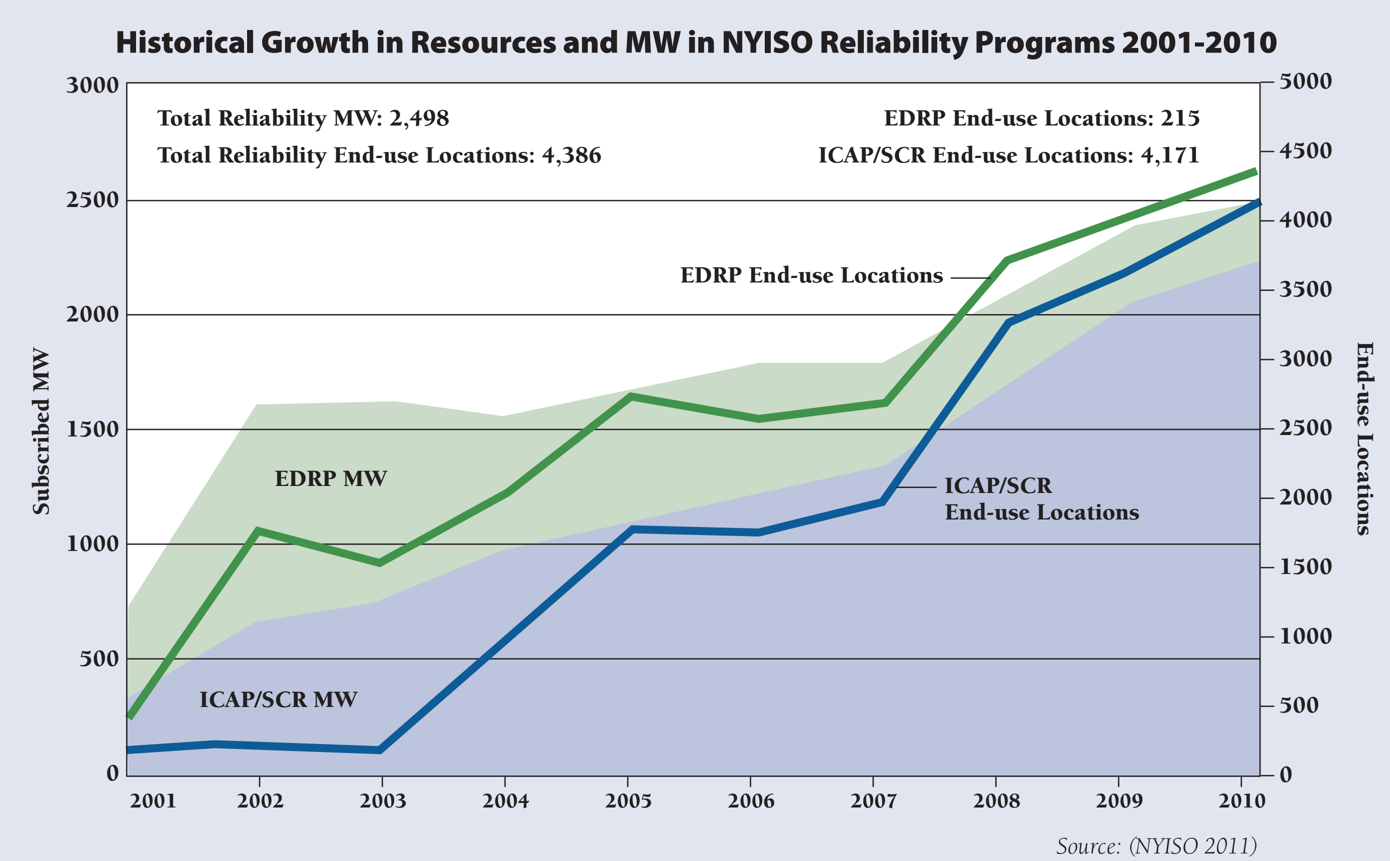
Year	Number of NYISO Event Days	Total Number of NYISO Event Hours	Avg. Number of NYISO Event Hours per Event Day
2001	4	23	5.8
2002	4	22	5.5
2003	2	22	11.0
2004	0	0	
2005	1	4	4.0
2006	5	35	7.0
2007	0	0	
2008	0	0	
2009	0	0	
2010	2	12	6.0
2011	2	11	6.5
Total	**20**	**129**	**6.5**

Source: (NYISO 2012b)

Demand response providers active in these NYISO programs indicate that the growth in participation stems from a number of logical reasons. The capacity market payment has been reasonably high and stable for a number of years, especially in the New York City area. The average monthly price over the past ten years is just over $7.50/kW, with a very regular seasonal pattern. Further, some demand response resources can participate in a program by the local utility that increases payment for demand response in New York City and its surrounding areas. Lastly, the rules around participation, baselines, and event data submittal have remained consistent over this period of time.

Demand response has played a critical role in preventing power outages during summer heat waves in New York. In the summer of 2006, NYISO peak load was on track to climb to an all-time record peak of more than 35,000 MW during the afternoon of August 2nd.

Figure 15

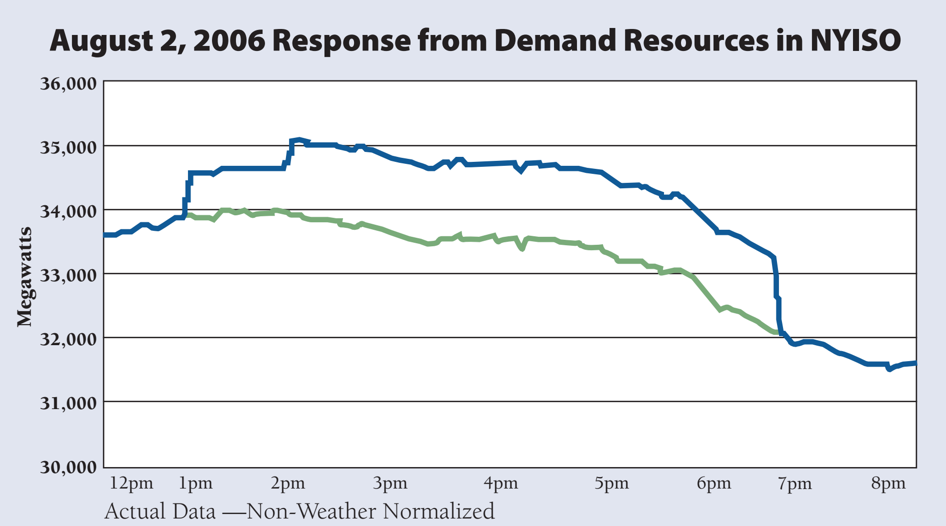


Source: (NYISO 2011)

More than 1,000 MW of demand response resources responded to NYISO instructions to curtail load that day, preventing the system load from rising above 34,000 MW, as shown in Figure 16. The lower line shows the actual system load, while the upper line shows the expected load without demand response.

Another heat wave in 2011 threatened to push system demand above 35,000 again, but demand response responded with more than 1,400 MW (NYISO 2012). Average hourly response has improved in recent years. In 2010, average hourly response, as measured using the installed capacity measure, was between 83 percent and 86 percent, while this increased to between 90 and 93 percent in 2011 (NYISO 2012). Peak load reductions during specific events are shown in Figure 17 on the following page.

Figure 16

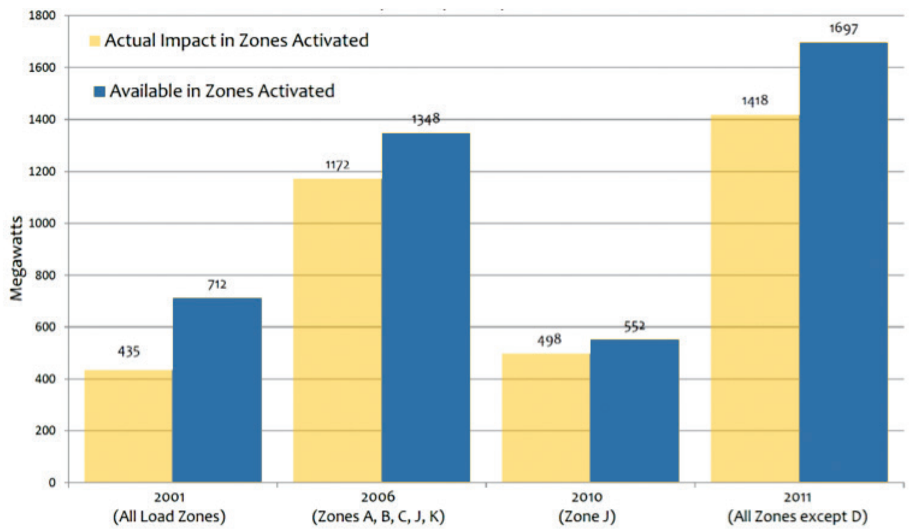


Source: (NYISO 2007)

Midwest Independent System Operator

The Midwest Independent System Operator (MISO) manages a multi-state region with an annual peak load of 105,000 MW. MISO designs and implements wholesale markets for the region and produces an annual system planning report. Unlike PJM and ISO-NE, each of which operates as a single balancing area, MISO operates its single multi-state

Figure 17



Source: (NYISO 2012)

market across a region incorporating numerous distinct balancing areas.

MISO has three demand response programs:

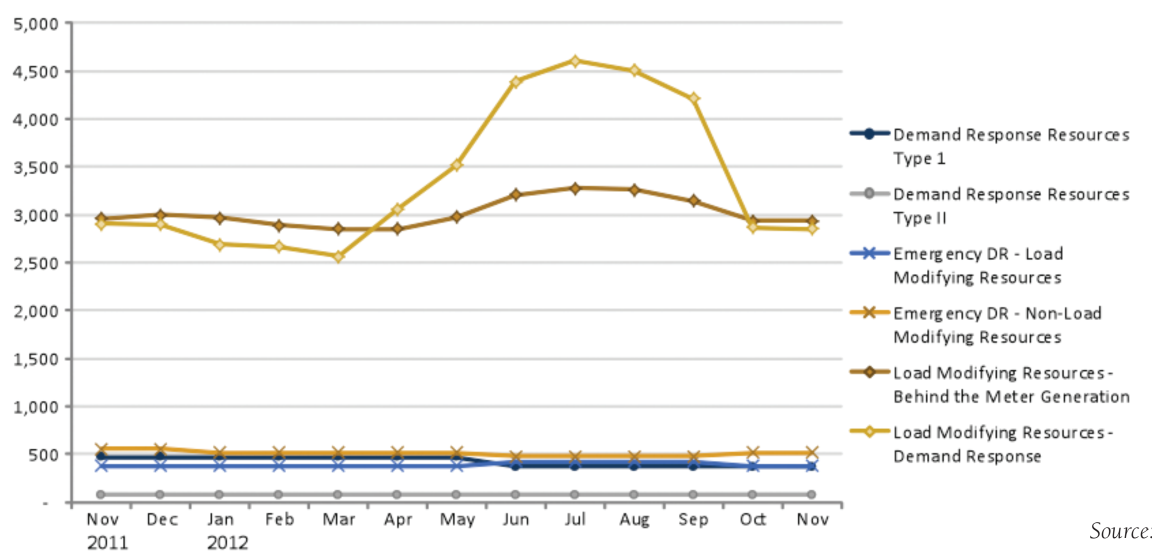
The Emergency Demand Response (EDR) program is for resources that maintain resource adequacy and reduce costs to all market participants. They are compensated based on the higher of their offer or the Locational Marginal Price for their zone. There are about 800 MW of EDR resources in MISO.

The Load Modifying Resources (LMR) program is a resource adequacy program that allows distribution utilities to include customer load reduction capabilities as a component of their monthly capacity obligation. These customers are paid based on their contracts with the local distribution utility. This is the largest component of demand response resources in MISO with over 6,000 MW, about half of which are behind-the-meter generation resources.

The Demand Response Resources (DRR) program is for resources that offer into the energy market, similar to the PJM Economic DR Program. Historically, the payment to demand response resources that participate in the energy market has been less than the full energy market prices. This lower payment has been one of the factors affecting participation levels; there are about 400 MW of demand response resources in MISO. MISO is in the process of complying with Order 745 that requires payments based on full energy market prices. See the discussion of order 745 in Section 6.

Figure 18 shows the current quantities of resources available to MISO as of December 2012.

Figure 18



Source: (MISO)

MISO currently operates a monthly voluntary capacity market (residual) similar to New England's capacity market prior to implementation of the Forward Capacity Market. Demand response is one of the resources that can qualify as a capacity resource. FERC has approved a MISO request to implement an annual capacity market (still voluntary) by June 2013. MISO is still considering whether to propose a mandatory, forward capacity market similar to PJM.

C. Non-Market Demand Response Capacity Resources

California Independent System Operator

The California Independent System Operator (CAISO) is a single-state system operator with a summer peak load of over 46,000 MW. CAISO has been the system operator since 1998 and manages the bulk power transmission system for California, schedules power flows, and performs system planning.

CAISO does not have a specific capacity market structure that would enable the system operator to have direct access to demand response resources. However, like most other regions of the United States, California distribution utilities have maintained traditional load interruption and load shedding programs during the transition from vertically integrated to restructured (and divested) entities.

Based on the most recent surveys, there are 923 MW of emergency resources signed up through various distribution utility programs. Many of these resources are large commercial and industrial customers that include agricultural pumping loads.

California also has 1,612 MW of demand response resources in economic programs that curtail load based on anticipated clearing prices in the real-time energy market. Some of these resources are on voluntary critical peak pricing programs that allow for nine to fifteen events per year.

The decision to activate demand response resources has remained with the individual distribution utilities. CAISO has not been able to integrate the use of demand response resources on a system-wide basis, nor has CAISO been able to forecast or anticipate activations until recently.

The recent changes adopted now provide CAISO, during the summer season, with a spreadsheet from each distribution utility that shows anticipated demand response activations for the current day and the next day, and the balance of demand response resources remaining for each day. The daily update is provided by 8:00 am each day to assist CAISO in its scheduling of resources.

While the current spreadsheet updates are a significant improvement from past practices, the updates are just anticipatory (non-binding) and a long way from the full integration with other market resources that CAISO manages on a daily basis.

As part of a multi-year effort to allow third party aggregators to offer demand response resources to CAISO, the California Public Utilities Commission issued a decision in December 2012 that resolved many issues related to the direct participation of demand response resources in the CAISO markets and established a timeframe for completing work on Energy Rule 24: Direct Participation of Demand Response.[22] In that decision, the California Public Utilities Commission defines the terms of service for demand response providers that seek to enroll or aggregate customers to participate as demand response resources. Energy Rule 24, scheduled for completion in 2013, will further define the rights and obligations of demand response providers. Eventually, demand response providers will be able to directly offer services to CAISO.

Similar to the Midwest ISO and Texas, CAISO has considered a forward market structure for capacity as part of its stakeholder discussions. However, no substantive program designs have advanced beyond the discussion stage. The most recent discussions occurred in February 2013 and can be accessed on CAISO's website.[23]

Current stakeholder discussions are focusing on opportunities for reserves and balancing resources, but until there is resolution of the role of CAISO in controlling the dispatch of demand response resources, there is unlikely to be any significant development of new programs.

22 Decision 12-11-045, issued December 4, 2012.

23 Materials from California's February 26, 2013 Long-term Resource Adequacy Summit can be found here: http://www.caiso.com/informed/Pages/MeetingsEvents/PublicForums/Long-TermRASummit.aspx

5. Demand Response as an Energy Resource

Demand response as an energy resource primarily refers to the ability of individual or aggregated customers to respond to the system need for reduced load during periods of high energy prices, as an alternative to more expensive supply-side resources generating during those hours or, potentially, system blackouts or brownouts. As described below, in markets where demand response has been able to participate in a manner that affects (lowers) the energy market clearing prices during these hours, the result has been significant cost savings to all customers. In addition, participation of demand response in these energy markets improves the ability of customers to reveal the value of "lost load" directly, which both in theory and practice can improve the functioning of competitive energy markets and help to mitigate potential market power abuse.

In addition to reducing load in high-priced hours, demand response may participate in the energy market through increasing load during low-priced hours. Extremely low-priced hours (and even negative prices) can occur when there is excess generation on the grid, such as when wind resources produce significant amounts of energy at night. Some demand response providers, particularly those with some form of storage, increase their demand during these periods, resulting in lower energy prices for those consumers as well as more efficient utilization of generation resources.

Although this capability to increase demand during low-priced hours clearly exists and is being explored in some pilot programs, we have not observed significant amounts of this behavior in the United States to date, and we do not expect it to be dispatched by system operators in the near future.[24] We include this form of demand response with other non-dispatchable, price-based programs, none of which are addressed in this report. This section reviews the more common behavior of reducing demand in response to price.

A. PJM Economic Demand Response

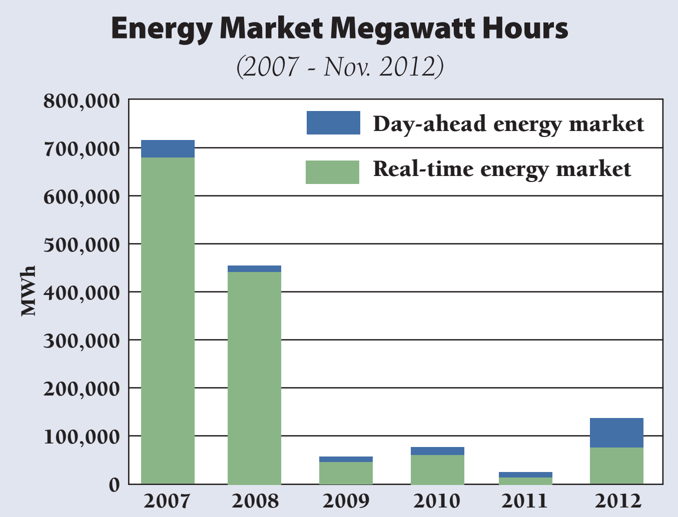
Demand resources have been participating in PJM wholesale energy markets since 2002. The initial design for the PJM Economic Load Response program was approved by FERC for three years (2002 through the end of 2004) and was then extended by FERC for three more years, through the end of 2007. The early program design provided a full energy market price payment to demand response providers whenever the price exceeded $75/MWh. PJM's full energy market price is based on locational marginal pricing. For prices below $75, PJM reduced the full energy market payment to reflect embedded generation, transmission, and distribution costs.[25]

PJM filed, and FERC approved, a modified compensation mechanism for 2008. That mechanism provided payment at full energy market price minus the assumed generation cost (but not transmission or distribution cost) under all market conditions. After that change, demand response resource participation in the energy market decreased steadily in 2008 and remained at a low level from 2009 through 2011. Figure 19, on the following page, shows the decline in 2008 and low levels until 2012, when PJM restored the full energy market price payment to demand response resources pursuant to FERC Order 745.

24 A notable exception would be changes in pumping load from certain hydroelectric stations, who routinely respond to prices. However, we would not classify such electric generation stations as typical end-use customers, and hence are outside our definition of demand response for the purpose of this report.

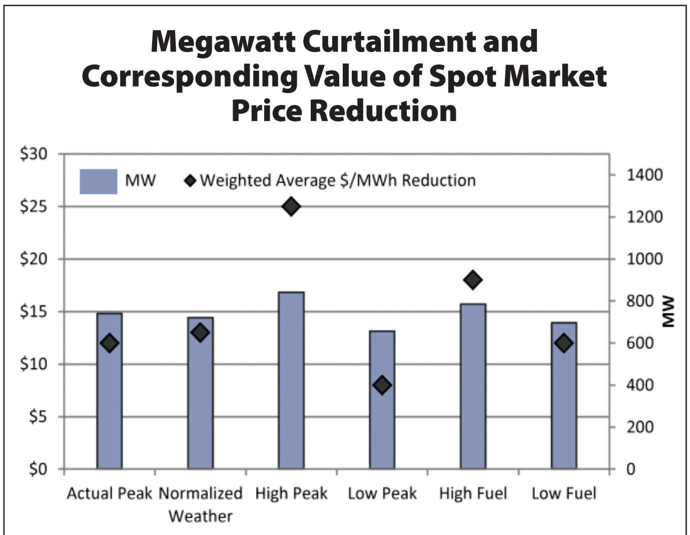
25 The energy market in PJM calculates locational marginal prices to reflect transmission constraints that may occur on the system. These "LMP" values, as they are called, are what we refer to as the full energy market price. Some of the graphics in this report use labels such as "LMP" or "locational marginal price."

Figure 19



Source: (PJM DSR Activity Reports)

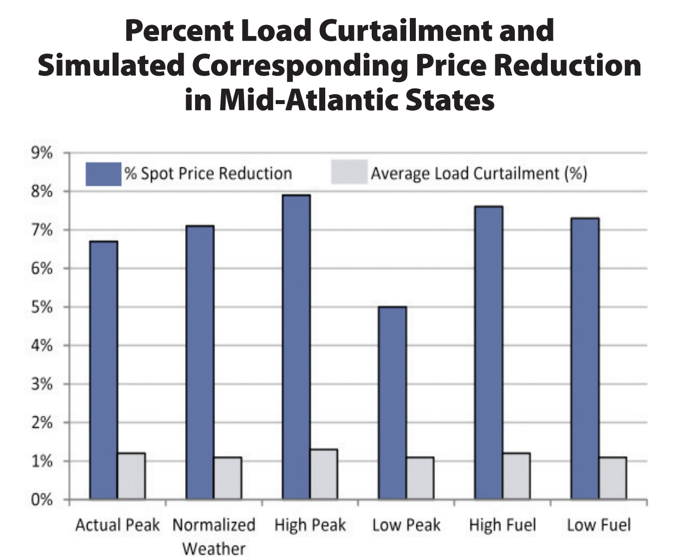
Figure 21



Source: (Brattle Group 2007)

In order to quantify the value of demand response in PJM's energy market, in 2007 the Brattle Group prepared a study that simulated the impact of curtailing three percent of each Mid-Atlantic zone's peak load during the zone's 100 highest hours per year.[26] A variety of market conditions were simulated, including high and low peak load cases, a range of fuel price cases, weather-normalized conditions, and actual peaks experienced in 2005. The results of the study showed that less than 2 percent load curtailment would reduce energy market prices between $8 and $25 per megawatt hour, or 5 to 8 percent on average, depending on market conditions, as shown in Figure 20, below. The estimated benefits to the Mid-Atlantic states ranged from $57 to $182 million per year, while addition benefits of $7 to $20 million would be felt in other PJM states (Brattle Group 2007).

Figure 20



Source: (Brattle Group 2007)

The actual curtailment levels (in megawatts) and dollar value of the market price reductions are shown in Figure 21, above.

Such impressive reductions in wholesale market prices and associated energy cost savings to all consumers are possible with only modest amounts of curtailments due to the steepness of the supply curve at high prices. This is of course even more true when fuel prices are high or there are a large number of peak periods, as shown in the "High Peak" and "High Fuel" cases.

We are not aware of any subsequent studies by Brattle or others that have expanded upon or updated the 2007 study.[27] However, a few earlier evaluations conducted by system operators do exist, which show a range of savings per megawatt hour of demand response. For example, savings from New York's Day Ahead Demand Response Program (DADRP) for the years 2001 to 2005 are shown in Table 9 below. The term "collateral savings" in the table refers to the savings accruing to load-serving entities who purchase electricity in the real-time and day-ahead markets, and who experience savings as a result of lower market clearing prices. The magnitude of the savings per megawatt hour is highly dependent on the slope of

26 The Mid-Atlantic region consists of zones in Pennsylvania, New Jersey, Delaware, Maryland, and Washington DC.

27 Access to specific hourly bid-stack data may be one barrier to such analyses; the Brattle Group study in 2007 was facilitated by PJM providing that access.

the supply curve at the point where the demand curve crosses. When the supply curve is very steep, demand response is more likely to have a significant price impact because of the displacement of high-cost generators than when the supply curve is relatively flat.

Table 9

Estimated Savings Due to Demand Response Depressing Market Prices in NYISO

Year	DADRP MWhs	Collateral Savings	Savings per MWh
2001	2694	$892,140	$331
2002	1468	$236,745	$161
2003	1752	$45,773	$26
2004	675	$8,996	$13
2005	2070	$109,789	$53
Total	**8659**	**$1,293,443**	**$149**

Sources: (Pratt, Cappers, and Anderson 2005); (Neenan et al. 2005)

B. Independent System Operator of New England

Price Response Program

One of the initial programs designed for demand response resources in New England was the Price Response Program. If day-ahead energy market prices were greater than $100/MWh, the ISO-NE would allow demand response to reduce load in real time during those time periods. Response from customers was voluntary, but if they did reduce during these interruption windows, the payment rate was the greater of the real-time energy price or $100/MWh.

Participation in the Price Response Program hit its peak in 2005 when energy prices in New England were high due to elevated natural gas prices following Hurricane Katrina in the Gulf Coast region of the country, a major hub for natural gas pipeline injection. During this time period, day-ahead energy prices frequently exceeded the $100/MWh threshold, and as such the opportunity for participation arose frequently. Yet even at this peak, participation never exceeded 200 MW.

As natural gas prices fell, day-ahead energy prices fell with them, and in recent years the interruption window for the Price Response Program has rarely opened. Price Response Program participation sank in accordance with this lack of opportunity.

Because the Price Response Program was a real-time voluntary action, these resources never actually "cleared" in the real-time energy market along with generation resources. They acted simply as a reduction in the amount of load to be purchased in real-time. The costs from program payments were not charged in the same manner as real-time energy market costs, but rather allocated to load on a pro-rata basis as an out-of-market charge.

Day-Ahead Load Response Program

From June 2005 through its expiration in June 2012, demand response in New England was also permitted to participate in the energy market via the Day-Ahead Load Response Program. After the day-ahead energy market cleared, the ISO would accept the offers of demand response resources in the Day-Ahead Load Response Program whose offers were less than the clearing price in the day-ahead energy market, but this process occurred afterwards, outside the clearing mechanism of the day-ahead energy market and did not impact its results. To the extent that any demand response was selected by the Day-Ahead Load Response Program and then proceeded to produce the reduction that cleared in the specified hours as obligated, the demand response would have an impact on what the real-time energy price would have been, but had no effect on the day-ahead energy price.[28]

Demand response resources that were selected by the Day-Ahead Load Response Program were paid the day-ahead clearing price for the amount selected. However, the rules permitted any resource that was selected to deliver more reduction in real-time than the amount obligated by the Day-Ahead Load Response Program. Any amount reduced from the baseline beyond that which cleared in the Day-Ahead Load Response Program was paid at the real-time energy price.

Predictably, participation in the Day-Ahead Load Response Program was at its peak during the summer of 2008, at a time when both the program was relatively mature, and energy prices were high. The maximum amount of Day-Ahead Load Response Program that was interrupted in any one hour was 884 MW in July 2010. Since that time, the maximum amount interrupted in any one hour has been much lower, once reaching 450 MW in July 2011, but otherwise never exceeding 310 MW.

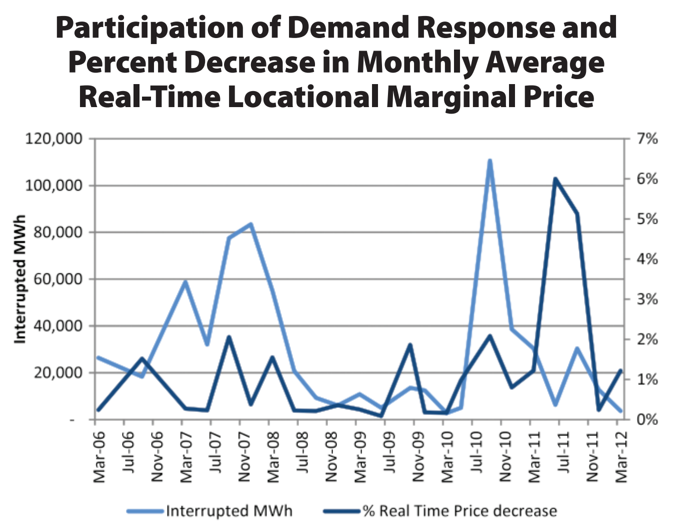
28 In New England, the day-ahead market accounts for approximately 93 percent of the total load on an annual basis from 2009-2011 (New England 2011 Annual Markets Report, page 40). Price reductions in the day-ahead market, in general, have a much greater impact on consumer costs than price reductions in the real-time market.

The payment rate in the Day-Ahead Load Response Program was the day-ahead energy price in the zone where the responding customers were located (Section III.E.2 of Appendix E to Market Rule 1). As in the case of PJM, real-time energy prices in New England incorporate locational marginal pricing to reflect transmission congestion between zones.

Because the Day-Ahead Load Response Program cleared outside of the day-ahead energy market, the costs were not allocated to load in the same manner as the day-ahead energy costs. As in the Price Response Program, they were allocated to load on a pro-rata basis as an out-of-market charge. However, all load does benefit from a small reduction in the real-time energy market prices. In Table G of their Semi-Annual Report on Load Response, the ISO-NE estimated the impact of demand response participation on real-time energy market prices. The value has generally ranged from $0 - $2/MWh, with a brief spike in the summer and fall of 2011 to just over $4/MWh.

To eliminate the effect of the changing value of the real-time price over time, we compared the percentage real-time price reduction with the amount of interrupted megawatt hours. No clear correlation is present. In those months where interruptions spiked, percent real-time price decreases may or may not spike, and large percent real-time price reductions are present in certain months where interruptions were moderate. This is an expected result, as we know that many other factors on the system will impact the real-time price, most notably the other resources that are in the supply stack at that time. Although it is difficult to predict the *amount* of the decrease, we note that the participation of demand response does always create a *decrease* in the real-time price.

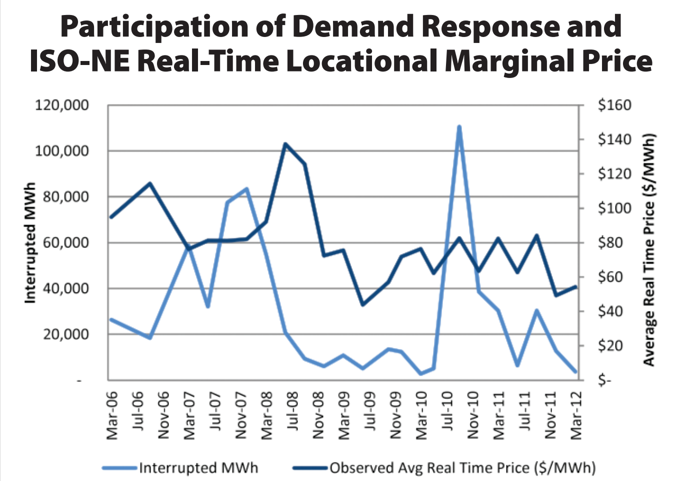
Figure 22



Source: (ISO-NE Semi-Annual Load Response Reports, Table G)

It is also difficult to draw a match directly between the amount of load interrupted and the value of the real-time locational marginal price. We cannot say clearly that more load will be willing to be interrupted when real-time locational marginal prices are the highest, as shown in Figure 23, below. Although interrupted load (as measured in interrupted megawatt hours per day) spiked in September 2010, it does not seem to be in response to dramatic change in real-time prices. Conversely, average monthly real-time prices were at their highest level in the summer of 2008, but demand response interruptions were falling.

Figure 23



Source: (ISO-NE Semi-Annual Load Response Reports, Table G)

We do not see a direct correlation between the participation of demand response and real-time energy market prices. It is clear that there are other factors affecting the decision of DR providers in New England to reduce load for a particular price.

In contrast to the capacity market programs, very little customer demand has been willing to participate in the region's energy market. That is, only a much smaller portion of the demand response that is willing to provide capacity in return for a monthly payment is also willing to offer in a price for which they would be willing to reduce load during the following day. It is unclear why this is the case. One might expect a purely financially rational customer to offer to reduce some amount of their demand every day, for a sufficiently high price. Perhaps the potential of reducing load in any set of hours where day-ahead prices spike is too risky for most customers or perhaps the specific rules of participating in the energy market create barriers to participation. We see a similar trend in PJM, where more than 90 percent of demand response revenue is derived from the capacity market.

There are reasons to suspect the gap between participation in the capacity markets and the energy markets is rational. The majority of demand response participation in recent years has been from DR providers aggregating multiple customers. The availability of a consistent monthly reserve payment in response for demand reduction or dispatchable on-site generation during a small number of mostly peak-load hours is a business model that has been successful. Likewise for other ancillary services that offer a consistent monthly reserve payment. However, similar demand response programs that offer unpredictable energy market prices have not been successful. Relying upon energy market prices that may or may not rise to sufficient levels on a regular basis has proven too risky a strategy for most demand response customers and their DR providers.

C. New York Independent System Operator Day-Ahead Demand Response

In New York State, demand response providers can participate in the day-ahead energy market via NYISO's Day Ahead Demand Response Program. Much like generation owners, demand response providers submit offers to curtail at a specified price one day in advance. The minimum offer price for demand response has varied over the years; it was $50/MWh for a number of years, and is currently at $75/MWh. These minimum offer price limits for demand response are intended to ensure that demand response only participates in hours where the demand reduction would be expected to occur for purely financial reasons; that is, to reduce load in high-priced hours. If the day-ahead offer clears, the demand response provider is expected to reduce load by the cleared amount in the specified hours. If they do not, they are charged for that amount of energy at the higher of the day-ahead or real-time price at their location.

In the most recent *Annual Report on Demand Response*, NYISO states that only one resource made offers in the Day Ahead Demand Response Program between September 2010 and August 2011. The offer was for only one day, and it cleared in only two hours, for an amount less than 3 MW. Although the number of hours has declined from that reported in prior years, the amount of participation – just under 3 MW – has remained consistent.

Table 10

NYISO Day Ahead Demand Response Total Scheduled Hours

Year	2007	2008	2009	2010	2011
Scheduled Hours	2,509	5,123	1,062	134	2

Sources: (NYISO Annual Report on Demand Response 2010-2011); (Annual Report on DSM Activity 2007-2009); (FERC Docket ER01-3001)

While the number of scheduled hours seems to fluctuate with average energy prices – 5,123 in 2008 and steadily dropping ever since – the amount of participation has remained very small at less than 10 MW, and in recent years only 2 MW on average for any one hour. It appears that very few customers are willing to participate when wholesale energy prices are quite low, and only very rarely peak to values at which they would be willing to interrupt. If these higher prices occurred with greater frequency, we would expect more demand response providers and more customers to explore the option of responding, and being paid for those interruptions.

D. Order No. 745

FERC Order No. 745 began as a Notice of Proposed Rulemaking in November 2009. FERC issued the final rule in March 2011 after numerous rounds of comments and two technical conferences. FERC Order No. 745 established several guidelines for demand response participation in energy markets:

1. Demand response can provide benefits similar to generation resources
2. Payments to demand response resources should be at the full market price of energy
3. A net-benefits threshold needs to be established
4. Demand response should be able to set the energy market clearing price, if it can provide the next available megawatt in economic order

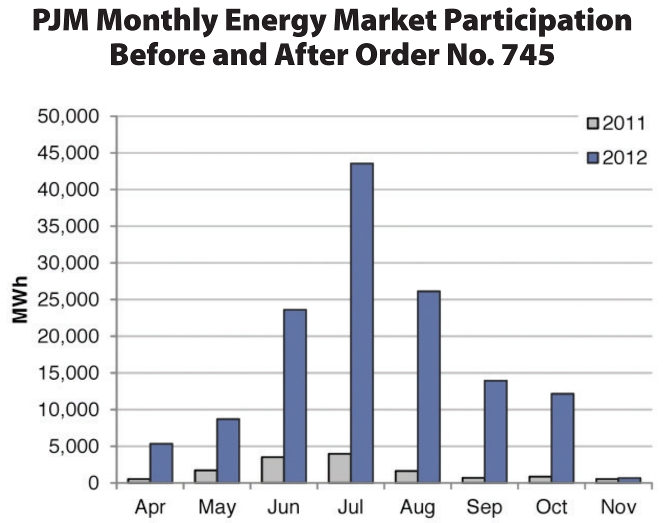
Although certain areas of the country were already paying demand response the market clearing price for participation in the energy market, full compliance with Order 745 will expand this practice to the entire country (excluding ERCOT in Texas, which is outside of FERC jurisdiction). This level of geographic scope has not yet been achieved. The order also took the first step towards evaluating those hours in which the participation by demand response reduces energy market costs for all ratepayers, both those that are demand response customers and those that are not.

One of the difficult issues that Order 745 resolved is that demand response compensation must be based on the full energy market clearing price in the relevant market. A wide range of opinions from noted economists were presented to FERC during its deliberations. FERC established that demand response resources provided net benefits to all energy market consumers by lowering the energy market clearing price whenever energy prices were above a certain threshold. FERC then reasoned that that any costs associated with the participation of demand response resources should be distributed among all energy market participants. The demand response resources should be compensated at the energy market clearing price, similar to generation resources.[29]

Most regions are still developing their compliance plans for Order 745 and it is uncertain what the long-term impacts will be. If regions with longstanding energy market participation offer any guidance, however, the amount of participation may be small relative to the amount of demand response willing to provide capacity services, at least until energy prices begin to rise again. The Northeast has shown that at prices in the range of $100/MWh or above, as they were in 2007 and 2008, numerous customers are willing to reduce demand for an energy payment. This was particularly true for PJM where there was substantial demand response participation in its Economic program when demand response resources were paid an incentive (the full energy market payment) whenever the market clearing price exceeded $75/MWh. New England also saw some participation when energy market prices were high. When PJM removed its incentive in late 2008, participation began to drop and continued at low levels for the next several years, a time that also saw substantial decreases in energy clearing prices.

However, since PJM implemented the Order 745 revisions to its Economic DR program and began compensating demand response providers at the full energy market price under essentially all hours, participation by demand response resources has rebounded, even though energy prices have been at historic lows in 2012. Figure 24, below, shows that dramatic increase in 2012 relative to 2011. New England implemented a transitional compliance with Order 745 beginning in June 2012. To date, participation rates are similar or lower than what was observed under the Day-Ahead Load Response Program in 2011. Full compliance with the order is scheduled for June 2017.

Figure 24



Sources: (PJM 2011 DSR Activity Report); (PJM 2012 DSR Activity Report)

29 This element of the FERC decision in Order 745 is being appealed to the Federal Circuit Court in Washington, DC.

6. Demand Response as a Provider of Ancillary Services

Ancillary services provide the resources to reliably maintain the balance between generation and load, and may include near-instantaneous regulation, load following or fast energy markets, and provision of reserves for contingency conditions. Due to the random variations in system load and generator output, these services are required to maintain reliability within a service area, and become especially critical as the proportion of variable resources on the grid increases.

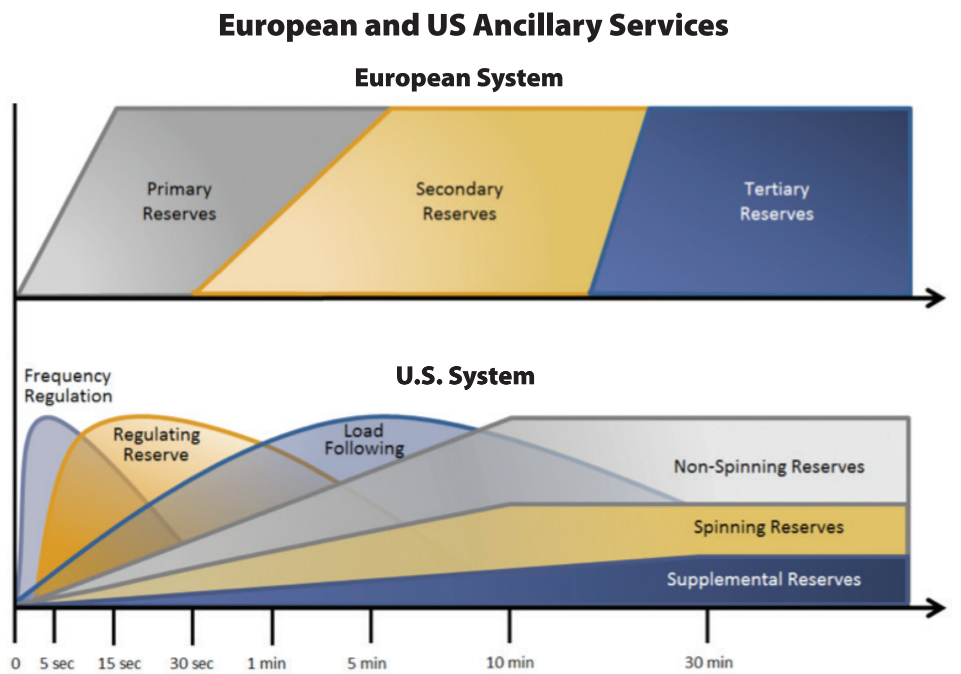
Historically, only generators provided these services, but today demand response is capable of providing many of these services as well, and often in a cost-effective manner. In regions where wholesale ancillary services markets exist, demand response resources receive payments in return for providing these services, while in areas without wholesale markets, vertically integrated utilities must themselves weigh the benefits against the costs. Figure 25 illustrates the various types, response times, and duration of ancillary services in Europe and the United States (although definitions of each service may differ slightly between regions).

Reserve resources are designed to respond to contingency events. Spinning reserves are capable of responding within 10 minutes to compensate for the loss of a large generator or transmission line, while non-spinning reserves respond within 30 minutes to replace the reserve capacity of a spinning reserve resource.

Balancing services such as frequency regulation, regulating reserves, and load-following services, refer to the quick adjustments made to compensate for the random fluctuations in system load and generator outputs. Regulation services must be performed on a near-instantaneous basis, while load-following is slower, occurring within approximately 10 minutes. In areas with wholesale markets, load-following is performed by the real-time energy market (Kirby 2007).

As is the case with generators, not all forms of demand response are capable of providing every form of ancillary service, as demand response resources vary in their ability to perform according to the differing requirements for response time, duration, and whether an increase in demand may also be required. Traditional dispatchable demand response resources tend to be well-suited to providing non-spinning reserves, as the 30 minute response

Figure 25



Sources: Adapted from Ruggero Schleicher-Tappeser diagram based on (NERC 2011) and (Amprion 2013).

time is easiest to accommodate. However, new types of automated demand response – including those with an element of storage – are increasingly providing regulation and load following functions. A brief description of the ancillary services that may be provided by demand response is given in Table 11, together with the average and maximum prices received in 2005 per megawatt of capacity provided for a one-hour duration. These prices are derived from California, ERCOT, and NYISO only, and may not be reflective of current market prices.

Frequency regulation and regulating reserves require fast, accurate response and are the most expensive ancillary services to provide, making them very attractive to resources that are capable of supplying such fast response and potentially providing a significant source of revenue to demand response providers while generating cost savings for system operators (Todd, et al. 2009).

A. Reserve Resources

ERCOT

Texas, similar to other regions of the United States, has a history of load interruption programs offered by vertically integrated utilities to their (usually) largest customers. The interruptible load programs might offer an overall lower rate, a special rebate mechanism, or a payment formula for each interruption. These types of programs are discussed in Section 4 of this report.

Table 11

Ancillary Services That May Be Provided by Demand Response

Service	Service Description			
	Response Speed	*Duration*	*Cycle Time*	*Price Range* (Average, Max) $/MW-hr*
Normal Conditions				
Frequency Regulation	Online resources, on automatic generation control, that can respond rapidly to changes in frequency.			
	<30 seconds	*Seconds to Minutes*	*Seconds to Minutes*	
Regulating Reserve	Online resources, on automatic generation control, that can respond rapidly to system-operator requests for up and down movements; used to track the minute-to-minute fluctuations in system load and to correct for unintended fluctuations in generator output.			
	4 Seconds to 5 minutes	*Minutes*	*Minutes*	*$35-$40 $200-$400*
Load Following	Similar to regulation but slower. Bridges between regulation service and hourly energy markets. This service is performed by the real-time energy market in regions where such a market exists.			
	~10 minutes	*10 min to hours*	*10 min to hours*	-
Contingency Conditions				
Spinning Reserve	Online generation, synchronized to the grid, that can increase output immediately in response to a major generator or transmission outage and can reach full output within 10 min.			
	Seconds to <10 min	*10 to 120 min*	*Hours to Days*	*$7-$7 $100-$300*
Non-Spinning Reserve	Same as spinning reserve, but need not respond immediately. Resources can be			
	<10 min	*10 to 120 min*	*Hours to Days*	*$3-$6 $100-$400*
Replacement or Supplemental Reserve	Same as supplmental reserve, but with a 30-60 min response time; used to restore spinning and non-spinning reserves to their pre-contingency status.			
	<30 min	*2 hours*	*Hours to days*	*$0.4-$2 $2-$36*

Prices are approximate ranges in $/MW-hr for 2005, including California, ERCOT, and New York.
Source: Adapted from (Kirby 2007), and (NERC n.d.)

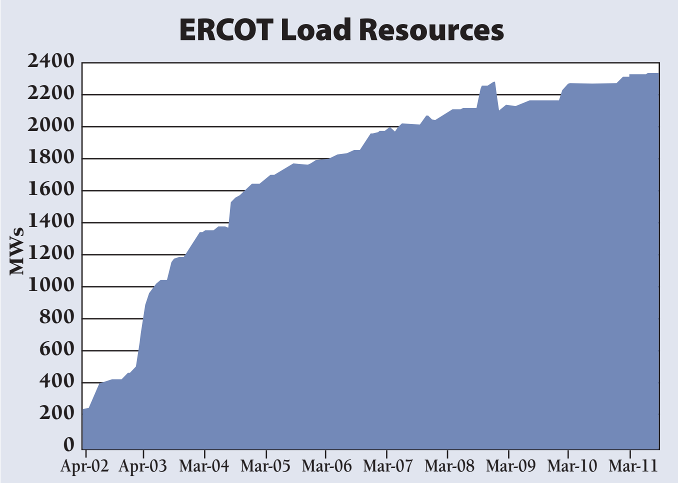
Those programs continue today for many transmission distribution utilities (the wires companies that remain after restructuring in Texas). In 1996, the Electric Reliability Council of Texas (ERCOT) became the Independent System Operator for 85 percent of Texas load.

We review two ERCOT programs that provide for demand response resource participation: A ten-minute spinning reserve program, Load Resources, and a ten-minute non-spinning reserve program, Emergency Interruptible Load Service.

Load Resources

One of the early services developed for responsive loads was essentially a spinning reserve service called "Load Acting as a Resource." From a start of 200 MW in May 2002, this program grew to over 1,600 MW of eligible resources by November 2004, although only a maximum of 0.50 percent of the spinning reserves requirement or 1,150 MW, is permitted to be cleared by the system operator (ERCOT 2005). This spinning reserve service could be activated either through an automatic frequency trip or directly dispatched by ERCOT operations. Each resource would need to respond (drop load) within ten minutes after dispatch and be able to return to 95 percent of its pre-event load within three hours of the event ending.

Figure 26



Source: (ERCOT Staff 2011)

The program continues today, although the name has been changed to Load Resources. Figure 26 shows the steady development of Load Resources from 2002 through 2011. The total quantity of eligible Load Resources currently exceeds 2,000 MW. These resources offer on a daily basis with ERCOT selecting the lowest cost offers up to the 1,150 MW limit.

As shown in Figure 27, below, Load Resources have been deployed 21 times between April 2006 and August 2011. In some of these cases Emergency Interruptible

Figure 27

ERCOT Load Resources

Day	Date	Time	Type of Deployment	Season	EILS TP
Mon	4/17/2006	15:34	EECP Step 2 Systemwide VDI	Spring	BH2
Tue	10/3/2006	17:37	Systemwide VDI for frequency restoration	Fall	BH3
Fri	12/22/2006	2:54	UF Event followed by VDI for frequency restoration	Winter	NBH
Mon	7/2/2007	19:38	Systemwide VDI for frequency restoration	Summer	BH3
Wed	9/5/2007	7:57	Systemwide VDI for frequency restoration	Summer	NBH
Wed	12/12/2007	1:56	Systemwide VDI for frequency restoration	Winter	NBH
Tue	2/26/2008	18:49	EECP Step 2 Systemwide VDI	Winter	BH3
Sun	3/16/2008	11:37	UF Event, frequency < 59.7 Hz	Spring	NBH
Mon	8/11/2008	17:14	Systemwide VDI for frequency restoration	Summer	BH3
Tue	12/16/2008	15:49	Systemwide VDI for frequency restoration	Winter	BH2
Sat	1/9/2010	10:32	Systemwide VDI for frequency restoration	Winter	NBH
Sat	5/15/2010	16:14	UF Event, frequency < 59.7 Hz	Spring	NBH
Wed	6/23/2010	15:20	UF Event followed by VDI to selected QSEs for frequency restoration	Summer	BH2
Fri	8/20/2010	15:28	Systemwide VDI for frequency restoration	Summer	BH2
Wed	11/3/2010	10:21	UF Event followed by VDI to selected QSEs for frequency restoration	Fall	BH1
Wed	2/2/2011	5:20	EEA Level 2A Systemwide VDI	Winter	NBH
Wed	3/23/2011	14:47	UF Event (partial), frequency dropped to near 59.7 Hz	Spring	BH3
Tue	4/5/2011	22:02	UF Event (partial), frequency dropped to near 59.7 Hz	Spring	NBH
Wed	5/19/2011	14:08	UF Event (partial), frequency dropped to near 59.7 Hz	Spring	BH2
Thu	8/4/2011	14:32	EEA Level 2A Systemwide VDI	Summer	BH2
Wed	8/24/2011	15:11	EEA Level 2A Systemwide VDI	Summer	BH2

- **Six of 21 deployments occurred during summer peak hours**
- **Eight of 21 deployments occurred during non-business hours**
- **ERCOT needs operational DR any time (not just on peak)**

EILS also deployed

Source: (Wattles 2011)

Load Service resources were also deployed.

Load Resources are compensated based on day-ahead offers of availability. A Load Resource may have small variations to the amounts of available responsive load on specific days. However, if the resource fails to provide its stated available amount on two occasions, it will be suspended from the program for six months (Anderson 2011).

Emergency Interruptible Load Service

After an April 2006 curtailment event, ERCOT and the Public Utilities Commission of Texas determined that additional load resources would help limit the potential for rolling black outs. In the fall of 2007, ERCOT implemented the Emergency Interruptible Load Service program. In this program, demand response resources are available to ERCOT for dispatch within ten minutes at a fixed contract price through a solicitation administered by ERCOT. The solicitations cover three different seasons and each solicitation has four categories of hours that can be selected. Until the spring of 2011, the program was capped at 1,000 MW and the annual cost could not exceed $50 million.[30]

ERCOT procured Emergency Interruptible Load Service resources for three seasonal time periods: the summer season of June 1 through September 30; the fall/winter season of October 1 through January 31; and the winter/spring season of February 1 through May 31. Each seasonal period is a separate procurement, with some demand response resources offering into all three procurements.

For each seasonal procurement, a demand response resource can offer for a specific set of operational hours:

- Business Hours 1 (~490 hours) 0900 through 1300, Monday-Friday;
- Business Hours 2 (~258 hours) 1400 through 1600, Monday-Friday;
- Business Hours 3 (~420 hours) 1700 through 2000; Monday-Friday;
- Non-Business Hours (~1900 hours) are all other hours

Demand response resources have the option to offer for some, or all, of the blocks of hours. Figure 28 through Figure 30 show the history of the seasonal Emergency Interruptible Load Service procurements by megawatts (bar columns) and price (lines).

The offers made by each demand response resource are the basis for payment. The demand response resource receives the offer price per megawatt for each hour that it is available. For Business Hours 1 DR resources in the

Figure 28

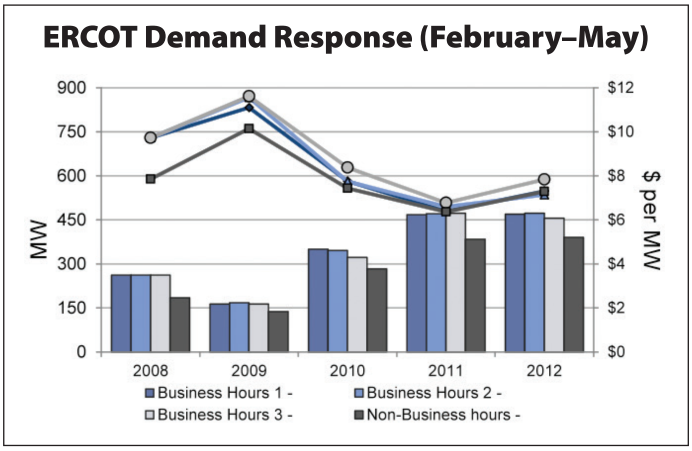


Figure 29

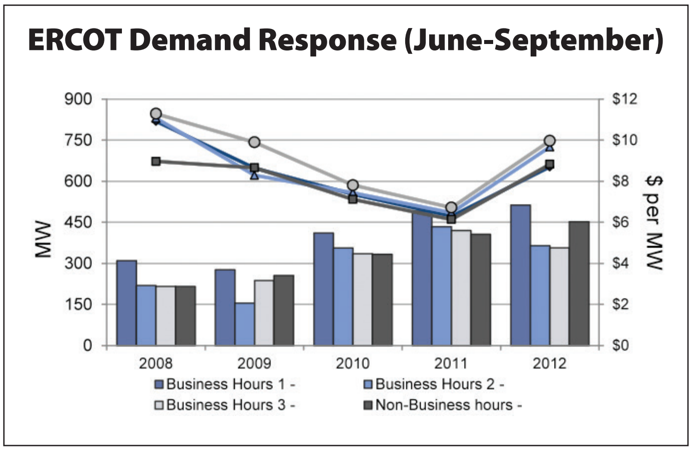
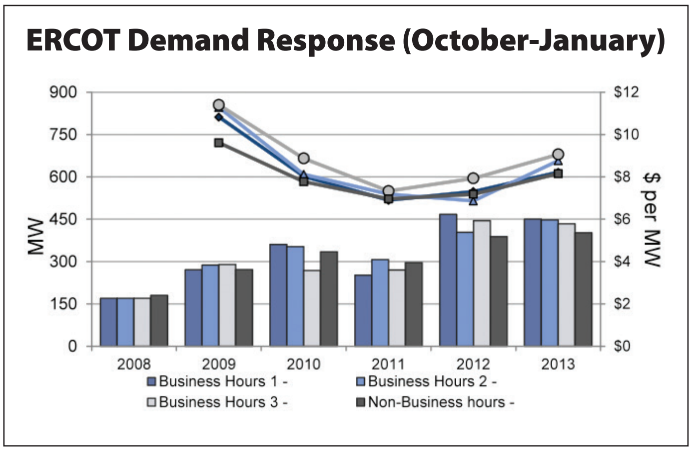


Figure 30



30 These limitations never became binding. The largest quantity purchased was under 600 MW and the annual costs never exceeded $25 million. (Rulemaking to Amend Subst. R. 25.507, Relating to Electric Reliability Countil of Texas (ERCOT) Emergency Interruptible Load Service (EILS) 2012)

summer 2012 season, the average offer price was $8.70/MW. Multiply that rate by the 420 hours for a total payment for the four month summer season of $3,600/MW. As a comparison, Non-Business Hours DR resources in the same season would earn a total of $16,900/MW. The five-to-one difference in payments is roughly equal to the five-to-one difference in hours. When demand response resources are activated, they do not receive any additional compensation.

February 2, 2011

During an extreme cold weather event on February 2, 2011, ERCOT first deployed its Load Resources early in the morning (5:20 am) as its responsive reserves dropped below 1,750 MW.[31] ERCOT requested 888.5 MW of Load Resources and 881.7 MW responded. Thirty minutes later, an additional 140 MW of Load Resources that were not committed also responded to the system-wide request from ERCOT operators.

At 5:48 am, ERCOT activated the obligated Emergency Interruptible Load Service resources (384 MW); these were the resources committed for non-business hours in the winter/spring season. At 8:53 am, ERCOT dispatched newly obligated Emergency Interruptible Load Service resources; these were the resources committed for the Business Hours 1 time period in the winter/spring season. Some additional Emergency Interruptible Load Service resources (83 MW) that were not obligated to respond also made themselves available. Due to the severity of system conditions (as more and more generators failed to operate for a variety of reasons) the Emergency Interruptible Load Service resources remained dispatched for 28 hours. The event ended at 10:00 am February 3. During that time, some Emergency Interruptible Load Service resources briefly resumed electricity consumption to maintain critical equipment or due to health concerns. Nonetheless, the performance of the Emergency Interruptible Load Service resources was exemplary.

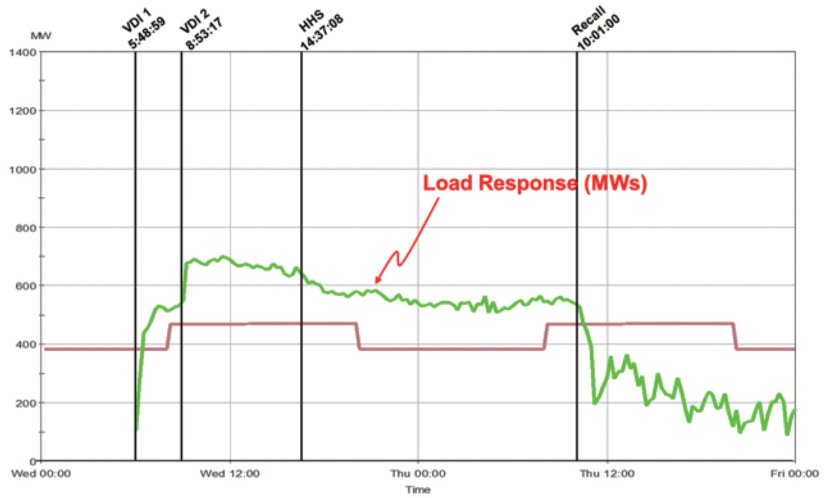
The average Emergency Interruptible Load Service obligation for the entire 28-hour event was 462.8 MW; the average actual load reduction for the entire event was 577.7 MW. Emergency Interruptible Load Service resources provided 135 percent of their obligated MW, on average, over the 28 hours. The peak reduction was 692.2 and occurred during the first few hours after the Business Hours 1 resources were activated on February 2.

The overall performance of the Emergency Interruptible Load Service resources on February 2, 2011 demonstrated that Emergency Interruptible Load Service resources are reliable during an emergency event in a non-traditional peak load season (winter). Although the deployment of these resources did not prevent subsequent load shedding, the magnitude of the load sheds was reduced by the approximately 500 MW of Emergency Interruptible Load Service resources that responded and, in essence, were paid to shed their load.

Because the February 2 event occurred at the start of the 2011 winter/spring season (February 1-May 31) and the entire eight-hour obligation had been exceeded, ERCOT made a supplemental Emergency Interruptible Load Service solicitation in March for the remaining portion of the winter/spring season for 2011. Without the supplemental solicitation, ERCOT would not have had access to any Emergency Interruptible Load Service resources for a possible system-wide event in April or May.

Figure 31

ERCOT Emergency Interruptible Load Service Deployment, February 2- 3, 2011



Source: (ERCOT 2012)

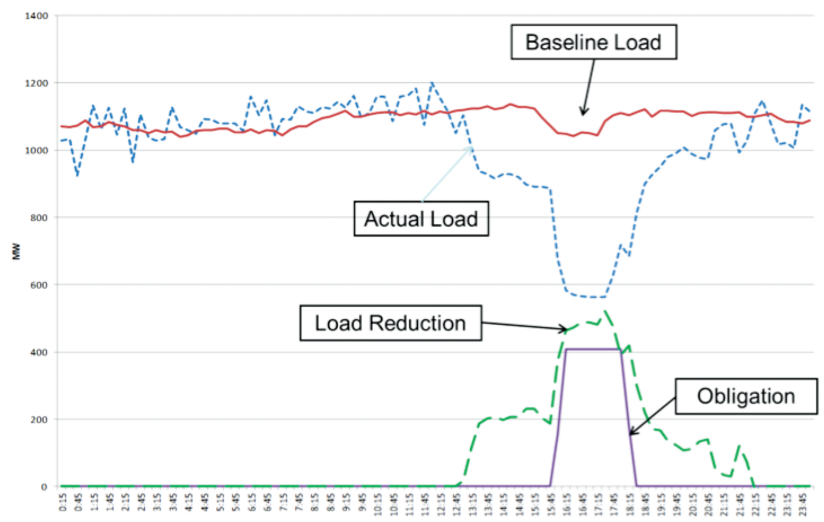
31 The primary cause for this event was determined to be the forced outages of a number of gas-fired combined cycle plants that were not fitted with adequate freeze protection equipment. Minimum freeze protection standards were subsequently imposed by ERCOT.

August 4, 2011

On August 4, 2011, a hot summer day, ERCOT experienced level 2 emergency system conditions that triggered the deployment of Load Resources and Emergency Interruptible Load Service resources. Both the Load Resources and the Emergency Interruptible Load Service resources performed as expected. Load Resources provided 863 MW of spinning reserves in response to the ERCOT dispatch signal (ERCOT Staff 2011). Emergency Interruptible Load Service resources exceeded their 400 MW obligation over the entire event and reached a maximum reduction of approximately 500 MW as shown in Figure 32.

Figure 32

ERCOT Emergency Interruptible Load Service Deployment, August 4, 2011



Source: (ERCOT 2012)

Changes to ERCOT's Emergency Demand Response Program

In the fall of 2011, after the events of February and August, ERCOT proposed changes to the rules governing the Emergency Interruptible Load Service program. The highlights of those changes include:

- Expanding the eligible resources to include distributed generation that can export to the grid and is not already an energy market resource
- An option to activate resources for more than eight hours during the seasonal contract period and provide additional compensation if those activations occur
- Retaining the 1,000 MW cap and $50 million annual compensation cap, while explicitly recognizing that ERCOT could request a modification to either cap as conditions warranted
- Modifying the performance test for qualified resources that are aggregations of many separate sites. The new test is the total performance of the aggregated amount, not each individual asset.
- Changing the name of the program from Emergency Interruptible Load Service to Emergency Response Service
- Developing a pilot program for Emergency Response Service resources that respond within 30 minutes of activation

The goal of these changes is to increase the quantity of resources available to ERCOT and to increase the flexibility/value of those resources to system operations.

The pilot program for resources capable of responding in thirty minutes or less is modeled after the current Emergency Response Service program in that resources commit to being available during a specific season across the same four segments of hours. One significant design change is that compensation for the thirty-minute resources will be based on the highest clearing price in the solicitation, not on each resource's offer price. Table 12 shows the results of the first solicitation for the

Table 12

Thirty-Minute Emergency Response Service Pilot

(July-September 2012 Procurement)

Time Period	Bus. Hrs. 1 HE 0900 – 1300, M-F except Holidays	Bus. Hrs. 2 HE 1400 – 1600, M-F except Holidays	Bus. Hrs. 3 HE 1700 – 2000, M-F except Holidays	Non-Bus. Hrs. All Other Hours
Capacity Offered	19.4 MW	16.25 MW	15.80 MW	9.5 MW
Capacity Procured	19.4 MW	16.25 MW	15.80 MW	9.5 MW
Number of Loads Procured (number of aggregations)	11 (5)	8 (4)	7 (3)	8 (3)
Clearing Price	$11.00	$16.00	$16.00	$11.00
Low $Offer Received	$5.00	$3.00	$5.00	$5.00
High $Offer Received	$11.00	$16.00	$16.00	$11.00
Highest $Offer Accepted	$11.00	$16.00	$16.00	$11.00

Source: Emergency Response Service Update – DSWG Presentation July 20, 2012 pilot program. It is interesting to note that most of the thirty-minute offers were from resources that had never participated in the ten-minute Emergency Response Service or Emergency Interruptible Load Service programs.

PJM

PJM operates a synchronized (spinning) reserve market, under which there are two primary types of resources: Tier 1 (economic) and Tier 2 (non-economic). Tier 1 resources are online units that are following economic dispatch but not operating at capacity, and are thus able to increase output within 10 minutes of a dispatch signal. Tier 2 resources may include demand response resources and are offered into the synchronized reserve market and cleared. Demand for Tier 2 reserves is determined by subtracting the amount of forecast Tier 1 available from each zone's reserve requirement. Recent demand has averaged 388 MW for all cleared hours.

Since August 2006, demand response has been permitted to participate in PJM's synchronized reserve market. Demand response has played a small, but significant role in this market, as it is one of the most cost-effective reserve resources and has resulted in lower prices than would otherwise have occurred (Monitoring Analytics, LLC 2012b).

In January through September 2012, demand response resources comprised 36 percent of all cleared Tier 2 synchronized reserves – an increase from 21 percent the previous year. In six percent of hours in which the synchronized reserve market was cleared, all megawatts cleared were demand response resources. The average price during the hours when all cleared resources were demand response was less than $1, while the average price for all hours was nearly $5 (Monitoring Analytics, LLC 2012b). Details on the average synchronized reserve market clearing price (SRMCP) and price when all reserves were demand response (labeled "DSR") is shown in Table 13 below.

Table 13

PJM Synchronized Reserve Clearing Prices and DR Participation

Year	Month	Average SRMCP	Average SRMCP when all cleared synchronized reserve is DSR	Percent of cleared hours all synchronized reserve is DSR
2010	Jan	$5.84	$2.03	4%
2010	Feb	$5.97	$0.10	1%
2010	Mar	$8.45	$2.01	6%
2010	Apr	$7.84	$1.86	17%
2010	May	$9.98	$1.68	15%
2010	Jun	$9.61	$0.74	9%
2010	Jul	$16.30	$0.79	7%
2010	Aug	$11.17	$0.93	12%
2010	Sep	$10.45	$1.15	12%
2011	Jan	$9.31	$0.10	0%
2011	Feb	$10.58	NA	0%
2011	Mar	$9.70	$2.04	2%
2011	Apr	$12.64	$1.84	10%
2011	May	$8.64	$1.71	14%
2011	Jun	$9.05	$1.18	10%
2011	Jul	$12.33	$0.62	6%
2011	Aug	$8.25	$0.78	7%
2011	Sep	$9.05	$1.73	15%
2012	Jan	$5.47	$1.71	11%
2012	Feb	$4.90	$1.78	24%
2012	Mar	$5.60	$1.40	6%
2012	Apr	$5.01	$0.91	4%
2012	May	$9.29	$0.54	2%
2012	Jun	$4.05	$0.43	1%
2012	Jul	$9.88	$0.10	0%
2012	Aug	$5.61	$0.60	1%
2012	Sep	$4.74	$1.23	2%

Source: (Monitoring Analytics, LLC 2012b)

Independent System Operator of New England

The ISO New England administers the Locational Forward Reserves Market that procures, by season, the majority of daily operating reserves needed to operate the system reliably. In this context, operating reserves refer to both synchronized and non-synchronized reserves that can be called upon in the event of a system contingency.

To date, demand response has not been eligible to participate in the reserves market in New England. Pilot programs have been completed, but the results were inconclusive, and market rule development for the inclusion of demand response has not yet been pursued.

New York Independent System Operator

NYISO describes its Demand Side Ancillary Services Program in its annual report on demand response activity. Since June 2008, NYISO has allowed demand response to provide operating reserves or regulation service, if the resources have the correct telemetry and response capabilities. Day-ahead offers must be submitted one day in advance, along with Day Ahead Load Response Program energy offers (for co-optimization) and the offer must include amount, price, and type of service must being offered: spinning reserves, non-spinning reserves, and/or regulation. Real-time offers can be submitted up to 75 minutes before the relevant hour.

However, there is no indication that demand response has actually participated in this program to date. In the most recent reports, NYISO indicates that there are no resources subscribed to the program and that they are still developing the detailed rules of participation.

B. Regulation and Load-Following Services

Balancing the normal fluctuations of energy supply and demand requires resources that can respond quickly. Balancing services (including near-instantaneous frequency regulation and slightly slower regulating reserves) and load-following services are designed for this purpose.[32]

Balancing Services for Renewable Energy Integration

Expansion of variable renewable capacity, particularly wind and solar energy, increases the need for flexible resources, particularly those with an element of storage, that are capable of maintaining the balance between generation and load under normal conditions. Regulation and load-following services differ from traditional demand response in that they are not activated in response to a contingency event, but rather enable the reliable functioning of the grid on a continuous basis. These services are gaining in importance as greater amounts of electricity are generated by renewables, and in several regions of the United States, demand response is beginning to be tapped as a potential low-cost provider of these services.

Loads that are best-suited to regulation and load-following are those that consume a large amount of energy (as opposed to operating primarily during peak hours), and particularly those that are coupled with a form of energy storage in order to permit the load to respond to system operator signals without impacting the load's primary function. These types of demand resources can be especially useful for increasing demand during periods of excess generation from wind and solar resources. Batteries represent a common mechanism for energy storage. Heating and cooling systems designed to absorb energy (as well as to heat and cool) are additional examples. Electric vehicles are likely to represent a large source of energy storage in the future, and this technology is being explored by a number of entities, including PJM through the Mid-Atlantic Grid Interactive Cars Consortium project from 2007 to 2010 (Carson 2012).

A growing trend in demand response is the use of thermal storage as a mechanism for providing regulation and load-following services to the grid. Thermal storage may be provided through hot water heaters, electric thermal storage heaters, ice-based air conditioning, and cold storage – all of which can be temporarily interrupted without much loss of functionality, due to the ability of the heat or cold to be retained by the appliance's thermal mass. Other forms of storage include compressed air storage (typically in rock formations) and water-based storage, whether in traditional grid-scale pumped storage hydro facilities or in distributed applications such as municipal or irrigation water systems that are capable of pumping large amounts of water during off-peak hours for later use.

Although experience with distributed thermal and water-based storage for regulation and load-following services is limited, preliminary evidence indicates that demand response balancing resources may offer a cost-

32 Load-following services are typically provided by the wholesale market in regions with real-time energy markets.

effective alternative to traditional generation, particularly as modern information technology continues to advance and penetrate the market. Moreover, some forms of demand response may be able to provide extremely fast and accurate regulation services, often exceeding that of traditional generation resources (Todd, et al. 2009). The size of the potential resource in both the residential and commercial/industrial sectors is vast due to the prevalence of electric water heating, space heating, air conditioning, and cold storage, but to date this resource remains largely untapped.

FERC Order 755 on Compensation of Regulation Services

The use of demand response for ancillary services, including regulation services, has been slow to spread in part due to market designs that implicitly or explicitly have favored traditional supply-side resources. Recent regulatory action on this front, particularly with regard to compensation of demand response resources, has begun to improve this situation. On October 20, 2011, FERC issued Order No. 755, regarding compensation of resources providing regulation service. FERC found that resources providing such services differ in their ramping ability and the accuracy of their response, yet compensation by system operators did not account for such differences. Thus FERC found rates unjust, stating that "current compensation methods for regulation service in RTO and ISO markets fail to acknowledge the inherently greater amount of frequency regulation service being provided by faster-ramping resources" (Order No. 755 2011). FERC therefore ordered system operators to base payment in part on the ability of each resource to respond to regulation signals quickly. Electric storage technologies such as batteries, thermal storage and mechanical flywheel storage can respond much more quickly than large power plants. FERC's order ensures that these resources would be paid a greater amount for the greater reliability value being provided to the system.

The development of compensation mechanisms that comply with Order No. 755 is still in progress, and other market participation rules have been changing rapidly over the course of the past few years. While there is some uncertainty as to what form the final market structures will take, lessons can be learned from the evolution of these rules and the degree to which they have fostered demand response participation. Examples of demand response's ability to provide regulation and load-following services in the United States are discussed in greater detail below, as well as brief commentary on the evolution of relevant market designs.

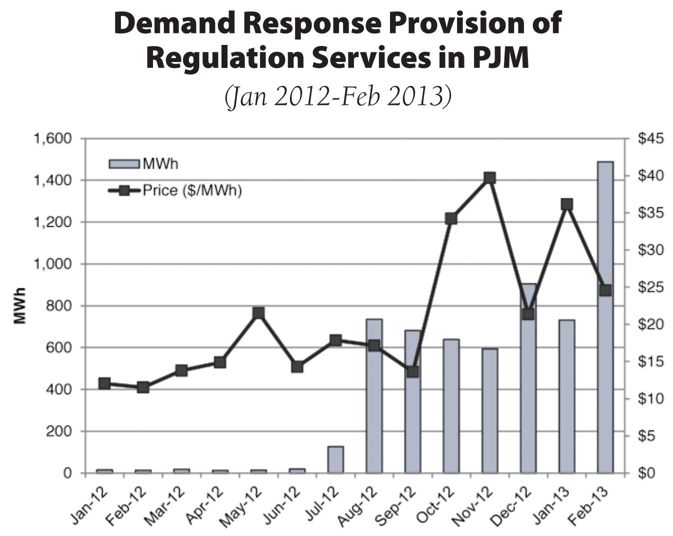
PJM

Despite being approved for participation in the regulation market in 2008, it was not until November 2011 that demand response resources were cleared in PJM. PJM's rules technically allowed up to 25 percent of the regulation requirement to be satisfied by demand resources, but until November 2011 other rules made participation of demand resources impractical.

PJM's rules that effectively prevented demand response from participating in the regulation market included a 1 MW minimum offer requirement, a prohibition on demand resources offering both economic and emergency demand reduction, and a restriction that demand response resources be represented by no more than one curtailment service provider. In particular, the 1 MW minimum size requirement acted as a barrier to entry for new demand response aggregators, especially in the residential sector where loads are small (Chatham, Baker and Miller 2012). The restriction requiring demand response to be represented by only one curtailment service provider further hindered the development of demand response, as some curtailment service providers are only active in a single market (such as the energy market), thus preventing the demand response resource from also participating in other electricity services markets. In November 2011, PJM members approved multiple rule changes, including reducing the minimum size requirement to 0.1 MW and allowing demand resources to be represented by more than one curtailment service provider.

Following the modification of these rules, demand response resources began to clear regulation. However, the amount of demand response that provides regulation is thus far quite small and has had little impact on the regulation market (Monitoring Analytics, LLC 2012). This is rapidly changing, however, as market participants gain greater experience and reach out to new customers. The amount of demand response participating in regulation has increased from 0 MW in 2011 to 5.8 MW providing nearly 1,500 MWh of regulation service as of February 2013. This growth is depicted in Figure 33 on the following page (McAnany 2013a, McAnany 2013b).

Figure 33

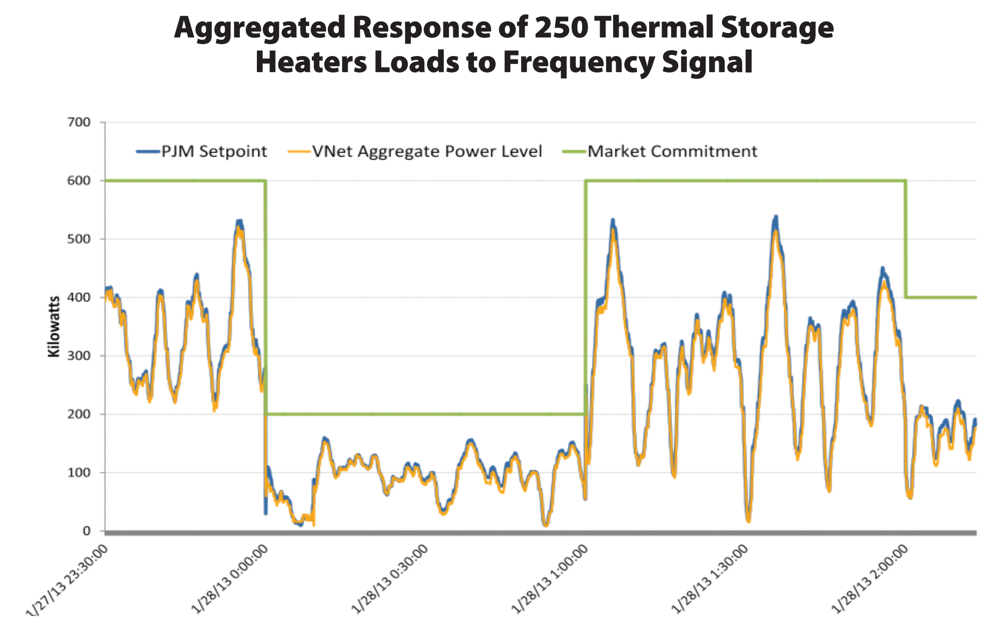


Source: (McAnany 2013a), (McAnany 2013b)

Payments have ranged from approximately $12 to $40 per megawatt hour, while the number of unique participating sites has increased steadily from only two in January 2012 to 52 in February 2013.

In addition, PJM is designing further market rule modifications, including compensation structures that reward participants based on the accuracy and speed of regulation response in order to better reflect the value of the resource as required by FERC Order No. 755. Such rule modifications may improve the incentives for demand response providers.

Figure 34



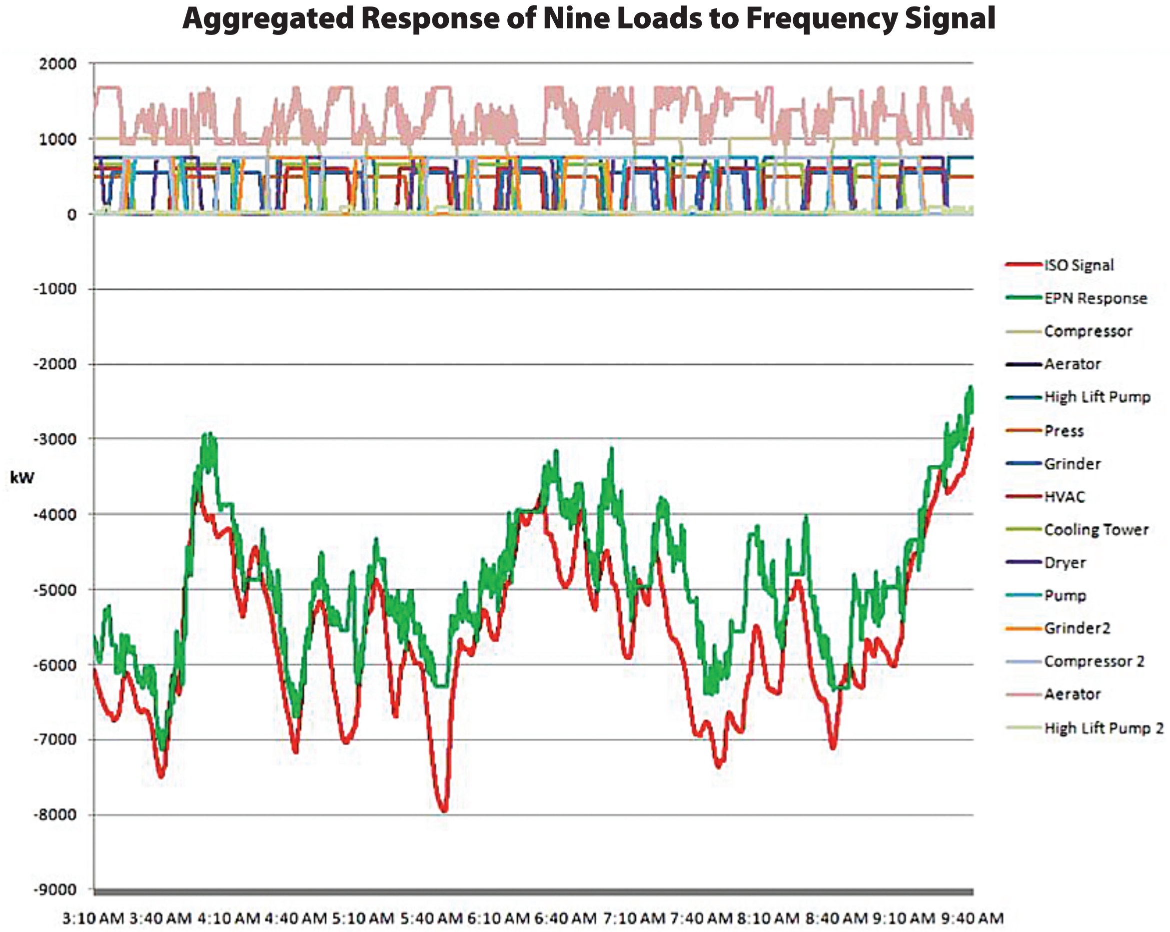
VCharge algorithm determines how much energy to buy and how much Regulation to provide based on the locational marginal price and the price of Regulation in each hour. Source: VCharge

Case Studies: VCharge and ENBALA

VCharge is one of the small number of pilot demand response providers of frequency regulation services that are now achieving commercial viability and expanding to other regions of the country. VCharge has aggregated 250 electric thermal storage heaters, both ceramic and hydronic, in 50 houses in a pilot for northeastern Pennsylvania. Aggregation enables customers to not respond when they face constraints, while still responding to the grid operator's requests for regulation services. VCharge operates this heater fleet commercially as a "Virtual Power Plant" that simultaneously buys energy during inexpensive hours and provides ancillary services to the grid operator. The vendor has both retrofitted existing electric thermal storage heaters with controls of its own design, and installed new electric thermal storage heaters for customers on the electric thermal storage tariff. VCharge supplies up to 600 kW of balancing services to PJM by responding to the area control error signal with a two-second response time. The large fleet enables VCharge to provide high fidelity in its tracking of the area control error signal as shown in Figure 34. Also, by leveraging its role as a licensed competitive energy supplier, VCharge is able to arbitrage energy prices, enabling it to provide the lowest-cost heating energy prices of approximately 40 such suppliers.

ENBALA Power Networks is another such curtailment service provider that, since November 2011, has aggregated customer loads to provide regulation services in PJM. ENBALA is partnering with Pennsylvania American Water, a large water and wastewater utility, to provide regulation services. Typical drinking water and wastewater plants have a peak load of 1.5 MW, with significant flexibility for pumping, aeration, and other processes in the short-term. When aggregated, these facilities can provide a sizeable amount of regulation service (Coulbeck 2012). Pennsylvania American Water currently provides approximately 400 kW of balancing services – either increasing or decreasing its energy consumption when

Figure 35



Source: (Coulbeck 2012)

requested. Providing regulation services has allowed the facility to earn enough revenue to reduce its total energy bill by two to three percent. Management at the water utility noted that the impacts from providing the services are negligible, stating "…we don't even know it is there. It is invisible to our operations."

As the number and diversity of loads aggregated increases, the precision of the response to the system operator's signal increases. This is demonstrated in Figure 35, where each individual load's response is shown in the top portion of the graph, while the lower green line shows correspondence of the nine aggregated loads' response to the system operator's signal (in red).

The success of the initial pilot program has led to plans to expand the program to additional pumps at the facility, as well as connect assets at other subsidiaries. ENBALA estimates that payments for such regulation services will range from approximately $35,000 - $50,000 per MW-year (ENBALA Power Networks 2012). The cost of the equipment installation on customer premises was estimated to be approximately $40,000 - $50,000 (Berst 2011).

Independent System Operator of New England

In November 2008, ISO-NE began conducting an Alternative Technology Regulation pilot program following FERC's Order 890 to remove barriers that prevent non-generating resources from providing regulation and frequency services. The pilot program's minimum size for entry was set at 0.1 MW, allowing small, innovative companies to participate, and providing a low-risk opportunity to evaluate new technologies. Over the course of the pilot program, the compensation mechanism for regulation resources has been revised in order to award payment based on the amount of regulation capacity the resource makes available as well as a "mileage payment," which compensates resources based on the performance of the resource. This compensation structure allows highly accurate and fast demand response resources to receive greater compensation than traditional generation resources that respond more slowly. Although participation in this pilot program was limited to just 13 MW, it has never been fully subscribed.

Through September 2011, maximum participation was only 3 MW.

Beacon Power was the largest pilot participant, and arguably the impetus for its creation. Beacon Power uses mechanical flywheels that absorb power off of the system to spin faster when the frequency regulation signal indicates this need. They can also slow their speed by pushing power onto the grid when signaled in that direction and the flywheels can perform these changes very quickly.

VCharge is another one of the small number of pilot participants and is now expanding to other regions of the country, as described in the section on PJM, above. In ISO-NE, VCharge partners with a local utility to both provide regulation service to the grid and purchase energy for the municipal utility's residential customers with electric thermal storage furnaces. These heaters utilize "SmartBricks" or "SmartHydro" – stacks of ceramic bricks or tanks of water with electric heating elements in them – to store up to 36 hours of energy. The heaters are remotely operated through a high-speed Internet connection to rapidly turn on or off based on the system operator's signal, typically within two or three seconds.

Currently VCharge operates nearly 1 MW of load in the ISO-NE pilot and is expanding into PJM and other markets. VCharge is currently also reviewing other types of storage mechanisms, including electric vehicles and ice-based air conditioning.

Through the ISO-NE pilot program, both Beacon Power and VCharge have shown that they can meet the technical requirements to provide small amounts of frequency regulation. In that sense, the pilot program has been a success. However, neither of these technologies has yet been scaled up in New England to meet a significant portion of the average hourly requirement of 60 MW of need for regulation in New England in 2011 (Internal Market Monitor of ISO-NE 2012).

It is also unclear if the business proposition can be successful for such alternative technologies in the frequency regulation market. The ISO-NE Internal Market Monitor publishes the regulation market supply stack in the 2011 Annual Market report, and it indicates that generation owners are willing to provide far more regulation than is needed at or near the current average price of $7/MWh. In other words, even small increases in regulation demand, should they occur, will not drive revenues for regulation suppliers in New England. The ISO-NE is working to comply with FERC Order 755 in 2015, and at that time additional revenues will be available to those resources who can respond more quickly (via "mileage payments") as the order requires. With another year or two of participation, it will become more clear if these revenues are sufficient to support such a business in New England.

Bonneville Power Administration

The Pacific Northwest is a region rich in hydropower, but without an organized power market. The Bonneville Power Administration, part of the US Department of Energy, owns the majority of the high voltage power lines in the region and provides transmission services to its customers such as interregional interconnections and maintaining electrical reliability and stability. The Bonneville Power Administration was originally created to market electricity from the Bonneville Dam, and now the agency markets approximately 35 percent of the electricity used in the Pacific Northwest from 31 federal hydro projects in the Columbia River Basin and several nonfederal power plants.

Wind power is growing rapidly in the Pacific Northwest, with the amount connected to BPA's transmission system increasing from less than 250 MW in 2002 to more than 4,500 MW in 2012. An additional 2,000 MW to 6,000 MW of wind capacity has been proposed to be added within the next fifteen years. Due to wind's rapid expansion, the Bonneville Power Administration is working to identify new cost-effective resources that will facilitate the integration of these and other variable resources, and is currently spearheading demand response pilot projects for regulation and load-following services.

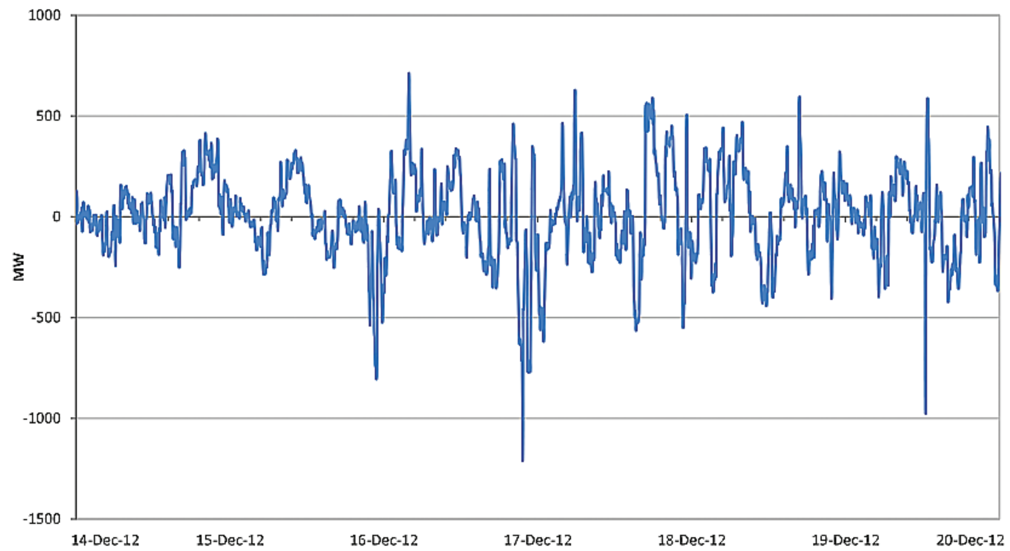
The Bonneville Power Administration has long used its vast hydro resources to provide load following services, but is now looking to demand response that is capable of increasing as well as decreasing load in under ten minutes. Such capabilities are needed every hour of the year, with an example of balancing services deployed shown in Figure 36.

In order to investigate the potential of demand response to provide load-following services, the Bonneville Power Administration conducted a two-year pilot from September 2010 to September 2012 to explore balancing capabilities of various loads, including water heating, space heating, and cold storage. The project was carried out by Ecofys in conjunction with multiple customer utilities, and with support from universities, national laboratories, the Northwest Power and Conservation Council, and private companies.

The balancing pilot program operated a 1.2 MW portfolio of assets at 130 customer sites spread over six

Figure 36

Example of Bonneville Power Administration Balancing Reserve Deployment in 2012



Source: Bonneville Power Administration

utilities, composed of refrigerated storage warehouses, commercial heating/ventilation/air-conditioning systems, electric water heaters, and electric thermal storage furnaces. These load types were chosen for evaluation in part because of the prevalence of these resources in the region and the relative low cost of installation.

Cold storage in particular is attractive due to the fact that more than 300 frozen food processing and storage facilities exist in the Pacific Northwest, and the economics of such projects tend to be favorable ($100-$500/kW upfront cost). Cold storage warehouses have the added benefit of controlling a large load in one facility (typically 200 kW of curtailment and 100-200 kW of increase), thereby reducing the number of parties involved in the coordination (Ecofys and Bonneville Power Administration 2012).

The technology and communication infrastructure used to integrate commercial and industrial cold storage facilities includes both a server for dispatching the event signal over the Internet and equipment located at each facility capable of monitoring and interfacing with the refrigeration system. This required the installation of hardware providing two-way communication regarding electricity consumption data in near real-time, and the ability to relay event signals to the centralized refrigeration control system (Ecofys and Bonneville Power Administration 2012).

Evaluation of load curtailment or increase was performed by comparing the demand data for the event to a baseline, which was developed from establishing the average shape of the participant's demand over the previous ten days, with a day-of adjustment in order to calibrate the baseline shape to the level of the event day demand (Ecofys and Bonneville Power Administration 2012).

Overall, the cold storage facilities successfully responded to the dispatch signals within ten minutes' notice and demonstrated capability of delivering the requested amount of response. However, the aggregate response level varied from event to event, due to the small number of facilities enrolled, highlighting the need for a larger portfolio of participating warehouses in order to smooth out resource variability. It is estimated that a portfolio of 10 MW or larger would compensate for the variability of individual resources due to seasonality, operation needs, and maintenance events (Ecofys and Bonneville Power Administration 2012).

Additional lessons were learned regarding seasonality of response, the need to coordinate response with peak demand penalties, and technical issues. The winter season reduced the ability of facilities to respond to dispatch signals, as cold storage facilities reduce their energy demand in the colder months. Complications were encountered when the program encouraged creation of new peak demand levels, which typically results in a pricing penalty for commercial customers. In some cases, technical difficulties were also encountered when attempting to curtail equipment that was already off, resulting in the perverse outcome of this equipment being switched on. In response to these problems, corrections to the control system and equipment were made in order to prevent future problems (Ecofys and Bonneville Power Administration 2012).

In the second phase of the project, the cold storage facilities delivered 101 percent of their curtailment goal amounts, but only 47 percent of their increase goal amount, in part due to a wiring problem that caused one facility to curtail load regardless of whether the signal was to increase or decrease load. Going forward, it is expected that performance incentives would motivate a faster resolution of such issues (Ecofys and Bonneville Power Administration 2012).

The pilot also tested hot water heaters, electric thermal storage furnaces, and commercial heating/ventilation/air-conditioning systems. The water heaters were capable of responding very quickly, but only provided a few kW per heater. In contrast, the electric thermal storage furnaces provide considerable energy storage, with peak input of 29 kW. However, the furnaces are highly seasonal and were not in place for testing for the bulk of the 2011/2012 winter. This portion of the pilot has been extended in order to gather better data. Finally, the heating/ventilation/air-conditioning systems were often unable to respond in less than 30 minutes due to the temperature control system, which relies on integral gain with the control loop, and slow thermostat update times (Ecofys and Bonneville Power Administration 2012).

Overall, the results of the project are encouraging and indicate that water heater, space heating, and cold storage loads are both capable and cost-effective resources for providing balancing services to the grid, while resulting in no reduction in the quality of service or comfort experienced by end-users. Moreover, water heating controls can also be used for the more traditional purposes of peak reduction and load shaping, thus enabling it to take advantage of multiple value streams. However, the pilot did highlight the need to align electricity pricing signals with the objectives of the load-following program, as well as the importance of forming a large enough portfolio of resources to smooth out individual load response variation, and the need for incentives to ensure proper equipment set-up and operation in order to quickly resolve technical problems (Ecofys and Bonneville Power Administration 2012).

The Bonneville Power Administration envisions that future demand response for balancing will be provided by a wide variety of assets, both large and small, and distributed geographically across the Pacific Northwest. The ideal asset will have low installation/enablement costs, short lead time for enablement, low operating and maintenance costs, easy measurement and verification of performance, and a long lifetime of assets. The Bonneville Power Administration has plans to explore additional pilots at data centers and other loads in the future. Additional utility pilots, particularly those focused on residential water and space heating, are ongoing. (Ecofys and Bonneville Power Administration 2012)

7. Lessons Learned and Future Directions

A. Lessons Learned

In the preceding sections of this report, we have described the numerous ways that demand response resources have participated as components of the US bulk power system. On a regional and sometimes state-by-state basis, demand response has provided substantial contributions to resource adequacy mechanisms as both a capacity and reserve resource. In wholesale markets, demand response has new opportunities to participate as an energy resource (both day-ahead and real-time). There are many new demand response applications being tested and developed that can provide specialized operational services (including load-following, frequency regulation, and special reserves) to system operators. We have found that demand response is reliable, can provide a significant amount of a region's resource adequacy needs, can achieve participation in market areas, can lower the cost of reliability, and, as one would expect, that the likelihood of financial reward matters.

Demand Response Can Perform Reliably

Table 14 is a summary of the overall performance of demand response capacity resources in New England that covers activations and tests over the last several years. While it is specific to New England, performance in other regions has shown similar results, some of which we have included in Section 4 of this report.

Demand Response Can Provide Significant Contributions to Resource Adequacy

Numerous regions around the country have acquired in excess of five percent of their resource adequacy requirement from demand response resources. As of 2010, PJM has achieved 10.5 percent of its peak load and New England has achieved 7.8 percent (FERC 2011). However, because resource adequacy is only one element of overall reliable service, both PJM and New England are examining ways to expand the demand response capacity product. As described in this paper, PJM has already created three demand response products to reflect the different capabilities of demand response resources; the three products have an increasing time period of availability and an increasing frequency with which the option can be called

Table 14

New England Capacity Resource Performance

	Real Time Demand Response		Real Time Emergency Generation	
Load Zone	**MW**	**Performance**	**MW**	**Performance**
Maine	278	100%	25	100%
New Hampshire	45	93%	33	98%
Vermont	33	100%	13	98%
Connecticut	261	72%	254	86%
Rhode Island	40	90%	56	88%
Southeastern Mass.	136	78%	37	86%
Western/Central Mass.	132	97%	577	94%
Northeastern Mass.	198	80%	78	89%
Total New England	1,124	86%	553	90%
Generation Fleet Average EFORd[33]		94.5%		
Quick Start Generation		Assumed 80% during planning		

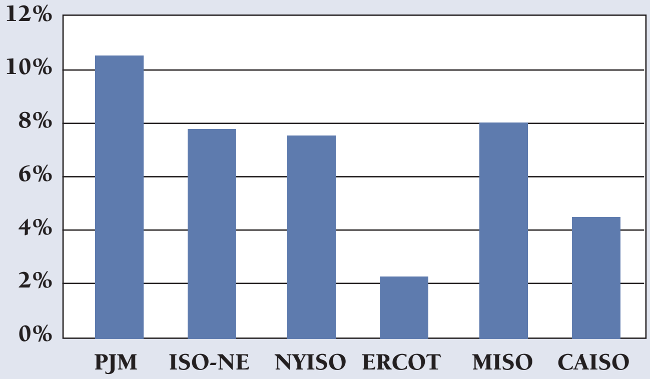
Source: (Scibelli 2012)

33 EFORd is the Equivalent Forced Outage Rate as measured when that particular unit is in demand.

(dispatch) for a correspondingly higher market payment. New England, as part of its Strategic Planning Initiative, is considering an incentive payment approach for all capacity resources (including demand response) based on operational performance (ISO-New England 2012).

Figure 37

Demand Response Resource Available as Percent of 2010 Peak Demand



Source: (FERC 2011)

These expanded services from demand response may also be procured in forward markets that are not capacity markets; they can be markets specifically designed for a particular service (such as fast start or frequent ramping). ERCOT has used demand response resources as a special category resource for system emergencies. It essence, the demand response in Texas helps maintain resource adequacy during times when the ERCOT bulk power system is under stress. Because the demand response share of spinning reserves is artificially capped at 50%, which is fully subscribed, it is likely there is greater potential for participation.

Markets Have Achieved Great Response

Markets are not necessary to recruit demand response. Demand response providers, like any other service provider, prefer a long-term contract at a favorable price as opposed to a one-year (or shorter) price guarantee that will fluctuate with market conditions from year-to-year. Further, integrating demand response into wholesale markets that are clearly designed with central station power plants in mind has proved difficult and complicated. Too often broad concepts for market rule modification are agreed upon in the early stages, only to discover later that the devilish details are unworkable. And yet, as the staff of the United States FERC reported in July 2011, the three regions of the country where demand response is integrated directly into the markets (PJM, ISO-NE, and NYISO) have some of the largest amounts of demand response (Figure 37).

This leads to an obvious question. Why? If there are reasons to expect traditional regulated market environments to be at least as good, if not inherently better at creating incentives for demand response, then why are the areas with demand response in wholesale markets generally acquiring more demand response as a percentage of peak load than other areas? We suggest several reasons.

Structure: In regions with traditionally regulated, vertically-integrated utilities, demand response programs may very well exist, but they are sourced by and delivered to the local electric utility. These utilities earn revenues based upon the quantity of electricity sold, and demand response erodes those sales, albeit only slightly. Perhaps the bigger effect is that reduced peak loads will result in reduced need to build peaking generation stations and the additional necessary transmission infrastructure to deliver that power. The utilities in these regions earn a return on revenue for construction of these facilities, and therefore haves a financial incentive to deter demand response as much as is politically possible.

Competitive Innovation: Independent DR providers whose sole business is to provide demand response have a greater financial incentive to sign up as many customers as possible. Utility providers can and have provided reliable demand response for many years, but they do not have the same financial incentives, and are not as aggressive as private DR providers.

The Money Matters

As a capacity resource, demand response relies on some form of regular capacity revenue, either from a market or a contract. For many demand response resources, the capacity payment is the only direct source of revenue.[34] In recent ISO-NE capacity market auctions for future delivery, a steady reduction in clearing prices (due largely to capacity over-supply) has led to some demand response resources withdrawing from the market. For the auction in 2014, the floor price will be removed and the clearing price is widely expected to drop below $1/kW-month. There may be additional demand response resources that choose to withdraw from the market due to that low price. This is what one would expect – demand response will behave like any other capacity resource, responding to changes in the value of capacity.

34 For some capacity programs, DR resources also receive an energy payment whenever DR is activated.

As an economic resource in the energy market, demand response is sensitive to energy market compensation. The PJM economic program provides a good example of this sensitivity. After steadily increasing participation from 2002 through 2007, PJM changed the compensation mechanism from full energy market price to energy market price minus the assumed cost of generation. PJM believed that reducing the compensation to demand response resources would provide appropriate (comparable to generation), economically efficient payments. After Order No. 745, PJM reinstated full energy market price compensation prior to 2012 and participation dramatically increased, *even though energy market prices were substantially lower in 2012 when compared to 2011*. This is shown in Figure 24 in Section 5 of this report.

It is likely that the methods for determining ancillary services compensation will also significantly affect whether demand response provides these services. If resources are rewarded based on the quality (including speed and accuracy) of response, demand response is likely to have a larger role in providing frequency response, load following, and other ancillary services.

In any electricity market, demand response is more likely to participate if the provider and/or the customer can expect a regular income. Capacity markets, ancillary services markets, and reserve contracts can all provide this type of reliable revenue. Energy markets can provide a steady revenue stream, but only if prices are sufficiently high and those prices occur regularly enough. If the availability of payment for service is rare and unpredictable, demand response – like any resource – will find the opportunity too risky and will withdraw.

B. Near-Future Opportunities

There are many promising new ways that demand response can contribute to overall system reliability. In preceding sections of this report, we described several pilots being implemented. Additionally, new program designs may improve the incentives and ability for demand response to provide greater services. Some of the more robust options for demand response development include:

New Developments in Demand Response

Across the country, several pilot programs are assessing the ability of demand response to provide ancillary services, while in PJM demand response has begun to participate fully in the ancillary services market. Ancillary services provided by demand response in PJM include regulation or load-following services using water heating, space heating, and water pumping, and an electric vehicle pilot that uses the battery as a storage device for regulation service (Carson 2012).

In the Pacific Northwest, the Bonneville Power Administration is conducting pilot programs designed to determine the ability of innovative demand response programs to provide balancing services for a region that has large quantities of hydropower and anticipates increasing quantities of variable wind resources.

Texas has a pilot program to test the performance of demand response resources that can respond to dispatch signals in thirty minutes or less. These new resources will be an addition to existing programs for demand response resources that are available in ten minutes or less and will use a similar program design.

As the foregoing suggests, much of the untapped potential value of demand response will be found in the growing need for flexibility services as the share of variable renewables grows in many markets.

Additional Design Options

PJM has implemented a three-tranche auction process for demand response that is tapping the potential for demand resources that offer greater value to the system than traditional, limited forms of "emergency response."

ISO-NE is considering design changes to its Forward Capacity Market that will provide incentives for all resources (including demand response resources) to be available during times of system need. An ISO-NE whitepaper provides an overview of the design elements for the new incentives. These Forward Capacity Market changes are being considered for implementation in 2015, at the earliest, for the 2018-2019 delivery year auction that will take place in March 2015 (ISO-NE 2012). In a previous whitepaper in June 2012, ISO-NE proposed adding a locational reliability element to Forward Capacity Market to procure more flexible resources (including demand response resources) in particular locations to address local reliability concerns.

MISO is exploring approaches to address balancing issues, which it states are the most common cause of scarcity pricing events on the MISO system. MISO's goal is to have more flexible resources, including demand response, on-line and available by providing appropriate incentives to attract resources that can provide flexible balancing services. MISO states that this will be less costly and more efficient than simply increasing the quantity of reserve and regulation resources that are required.

CAISO is working to resolve some jurisdictional issues regarding the provision of demand response resources by aggregators and the ability of CAISO to dispatch demand response resources. CAISO anticipates that renewable resources (largely wind and solar photovoltaics) will continue their rapid development in order to satisfy California's 33 percent renewable standard. One option that CAISO is considering is a flexible ramping program that would include demand response resources that can provide short-term balancing services for the frequent swings in energy supply needed as wind and solar resources are self-dispatched. CAISO is also exploring options for long-term support for flexible resources to meet resource adequacy needs. The theme for these efforts is expressed in the phrase "the right resource, in the right place, at the right time" (Marnay, et al. 2001).

Texas has an open proceeding on resource adequacy and is considering several recommendations from a Brattle Group report (Newell, et al. 2012) that evaluates alternative approaches for meeting resource adequacy goals. The two approaches that the Brattle Group and stakeholders reviewed in 2012 are (1) the introduction of a forward capacity market (similar to the PJM approach) and (2) an expansion of the demand response reserve programs combined with more active demand response resource participation in the energy market as price-responsive load (S. Newell 2012). A third approach proposed in December 2012 suggested creating a separate market for operating reserves that would not suppress energy market prices when activated (Hogan 2012).

C. Key Challenges

Despite the success to date, we see some key challenges preventing demand response from achieving its full potential in the United States.

Baseline Determinations and Infrastructure Requirements

As was explained in Section 2, demand response involves changes in normal electricity usage in response to a dispatch signal from the system operator. It is required, but costly, to install the metering, data retention, communications, and reporting infrastructure to establish and maintain customer baselines, and to accurately measure and report the amount of interruption during an event. However, ISOs, RTOs, and utilities use somewhat differing baseline methodologies to measure and verify the load impacts of demand response resources participating in wholesale markets.

It is argued by a number of demand response aggregators, large customers, and retail suppliers that this lack of a uniform baseline protocol increases transaction costs and presents barriers to participation of demand response in wholesale markets. Developing common measurement and verification standards would better enable third parties and demand response aggregators to provide demand response in multiple regions, as well as improve the ability of firms located in different geographic locations to standardize their demand response behavior. Currently, the North American Energy Standards Board, with support of FERC, is developing market standards for the measurement and verification of demand response contributions (Market Committee of the ISO/RTO Council 2007). In addition, excessively high infrastructure requirements should be avoided. As metering accuracy increases, so do costs. The importance of accurate and reliable customer baselines cannot be understated. With hundreds or thousands of customers aggregated together into a portfolio of demand response resources, trust in baselines can become fragile, and a few examples of bad actors will erode confidence further. Setting appropriate baseline methodologies in the beginning will avoid future troubles.

Regional Versus Distribution Utility Dispatch

Although the history of demand response resource participation began with individual utility programs, there is much greater value from demand response resources that can be directly dispatched by a regional system operator. Experience in California demonstrates that the inability of CAISO to directly control the dispatch of demand response has limited the growth and flexibility of demand response resources. The entity that dispatches the entire system (CAISO) is not the same entity that procures and dispatches demand response. This causes confusion.

In CAISO, only demand response resources that can participate in utility-designed programs are being developed. Third party aggregators who might prefer to offer in a statewide procurement are forced to negotiate with each distribution utility or respond to multiple, different utility requests for competitive bids. In the Midwest, many distribution utilities will not voluntarily permit DR aggregators to solicit their customers to participate as aggregated load resources; the utility prefers to determine when to dispatch load resources, particularly if the distribution utility is still a vertically-integrated utility. In some states, utility commissions have adopted rules that prohibit DR aggregators from enrolling demand response customers without the permission of the local utility.

Transitioning demand response resources from utility programs to system operator-controlled programs improves the effectiveness and efficiency of power system operations. The ability to reflect the value of demand response in program compensation mechanisms is an important element of finding a resolution. We recognize, however, that the transition of control is not easily achieved.

Market Barriers

Most wholesale markets were originally designed to procure energy and other reliability services from a relatively small number of large, central station power plants. The rules and administrative requirements of participating in those markets reflect this perspective. Although changing, market rules still struggle with how to incorporate the aggregation of hundreds or eventually thousands of customers into a single market "resource" that is being coordinated by a demand response aggregator. Minimum size requirements have deterred demand response participation in some regions, and metering and telemetry requirements may too stringent, making participation uneconomic. Market and system planners are accustomed to representing supply in discrete points with specific locations, and often require market participants to represent themselves in this manner in order to participate. Demand response struggles to fit this mold.

The design of the markets and the administrative details of participating in them will need to continue to evolve to move from hundreds of coordinated supply resources to an operational and financial system that can support thousands of smaller demand resources whose locations are spread around the region and constantly changing as one customer exits the aggregation and others join.

The value the market places on electric services will also need to change from one that places value on large, discrete, slow supply resources that have limited dispatch capability to one that values a large number of fast responding, smaller resources.

Institutional Resistance

Many areas in the United States have seen substantial growth in demand response over the past decade. We have focused upon these regions in our report. In other regions where the growth of demand response resources has lagged, there are often regulatory and corporate structures that provide incentives to build large, costly, long-life infrastructure projects and sell as many kilowatt hours as possible for maximum revenue. Perhaps regulatory bodies in these regions are not as comfortable changing the traditional vertically-integrated utility model, and consider the potential benefits of energy services from private companies to be too risky for their consumers. Although divestiture of generation resources is not a prerequisite for the development of demand response resources, in areas where divestiture has occurred, demand response providers are active and can develop resources that the traditional utility has not. Independent DR-focused providers must be allowed to participate for demand response to flourish. FERC Order No. 745, once fully implemented, will provide a consistent basis for demand response providers to participate in wholesale energy markets. It remains to be seen just how soon, and how strongly, this participation will occur.

8. Conclusions

Demand response resources are capable of providing numerous services/products that can enhance the efficiency and reliability of bulk power systems. These services span the range of resource adequacy, energy, and ancillary services.

Demand response has proven that it can reliably provide energy in times of high prices or high loads, reserves to support contingencies, and balancing services. However, like any business, there are upfront capital costs. For demand response aggregators, the costs of setting up the business, telemetry and metering requirements, and ongoing interactions with so many customers may be substantial. The business won't work without a robust investment case. As such, the growth of demand response has been strongest where a steady monthly payment is available, and where multiple streams of revenue are present to support different types of loads and different types of customers. Relying solely on occasional high-priced events is a business model that may prove too risky for the typical demand response resource. Capacity markets offer one model; developing service markets can provide additional options.

The most fruitful demand response exists not in narrowly-defined programs with specific amounts to be procured. Rather, regions that do not limit demand response resources and allow demand response to participate in multiple types of services (energy, reserves, regulation) have demonstrated greater participation by demand response resources.

Independent demand response aggregators whose sole business is to provide demand response have a greater financial incentive to sign up as many customers with load reduction capabilities as possible. Utility providers can and have provided reliable demand response for many years, but they often have conflicting financial incentives, and have not been as aggressive as private DR providers. Shifting traditional utility compensation to a more nuanced compensation system that can recognize cost-effective demand reductions will help align incentives towards a more efficient overall use of resources.

Large customers are easier to sign up than small customers because of the amount and variety of types of load available for reduction from a single facility. Few DR providers have even approached the residential market to date. However, the technology to provide small amounts of demand response from a very large number of residential customers is close at hand and may only be waiting for regulatory acceptance and small accommodations from system operators and regulatory agencies. Numerous pilot programs have explored this possibility of late, but widespread implementation is still several years away.

The ability for storage-type demand response that can ramp in both directions to both reduce load and also absorb excess generation is a new and developing area of demand response. It has been proven reliable for regulation and load following services at a small scale, and technically it can scale up. Historically, power plants have provided these services at a very low cost. Simply asking demand response resources to replace traditional generation services may not be enough. A re-determination of the value of regulation and load following services, as well as new services such as multi-interval ramping, may need to occur in parallel, particularly while considering likely future resource mixes that may include greater quantities of variable resources. Whether or not provision of these services is financially viable in the face of competition from power plants remains to be seen, but compensation that recognizes the value of speed and accuracy of response, and pays appropriately, will help.

Finally, regulatory support at both the state and federal level has been a critical element to enable greater participation of demand response. Change is always difficult, and some in the United States electric industry remain skeptical and resistant to the full incorporation of demand response resources. The consistent attention to these issues by the Federal Energy Regulatory Commission has proven essential to the success of demand response resources to date.

Works Cited

Anderson, Kenneth W. "Frequency Response as a Tool for Grid Reliability - ERCOT Demand Response." *2011 NARUC Winter Committee Meetings.* Washington, DC: NARUC, 2011.

Berst, Jesse. "The Latest Breakthrough in the Evolution of Demand Response." *Smart Grid News,* October 26, 2011.

Binz, Ron, Richard Sedano, Denise Furey, and Dan Mullen. *Practicing Risk-Aware Electricity Regulation: What Every State Regulator Needs to Know.* Boston: CERES, 2012.

Brattle Group. *Quantifying Demand Response Benefits in PJM.* PJM and MADRI, 2007.

Cappers, Peter, Charles Goldman, and David Kathan. *Demand Response in U.S. Electricity Markets: Empirical Evidence.* Ernest Orlando Lawrence Berkeley National Laboratory, 2009.

Carson, Phil. "The Business of V2G: Technology and Business Case go Hand in Hand." *Intelligent Utility,* September/October 2012.

Chatham, Bob, George Baker, and Jessica Miller, interview by Melissa Whited. *VCharge Interview* (December 13, 2012).

Coulbeck, Rob. "Demand Side Regulation for a Smart Water Grid." 2012.

Ecofys and Bonneville Power Administration. *TI220 Project Evaluation Report: Smart End-Use Energy Storage And Integration of Renewable Energy.* Bonneville Power Administration, 2012.

Electric Energy Market Competition Task Force. "Report to Congress on Competition in Wholesale and Retail Markets for Electric Energy." 2007.

ENBALA Power Networks. "ENBALA Power - Case Studies." ENBALA Power Networks. 2012. http://www.enbala.com/media/casestudies/Pennsylvania%20American%20Water%20Connects%20to%20the%20Smart%20Grid.pdf.

ERCOT. "2011 EILS Deployments." QMWG. 2012.

—. "Load Participation Update, January 21." Demand Side Working Group Monthly Meeting. ERCOT, 2005.

ERCOT Staff. "Load Resource Participation Update, October 28, 2011." ERCOT, 2011.

FERC. *Assessment of Demand Response & Advanced Metering Staff Report - Novemeber 2011.* Federal Energy Regulatory Commission, 2011.

FERC. *Assessment of Demand Response and Advanced Metering.* Federal Energy Regulatory Commission, 2006.

FERC. *National Action Plan on Demand Response.* Docket No. AD09-10, Washington, DC: Federal Energy Regulatory Commission, 2010.

—. *Reports Demand Response & Advanced Metering.* March 23, 2012. http://www.ferc.gov/industries/electric/indus-act/demand-response/dem-res-adv-metering.asp (accessed October 30, 2012).

FERC Staff. *A National Assessment of Demand Response Potential.* Federal Energy Regulatory Commission, 2009.

Fryer, Lynn, Steven Schiller, Andrea Geanacopoulos, and Harlan Coomes. "Demand Responsive Load Management: From Programs to a Demand Response Market." *ACEEE Summer Study Proceedings.* ACEEE, 2002.

Goldman, Charles, Joseph Eto, and Galen Barbose. *California Customer Load Reductions during the Electricity Crisis: Did they Help to Keep the Lights On?* Berkeley: Ernest Orlando Lawrence Berkeley National Laboratory, 2002.

Gottstein, Meg, and Lisa Schwartz. *The Role of Forward Capacity Markets in Increasing Demand-Side and Other Low-Carbon Resources.* Montpelier, VT: Regulatory Assistance Project, 2010.

Hirst, Eric, and Charles Goldman. "Creating the Future: Integrated Resource Planning for Electric Utilities." *Annual Review of Energy and the Environment 16* (1991): 91-121.

Hogan, William. *Electricity Scarcity Pricing Through Operating Reserves: An ERCOT Window of Opportunity.* PUCT Proceeding to Ensure Resource Adequacy in Texas, Public Utility Commission of Texas, 2012.

Internal Market Monitor of ISO-NE. *2011 Annual Markets Report.* ISO New England, 2012.

ISO-NE. *FCM Performance Incentives.* White Paper, ISO New England, 2012.

ISO-NE. *Semi-Annual Load Response Reports, Docket ER03-345-*.* FERC Compliance Filing, ISO New England, 2006-2012.

ISO-New England. *FCM Performance Incentives.* ISO-NE White Paper, ISO-NE, 2012.

ISO/RTO Council. (2011) *ISO RTO Operating Regions.* Available at: http://www.isorto.org/site/c.jhKQIZPBImE/b.2604471/k.B14E/Map.htm

Isser, Steve. *Demand Response Measurement and Verification Literature Review.* North American Energy Standards Board, 2008.

Joskow, Paul. "California's Electricity Crisis." *Oxford Review of Economic Policy* 17, no. 3 (2001).

Kirby, Brendan. *Ancillary Services: Technical and Commercial Insights.* Wartsila North America Inc., 2007.

Levy, Roger, Sila Kiliccote, and Charles Goldman. "Demand Response: Lawrence Berkeley National Laboratory - Smart Grid Technical Advisory Project." *NARUC Webinar.* National Association of Regulatory Utility Commissioners, 2011.

Liu, Shucheng. "System Flexibility for Integrating 33% Renewable Generation in California ISO." *Clean Energy Regulatory Forum Workshop 4, November 8-9.* California ISO, 2012.

MacDonald, Jason, Peter Cappers, Duncan Callaway, and Sila Kiliccote. "Demand Response Providing Ancillary Services." *Grid-Interop Forum 2012.* Irving, TX: GridWiseArchitecture Council, 2012.

Market Committee of the ISO/RTO Council. *Harnessing the Power of Demand.* IRC ISO/RTO Council, 2007.

Marnay, Chris, Kristina Hamachi, Mark Khavkin, and Afzal Siddiqui. *Direct Participation of Electrical Loads in the California Independent System Operator Markets During the Summer of 2000.* Environmental Energy Technologies Division Report LBNL-47897, Berkeley, CA: Lawrence Berkeley National Laboratory, 2001.

McAnany, James. "2012 Load Response Activity Report March 2013." PJM, 2013a.

—. "2013 Load Response Activity Report March 2013." PJM, 2013b.

Monitoring Analytics, LLC. 2011 *State of the Market Report for PJM.* Monitoring Analytics, LLC, 2012.

Monitoring Analytics, LLC. *State of the Markets Report for PJM: January through September 2012.* Monitoring Analytics, LLC, 2012b.

Neenan, Bernard, Donna Pratt, Peter Cappers, Jeremey Anderson, Lydia Scholle-Cotton, and Kim Butkins. *2004 NYISO Demand Response Program Evaluation.* Report Prepared for the NYISO, Neenan Associates, 2005.

NERC. *Demand Response Availability Data System (DADS): Phase I and II Final Report.* Princeton, NJ: North American Electric Reliability Corporation , 2011.

—. "NERC Ancillary Services Matrix." *North American Electric Reliability Corporation.* n.d. http://www.nerc.com/docs/pc/ivgtf/NERC_ancillary_services%20ERCOT%20IESO%20NYISO%20MISO%20PJM%20SPP%20WECC%2012%2014.pdf (accessed January 14, 2013).

Newell, Sam. "Resource Adequacy in ERCOT: 'Composite' Policy Options." Prepared for the Public Utilities Commission of Texas, October 25. 2012.

Newell, Samuel, Kathleen Spees, Johannes Pfeifenberger, Robert Mudge, Michael DeLucia, and Robert Carlton. *ERCOT Investment Incentives and Resource Adequacy.* The Brattle Group, 2012.

NYISO. *2007 Power Trends.* Rensselaer, NY: New York Independent System Operator, 2007.

NYISO. *2011 Annual Report on Demand Response Programs, Errata Filing.* FERC Docket ER01-3001-000, NYISO, 2012c.

NYISO. *Annual Report on Demand Response.* New York Independent System Operator, 2011.

NYISO. *Power Trends 2012: State of the Grid.* Rensselaer, NY: New York Independent System Operator, 2012.

NYISO. *Special Case Resources: Evaluation of the Performance and Contribution to Resource Adequacy.* New York Independent System Operator, 2012b.

Order No. 755. 137 FERC, 61,064. Docket Nos. RM11-7-000 and AD10-11-000; Order No. 755; Frequency Regulation Compensation in the Organized Wholesale Power Markets (Federal Energy Regulatory Commission, October 20, 2011).

Pfeifenberger, J.P., G.N. Basheda, and A.C. Schumacher. "Restructuring Revisited." *Public Utilities Fortnightly,* 2007: 64-69.

Pfeifenberger, Johannes, and Attila Hajos. "Demand Response Review." *Presentation to AESO.* The Brattle Group, 2011.

PJM. *Load Management Performance Report 2010/2011.* PJM, 2010.

—. "PJM Estimated Demand Response Activity July 16-18 2012." *PJM Hot Days Report.* PJM, 2012.

Pratt, Donna, Peter Cappers, and Jeremy Anderson. "NISO 2005 Demand Response Program Evaluation Results." *Presentation by Neenan Associates.* NYISO, 2005.

Rulemaking to Amend Subst. R. 25.507, Relating to Electric Reliability Countil of Texas (ERCOT) Emergency Interruptible Load Service (EILS). Project 39948 (PUCT, March 22, 2012).

Scibelli, Maria. "Proposed Installed Capacity Requirement Values & Appendix of Assumptions for the 2013/14 ARA3, 2014/15 ARA2, 2015/16 ARA1." *ISO-NE September 19 RC Meeting.* ISO-New England, 2012.

Smith, Doug. "Active Demand Resource Performance: Preliminary Audit and Dispatch Event Results, June and July 2011." ISO-NE, 2011.

Taniwha, Michael. "Summary of Capacity Deficiency Events - June & July 2010." *Reliability Committee Meeting.* ISO New England, 2010.

Todd, D., M. Caufield, B. Helms, M. Starke, B. Kirby, and J. Kueck. *Providing Reliability Services through Demand Response: A Preliminary Evaluation of the Demand Response Capabilities of Alcoa Inc*. Washington, DC: US Department of Energy, 2009.

US Department of Energy. *Benefits of Demand Response in Electricity Markets and Recommendations for Achieving Them.* A Report to the United States Congress, Washington, DC: US Department of Energy, 2006.

Wattles, Paul. "Demand Response in ERCOT Operations." *Clean Air through Energy Efficiency Conference*. 2011.

Weare, Christopher. *The California Electricity Crisis: Causes and Policy Options.* Public Policy Institute of California, 2003.

Wellinghoff, Jon, David Morenoff, James Pederson, and Tighe Mary Elizabeth. *Creating Regulatory Structures for Robust Demand Response Participation in Organized Wholesale Electric Markets.* ACEEE Summer Study Article, American Council for an Energy Efficient Economy (ACEEE), 2008.

York, Dan, and Martin Kushler. *Exploring the Relationship Between Demand Response and Energy Efficiency.* Report Number U052, Washington, DC: American Council for an Energy-Efficient Economy (ACEEE), 2005.

Terminology Appendix

Ancillary services: Services that help the system operate continuously within required parameters (e.g., frequency and voltage range), including the ability to recover energy balance after significant unplanned changes in supply and demand. In this report, ancillary services are defined as follows, although regional definitions may vary:

Service	Service Description		
	Typical Response Speed Required	*Duration*	*Cycle Time*
Normal Conditions			
Frequency Regulation	Online resources, on automatic generation control, that can respond rapidly to changes in frequency.		
	<10 seconds	*Seconds to Minutes*	*Seconds to Minutes*
Regulating Reserve	Online resources, on automatic generation control, that can respond rapidly to system-operator requests for up and down movements; used to track the minute-to-minute fluctuations in system load and to correct for unintended fluctuations in generator output. Response times are typically <1 minute.		
	4 Seconds - 1 minute	*Minutes*	*Minutes*
Load Following	Similar to regulation but slower. Bridges between regulation service and hourly energy markets. This service is performed by the real-time energy market in regions where such a market exists.		
	5 - 10 minutes	*10 min to hours*	*10 min to hours*
Contingency Conditions			
Spinning Reserve	Online generation, synchronized to the grid, that can increase output immediately in response to a major generator or transmission outage and can reach full output within 10 min.		
	Seconds to <10 min	*10 to 120 min*	*Hours to Days*
Non-Spinning Reserve	Same as spinning reserve, but need not respond immediately. Resources can be offline but still must be capable of reaching full output within the required 10 min.		
	<10 min	*10 to 120 min*	*Hours to Days*
Replacement or Supplemental Reserve	Same as supplmental reserve, but with a 30-60 min response time; used to restore spinning and non-spinning reserves to their pre-contingency status.		
	<30 min	*2 hours*	*Hours to days*

Balancing services: A particular class of ancillary service that involves purchases and sales of energy made by the system operator close to real time that are necessary to correct current or expected imbalances between supply and demand for each trading period. Generally occur after bi-lateral physical markets have closed (gate closure).

Capacity markets: Encompasses the range of capacity payment mechanisms designed to remunerate market participants for committing a volume of firm capacity to generate power or reduce demand by an equivalent amount during hours of system peak demand.

Demand response (or 'responsive demand'): Customer loads that can be modulated up or down in real time in response to wholesale market conditions. Market signals may be expressed through wholesale prices, frequency or voltage fluctuations, or through arrangements allowing direct control by the system operator or a third-party aggregator.

Demand-side resources: The full range of customer-based resources (end-use energy efficiency, demand-response and customer-sited generation) that reduce energy needs at various times of the day and year—across some or many hours.

Dispatch: Unit commitment day ahead and adjustment to the output of system resources in line with real time changes in the level of demand.

Firm capacity: The volume of megawatts guaranteed to be available to provide energy to the system at any moment in time.

Frequency regulation: Reliability service provided by reserves on automatic generation control that can start immediately and provide full availability within seconds in response to changes in system frequency. This type of reserve is referred to as "primary reserve" or "primary response" in Europe (involving inertia and speed regulators).

Load following: The ability of supply or demand resources to follow net demand, ramping up and down as necessary. In the United States and elsewhere, the capability of load following can be used in markets in addition to regulating reserves which are generally equivalent to "secondary reserves" as used in the EU.

Net demand: Demand for energy not already served by the output of variable renewables.

Primary reserve: See "frequency regulation" above.

Ramping: The capability of a supply or demand resource to ramp up or ramp down as required. In the United States and elsewhere, the capability of ramping can be used in markets in addition to regulating reserves which are generally equivalent to "secondary reserves" as used in the EU.

Reliability: Ability to meet the electricity needs of customers connected to the system over various timescales even when unexpected equipment failures or other factors reduce the amount of available electricity. Consistent with current industry practice, 'reliability' can be broken down into two general categories—resource adequacy and system quality.

Resource adequacy: Enough of the right kinds of resources to match demand and supply across time and geographic dimensions and deliver an acceptable level of reliability. Traditionally a "volume-based" standard based on the amount of firm capacity available to meet system peak demand.

Regulating reserve (or Regulation): Reserves - of centralized automatic control and generally starting after 30 seconds with full availability within minutes - used to bring back frequency or interchange programs to target. This type of reserve is generally equivalent to fast "secondary reserves" as used in the EU. Secondary reserves can be provided by increased fuel input on part-loaded thermal plant, hydro or pumped storage or by fast-start plant at standstill. They can also be used for load following or the maintenance of interchange programs. In the United States and elsewhere, the capabilities of ramping and load following can be used in markets in addition to regulating reserves.

Secondary reserve: See "regulating reserve" above.

Spinning reserve: The spinning reserve is the online but unused capacity that is synchronized to the grid and which can be activated on decision of the system operator. Non-spinning reserve is off-line generation capacity that can be ramped to capacity and synchronized to the grid.

System operator(s): Entities authorized to perform planning, operational or investment-related functions in power markets (e.g., system administrators, planning authorities).

System peak demand: Highest instantaneous level of total energy demand on the power system over a given period of time (e.g., daily peak, seasonal peak, annual peak).

System quality: Short-term, reliable operation of the power system as it moves electricity from generating sources to retail customers, including the ability of the system to withstand unanticipated disturbances or imbalances in the system. Balancing and ancillary services contribute to system quality.

Tertiary reserve: Reserves involving manual change in the dispatching and unit commitment used for load following, to restore the secondary control reserve, to manage eventual congestions and to bring back the frequency and the interchange programs to their target if the secondary control reserve is not sufficient. This type of reserve, which generally takes from around 15 minutes to one hour at most to come online, is often referred to as "operating", "supplemental" or "replacement" reserves in the United States and elsewhere. Tertiary reserves can be provided by a range of sources, including synchronized part-loaded conventional generation, warm plant at standstill or fast start generation.

Variable renewables: A power system resource using a primary renewable energy source that cannot be controlled (e.g., solar- and wind-powered generation). Such resources can be curtailed if needed and to varying degrees available capacity can be held as reserve; however, their availability is significantly less controllable than conventional thermal generation.

The Regulatory Assistance Project (RAP) is a global, non-profit team of experts focused on the long-term economic and environmental sustainability of the power and natural gas sectors. We provide technical and policy assistance on regulatory and market policies that promote economic efficiency, environmental protection, system reliability, and the fair allocation of system benefits among consumers. We work extensively in the US, China, the European Union, and India.
Visit our website at **www.raponline.org** to learn more about our work.

RAP

The Regulatory Assistance Project

Beijing, China • Berlin, Germany • Brussels, Belgium

Montpelier, Vermont USA • New Delhi, India

Rosenstrasse 2 • D – 10178 Berlin • *phone:* +49 30 2844 901-21

rue de la Science 23 • B - 1040 Brussels • *phone:* +32 2 894 9300

www.raponline.org