Application of San Diego Gas & Electric Company (U-902-E) for Adoption of an Advanced Metering Infrastructure Deployment Scenario and Associated Cost Recovery and Rate Design.

Application 05-03-015
Exhibit No.: _____________

CHAPTER 7
CAPACITY & ENERGY VALUE OF AMI-ENABLED DEMAND RESPONSE

JULY 14, 2006 AMENDMENT

Prepared Supplemental, Consolidating, Superseding and Replacement Testimony
of
JOHN C. MARTIN

SAN DIEGO GAS & ELECTRIC COMPANY

BEFORE THE PUBLIC UTILITIES COMMISSION OF THE STATE OF CALIFORNIA

July 14, 2006

TABLE OF CONTENTS

CHAPTER 7

CAPACITY & ENERGY VALUE OF AMI-ENABLED DEMAND RESPONSE

JULY 14, 2006 AMENDMENT

Prepared Updated Supplemental, Consolidating, Superseding and Replacement Testimony

of

JOHN C. MARTIN

SAN DIEGO GAS & ELECTRIC COMPANY

I. INTRODUCTION

A. Scope and Purpose

The purpose of this *amended* testimony is to refresh my March 28, 2006 testimony to include material information which will impact my (Chapter 7) testimony in which I show that \$85/kW-Year is an appropriate levelized fixed cost for generation capacity avoided by advanced metering infrastructure (AMI) enabled demand response (DR) for San Diego Gas & Electric Company (SDG&E). Second, my testimony describes the avoided energy prices used to value the energy savings associated with the AMI enabled demand response. My testimony also discusses general approaches used to value DR, as well as other benefits which have not been quantified elsewhere in SDG&E's business case. This July 14th amendment has three major changes. I incorporate the new demand response MWs from Dr. George's testimony (Chapter 6) which is reflected in my Table JCM 7-2 and in sections II.D.1 Reduced Demand Volatility and Planning Reserves. Section II.D.1 also contains a correction to my \$/kW-year calculation. Section II.D.3 Additional Reliability Value is updated to reflect SDG&E's new Programmable Controllable Thermostats (PCT) retrofit program. Minor changes include corrected reference dates. This testimony consolidates, supersedes, and replaces all previous direct and supplemental testimony filed by me or by any other SDG&E witness testifying in this docket, on the topics covered herein.

B. Summary of Testimony

SDG&E uses the $85/kW-year as the levelized value for avoided capacity of AMI enabled demand response. This avoided capacity value was independently derived but is also consistent with Administrative Law Judge and Assigned Commissioner direction regarding an AMI business case analysis framework[1]. SDG&E uses this value because it is representative of the levelized fixed cost for new gas combustion turbines (CT) that can be displaced by AMI enabled demand response. A CT does have a market energy benefit beyond the AMI enabled demand response envisioned in this application, but, on the other hand, demand response has additional volatility, efficiency, reliability and other benefits that generation capacity can not provide. Thus, on balance, SDG&E believes the value of these additional demand response benefits outweigh the market energy benefits of a CT.

C. Background on Demand Response Benefits

The challenge of valuing AMI enabled demand response has been apparent since the California Public Utilities Commission (CPUC) began its proceeding on policies and practices for advanced metering, demand response, and dynamic pricing (Demand Response Proceeding), R.02-06-001. The specific quantification of the value of demand response was not, however, resolved in that proceeding. In D.05-11-009, the Commission staff was ordered to prepare draft protocols for estimating impacts for demand response programs, and to prepare a proposed rulemaking. The avoided cost Rulemaking (R.04-04-025) is also grappling with valuing demand response. Unfortunately, these protocols and valuations are not currently available for guidance in this instant application, nor are they likely to be available in the near future.

Fortunately, valuable guidance is provided by a recently published U.S. Department of Energy (DOE) report to Congress, entitled <u>Benefits of Demand Response in Electricity Markets and Recommendations for Achieving Them</u>

[1] Appendix B – Derivation of Capacity and energy values for on and off peak periods, ALJ and ACR ruling dated July 21st 2004, R.02-06-001.

(DOE report).[2] This report frames many of the pertinent demand response valuation issues in its Appendix B: Economic and Reliability Benefits of Demand Response. This Appendix B is attached to my testimony as Attachment JCM-7.1.

The DOE Report Appendix B categorizes demand response benefits as short-term and long term market impacts. The report also discusses reliability benefits associated with emergency demand response. The following section briefly summarizes the impacts detailed in the report's Appendix B as well as additional benefits not addressed in the DOE Report.

1. Short Term Benefits – Reduced Supply Costs and Market Prices

Supply costs are directly reduced when consumers reduce energy consumption in response to critical peak prices or rebates. The utility or load serving entity (LSE) avoids having to generate or purchase the reduced energy. Customers bills are lowered by the energy reduction (kWhs) times the critical peak price or rebate.

In a wholesale market structure, market prices are reduced during critical peak periods. Demand response causes a shift and/or a slope change in the demand curve, thus moving the market clearing price down the supply curve. This market clearing price shift can be considerable at high loads when the supply curve tends to be very steep. Lower market clearing prices reduce the cost of all energy purchased in the market. The magnitude of this reduction is dependent on many factors including: how supply and demand are integrated into the market bidding mechanisms, how market clearing prices are determined, and the quantity of energy traded in the market.

This reduction in market prices can result in a "rent" transfer from suppliers to consumers. This short term transfer is a benefit to customers, but is not an efficiency improvement and does not cause a social welfare gain.

[2] U.S. Department of Energy, Benefits of Demand Response in Electricity Markets and Recommendations for Achieving Them, A report to the United States Congress Pursuant to Section 1252 of the Energy Policy Act of 2005, February 2006.

The argument states that in the long run, such market distortions can not be maintained and can result in higher prices.[3]

2. Long Term Benefits – Reduced Capacity Requirements

Generation capacity can be avoided or deferred when AMI enabled demand response reduces consumption during peaks hours, thus reducing peak system demand. This avoided generation capacity is typically valued as the marginal cost of a CT. The DOE report states that:

> "By convention, marginal capacity is assumed to be a 'peaking unit', a generator specifically added to run in relatively few hours per year to meet peak system demand. Currently, peaking units are typically natural gas turbines with annualized capital costs on the order of $75/kilowatt-year (kW-year) (Orans et al. 2004, Stoft 2004)." (Appendix B, p. 74)

Note that the $75/kW-year cited above may be an appropriate 2004 average national valuation of avoided capacity, but, later in my testimony, I demonstrate that $85/kW-year is the appropriate levelized, 2006 fixed capital cost for a gas turbine in California. The benefits associated with reduced system peak for Transmission and Distribution are detailed in Mr. Lee's testimony (Chapter 4).

3. Reliability Benefits

Load reductions associated with demand response can provide incremental reliability benefits to the electrical system. The DOE report defines emergency demand response programs as those programs that provide incremental reliability benefits at times of unexpected shortfalls in generation reserves, beyond demand response programs that provide capacity benefits (such as the CPP rate and the Peak Time Rebate proposed by SDG&E in this application). The DOE report states that this reliability benefit should be valued as follows:

> "Economists define the concept of value of lost load (VOLL) as the proper measure of improved reliability, since it reflects customer's marginal value for electricity under these circumstances. The product of VOLL and the

[3] For a further discussion on this subject see the textbox "Distinguishing Societal Benefits from Rent Transfers" in the DOE Report Appendix B, at p. 71.

> expected un-served energy (EUE), the load that otherwise would not have been served, monetizes the value of load curtailments." (Appendix B, p. 82)

Currently, SDG&E's AMI business case does not include a specific emergency demand response program, but SDG&E's proposed AMI system can, however, support such programs in the future.

4. Societal Benefits – Pricing Efficiency Gains

According to economic theory, demand response provides societal benefits by improving pricing efficiency. Electricity, like other resources, is used most efficiently when consumer prices reflect marginal cost of supply. Typical retail electric pricing (e.g., flat rates, inverter tiers or time-of-use) can cause a mismatch between the prices customers pay for electricity versus the cost to supply the electricity. This mismatch causes greater than socially optimal usage during peak periods when supply costs are greater than retail prices. Demand response reduces this mismatch between retail electricity prices and marginal supply costs, thus increasing pricing efficiency, which benefits society as a whole.

5. Other Benefit Considerations

Several other considerations are important when evaluating the capacity and energy benefits of AMI enabled demand response. Most of these considerations are not specifically included in the DOE report Appendix B. They include, but are not limited to, CT net energy, planning reserves, line losses, emissions, demand volatility, and retail rate design flexibility.

The gas CT market energy benefit must be considered when evaluating the value of demand response. Since the cost of a natural gas turbine is the appropriate marginal capacity value for demand response, the analysis must consider the benefits a CT can provide that can not be provided by SDG&E's proposed demand response rates and programs. SDG&E's CPP and Peak Time Rebate program are expected to operate 91 hours per year (i.e., the seven hour critical peak period times the thirteen average critical peak events per year). However, the proposed rebate program is not limited to 91 hours per year, nor limited to the summer season only. Furthermore SDG&E's AMI system can support many demand response rates and program designs including real-time or hourly pricing, which have the potential to provide significant additional demand response benefits. A market based CT could be operated almost any hour of the year, but is likely only to run when profitable. This value is discussed in detailed later in my testimony.

Planning reserves can be reduced when system peak (maximum) demand is lowered by demand response. Typically, planning reserves are generation reserves above and beyond the anticipated maximum demand or historic system demand. These reserves are intended to provide "standby" generation to cover unexpected losses of generator or transmission resources, or unexpected demand increases. SDG&E's CPP rate and Peak Time Rebate are equivalent to generation plus planning reserves.

The CPUC acknowledges this fact by requiring LSEs to procure sufficient resources to meet a 15-17 percent planning reserve margin (D.04-01-050). In addition, the CPUC recognizes that dispatchable demand response, over which the LSE has dispatch control, should be counted as a resource (D.04-10-035, Finding of Fact 9). Because it is not debited from load forecasts, reserve requirements should not be imposed for demand response counted as resources (D.04-10-035, Conclusion of Law 18). Therefore, one MW of demand response resource is equivalent to 1.15 MWs of generation resource for valuation purposes.

Distribution line losses are added to demand response MW estimates to produce equivalent MW as a generation resource for valuation purposes. Transmission line losses are not included in SDG&E's valuation since most estimates of generation capacity value reduce the gross plant output by transmission line losses. The 2004 Market Price Referent uses an implied 1.43% transmission loss factor, based on a Generator Meter Multiplier (GMM) of 98.57% (D.05-12-042, p. 45). SDG&E's distribution line losses are summarized in Dr. George's testimony in Chapter 6 (see Table SSG 6-19).

Reduced emissions may result when older, less efficient generation units used to supply peak generation are avoided. Less efficient generation requires more fuel to produce a unit of output; therefore, they generally have the highest variable cost and are the last units dispatched to balance system supply and demand. Furthermore, the more fuel used per unit of output generally results in more pollution per unit of output (NO_x, SO_x, and CO_2). This benefit must be reduced by any consumption increases during lower price, off-peak periods because more efficient base-load and intermediate-load generators would operate more hours.[4]

Reduced demand volatility is a benefit of demand response during times of high demand such as system peaks. Lower demand volatility makes it easier for generation resources to balance the electrical system. Therefore, over time there is the potential to reduce the Planning Reserve margin from the current 15-17 percent discussed above. This is another potential form of avoided capacity provided by AMI enabled demand response.

Additional rate design flexibility is an important benefit of AMI. SDG&E's CPP and Peak-Time rebate are only two possible demand response rate designs possible. AMI opens the possibility for designs such as Real-Time pricing, reliability rates and programs, as well as other designs that provide opportunities to reduce generation costs and improve pricing efficiency.

II. SDG&E'S VALUATION OF SUPPLY BENEFITS OF DR

This section provides the estimated values, sources, and justification used by SDG&E to value supply benefits. The majority of the benefits are capacity and energy related. My testimony also addresses other benefits attributable to AMI enabled demand response.

[4] Holland, Stephen, Mansur, Erin T. The Distributional and Environmental Effects of Time-Varying Prices in Competitive Electricity Markets, CSEM WP-143, May 2005. This paper describes two benefits of Real Time Pricing (RTP) -- reduced volatility and reduced prices. These benefits, however, can increase emissions. SDG&E's CPP rates and Peak-Time rebate do not have as strong an incentive as RTP to increase consumption during low supply cost hours. Therefore, the decrease in volatility and price are likely to be less than RTP, which may result in different net emission impact.

A. The Value of Avoided Generation Capacity

SDG&E values avoided generation capacity at $85/kW-Year, which represents the levelized fixed costs of a CT generator. I utilize this value for the entire analysis period (2006 through 2038). This fixed valuation provides the same benefits as a 2006 capacity value of $60/kW-year growing at a 2.5% inflation rate over the analysis period.

My $85/kW-year value is based on the 2004 Market Price Referent (MPR)[5] capital and fixed O&M costs and the E3 Financial Model for Capacity Costs as adopted in D.05-04-024,[6] modified for a twenty-five year life. The 2004 MPR provides values for baseload and peaking proxy plants for the 2004 Renewables Portfolio Standard (RPS) solicitation. The 2004 MPR represents the presumptive cost of electricity from a natural gas-fired baseload or peaker plant (D.04-06-015). The E3 Financial Model calculates fixed costs for a generator and is part of its methodology to evaluate energy efficiency programs. At this time, E3's overall methodology used to evaluate energy efficiency programs is in the process of being updated for consideration as a valuation approach for evaluating Demand Response. Nevertheless, even without the benefit of an updated review, I believe the E3 Financial Model for Capacity Costs is entirely appropriate for calculating fixed CT costs for the purposes of evaluating this application.

The inputs from the 2004 MPR include CT construction costs and fixed O&M costs. The construction cost is $556 per kW of capacity, in 2005 dollars. The fixed O&M is $12.10/kW-year, in 2005 dollars. These values are escalated by 2.5% for 2006 dollar values in the E3 Financial Model.

The E3 modeling assumptions are those approved in D.05-04-024, with an adjustment for the life of a CT. The E3 model assumes a 20 year life for a CT, however, I use a 25 year life which more appropriately reflects the useful life of a CT. A longer life results in a lower capacity cost. The modeling input assumptions are summarized in Table JCM 7-1. The model produces a levelized

[5] Resolution E-3942, July 21st 2005.
[6] Interim Option on E3 Avoided Cost Methodology, CPUC D.05-04-024, April 7, 2005.

fixed cost of a CT at $85.84/kW-Year. For demand response valuation purposes $85/kW-Year is used.

Table JCM 7-1

E3 Financial Model for Capacity Costs Inputs for SDG&E's AMI Case	
Operating Data	
Lifetime (yrs)	25
Plant Costs	
In-Service Cost ($/kW)	$556.00
Fixed O&M ($/kW-yr.)	$12.10
Property Tax (%)	1.20%
Property Tax ($/kW-yr.)	$6.67
Insurance (%)	0.60%
Insurance ($/kW-yr)	$3.34
Financing	
Debt-to-Equity	70.00%
Debt Cost	6.50%
Equity Cost	12.00%
Marginal Tax Rate	40.75%
Other Inputs	
Cost Basis Year	2005
Resource Balance Year	2006
Generation Capital Inflation	2.50%
Fixed O&M Escalation	2.50%
Financing Period	25
Book Life	25

My evaluation is within the range of a number of other reasonable estimates of a levelized fixed cost for a CT. For example, the Administrative Law Judge and Assigned Commissioner direction regarding an AMI business case analysis framework has the same value[7]. That $85 value is based on the California Energy Commission's (CEC) Comparative Cost Study[8]. The CEC report provides a basic understanding of certain fundamental attributes that are generally considered when evaluating the cost of building and operating different electricity generation technology resources.

A slightly higher levelized fixed capacity is derived from the 2004 MPR. The purpose of the 2004 MPR was to provide values for baseload and peaking proxy plants for the use in the 2004 Renewables Portfolio Standard (RPS) solicitation.

[7] Appendix B – Derivation of Capacity and energy values for on and off peak periods, ALJ and ACR (R.02-06-001), July 21st 2004.
[8] Comparative Cost of California Central Station Electricity Generation Technologies, CEC Staff Report (100-03-001), August 2003.

These costs were estimated with a financial cash flow model prepared by Southern California Edison (SCE) and reviewed by the Energy Division and parties to that proceeding. The financial model includes a calculation for fixed capacity component, including just under $7 Million to cover CT interconnection costs, project contingency, environmental review and mitigation, as well as permitting costs[9] This model produces a levelized fixed revenue requirement of $87.19/kW-year in 2005 dollars, which includes all fixed costs to maintain a CT over its assumed 20 year life (in my analysis, however, I use a twenty-five year life to better reflect expected a CT's operational life). All these capital costs can be avoided by AMI enabled demand response.

This capacity cost estimation work continues at the CPUC with the recently adopted 2005 MPR[10]. The guidelines from this decision will likely result in avoided CT capacity values higher then $85/kW-year. The 2005 MPR uses a higher cost of capital (D.05-12-042, page 40 - 41), and updates California construction costs (D.05-12-042, page 28).

SDG&E's nominal levelized $85/kW-year capacity value is used to estimate the net present value for the AMI capacity benefits. These results are presented in Dr. George's testimony in Chapter 6, with a total NPV of $243.7 million. Table JCM 7-2 provides a sensitivity analysis of hypothetical real economic carrying charges and inflation rates for escalation. The results show that for SDG&E, the NPV (discounted at SDG&E's weighted average cost of capital of 8.23%) of capacity benefits using my $85/kW-year with no inflation is equivalent to a real economic carrying charge of $60 in 2006, escalated at a 2.5% annual inflation rate. In other words, whether $85 flat or $60 escalated for inflation is used, the avoided capacity value is essentially the same.

---

[9] Revised 2004 Market Price Referent (MPR) Staff Report, CPUC Energy Division, February 10th 2005, page 8 of 14.
[10] Interim Opinion Adopting Methodology for 2005 Market Price Referent, CPUC (D.05-12-042), December 15, 2005.

Table JCM 7-2

Sensitivity of AMI Generation Capacity Value to Avoided Cost and Inflation (NPV 2006 $Millions)							
		Avoided Cost ($/kW-Year)					
		60	65	70	75	80	85
	0.0%	$ 172	$ 186	$ 200	$ 215	$ 229	$ 243
Annual	0.5%	$ 184	$ 200	$ 215	$ 230	$ 246	$ 261
Inflation	1.0%	$ 198	$ 214	$ 231	$ 247	$ 264	$ 280
Rate	1.5%	$ 213	$ 230	$ 248	$ 266	$ 284	$ 301
	2.0%	$ 229	$ 248	$ 267	$ 286	$ 305	$ 324
	2.5%	$ 247	$ 267	$ 288	$ 308	$ 329	$ 349

An avoided capacity value of $85/kW-year represents a reasonable levelized valuation of a CT's avoided fixed capacity costs that can be displaced by AMI enabled demand response. This valuation is supported by the CEC's Comparative Cost Study, Commission direction regarding an AMI business case analysis framework, and the Commission's MPR decisions.

B. The Value of Avoided Energy

SDG&E prepared marginal summer energy costs by rate period for both Residential and C&I customers, respectively. These marginal costs are based on the SP-15 component of a multi-area run for the entire Western Electric Coordinating Council (WECC), using Global Energy's [formerly Henwood] base case assumptions for load and planning resource additions over a 20-year period (2005-2025). Originally, the costs were prepared for SDG&E's 2004 Long Term Resource Plan filing. Hourly prices were derived from the forecast's on-peak and off-peak prices using SDG&E's historical load. The hourly prices were averaged into time-of-use periods for CPP and Non-CPP days, based on the separate Residential and Commercial TOU periods and seasons. Values for 2005 and 2022 were used to interpolate interim years and estimate future year values. The Commercial & Industrial and Residential avoided energy prices are presented in Table JCM 7-3

Table JCM 7-3

AMI Avoided Energy Prices ($/MWhr)												
	Commercial & Industrial							Residential				
	CPP Days			Non-CPP Days				CPP Days		Non-CPP Days		
Year	**Peak**	**Semi**	**Off**	**Peak**	**Semi**	**Off**	**W-end**	**Peak**	**Off**	**Peak**	**Off**	**W-end**
2006	80	53	36	54	45	31	40	80	45	53	39	40
2007	83	56	37	57	47	32	42	83	47	55	40	42
2008	86	58	38	59	49	33	44	86	48	57	42	43
2009	89	60	40	62	51	34	46	89	50	60	43	45
2010	93	62	41	64	53	36	48	93	52	62	45	47
2011	97	64	42	67	55	37	50	97	54	65	47	50
2012	101	67	44	70	57	39	52	101	56	68	49	52
2013	105	69	46	73	60	40	55	105	58	71	51	54
2014	109	72	47	76	62	42	57	109	60	74	53	56
2015	113	74	49	80	65	44	60	113	62	77	55	59
2016	118	77	51	83	67	45	63	118	65	80	57	62
2017	122	80	52	87	70	47	65	122	67	84	59	64
2018	127	83	54	90	73	49	68	127	69	87	61	67
2019	132	86	56	94	76	51	72	132	72	91	64	70
2020	137	89	58	99	79	53	75	137	75	95	66	74
2021	143	92	60	103	82	55	78	143	77	99	69	77
2022	148	96	62	107	85	57	82	148	80	103	72	80
2023	154	99	65	112	89	60	86	154	83	108	74	84
2024	160	103	67	117	93	62	90	160	86	112	77	88
2025	167	107	69	122	96	64	94	167	89	117	80	92
2026	173	111	72	127	100	67	98	173	93	122	83	96
2027	180	115	74	133	104	70	103	180	96	128	87	100
2028	187	119	77	139	109	72	108	187	99	133	90	105
2029	195	124	80	145	113	75	113	195	103	139	94	109
2030	203	128	83	151	118	78	118	203	107	145	97	114
2031	211	133	86	158	122	81	123	211	111	151	101	119
2032	219	138	89	164	127	85	129	219	115	157	105	125
2033	228	143	92	172	132	88	135	228	119	164	109	130
2034	237	148	95	179	138	92	141	237	123	171	114	136
2035	246	154	98	187	143	95	148	246	128	179	118	142
2036	256	160	102	195	149	99	155	256	132	186	123	149
2037	266	166	106	204	155	103	162	266	137	194	128	155
2038	277	172	109	212	162	107	169	277	142	203	133	162

C. Gas CT Market Energy Benefit

I estimate a CT market energy benefit at $22.89/kW-year. By benefit, I mean the price received by the generator less variable costs. The methods used to estimate this value is detailed later in this section. This benefit is a valuation deduction representing the energy benefit that a CT might provide, and should be deduced from the capacity and energy benefit of AMI enabled DR, for a fair comparison.

My gross market energy benefit of a CT is dependent on many assumptions including CT efficiency and market prices for both electric energy and delivered natural gas. This value has been estimated in several other proceedings before the CPUC from a low of $8.76/kW-year in TURN's testimony on Southern California Edison Co.'s Marginal Cost[11], to a high of $57.33/kW-year in TURN's analysis of PG&E's AMI case[12].

My calculation of a CT's gross market energy benefit is based on CT operational assumptions from the 2004 MPR. These operational assumptions include: A heat rate of 9,662 BTUs/kWh adjusted to 10,000 BTUs/kWh for ramping, start-ups, and temperatures above 59 degrees Fahrenheit; variable O&M of $9.68/MWhr in 2005; and an O&M escalation rate of 2.5% per year. The hourly energy prices are the same as those used to derive the avoided energy prices forecast described earlier in this testimony. The monthly gas prices are those associated with the energy price forecast, representing a Southern California border price. An intrastate transportation and municipal surcharge of $0.40/MMBTU is added to the Southern California border price to arrive at a delivered gas price. This model was run for each hour of the year in 2006 through 2025. The result is a total NPV of $286,703/MW, from this result a real 2006 value of $22.89/kW-year is estimated assuming 2.5% escalation per year.

D. Additional Value of AMI Enabled Demand Response

In this section I discuss several additional benefits that SDG&E's AMI system can achieve. These benefits are discussed at a conceptual level with estimates presented to illustrate potential values. These values are in addition to my earlier capacity and energy valuations. Three promising values of AMI enabled demand response discussed below are reduced demand volatility, rate design flexibility, and additional reliability. Furthermore, other potential option benefits may exist that I have not identified. On balance these benefits will likely provide sufficient value to offset the gas CT market energy benefit.

---

[11] Marcus, W., Florio, M. Electric Marginal Cost and Revenue Allocation for Southern California Edison Company, TURN, January 20, 2006 (p. 34)

[12] Nahigian, J, Shilberg, G, Marcus, W. Analysis of PG&E's Proposed Advanced Metering Infrastructure Application, TURN, January 18, 2006 (Table 8 p. 87)

1. Reduced Demand Volatility and Planning Reserves

Reduction of demand volatility is a demand response benefit documented by several studies including Holland & Mansur and Borenstein[13]. A long term benefit of reduced demand volatility is the possibility of reducing the level of planning reserves (currently 15% to 17% of system peak), since less generation will be needed to cover the reduced uncertainty in peak system demand and overall demand volatility. For instance, if demand response could reduce planning reserves by 1% (e.g., from 15% to 14%), then a significant quantity of generation capacity would not need to be constructed. For SDG&E, with a current system peak demand of about 4,000 MWs, this would represent significant long term generation capacity avoidance.

I estimate a 2006 value of avoiding 1% of planning reserves at $1.51/kW-Year, based on a 2011 peak reduction of 219 MWs (Table SSG 6-5) Dr. George's testimony, Chapter 6 and a $5.2 Million NPV total avoided planning reserve benefit between 2012 through 2038. This benefit is based on a bundled average peak customer demand in 2006 of 3,615 MW and assuming a 1% reduction in reserves starting in 2012. The avoided planning reserve reduction is estimated to be 0.82 MW in 2012 and grows with load growth. These avoided planning reserves are valued using the same 2004 MPR capacity values and 2.5% inflation I use earlier in my testimony.

2. Increased Rate Design Flexibility

A promising value of AMI enabled demand response is the rate design flexibility enabled by the system. SDG&E's proposed CPP rate and Peak-Time Rebate are just two of many possible rate designs to achieve demand response. As SDG&E becomes more familiar with these rates and the capabilities of the AMI system, subsequent rate design improvements will undoubtedly come to light. Indeed, the full panoply of rate design possibilities, including real time pricing (RTP), is enabled by AMI. RTP in

[13] Borenstein, Severin The Long-Run Efficiency of Real-Time Electricity Pricing, Center for the Study of Energy Markets (CSEM WP 133r) February 2005.

particular has economic efficiency benefit above and beyond CPP and Peak-Time rebate. I estimate this value to be an incremental $13.79/kW-Year.

This benefit is estimated for RTP by the Pacific Northwest National Laboratory on a national level[14], and by Borenstein for California. Borenstein's The Long-Run Efficiency of Real-Time Electricity Pricing (Efficiency Paper) referenced earlier is used as the basis of this valuation and is included as Attachment JCM-7.2. The method to estimate this value is discussed later in this section. This Efficiency Paper is best summarized by quoting from the Abstract:

> Retail real-time pricing (RTP) of electricity – Retail pricing that changes hourly to reflect the changing supply/demand balance – is very appealing to economists because it "sends the right price signals." Economic efficiency gains from RTP, however, are often confused with the short-term wealth transfers from producers to consumers that RTP can create. Abstracting from transfers, I focus on the long-run efficiency gains from adopting RTP in a competitive electricity market. Using simple simulations with realistic parameters, I demonstrate that the magnitude of efficiency gains from RTP is likely to be significant even if demand shows very little elasticity. I also show that "time-of-use" pricing, a simple peak and off-peak pricing system, is likely to capture a very small share of the efficiency gains that RTP offers."

The main analysis of the paper is an evaluation of RTP benefits relative to flat rate retail pricing. The absolute results are not comparable to energy and capacity values calculated earlier in this Chapter, because the Efficiency Paper uses different assumptions regarding annual generation capital cost and variable costs, among other assumption differences. However, for evaluating the relative economic efficiency benefit difference of RTP compared to CPP and Peak-Time rebate, the paper is useful.

The Efficiency Paper includes estimates for long term efficiency gains (CS Change of Customers on RTP) and MW impacts (Equilibrium Capacity),

[14] Baer, W., Fulton, B., Mahnovski, S. Estimating the Benefits of the GridWise Initiative, Phase I Report, Pacific Northwest National Laboratory (TR-160-PNNL), May 2004.

detailed in Tables 2 & 3 of the paper. These benefits are estimated for various price elasticities and share of customers (participation) on RTP. The elasticities range from -0.025 to -0.500, and the participation ranges from 33.3% to 66.6% to 99.9%. I use the paper's results for a scenario with an elasticity of -0.05 and a participation share of 66.6%. These results are a conservative generalization of SDG&E's AMI business case scenario for the CPP rate and the Peak Time Rebate program.

The elasticities for SDG&E's customers are detailed in Dr. George's testimony (Chapter 6) in Table SSG 6-11 for Residential customers and Table SSG 6-14 for C&I customers. SDG&E characterizes elasticities differently than the Efficiency Paper. SDG&E uses both an elasticity of Substitution and a Daily elasticity in combination to estimate demand response impacts. The Efficiency Paper uses a more general constant elasticity to estimate demand response. By examination of SDG&E's elasticities, I reasonably deduce that the -0.05 elasticity level in the Paper is an underestimate of SDG&E's combined elasticities.

The participation rates for SDG&E's CPP rate and Peak Time Rebate program are detailed in Mr. Gaines' testimony (Chapter 5). SDG&E participation ranges from about 70% for the Residential Peak-Time rebate to about 75% for the Large C&I customers on CPP. Thus, the 66.6% participant level from the Efficiency Paper is fairly conservative.

The paper estimates Annual Consumer Surplus change from a Flat rate to RTP to be $267.6 million per year for participants in the State of California in the scenario with an elasticity of -0.05 and 66.6% participation (Table 3 of the paper). I believe these are reasonable assumptions consistent with SDG&E's case. Normalizing this value, I calculate an efficiency gain of $38.71/kW-Year. This represents the reduction in capacity and energy costs for RTP participants moving from inefficient flat rates

This efficiency value must be discounted to adjust for the fact that it is based on flat rates and RTP. SDG&E's customers currently are on inverted tier rates for Residential customer, flat rates for Small C&I Customers, and

TOU rates for Medium and Large C&I customers. SDG&E's AMI rate design is the Peak Time Rebate program for Residential customers, TOU and the Peak Time Rebate for Small C&I customers, and the CPP rate for the Medium & Large C&I customers. The adjustments are detailed below to net out benefits attributable to flat rates, TOU rates, and CPP rates.

First, a discount is applied to reflect that a portion of the efficiency gain can be achieved simply by implementing monthly-changing flat rates. Monthly-changing flat rates do not require AMI metering technology. On page 14 of the Holland and Mansur paper, the authors conclude that "fully 30% of the deadweight loss is eliminated by allowing flat rates to vary monthly." Although elimination of deadweight loss is a small part of overall capacity and energy savings, SDG&E reduces the RTP efficiency gain by 30% to provide an estimate of cost reductions possible with monthly-changing flat rates.

Second, a discount is applied to reflect that efficiency gains are less for Medium & Large customers on TOU rates. TOU rates may already capture a share of the benefits obtained from RTP. The Efficiency Paper dedicates section IV to this topic. Efficiency gains for two-period TOU rates are estimated using three approaches, and summarized in Table Six of the paper. I calculate a 24% efficiency gains for TOU rates, compared to a flat rate, this leads to a 10% discount to net out efficiency gains from SDG&E customers currently on TOU rates.

Finally, I apply an adjustment to account for the fact that CPP rates and Peak-Time Rebates are more efficient than TOU rates. This adjustment is not directly estimated in the Efficiency Paper. However, at page 20 the author does state that his "preliminary analysis suggests CPP could capture a much greater share of the RTP efficiency gains than could TOU." Therefore, I assume that the CPP rate combined with Peak-Time Rebate will likewise capture a much greater share of the RTP efficiency gains than could TOU. Specifically, I assume that the AMI enabled demand response is twice as efficient as that from TOU rates. Therefore I include a 24% adjustment to represent the efficiency gains from CPP and the Peak-Time Rebate beyond those achieved by TOU.

The net effects of these discounts and adjustment are summarized in Table JCM 7-4. I begin with the $38.71/kW-year benefit of RTP over flat rates. This is discounted by a 30% ($11.61/kW-year) for monthly varying rates, discounted by 10% ($4.02/kW-year) for customers on TOU, and adjusted by 24% ($9.29/kW-year) for the efficiency differences between TOU and SDG&E's CPP rate and Peak-Time Rebate.

Table JCM 7-4

Estimated Capacity & Energy Benefits For RTP above and beyond CPP			
Annual Welfare Gain from RTP over Flat Rate		$	267,647,344
MW Reduction from RTP			6,914
Gross RTP Value per MW-Year		$	38,711
Gross RTP Value per kW-Year		**$**	**38.71**
Discount for Monthly Varing Prices	30%	$	(11.61)
Discount for TOU	10%	$	(4.02)
Adjustment for CPP/Rebate	24%	$	(9.29)
Net Value per kW-Year		**$**	**13.79**

Rate design flexibility has economic efficiency value. RTP can add benefits above and beyond the CPP rate and Peak-Time Rebate, estimated based on the Borenstein paper at $13.79/kW-Year.

3. Additional Reliability Value

Reliability value relates to opportunities for demand response programs that provide benefits discussed above in the Reliability Benefit section of my DOE Report summary. These Reliability benefits can be achieved with programs that encourage customers to reduce consumption or demand on short notice, and/or with technology such as Programmable Controllable Thermostats (PCT), automated energy management systems, and other future technological innovations. Such programs and technology are described by Mr. Gaines in Chapter 5 and Mr. Prushki in Chapter 11.

The DOE report values reliability programs as the decreased likelihood of a forced outage times the value of lost load (DOE Appendix B, p. 82). Value of lost load (VOLL) is a measure of how customers value electric reliability, or how much they are willing to pay to avoid an un-planned outage. The DOE Appendix states that "the accepted industry practice is to adopt a VOLL of $2-5 / kilowatt-hour (kWh), which represents an average value across the entire market." (p. 83). For illustration, SDG&Es PCT retrofit program provides over 34 MWs of load reduction in 2013.[15] Assuming the PCT retrofit program can on average avoid one hour of rolling black-outs each year, this would result in an annual value of $55,306 to $138,266 in 2013, or a 2006 present value for the period 2011 to 2038 from $0.7 to $1.8 million, at SDG&E's WACC of 8.23%. I estimate a 2006 reliability value at $.021 to $0.53/kW-Year, based on a 2011 peak reduction of 219 MWs.

4. Additional Unique Benefits of AMI

The list of additional unique benefits relating to AMI and AMI enabled demand response (which have not otherwise been addressed by other SDG&E witnesses) is potentially large and difficult to quantify. Their value is, nevertheless, undoubtedly greater than zero. These other unique benefits may be explored by the CPUC in demand response valuation workshops and in Phase 3 of R.04-04-025, and by the current research effort of the Demand Response Research Center. Additional benefits may include peak fuel diversity, reduction in market power of generators, smart home integration, and other demand side management innovations.

III. CONCLUSION

I identify several benefits of AMI enabled demand response, of which avoided CT Fixed Capacity and Energy costs are the main benefits. These benefits must be netted against potential market profits achievable with a CT. However, AMI enabled demand response provides unique benefits that a CT cannot provide, including reduced demand volatility, pricing flexibility, and improved reliability options. When CT market profits are weighed against these unique AMI and demand response benefits, on balance, they have a cancelling effect. Thus, my levelized fixed capacity valuation of $85/kW-Year

---

[15] The 34.3 MW peak reduction is based in the incremental price response due to the PCT. Reliability dispatch that raises the PCT set point with no override would likely produce even greater MW potential.

and my avoided energy benefits capture the net relevant factors and are appropriate for calculating the demand response benefits derived from AMI.

This concludes my testimony.

IV. QUALIFICATIONS OF JOHN C. MARTIN

My name is John C. Martin. My business address is 8326 Century Park Court, San Diego, California 92123. I am employed by San Diego Gas & Electric Company ("SDG&E") as a Business Economics Advisor in the Electric Measurement and Advanced Metering Department. In my current position, I am responsible for providing analysis associated with Advanced Metering Infrastructure (AMI).

I have over 17 years of energy industry experience. My current duties focus on costs and benefits associate with the capabilities of AM. This work draws upon my broad experience in the electricity and oil industry. My prior electricity work experience includes demand response program and tariff development, electricity trading and scheduling, demand side management program evaluation and load research of customer energy use. My duties also utilize my financial analysis experience in the oil refining, trading, and marking industry.

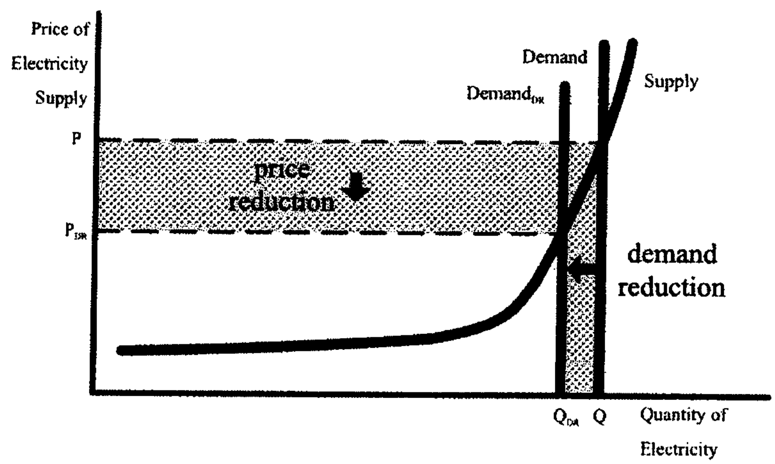
My education is in the general area of resource economics. I graduated from Cornell University in 1988 with a master's degree in agricultural economics. My bachelors of Science degree was granted by Purdue University in 1984 in business and farm management. This is my first opportunity to testify before the California Public Utilities Commission.

ATTACHMENT

JCM 7-1

BENEFITS OF DEMAND RESPONSE IN ELECTRICITY MARKETS AND RECOMMENDATIONS FOR ACHIEVING THEM

A REPORT TO THE UNITED STATES CONGRESS PURSUANT TO SECTION 1252 OF THE ENERGY POLICY ACT OF 2005



February 2006

U.S. Department of Energy

APPENDIX B. ECONOMIC AND RELIABILITY BENEFITS OF DEMAND RESPONSE

This Appendix provides a more detailed conceptual discussion of the economic and reliability benefits of demand response than was included in Section 3. First, short-term market impacts are described, drawing on economic theory to show how demand response can result in improved economic efficiency, and distinguishing how these benefits are manifested under different market structures. Next, long-term economic benefits from avoided capacity investments are discussed along with issues in designing and implementing programs designed with this goal in mind. Differences in how short-term and long-term economic benefits are realized and passed on to consumers are then described for vertically integrated utilities and regions with ISO/RTO spot markets. Finally, reliability benefits are described along with concepts used to value them.

Short-Term Market Impacts: Supply Costs and Market Prices

This section provides a detailed discussion of how customer load reductions lower energy supply costs in the short term. First, the basic source of short-term market benefits—improved economic efficiency brought about by allowing consumers to make electricity usage decisions based on marginal, rather than average, supply costs—is described. Differences in how these benefits are manifested in regions with differing market structures are then discussed.

Societal Benefits

In evaluating policies or structural changes that impact how markets work, economists distinguish between societal gains, which benefit everyone, and financial flows that involve gains by some at the expense of others, called transfers. In the absence of a way to weigh the relative impact on individuals of gains and losses (i.e., a change in utility), economists argue that policies should primarily be judged on their net outcome, which is defined by the level of societal benefits (see the textbox below).

Demand response produces societal benefits, which are resource savings, by reducing the gap between time-varying marginal supply costs and retail electricity rates based on average costs. Economic theory asserts that the most efficient use of resources occurs when consumption decisions are based on prices that reflect the marginal cost of supply. In a competitive market, this is defined by the intersection of a good's supply and demand curves (see Figure B-1). In electricity markets, the marginal electricity supply curve is constructed by ordering generators from lowest to highest operating costs (often referred to as "merit order").[68] Due to the technical characteristics of electricity generation equipment, the supply curve—the upward curving line in Figure B-1—tends

[68] Certain generators may be required to run, regardless of their marginal operating costs, to maintain reliability in areas with constrained generating and/or transmission capacity, which limits the ability of least-cost resources to serve local demand.

to increase very steeply at its upper end.[69] This means that when demand approaches the industry's installed capacity, each additional increment of demand imposes increasingly more cost than the previous one. In other words, the marginal cost of electricity becomes most sensitive to changes in demand when demand is already high.[70]

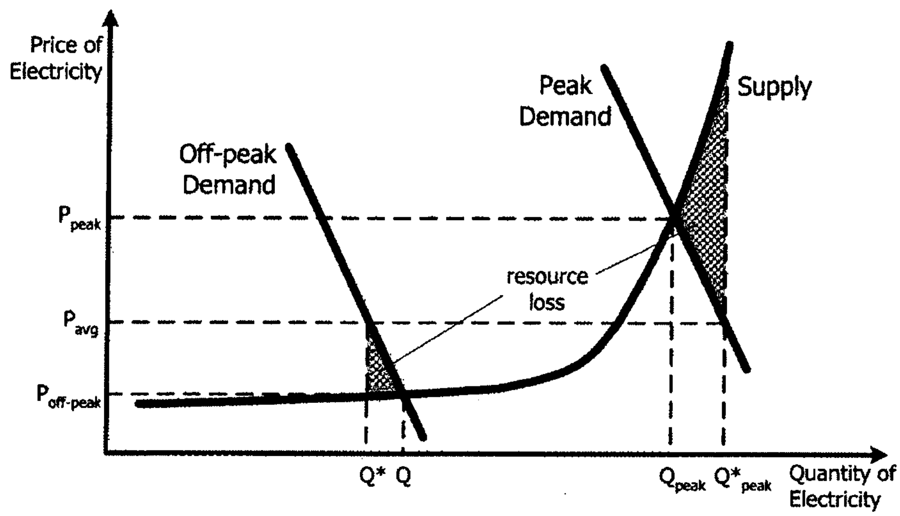


Figure B-1. Inefficiencies of Average-Cost Pricing

Like most goods, the demand for electricity exhibits declining marginal value (i.e., the marginal value of additional consumption declines as consumption increases). Electricity demand is characterized by a downward-sloping line, regardless of how electricity is priced. But, if the price that consumers pay never varies, demand appears to be perfectly inelastic, and is characterized by a vertical line. Moreover, consumers' demand for electricity also depends on the time of day, with more usage typically occurring during the "peak" afternoon and early evening hours and less at other times. This phenomenon is driven by the economic activity of businesses and residential customer lifestyles and usage patterns, but is also influenced by electricity rates that are the same throughout the day. For simplicity, the two lines labeled "peak" and "off-peak" in Figure B-1 represent consumer demand.

The most efficient pricing and usage of electricity is determined by the intersection of the supply and demand curves in Figure B-1. In other words, during off-peak periods, the efficient price of electricity should equal $P_{off\text{-}peak}$ and consumers would use an amount of

---

[69] The long, flat portion of the electricity supply curve represents "base-load" power plants, such as nuclear, hydroelectricity and coal plants that have very low operating costs and are run most hours of the year. Base-load plants are typically large with similar characteristics. The steeply inclining portion of the supply curve represents "peaking" plants that are used to meet peak demand needs and may be run only a few hours per year. These plants are typically natural gas- or oil-fired combustion turbines that are less expensive to build than most base-load technologies but have higher operating costs. Peaking plants are typically smaller units with varied operating characteristics.

[70] High demands do not always lead to high prices. If the entire portfolio of capacity is available, then the marginal unit may be relatively low cost. The steepest part of the supply curve is encountered when demands are especially high (e.g. a heat wave) or generation is short due to forced outages, or both.

electricity equal to Q, and during peak hours, the efficient price should equal P_{peak} and consumers would use Q_{peak} units of electricity. However, most consumers currently pay electricity tariffs that reflect average, rather than marginal, electricity supply costs; this is represented by P_{avg} in Figure B-1. Actual usage therefore reflects the intersection of the demand curves with this average price, resulting in less than the social optimal usage in off-peak periods (Q*) and more than the social optimal usage in peak periods (Q^*_{peak}) relative to the optimally efficient system.

Distinguishing Societal Benefits from Rent Transfers

Economists make a distinction between *transfers*—the benefits of a policy initiative that amount to gains for some at the expense of others—and *social welfare gains* that inure to society as a whole. Social welfare gains are desirable because they derive from efficiency improvements that benefit all market participants. These benefits provide a strong rationale for policymakers to invest consumers' money in initiatives to realize such gains. Transfers result in some market participants being better off than others. In the case of demand response, lower market prices reduce revenue to suppliers and lower costs to consumers. The economists' task is to quantify the relative marginal gains and losses to the individuals involved.

Some economists caution that treating market price reductions as benefits is misleading, and may result in policies that undermine, rather than enhance, market efficiency (Ruff 2002). Specifically, they contend that using the bill savings from price reductions, which largely amount to transfers, to justify demand response incentive payments to customers actually raises electricity prices in the long term. They contend that merchant generators count on the profits (called scarcity rents) realized when prices are high to recoup their capital costs and achieve the rate of return their investors require. If these profits are reduced because policymakers use them to justify customer curtailment incentives, then investors will become more skeptical and require higher returns, which, the argument concludes, results in higher prices in the long run.

This is the basis for many of the objections to allowing customers to bid load curtailments as resources into ISO/RTO spot markets, called "demand bidding as a resource." However, other economists contend that if demand response moves the wholesale market to greater economic efficiency and the result is a more appropriate supply and demand balance, then the elimination of those artificial rents to generators corrects a market distortion and prevents investments that are not needed based on how customers value electricity.

Another objection to demand bidding raised by some economists is their claim that customers on default service have no right to the energy, since the utility rates require that it be served, but do not give the customer any contractual rights to that supply. This could be corrected by requiring that in order to bid curtailments into spot energy markets, the customer would have to demonstrate that it has contractual rights to that power. As an alternative, these critics propose "self-financing" demand response whereby the inherent savings from avoiding paying high market prices is the inducement for customers to curtail, and no payment has to be made to achieve that result (Braithwait 2003).

These arguments have only been raised for demand response programs that allow customers to offer curtailments as resources in centrally organized spot markets. Yet, substantially the same transactions characterize demand bidding and CPP programs run by vertically integrated utilities.

Economists refer to the inefficiencies that arise when retail prices do not reflect marginal supply costs as "dead-weight losses" or resource losses (i.e., the loss of societal welfare when resources are not used optimally). The resource losses from average cost pricing are illustrated by the shaded triangles in Figure B-1. In the off-peak period, electricity that would have value to consumers if it were priced according to its marginal supply cost is not consumed—this represents a loss to society in economic activity that would have occurred but did not. In the peak period, consumers that do not pay the full marginal cost of power consume excessive amounts of electricity at a cost in excess of the value it provides them. Because this occurs at the steeply inclining portion of the electricity supply curve, these costs can be substantial.[71]

The short-term market-impacts benefit of demand response lies in reducing or eliminating this resource loss, thereby improving net social welfare. The combined resource loss from all peak and off-peak hours—and thus the potential for short-term demand response benefits—depends on how widely average and marginal electricity costs vary. For example, in a tightly constrained market, where peak demand is often very close to supply limits, the potential short-term efficiency benefit from implementing demand response can be substantial.

Supply Cost and Market Price Impacts in Regions with Differing Market Structures

Short-term market impacts are illustrated for vertically integrated utilities in Figure B-2. The supply curve typically reflects the utility's supply costs, including its own generation plants and any incremental wholesale power purchases. If demand is forecast to be Q, then a demand reduction that moves consumption to Q_{DR} results in an avoided utility supply cost equal to the shaded area in Figure B-2.

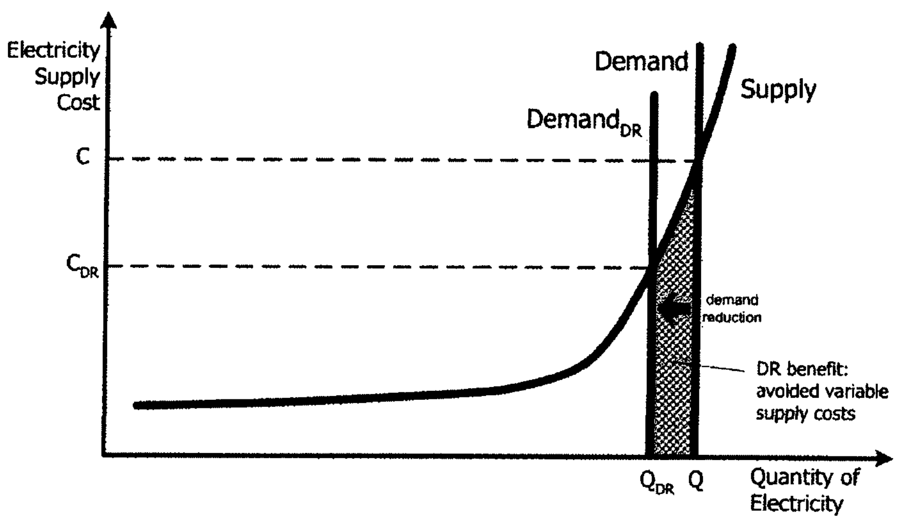


Figure B-2. Impact of Demand Response on Vertically Integrated Utility Supply Costs

The same load reduction produces more extensive impacts in regions with organized wholesale markets because of the way these wholesale markets are designed. The supply curve is developed by arranging generators' offer bids in merit order from lowest to

[71] Electricity pricing that does not reflect supply costs results in societal losses both when costs are high, and when they are low. However, the extent of these losses is greater at elevated supply costs, and therefore correcting prices in these periods has captured the attention of policymakers and market designers.

highest. Because of competition among generators, generators' offer bids reflect their marginal operating and maintenance costs and in some circumstances additional margins to recover fixed costs. LSEs also bid their expected load requirements into the market, producing a demand curve.[72] The bid price of last generator needed to serve the LSE's purchases sets the market clearing price for the whole market. This means that a demand reduction from Q to Q_{DR} not only provides the avoided variable cost savings observed for vertically integrated utilities (the shaded area to the right in Figure B-3), but it also lowers the price of all other energy purchased in the market. This second market impact, represented by the shaded rectangle in Figure B-3, is dependent on the level of price reduction—the difference between P and the new price P_{DR}—and the amount of energy bought in the applicable market. LSEs typically commit their expected energy requirements with a mix of bilateral forward contracts with generators and purchases in day-ahead and real-time markets. This is represented by the dotted line in Figure B-3. The extent of customer savings from price reductions thus depends on how much energy is purchased in spot markets.[73]

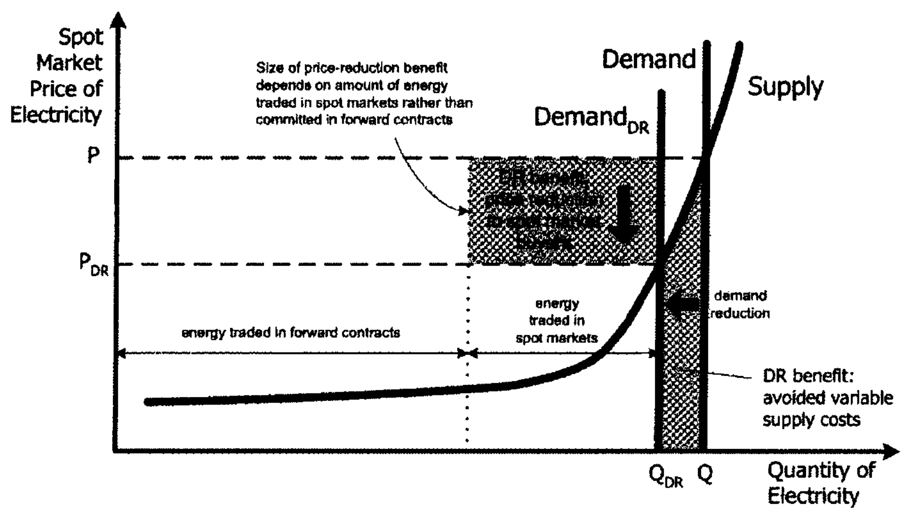


Figure B-3. Impact of Demand Response in Regions with Organized Wholesale Markets

In regions with organized wholesale markets, if, over time, customers routinely respond to high prices by curtailing or shifting loads, then additional, longer-term savings will result. Thus, if demand response consistently reduces market prices and volatility, bilateral contract prices will also drop over time, as reduced price risk in day-ahead and real-time markets pushes longer-term contract prices down. This is because LSEs may be willing to pay less for hedged forward contracts and will buy instead from the spot market if generators do not offer lower forward contract prices. In this way, lower energy

---

[72] In this example, demand is represented by a vertical line for simplicity (i.e., it is presumed to be fixed). Currently, most LSEs bid fixed quantities of electricity in spot markets, so this characterization is appropriate.

[73] 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.

prices resulting from short-term demand response market impacts can eventually extend to the entire market.[74]

Long-term Market Impacts: Capacity Benefits

The long-term market impacts of demand response hinge on reducing the *system peak demand*—the highest instantaneous usage by consumers in a particular market. Reducing system peak demand can avoid or defer the need to construct new generating, transmission and distribution capacity, resulting in savings to consumers. This applies for both vertically integrated utilities and organized wholesale markets, although capacity costs are allocated differently. This benefit can be specifically elicited from customers through capacity-based demand response programs (e.g., DLC, I/C rates or ISO/RTO capacity based programs) or may result from consistent load reductions from price-based demand response options (e.g., RTP). For example, in a capacity-based demand response program, load reductions timed to reduce load from a level that otherwise would have established the system maximum demand can yield large benefits for all consumers. Historical system maximum demand, adjusted for planned reserves, establishes ongoing generating capacity requirements, usually on an annual or semi-annual basis. For example, if the maximum demand served in a control area during the past summer was 5,000 MW, then that demand would serve as the basic capacity target for the next summer, to which an additional reserve margin (e.g., 18%) would be added.[75] If the existing infrastructure were insufficient to serve the resulting 5,900 MW capacity requirement, additional capacity would be necessary. Since generating capacity is expensive, ranging from about $50,000 to over $100,000 per MW-year (depending on the type and location of generating units), demand response that displaces the need for new infrastructure can produce substantial avoided cost savings.

Demand response programs designed to reduce capacity needs are valued according to the marginal cost of capacity. By convention, marginal capacity is assumed to be a "peaking unit", a generator specifically added to run in relatively few hours per year to meet peak system demand. Currently, peaking units are typically natural gas turbines with annualized capital costs on the order of $75/kilowatt-year (kW-year) (Orans et al. 2004, Stoft 2004). Thus, if demand response programs avoid 100 MW of generating capacity, the avoided capacity cost savings would be $7.5 million per year in this example. If the total program costs were $50/kW-year, including incentive payments to participating customers, then other customers realize the rest as savings (e.g., $2.5 million per year in this example), which may eventually be reflected in lower rates and bills. As long as there is some sharing of benefits, all customers benefit from others' participation in a capacity demand response program.

---

[74] Whether or not savings from short-term market price impacts and reduced forward contract prices brought about by incentive-based demand response programs should be treated as societal benefits is a subject of controversy (see the textbox on "Distinguishing Societal Benefits from Rent Transfers", earlier in this Appendix).

[75] Reserve margins vary in electricity markets across the U.S., but are typically 15-18%.

Transmission and distribution system capacity investments are also capital-intensive, and demand response that reduces local maximum demand in areas nearing infrastructure capacity can also provide significant avoided cost savings.

Realizing Capacity Benefits: Establishing and Reducing System Peak Demand

Capacity-based demand response programs are designed to replace generation investments and participants receive up-front capacity payments tied to this avoided cost. To realize this benefit and justify making the capacity payments, system operators must be able to dispatch curtailments that actually avoid building new capacity. This is accomplished in one of two ways: (1) predicting when system peak demand will exceed historic levels and dispatching load reductions accordingly or (2) dispatching curtailments when a designated peaking generation unit would otherwise be in service.

Dispatching demand response to avoid increasing system peak demand involves predicting when peak demand is likely to exceed historic levels absent any curtailments. Electric systems are generally either winter or summer peaking, meaning that annual demand is seasonal. However, demand can exceed historic peak levels several times during the peak season, which may span several months. To ensure that a capacity program truly does reduce peak demand, operators may need to dispatch the program several times during the peak season to account for forecast error. For participating customers, multiple curtailment obligations can be burdensome. To improve the attractiveness of capacity programs to customers, limits are sometimes are placed on how many curtailments can be called in a particular season.

The alternative method is to dispatch capacity-based demand response programs when an existing plant designated to meet peak demand would be needed to serve expected demand, absent any curtailments. This practice is somewhat more straightforward in regions with organized wholesale markets because transparent market rules direct dispatch operations. However, vertically integrated utilities have similar unit dispatch rules that could be used. Here too, limits may be placed on how frequently curtailments are called for.

Both methods of dispatching demand response to realize capacity value require provisions for periodic testing of customer response as well as penalties for non-performance. Testing is necessary to certify that customers truly have the capability to deliver the contracted curtailments on an on-going basis. Penalties serve to reinforce their obligation to be available and deliver load reductions when called. However, establishing appropriate penalty levels can be challenging. Increased penalty levels make demand response commitments more reliable and more valuable to the system operator, but are likely to reduce the amount of demand response committed by customers.[76] Program designers must balance the attractiveness of the program to customers against the potential consequences of forced outages that affect a large number of customers at costs well in excess of the avoided cost payment participating customers receive.

Because the avoided capacity cost savings calculation is prospective, so is the value of a capacity-based demand response program. This raises issues in forecasting the timing of system peak demand, or the highest 10-30 load hours of the year, so that calls for demand reductions actually moderate system maximum demand as designed. Since forecasting involves errors, program administrators/sponsors must make provisions to ensure the

[76] One useful strategy may be to recruit larger numbers of customer participants by dropping or reducing penalties for non-performance. Even though each customer is a less reliable source of demand response in the absence of penalties, the larger number of participants could increase the total expected demand response. The adoption of such a strategy would require evaluation of accumulated experience on the effect of various levels of penalties on customer performance.

demand response program is called often enough to effectively lower the forecast of system peak demand (see the textbox above).

Timing and Distribution of Market Impacts of Demand Response

Differences in market structure influence the timing and distribution of short-term and long-term market impacts of demand response in important ways. These differences are illustrated in this section by tracing the market impacts and resulting benefits of demand response in two types of market structure: 1) "vertically integrated systems", in which a vertically integrated utility with a retail monopoly franchise engages in some wholesale market transactions but operates in a region without an ISO or RTO, and 2) regions with organized wholesale markets in which ISOs/RTOs administer spot markets and retail competition is enabled at the state level. These illustrative combinations of retail and wholesale market structures reflect the current situation in many states or regions, although other retail/wholesale market structures are prevalent in the U.S.[77]

In this section, the examples suggest that the market impacts of demand response within organized spot markets produce benefits in a *shorter* timeframe than those for a vertically integrated, monopoly utility.

Market Impacts of Demand Response for Vertically Integrated Utilities

Vertically integrated utilities are responsible for making capacity investment decisions (whether to build new generation itself or to purchase supply contracts from other sources such as independent power producers), subject to regulatory oversight and approval, and for planning and operating the electricity grid and ensuring reliability. Retail rates are determined administratively, based on the average cost of supplying all three major facets of electricity production and delivery—production, transmission and distribution—and expected sales volumes. Embedded in retail rates are marginal costs to supply power, such as fuel, operating and maintenance costs, as well as a return on investment for un-depreciated utility-owned generation.

The economic impacts of demand response for a vertically integrated utility operating with a retail monopoly franchise are depicted in Figure B-4. Short-term demand response benefits may be traced as follows:

- Depending on the timing and type of demand response option, customers' load changes may be integrated into the utility's scheduling and dispatch decisions on a day-ahead or near-real-time basis.
- Changes in load (e.g., reductions in usage during high-priced peak periods) offset a portion of usage that otherwise would have been met by production from high-

[77] For example, utilities in some states are still vertically integrated and retain a retail monopoly franchise but are part of an organized regional wholesale market administered by an ISO or RTO (e.g., some parts of MISO, Vermont).

operating-cost power plants or purchases during the load response event (see Figure B-2).[78]

- This lowers the average variable electricity cost, which should be manifested eventually as customer bill savings through lower regulated electricity rates.

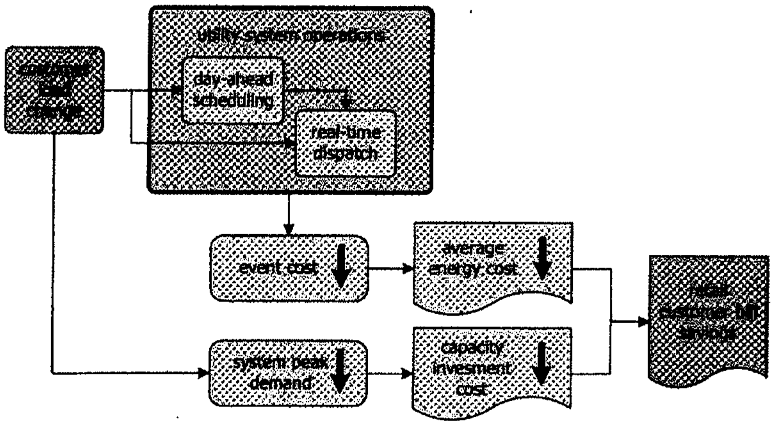


Figure B-4. Market Impacts of Demand Response for Vertically Integrated Utilities

The utility's return on capacity investments is recovered separately from its marginal costs to produce or purchase electricity and operate the electric grid. Thus, in vertically integrated systems, in the absence of a mechanism to reveal marginal capacity or reliability costs in unit operating costs, the short-term market impacts of demand response are limited to efficiency improvements in operating costs (including energy production and purchase costs) alone.[79]

In the long term, demand response that reduces peak demand growth directly averts the need for utilities to build more power plants, power lines and other capacity-driven infrastructure or to buy new capacity and energy from other suppliers (see Figure B-4). Because capacity investments are usually fully recovered—along with a pre-established return on investment—through higher retail electricity rates, these long-term benefits are realized over a multi-year period and can result in significant savings to consumers.

In vertically integrated, stand-alone utility systems, demand response is most useful to improve generation and transmission asset usage, avoid new capacity construction or purchases, and create more flexibility to assure reliable system operations. This influences the types of demand response programs preferred by vertically integrated utilities, as well as how they value and compensate demand response program participants.

[78] The converse is true for increases in load at times when the marginal cost of electricity is lower than the average retail price.

[79] Some utilities quantify the marginal value of reliability in their RTP tariffs quoting hourly prices to participants for changes in their usage from an established base amount; those hourly prices contain an explicit ($/kWh) marginal reliability (outage cost) element to reflect exigent reserve conditions (Barbose et al. 2004)

Market Impacts of Demand Response in Regions with Organized Wholesale Markets

About 60% of U.S. load is served by utilities or load serving entities that operate in regions with wholesale markets administered by ISOs/RTOs. Retail competition is also allowed in many of the states in these regions. These last-price wholesale electric commodity markets pay all competitively dispatched load a price determined by the last successful bid, which also sets the market clearing price. The market clearing price covers operating or production costs for the dispatched load (if each generator bids at least its marginal supply cost). If supply is very tight relative to demand, spot market energy prices will rise as more expensive units set the market clearing price. As a result, all units get the higher price, which includes creating "scarcity rents" for suppliers with costs below that of the marginal, price-setting unit.[80] Accordingly, spot energy prices serve as signals about whether additional supply- or demand-side capacity investments are needed, and what level of return to expect.

Three organized markets (NYISO, PJM, and ISO-NE) have established capacity payment mechanisms to create an additional stream of revenues for generators to recoup their investment costs. LSEs are required to purchase capacity in these markets to meet the expected peak demand of the customers they serve.

The impacts of demand response in an organized wholesale spot market are depicted in Figure B-5.[81]

The short-term market impacts of specific demand response events can be traced as follows:

- Depending on the timing and type of demand response option, customers' load changes may be integrated into day-ahead or real-time energy markets [as indicated by the arrows at the top of Figure B-5).
- Reductions in load during high-priced peak periods move marginal usage down the electricity supply curve (see Figure B-3), lowering market clearing prices during the demand response event (the event price in Figure B-5).
- This lowers LSEs' purchasing costs in the applicable wholesale market during the event. These savings may be captured by the LSE initially, but ultimately a significant share should be passed on to their customers (LSE event energy cost in Figure B-5).[82]

[80] This argument assumes that generators must recovery all of their revenue requirements and variable running costs, from energy sales at spot market prices. Some markets impose capacity requirements on LSEs that constitute a form of investment cost recovery for generators selling in those markets.
[81] The Midwest ISO (MISO), ERCOT and the California ISO (CAISO) all do not operate capacity markets.
[82] In some states, public utility commissions have adopted tariffs that specify the percent of savings that a regulated LSE providing default service must pass on to their customers. Eventually, competitive pressures should motivate LSEs to pass a significant portion of purchase cost savings to their customers.

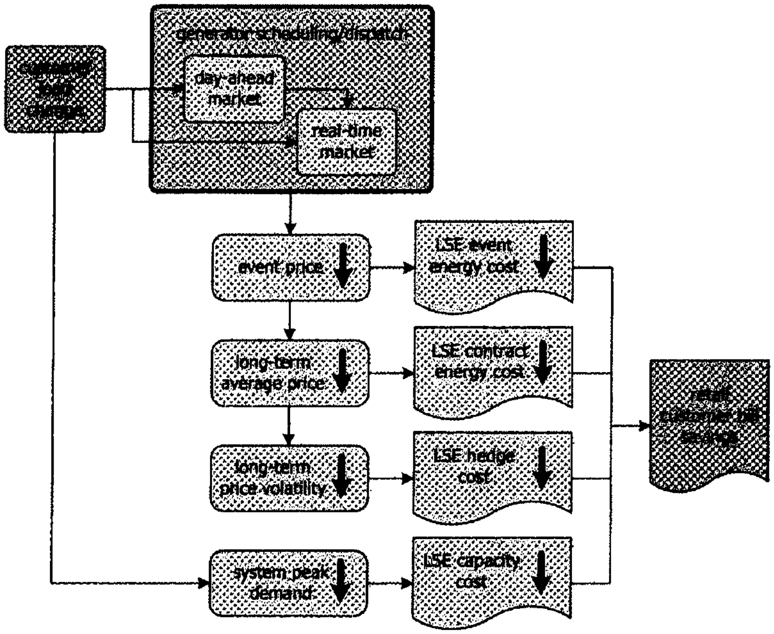


Figure B-5. Market Impacts of Demand Response in Regions with Organized Wholesale Markets

In regions with organized spot markets, demand response can produce cascading positive market impacts in the medium or long-term, realized over months or years (see Figure B-5):

- Reduced average market clearing prices can reduce forward contract costs for LSEs; these savings are then passed on to their customers (LSE contract energy cost in Figure B-5)
- Reduced volatility in market clearing prices puts downward pressure on risk premiums incorporated into hedged pricing products offered by competitive LSEs (LSE hedge cost in Figure B-5) and may lower transaction prices
- Lower forecast peak demand, resulting from demand response, also reduces LSEs' capacity acquisition requirements (LSE capacity cost in Figure B-5).

Long-term market impacts are less clear in organized wholesale and competitive retail markets compared to a vertically integrated utility system. A vertically integrated utility is allowed to directly pass through its capacity investment to customers in rates and likely most of its purchased energy and capacity costs as well; savings realized from demand response that avoids "uneconomic" investments or expenditures for peaking capacity are a direct source of cost savings to customers. In contrast, in organized spot markets, investment risk for new resources is assumed by the private sector. The combination of lower market clearing prices and reduced capacity requirements will dampen capacity investment signals, which should reduce construction of unneeded new power plants.

In summary, because organized spot markets use energy market clearing prices to pay generators for operating, but often only a fraction of the committed capacity costs, the long-term capacity savings benefits of demand response may not be fully monetized and paid to demand response providers. Because the spot market valuation of demand response is linked to wholesale market clearing prices (for energy and capacity) rather than avoided capacity costs, this creates different payment streams and priorities between the two market structures. Policymakers need to recognize these differences in designing demand response options and evaluating benefits derived from market impacts under these different market structures.

Reliability Benefits

In addition to improving the efficiency of electricity markets, demand response can provide value in responding to system contingencies that compromise the dispatcher's ability to sustain system-level reliability, and increase the likelihood and extent of forced outages. Electric systems in the U.S. conduct long-term planning exercises to specify the level of resources required to serve the system's anticipated maximum load reliably in the long term. Typically, planning reserve margins are 15-18% of historic maximum system demand.

System operators arrange for some of the available generation resources to serve as reserves to cover real-time load-serving requirements and avoid outages; operating reserves of 5-7% of forecast demand must be maintained at all times. The system operator typically uses standby generators, ready to be run in less than 30 minutes, to deal with abrupt changes in load or unexpected loss of generator or transmission availability. Demand-response based load reductions can be used to replace some of this stand-by generation to rebalance load and supply.

Demand response can supplement system reliability by providing load curtailments that help restore reserves, providing incremental reliability benefits to the system.[83] Customers participating in emergency demand response programs receive incentive payments for reducing load when called upon by the system operator. They receive no up-front capacity payments in some program designs because they are not counted on as system resources for planning purposes. Instead, they are supplemental resources, the need for which is not foreseeable, or even likely, but possible. They represent an additional resource for reliability assurance, distinct from capacity-based demand response programs (see the textbox below).

[83] The capacity they provide can be particularly valuable if located in what operators call "load pockets", localized areas with a shortage of available resources to serve load when a generator is out of service.

Roles of Capacity and Emergency Demand Response Programs

Emergency demand response programs provide benefits distinct from capacity-based demand response programs. In capacity programs, customers are paid incentives based on the avoided cost of new generation capacity and are counted among planned reserves. As such, they become part of the overall portfolio of resources assembled to meet system reserve requirements. Capacity-based demand response does not provide incremental system reliability—it supplants conventional resources in meeting established reliability goals, simply replacing what a generator that was not built would have provided.

In contrast, emergency demand response programs provide incremental reliability benefits at times of unexpected shortfalls in reserves. When all available resources, including capacity demand response programs, have been deployed and reserve margins still cannot be maintained, curtailments under an emergency demand response program reduce the likelihood and extent of forced outages. Load curtailments under emergency demand response programs are therefore valued according to their impact on system reliability.[84]

System operators generally dispatch emergency demand response programs only after exhausting all available capacity and operating reserves. When operating reserves are called upon to go from standby status to actually producing energy to serve load, the level of remaining operating reserves drops if additional replacement resources are not available. This is analogous to a consumer drawing down savings to pay an unexpected bill, leaving them more vulnerable to consequences from further unanticipated expenses.

System operators can reduce this vulnerability by asking emergency program participants to curtail load, thereby reducing system demand and operating reserve requirements. This means that some generating resources can revert to their standby status and be ready for another contingency event, and can be likened to a cash infusion to restore savings in the consumer analogy. The curtailment allows the operator to maintain reliability at prescribed or target levels (Kueck et al. 2001). At the margin, this form of demand response provides value, although it is not priced in any market.

Figure B-6 illustrates this impact, and provides a way to estimate these reliability benefits. The portrayed system has been scheduled to provide D_1 units of energy (including required reserves) at a price of P_1 at a specific time.[85] As the delivery time approaches, a system contingency arises that effectively pushes the supply curve to the left (e.g., a generator outage) or customer demand to the right (e.g., an unexpected surge in demand, as portrayed in the figure by the move from D_1 to D_2), so that supply and demand no longer intersect. This reserve shortfall is represented by the demand curve D_2. Activating an incentive-based demand response program initiates customer demand reductions that bring system demand back to D_1, thereby eliminating the reserve shortfall.

[84] It is possible that an emergency demand response program, while not explicitly designed to fulfill capacity requirements , may nonetheless be capable of providing some level of capacity benefits as well.
[85] In this example, customer demand is represented by a vertical line, because in a reliability event, which occurs within minutes or seconds of power delivery, demand may be viewed as fixed.

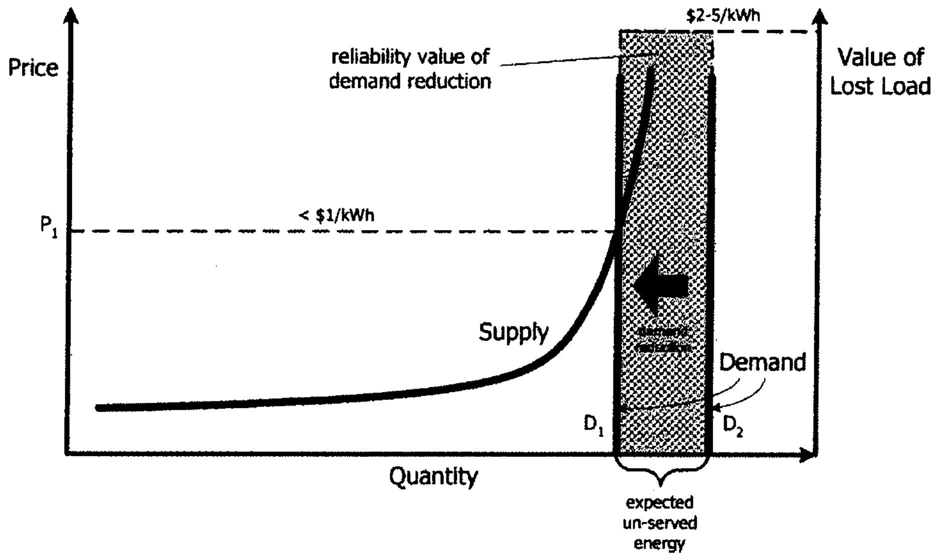


Figure B-6. Valuing the Reliability Benefits of Demand Response

While the price of served energy is determined by market conditions (P_1 in Figure B-6), the value of the demand reduction is defined by the decreased likelihood of a forced outage. Economists define the concept of *value of lost load* (VOLL) as the proper measure of improved reliability, since it reflects customer's marginal value for electricity under these circumstances. The product of VOLL and the *expected un-served energy* (EUE), the load that otherwise would not have been served, monetizes the value of the load curtailments (see the textbox below). This is represented by the shaded rectangle in Figure B-6 in the case where the curtailed load corresponds exactly to the amount of expected un-served energy.

Emergency demand response programs can provide low-cost, incremental resources to preserve reliability in various market structures; at present, the most prominent examples are implemented by the Northeast ISOs.

Value of Lost Load and Expected Un-Served Energy

"Value of lost load" (VOLL) is a measure of how customers value electric reliability, or what they would be willing to pay to avoid a loss of service. It varies among customers but is almost always greater than the retail price of electricity because customers incur costs from being disconnected without notice. Customer values factored into VOLL include inconvenience or discomfort, loss of sales or productivity (e.g., at retail premises or factories), large cleanup and restart costs (e.g., at pharmaceutical companies), and overtime costs to make up for lost production. Given the wide range of customer circumstances and difficulties in predicting which customers will be affected by a particular outage, the accepted industry practice is to adopt a VOLL of $2-5/kilowatt-hour (kWh), which represents an average value across the entire market.

"Expected un-served energy" (EUE) is a measure of the magnitude of a reserve shortfall. It takes into account the change in the likelihood of a curtailment and the consequences of such an event: how much load would have been forced off-line by dispatchers in such circumstances if the curtailments had not been undertaken. NYISO concluded that during the service restoration effort following the 2003 northeast blackout, demand response curtailments reduced forced outages kWh for kWh, because they enabled smoother service restoration. However, under other, less extreme conditions, curtailments were found to produce less than proportional reductions in EUE (NYISO 2003).

ATTACHMENT

JCM 7-2

CSEM

Center for the Study of Energy Markets

CSEM WP 133r

The Long-Run Efficiency of Real-Time Electricity Pricing

Severin Borenstein

Revised February 2005

This paper is part of the Center for the Study of Energy Markets (CSEM) Working Paper Series. CSEM is a program of the University of California Energy Institute, a multi-campus research unit of the University of California located on the Berkeley campus.

UCEI
University of California Energy Institute

2547 Channing Way
Berkeley, California 94720-5180
www.ucei.org

The Long-Run Efficiency of Real-Time Electricity Pricing

by

Severin Borenstein[1]

Revised February 2005

Abstract: Retail real-time pricing (RTP) of electricity – retail pricing that changes hourly to reflect the changing supply/demand balance – is very appealing to economists because it "sends the right price signals." Economic efficiency gains from RTP, however, are often confused with the short-term wealth transfers from producers to consumers that RTP can create. Abstracting from transfers, I focus on the long-run efficiency gains from adopting RTP in a competitive electricity market. Using simple simulations with realistic parameters, I demonstrate that the magnitude of efficiency gains from RTP is likely to be significant even if demand shows very little elasticity. I also show that "time-of-use" pricing, a simple peak and off-peak pricing system, is likely to capture a very small share of the efficiency gains that RTP offers.

[1] Director of the University of California Energy Institute (www.ucei.org) and E.T. Grether Professor of Business Administration and Public Policy at the Haas School of Business, U.C. Berkeley (www.haas.berkeley.edu). Email: borenste@haas.berkeley.edu. My thanks to Carl Blumstein, Jim Bushnell, Ali Hortacsu, Ed Kahn, Erin Mansur, Karen Notsund, Celeste Saravia, Ralph Turvey, Bert Willems, Frank Wolak , Catherine Wolfram, participants in seminars at U.C. Berkeley, U.C. Santa Barbara, U.C. Irvine, Wharton, and University of Florida, and two anonymous referees for valuable comments. Meredith Fowlie and Amol Phadke provided excellent research assistance. This work grew directly from related research with Stephen Holland. Many hours of valuable discussion with Stephen have shaped my thinking on RTP issues, though he bears no responsibility for any errors in this paper.

Over the last few years, a great deal has been written about time-varying retail pricing of electricity. Many authors, myself included, have argued that real-time retail electricity pricing (RTP) – retail prices that change very frequently, *e.g.,* hourly, to reflect changes in the market's supply/demand balance – is a critical component of an efficient restructured electricity market. During the California electricity crisis in 2000-2001, RTP boosters pointed out its value in reducing the ability of sellers to exercise market power. While nearly all economists have supported RTP conceptually, Ruff (2002) among others has argued that it is important to distinguish between RTP's long-run societal benefits and the short-run wealth transfers it might bring about. In particular, the reductions in market power primarily prevent a short-run wealth transfer from customers to generators, though the transfers can still be quite large.

In this paper, I estimate the magnitude of the potential long-run societal gains from RTP, abstracting from market power issues and short-run wealth transfers in general. I do this by formulating a model of competitive electricity generation with demand and production costs based on actual data from U.S. markets. I solve computationally for the model's long-run competitive equilibrium, with the results indicating the amount of each possible type of capacity that would be built, the prices that would be charged to customers on RTP and on flat-rate service, and the total social surplus that would be generated by the system. The model also allows estimation of the transfers that would occur among customers if customers on RTP had demands that were (absent RTP) peakier or flatter than customers not on RTP.

The estimates indicate that RTP would substantially reduce peak electricity production and thereby reduce the use of low-capital-cost/high-variable-cost peaker generation. The social gains from RTP for at least the largest customers in the system are estimated to far outweigh reasonable estimates of the metering cost. The magnitudes of the social gain are sensitive to the demand elasticity that is assumed, but the results indicate that even with quite small elasticities, the benefits are substantial.

Section I presents the economic model that is the basis for simulations. Section II explains the data used in the simulations and the process used to compute long-run equilibria. The results of the simulations are presented and their implications discussed in Section III. In section IV, I carry out a similar analysis on a much simpler pricing system, time-of-use (TOU) pricing, in which there are simple peak and off-peak periods, with the prices differing between periods, but being held constant for months or even years at a time. Section V discusses a number of factors that are omitted from the simulations and suggests how those factors are likely to affect the results. I conclude in Section VI.

I. Model of Long-Run Competition in Electricity Markets

The model that is the basis for the simulations is adapted from Borenstein and Holland (revised 2004a, hereafter BH).[2] It assumes a simple competitive wholesale and retail market structure. The retail structure is identified only by the way in which it charges end-use customers for electricity, using a flat rate or RTP. The price(s) charged to each group allow the retailer to exactly break even on service to that group. As in BH, this reflects the outcome of competition among many retail providers, but it also could be interpreted as a single regulated retail provider that is required to exactly cover its costs and required not to cross-subsidize between flat-rate and RTP customers. Following BH, I assume for simplicity that retailers have no other transaction costs.

I assume free-entry of generators of three different types. Generation exhibits no scale economies, with each generation unit having a capacity of one megawatt. The types of generation differ in their fixed and variable costs, higher fixed costs being associated with lower marginal cost of production. For generator type j, annual generator costs are modeled a fixed cost plus variable costs that are linear in the number of megawatt-hours produced during the year, $TC_j = F_j + m_j \cdot MWh_j$. Startup costs and restrictions on ramping are not considered, an issue discussed in section V. Parameters used for this and all other aspects of the simulations are discussed in the next section.

Demand is modeled as constant elasticity, using a range of possible elasticities. Within any one simulation, demand is first assumed to have the same elasticity in all hours. I

[2] A slightly different version of this model with continuous marginal cost functions is in Borenstein & Holland forthcoming.

then consider the effect of demand elasticity varying positively or negatively with the level of demand. The level of demand in each hour is taken from the distribution based on the actual levels of demand in various US electricity regions, as explained in the following section. Cross-elasticities across hours are assumed to be zero, another issue discussed in section V.

Some proportion of customers, α, are on real-time pricing, and the remainder are on flat-rate service. I assume that all customers have identical demand up to a scale parameter. Thus, following BH, if the total demand in hour h is $D_h(p_h)$ and the flat-rate service customers are charges $\bar{p}$ in every hour, the wholesale demand is

$$\tilde{D}_h(p_h, \bar{p}) = \alpha \cdot D_h(p_h) + (1 - \alpha) \cdot D_h(\bar{p}). \qquad [1]$$

In this case, demand is modeled as constant elasticity, $D_h(p_h) = A_h \cdot p_h^{\epsilon}$.

Under these assumptions, for any set of installed baseload, mid-merit, and peaker capacity, K_b, K_m, K_p, there is a unique market-clearing wholesale price in each hour, provided that total installed capacity exceeds demand from flat-rate customers in every hour, $K_b + K_m + K_p > (1 - \alpha) \cdot D_h(\bar{p}) \ \forall \ h$. In the following section, I discuss the algorithm for finding the short-run equilibrium for any set of installed capacity and the long-run equilibrium allowing capacity to vary. In presenting the algorithm, I demonstrate that there is a unique long-run equilibrium.

In addition to establishing long-run equilibria for any $0 \leq \alpha < 1$, it will be important, as a baseline, to determine an equilibrium with no customers on RTP. The model above is not applicable to a market with no RTP customers, because without RTP there is no short-run demand elasticity, so in order to meet demand in all hours, sufficient capacity must be built so that the market always clears "on the supply side," *i.e.*, at a price no greater than the marginal generation cost of the technology with the highest marginal cost. Such an organization requires some sort of additional wholesale payment to generation in order to assure that demand does not exceed supply in any period and, at the same time, that generators' revenues exceed their variable costs over a year by an amount sufficient to cover their fixed costs.

It is straightforward to show that the annual capacity payment that assures sufficient generation and the optimal mix of generation is equal to the annual fixed costs of a unit of peaker capacity. To avoid distorting the mix of capacity, this payment is made to all units of capacity, regardless of type.[3] The payment is financed by increasing the price of the flat-rate electricity service until it generates sufficient revenue to cover the capacity payments. That is how simulation of the baseline flat-rate service is implemented in the following section.

II. Data, Model Details and Solution Algorithm

The value of the simulation results depends on the realism of the underlying assumptions. In this section, I describe in detail the modeling of demand and supply, and then the algorithm for finding the long-run competitive equilibrium. I first present the details of the model, and then discuss the data used to parameterize the model.

Demand, Supply and Equilibrium Modeling

Within each hour, each customer's demand is modeled as constant elasticity. Each customer i is assumed to have a demand that is simply a fixed proportion, γ_i, of total demand. In the base simulations, I assume that total demand has the same elasticity in all hours, but this is later relaxed to allow elasticity to vary positively or negatively with the overall demand level.

The aggregate demand function for hour h can be specified as $D_h(p_h) = A_h \cdot p_h^{\epsilon_h}$, where elasticity may or may not vary by hour depending on the simulation run. For any share of demand on RTP, α, the demand from customers on RTP is then $D_h(p_h) = \alpha \cdot A_h \cdot p_h^{\epsilon_h}$ and the demand function for customers on flat rate service is $D_h(\bar{p}) = (1-\alpha) \cdot A_h \cdot \bar{p}^{\epsilon_h}$. The aggregate demand in the wholesale power market is then $\tilde{D}_h(p_h, \bar{p}) = \alpha \cdot A_h \cdot p_h^{\epsilon_h} + (1-\alpha) \cdot A_h \cdot \bar{p}^{\epsilon_h}$.

Given an elasticity for a certain hour, ϵ_h, and the assumption of a constant-elasticity

[3] This would also be the outcome if the wholesale price exceeded the marginal cost of the peaking generation only in the highest demand hour of the year, and the price in that hour was equal to the marginal cost of the peaker plus its annual fixed cost.

functinal form, demand is fully specified by A_h, the scale parameter. A_h is determined by any one price/quantity point on the demand curve, which I refer to as the demand "anchor point" for the hour. I assume that at a given constant price (discussed next), the anchor quantity demanded takes on a distribution equal to the actual distribution of quantities demanded from a certain electricity control region.

The constant price used to specify the anchor points is chosen to be the price that would allow producers to break even if it were charged as a flat retail price to all customers. This is not the actual flat rate (or time-of-use rate) that was charged to customers during the observed period from which the demand distribution data are taken. The difference, however, will not substantially change the results for two reasons. First, at the low elasticities I consider in the simulations, a change of 10%-20% in the base flat rate that I assume (which is the magnitude of the potential difference between the rate assumed and the actual flat rate in use) will change quantity demanded very little. Second, and more important, the overall level of base demand is just a scale factor in the simulations. The value of using an actual distribution comes from accurately representing the *shape* of the distribution; that changes negligibly with the assumption made about the level of the flat retail rate.

Once the wholesale demand function has been specified each hour, that can be combined with the production technologies to calculate the long-run equilibrium capacity of each technology type. Note that from any given baseload, mid-merit, and peaker capacities, K_b, K_m, K_p, one can determine a short-run industry supply function and therefore wholesale prices for each hour. From those prices, one can calculate the profits of owners of each technology type. In the long-run each technology type is built to the point that one more unit of that capacity would cause profits of all owners of the capacity to be negative. So, the goal is to identify the mix of capacity that causes this condition to hold for all three technologies simultaneously.

At first, this might seem difficult, and it might seem that there could be multiple long-run equilibria or none, but in fact there is a unique technology mix that satisfies this condition. To see this, begin with the peaker technology which, if it is used at all, will be used in the highest demand hour. It is straightforward to find a unique long-run equilibrium if supply is restricted to use only the peaker technology. One simply expands the quantity of peaker capacity, recalculating the associated short-run equilibrium with each increment in capacity, until expansion of capacity by one more unit, causes profits to go negative. Call the capacity level that satisfies this condition K_{tot} since that will generally turn out to be the equilibrium total amount of capacity.

In this peaker-only equilibrium, all rents to generators are earned when production quantity is equal to K_{tot}. In hours with lower equilibrium quantity, price must be equal to peaker marginal cost. Now, begin substituting mid-merit capacity for peaker capacity. Once built, the mid-merit capacity will all be used in any given hour before any of the peaker capacity is used; it is lower on the supply function than the peaker capacity, The key is to recognize that substituting mid-merit for peakers units, holding total capacity constant, does not change the rents earned by the remaining peaker units. In fact, so long as one peaker unit remains, the rents it earns are unchanged by substituting lower-MC technologies for the other units.[4]

Continuing to substitute mid-merit for peaker units will drive down the equilibrium profits of mid-merit units until one more unit would drive the profits of all mid-merit units to be negative. Call the largest capacity of mid-merit units that still earns positive profits, K_{bm} because this will generally turn out to be the total of the baseload and mid-merit capacity. Next, begin substituting baseload capacity for mid-merit units. Note that this does not change the rents to mid-merit units. Continue this substitution until one more baseload unit would drive baseload profits negative. This is K_b. Then, $K_m = K_{bm} - K_b$ and $K_p = K_{tot} - K_m - K_b$. These are the unique long-run competitive equilibrium capacity levels for a given set of available technologies, share of customers on RTP (α), and flat rate ($\bar{p}$).[5]

[4] This description assumes that equilibrium capacity investment includes at least one unit of each type of capacity. If peaker capacity is dominated by mid-merit or baseload for even the least utilized peaker unit, or if mid-merit is dominated by baseload for the least utilized mid-merit unit, then the same process is followed omitting the dominated technology.

[5] These searches were done inefficiently from a computing standpoint, as grid searches with a 1 mW

This equilibrium, however, may not satisfy the retailer breakeven condition, so one must calculate the profits retailers earn on flat rate customers in this equilibrium. If it is not zero, then one adjusts $\bar{p}$ up or down and resimulates capacity. When the resulting equilibrium yields zero profits for retailers as well as generators, this is the unique long-run competitive equilibrium in the generator and retailer markets given the set of available technologies and share of customers on RTP (α). Using this supply function, one can then calculate the equilibrium distribution of prices, loads (quantities), and the consumer surplus for each group.[6]

Data Inputs for Simulation

The critical inputs for the simulation are a load profile, demand elasticities, and cost characteristics of the production technologies.

The load profile determines the distribution of quantity demand and the flat rate when all customers are on flat-rate service, as described in the previous section. For the simulations presented in here, I use five years of hourly demand data from the California Independent System Operator, 1999 through 2003.[7] This period includes both relatively cool summers and quite hot summers.[8] As pointed out earlier, the importance of the load distribution used is in the shape of the load duration curve, not the overall size of the loads. It appears that load duration curves don't differ that much in shape from one control area to another.

Electricity demand elasticities are a subject of nearly endless contention. The relevant

---

grid over a very wide range of possible capacity quantities. They still converged quite quickly on a desktop PC.

[6] The updating algorithm for $\bar{p}$ was to always reset it to the level that would have broken even given the prior iteration's quantities demanded by flat-rate customers and the wholesale prices from the current iteration. This usually converged in two to four iterations on $\bar{p}$.

[7] I adjust the baseline hourly demand data for the fact that about half of all demand is on time-of-use rates (TOU). I do this by assuming that the elasticity of demand with respect to TOU price variation is -0.1 and that the price ratios among TOU periods are equal the average ratios in the TOU rate schedules offered by Pacific Gas & Electric and Southern California Edison.

[8] I've carried out the same analysis using datasets from the ECAR (upper midwest) and NPCC (New England) regions with very similar results.

Table 1: Generation Costs Assumed in Long-Run RTP Simulations

Generation Type	Annual Capital Cost	Variable Cost
Baseload	\$155,000/MW	\$15/MWh
Mid-merit	\$75,000/MW	\$35/MWh
Peaker	\$50,000/MW	\$60/MWh

elasticity would be a short-run elasticity, but still recognizing that customers would know well in advance that prices would be volatile. The actual elasticity will depend in great part on technology, as automated response to price changes will surely become easier over time. I simulate for a fairly wide range of elasticities from -0.025 to -0.500. The range -0.025 to -0.150 illustrates that likely impact of RTP in the short run and under current available technologies for demand response. Probably the two most current and relevant sources for elasticity estimates, Patrick and Wolak (1997) and Braithwait and O'Sheasy (2002), derive estimates that span this range. In the longer run, however, real-time demand response will become easier to automate and larger elasticities might be expected, so I include results using -0.3 and -0.5 as well. All demand levels are calculated based on the full retail price, which is assumed to be the cost of power plus \$40/MWh for transmission and distribution (T&D).[9]

The assumptions about production technology are presented in Table 1. They are intended to represent typical capital and variable costs of baseload, mid-merit, and peaker technologies, corresponding roughly to coal, combined-cycle gas turbine, and combustion turbine generation. The numbers were derived from conversations with industry analysts. The variable costs depend on fuel prices, and are meant to include variable O&M.[10] The

[9] I assume that the T&D charge is not time-varying. T&D could also be subject to real-time pricing if capacity constraints become binding at some times.

[10] For these costs, the price of natural gas is assumed to be \$4.25/MMBtu and variable O&M is assumed to be \$1/MWh.

annual fixed costs are more difficult to determine precisely in part because they depend on the cost of capital and on the rate of economic depreciation of the plant. These figures appear to be in what most industry analysts would consider to be a reasonable range.

Two further comments on plant costs are warranted. First, the results are not particularly sensitive to the exact cost assumptions on the baseload and mid-merit technology. The different effects of RTP under varying assumptions on elasticity and the share of customers on RTP are driven mostly from changes in the amount of peaker capacity that is built. In future versions, I will include a range of cost assumptions. Second, this paper presents an easily-replicated algorithm for analyzing the long-run effect of introducing demand elasticity. For whatever cost assumptions the policy analyst believes are appropriate, this technique can be used to analyze the long-run implications.

III. Simulation Results and Implications

The first line of Table 2 presents the equilibrium flat rate ($79.68/MWh, which includes $40/MWh for transmission and distribution), as well as the capacity that is utilized in efficiently providing the demand under the flat rate, and the total energy consumed and cost of that energy. The remainder of the table presents the equilibrium capacities and information about equilibrium price distributions under scenarios with varying proportions of customers on RTP and with those customers exhibiting various demand elasticities. Within each simulation, demand has the same elasticity in all hours.

It is apparent from Table 2 that with even moderate demand elasticity, RTP will significantly change the composition of generation, as indicated in columns F,G,H and I. The greatest effect will be a large decline in the amount of installed peaker capacity (column H). Mid-merit capacity (column G) would likely also decline and baseload capacity (column F) would increase, though these changes would be small in comparison to the potential for drastic reductions in peaker capacity. Figure 1 shows the load duration curves for simulations with varying elasticities and one-third of customers on RTP.[11] The highest

[11] A load duration curve shows the number of hours (horizontal axis) in which the quantity demanded will be at least a certain level (vertical axis).

curve, representing all customers on a flat-rate tariff, has one hour in the upper left corner in which quantity hits 46928, which is the highest quantity demanded when all customers are on flat rates. Note that the other curves, representing differing demand elasticities for the one-third of demand on RTP, flatten out at different load levels, with lower peak demand levels associated with greater demand elasticity. For demands in these regions, the market clears "on the demand side," *i.e.,* on the vertical portion of the supply curve (constant quantity, varying price). This illustrates the effect shown in column I in table 2: RTP has a very significant effect on the total capacity needed because for the highest demand periods, the market equilibrates by raising price rather than building additional generation capacity that is used for only a few hours per year.

A question that frequently arises with RTP is how high prices could get and whether "bill shock" during a high-price month would undermine the program. This concern, of course, is greatly mitigated by forward contracts and other financial instruments, as explained in Borenstein (forthcoming). Customers that hold fixed-quantity forward contracts can eliminate most price risk without reducing the strong price incentives on marginal purchases.

Setting aside hedging instruments, however, it is apparent from Table 2, columns J,K,L and M, that an RTP program could yield very high prices for a few hours. With very inelastic demand, the prices would be extremely high in some hours. The reason for these high prices are shown in column K, which shows the total number of hours during the 5-year period in which all capacity was used and, thus, the price was above the marginal cost of a peaker plant. With extermely inelastic demand, the peaker plants must recover all of their fixed costs over just a few hours, so spectacular price spikes are dictated. But taken in the context of the annual bill, even the very high prices seem more manageable. With a demand elasticity of -0.1, column L shows that the highest price hour would amount to 4.2% of the annual bill. Column M indicates that the 10 most expensive hours of the 5-year period, if they all occured in a single month, would account for about 22% of the annual bill. Although these amounts would be substantial in monthly bills, the suggestion that a customer would find that half or more of its *annual* bill occurs in two or three hours is not consistent with my findings.[12]

Before leaving table 2, it is worth pointing out that RTP is not an energy conservation program. In these simulations, the aggregate energy consumed actually increases slightly (0%-2%), though that this could be due to the constant-elasticity demand function; in theory, total quantity consumed could increase or decrease. By lowering off-peak prices and lowering overall average prices,there is a real possibility that RTP would stimulate increased aggregate consumption of electricity.

The overall effect of RTP on social welfare is presented in Table 3. Because I use constant-elasticity demand curves, for which total consumer surplus is undefined, I evaluate the effects by calculating the *change* in consumer surplus from the flat-rate tariff consumer's faced before RTP was introduced. Thus, the equation for aggregate change in consumer surplus over the H hours simulated is:

$$\Delta CS = (1-\alpha)\sum_{h=1}^{H}\frac{A_h}{\epsilon+1}\cdot(\hat{P}^{\epsilon+1}-\bar{P}^{\epsilon+1})+\alpha\sum_{h=1}^{H}\frac{A_h}{\epsilon+1}\cdot(\hat{P}^{\epsilon+1}-P_h^{\epsilon+1}) \qquad [2]$$

where $\hat{P}$ is the flat rate prior to introduction of RTP and $\bar{P}$ is the flat rate in equilibrium after α share of demand is on RTP. The A_h for each hour are set so as to include the actual quantity demanded at a price of $\hat{P}$, as described earlier. The annual average ΔCS is shown in column C of table 3. Columns E and G break out that number into the two terms in equation [2], which represent the change in surplus, still compared to having everyone on flat rate, for customers who stay on flat rate (column E) and for customers who move to RTP (column G).

It is immediately clear that the surplus gains from real-time pricing are substantial, even if demand of customers on RTP is quite inelastic. With an elasticity of only -0.025, the surplus gain from putting one-third of demand on RTP, shown in column C, is over \$100 million per year. To give these figures some context, in 2001 the state of California appropriated \$35 million as a *one time* cost of installing real-time meters for the largest

[12] Note that unlike the surplus comparisons I make below, this comparison is to the total bill including non-energy (T&D) components of the bill. This seems appropriate given that the concern is bill shock. Roughly half of the total bill is energy and the remainder is T&D.

customers in the state, representing slightly under one-third of total demand. That isn't the only cost of switching these customers to RTP, since billing systems must be changed as well, but there are also other benefits to the meters, including remote meter reading that can yield big labor savings. Nonetheless, as shown in column D, the savings are still a fairly modest share of the total energy cost for the system, less than 10% for all but the most optimistic case, and quite possibly less than 5%. Still, as discussed in section V, the long-run energy market impact analyzed here is only one part of the value of RTP.

It is also clear that the total surplus gains from RTP are highly non-linear in both the elasticity of demand and the share of demand that is on RTP. There is diminishing returns to both greater elasticity and a greater share of demand on RTP. For most elasticities, putting one-third of demand on RTP achieves more than one-half the benefits of putting all demand on RTP. For any given $\alpha > 0$, a demand elasticity of -0.05 generates more than half the benefits of a demand elasticity of -0.15.

Decomposing the change in total surplus reveals two effects that BH demonstrate theoretically. First, column E shows that flat-rate customers are made better off by other customers moving to RTP. Column F calculates the "per capita" benefit for a hypothetical customer who makes up 0.001% of the total demand ($D_h(p_h)$) in any given hour.[13] This customer on flat rate billing benefits as an increasing share of other customers moves to RTP. This effect is frequently argued by parties who advocate subsidizing RTP participants.

A second effect, however, suggests that policy is not always wise: as demonstrated theoretically by BH, customers moving to RTP harm other customers who are already on RTP. This is shown numerically in column H, which presents the "per capita" benefit of a customer (again representing 0.001% of total demand) on RTP when the total share of customers on RTP is the α in column B. We see that the benefits to a customer on RTP decline as more customers switch to RTP. In fact, the overall externality from a group of

---

[13] This would be a customer with a peak demand of about 450kW. In California, there were approximately 8,000 customers of at least this size during the sample period.

customers moving to RTP can be positive or negative, as shown in column J.[14]

Elasticity Varying with Demand Level

In the simulations presented thus far, the elasticity of demand has been the same in all periods, the case in which BH show that the equilibrium flat rate will be equal to the optimal flat rate. BH also show that if demand elasticity is greater in high-demand periods than in low-demand periods, the equilibrium flat rate will be below its optimal level. BH demonstrate that in that case it is theoretically possible that moving more customers on to RTP could lower long-run equilibrium total surplus.

I simulate this case by allowing elasticity of demand to vary with the level of demand, where the level is indicated by the quantity demanded if all customers were charged the flat rate.[15] The elasticity of demand varies linearly with demand level, in this case from 50% of the original demand elasticity for the lowest demand level to 192% of the original demand elasticity for the highest demand level. These boundaries were chosen so that the demand-weighted average elasticity is equal to the original demand elasticity in order to allow some comparability to the previous simulations.

Omitting a few of the columns, table 4 presents results comparable to tables 2 and 3, but for a simulations in which demand is more elastic at higher demand levels. In fact, the introduction of RTP yields greater benefits in this case than the base case in which elasticity is the same in all periods. The reason is clear from looking at the equilibrium capacities. Elasticity in the peak periods is what drives the reduction in peaker capacity when customers move to RTP. This effect is larger when demand elasticity is greater in the peaks. So, having greater elasticity in peak periods means both greater demand response when there is more demand and a larger change in the equilibrium level of capacity, both of which contribute to a greater surplus gain from moving to RTP.

---

[14] BH show that the net externality from a *marginal* change in α is zero when demand in all periods has the same elasticity. There is a non-zero net externality in the cases shown here because the change is not incremental: Some of the externality of any one customer switching to RTP is captured by other customers in the switching group, so is not an externality from the group as a whole.

[15] As explained above, this is by assumption the actual CAISO load during each hour.

Table 5 presents the opposite case, in which demand is more elastic in low-demand periods than in high demand periods. The elasticity of demand varies linearly with demand level, in this case from 127% of the original demand elasticity for the lowest demand level to 50% of the original demand elasticity for the highest demand level. These boundaries were again chosen so that the demand-weighted average elasticity is equal to the original demand elasticity.

BH demonstrate that when elasticity is greater in low demand periods, the equilibrium flat rate will be above optimal and increasing the share of customers on RTP must necessarily increase total surplus. Nonetheless, the surplus gains in this case are smaller than in the base case, and much smaller than in the case in which demand is more elastic at peak times. The result follows intuitively after recognizing that inelastic demand during peak times means that RTP has less effect of reducing the amount of peaker capacity necessary to meet demand.

The Efficiency of RTP with Heterogeneous Customers

Throughout this analysis, I have assumed that all customers have identical demand patterns. Technically, this means that each customer's demand function is a fixed proportion of the aggregate demand function, $D_{hi}(p_h) = \gamma_i \cdot A_h \cdot p_h^{\epsilon_h}$.[16] One might ask how the results would change if customers differed in their demand patterns.

I do not carry out a complete exploration of this complex topic, but a few observations are useful. First, if the customers switching to RTP are chosen randomly from the population as a whole, and each customer is small relative to the aggregate demand, then the results presented here will apply. The aggregate wholesale demand will still be approximately $\tilde{D}_h(p_h, \bar{p}) = \alpha \cdot A_h \cdot p_h^{\epsilon_h} + (1 - \alpha) \cdot A_h \cdot \bar{p}^{\epsilon_h}$.

More interesting, however, is the recognition that the RTP adopters are likely to differ from the population on average in two important ways. First, they are likely to have

[16] Note that this means that the demand *function* is a fixed proportion of the aggregate demand *function*. Because different customers face different prices in a given hour – depending on whether they are on a fixed-rate tariff or RTP – this does not mean that a given customer will consume the same share of total system quantity in all hours.

demand profiles that, even absent any adjustment to RTP prices, are less peaky at high-demand times than the aggregate demand. These are the customers who cross-subsidize the peaky-demand customers when all are under a common flat-rate tariff. RTP gives them an opportunity to reduce or end this cross-subsidy. Second, the RTP adopters are likely to be more able to respond to high peak prices by reducing consumption, *i.e.*, to have demand that exhibits more price-elasticity in response to peak prices.

While this heterogeneity has obvious and important implications for the wealth transfers that RTP would effect, it also has potential implications for the efficiency of RTP. To the extent that the RTP adopters exhibit less peaky demand (but still the same demand *elasticity* in each hour as all other customers), this selection of customers moving to RTP would reduce the efficiency gains from the change. This is because the RTP adopters would in aggregate be a smaller proportion of total demand at peak times than at other times. The primary efficiency gains come from price-responsive demand reduction at peak times, so the potential for gains from such response is reduced if RTP adopters have relatively less demand at those times.

The fact that RTP adopters are likely to be more able to respond to high prices, however, will tend to improve the efficiency gains from RTP. If RTP adopters have the same peakiness in their demands as the system aggregate, analyzed for instance at the original flat-rate tariff, but have greater elasticity, then the gains from RTP would be greater than suggested by the previous calculations. The RTP adopters would simply have higher demand elasticity, so one would want to use a different row of the tables than if RTP adpters were representative of the overall demand elasticity of all customers.

IV. Is Time-of-Use Pricing a Good Substitute for RTP?

Though RTP has not been implemented in many electricity systems, the alternative assumption I've made thus far – that all customers are on flat rates – is also not accurate. In fact, in nearly all systems, prices for some customers vary over time, but in a pre-set manner. These "time-of-use" (TOU) pricing systems generally include peak/shoulder/off-prices that are set months in advance and are in effect for fixed hours of each week.

For example, Pacific Gas & Electric's basic TOU rate for small commercial customers in summer 2004 was 30.0¢/kWh during peak hours (noon-6pm on non-holiday weekdays), 13.9¢/kWh during shoulder hours (8am-noon and 6pm-11pm on non-holiday weekdays) and 8.7¢/kWh during off-peak (all other) hours. Thus, a worthwhile question to ask is how much of the welfare gains I've identified from RTP are captured with simple TOU pricing.

Unfortunately, while the allocation of costs in a flat-rate or an RTP system is straightforward, this is not necessarily the case in "middle" cases such as TOU or seasonally varying prices. To illustrate, consider a simple example with one L-shaped production technology and two TOU pricing periods. Assume that the utility must break even overall, exactly covering its fixed plus variable costs, and that it must build enough capacity to meet the highest quantity demanded. These seem like minimal constraints, but this problem in many cases still has no solution.

One approach is to allocate all of the fixed costs of capacity to the "peak" period, when the full capacity is used in at least one hour, and set the price during the off-peak period equal to the variable production cost. This approach is appealing because it mimics the outcome that would optain if the prices were equal to the weighted-average (competitive) wholesale price during a TOU period (assuming that the wholesale demand exhibited just the slightest bit of elasticity so prices were not indeterminate in the peak hour). Even in this case, however, things do not work out simply. If the "off-peak" period has even one high-demand hour, then with a sufficiently high retail demand elasticity and peak-period price, the highest quantity demanded hour for the system could occur during the off-peak period, making it effectively the peak period.[17] One solution is to constrain the prices so that the maximum quantity always occurs during the designated peak period, but this is just artificially constraining the peak/off-peak price difference to the level that (nearly)

[17] This is obviously related to the "shifting peaks" problem which was identified in the early peak-load pricing literature of the 1950s and then, with the help of a Lagrangian multiplier, solved. See, for instance, Steiner, 1957. The present problem, however, does not disappear so easily. A first-best solution cannot be implemented, because there is not a complete set of prices for every demand state.

equalizes demand in the highest demand peak-period hour and the highest-demand off-peak period hour.

Another approach is to allocate the fixed cost of a unit of capacity equally over all periods in which that unit is used. For example, the cost of capacity used at the minimum quantity time would be spread over all hours, because that capacity is used in all hours, while the cost of the last unit of capacity, used only at the maximum demand time, would be borne entirely by the consumers in that period. This greatly reduces (though does not eliminate) the peak-switching problem, but it also greatly dampens the price swing across TOU periods. It does have the populus appeal that only, and all, those who use a given unit of capacity pay for it.

I have tried three different approaches to constructing a TOU pricing scenario that could then be compared to the RTP and the flat rate scenarios. The first, which I call the "quasi wholesale" scenario attempts to mimic the weighted average competitive wholesale price with all capacity costs of the peaker capacity allocated to the hour in which quantity demanded is highest.[18] This has a solution for low demand elasticities, but does not have a solution if demand elasticity is too large. With the five years of California data that I am using, 'too large" is an elasticity greater in absolute value than 0.05.

The second approach is the "cost-share" scenario in which the allocation of capacity costs is determined by the number of hours in which a given unit of capacity is used during each of the TOU periods.[19] With the data I am using, this produces a solution for all elasticities up to 0.5 in absolute value.

[18] Baseload, mid-merit, and peaker capacities are set to minimize total production costs for a given peak, shoulder, and off-peak price, which determine quantity demanded in each hour. The competitive wholesale price for each hour is then calculated using those demand quantities (without elasticity). Then the TOU prices during each period are reset to be weighted-average wholesale prices during the period. This iteration continues until a fixed point is found.

[19] Baseload, mid-merit, and peaker capacities are set to minimize total production costs for a given peak, shoulder, and off-peak price, which determine quantity demanded in each hour. The allocation of the fixed capacity costs is then determined by the quantities demanded in each period and the levels of each type of capacity. Based on this allocation, the TOU prices during each period are reset to cover each period's variable costs plus share of capacity costs. This iteration continues until a fixed point is found.

The third approach is a "fixed-ratio" scenario in which the ratios of peak to shoulder and off-peak prices are set exogenously and then the prices and capacity are set in much the same way as in the flat-rate simulation described in the previous section. The price ratios were set to a level that reflects the average of the (fairly similar) pricing structures used by Pacific Gas & Electric and Southern California Edison, the two major utilities in California. This yielded a solution for elasticities up to 0.15 in absolute value.

For all three scenarios, all prices were allowed to vary between winter and summer as well. In particular, similar to the utilities' actual TOU rate structures, there were two prices in the winter, a peak price that was in effect 8am-9pm on non-holiday weekdays, and an off-peak price that was in effect on at all other times in the winter. In the summer, there were three TOU periods: Peak was noon-6pm on non-holiday weekdays; Shoulder was 8am-noon and 6pm-11pm on non-holiday weekdays; Off-peak was in effect at all other times.[20] Summer was defined as June-September and winter was defined as October-May.

In the simulations I present, the prices change by season, but not year-to-year. The summer peak price, for instance, is the same in all years. This is meant to reflect the fact that the year-to-year variation during this period is mostly not predictable growth, but idiosyncratic weather variation that would not be predictable at the time that the TOU prices were set for each time period.

The welfare results of these simulations are presented in table 6, with the figures for RTP also presented for comparison. The conclusion is clear: TOU rates capture a small share of the benefits that would be obtained from RTP. Even the most efficient form of TOU ("quasi-wholesale"), which generates peak to off-peak price ratios well above those observed in actual TOU programs, captures only one-quarter or less of the RTP gains for those elasticities for which it is feasible. Using actual fixed-ratios of prices, the gains also seem to get up to about one-quarter of RTP before those price ratios become infeasible at higher elasticities.

---

[20] For the fixed-ratio scenario, all prices were fixed as a proportion of the summer peak price. The proportions were summer/shoulder 57.4%, summer/off-peak 45.0%, winter/peak 61.9%, and winter/off-peak 47.7%.

I should note, however, that there is a critical assumption in these calculations, that elasticity of demand in responding to long-run TOU prices is the same as the elasticity in response to RTP prices. Put differently, one can think of RTP prices as decomposable into different averages for TOU-like periods and deviations from those averages in any given hour. The underlying assumption is that customers would be equally responsive to the variations in averages as to the deviations from those averages in a particular hour.

In reality, elasticity with respect to short-term fluctuations could be lower or higher than with respect to longer-term predictable average price differences. One could argue that the short-term less-predictable deviations are more difficult to respond to because of the lack of advanced notice. For instance, companies could not reschedule work shifts based on a price spike that becomes apparent only hours before it actually occurs. In the extreme, if the only electricity-consumption modifications that a customer could make would be the result of months-ahead planning, then RTP offers a much smaller advantage over TOU.[21] The elasticity of demand with respect to deviations from months-ahead expected price for a given hour would be virtually zero.

On the other hand, there may be short-duration adjustments that a firm could do to respond to a price spike that they could not maintain for a longer period. For instance, if a company knew that a heat spell is driving prices to very high levels today, but will likely break by tomorrow, it could possibly shift some electricity-intensive activity to tomorrow. The potential for these sort of short-term adjustments suggests that the elasticity could be greater for short-term deviations than for long-term average price differences.

The relatively small efficiency gains from TOU pricing are quite intuitive when one recalls that the inefficiency from non-optimal pricing in a given hour goes up in proportion to the square of the deviation of price from marginal cost. Thus, the most costly "mistakes" occur during the times when prices deviate most from the mean during a given TOU period. Intuitively, then it would be more effective to attain a given average price within a certain

[21] RTP would still offer better granularity of prices, as the 3-4pm expected price for a day six months hence would differ at least slightly from the 2-3pm expected price for that same day.

TOU period by having very high retail prices during the few hours with highest wholesale prices and slightly lower retail prices at all other times, thereby substantially mitigating the largest pricing mistakes rather than addressing slightly a larger number of small price mistakes. A program that roughly takes this approach, called "Critical Peak Pricing" is currently being tested in California and elsewhere. My preliminary analysis suggest CPP could capture a much greater share of the RTP efficiency gains than could TOU.

Still, the TOU results do make clear that the gains I've claimed from RTP in the previous section were slightly overstated. The baseline from which most systems begin is with 50% or more of total demand on TOU, including most customers that would be initially put on RTP if only a share of customers were moved to RTP. Thus, the gains from moving these customers to RTP should be scaled down by between 15% and 25% (using the assumption that elasticity of demand is the same for longer-term changes as for shorter-term price varitions).

V. Limitations of the RTP Simulation Model

Though these simulations are useful in giving an idea of the potential gains from RTP, they don't take into account all aspects of electricity markets. Incorporating many of these characteristics will be challenging, but it is clear even without that additional analysis that these simulations are likely to understate the benefits of RTP.

The most important area of omission is the stochastic elements of supply and demand. The model does not incorporate the unpredictability of demand or the probabilistic outages of generation supply. Currently, responses to these stochastic elements of the supply/demand balance are addressed almost entirely with supply adjustment. Unless, short-run demand adjustment is impossible, which there is increasing evidence is not the case, responding entirely on the supply side is clearly not the most efficient way to address such outcomes.

Including RTP in system balancing will further enhance system efficiency. It seems almost certain that RTP would decrease system peak loads, so using standard proportional reserve rules, it would reduce the amount of reserve capacity needed and the payments for that capacity. More importantly, RTP would increase the responsiveness of demand to system stress and thus would reduce the level of reserves needed for any given level of demand. In economic terms, RTP would not just shift demand to the left at peak times, it would make demand more price elastic, so more balancing could be accomplished with less supply-side adjustment. Likewise, incorporating generator outages raises the benefits of demand responsiveness by reducing the need to compensate for a generator outage completely on the supply side.

Assuming competitive supply, an upper bound on the "reserves cost" savings from RTP is the total cost of reserve payments. In most systems, operating reserves average 5-10% of energy costs. Planning reserves costs may be covered by energy and operating reserve payments, or they may require additional payments, which would also be subject to reduction through use of RTP. RTP is likely to reduce these costs by a significant amount, but much of these costs will remain for a long time. Nonetheless, the benefits from RTP are likely to be underestimated from the simulations presented, because they do not incorporate the benefits from reduced need for reserves.

Closely related to reserves costs are the effect of non-convexities in operation of plants and lumpiness in the size of plants. As discussed in detail by Mansur (2003), generation units do not costlessly or instantly switch from off to full production. There are start-up costs and "ramping" constraints (on the speed with which output can be adjusted). These constraints make it more costly to adjust supply to meet demand fluctuations. As with reserves, RTP would allow some of this adjustment to occur on the demand side in a way that would enhance efficiency. Similarly, I have assumed the plants can be scaled to any size at the same long-run average cost. If this were not the case, then there would be greater mismatches between demand and the capital stock. In conventional electricity systems, these mismatches have been handled by over-building and then either selling excess production on the wholesale market or leaving excess capacity idle. Having the additional option of demand-side adjustment could only lower long-run costs.

The simulations also have ignored market power issues, instead assuming that free entry would bring a completely competitive market over the longer run. As has been discussed elsewhere,[22] demand elasticity introduced by implementing RTP reduces the incentive of sellers to exercise market power. However, it is unclear how much incremental inefficiency the exercise of market power itself introduces in a flat-rate system, since it simply changes the flat retail rate that is charged in all time periods. In fact, Borenstein and Holland's analysis (forthcoming) suggests that if the equilibrium flat rate is less than the surplus-maximizing flat rate, $\bar{p}^e < \bar{p}^*$, seller market power could increase efficiency. In a full RTP system, market power could not increase efficiency. Thus, it is difficult to analyze the bias from excluding seller market power.

The demand system I've analyzed departs from reality by assuming all cross-elasticities are zero. Simulation with a complete matrix of own- and cross-elasticities would increase the complexity substantially. Still, if demands are generally substitutes across hour, it seems very likely that incorporation of cross-elasticities would increase the gains from RTP. Essentially, RTP increases efficiency by reducing the volatility of quantity consumed and increasing the utilization rate of installed capacity. Holding constant own-price elasticities, increasing cross-price elasticities from zero to positive (substitutes) will tend to further reduce quantity volatility by increasing off-peak quantity when peak prices rise and reducing peak quantity when off-peak prices fall.

Finally, the simulations take a constant \$40/MWh charge for transmission and distribution (T&D). This is based on the historical recovery of the costs of these services, which are provided by a regulated monopoly. To the extent that minimum efficient capacity scale for T&D implies that they are never capacity constrained, introducing time-varying prices of these services would not improve efficiency. That may be the case with most local distribution, but transmission lines frequently face capacity constraints. By ignoring these constraints and holding the T&D cost per MWh constant, the simulations understate the potential gains for RTP that could also reflect time-varying (opportunity) cost of transmission, which are already reflected to varying degrees in wholesale electricity markets.

[22] See Borenstein and Bushnell (1999) and Bushnell (forthcoming).

VI. Conclusions

Real-time electricity pricing has tremendous appeal to economists on a theoretical level, because it has the potential to improve welfare by giving customers efficient consumption incentives. The theoretical analysis, however, does not indicate how large the gains from RTP are likely to be. With a simple simulation exercise, I have tried to generate some numbers to go with the theory. This is obviously just a first cut, but the results suggest a number of likely findings:

– The benefits of RTP are likely to far outweigh the costs for the largest customers.

– The incremental benefits of putting more customers on RTP are likely to decline as the share of demand on RTP grows. At the same time, the costs of increasing the share of demand on RTP increases as the size of each customer declines. Thus, while there seems to be clear net social value from putting larger customers on RTP, the additional gains from putting smaller customers on RTP may not justify the cost. A factor weighing against this conclusion is that small customers are thought by many electricity analysts to be the most price responsive. If that is true, then the argument for RTP metering of them is, of course, strengthened. Further analysis of both the costs and benefits is needed.

– Time-of-use rates are a very poor substitute for RTP. Roughly speaking, TOU rates capture only 20% of the efficiencies of RTP, though this finding has the caveat that it assumes as high an elasticity for response to short-run price variation as long-run differences in average prices.

The findings of this study must be viewed as preliminary. A number of factors have not been addressed in the analysis thus far, though incorporating them seems likely to lead to larger estimated gains from RTP. Incorporation of these factors into the analysis is not particularly complex. A larger barrier is likely to be the data necessary to permit reliable estimates of demand elasticities and supply flexibility, which I've shown have very large impacts on the efficiency gains.

Finally, it is worth pointing out that RTP is being adopted in a number of places in the U.S. and elsewhere. The programs are relatively young (up to 15 years old), but there are already a number of examples of programs with which both the utilities and the customers are quite happy, and that have documented both peak-demand reductions and reduced need for peeking capacity. For a very thorough description of voluntary dynamic pricing programs in the U.S., see Barbose, Goldman and Neenan (2004). The large RTP programs operated by Georgia Power, Gulf Power and Niagra Mohawk should be of great interest to those evaluating the efficacy of RTP.

REFERENCES

Barbose, Galen, Charles Goldman and Bernie Neenan. "A Survey of Utility Experience with Real-Time Pricing," Lawrence Berkeley National Laboratory working paper LBNL-54238, December 2004. Available at http://eetd.lbl.gov/ea/ems/reports/54238.pdf.

Borenstein, Severin. "Time-Varying Retail Electricity Prices: Theory and Practice," in Griffin and Puller, eds., *Electricity Deregulation: Choices and Challenges*, Chicago: University of Chicago Press, forthcoming.

_________ and James B. Bushnell, "An Empirical Analysis of Market Power in a Deregulated California Electricity Market, *Journal of Industrial Economics*, September 1999, **47**(3).

Borenstein, Severin and Stephen P. Holland. "Investment Efficiency in Competitive Electricity Markets With and Without Time-Varying Retail Prices," CSEM Working Paper CSEMWP-106, University of California Energy Institute, revised July 2003(a). Available at http://www.ucei.org/PDF/csemwp106.pdf.

Borenstein, Severin and Stephen P. Holland. "On the Efficiency of Competitive Electricity Markets With Time-Invariant Retail Prices," forthcoming, *RAND Journal of Economics.*

Bushnell, James B. "Looking for Trouble: Competition Policy in the U.S. Electricity Industry," in Griffin and Puller, eds., *Electricity Deregulation: Choices and Challenges*, Chicago: University of Chicago Press, forthcoming.

Braithwait, Steven D. and Michael O'Sheasy. 2002. "RTP Customer Demand Response – Empirical Evidence on How Much Can You Expect," in Faruqui and Eakin eds. *Electricity Pricing in Transition*, Boston, MA: Kluwer Academic Publishers.

Patrick, Robert H. and Frank A. Wolak. 1997. "Estimating the Customer-Level Demand for Electricity Under Real-Time Market Prices," Stanford University working paper, available at ftp://zia.stanford.edu/pub/papers/rtppap.pdf.

Ruff Larry E. 2002. "Demand Response: Reality versus 'Resource' " *The Electricity Journal*, December, pp. 10-24.

Steiner, Peter O. "Peak Loads and Efficient Pricing." *Quarterly Journal of Economics*, Vol. 72 (1957), pp. 585-610.

TABLE 2 -- Capacity, Price and Quantity Effects of RTP

A	B	C	D	E	F	G	H	I	J	K	L	M
Elas-	Share on	Total Annual	Total Annual	Flat	EQUILIBRIUM CAPACITY (MW)				PRICE DURATION CURVE			
ticity	RTP	Energy	Energy	Rate	Base-	Mid-		Total	Peak	Avg Hrs	Pctg	Pctg
		Consumed	Bill	($/MWh)	Load	Merit	Peaker		Price	per year	of annual	of annual
		(MWh million)	($ million)						($/MWh)	at Peak	bill from	bill from
All On Flat Rate										Quantity	top 10	top hour
---	0.000	231,095,835	9,170,521,267	79.68	26984	5384	14560	46928		(of 8760)	hours	in sample
											in sample	
Some On RTP												
-0.025	0.333	231,405,274	9,048,736,469	79.65	27028	5341	12038	44407	90772	4	60.8%	22.0%
-0.025	0.666	231,691,153	8,945,358,991	79.47	27074	5258	10014	42346	45292	30	44.0%	10.5%
-0.025	0.999	231,933,022	8,871,844,654	79.22	27118	5184	8603	40905	19505	67	23.5%	4.4%
-0.050	0.333	231,711,476	8,958,308,425	79.52	27075	5258	10251	42584	54052	25	48.3%	12.6%
-0.050	0.666	232,212,872	8,826,608,560	79.08	27169	5113	7732	40014	11890	97	15.5%	2.6%
-0.050	0.999	232,625,430	8,739,797,966	78.84	27256	4974	6176	38406	4405	157	6.6%	0.9%
-0.100	0.333	232,326,272	8,848,470,458	79.18	27178	5116	8074	40368	18834	84	21.6%	4.2%
-0.100	0.666	233,214,051	8,689,157,551	78.73	27361	4837	5211	37409	3038	206	4.7%	0.6%
-0.100	0.999	233,932,035	8,572,879,071	78.48	27531	4556	3364	35451	1321	348	2.2%	0.3%
-0.150	0.333	232,953,671	8,780,950,177	78.97	27284	4978	6733	38995	9302	132	11.8%	2.0%
-0.150	0.666	234,209,237	8,594,016,415	78.53	27554	4558	3568	35680	1577	328	2.5%	0.3%
-0.150	0.999	235,202,138	8,455,153,100	78.24	27799	4154	1573	33526	752	556	1.3%	0.1%
-0.300	0.333	234,955,611	8,659,285,409	78.68	27612	4564	4266	36442	3505	264	4.8%	0.7%
-0.300	0.666	237,327,726	8,409,265,790	78.12	28133	3759	547	32439	647	682	1.0%	0.1%
-0.300	0.999	238,825,409	8,238,485,575	77.59	28606	1786	0	30392	340	1891	0.6%	0.1%
-0.500	0.333	237,926,466	8,576,661,386	78.47	28062	4026	2361	34449	2302	438	3.1%	0.4%
-0.500	0.666	241,571,384	8,289,753,486	77.58	28942	1445	0	30387	370	2006	0.6%	0.1%
-0.500	0.999	243,110,229	8,139,900,836	76.73	28986	0	0	28986	209	5498	0.3%	0.0%

TABLE 3 -- Welfare Effects of RTP

A	B	C	D	E	F	G	H	I	J
Elas-ticity	Share on RTP	Annual Total Surplus Change from All on Flat ($)	Annual TS Change as percentage of original energy bill	Annual CS Change of Customers on Flat Rate ($)	Annual CS change "per customer" on Flat Rate ($)	Annual CS Change of Customers on RTP ($)	Annual CS change "per customer" on RTP ($)	Annual Incremental Surplus to Switchers ($)	Annual Incremental Externality ($)
-0.025	0.333	112,060,365	1.2%	4,602,394	69	107,457,971	3,227	107,457,971	4,602,394
-0.025	0.666	205,800,109	2.2%	16,195,248	485	189,604,862	2,847	92,504,684	1,235,061
-0.025	0.999	271,333,946	3.0%	107,052	1,071	271,226,894	2,715	74,262,205	-8,728,369
-0.050	0.333	196,836,537	2.1%	24,879,553	373	171,956,984	5,164	171,956,984	24,879,553
-0.050	0.666	314,219,558	3.4%	46,572,214	1,394	267,647,344	4,019	121,402,546	-4,019,525
-0.050	0.999	388,316,857	4.2%	194,639	1,946	388,122,219	3,885	82,941,297	-8,843,997
-0.100	0.333	302,262,176	3.3%	77,399,306	1,160	224,862,870	6,753	224,862,870	77,399,306
-0.100	0.666	439,987,363	4.8%	73,366,291	2,197	366,621,072	5,505	144,668,903	-6,943,716
-0.100	0.999	537,284,137	5.9%	276,546	2,765	537,007,592	5,375	105,855,899	-8,559,124
-0.150	0.333	370,238,483	4.0%	108,757,099	1,631	261,481,384	7,852	261,481,384	108,757,099
-0.150	0.666	530,960,593	5.8%	89,145,379	2,669	441,815,214	6,634	166,610,585	-5,888,475
-0.150	0.999	647,620,518	7.1%	333,189	3,332	647,287,329	6,479	126,883,966	-10,224,041
-0.300	0.333	509,388,631	5.6%	154,467,302	2,316	354,921,329	10,658	354,921,329	154,467,302
-0.300	0.666	730,577,275	8.0%	120,644,221	3,612	609,933,053	9,158	227,848,668	-6,660,025
-0.300	0.999	888,877,347	9.7%	484,978	4,850	888,392,369	8,893	175,847,779	-17,547,706
-0.500	0.333	641,472,723	7.0%	187,262,169	2,808	454,210,554	13,640	454,210,554	187,262,169
-0.500	0.666	922,328,312	10.1%	162,892,786	4,877	759,435,525	11,403	286,227,054	-5,371,466
-0.500	0.999	1,098,811,460	12.0%	687,144	6,871	1,098,124,316	10,992	203,636,356	-27,153,207

Table 4 - Larger elasticity with higher demand

A	B	C	D	E	F	G	H	I	J	K	L	M
Elas-ticity	Share on RTP	Total Surplus Change from All on Flat	CS Change of Customers on Flat Rate	CS Change of Customers on RTP	Total Energy Consumed (MWh)	Total Energy Bill ($)	TS Chg as pctg of orig energy bill	Flat Rate ($/MWh)	CAPACITY Base-Load	Mid-Merit	Peaker	Total
All On Flat Rate												
---	0.000				231,095,835	9,170,521,267		79.68	26984	5384	14560	46928
Some On RTP												
-0.025	0.333	185,012,684	24,753,508	160,259,175	231,299,627	8,960,830,174	2.0%	79.52	27028	5320	10396	42744
-0.025	0.666	290,445,187	45,902,053	244,543,133	231,410,802	8,834,577,684	3.2%	79.09	27073	5246	8031	40350
-0.025	0.999	354,501,597	190,899	354,310,698	231,466,667	8,752,105,413	3.9%	78.85	27116	5167	6611	38894
-0.050	0.333	279,563,402	77,004,082	202,559,320	231,471,033	8,851,986,763	3.0%	79.18	27077	5248	8324	40649
-0.050	0.666	397,074,357	71,576,380	325,497,977	231,606,958	8,700,275,606	4.3%	78.75	27167	5093	5720	37980
-0.050	0.999	476,254,443	266,271	475,988,172	231,655,045	8,591,255,016	5.2%	78.53	27251	4946	4036	36233
-0.100	0.333	381,877,312	126,459,616	255,417,696	231,802,179	8,731,108,942	4.2%	78.86	27181	5105	6141	38427
-0.100	0.666	528,647,468	97,383,322	431,264,146	231,966,565	8,530,420,782	5.8%	78.42	27357	4798	3087	35242
-0.100	0.999	631,606,056	359,894	631,246,162	231,970,742	8,382,906,188	6.9%	78.12	27520	4497	1241	33258
-0.150	0.333	450,449,991	151,066,725	299,383,266	232,154,887	8,651,179,570	4.9%	78.70	27288	4958	4768	37014
-0.150	0.666	621,625,868	115,027,545	506,598,324	232,347,230	8,412,227,233	6.8%	78.19	27548	4510	1472	33530
-0.150	0.999	742,093,305	430,021	741,663,284	232,300,973	8,239,591,317	8.1%	77.82	27782	3716	0	31498
-0.300	0.333	588,538,248	191,696,925	396,841,323	233,366,127	8,503,667,355	6.4%	78.44	27616	4534	2329	34479
-0.300	0.666	812,804,128	157,464,303	655,339,825	233,634,657	8,196,068,488	8.9%	77.65	28123	2673	0	30796
-0.300	0.999	951,596,870	603,932	950,992,938	233,361,253	8,008,016,774	10.4%	77.08	28560	603	0	29163
-0.500	0.333	713,407,144	215,805,278	497,601,865	235,355,289	8,401,517,798	7.8%	78.29	28058	3994	577	32629
-0.500	0.666	971,233,473	198,634,806	772,598,667	235,752,041	8,063,581,334	10.6%	77.13	28911	343	0	29254
-0.500	0.999	1,112,675,961	788,031	1,111,887,930	235,191,765	7,915,433,819	12.1%	76.31	28304	0	0	28304

Table 5 - Smaller elasticity with higher demand

A	B	C	D	E	F	G	H	I	J	K	L	M
Elas-ticity	Share on RTP	Total Surplus Change from All on Flat	CS Change of Customers on Flat Rate	CS Change of Customers on RTP	Total Energy Consumed (MWh)	Total Energy Bill ($)	TS Chg as pctg of orig energy bill	Flat Rate ($/MWh)	CAPACITY Base-Load	Mid-Merit	Peaker	Total
All On Flat Rate												
---	0.000				231,095,835	9,170,521,267		79.68	26984	5384	14560	46928
Some On RTP												
-0.025	0.333	60,973,382	0	60,973,382	231,456,637	9,109,076,008	0.7%	79.68	27028	5359	13169	45556
-0.025	0.666	121,784,247	2,460,451	119,323,796	231,815,267	9,047,779,004	1.3%	79.65	27074	5268	11952	44294
-0.025	0.999	177,024,608	22,150	177,002,458	232,164,082	8,992,105,924	1.9%	79.58	27119	5192	10819	43130
-0.050	0.333	118,437,413	4,494,537	113,942,876	231,821,793	9,052,836,335	1.3%	79.65	27074	5268	12051	44393
-0.050	0.666	220,414,820	14,073,842	206,340,978	232,522,183	8,950,840,222	2.4%	79.50	27168	5124	9953	42245
-0.050	0.999	297,190,391	98,205	297,092,186	233,161,482	8,873,523,930	3.2%	79.26	27259	4993	8381	40633
-0.100	0.333	210,667,653	21,371,969	189,295,684	232,564,028	8,965,285,145	2.3%	79.54	27171	5121	10221	42513
-0.100	0.666	349,596,306	43,941,433	305,654,873	233,890,389	8,828,532,614	3.8%	79.11	27360	4859	7407	39626
-0.100	0.999	444,180,558	193,792	443,986,766	235,055,299	8,734,565,308	4.8%	78.84	27536	4592	5607	37735
-0.150	0.333	280,447,527	47,059,933	233,387,593	233,332,159	8,903,748,804	3.1%	79.37	27273	4989	8874	41136
-0.150	0.666	439,799,926	62,288,190	377,511,736	235,253,128	8,749,697,717	4.8%	78.87	27554	4587	5783	37924
-0.150	0.999	552,454,609	246,286	552,208,323	236,918,971	8,639,981,640	6.0%	78.62	27808	4200	3771	35779
-0.300	0.333	423,048,000	103,852,046	319,195,955	235,774,058	8,796,788,098	4.6%	79.01	27596	4583	6365	38544
-0.300	0.666	638,364,596	92,432,780	545,931,816	239,438,461	8,602,426,632	7.0%	78.49	28135	3806	2652	34593
-0.300	0.999	801,790,577	352,184	801,438,393	242,528,799	8,454,748,414	8.7%	78.16	28629	3056	274	31959
-0.500	0.333	557,558,695	141,926,277	415,632,419	239,327,549	8,728,896,973	6.1%	78.76	28041	4044	4376	36461
-0.500	0.666	842,778,837	118,889,088	723,889,749	245,404,704	8,497,863,423	9.2%	78.15	28938	2778	101	31817
-0.500	0.999	1,049,404,926	527,912	1,048,877,014	249,300,982	8,341,841,585	11.4%	77.41	29690	0	0	29690

TABLE 6 -- Welfare Effects of RTP vs TOU pricing

A	B		C	D	E	F
Elas-	Share on		ANNUAL TOTAL SURPLUS CHANGE VS FLAT RATE			
ticity	RTP/TOU			"Quasi-wholesale"	Actual TOU	"Cost-share"
			RTP	TOU	price ratios	TOU
-0.025	0.333		112,060,365	16,269,127	10,657,394	6,928,165
-0.025	0.666		205,800,109	32,538,254	21,314,789	13,856,330
-0.025	0.999		271,333,946	48,807,381	31,972,183	20,784,495
-0.050	0.333		196,836,537	32,226,253	21,322,177	13,683,652
-0.050	0.666		314,219,558	64,452,506	42,644,355	27,367,305
-0.050	0.999		388,316,857	96,678,759	63,966,532	41,050,957
-0.100	0.333		302,262,176	N/A	42,006,103	26,159,344
-0.100	0.666		439,987,363	N/A	84,012,206	52,318,689
-0.100	0.999		537,284,137	N/A	126,018,309	78,478,033
-0.150	0.333		370,238,483	N/A	61,775,434	37,387,646
-0.150	0.666		530,960,593	N/A	123,550,868	74,775,291
-0.150	0.999		647,620,518	N/A	185,326,302	112,162,937
-0.300	0.333		509,388,631	N/A	N/A	65,167,555
-0.300	0.666		730,577,275	N/A	N/A	130,335,110
-0.300	0.999		888,877,347	N/A	N/A	195,502,666
-0.500	0.333		641,472,723	N/A	N/A	92,710,676
-0.500	0.666		922,328,312	N/A	N/A	185,421,352
-0.500	0.999		1,098,811,460	N/A	N/A	278,132,028

FIGURE 1: Load Duration Curve with Varying Demand Elasticities of RTP Customers
(1/3 of total demand on RTP, 2/3 on flat-rate tariff)