of demand response program, which included categories for direct load control (DLC) and interruptible/curtailable (I/C) rate programs.

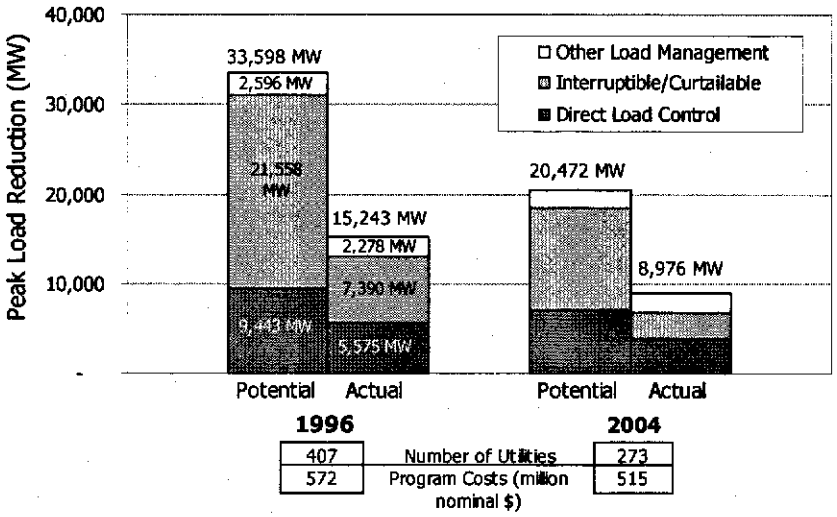

Figure 2-1. Existing U.S. Demand Response Potential

Figure 2-1 summarizes information on potential and actual peak reductions and program costs for 1996 and 2004.[18] Several trends are worth noting:

- Demand response potential in 2004 was about 20,500 MW, 3% of total U.S. peak demand. Actual delivered peak demand reduction was about 9,000 MW, about 1.3% of total peak (NERC 2005).
- Total potential load management capability has fallen by 32% since 1996. Factors affecting this trend include fewer utilities offering load management services (407 utilities in 1996 to 273 in 2004), declining enrollment in existing programs, the changing role and responsibility of utilities, and the increase in installed capacity The DSM information reported by industry participants to EIA does not fully reflect current demand response activity levels.[19]
- Actual peak reductions are affected by the available installed load reduction capability (i.e., the demand response potential), whether utilities or ISOs/RTOs called program events, and the extent to which enrolled participants respond during events.

[18] 1996 is both the year with the highest potential load reduction capability and the last year for which disaggregated information on demand response program type is available; 2004 is the most recent year of reported data.

[19] For example, utilities do not systematically report information on customer participation in optional "price-based" demand response programs (e.g. RTP, CPP, and TOU) and competitive retailers do not report the types and mix of contracts/products provided to retail customers. It is unlikely that all industry participants enrolled in ISO demand response programs are reporting their demand response activities.

- In 2004, utilities reported spending about $515M on load management programs; this represents about a 10% decrease from the early to mid-1990s.
- Although not shown explicitly in Figure 2-1, residential and industrial customers account for the bulk of actual peak load reductions (32% and 50% respectively) in 2004.

Market Structures for Electricity Production in the U.S.

Historically, the U.S. electric power industry has relied heavily on a market structure based on vertically integrated utilities that planned and operated electric generation, transmission and distribution systems on an integrated basis. Investor-owned utilities have an obligation to provide reliable service to customers in established, franchise service territories and are subject to regulation as a monopoly by state public utility commissions that set retail rates and review major capital investments and utility operations.

During the last decade, federal legislation (e.g., Energy Policy Act of 1992) and various Federal Regulatory Energy Commission (FERC) orders have helped create more competitive wholesale power markets with mandated open transmission access. Today almost every load-serving entity in the nation purchases some portion of its supply from these wholesale power markets, whether through bilateral contracts or in an organized spot market. Organized spot markets for wholesale electricity, operated by RTOs or ISOs) exist in the Northeast, Mid-Atlantic, much of the Midwest, and in Texas and California. ISOs/RTOs are typically responsible for maintaining grid reliability by overseeing and operating the high-voltage bulk power system and coordinating electricity generation, operating bid-based markets for spot energy (e.g. real-time, day-ahead, or ancillary services), and conducting long-term regional planning to identify system upgrade and expansion needs and overseeing capacity markets (in some cases).

In those states and regions without an ISO or RTO, electricity is delivered and transacted primarily by vertically integrated utilities through self-generation and bilateral contracts with significant state regulatory oversight of resource planning and rates.

Retail competition has been established in 18 states, which give customers additional choices in the supply and pricing of electricity. In these states, there have also been significant changes in the roles and responsibility of utilities (e.g. divesting of some generation assets, separation of competitive retail service function from transmission and distribution services which remain regulated).

A significant number of customers (20-25% of U.S. electric load) are also served by rural electric cooperatives or public power (municipal or public utility district) utilities. These entities have structural characteristics that are similar to vertically integrated utilities in that they typically have an obligation to serve customers in an established franchise service territory and many own generation, transmission and distribution assets, but their governance structure differs in that they are overseen by local authorities and boards. In a few states they are also regulated at the state level. Some public power utilities and rural cooperatives purchase some or all of their power requirements from vertically integrated utilities, generation and transmission cooperatives, power marketing authorities, or through wholesale markets and in some cases have developed load management resources to a greater extent than investor-owned utilities (Kexel 2004).[20]

[20] For some rural cooperatives, the primary reason for implementing load management programs was to reduce billed demand charges to the member cooperatives themselves and to reduce the capacity requirements of their Generation and Transmission cooperatives (Kexel 2004).

The Role of Demand Response in Electric Power Systems

In assessing the benefits of demand response, it is important for policymakers to be cognizant of the physical infrastructure and operational requirements necessary to construct and reliably operate an electric power system as well as regional differences in market structure and industry organization (see the previous textbox).

In all market structures, the management of electric power systems is largely shaped by two important physical properties of electricity production. First, electricity is not economically storable, and this in turn requires maintaining the supply/demand balance at the system level in real time. Mismatches in supply and demand can threaten the integrity of the electrical grid over extremely large areas within seconds. Second, the electric power industry is very capital intensive. Generation and transmission system investments are large, complex projects with expected economic lifetimes of several decades that often take many years to develop, site and construct.

These features of electric power systems necessitate management of electricity on a range of timescales, from years (or even decades) for generation and transmission planning and construction, to seconds for balancing power delivery against fluctuations in demand (see Figure 2-2). Decisions are made at several junctures along this timeframe. Generally speaking, the amount of load committed at each juncture declines as the time horizon approaches power delivery. For example, 70-80% of supplied load is often committed through forward energy contracts, months or even years before it is delivered. The amount of power arranged on a day-ahead basis varies, but is typically 10-25% of total requirements. In most cases, less than 5% of supply is committed in the last two hours before its delivery.

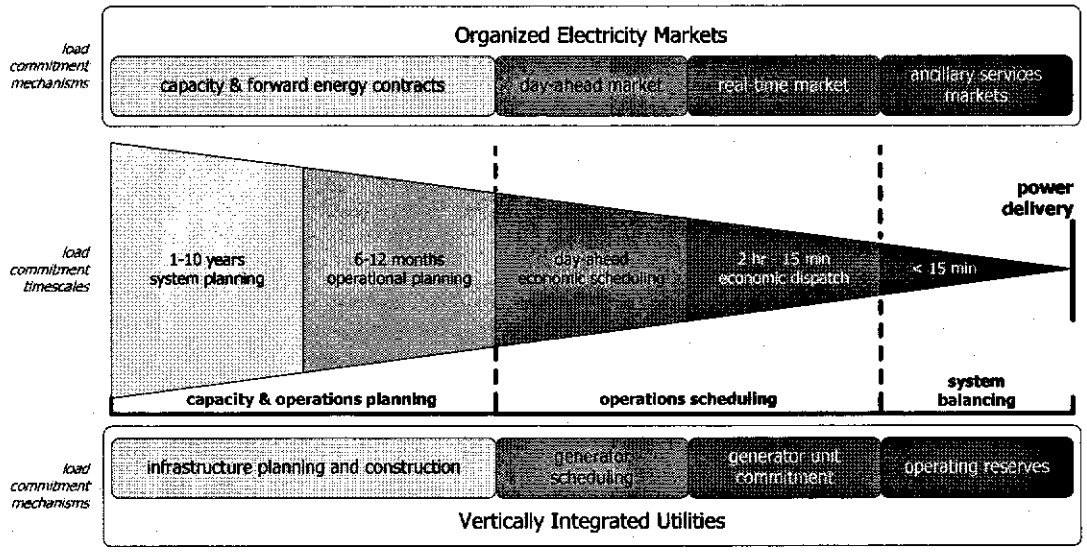

Figure 2-2. Electric System Planning and Scheduling: Timescales and Decision Mechanisms

The major infrastructure planning and operational power delivery decision timeframes are similar in regions with organized wholesale markets and in vertically integrated

systems, although the mechanisms for committing energy supply responsibilities differ (see Figure 2-2). In states with retail competition, default service providers and competitive retailers often have a much shorter horizon for acquiring resources than a vertically integrated utility in a state without retail competition.

- *Capacity and operations planning* includes long-term investment and planning decisions. Capacity, or system, planning involves assessing the need for and investing in new generation, transmission and distribution system infrastructure over a multi-year time horizon. Operations planning involves scheduling available resources to meet expected seasonal demand and spans a period of months. In vertically integrated utility systems, these investments are typically evaluated in a utility resource planning process, subject to state regulatory review. In regions with organized wholesale markets, responsibility for these activities is more diffuse. An ISO or RTO engages in a long-term transmission planning process, while distribution utilities retain responsibility for distribution system planning and operations. ISO-administered energy and capacity markets (in some areas) determine the scheduling and operation of available resources to meet daily and seasonal needs and also provide price signals for investments in new generation plants. Utilities and competitive retail suppliers, collectively referred to as load-serving entities (LSEs), contract with generators to meet forward energy requirements.
- *Operations scheduling* refers to the process of determining which generators operate to meet expected near-term demand. This typically involves making day-ahead commitments based on the next day's forecasted demand, with adjustments made in a period of hours down to 15 minutes to account for discrepancies in day-ahead and day-of demand forecasts as well as to account for any unexpected generation plant outages or transmission line problems. Day-ahead and real-time markets administered by ISOs or RTOs fulfill these responsibilities in regions with organized wholesale markets, using generator (or demand resource) offers as the mechanism for scheduling resources for dispatch. Vertically integrated utilities evaluate and schedule generation plants on a merit order basis ranked according to their variable operating costs.
- *System balancing* refers to adjusting resources to meet last-minute fluctuations in power requirements. In regions with organized wholesale markets, resources offer to provide various ancillary services, such as reactive supply and voltage control, frequency-responsive spinning reserves, regulation, and system black-start capability that are necessary to support electrical grid operation.[21] Vertically integrated utilities typically provide ancillary services as part of their integrated operation of the power system.

Ultimately, supply resources are valued according to the timescale of their *commitment* or *dispatch*. Yet because electricity is not storable, its *delivery* to consumers—the goal

[21] Reserves are a type of ancillary service for which ISO/RTO markets have been established in regions with organized wholesale markets. Generators (and loads) bid their availability to supply backup power with varying degrees of notice (usually from 30 minutes down to 10 minutes). Other types of ancillary services are typically contracted for directly by ISOs or RTOs.

around which power systems are constructed and managed—occurs in real-time, regardless of when it was committed and priced.

Demand response options can be deployed at all time scales of electricity system management.

Demand response options can be deployed at all timescales of electricity system management (see Figure 2-3) and can be coordinated with the pricing and commitment mechanisms appropriate for the timescale of their commitment or dispatch.[22] For example, demand response programs designed to alert customers of load response opportunities on a day-ahead basis should be coordinated with either a day-ahead market or, in a vertically integrated market structure, with the utility's generator scheduling process. Like generation resources, the actual *delivery* of customer load reductions occurs in real time.

Energy efficiency is a demand-side resource that can be integrated and valued as part of the system planning process and time horizon (Figure 2-3). Though not dispatchable, energy-efficiency measures often create permanent demand-reduction impacts as well as electricity savings.

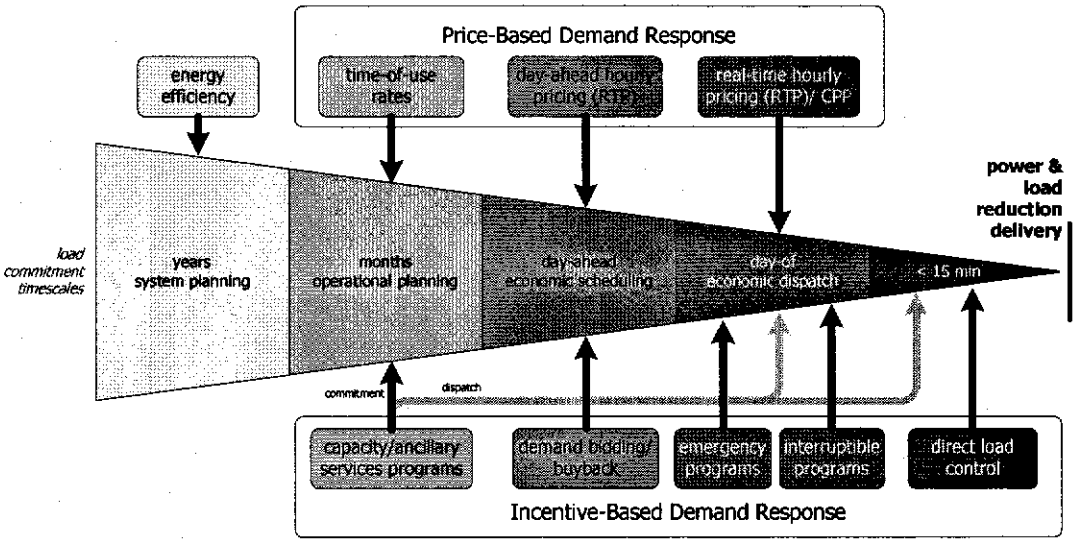

Figure 2-3. Role of Demand Response in Electric System Planning and Operations

If utility resource planners and system operators have a good sense of how their customers respond to changes in the price of electricity, price-based demand response options may be incorporated into system planning at different time scales (Figure 2-3):

[22] In some cases, demand response resources have been included in a Request for Proposals (RFP) process designed to alleviate short-term (e.g., 3-4 years), localized transmission capacity constraints. For example, ISO-NE issued an RFP for demand relief over four years in Southwest Connecticut, where construction of transmission capacity was delayed (Platts 2004), and Bonneville Power Administration issued an RFP for demand reduction, energy efficiency and distributed generation options to defer new transmission investments on a five-year timescale in 1994.

- *TOU rates*, which reflect diurnal and seasonal variations in electricity costs but are fixed months in advance, may be valued and integrated as part of operations planning.
- *RTP* provides hourly prices to customers with day-ahead or near-real-time notice, depending on the tariff design.[23] In wholesale markets with ISOs/RTOs, RTP prices are typically indexed to transparent, location-based, day-ahead or real-time hourly energy market prices; absent an organized spot market, utilities establish RTP "prices" based on the utility's marginal procurement costs.
- *CPP rates* are essentially TOU rates with the addition of a critical peak price that is called on a day-of basis.

Incentive-based demand response programs may be introduced at virtually all timescales of electric system management (Figure 2-3):

- *Capacity programs* involve load reduction commitments made ahead of time (e.g., months), which the system operator has the option to call when needed. The call option is usually exercised with two or less hours of notice, depending on the specific program design. Participants receive up-front capacity payments, linked to capacity market prices, from entities that otherwise would need to purchase comparable levels of generation to satisfy capacity reserve obligations.
- *Ancillary services programs* also involve establishing customer load commitments ahead of time. Customers whose reserve market bids are accepted must then be "on call" to provide load reductions, often with less than an hour's notice.[24]
- Load reductions from *demand buyback* or *bidding programs* are typically scheduled day-ahead, and incentive payments are valued and coordinated with day-ahead energy markets.
- *Emergency programs* are reliability-based, and payments for load reductions are often linked to real-time energy market prices (in regions with organized wholesale markets) or values that reflect customer's outage cost or the value of lost load. Program events are usually declared within 30 minutes to 2 hours of power delivery.
- *DLC programs* are typically reliability-based and can be deployed within minutes because the utility or system operator triggers the reduction directly, without waiting for a customer-induced response.[25]

---

[23] In some states (e.g., New Jersey, Maryland, Pennsylvania), RTP tariffs have been implemented that are indexed to real-time markets that do not communicate prices until after the fact. No studies assessing observed price response from this tariff design have been conducted. It is conceivable that customers look to near real time prices or day-ahead market prices posted by the PJM Interconnection, as a proxy and adjust their usage accordingly (Barbose et al. 2005).

[24] See Kirby (2003) and Kueck et al. (2001) for more information on customer load participation in ancillary services markets.

[25] DLC can also be used by LSEs to mitigate the impact of high wholesale market prices or manage system-demand related charges.

How Do Customers Accomplish Demand Response?

There are significant challenges in matching customers' preferences for demand response program features to system characteristics that drive value. From the customer perspective, investments in demand response and energy efficiency are both DSM strategies that can be used to manage energy costs. Participation in DSM programs (or making DSM investments) involves a series of decisions (see Figure 2-4).

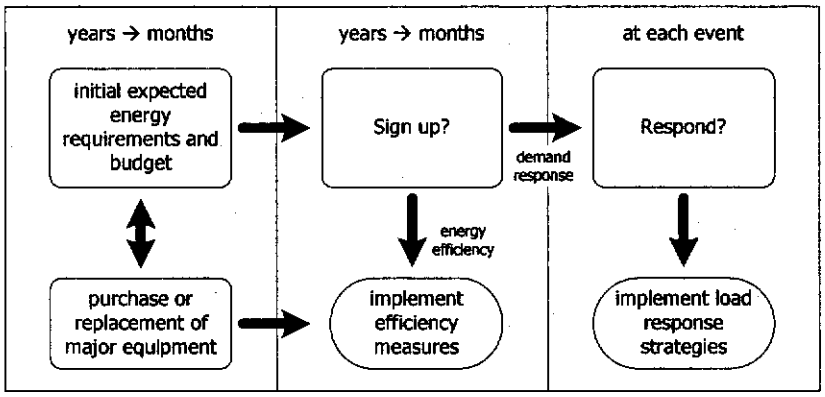

Figure 2-4. Customer Decisions for Demand-Side Management

First, customers implicitly or explicitly determine an initial energy budget based on their expectations of current and future average electricity prices and their household or facility energy needs (see Figure 2-5). The timeframe for this decision (or expectation) is typically monthly or annual, and decisions about purchasing or replacing major energy-using equipment may be made at the same time (see Figure 2-4). The decision-making process may be somewhat different for residential and small commercial customers, who may have a less formalized notion of their usage needs and budget than for large commercial or industrial facilities that may include energy costs as part of a specific operating budget.[26] Larger demand-metered customers are also more likely to be concerned with managing their peak demand in response to demand charges, which are typically included in their electricity tariffs.

Customer participation in demand response options involves *two* important decisions: whether or not to sign up for a voluntary program or tariff (or remain on the option in the case of a default tariff) and, subsequently, whether or not to respond to program events or adjust usage in response to prices as they occur (see Figure 2-4). This is in contrast to traditional energy-efficiency programs, in which customers invest in high-efficiency equipment in response to an existing program offered by a utility, state agency, or public benefits administrator that provides information, technical assistance and/or financial incentives.[27] In most cases energy-efficiency measures, once installed, continue to reduce

[26] This characterization of the customer decision process is more applicable to large, sophisticated, customers. There is a portion of the customer base, particularly many residential and small business customers that have limited understanding of their energy usage patterns and existing tariffs.
[27] Many customers also decide to invest in high efficiency equipment or measures based solely on their own internal economic decision criteria, apart from publicly funded programs.

energy usage over a multi-year economic lifetime, usually without much ongoing customer attention.[28] Compared to the initial usage and budget decision, which is relatively simple and familiar to customers, customers' decisions to enroll in demand response programs and to respond during events can be quite complex.

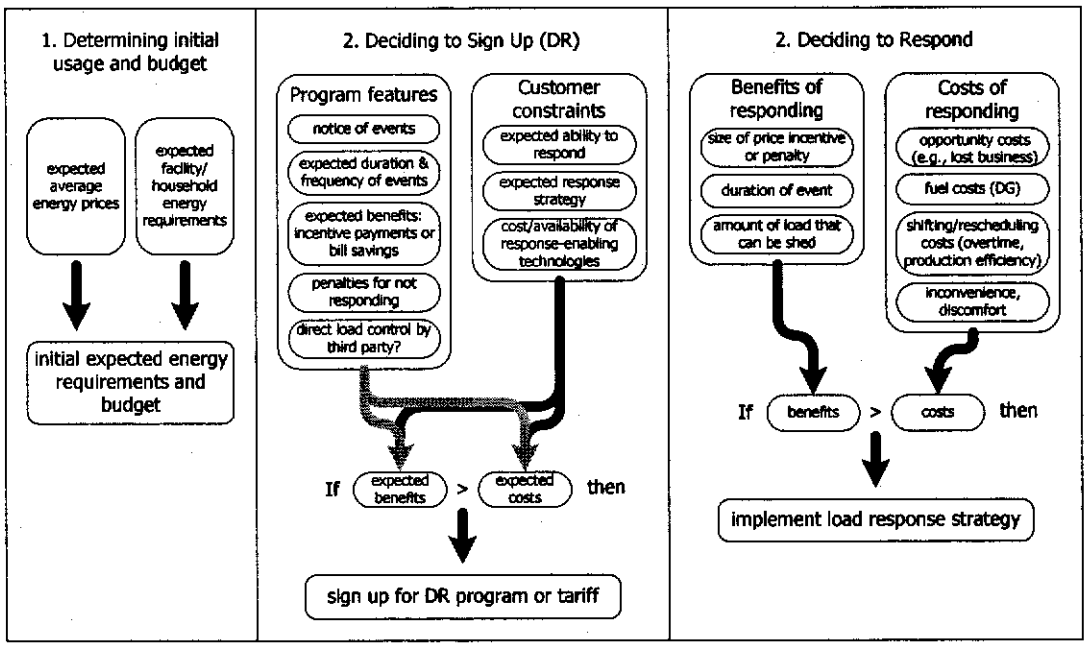

Figure 2-5. Factors Affecting Customer Decisions About Demand Response

The decision to sign up for demand response options involves evaluating offered program or tariff features and weighing the *expected* costs and benefits (see Figure 2-5). A demand response program may specify key parameters of interest to customers (e.g., maximum number of emergency events, payment if event is called), although there is significant uncertainty about the probability and timing of emergency events for the customer.

Ultimately, uncertainties in the costs and benefits of program participation represent risks to customers that may pose significant barriers to their signing up. For example, under RTP, future hourly prices are uncertain, making the benefits of participation difficult to predict.[29]

---

[28] Some energy-efficient equipment does require ongoing commissioning or maintenance to ensure energy savings continue to be realized over time, or savings may be affected by changes in customer usage of the equipment. Nonetheless, most energy-efficiency investments produce at least some level of savings over a period of years without further customer attention.

[29] However, the most popular form of RTP, two-part RTP, provides some financial protection against unexpectedly high prices, and the primary driver of participation is likely the expectation of *lower* average prices than under a standard tariff. Experience at successful programs (e.g., Georgia Power and Duke Power Company) has shown that some customers reduce load substantially during hours of high prices. Thus, RTP customers have the possibility of achieving bill savings from both lower prices overall, and from responding to high prices when they occur.

The relative certainty of a benefit stream may be as important to customers as the benefits themselves.

Potential participants in emergency demand response programs also face uncertainty about the number of demand response events in which they will be able to achieve benefits, and the payments they will receive when the events occur. Only in capacity-related demand response programs are up-front payments typically provided, in return for which customers agree to curtail on short notice when notified. The relative *certainty* of a benefit stream may be as important as the incentive payments themselves. While certain up-front investments, such as programmable thermostats, energy management systems or onsite generation equipment, may make responding easier, uncertainties about the benefits of responding can make these investment decisions difficult to justify.

Once enrolled, customers must decide whether or not to respond as events arise (see Figure 2-5). The benefits of responding are dependent on the actual financial incentive payment that applies to the given event (including the penalty for not responding), the number of hours that the event extends for, the amount of load the customer can shed, and may also include such considerations as the desire to help others by keeping the electric system secure.[30]

Customers may adopt one or more of three basic load response strategies (see the textbox below) and will assess the actual costs of responding in a specific situation. Their costs of responding depend in part on the type of response strategy undertaken. For example, customers who forego usage without making it up later incur costs due to lost productivity or foregone amenity. Customers that shift or reschedule their energy usage may incur costs from labor rescheduling, overtime pay or productivity losses from adjustments to their production process. If onsite generation is used to respond, fuel and maintenance costs are incurred. For any response strategy, inconvenience or discomfort to building occupants or tenants are likely to be important considerations and may be an important part of the cost-benefit decision, even if they are not directly monetized.

[30] Note that customers in DLC programs often do not have the choice about whether or not to respond during emergency events. Rather, their choices are focused on the decision to enroll or continue to participate in the program.

Types of Customer Load Response

Customers participating in demand response options may respond to high prices or program events in three possible ways:

- *Foregoing:* involves reducing usage at times of high prices or demand response program events without making it up later. For example, a residential customer might turn off lights or turn up the thermostat on an air conditioner during an event, or a commercial facility might turn off office equipment. In both cases, a temporary loss of amenity or comfort results.
- *Shifting:* involves rescheduling usage away from times of high prices or demand response program events to other times. For example, a residential customer might put off running a dishwasher until later in the day, or an industrial facility might reschedule a batch production process to the prior evening hours or the next day. The lost amenity or service is made up either prior to or at a subsequent time.
- *Onsite generation:* some customers may respond by turning on an onsite or backup emergency generator to supply some or all of their electricity needs. Although the customer may have little or no interruption to their electrical usage, their net load and requirements on the power system is reduced.

Load response strategies may be enhanced with technologies and techniques that allow for fully automated demand response. Pilot projects have demonstrated this potential (Piette et al. 2005), although few customers have yet adopted fully automated demand response.

SECTION 3. BENEFITS OF DEMAND RESPONSE

EPACT requires DOE to identify the benefits of demand response in this report. This section addresses this requirement with a conceptual discussion of the various benefits of demand response, how they are derived, to whom they accrue and how to correctly ascribe value to them. The latter is important to policymakers and utilities in determining how much and what types of time-varying rates and demand response programs to include in their resource portfolios.

The following considerations underlie this discussion of demand response benefits:

- *Customers adjust their electricity usage from typical levels in expectation of receiving benefits.* These benefits must be tangible and sufficient to compensate them for the costs they incur to provide demand response, or else they will not respond.
- *Customers and program administrators incur costs in achieving demand response.* Thus, any discussion of benefits must also define and recognize costs, and quantitative assessments should identify net benefits.
- Policymakers should consider the distributional impacts—*who bears the costs and who receives the benefits*—in designing and evaluating demand response strategies.
- *The durability of benefits must be taken into account*; short-term impacts should be distinguished from long-term impacts that provide benefits over a multi-year period.
- There are important *differences in the timing and distribution of demand response benefits* for vertically integrated utilities in states without retail competition compared to regions with organized wholesale markets and retail competition.

This section begins by identifying and discussing the costs of enabling and implementing demand response. Demand response benefits are then discussed, looking at benefits to participants, collateral benefits (which include economic and reliability benefits enjoyed by some or all market participants), and other benefits that are not easily quantifiable. Appendix B provides a more detailed discussion of collateral benefits, including a discussion of differences in the timing and flow of benefits in different market structures.

Demand Response Costs

The costs of realizing demand response can be distinguished as *participant* and *system* costs (see Table 3-1). Individual customers that curtail usage incur participant costs. Demand response program administrators incur system costs to create the infrastructure required to launch and support demand response, including providing incentive payments to customers. System costs may be recovered from ratepayers (either all ratepayers or designated classes of customers) or, in some cases, through "public benefits" charges on

their electric bills. Cost recovery decisions are typically made with oversight from state regulatory agencies.

Table 3-1. Costs of Demand Response

Type of Cost		Cost	Responsibility/ Recovery Mechanism
Participant costs	Initial costs	Enabling technology investments	Customer pays; incentives may be available from public benefit or utility demand response programs to offset portion of costs
		Establishing response plan or strategy	Customer pays; technical assistance may be available from public benefits or utility demand response programs
	Event-specific costs	Comfort/inconvenience costs	Customer bears "opportunity costs" of foregone electricity use
		Reduced amenity/lost business	
		Rescheduling costs (e.g., overtime pay)	
		Onsite generator fuel and maintenance costs	
System costs	Initial costs	Metering/communications system upgrades	Level of costs and cost responsibility vary according to the scope of the upgrade (e.g., large customers vs. mass market), the utility business case for advanced metering system or upgrades, and state legislation/policies
		Utility equipment or software costs, billing system upgrades	Utility typically passes cost through to customers in rates
		Customer education	Ratepayers, public benefits funds
	Ongoing program costs[1]	Program administration/management	Costs are incurred by the administering utility, LSE or ISO/RTO and are recovered from ratepayers
		Marketing/recruitment	
		Payments to participating customers	
		Program evaluation	
		Metering/communication[2]	

[1] Ongoing program costs apply for incentive-based demand response programs and optional price-based programs only. For default-service time-varying pricing, ongoing costs are equivalent to any other default-service tariff offering.

[2] Metering/communications costs can include dedicated wire or wireless lines leased from a third-party telecommunications provider and costs to communicate pricing or curtailment information to customers or their energy services suppliers.

Customers undertaking load reductions may incur *initial* as well as *ongoing* costs to respond (see Table 3-1):

- *Initial costs* are incurred before a particular demand response behavior or action can be undertaken. They include devising a load response strategy that takes costs and benefits into account, and investing in enabling technologies to assist with load response. Enabling technologies include devices, such as "smart" thermostats, peak load controls, energy management control or information systems fully integrated into a business customer's operations, and onsite generators deployed as backup to network service. Policymakers may find it appropriate to invest in customer education and/or technology rebate programs, using ratepayer or public

benefits funds, to defray some of participating customers' initial costs, especially if they are barriers to the achievement of demand response potential.

- *Ongoing costs* are incurred by customers when they respond to high prices or demand response program events. These costs may be measurable financial costs (e.g., lost business activity, rescheduling costs such as employee overtime pay, fuel and maintenance costs from operating onsite generation) or more abstract measures of the value of electricity (e.g., the inconvenience or discomfort associated with load reductions).

Various system-wide costs are incurred in implementing demand response, which should be considered in assessing cost-effectiveness.

A variety of *system-wide costs*, which may be passed through to ratepayers or borne by utility or LSE shareholders, are associated with implementing demand response and require consideration in evaluating benefits. These include *initial costs* as well as *ongoing costs* for certain demand response options (see Table 3-1).

Initial costs can be organized into several functional categories, as follows:

Metering and communication system upgrade costs can present a significant barrier to widespread implementation of price-based DR.

- *Metering/communication system upgrade costs.* Customer retail rates typically charge only for the monthly volume of energy consumed, and for larger customers for maximum monthly demand. Time-varying tariffs (e.g., RTP, CPP) requires chronological measurement of energy usage or demand. This is typically accomplished by installing advanced metering systems (AMS) that measure and store energy usage at intervals of one hour or less and include communication links that allow the utility to remotely retrieve current usage information whenever need.[31] Metering and communications system upgrade costs depend on the existing technology as well as the applicable customer classes. Because the aggregate costs may be substantial, they can present a significant barrier to widespread implementation of time-varying tariffs especially for small and medium-sized customers and often raise cost responsibility and recovery issues. Advanced metering issues are discussed in the textbox below.
- *Utility billing system* upgrades may be necessary for some demand response options (e.g., RTP, CPP) because most legacy systems are not equipped to handle time-varying costs or usage. Pricing hourly (RTP), or having provision to price some hours differently (CPP), requires changing the way metered data are collected, processed, and stored.[32]

[31] Note that for some pricing applications (e.g., TOU rates) only usage by daily pricing period (peak and off-peak) needs to be recorded.

[32] RTP (and/or CPP) rates significantly increase the amount of usage data that must be collected (i.e., from two to four observations of customer demand and energy usage per month to at least 720 observations).

Advanced Metering to Support Price-Based Demand Response

Advanced metering is a key technology that enables many utility and customer functions. This textbox addresses four key questions regarding the role and cost of advanced metering.[33]

What is the relationship between price-based demand response and advanced metering? Price-based demand response (e.g., RTP or CPP) requires a tariff that links what the customer pays to the hourly wholesale costs of power. Advanced metering provides utilities with the capability to collect hourly interval or more frequent usage data, which is necessary to support RTP or CPP tariffs.

What is advanced metering? There are three basic types or classes of meters.

- *Conventional "kilowatt-hour" (kWh) meters* account for more than 90% of the current meter population. They record cumulative energy usage and are usually read once each month during an on-site visit by a utility employee.
- *Automated meter reading systems (AMR)* add a low power transceiver, a communication link, to a conventional kWh meter. The transceiver allows the meter to be read from a utility vehicle that drives by the customer site. These meter systems are usually limited by communication capability to collecting a single cumulative kWh reading. AMR speeds up the metering reading function and reduces utility personnel costs.
- *Advanced metering systems (AMS)*, also referred to as *advanced metering infrastructure (AMI)*, provide two features that distinguish them from conventional and AMR systems: (1) the capability to measure and store energy usage at intervals of one hour or less and, (2) a communication link that allows the utility to remotely retrieve current usage information to support customer billing and other utility operational functions.

Aren't advanced meters expensive? Advancements in communications and solid-state technology have reduced the cost of AMI to about $100 per meter if deployed system-wide. Costs to enhance and/or upgrade utility customer information and billing systems are extra. Several recent studies suggest that per-meter hardware and installation costs for advanced metering systems may be comparable to the cost of a new AMR system (King 2004).

What factors should be considered when evaluating the costs and benefits of advanced meters? Advanced metering (AMI) evaluations should consider three major categories of cost and benefit impacts:

- *Utility Operational Impacts:* AMI is first and foremost a technology for automating and improving basic utility operations. Interval metered customer usage data is essential to support billing, outage management, complaint resolution, forecasting, real-time dispatch, rate design and other utility functions. Benefits such as reductions in theft that do not impact utility revenue requirements also need to be addressed. Operational savings alone economically justified all 13 major AMI installations undertaken in North America through 2005. Utility business case analyses should account for the net impact of forecasted operational savings in estimating changes in the utility's revenue requirement from AMI deployment.
- *Demand Response Impacts:* AMI enables RTP, CPP and other forms of performance-based demand response.
- *Societal Impacts:* Societal impacts include improved customer service, environmental, equity and other benefits from more efficient utility operation.

Billing invoices must also be expanded to provide detailed, hour-by-hour accounting. Some utilities and load serving entities can accommodate these new pricing schemes at moderate cost if their existing billing systems are compatible with detailed usage accounting, while others may need to completely revamp or replace their entire billing systems (depending on the number of customers eligible for RTP or CPP).

[33]*For more information on Advanced Metering Infrastructure, see* http://www.energetics.com/madri/toolbox/.

- *Customer education* about the time-varying nature of electricity costs, potential load response strategies, and available retail market choices is often included in the rollout of demand response options.

Ongoing costs, including program administration and operation, marketing, evaluation, and customer recruitment costs, apply to incentive-based demand response programs and optional pricing tariff options that are offered in addition to customers' standard electricity tariff. For incentive-based demand response programs, additional costs also include payments to participating customers. For most default-service price-based options, there are no incremental ongoing costs relative to any other default-service tariff. However, depending on the type of metering/communication infrastructure used, ongoing equipment operation or leasing costs may apply.

Benefits of Demand Response

The benefits of demand response can be classified into three functional categories: *direct*, *collateral* and *other* benefits (see Table 3-2). Direct benefits accrue to consumers that undertake demand response actions, and collateral and other benefits are enjoyed by some or all groups of electricity consumers. Direct and collateral benefits can be quantified in monetary terms. Other benefits are more difficult to quantify and monetize.

Participant Benefits

Customers who adjust their electricity usage in response to prices or demand response program incentives do so primarily to realize *financial* benefits. In addition, they may be motivated by implicit *reliability* benefits (see Table 3-2).

- *Financial benefits* include cost savings on customers' electric bills from using less energy when prices are high, or from shifting usage to lower-priced hours, as well as any explicit financial payments the customer receives for agreeing to or actually curtailing usage in a demand response program.
- *Reliability benefits* refer to the reduced risk of losing service in a blackout. This benefit may be associated with an internalized benefit, in cases where the customer perceives (and monetized) benefits from the reduced likelihood of being involuntarily curtailed and incurring even higher costs, or societal, in which the customer derives satisfaction from helping to avoid widespread contingencies. Both are difficult to quantify but may nonetheless be important motivations for some customers.

The level of direct benefits received by participating customers depends on their ability to shift or curtail load and the incentives afforded by time-varying electricity prices and any additional program incentives that are offered.

Collateral Benefits

Demand response, through its impacts on supply costs and system reliability, produces *collateral benefits* that are realized by most or all consumers (see Table 3-2). It is these collateral benefits, which have system-wide impacts, that provide the primary motivation for policymakers' interest in demand response.

Table 3-2. Benefits of Demand Response

Type of Benefit	Recipient(s)	Benefit		Description/ Source
Direct benefits	Customers undertaking demand response actions	Financial benefits		• Bill savings • Incentive payments (incentive-based demand response)
		Reliability benefits		• Reduced exposure to forced outages • Opportunity to assist in reducing risk of system outages
Collateral benefits	Some or all consumers	Market impacts	Short-term	• Cost-effectively reduced marginal costs/prices during events • Cascading impacts on short-term capacity requirements and LSE contract prices
			Long-term	• Avoided (or deferred) capacity costs • Avoided (or deferred) T&D infrastructure upgrades • Reduced need for market interventions (e.g., price caps) through restrained market power
		Reliability benefits		• Reduced likelihood and consequences of forced outages • Diversified resources available to maintain system reliability
Other benefits	• Some or all consumers • ISO/RTO • LSE	More robust retail markets		• Market-based options provide opportunities for innovation in competitive retail markets
		Improved choice		• Customers and LSE can choose desired degree of hedging • Options for customers to manage their electricity costs, even where retail competition is prohibited
		Market performance benefits		• Elastic demand reduces capacity for market power • Prospective demand response deters market power
		Possible environmental benefits		• Reduced emissions in systems with high-polluting peaking plants
		Energy independence/ security		• Local resources within states or regions reduce dependence on outside supply

Collateral benefits can be categorized functionally as *short-term* and *long-term market impacts* as well as *reliability* benefits:

- *Short-term market impacts* are the most immediate and easily measured source of financial benefits from demand response. Broadly speaking, they are savings in variable supply costs brought about by more efficient use of the electricity system, given available infrastructure. More efficient resource use, enabled by building better linkages between retail rates and marginal supply costs, translates to short-term bill savings to consumers from avoided energy and, in some cases, capacity costs. Where customers are served by vertically integrated utilities, short-term benefits are limited to avoided variable supply costs. In areas with organized spot markets, demand response also reduces wholesale market prices for all energy

traded in the applicable market. Reductions in usage during high-priced peak periods result in a lower wholesale spot market clearing price. The amount of savings from lowered wholesale market prices depends on the amount of energy traded in spot markets, rather than being committed in forward contracts.[34]

- *Long-term market impacts* hinge on the ability of demand response to reduce system or local peak demand, thereby displacing the need to build additional generation, transmission or distribution capacity infrastructure. Because the electricity sector is extremely capital-intensive, avoided capacity investments can be a significant source of savings. However, for demand response resources to reduce capacity costs, it must be available and perform reliably at high-demand periods throughout the year because it is displacing other capacity resources.

Demand response also provides reliability benefits, reducing the probability and severity of forced outages.

- *Reliability benefits* refer to reducing the probability and severity of forced outages when system reserves fall below desired levels.[35] By reducing electricity demand at critical times (e.g., when a generator or a transmission line unexpectedly fails), demand response that is dispatched by the system operator on short notice can help return electric system (or localized) reserves to pre-contingency levels.[36] These reliability benefits can be valued according to the amount of load that demand response load reductions removed from the risk of being disconnected and the value that consumers place on reliable service (the "value of lost load").

Appendix B provides a more detailed discussion of the collateral benefits of demand response to assist policymakers' understanding of economic efficiency gains, avoided capacity benefits and capacity program design and valuation issues, the impact of different market structures on the timing and distribution of short-term and long-term demand response benefits, and the identification and valuation of reliability benefits.

---

[34] Many load-serving entities currently purchase a substantial portion of their electricity in ISO-administered spot energy markets. In New York, a state with organized wholesale markets and retail competition, over 50% of electricity is traded in day-ahead and real-time spot markets, with the rest settled in forward contracts. In New England, about 40% of the electricity volume is traded in ISO-NE's spot markets, with about 60% committed in forward contracts.

[35] At times, system dispatchers are faced with either shutting off load to parts of the system, or risk an outage that affects many more customers and load. The loads that are shut off depend on exigent circumstances. Demand response reduces load and thereby lowers the likelihood of the need to impose forced outages. It also reduces the amenity impact of a given level of load shedding because it is distributed among customers according to their willingness and ability to curtail (given appropriate incentives) rather than, for example, cutting off all customers and all load served by a given substation.

[36] Dispatchable demand response resources include direct load control programs, interruptible/curtailable rates and emergency demand response programs. Reliability benefits derive from curtailments undertaken when all available generation has been exhausted and only load reductions can serve to restore system reliability to acceptable levels.

Other Benefits

Demand response can provide several *other benefits* that accrue to some or all market participants but are not easily quantified or monetized:

- *More robust retail markets.* In competitive retail markets, default-service RTP can stimulate innovation by retail suppliers (Barbose et al. 2005), and ISO/RTO-administered demand response programs can provide value-added opportunities for marketers (Neenan et al. 2003).
- *Improved choice.* Demand response can provide expanded choices for customers in varying retail market structures (e.g. states with or without retail competition) through additional options to manage their electricity costs.

Demand response can reduce the potential for generators to exert market power by with-holding supply.

- *Market performance benefits.* Demand response can also play an important role in mitigating the potential for generators to exert market power in wholesale electricity markets by withholding supply in order to cause prices to increase. Price-responsive demand mitigates this potential because demand reductions in response to high prices increase suppliers' risk of being priced out of the market. Demand response can provide this "market performance" benefit even if it is rarely exercised because the *prospect* of demand response may be a sufficient deterrent to prevent generators from attempting market manipulation.
- *Possible environmental benefits.* Demand response may provide environmental benefits by reducing the emissions of generation plants during peak periods. It may also provide overall conservation effects, both directly from demand response load reductions (that are not made up at another time) and indirectly from increased customer awareness of their energy usage and costs (King and Delurey 2005). However, policymakers should exercise caution in attributing environmental gains to demand response, because they are dependent on the emissions profiles and marginal operating costs of the generation plants in specific regions.[37] Emission reductions during peak periods need to be balanced against possible increases in emissions during off-peak hours as well as from increased use of onsite generation.

[37] See Holland and Mansur (2004) for an analysis of regional differences in the impacts of load response on net power plant emissions, and Keith et al. (2003) for an analysis of impacts of demand response resources on net power sector emissions in New England.

SECTION 4. QUANTIFYING DEMAND RESPONSE BENEFITS

Quantifying the potential nation-wide benefits of demand response, as EPACT charges DOE to accomplish, is a large and complex undertaking and involves several functional aspects:

- *Demand Response Options*—the types of time-varying rates and demand response programs that are currently offered (or potentially available);
- *Customer Participation*—the likelihood that a customer will choose to take part in the program;
- *Response*—documenting and quantifying participants' current energy usage patterns, and determining how participants adjust that usage in response to changes in prices or incentive payments;
- *Financial Benefits*—developing methods to quantify the level and distribution of short-and long-term resource savings of load response under varying market structures;
- *Other Benefits*—identifying and quantifying any additional benefits provided by demand response resources (e.g., improved reliability); and
- *Costs*—establishing the costs associated with achieving demand response.

Given differences in market structure among states, the lack of a uniform method to measure demand response benefits and significant data limitations and gaps, which could not be overcome in the time allotted for completion of this report, DOE has chosen to take a different approach to meet its mandate.[38]

DOE's approach in meeting its EPACT mandate is to summarize and compare the results of recent studies that quantified demand response benefits.

DOE's approach is to summarize and compare the results of a number of recent studies that have attempted to quantify demand response benefits under a variety of contexts and scopes and for different regions or markets. Results are used as a basis for recommendations that can guide future efforts to quantify demand response benefits at the regional market level.

This section begins by summarizing the results of recent studies of the intensity of customer response to time-varying pricing and other demand response programs to establish the extent to which participants adjust their usage in response to price changes or incentive payments. Then, ten selected studies of demand response benefits are reviewed to assess and compare the impact of varying demand response mechanisms, study methodologies, and wholesale and retail market structure. The estimates of demand response benefits are normalized to provide insight into the importance of some factors in

[38] A comprehensive study quantifying the national benefits of demand response would have to account for different types of demand response (e.g., time-varying tariffs, incentive-based demand response programs).

determining the level of benefits attributed to demand response. Finally, recommendations on practices, protocols, and standards for improving estimates of the benefits of demand response are summarized.

Intensity of Customer Demand Response

To quantify demand response benefits in aggregate, two key inputs are: (1) measures of customer acceptance and participation rates in dynamic pricing and demand response programs, and (2) measures of the extent to which individual customers curtail load in response to either time-varying prices or demand response program incentive payments i.e. intensity).

With respect to the first input, a number of studies have characterized drivers to customer participation as part of evaluations of demand response programs or pilot tariffs. Important factors in the customer's decision to enroll and participate include the level and type of incentives offered, program requirements and conditions (e.g., notice, duration, and frequency of curtailments), customer assessment of risks and value proposition (e.g., financial consequences for failure to curtail loads), and effectiveness of program design and implementation (e.g., marketing, customer education and information, technical assistance).

With respect to the second input, a relatively large number of studies characterize the extent to which customers respond to dynamic prices and demand response programs. Results are typically reported in terms of two measures (or indicators): 1) price elasticity or 2) absolute or relative load impact (e.g., kilowatt [kW] or percent load reduction).

Customer Response to Time-Varying Prices

Price elasticity is a normalized measure of the intensity of customers' load response to prices.

A price elasticity provides a normalized measure of the intensity of customers' load changes in response to price circumstances especially for time-varying rates or demand response programs that induce load modifications directly in response to price changes. It is defined as the percentage change in usage for a one-percent change in price, and takes on values of zero and above, in absolute terms.[39] For example, if a customer's price elasticity is 0.15, then a doubling (100% change) of price results in a 15% reduction in electricity usage, other things equal. Higher elasticity values

[39] This definition is for own-price elasticity, which is always *negative*; usage goes down as price goes up. There are several variations on the concept of price elasticity that relate to different aspects of the full consequences of the change in usage. For example, a cross-price elasticity measures the consequences of reduced electricity usage on other goods. If a customer buys less electricity, then it has more money to spend on other goods and services. A substitution elasticity characterizes how a customer shifts the use of electricity in one period of the day to another (e.g. peak versus off-peak) in response to price differences between the two periods. A substitution elasticity can have a *positive* value (or zero). The discussion in this section reports elasticity values on an absolute basis, with the sign always positive, to emphasize the differences in results among studies. Appendix C provides a more complete and technically accurate characterization of the study results.

translate into increased price response by customers. Price elasticity is a useful measure because it allows for comparison of the load response of customers facing different prices.

Figure 4-1 summarizes the results of studies that estimated the price response exhibited by customers that participated in voluntary programs that involved time-varying prices (see Appendix C for more detailed information):

- several existing RTP programs available to larger industrial and commercial customers that have been operating for many years;
- an ongoing residential real-time-pricing (RTP) pilot;
- the California CPP pilot conducted in 2003-4; and
- pooled results of five residential TOU pilots conducted in the late 1970s.[40]

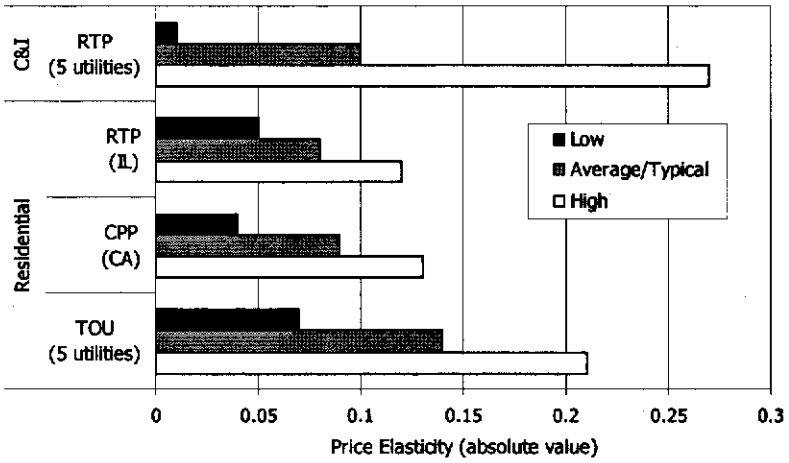

Figure 4-1. Customer Response to Time-Varying Prices: Price Elasticity Estimates

For each study, the low, average (or typical), and high estimates of price response are illustrated, although the interpretation of the low to high range values varies somewhat across studies. For example, the range in price elasticity values for a residential RTP pilot in Illinois are attributed to demographic differences within the pilot group, while for a pilot CPP program in California, the range in elasticity values primarily reflects climatic differences and saturation of air conditioning equipment among participant groups. For the residential TOU studies, the range of elasticity values reflects results across the five pilots.

Average price elasticities among the studies are fairly similar, ranging from 0.08 to 0.14 (in absolute value). The average elasticity value for RTP for large industrial and commercial customers (0.10) represents a typical value reported by several studies. The low and high elasticity values for commercial and industrial RTP customers exhibit the largest variation (i.e., 0.01 to 0.27) and reflect differences in the price responsiveness of

[40] See Appendix C for a more in-depth description of these studies and their results.

various market segments. Studies of large customers' response to RTP consistently find large differences in price elasticity across business categories. For example, a recent study of about 150 customers on RTP at Niagara Mohawk reported average elasticities of 0.16 for manufacturing customers, 0.10 for government/education customers, 0.06 for commercial/retail and 0.04 for healthcare facilities (Goldman et al. 2005).

The Residential RTP study (Illinois) reported similar price elasticities as the California residential CPP study (i.e., 0.08 to 0.09); both studies were conducted during a comparable time period (2004) but in different markets. Studies of residential customer response to time-varying prices often report that price elasticity is driven in part by the number of electricity devices present in the home. Climate also has a discernable affect, as do occupant characteristics and circumstances that affect when they are home and likely to be able to shut off devices or reduce usage.

Customer Response to Load Control Programs

Over one hundred U.S. utilities report that they currently offer residential or small commercial DLC programs that primarily target customers with air conditioning or domestic water heating load-control devices (EIA 2004).[41] A number of these programs have conducted relatively recent measurement and evaluation studies with results that are publicly available.

In some demand response programs (e.g., where customers do not directly respond to prices), their response is typically measured by the amount of load reduced.

For DLC programs and other types of demand response programs where customers are not directly responding to a price, the intensity of customers' response is typically measured in terms of an absolute or relative load impact (e.g., kW of load curtailed or percent of the customer's total load that is curtailed, either through equipment cycling or shedding).

Figure 4-2 summarizes reported load reduction estimates for large groups of customers with water heating load controls and various types of control strategies for air conditioning equipment (e.g., cycling the device on and off at a specified time interval, shutting the device off for a period of time, or resetting a thermostat set point) [see Appendix C for more detailed information].[42] Residential water heating DLC programs have typically yielded load reductions in the range of 0.3 to 0.6 kW per house; the magnitude and timing of the load impact depends on household and equipment size, ground water temperature and household usage patterns. DLC programs targeting residential air conditioning (A/C) have reported load reductions ranging from approximately 0.4 to 1.5 kW per customer over the course of an event. The magnitude of the load reduction per customer can strongly depend on climate, the control strategy deployed (e.g. 100% shed, duty cycling, thermostat reset) and the customer's air

[41] Demand-side management efforts include energy efficiency and/or load management programs.
[42] The results indicate the range of possible load impacts, although the values across studies are not readily comparable because of differences in program design features, cycling strategies, and climate.

conditioning usage levels absent load control. This is illustrated in Figure 4-2 by several studies that reported low and high load reduction values based on testing different cycling strategies at various temperature levels.

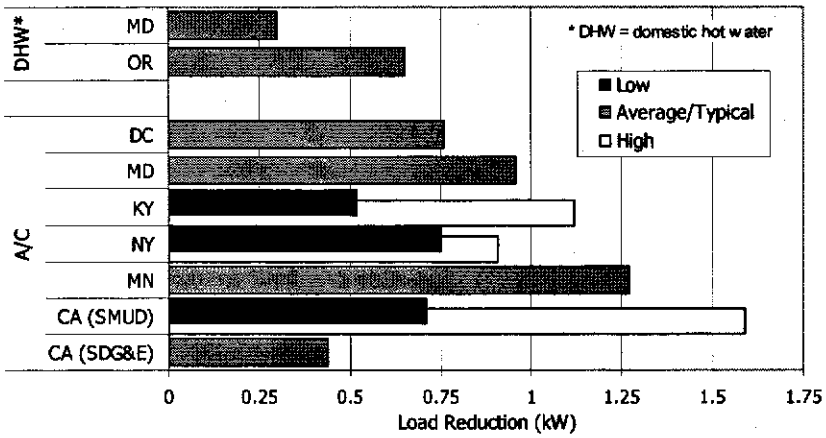

Figure 4-2. Estimated Load Impacts from Direct Load Control Programs

Impact of Enabling Technologies on Price Response

Studies of pilot programs combining pricing with enabling technologies provide important insights on the technical potential for demand response.

Some utilities have offered pilot programs targeted to mass-market customers that integrate CPP with enabling technology, specifically load control devices that receive price signals and can be programmed by customers to reduce A/C or other loads during critical peak periods (see Figure 4-3 and Appendix C). Several of these programs have obtained promising results. For example, in Florida, Gulf Power reported average load reductions of 40% during critical peak periods for groups of customers that could control multiple loads (e.g. A/C, water heating, pool pumps) (Levy Associates 1994). In California, a recent Statewide Pricing Pilot (SPP) sought to quantify the impact of "smart thermostats" with critical peak prices. The average load reduction of 220 residential customers with smart thermostats during critical peak days was approximately 0.64 kW, a 27% reduction during peak periods, approximately two-thirds of which was attributed to use of the smart thermostat. Among the 235 small business customers in the California SPP, the average peak period load reduction was about 14%, although the relative impact of the enabling technology was even more pronounced. These studies may reveal the technical potential for demand response in certain market segments when time-varying pricing is combined with enabling technology.

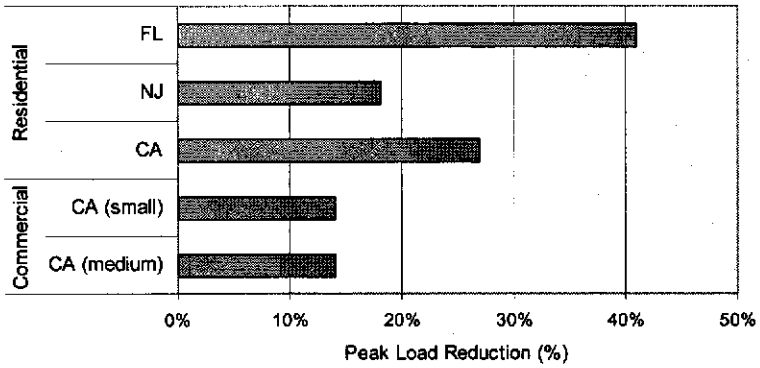

Figure 4-3. Load Response from Critical Peak Pricing and Demand Response Enabling Technologies

Summary

The following key findings and lessons can be drawn from this review of studies that examine customer response to time-varying prices and different types of demand response programs:

- Many initiatives have been undertaken that quantify the price-to-quantity relationship for various types of dynamic pricing and demand response programs. These data are critical because policymakers require price elasticity and load impact estimates as an input in estimating the benefit of specific demand response programs.
- Based on several of the more comprehensive studies, it is reasonable to assume that a group of large customers participating in well-designed RTP tariffs respond with a substitution elasticity of around 0.10 on average, which means that when peak prices rise by 50%, these customers will lower or shift their load to other times of the day by 5%.[43]
- Elasticities for groups of residential customers enrolled in TOU rates with significant differentials in peak to off-peak prices (e.g. factor of three) are also about 0.10—0.15.
- A small number of studies of residential customers on CPP rates, with very high critical peak prices ($.50/kWh or higher) report that that customers reduce load by an even greater amount than is reported in other studies for TOU. The recent California pilot, where the two designs were tested side-by-side, reports that the difference is almost a factor of two. However, the difference may be due to the large price differences between the two rate offerings.[44]

[43] The ability of customers enrolled in RTP tariffs to respond to prices is varied. Several studies report that 65-75% of the total measured price response is provided by about 20% of the customers on RTP rates.
[44] Two customers with identical price response capability (price elasticity) may exhibit different levels of load response if they face vastly different prices. This is because the nature of the response may increase with the nominal level of prices. The price elasticities estimated for TOU rates may be smaller than for CPP rates, because the customers never faced the higher CPP prices.