Demand Response Assessment

INTRODUCTION

This appendix provides more detail on some of the topics raised in Chapter 4, "Demand Response" of the body of the Plan. These topics include

1. The features, advantages and disadvantages of the main options for stimulating demand response (price mechanisms and payments for reductions)
2. Experience with demand response, in our region and elsewhere
3. Estimates of the potential benefits of demand response to the power system

PRICE MECHANISMS

Real-time prices

The goal of price mechanisms is the reflection of actual marginal costs of electricity production and delivery in retail customers' *marginal* consumption decisions. One variation of such mechanisms is "real-time prices" -- prices based on the marginal cost of providing electricity for each hour. This does not mean that every kilowatt-hour customers consume needs to be priced at marginal cost. But it does mean that consumers need to face the same costs as the power system for their *marginal* use.

Real-time prices, if we can devise variations that are acceptable to regulators and customers, have the potential to reach many customers. Real-time prices can give these customers incentives that follow wholesale market costs very precisely every hour. Once established, real-time prices avoid the transaction costs of alternative mechanisms. For all of these reasons, the potential size of the demand response from real-time prices is probably larger than other mechanisms.

However, real-time prices have not been widely adopted for a number of reasons:

1. Most customers would need new metering and communication equipment in order to participate in real-time pricing. Currently, most customers' meters are only capable of measuring total use over the whole billing period (typically a month). Real-time prices would require meters that can measure usage in each hour. Also, some means of communicating prices that change each hour would be required. It's worth noting that more capable meters are also necessary for alternatives such time-of-use metering, and for such programs as short term buybacks and demand side reserves.
2. Currently, there is no source of credible and transparent real-time wholesale prices for our region. Any application of real-time retail prices will need all parties' trust that the prices are fair representations of the wholesale market. The hourly prices from the California PX were used as the basis for some deals in our region until the PX was closed in early 2001, but prices from a market outside our region were regarded as less-than-ideal even while they were still available. Now the Cal PX is closed, and a credible

regional source is needed. This is a problem that affects many of the other mechanisms for demand response[1] as well.

3. Some customers and regulators are concerned that real-time prices would result in big increases in electricity bills. While the argument can be made that such increases would be useful signals to consumers[2], the result could also be big decreases in bills. In either case, however, many customers and regulators are concerned with questions of unfair profits or unfair allocation of costs if real-time prices are adopted. The Council shares this concern.
4. Even if price increases and decreases balance over time, the greater volatility of real-time prices is a concern. Customers are concerned that more volatile prices will make it hard for them to plan their personal or business budgets. Regulators are concerned that more volatile prices will make it a nightmare to regulate utilities' profits at just and reasonable levels. The volatility is moderated if the real-time pricing applies only to marginal consumption, but it is still greater than consumers are used to.
5. Some states' utility regulation legislation constrains the definition of rates (e.g. rates must be numerically fixed in advance, not variable based on an index or formula).

With time, some of these issues can probably be solved, making real-time prices more practical and more acceptable to customers and regulators. For example:

Metering and communication technology has improved greatly. New meters not only offer hourly metering and two-way communication but also other features, such as automatic meter reading and the potential for the delivery of new services, that may make their adoption cost-effective.

Customers and regulators' concerns with fairness and volatility may be relieved by such variations of real-time prices as the Georgia Power program. That program applies real-time prices to <u>increases</u> or <u>decreases</u> from the customer's base level of use, but applies a much lower regulated rate to the base level of use itself. Compared to application of real-time prices to the total use of the customer, this variation reduces the volatility of the total bill very significantly.

Concerns with fairness may also moderate, as it is better understood that "conventional" rates have their own problems with fair allocation of costs among customers.

Time-of-use prices

We could think of "time-of-use prices" -- prices that vary with time of day, day of the week or seasonally -- as an approximation of real-time prices. Time-of-use prices are generally based on the expected average costs of the pricing interval (e.g. 8 a.m. to 6 p.m. January weekdays).

While time-of-use prices, like real-time prices, require meters that measure usage over subintervals of the billing period, they have some advantages over real-time prices. A significant advantage of time-of-use rates is that customers know the prices in advance (usually for a year or

[1] For example, participation in short term buyback programs is enhanced when customers have confidence that their payments are based on a price impartially determined by the wholesale market rather than simply a payment the utility has decided to offer.

[2] For example, bills might rise for those customers whose use is concentrated in hours when power costs are high. While those customers would be unhappy about the change, their increased bills could be seen as an appropriate correction of a traditional misallocation of the costs of supplying them -- traditional rates shifted some of the cost of their service to other customers. Real-time prices would also increase the bills of all customers in years like 2000-2001, when wholesale costs for all hours went up dramatically. While customers are never happy to see bills rise, the advantage of such a <u>prompt</u> rise in prices would be a similarly prompt demand response, reducing overall purchases at high wholesale prices. This is a better result than the alternative of raising rates <u>later</u> to recover the utilities' wholesale purchase costs, after the costs have already been incurred.

more). This avoids the necessity of communication equipment to notify customers of price changes. It also makes bills more predictable, which is desirable to many customers and regulators.

A significant disadvantage, compared to real-time prices, is that prices set months or years in advance cannot do a very good job of reflecting the real-time events (e.g. heat waves, droughts and generator outages) that determine that actual cost of providing electricity. As a result, time-of-use pricing as it has usually been applied cannot provide efficient price signals at the times of greatest stress to the power system, when customers' response to efficient prices would be most useful.

"Critical peak pricing" is a variant of time-of-use pricing that could be characterized as a hybrid of time-of-use and real-time pricing. This variant leaves prices at preset levels, but allows utilities to match the timing of highest-price periods to the timing of shortages as they develop; these variations provide improved incentives for demand response.

Time-of-use prices will affect customers differently, depending on the customers' initial patterns of use and how much they respond to the prices by changing their patterns of use. While customers whose rates go up will be inclined to regard the change as unfair, regulators can mitigate such perceptions with careful rate design and making a clear connection between cost of service and rates.

PAYMENTS FOR REDUCTIONS

Given the obstacles to widespread adoption of pricing mechanisms, utilities have set up alternative ways to encourage load reductions when supplies are tight. These alternatives offer customers payments for reducing their demand for electricity. In contrast with price mechanisms, which vary the cost of electricity to customers, these offers present the customers with varying prices they can receive as "sellers". Utilities have offered to pay customers for reducing their loads for specified periods of time, varying from hours to months or years.

Short-term buybacks

Short-term programs can be thought of as mostly load shifting (e.g. from a hot August afternoon to later the same day). Such shifting can make investment in a "peaking" generator[3] unnecessary. The total amount of electricity used may not decrease, and may even increase in some cases, but the overall cost of service is reduced mostly because of reduced investment in generators and the moderating effect on market prices. Short-term programs can be expected to be exercised and have value in most years, even when overall supplies of energy are plentiful.

Generally, utilities establish some standard conditions (e.g. minimum size of reduction, required metering and communication equipment, and demonstrated ability to reduce load on schedule) and sign up participants before exercising the program. Then, one or two days before the event:

1. The utility communicates (e.g. internet, fax, phone) to participating customers the amount of reduction it wants and the level of payment it is offering.
2. The participants respond with the amount of reduction they are willing to contribute for this event.

[3] A generator that only runs at peak demands and is idle at other times.

3. The utility decides which bids to accept and notifies the respondents of their reduction obligation.
4. The utility and respondents monitor their performance during the event, and compensation is based on that performance.

Generally participants are not penalized for not responding to an offer. However, once a participant has committed to make a reduction there is usually a penalty if the obligation is not met.

Both BPA and PGE regarded their Demand Exchange programs as successful. Between the two programs, participating customers represented nearly 1,000 megawatts of potential reductions. Actual reductions sometimes exceeded 200 megawatts.

As the seriousness of the supply shortage of the 2000-2001 period became clearer, the participation in both utilities' Demand Exchange programs declined, but largely because customers who had been participating negotiated longer-term buybacks instead.

These programs require that customers have meters that can measure the usage during buyback periods. The programs also require that the utility and customer agree on a base level of electricity use from which reductions will be credited. The base level is relatively easy to set for industrial customers whose use is usually quite constant. It's more complicated to agree on base levels for other customers, whose "normal" use is more variable because of weather or other unpredictable influences.

Longer-term buybacks

Longer-term programs, in contrast to short-term buybacks, generally result in an overall reduction of electricity use. They are appropriate when there is an overall shortage of electricity, rather than a shortage in peak generating capacity.

Most utility systems, comprised mostly of thermal generating plants, hardly ever face this situation. If they have enough generating capacity to meet their peak loads, they can usually get the fuel to run the capacity as much as necessary. The Pacific Northwest, however, relies on hydroelectric generating plants for about two-thirds of its electricity. In a bad water year we can find ourselves with generating capacity adequate for our peak loads, but without enough water (fuel) to provide the total electricity needed.

This was the situation in 2000-2001, and the longer-term buybacks that utilities negotiated with their customers were reasonable responses to the situation. We faced an unusually bad supply situation in those years, however. We shouldn't expect to see these longer term buybacks used often even here in the Pacific Northwest, and hardly ever in other regions with primarily thermal generating systems.

Generally, buybacks avoid some of the problems of price mechanisms, and they have been successful in achieving significant demand response. Utilities have been able to identify and reach contract agreements with many candidates who have the necessary metering and communication capability. . The notification, bidding and confirmation processes have worked. Utilities in our region have achieved short-term load reductions of over 200 megawatts. Longer-term reductions of up to 1,500 megawatts were achieved in 2001 when the focus changed because of the energy shortages of the 2000-2001 water year.

In principle, the marginal incentives for customers to reduce load should be equivalent, but buybacks have some limitations relative to price mechanisms. Buybacks generally impose transaction costs by requiring agreement on base levels of use, contracts, notification, and explicit compensation. The transaction costs mean that they tend to be offered to larger customers or easily organized groups; significant numbers of customers are left out. Transaction costs also mean that some marginally economic opportunities will be passed--there may be times when market prices are high enough to justify some reduction in load, but not high enough to justify incurring the transaction cost necessary to obtain the reduction through a buyback.

Demand side reserves

Another mechanism for achieving demand response is "demand side reserves," which can be characterized as options for buybacks.

The power system needs reserve resources to respond to unexpected problems (e.g. a generator outage or surge in demand) on short notice. Historically these resources were generating resources owned by the utility and their costs were simply included in the total costs to be recovered by the utility's regulated prices. Increasingly however, other parties provide reserves through contracts or an "ancillary services" market. In such cases, the reserves are compensated for standing ready to run and usually receive additional compensation for the energy produced if they are actually called to run.

The capacity to reduce load can provide much the same reserve service as the capacity to generate. The price at which the customer is willing to reduce load, and other conditions of his participation (e.g. how much notice he requires, maximum and/or minimum periods of reduction) will vary from customer to customer. In principle, customers could offer a differing amount of reserve each day depending on his business situation.

The California Independent System Operator administers an ancillary services market that has used demand side reserves in some cases. Their early experience has been that most load cannot be treated the same as generating reserve in every detail, but that demand side reserve can be useful. Analysis of their experience is continuing.

The metering and communication equipment requirements, and the need for an agreed-upon base level of use, are essentially the same for demand side reserve participants as for short-term buyback participants. Demand side reserve programs may have a potential advantage to the extent that they can be added to an existing ancillary services market, compared to setting up stand-alone buyback programs.

Payments for reductions -- interruptible contracts

Utilities have negotiated interruptible contracts with some customers for many years. An important example of these contracts was Bonneville Power Administration's arrangement with the Direct Service Industries (DSI), which allowed BPA to interrupt portions of the DSI load under various conditions. In the past, these contracts have usually been used to improve reliability by allowing the utility to cut some loads rather than suffer the collapse of the whole system. Those contracts were used very seldom. Now these contracts can be seen as an available response to price conditions as well as to reliability threats. We can expect that participants and utilities will pay close attention to the frequency and conditions of interruption in future contracts, and we can imagine a utility having a range of contract terms to meet the needs of different customers.

Payments for reductions -- direct control

A particularly useful form of interruptible contract gives direct control of load to the utility. Part of BPA's historical interruption rights for DSI loads was under BPA direct control. Not all customers can afford to grant such control to the utility. Of those who can, some may only be willing to grant control over part of their loads. Direct control is more valuable to the utility, however, since it can have more confidence that loads will be reduced when needed, and on shorter notice. Advances in technology could mean expansion of direct control approaches. The ability to embed digital controls in residential and commercial appliances and equipment make it possible to, for example, set back thermostats somewhat during high cost periods. While the individual reductions are small, the aggregate effect can be large. Consumers typically have the ability to override the setbacks. Puget Sound Energy carried out a limited test of controlling thermostat setback. Most consumers were unaware that any setback had occurred. The adoption of advanced metering technologies for other reasons will facilitate the use of direct control.

SUMMARY OF ALTERNATIVE MECHANISMS

Table H-1 summarizes the alternative mechanisms and some of their attributes. Staff has offered subjective evaluations of each mechanism to stimulate comment and discussion.

Table H-1: Types of Demand Response Programs and Attributes

Type of Program	Primary Objective: Capacity or Energy?	Time span	Size of Potential Resource	Flexible for Customer?	Flexible for Utility?	Predictable, Reliable Resource for Utility?
Real-time Prices	Both	One hour to several hours	+++ (depending on extent applied)	++	++	-
Time-of-use Prices	Capacity	Several hours	++	++	--	-
Short Term Buybacks	Capacity	Several hours (possibly more)	++	++	+	+ (once customer committed)
Long Term Buybacks	Energy	Several months	+	--	--	+++
Standing Offer (e.g. 20/20)	Energy	Several months	+	++	--	-
Demand side reserves	Capacity	Hours or longer	+	++	++	+
Interruptible Contracts	Capacity	Hours or longer	+	--	++	++
Direct Control	Capacity	Minutes, Hours or longer	+	---	+++	+++

For example, staff's evaluation suggests that time-of-use prices:

- have significant potential for load reduction, but somewhat less than real-time prices;
- have the primary objective of reducing capacity requirements;
- are flexible for the customer -- the customer can decide how to respond depending on his real time situation;
- are relatively inflexible for the utility -- it is committed to the price structure in advance for an extended period;
- is not a very predictable resource for the utility – customers' response may vary from one day to the next (although more experience may help the utility predict that response more accurately).

Or, long term buybacks:

- have significant potential for load reduction, but less than time-of-use prices;
- have the primary objective of reducing energy requirements;
- are relatively inflexible for both customer and utility (because they are both committed to the terms of the buyback over a long term)
- are a predictable resource for the utility (once the contract is signed).

EXPERIENCE

Experience with demand response is growing constantly, so that any attempt to describe it comprehensively is likely to be incomplete and is certain to go out of date quickly. Rather than attempt a comprehensive account, this section presents a number of significant illustrations of experience around the U.S.

RTP Experience

Georgia Power

Georgia Power has 1,700 customers on real-time prices. These customers, who make up about 80 percent of Georgia Power's commercial and industrial load (ordinarily, about 5,000 megawatts), have cut their load by more than 750 megawatts in some instances. The program uses a two-part tariff, which applies real-time prices to <u>increases</u> or <u>decreases</u> from the customer's base level of use, but applies a much lower regulated rate to the base level of use itself. As a result, the total power bills don't vary in proportion to the variation of the real-time prices, but customers do have a "full strength" signal of the cost of an extra kilowatt-hour of use (and symmetrically, the value of a kilowatt-hour reduction in use).

Duke Power

Duke Power has a similar two-part tariff that charges real-time prices to about 100 customers with about 1,000 megawatts of load. Duke has observed reductions of 200 megawatts in these customers' load in response to hourly prices above 25 cents per kilowatt-hour.

Niagara Mohawk

Niagara Mohawk has a one-part real-time price tariff that charges real-time prices for all use of its largest industrial customers. More than half of the utility's original customers in this class have moved to non-utility suppliers, and many of those remaining have arranged hedges to reduce their vulnerability to volatility of real-time prices.

Critical Peak Pricing Experience

Gulf Power

Gulf Power offers a voluntary program for residential customers that includes prices that vary by time of day along with a programmable control for major electricity uses (space heating and cooling, water heating and pool pump, if present). While this program mostly falls in the "time-of-use pricing" category to be described next, it has an interesting component that is similar to real-time pricing--"Critical" price periods:

The Critical price (29 cents per kilowatt-hour) is set ahead of time, like the Low (3.5 cents), Medium (4.6 cents) and High (9.3 cents) prices, but unlike the other prices, the hours in which the Critical price applies are not predetermined. The customer knows that Critical price periods will total no more than 1 percent of the hours in the year, but not when those periods will be, until 24 hours ahead of time. Gulf Power helps customers program their responses to Critical periods ahead of time, although they can always change their response in the event.

Customers appear very satisfied by this Gulf Power program. Customers in the program reduced their load 44 percent during Critical periods, compared to a control group of nonparticipants.

TOU Experience

The Pacific Northwest

Puget Sound Energy offered a time-of-use pricing option for residential and commercial customers. There are about 300,000 participants in the program. PSE's analysis indicates that this program reduced customers' loads during high costs periods by 5-6 percent. However, analysis showed that most customers paid slightly more under time-of-use pricing than they would have under conventional rates. PSE has ended the program, though a restructured program might be proposed later if careful analysis suggests it would be effective.

In Oregon, time-of-use pricing options have been offered to residential customers of Portland General Electric and PacifiCorp since March 1, 2002. So far about 2,800 customers have signed up, and early measures of satisfaction are encouraging, but data are not yet available on any changes in their energy use patterns.

California

Time of use rates are now required for customers larger than 200 kilowatts, and critical peak pricing is available for those customers. The effect of the critical peak prices on customers who have selected that option is estimated to provide a load reduction potential of about 16 megawatts in 2004.

A pilot program testing the effectiveness of critical peak pricing for residential customer is completing its second year. Analysis of the first year's experience estimated own price elasticities of peak demand in the –0.1 to –0.4 range, similar to the results of the Electric Power Research Institute study described below.

There have been many other time-of-use pricing programs elsewhere in the U.S. Rather than describe a number of examples, it should suffice to say that a study funded by the Electric Power Research Institute concluded that 25 years of studies indicated that "peak-period own-price elasticities range from -0.05 to -0.25 for residential customers, and -0.02 to -0.10 for commercial and industrial customers." Stripped of the jargon, this means that a time-of-use rate schedule that increases peak period rates by an assumed 10 percent would lead to a 0.5 to 2.5 percent reduction in residential peak use, and a 0.2 to 1.0 percent reduction in commercial and industrial peak use. While the assumed 10 percent rate increase is only illustrative, it is not exaggerated; PSE's peak time rates are about 10 percent higher than its average rates, and PGE's peak time rates are 67 percent higher than its average rates.

Short-term Buyback Experience

The historical experience with demand response is limited, and most of it is from short-term situations of tight supply and/or high prices (i.e. episodes of a few hours in length). Therefore we'll examine the potential for short-term demand response first, and turn to longer-term demand response later.

Pacific Northwest

B.C. Hydro offered a form of short-term buyback as a pilot program quite early -- in the winter of 1998-1999. The utility offered payment to a small group of their largest customers for reductions in load. The offer was for a period of hours when export opportunities existed and B.C. Hydro had no other energy to export. Compensation was based on a "share the benefits" principle, sharing the difference between the customers' rates and the export price equally between B.C. Hydro and the customer.

The program was exercised once during the pilot phase, realizing about 200 megawatts of reduction. The overall evaluation of the program was positive and it has been adopted as a continuing program by B.C. Hydro.

Bonneville Power Administration, Portland General Electric and some other regional utilities offered another form of short-term buyback beginning in the summer of 2000. This program was called the Demand Exchange. The Demand Exchange was mostly limited to large industrial customers who had the necessary metering and communication equipment and who had demonstrated their ability to reduce load on call. Participating customers represented over 1,000 megawatts of potential reductions, and over 200 megawatts of reductions were realized in some events.

An exception to the focus on large customers was the participation of Milton-Freewater Light and Power, a small municipal utility with about 4,000 customers. Milton-Freewater participated by controlling the use cycles of a number of their customers' residential water heaters.

California

Investor-owned utilities in California have over 1,600 megawatts of demand response available in June 2004. Over 1,000 megawatts of that total are in interruptible contracts, with about 300 megawatts in air conditioning cycling and smart thermostat programs, about 150 megawatts in demand bidding programs and the remainder in critical peak pricing and backup generation programs.

The California Independent System Operator (CAISO) has reduced its demand response programs in recognition of the programs offered by California utilities and the California Power Authority. The CAISO continues its "Participating Load Program (Supplemental and Ancillary Services)," which includes demand reductions as a source of supplemental energy and ancillary services (non-spinning reserves and replacement reserves). In this program demand reductions are bid into the ancillary services market similarly to generators' capacity and output.

The California Power Authority offers a variant of interruptible contract, with capacity payments every month based on the customer's commitment to reduce load, and energy payments based on actual reductions when the customer is called upon to do so. In June of 2004 this program was estimated to have a demand reduction capability of over 200 megawatts.

New York Independent System Operator

The New York Independent System Operator (NYISO) has three demand response programs, the Emergency Demand Response Program (EDRP), the Day-Ahead Demand Response Program (DADRP) and Installed Capacity Special Case Resources (ICAP SCR).[4]

The EDRP is, as the name suggests, an emergency program that is exercised "when electric service in New York State could be jeopardized." Participants are normally alerted the day before they may be called upon to reduce load; they are usually notified that reductions are actually needed at least 2 hours in advance. Participants are expected, but not required, to reduce their loads for a minimum of four hours, and are compensated at the local hourly wholesale price, or $500 per megawatt hour, whichever is higher. Reductions are calculated as the difference between metered usage in those hours and the participants' calculated base loads (CBLs), which are based on historical usage patterns.

The DADRP allows electricity users to offer reductions to the NYISO in the day-ahead market, in competition with generators. If the reduction bid is accepted, the users are compensated for reductions based on the area's marginal price. The users are obligated to deliver the reductions and are charged the higher of day-ahead or spot market prices for any shortfall in performance.

The ICAP SCR program pays qualified electricity users for their <u>commitment</u> to reduce loads if called upon during a specified period, "during times when the electric grid could be jeopardized." Users receive additional payments when they are actually called and deliver reductions, at rates up to $500 per megawatt hour. Qualified electricity users cannot participate in both the EDRP and the ICAP SCR at the same time, and ICAP SCR resources are called first.

During the summer of 2003, these NYISO programs resulted in the payment of more than $7.2 million to over 1,400 customers, who reduced their peak electricity loads by 700 megawatts.

PJM Interconnection

PJM Interconnection is the regional transmission operator of a system that covers 8 Mid Atlantic and Midwestern states and the District of Columbia. It serves a population of about 35 million, with a peak load of about 85,000 megawatts. PJM has operated demand response programs for several years.

PJM's demand response programs are categorized as "Emergency" and "Economic" options. PJM takes bids from end-use customers specifying reduction amounts and compensation

[4] For more details, see http://www.nyiso.com/services/documents/groups/bic_price_responsive_wg/demand_response_prog.html

requirements for the next day. These bids are considered alongside bids from generators, and demand reduction bids can set the market clearing "locational marginal price" (LMP, the marginal cost of service for each zone in the system) in the same way as a generator's bid. Load reductions in their "Emergency" category are paid at each hour's LMP, or $500 per megawatt-hour, whichever is greater. Load reductions in their "Economic" category are paid the LMP less the retail rate if the LMP is less than $75 per megawatt-hour, or the whole LMP if it is higher than $75 per megawatt-hour.

PJM also has an "Active Load Management" (ALM) program that compensates customers for: allowing PJM to have direct control of some loads; committing to reduce loads to a specified level; or committing to reduce loads by a specified amount.

In total PJM demand response programs had over 2,000 megawatts of potential load reductions participating in 2003, and over 3,500 megawatts of potential load reductions in 2004.

ISO New England

The Independent System Operator (ISO) of the New England Power Pool operates the electrical transmission system covering the 6 New England states, with a population of 14 million people and a peak load of over 25,000 megawatts. Its demand response programs had 400 megawatts of capacity in 2004, about double the capacity in 2002.

ISO New England demand response programs share some features with those of the NYISO and PJM, in that they fall into "economic" and "reliability" categories. The "economic" category is voluntary -- qualified customers[5] are notified when the next day's wholesale price is expected to be above $.10 per kilowatt-hour for some period. They can voluntarily reduce their load during that period and be compensated at the greater of the real time wholesale price, or $.10 per kilowatt-hour. Their reduction is computed based on their recent load history, adjusted for weather conditions. There is no penalty for choosing not to reduce load for these customers.

In the "reliability" category customers can commit to reducing load at the call of the ISO, and be compensated based on the capacity they have committed and the energy reduction they actually deliver when called upon. The compensation for capacity (ICAP) is based on a monthly auction. The compensation for energy is the greater of the real time price or a minimum of $.35 or $.50 per kilowatt-hour, depending on whether the customer is committed to responding in 2 hours or 30 minutes, respectively. If a customer does not deliver the committed reduction it is compensated for energy reduction based on the actual performance, but the ICAP payment is reduced to the level of delivered reduction. The ICAP payment remains at that reduced level until another load reduction event; the customer's performance in that event resets the ICAP level higher or lower.

ISO New England recently issued a request for proposals to remedy a localized shortage of generation and transmission in Southwest Connecticut. It selected a combination of resources that included demand response amounting to 126 megawatts in 2004 and rising to 354 megawatts in 2007. These resources were called on in August of 2004 and delivered over 120 megawatts within 30 minutes. In that event, roughly another 30 megawatts of load reduction were realized elsewhere in ISO New England's territory.

[5] Customers with the ability to reduce loads by 100 kilowatts, with appropriate metering and communication equipment.

Longer-term Buyback Experience

As high wholesale prices and the drought in the Pacific Northwest continued, utilities began to negotiate longer-term reductions in load with their customers. BPA found the largest reductions, mostly in aluminum smelters but also in irrigated agriculture. Idaho Power, PGE, the Springfield Utility Board (SUB) and the Chelan Public Utility District negotiated longer-term reductions with large industrial customers. Idaho Power, Grant County Public Utility District and Avista Utilities negotiated longer-term reductions with irrigators. The total of these buybacks varied month to month but reached a peak of around 1,500 megawatts in the summer of 2001.

There were also "standing offer" buybacks offered by several utilities in 2001. Most of these offers were to pay varying amounts for reductions compared to the equivalent billing period in 2000. The general structure of these offers was a further savings on the bill if the reduction in use was more than some threshold. For example, a "20/20" offer gave an additional 20 percent off the bill if the customers' use was less than 80 percent of the corresponding billing period in 2000. Since the customer's bill was reduced more or less proportionally to his usage already, this amounted to roughly doubling his marginal incentive to save electricity. Utilities usually reported that many customers qualified for the discounts. However, attributing causation to the standing offers vs. quick-response conservation programs many utilities were running at the same time vs. governors' appeals for reductions, etc. is very difficult.

The Eugene Water and Electric Board had a standing offer that based its incentives more directly on current market prices. From April through September of 2001, 29 of EWEB's larger customers were paid for daily savings (compared to the corresponding day in 2000) based on the daily Mid-Columbia trading hub's quotes for on-peak and off-peak energy. Customers reduced their use of electricity by an average of 14 percent, and divided a total savings of $6.5 million with the utility.

ESTIMATES OF POTENTIAL BENEFITS OF DEMAND RESPONSE

Potential size of resource

One way to arrive at a rough estimate of short-term demand response is to use price elasticities[6] that have been estimated based on response to real-time prices elsewhere. Though we're unlikely to rely on real-time prices, at least in the near future, the other instruments we've described can provide similar incentives[7], resulting in similar demand reductions.

Price elasticities have been estimated based on data from a number of American and other utilities. The elasticities vary from one customer group and program to another, from near zero to greater than -0.3. For example, we can assume, conservatively:

1. a -0.05 elasticity as the lower bound of overall consumer responsiveness,
2. a $60 per megawatt hour average cost of electricity divided equally between energy cost and the cost of transmission and distribution
3. a $150 per megawatt hour cost of incremental energy at the hour of summer peak demand, and

[6] Price elasticity is a measure of the response of demand to price changes -- the ratio of percentage change in demand to the percentage change in price. A price elasticity of –0.1 means that a 10 percent increase in price will cause a 1 percent decrease in demand.
[7] For example, a customer with conventional electricity rate of $0.06 per kilowatt-hour might get a buyback offer of $0.15 per kilowatt-hour in a given hour. A real-time price of $0.21 per kilowatt hour would offer a similar incentive to reduce use in that hour -- in either case he is better off by $0.21 for each kilowatt hour reduction.

4. a 30,000 megawatts regional load at that hour.

For these conditions, the amount of load reduction resulting from real-time prices would be 1,603 megawatts[8]. Actual elasticities could well be larger and actual prices seem quite likely to be higher on some occasions. In either of these cases, the load reduction would be increased.

This very rough estimate could be refined, although the basic conclusion to be drawn seems clear – even if this estimate is wrong by a factor of 2 or 3, the potential is significant, and demand response should be pursued further.

The Value of Load Reduction (avoided cost)

The primary focus of analysis was the estimation of costs avoided by demand response. These avoided costs establish the value of demand response, and provide guidance for incentive levels in demand response programs.

We used three different approaches to the estimation of avoided cost. Each of these approaches has shortcomings, but together they suggest very strongly that development of demand response will reduce total system cost and reduce risk.

The first two of these estimates focus on the costs of meeting peak loads of a few hours' duration ("capacity problems"). These are not the only situations in which demand response can be useful, but they are the most common. These estimates address the net power system costs of serving incremental load, in a world of certainty.

If our region faced a fully competitive power market, the cost avoided by demand response would be the hourly price of power in that market. Over the long run, hourly prices at peak hours should tend to approach the fully allocated net cost of peaking generators built to serve those peak hours' loads. Even if prices are capped and the construction of peaking generators is encouraged by incentives such as capacity payment, the system costs avoided by load reductions should tend toward the net cost of a new generator. Approaches 1 and 2 estimate these net costs using contrasting methodologies.

Approach 1: Single utility, thermal generation

Approach 1 assumes that the power system is a single utility with an hourly distribution of demands similar to the Pacific Northwest. Further it assumes that the generating system is made up of thermal generators, with marginal peaking generators that are new single cycle combustion turbines or "duct firing" additions to new combined cycle combustion turbines. The assumed costs and other characteristics of these generators are taken from The NW Power Planning Council's standard assumptions for new generating resources.[9]

---

[8] Using the convention that the percentage changes in demand and price are $\ln(D_2/D_1)$ and $\ln(P_2/P_1)$, respectively, we can calculate the new demand $D_2 = \exp(-0.05*\ln(180/60) + \ln(30{,}000)) = 28{,}397$ megawatts. The reduction from the initial peak demand of 30,000 megawatts is 1,603megawatts.

[9] These assumptions are documented in the *Northwest Power Planning Council New Resource Characterization for the 5th Power Plan.* The duct firing and simple cycle combustion turbine generators cited in this paper are covered in sections on "Natural Gas Combined Cycle Gas Turbine Power Plants" and "Natural Gas Simple Cycle Gas Turbine Power Plants." These documents are available on request from the Council--contact the author.

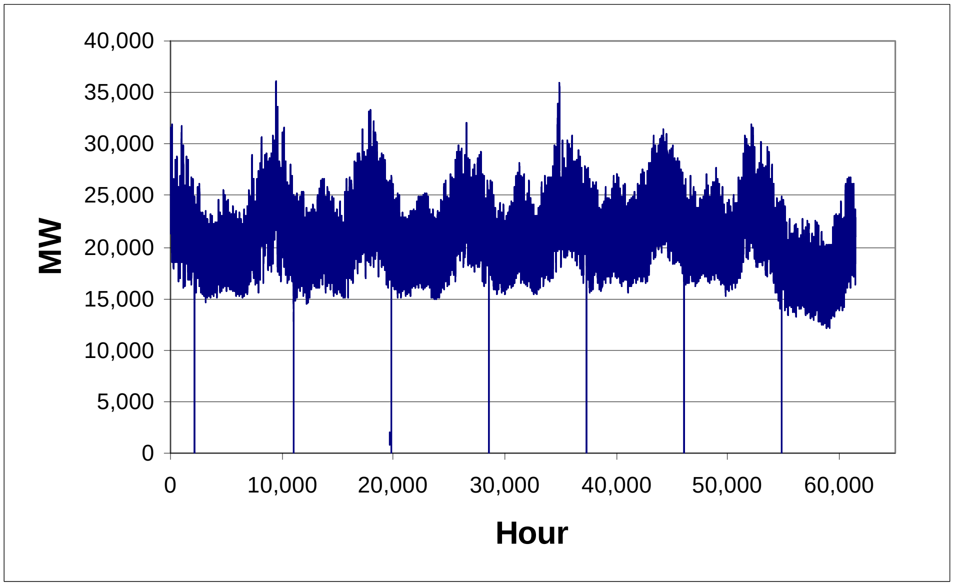

Figure H-1: Pacific Northwest Hourly Loads 1995-2001

In our assumed utility the cost of serving each increment of load depends on how many hours per year that load occurs. We must therefore examine the hourly distribution of loads. The Pacific Northwest hourly loads shown in Figure H-1 are loads from January 1, 1995 through December 31, 2001. The loads demonstrate that the Pacific Northwest is a winter-peaking system. The highest hourly load in the 7-year period shown is 36,118 megawatts in hour 8 of February 2, 1996 (hour 9536), and loads reach nearly 36,000 NW in several hours in December of 1998 (between hours 34,808 and 34,834). There is considerable year-to-year variation in peak loads; peak loads were below 32,000 megawatts in 1995, 1999 and 2000.

When we rearrange the same data, by ordering hourly loads from highest to lowest, we form a "load duration curve" shown in Figure H-2. Figure H-3 shows the first 700 hours in Figure 2, that is, the highest 700 hourly loads. These data let us focus on the amount of generating capacity that is used just a few hours each year to serve the highest loads.

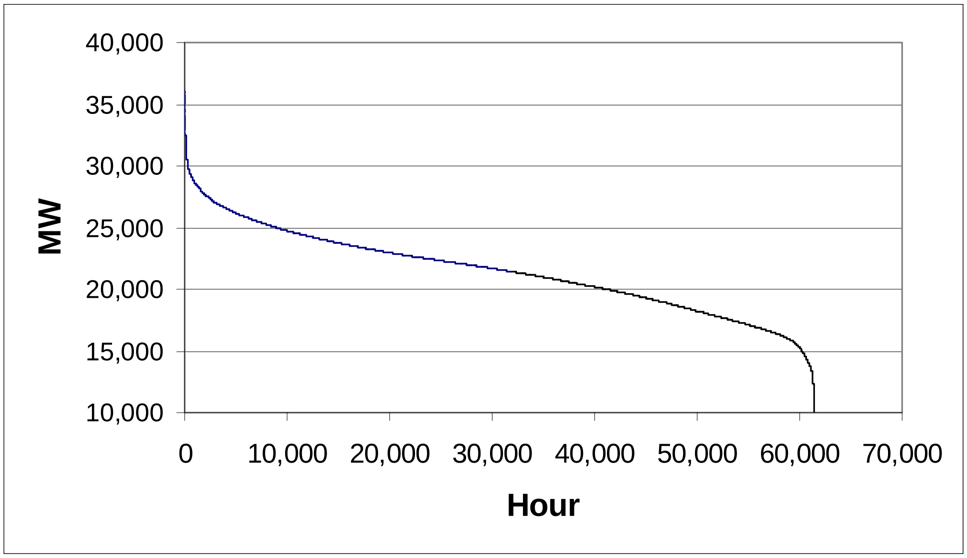

Figure H-2: Pacific Northwest Load Duration Curve 1995-2001

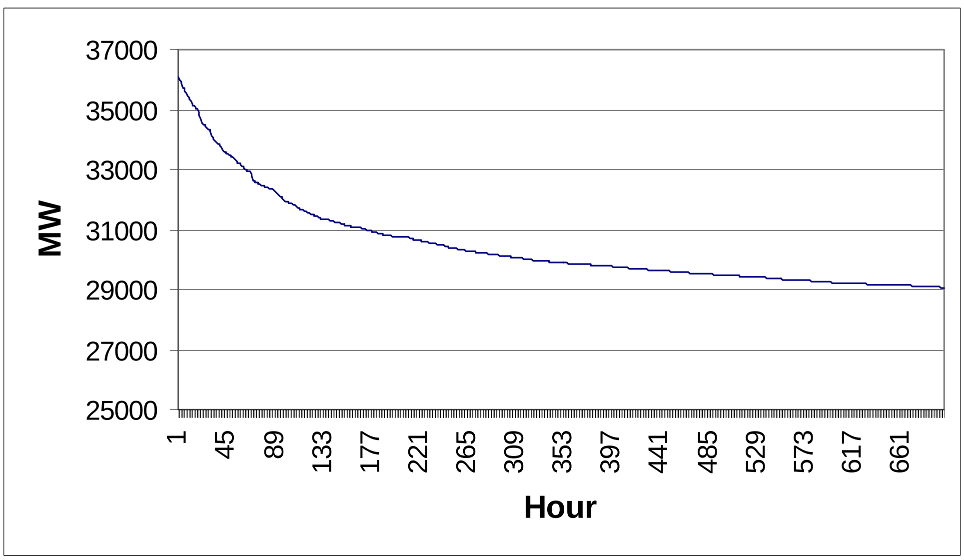

Figure H-3: Loads of Highest 700 hours 1995-2001

Referring to the data underlying Figure H-3, the highest load in the 7-year period is 36,118 megawatts. Of that peak load, 500 megawatts of load needs to be served only 7 hours (1 hour per year on average), 1,563 megawatts of load is served only 21 hours (3 hours per year on average), 3,500 megawatts is served 70 hours (10 hours per year on average), and so forth.

What does it cost to serve this load? Since incremental generators necessary to serve the load operate for different numbers of hours per year, each one has its own cost per megawatt-hour, declining as hours of operation per year increase. Let's look at two levels of use, 10 hours per year and 100 hours per year.

Based on the Council's generating cost data base, the cost of new[10] peaking generators used 10 hours per year is $6,489 per megawatt hour ($6.49 per kilowatt hour) for duct burner attachments on combined cycle combustion turbines, and $11,442 per megawatt hour ($11.44 per kilowatt hour) for simple cycle combustion turbines. The generators operating less than 10 hours will of course have even higher costs per megawatt-hour than these estimates.

The 700th highest hour's load in Figure 3 is 29,076 megawatts. This means that there are 3,542 megawatts of load that need to be served more than 10 hours but less than 101 hours per year. The same Council cost data cited above indicate that new peaking generators that are used 100 hours per year cost $677 per megawatt hour ($0.68 per kilowatt hour) for duct firing and $1,179 ($1.18 per kilowatt hour) for simple cycle combustion turbines. That means that serving peak loads between 29,076 megawatts and 32,618 megawatts by building and operating new peaking generators costs between $0.68 per kilowatt hour and $11.44 per kilowatt hour, depending on which type of generator is used and whether its hours of use are closer to 10 hours per year or 100 hours per year. All of these costs are much higher than retail electricity prices, which run in the $0.05-0.10 per kWh range in our region.

To summarize, the assumption of a single utility, Pacific Northwest hourly loads and new thermal resources leads to the conclusions:

1. The highest 70 hourly loads in the 1995-2001 period require about 3,500 megawatts of peaking generation to serve. Load reductions that made it unnecessary to serve these loads would save at least $6.49 per kilowatt-hour.
2. The next highest 630 hourly loads in the 1995-2001 period require about 3,542 megawatts of peaking generation to serve. Load reductions that made it unnecessary to serve these loads would save between $0.68 and $6.49 per kilowatt-hour.

Limitations of this analysis

This analysis used simplifying assumptions that let us focus on the concepts involved, but excluded some features of the real world, possibly influencing the results. What assumptions deserve consideration for a more refined analysis?

Hydroelectric resources

The initial analysis assumed that the generating system was made up entirely of thermal resources. In fact, hydroelectric generators provide more than half of the electrical energy of the Pacific Northwest power system. Hydroelectric resources look like baseload generators in some respects--their cost structure is high capital cost/low variable cost, like nuclear plants.

But in other respects, hydro resources lend themselves to use as peaking resources. Their output can vary quickly to follow loads' short-term variation. Our hydro system was built with a lot of generating capacity to take advantage of years when more-than-normal precipitation makes more energy production possible. By using their reservoirs, hydro resources can even store energy generated by baseload thermal units and release it to meet peak loads, within limits.

Finally, the total energy available from the hydro system varies, depending on variation in seasonal and annual precipitation. In our power system a thermal peaking generator may operate

[10] Operating an existing peaking plant, once the fixed costs are incurred, is much cheaper. The greatest savings offered by demand response is as an alternative to building a new generating plant, avoiding the generator's fixed cost.

more like a baseload plant in bad water years, because of a shortage in energy from the hydro system.

These considerations make it desirable to reflect hydro resources' effects in our analysis.

Trade between systems with diverse seasonal loads

The initial analysis assumed that generation served a single utility with an hourly distribution of loads like the Pacific Northwest. Actually, our transmission system links us to other systems (most notably California) that have different load distributions. In the real world peaking generators may very well run to meet winter peak loads in our region, and also to help meet summer peak loads in California. This would tend to increase the use of each peaking generator, spreading its fixed cost over more hours and reducing the average cost of meeting peak loads.

Operational savings of new units

The marginal effect of a new peaking generator added to an existing system to meet peak loads is more complex than we assumed in the initial analysis. The new unit, if it is more efficient than older units, will be operated ahead of them. The result could be that the new unit is operated not just to cover growth in peak loads, but also to reduce operating costs by replacing older units' production. In this case the net cost of meeting incremental peak load is not the fixed and operating costs of the new unit, as we assumed in the initial analysis, but rather the fixed cost of the new unit minus the net operational savings that it makes possible for the system as a whole.

Approach 2: AURORA® simulation of Western power system

The Council uses a proprietary computer model, AURORA®,[11] to project electricity prices and to simulate other effects of changes in the development and operation of the power system. AURORA® simulates the development and operation of the power system of the Western United States and Canada. It takes account of interaction between hydro and thermal generators, trade among the various regions, and the operational interaction among plants of different generating efficiencies; that is, it allows a more realistic set of assumptions than we adopted in Approach 1. We used AURORA® to refine our initial estimate of the net cost of serving incremental peak load.

Our analytical approach was to begin with the Council's baseline projection, noting the amount of electricity service that is projected by AURORA® and the generating costs of the power system. Then we varied the amount of generating capacity, and simulated the operation of the power system again, noting the changes in electricity service and generating costs. We focused on the year 2010 because we appear to have a surplus of generating capacity at the present, and by 2010 AURORA® has arrived at something like equilibrium between supply and demand.

In order to vary the amount of generating capacity, we varied the operating reserve requirements simulated by AURORA® across three levels--6.5 percent, 15 percent and 25 percent. We performed the experiment twice with the same three generating portfolios: once assuming energy output from the Pacific Northwest hydro system based on average precipitation, and again with Pacific Northwest hydro energy based on "critical" precipitation.[12]

---

[11] The AURORA® Energy Market Model is licensed from EPIS, Inc.

[12] "Critical" water is used in the Pacific Northwest as the basis of the energy that can be counted as "firm" from the hydro system. Critical water is based a series of bad water years in the 1930s.

The result was three levels of costs and levels of service for average water and three levels of costs and levels of service for critical water, shown in Table H-2.

Table H-2: West-wide Change in Costs and Service from AURORA® Simulations - 2010

Case	Change in System Costs ($thousands)	Change in Electricity Service - megawatt hour	Cost of Change in Service $ per megawatt hour ($ per kilowatt hour)
6.5% -15% Reserve (Average Water)	1,190,262	1,157,188	1029 (1.03)
15% - 25% Reserve (Average Water)	2,467,836	168,793	14,621 (14.62)
6.5% - 15% Reserve (Critical Water)	1,113,170	2,144,813	519 (0.52)
15% - 25% Reserve (Critical Water)	2,420,030	580,653	4,168 (4.17)

Given that Approach 2 is much different in structure and assumptions than Approach 1, it's not surprising that the estimated costs of incremental service are different. However, both approaches show that at high levels of service the cost of serving incremental load can be well over $1,000 per megawatt hour ($1.00 per kilowatt hour). Put another way, both approaches suggest that the power system could save well over $1.00 per kilowatt-hour if it could avoid serving the highest peak loads. In both approaches the cost of serving incremental load rises as we serve the last few hours of the highest peak loads (the highest 10 hours in Approach 1, the highest operational reserves in Approach 2).

Approach 2 lets us examine the effects of variation in output from the hydroelectric system on the results. Other factors equal, overall system costs are higher when we assume critical water than when we assume average water. However, with critical water, less energy is available from the Pacific Northwest hydroelectric system and generators run more hours, spreading their fixed cost and reducing the cost of incremental service per megawatt-hour. Table H-2 doesn't show this, but the absolute levels of service are lower with critical water. The general pattern noted above, of incremental costs rising at higher operational reserves, persists with critical water.

The Council's AURORA® analysis treats the power system of the western U.S. and Canada as made up of 16 regions, with four of these regions corresponding to the Pacific Northwest. Table H-2 shows the total results of all 16 regions, but we also examined the results for the Pacific Northwest, shown in Table H-3.

Table H-3: Pacific Northwest Change in Cost and Service from AURORA Simulations - 2010

Case	Change in System Costs ($thousands)	Change in Electricity Service MWh	Cost of Change in Service $ per megawatt hour ($ per kilowatt hour)
6.5% -15% Reserve (Average Water)	-2,112	328,705	-6 (-0.01)
15% - 25% Reserve (Average Water)	7,346	50,386	146 (0.15)
6.5% - 15% Reserve (Critical Water)	29,756	596,896	50 (0.05)
15% - 25% Reserve (Critical Water)	131,323	112,299	1,169 (1.17)

These results are markedly different than the results for the whole West. The costs of incremental service shown in the last column are much lower than in Table H-2, and even include a negative cost. This seemed unreasonable at first, but after more examination of the detailed results it became clear that the Pacific Northwest added relatively less generating capacity in response to the increased reserve requirements than did the West as a whole.

This is because the heavily hydroelectric power system of the Pacific Northwest already had relatively high reserves. Our hydro system was built with such reserves to cover the variation in river flows as well as concern about serving peak load. The result is that the Pacific Northwest had to invest relatively little fixed cost to meet the 15 percent and 25 percent operational reserve. At the same time, the extra generating reserves throughout the West drove market prices of wholesale electricity down. The Pacific Northwest could reduce operational costs by taking advantage of increased opportunities to buy energy from neighboring regions. These operational cost savings partially offset (and in the "6.5% -15% Reserve (Average Water)" case, more than offset) the increased fixed costs due to new generator investments in the Pacific Northwest.

This example illustrates a more general issue, which is: any region (or utility) will benefit if it can depend on its neighbors' reserves while avoiding some of the fixed costs of those reserves. The temptation for each party to lean on others' reserves will tend to discourage everyone from making such investments, and tend to leave the whole system with less-than-optimal reserves.

What's the implication of this issue for demand response? Avoidance of fixed costs is the main incentive for leaning on neighbors' reserves. To the extent we can identify lower-fixed-cost alternatives to provide reserves, we reduce this incentive. To the extent that demand response comes to be seen as a proven alternative to building peaking generators, the very low fixed cost of demand response would make it less risky for each party to cover its own reserve needs, and more likely that total system reserves are adequate.

Approach 3: Portfolio Analysis of Risk and Expected Cost

Approaches 1 and 2 estimated the avoided cost of serving known loads with known resources. In fact, loads are uncertain because we don't know future weather and economic growth, and the capability of our generating resources is uncertain because of unplanned outages, variation in rain and snowfall, among other factors. In addition, the region's utilities buy and sell into an electricity market that includes the western U.S. and Canada, making market prices a further source of uncertainty. For these and other reasons, the Council adopted a long-term portfolio analysis in formulating the Fifth Power Plan. Approach 3 used the Council's portfolio analysis model to make a third estimate of the value of demand response to the system.

The Council's portfolio methodology is described in Chapters 6 and 7 of the Plan, and in more detail in Appendix L. To evaluate the effect of demand response on risk and expected cost, the Council's portfolio model was run with and without demand response, and the resulting shift in the efficient frontier of portfolios was analyzed. This analysis was described briefly in Chapter 7.

For the "with" demand response portfolio analysis, Council staff assumed a block of 2,000 megawatts of load reduction is available by 2020, with an initial fixed cost of $5,000 per megawatt, a maintenance cost of $1,000 per megawatt per year and a variable cost of $150 per megawatt-hour when the load reduction is actually called upon.[13] The "without" demand response assumed that no demand response is available.

The portfolio model simulated 750 20-year futures with demand response available 16 years in each future. Demand response was used in 83 percent of years in which it is available, but the amount of demand response used is usually quite small. In 85 percent of the years in which demand response is used, it is used less than 0.1 percent of its capability (i.e. less than 9 hours per year). According to the portfolio model's simulations, demand response is used more than 10 percent of its capability (equivalent to about 870 hours per year) in about 5 percent of all years.

The effect of removing demand response on the efficient frontier is demonstrated in Figure H-4. The efficient frontier is shifted from the "Base Case" up and to the right to "No Demand Response," reflecting increases in both expected cost and risk. The amount of the shift varies along the frontier, but in general the loss of demand response increases expected cost by more than $300 to more than $500 million for constant levels of risk. Expressed another way, the loss of demand response increases risk in the range of $350 to $650 million at given levels of expected cost. These increases in expected cost and risk are largely due to increased purchases from the market at times of high prices and to the cost of building and operating more gas-fired generation.

[13] This assumption is simpler than reality, since the variety of load reduction opportunities mean that there is really a supply curve for demand response, with more response available at higher costs.

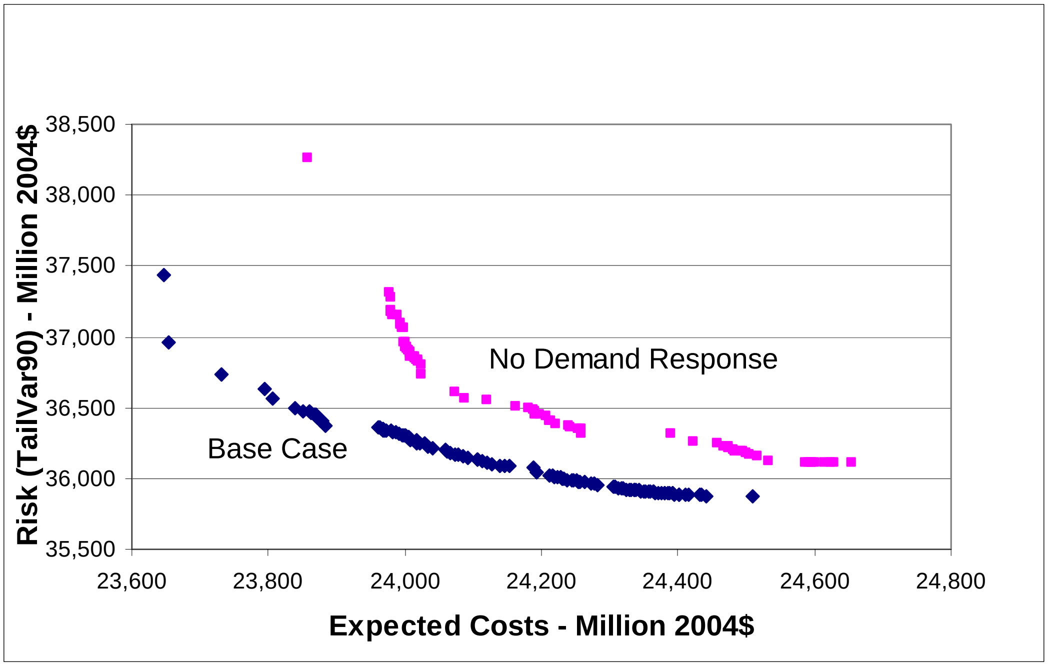

Figure H -4: Effect of Demand Response on Efficient Frontier

Summary of Analysis on Value of Load Reduction

Each of the approaches to estimating the value of load reduction has its own strengths and limitations, but the general conclusions are quite robust: Demand response offers very significant potential value to the region. As laid out in Chapter 4 and in the Action Plan, there are a number of areas that need further experience and analysis in order for the region to realize that potential value, but the analysis presented here is evidence that the effort to acquire that experience and perform that analysis is very worthwhile.

r:\dw\ww\fifth_plan\push to the final\prepub\appendix h (demand response) (pp).doc