Priority Pricing for Clean Power under Uncertainty

Hung-po Chao, Shmuel Oren and Robert Wilson[1]

Abstract

Purpose of Review Climate change has become a defining issue on a global scale. With declining costs, renewables have expanded rapidly to become the preferred clean energy resources for electricity generation. With salient features of zero marginal cost and supply uncertainty, renewables present unique challenges to the electricity market design and price formation for long-run efficiency. The purpose of this paper is to review the impact of metering technology on demand-side management and price formation. We compared three cases, integrated resource planning, ex-ante linear pricing, and priority service pricing, with different levels of metering technology. The cases are clarified with numerical examples and the results are compared in terms of capacity level, capacity cost, service reliability, and social welfare.

Recent Findings Traditionally, the electric utility industry has relied on integrated resource planning and central system operations to meet the growing demand and maintain a standard level of service reliability. During the past two decades, the liberalization of electricity markets has transformed the traditional industry practice with the introduction of market-based locational marginal pricing for improved economic efficiency. As evidenced by recent research, a growing consensus has emerged that the future transition to renewables with zero marginal cost and supply uncertainty would present fundamental challenges to the current market designs.

Summary In this paper, we examine demand-side management with priority service addressing price formation and financial viability of merchant investments from a risk perspective to guide practical pricing and investments policy decisions in the presence of supply uncertainty. To illuminate the issue, we study a few highly simplified cases for a clean power economy in the environment of a large remote island running solely on renewable energy sources, e.g., solar and winds. To meet the local electricity needs, we address several key issues including: How would the system operator keep the lights on when demand fluctuates continuously over days, weeks and seasons, but supply is unpredictable and difficult to control? How should the price of electricity be set when the short-run marginal cost is zero? How would the market attract investments to meet the growing demand in the long term? What would be the impacts on the service reliability? To address practical implementation issues, we discuss a stochastic auction-based market platform that enables innovative demand-side management harnessing flexibility via an end-to-end business model, in ways that flexible demand devices (e.g., hot water heaters, air conditioners, energy storage etc.) on the customer end are aggregated into a "virtual power plant" submitted by an aggregator through a supply function offer into the wholesale market as demand reserve.

Keywords Efficient price formation, metering technology, demand-side management, priority service, financial risk management, integrated planning, clean energy

[1] Corresponding co-author: Hung-po Chao at Hungpo.chao@et-analytics.com.

1. Introduction

Electricity is essential for the modern information economy. During the past decade, with accelerating momentum, climate change has become a defining issue on a global scale.[2] With declining costs, renewables have expanded rapidly, increasing from virtually nothing to about 30% of electricity demand in the U.S. As the costs of winds and solar have declined rapidly to a level rivaling fossil fuels, the resource mix for electricity generation is shifting toward renewables. Renewables have become the preferred clean energy sources for electricity generation surpassing coal for the first time in April 2019. (EIA, 2019)

Over the past two decades, market liberalization has led to competitive markets with locational marginal pricing, complementing the traditional centralized system planning and operational control process, which has been the standard industry practice for electric utilities over the past century. (EPRI 2004; Chao, Oren, and Wilson) Competitive markets have produced low price outcomes, attracted new investments, stimulated innovations and lowered the entry barrier for new technologies such as renewable energy. Nonetheless, the high penetration of renewable energy resources, with distinctive attributes of supply uncertainty and zero marginal cost, presents unique challenges for efficient pricing mechanism at a fundamental level. (Joskow, 2019; Hogan, 2021)

To address the challenges, we review the literature on the peak-load pricing that lays the welfare foundation for efficient price formation and address financial viability of merchant investments from a risk management perspective.[3] In the landmark paper, Boiteux (1949) addresses the basic conundrum for an electric utility: to meet the peak demand requires such capacity investments that would result in excess capacity sitting idle and under-utilized during the off-peak periods. Peak-load pricing charges customers based on the marginal cost of production which is higher during peak hours and lower during off-peak hours. The early literature linked peak-load pricing to the principle of marginal cost pricing (which was under active debate at the time across Europe and the U.S.) emphasizing the advantages of fair cost allocation, revenue sufficiency and economic efficiency. This theory of spot pricing based on marginal cost has played a significant role in the development of wholesale electricity markets in the U.S. However, one of the most significant limitations of spot pricing is the lack of a proper recognition of supply uncertainty.[4]

[2] According to a 2013 UN IPCC report the average global temperature increased by 0.85 °C from 1880 to 2012.
[3] Peak-load pricing refers to the pricing strategy for economically non-storable commodities whose demand fluctuates cyclically. For a literature survey, see Crew, Fernando & Kleindorfer (1995).
[4] For example, Crew, Fernando & Kleindorfer (1995) commented that "a significant drawback of this existing literature on real-time pricing is that it (almost by definition) has no provision for quantity rationing, assuming instead that the market would clear through price adjustment under all contingencies."

In a seminal paper, Weitzman (1974) exposited that in the presence of supply uncertainty, either price or quantity rationing alone is no more than a "second-best" instrument and the optimal choice between the two depends on the shapes of the cost and benefit functions. At a fundamental level, supply uncertainty is closely tied to reliability risk management in ways that are asymmetric with demand uncertainty. In the history of the power industry, large system outages are rarely not directly connected to uncertain supply failures in ways that inflict the complex externalities for a public good. In the absence of supply uncertainty, reliability risk management would be much simpler and more compatible with the conventional theory of spot pricing.

Undeniably, the theory of marginal cost pricing is simple, elegant and powerful in providing a first-best solution. But by ignoring the essential elements of supply uncertainty and the need for quantity rationing, this branch of the peak-load pricing literature has relatively little to offer on reliability risk management in electricity markets. Under conditions of zero marginal cost and supply uncertainty, any pricing mechanism has to be ex ante and quantity rationing could not be ignored. We believe that supply uncertainty is one of the most significant challenges for the literature of peak-load pricing for guiding pricing and investments policy decisions in ways that would enhance efficient price formation and assure revenue sufficiency. A key challenge to guide practical pricing and investments policy decisions is to recognize the need for quantity rationing in the presence of supply uncertainty.

Chao (1983) introduced the basic ex ante pricing framework for the peak-load pricing problem under demand and supply uncertainty and it was generalized by Crew, Kleindorfer and Fernando (1993). Chao and Wilson (1987) introduced the basic nonlinear pricing framework for priority services, an incentive mechanism that allows consumer to self-select service reliability options and it was generalized by Chao, Oren, Smith and Wilson (1988), Wilson (1989), Wilson (2002) and Chao (2012). Priority service pricing facilitates demand-side management offering a menu of contingent contracts for distribution of uncertain supplies. Priory service prices are expectations of ex post spot prices for comparable services. Customers' selections reveal the benefit of capacity expansion. By providing a general framework for price and quantity rationing, priority service pricing is Pareto superior to ex ante pricing. Priority service pricing can be implemented under alternative market organizations, e.g., in retail markets via buying compensatory insurance or in wholesale markets via selling demand reserves. Under conditions of supply uncertainty and zero marginal cost, priority service pricing achieves first-best ex post efficiency and assures revenue sufficiency for merchant investments. Recent developments driven by the rapid penetration of renewable resources, storage and edge technologies brought about new opportunities on the regulatory side culminating with FERC Order 2222 (2020) that aims at facilitating the participation of demand management and distributed energy resources in the wholesale market. This opened the door for new implementation approaches for priority services through wholesale market participation of demand-backed resources participating in reserve market.

To illuminate the challenge, we examine a few highly simplified cases for a clean power economy. Imagine a large remote island that runs solely on renewable energy sources, e.g., solar and winds, to meet the local electricity needs. Some of the obvious challenges that the island faces are: How would the system operator keep the lights on when demand fluctuates continuously over days, weeks and seasons, but supply is unpredictable and difficult to control? How should the price of electricity be set when the short-run marginal cost is zero? How would the market attract investments needed to meet the growing demand in the long term? What would be the impacts on the service reliability?

To address practical implementation issues, we discuss a stochastic market auction platform that enables innovative demand-side management with priority service via an end-to-end business model, in ways that flexible demand devices (e.g., hot water heaters, air conditioners, energy storage etc.) on the customer end are aggregated into a "virtual power plant," and an aggregator for distributed resources addresses the financial risks for curtailed energy on the supply side. In essence, priority service is implemented through curtailment options that can be employed by the service provider on the demand side. Then, the aggregator offers the virtual power plant through a supply function into the wholesale market as demand reserve. Building on the literature before the restructuring of wholesale electricity markets, an early developments of curtailable service contracts as real options or callable forward contracts and the pricing strategy of such options was discussed by Gedra and Varaiya (1993) and Kamat and Oren (2002) in the context of vertically integrated utilities to improve energy efficiency through product differentiation in retail electricity supply. Recent developments on the regulatory side culminating with FERC Order 2222 (2020) open the door for new implementation approaches for virtual power plant based on statistically verified performance.

In this paper, we examine three cases with evolving metering technology, 1) integrated resource planning with no meters, 2) ex ante pricing with traditional time-of-use meters and 3) priority service pricing with smart meters/contracts. Historically, the evolution of metering technology affects the development of pricing mechanism and demand-side management in electricity markets. In 1881, Thomas Edison's DC electric meter was patented. Although it was in use until the end of the 19th century, the DC meter has some severe limitations because meter reading was a difficult task for the utility and an impossible one for the customer. As a result, Edison was later forced to charge for electricity based on counting the number of lamps that each customer uses. Near the end of the 19th century, the AC induction meter was invented and has become the standard metering technology which is still in use today. Over time, the induction meter has improved significantly and now it features time-of-use, the maximum demand, and remotely controlled switching capability. In the 1970s, with the advances in electronics, the manufactures started to introduce electronic registers and automatic meter reading devices. Smart electronic, digital meters were born during this period after the introduction of the integrated circuit. The modern internet technology has given a new impetus to the development of smart meters with increasing flexibility and functionality. Alternative metering & verification methodologies would become an important issue.

The remaining sections of the paper are organized as follows: Section 2 presents the basic model formulation and assumptions. Section 3 examines the case of integrated resource planning with no meters. Section 4 examines the case of ex ante pricing with traditional time-of-use meters allowing partially responsive demand-side management. Section 5 examines the case of nonlinear pricing of priority service options with smart meters that allow fully responsive demand-side management. Section 6 addresses implementation and financial risk management issues. Section 7 provides a summary of findings.

2. Basic model structure and assumptions

In this section, we describe the basic model structure and assumptions for the cases that will be examined in subsequent sections. We adopt the framework of social welfare economics focusing on reliability, economic and financial risks. Figure 1 shows the time sequence of events in a decision process in which each square is a decision node and each circle is an uncertainty node.

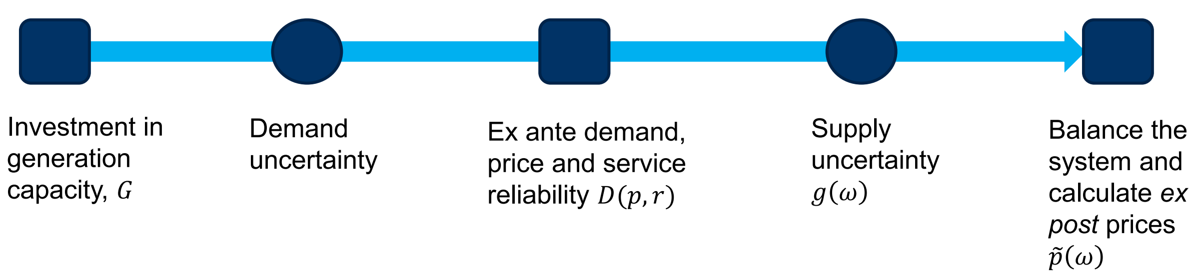


Figure 1 Time sequence of decision process

In Figure 1, the event starts with long-term investment on the left side and ends with real-time system dispatch on the right side. In the long run, we wish to determine the capacity investment in renewable under demand and supply uncertainties. The common objective of these decisions is to maximize the expected social welfare that equals the sum of the consumer and producer surplus. Our focus is within a market environment, where investors bear the financial risks of making entry and investment decisions in the presence of uncertainty about future states of demand and supply, while recognizing that external to the market, long-term bilateral contracts and financial products could facilitate mutual hedging between buyers and sellers on their financial and price risks.

In the short run, we assume that the available generating capacity is fixed though its output is uncertain. Traditionally, short-run economic efficiency is largely derived from the marginal cost differences arising from the diversity of fuels and technologies. With zero marginal cost, short-run efficiency would necessarily depend on the differences in the value of service and outage costs among consumers. Smart metering plays an essential role for effective demand-side management facilitating system operation in balancing demand and supply to maintain reliability. Under supply uncertainty, market price determination is necessarily ex ante before the full resolution of uncertainty about available supply. In practice, a system operator needs to exert physical control during the last ten minutes before actual dispatch. Shortage events that require load shedding cannot be completely eliminated when excess demand persists. Quantity rationing is unavoidable.

Assumptions

To focus on fundamental issues, we adopt some simplifying assumptions. First, noting that the renewable generation technologies (e.g., winds and solar) produce energy output with zero marginal cost, we let $C(G)$ denote the capacity cost function, where G is the generation capacity. The energy output from the generation capacity is a random variable $g(\omega) \in [0, G]$ with a probability distribution function, $F(g)$.[5] Second, on the demand side, we adopt the convention of representative demand function assuming that consumers have the same demand profile up to a scale; the demand function, $D(p) \in [0, \hat{Q}]$, where p is the price, is stationary, and $\hat{Q}$ is the maximum demand level. We let $r(q)$ denote the service reliability measured by the probability of load being served, $u(q)$ denotes the marginal utility function which is interpreted as the gross benefit for the q-th consumption unit from consuming one unit of energy, $w(q)$ denote the marginal disruption cost incurred when the service is interrupted including, for example, the stress and inconvenience when people get caught in the elevator causing, or wastages incurred when manufacturing processes are interrupted abruptly. When a shortage event occurs, the outage cost for each consumtion unit is given by, $v(q) = u(q) + w(q)$, which equals the sum of the foregone benefit of consumption and the disruption cost. Last, we ignore the income effects and other externalities.

Social welfare function

The social welfare function is written as follows,

$$\begin{aligned} SW &= \max_{G, r(\cdot)} E\left[\int_0^{\hat{Q}} \{r(q)u(q) - [1 - r(q)]w(q)\}dq\right] - C(G) \\ &= \max_{G, r(\cdot)} E\left[\int_0^{\hat{Q}} \{r(q)v(q) - w(q)\}dq\right] - C(G) \end{aligned} \tag{1}$$

[5] Throughout this paper, we adopt the convention that $\omega \in \Omega$ represents a random variable with a probability space, $\Omega \times \mathcal{B} \times \mathcal{P}$, where Ω is the state space, $\mathcal{B}$ is a Borel-field on Ω and $\mathcal{P}$ is the probability measure.

For simplicity, we assume that the investment cost function is linear, i.e., $C(G) = kG$, and the probability density for supply is uniform in the interval $[0, G]$, i.e., $F(g) = g/G$. In general, the parameter k is interpreted as the marginal capacity cost. Further, we assume that the marginal utility, marginal outage cost and marginal disruption cost functions are linear: $u(q) = \hat{u}\left(1 - q/\hat{Q}\right)$, $v(q) = \hat{v}\left(1 - q/\hat{Q}\right)$ and $w(q) = \hat{w}\left(1 - q/\hat{Q}\right)$. The analysis can be extended with no loss of generality to the case with continuously differentiable convex cost functions and continuously differentiable concave utility functions.

3. Integrated resource planning with no meters

In this section, as a benchmark case, we assume that consumers do not have access to conventional meters for recording kilowatt-hour energy consumption. The system planner performs integrated resource planning to determine the capacity level based on peak demand forecast and the cost is allocated to consumers through fixed fees.

We let $\hat{Q}$ denote the peak demand and $\bar{r}(G)$ denote the uniform service reliability, a function of the capacity level, G. We let $V\left(\hat{Q}\right)$ denote the total outage cost, which equals the integral of $v(q)$ over the range $\left[0, \hat{Q}\right]$, and similarly, $W\left(\hat{Q}\right)$ is the total disruption cost, which equals the integral of $w(q)$ over the same range. The social welfare maximization problem could be written:

$$SW = \max_{G,\bar{r}} \bar{r}(G)V\left(\hat{Q}\right) - W\left(\hat{Q}\right) - C(G) \tag{2}$$

The first-order condition for the welfare maximization problem (2) is as follows,

$$\bar{r}'(G)V\left(\hat{Q}\right) = k \tag{3}$$

Equation (3) indicates that at the optimal capacity investment level, the incremental benefit of investment from reduced outage cost equals the marginal capacity cost. We define the value of lost load (VOLL) as the average outage cost,

$$VOLL \equiv \frac{V\left(\hat{Q}\right)}{\hat{Q}} = \frac{\hat{v}}{2} \tag{4}$$

When there is excess supply, we assume that disposal is free. When there is excess demand, we assume that the system operator could curtail demand through random or rotating load shedding. Thus, the service reliability for each consumer is as follows,

$$\bar{r}(G) = E\left[Min\left\{\frac{g(\omega)}{\hat{Q}}, 1\right\}\right] = \int_0^G Min\left\{\frac{g}{\hat{Q}}, 1\right\} dF(g) = 1 - \frac{\hat{Q}}{2G} \tag{5}$$

We define the loss of load probability as $LOLP(G) = 1 - \bar{r}(G)$.

Solving (3) with (5) and (6), we obtain.

Proposition 1: The optimal capacity investment level is as follows

$$G^* = \left[\frac{VOLL}{2k}\right]^{1/2} \hat{Q} \tag{6}$$

and the loss of load probability is given by

$$LOLP(G^*) = \left[\left(\frac{1}{2}\right)\frac{k}{VOLL}\right]^{1/2} \tag{7}$$

Optimal capacity investment requires accurate estimates of peak demand level, the value of lost load and the marginal cost of new capacity. In (6), the optimal capacity level increases linearly with the peak demand, and it reflects a tradeoff between the outage cost and the capacity cost increasing with the value of lost load but decreases with the marginal capacity cost. On the other hand, the optimal LOLP target in (7) increases with the marginal capacity cost but decreases with the value of lost load.

With capacity planning, consumers collectively bear the economic risks of over- and under-investments. Under the conditions of no market power, and if physical performance of new capacity can be assured through contractual arrangement, the financial risks could be allocated to competing generators through a competitive capacity procurement auction.

Example

We consider a simple example below with two time periods (peak and off-peak). Table 1 provides the model parameters and numerical assumptions which will be used for the other examples in later sections.

Table 1. Model parameters and assumptions

Parameter	value
Maximum consumption value, $\hat{u}$	2500
Maximum disruption cost, $\hat{w}$	7500
Maximum outage cost, $\hat{v}$	10000
Peak demand $\hat{Q}_1$	150,000
Off-peak demand $\hat{Q}_2$	100,000
Fraction of peak period, θ	0.5
Marginal cost of capacity, k	30

Table 2 shows the results for the case with integrated resource planning. It shows that when the likelihood of system blackout α decreases, the capacity investment, capacity margin and capacity cost decrease and the social surplus increases.

Table 2. Results for Case 1 - integrated resource planning

	Case 1
Capacity level, G	1,163,687
Capacity margin	776%
Capacity cost, kG	34,910,600
Service reliability, $\bar{r}(G)$	0.946
Social surplus	86,428,800

4. Ex ante pricing with traditional TOU meters

In this section, we examine pricing mechanism assuming that consumers have access to the conventional meters (including time-of-use meters). Under supply uncertainty, pricing is inherently ex ante and "second-best" efficient.

We let $D(p)$ be the demand function where p is the price. The uniform service reliability is denoted by $\bar{r}$. Suppose that the demand is set at a level that maximizes the consumer surplus as follows,

$$D(p) = \underset{q}{\operatorname{argmax}}\, \bar{r}V(q) - W(q) - pq \tag{8}$$

The first-order condition is as follows,

$$\bar{r}v(D(p)) - w\big(D(p)\big) = p \tag{9}$$

Solving (9) yields the ex-ante demand function,

$$D(p) = \hat{Q}\left[1 - \frac{p}{\bar{r}\hat{v} - \hat{w}}\right] \tag{10}$$

Note that the ex-ante demand function in (10) is conditioned on the expectation of the service reliability, $\bar{r}$. If $\bar{r} = 1$, then the ex-ante demand function is the same as the ex-post demand function that equals to the inverse marginal utility function, $D(p) = u^{-1}(p)$.

The value of lost load is defined as the average outage cost,

$$VOLL(p) \equiv \frac{V\big(D(p)\big)}{D(p)} = \hat{v}\left(1 - \frac{D(p)}{2\hat{Q}}\right) \tag{11}$$

We assume that the expected service reliability is a function of price and the capacity level, denoted by $\bar{r}(p, G)$. As in Section 3, we consider two possible shortage events: A) random load shedding and B) total system collapse. With random load shedding, the service reliability for each consumer is given by,

$$\bar{r}(p, G) = E\left[Min\left\{\frac{g(\omega)}{D(p)}, 1\right\}\right] = 1 - \frac{D(p)}{2G} \tag{12}$$

The loss of load probability is defined as $LOLP(p, G) = 1 - \bar{r}(p, G)$.

The problem of social welfare maximization can be stated as follows,

$$SW = \max_{p,G} \bar{r}(p, G)V(D(p)) - W(D(p)) - C(G) \tag{13}$$

The first-order optimality condition can be written as follows

$$\bar{r}_G(p, G)V\big(D(p)\big) = C'(G) = k \tag{14}$$

$$\big[\bar{r}(D(p), G)v\big(D(p)\big) - w\big(D(p)\big)\big]D'(p) + \bar{r}_p(p, G)V\big(D(p)\big) = 0 \tag{15}$$

Using (9), we simplify (15) and obtain,

$$p = -\bar{r}_p(p, G)V\big(D(p)\big)/D'(p) \tag{16}$$

The optimal capacity level and the optimal ex ante price can be obtained by solving (14) and (16) jointly. The results are summarized below.

Proposition 2: The optimal capacity level, the reliability target and the ex-ante price are as follows,

$$G^* = \sqrt{\frac{VOLL(p^*)}{2k}} D(p^*) \tag{17}$$

$$p^* D(p^*) = kG^* \tag{18}$$

Note that the optimal capacity level increases linearly with the peak demand at a rate that increases with the ratio between the value of lost load and the marginal capacity cost. Equation (18) shows that the total revenue equals the total capacity cost. The revenue sufficiency condition will be preserved if the capacity cost function is convex.

With non-convex capacity cost function, a quasi-equilibrium can be obtained by using the convex hull relaxation $\check{C}(G)$ of $C(G)$ and $k = \check{C}'(G)$. The quasi-equilibrium price yields minimum revenue deficiency and incentive compatibility when the number of consumers increases to infinity.[6]

Example

Table 3 compares the results of Case 2 with ex ante pricing and Case 1 with integrated resource planning.

Table 3. Results for Case 2 with ex ante pricing in comparison with Case 1

		Case 1	Case 2
Optimal capacity level		1,163,687	1,053,094
Capacity margin		776%	756%
Capacity cost		34,910,600	31,592,818
Service reliability		0.946	0.950
Demand	Peak		123,022
	Off-peak		88,725
Price	Peak		344.58
	Off-peak		234.38
Market revenue	Peak		21,195,275
	Off-peak		10,397,692
	Total		31,592,967
Social surplus		86,428,800	89,237,359

Table 3 shows that the standard patterns of peak and off-peak price-quantity profiles. The total market revenue covers the capacity investment cost. Relative to the benchmark Case 1, the service reliability and the social surplus are increased by 0.4% and 3.2% respectively.

Ex ante pricing vs spot pricing

An ex-ante price is set in advance before the demand and supply uncertainties are completely resolved. Ex ante pricing achieves second-best efficiency recognizing the inherent informational imperfection in electricity markets. Ex ante pricing includes time-of-use rate, but it differs from spot pricing. Ideally, ex post marginal cost pricing would achieve Pareto efficiency and perfect price rationing after the complete resolution of demand and supply uncertainties. In practice, spot pricing presumes supply certainty yielding capacity investment plans based on the expected or de-rated capacity value adjusted with added capacity margin, a practice that is suboptimal.

With supply certainty, the known results in optimal pricing and investment conditions are as follows:

[6] See Chao (2019)

$$G^* = D(p^*) \tag{19}$$

$$p^* = k \tag{20}$$

Equation (19) indicates that the capacity investment should equal the peak demand, and Equation (20) indicates that the price would recover the marginal capacity cost when the marginal variable cost is zero. To allow for the uncertainty of supply availability, the optimal capacity margin is 100%. If the supply uncertainty is ignored, the spot pricing model could lead to under-investment and revenue deficiency. Table 4 compares ex ante pricing and spot pricing results under supply uncertainty.

Table 4. Ex ante pricing vs spot pricing

		Ex ante pricing	Spot pricing
Optimal capacity level		1,053,094	292,800
Capacity margin		756%	100%
Capacity cost		31,592,818	8,784,000
Service reliability		0.950	1.000
Demand	Peak	123,022	146,400
	Off-peak	88,725	100,000
Price	Peak	344.58	60.00
	Off-peak	234.38	0
Market revenue	Peak	21,195,275	4,392,000
	Off-peak	10,397,692	0
	Total	31,592,967	4,392,000
Social surplus		89,237,359	151,804,000

5. Priority service pricing with smart meters

In this section, we examine the design of efficient nonlinear pricing of priority services with smart meters. We proceed by describing the basic components of consumer choice, efficient rationing, and menu design underlying the problem of social welfare maximization.[7]

Consumer choice

Suppose that consumers choose from a price menu of service reliability options denoted by $\mathcal{M} = \{(r, p(r))\}$ where $p(r)$ is a nonlinear price schedule of service reliability. We assume that for each consumption unit of q, the consumer chooses the service reliability $\hat{r}(q)$ that maximizes its expected net benefit:

[7] The basic arguments herein largely follow the more theoretical developments in Chao and Wilson (1987), Wilson (1989) and Chao (2012).

$$\max_{r} rv(q) - w(q) - p(r) \tag{21}$$

The optimal self-selection depends on the consumption unit's outage cost, $\hat{r}(v(q))$, which satisfies the first-order optimality condition,

$$p'\big(\hat{r}(v(q))\big) = v(q) \tag{22}$$

Equation (22) provides the incentive compatibility condition for the price schedule.

The consumer surplus is the integral of the net benefits for all consumption units. The quantity $\hat{Q}$ is the total number of consumption units in the market. The following equality serves as the boundary condition for (22):

$$\hat{r}(v(Q))v(Q) - w(Q) - p\big(\hat{r}(v(Q))\big) = 0 \tag{23}$$

Expression (23) provides the individual rationality condition for the ex-ante demand schedule.

Efficient rationing plan and menu design

Priority service pricing implements an efficient rationing plan that assigns service reliability in an increasing order of the outage cost so that a consumption unit with higher outage cost will receive a higher priority and service reliability than those units with lower outage costs. A key idea of menu design is to design an incentive compatible price schedule so that consumers will be motivated to self-select according to the efficient rationing. To implements the efficient rationing plan, the service reliability assignment under priority service can be written as,

$$\hat{r}(v(q)) = r(q, G) = Pr\{g(\omega) \geq q\} = 1 - \frac{q}{G} \tag{24}$$

Using (24), we solve (22) and (23) and obtain the price schedule,

$$p(r) = \frac{\hat{v}\hat{Q}}{2G}\left[1 - (1 - r)\frac{G}{\hat{Q}}\right]^2 \tag{25}$$

Substituting (24) into (25) yields,

$$p\big(\hat{r}(v(q))\big) = \frac{\hat{Q}v(q)^2}{2\hat{v}G} \tag{26}$$

Social welfare maximization

Invoking (24), we state the optimal capacity planning problem as follows,

$$SW = \max_{G} \int_0^{\hat{Q}} \left\{\left(1 - \frac{q}{G}\right)v(q) - w(q)\right\} dq \; - C(G) \tag{27}$$

The first-order optimality condition is given by,

$$\int_0^{\hat{Q}} \left\{\frac{q}{G^2} v(q)\right\} dq \; - k = 0 \tag{28}$$

Solving (28) yields the optimal capacity level,

$$G^* = \sqrt{\frac{\hat{v}}{6k}\hat{Q}} \tag{29}$$

By integrating (26) over $[0, \hat{Q}]$ and using (29), we obtain,

$$\int_0^{\hat{Q}} p\big(\hat{r}(v(q))\big) dq = kG^* \tag{30}$$

Equation (30) indicates that the market revenue covers the capacity investment cost.

The above results are summarized as follows.

Proposition 3. The optimal price schedule provided by (25) – (26) and the optimal capacity investment plan (29) – (30) achieve efficient allocation and revenue sufficiency.

Example

In the following, we compare the three cases, 1) integrated resource planning, 2) ex ante pricing and 3) priority service pricing.

Table 5 shows that among the three approaches, priority service pricing produces the first-best result with the lowest capacity level, the lowest capacity margin, and the lowest capacity cost and the highest social surplus. Under priority service pricing, the optimal capacity level is determined by customers' selections that reveal the benefit of capacity expansion. Both priority service pricing and ex ante pricing collect sufficient revenue to pay for the capacity cost. Compared with the benchmark case, ex ante pricing and priority service pricing raise the social surplus by 3.25% and 14.82%, respectively. Priority service yields perfect service reliability which reflects not only efficient rationing but also customer self-selection obviating involuntary service curtailment, a defining signature of unreliable service. In essence, priority service transforms service reliability from a common public good to differentiated private products.

Table 5. Integrated resource planning, ex ante pricing and priority service pricing

	Integrated resource planning	Ex ante pricing	Priority service pricing
Capacity level	1,163,687	1,053,094	950,146
Capacity margin	676%	756%	533%

Capacity cost		34,910,600	31,592,813	28,504,385
Service reliability		0.946	0.950	1.000
Demand	Peak	150,000	123,022	150,000
	Off-peak	100,000	88,725	100,000
Price	Peak		344.58	0 - 789.35
	Off-peak		234.38	0 - 526.23
Energy market revenue	Peak		21,195,288	19,733,806
	Off-peak		10,397,696	8,770,580
	Total		31,592,984	28,504,386
Social surplus		86,428,800	89,237,359	99,241,229

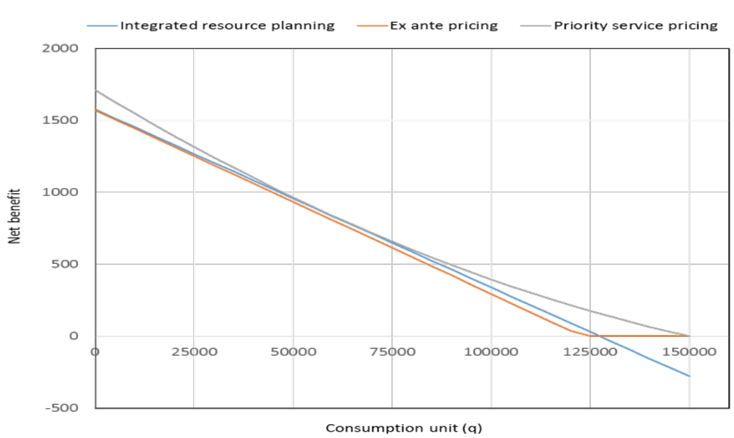
Table 6 illustrates consumer's valuation of priority service options in order to select one that maximizes the net benefit according to (21) as shown along the diagonal elements in the matrix located at the lower-right corner.

Table 6. Consumer self-selection from priority service menu

			r	0.874	0.905	0.937	0.968	1.000
			$p(r)$	32	126	284	505	789
q	$v(q)$	$w(q)$		$rv(q) - w(q) - p(r)$				
120	2,000	1,500		216	184	90	-68	-289
90	4,000	3,000		463	495	463	369	211
60	6,000	4,500		711	805	837	805	711
30	8,000	6,000		958	1116	1211	1242	1211
0	10,000	7,500		1205	1426	1584	1679	1711

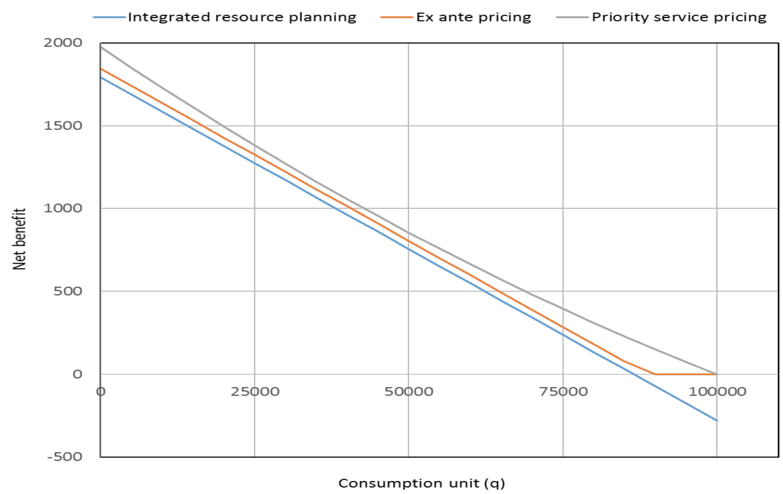
Priority service pricing Pareto dominates ex ante pricing with random rationing. Figures 2 and 3 show that priority service pricing yields higher consumer net benefit than integrated resource planning and ex ante pricing for every consumption unit in both peak and off-peak periods.

Figure 2. Consumer net benefits in the peak period



Figures 3 shows that consumers are unambiguously better off under ex ante pricing in the off-peak period than integrated resource planning. This is largely because peak-load pricing shifts the burden of capacity cost recovery from off-peak consumers to peak consumers, who were cross-subsidized under integrated resource planning.

Figure 3. Consumer net benefits in the off-peak period



6. Implementation and financial risk management issues

The priority service menu $\mathcal{M} = \{(r, p(r))\}$ can be implemented through alternative market organizations that offers customers choices of unbundled service reliability options. (EPRI, 1986, 2004) Consumers with a high outage cost would choose a higher service reliability option and pay a higher premium than those who are willing to take lower service priorities with more frequent interruptions in exchange for lower rates. Risk averse customers could procure compensatory insurance to hedge against the financial consequences of interruptions in a way that the insurance premium equals the sum of the priority service charge and actuarial risk based on the insurance coverage selected, yielding efficient risk sharing and efficient rationing simultaneously.

Alternatively, priority service on the demand side can be implemented through curtailment options procured and exercised by the service provider as a load modification strategy. A third-party aggregator can bundle such options into a virtual power plant represented by a supply function for curtailed energy. Regardless of the contractual arrangement used to enlist curtailable demand, the resulting portfolio of demand side resources underlying the virtual power plant consists of curtailable capacity with various curtailment options characterized by attributes such as frequency, duration, and cost. Such curtailable capacity can be segmented into priority tranches differentiated by marginal cost per kW per hour and sorted in merit order. An important aspect in the implementation of curtailable electricity service options is a "nameplate capacity" of curtailable devices, an observable and verifiable quantity by which the contract can be defined. However, the benefit to the system is determined by the curtailed energy consumption which constitutes load relief and avoided generation. Unfortunately, the curtailed energy at small scale on device level is practically unobservable due to prohibitive metering cost. Furthermore, even if energy consumption could be metered right before curtailment it is still unknown how much energy would have been consumed beyond that point, especially for thermostatically controlled loads, so that the avoided energy consumption over the curtailment interval remains uncertain. To address this issue, we introduce the concept of "yield" as a random variable characterizing the fraction of curtailed capacity representing avoided energy consumption, The yield associated with a curtailable unit of contracted capacity is characterized by a probability distribution estimated from load data, conditioned on the priority tranche and observable characteristics such as device type, geographic location, time of day and ambient temperature.

An aggregator that participates in the wholesale market by offering a supply function specifying curtailed energy as function of wholesale price has at his/her disposal a portfolio of curtailable capacities organized into tranches defined by a capacity limit in each time interval and marginal cost per curtailed kW per hour. The aggregator must then manage the delivery risk by selecting the capacity curtailed in each priority tranche and quantity of curtailable energy offered as function of the wholesale price so as to maximize net profit given the wholesale price. The aggregator profit consist of the offered energy times the wholesale price less the cost of the curtailed capacity less a shortfall penalty imposed on any shortfall between the awarded energy offer and the realized energy curtailment given the realized uncertain yields.

The aggregator's problem described above is formulated as a conventional Revenue Management problem (a variant of the well-known Newsboy Problem) similar to the formulation used in the airline industry for allocating seat classes to uncertain demand segments (See Ozalp and Phillips 2012). Let us assume that the aggregator holds a portfolio of N contracted demand side resource types denoted by an indexed *i* and characterized by the following parameters specific to the offer interval t.

c_i - Cost per MW capacity per hour of curtailment of resource *i*

K_i - Available kW capacity of resource type *i*

k_i - Committed MW capacity of resource type *i*

y_i - Energy yield in kWh per hour of one kW capacity of resource type *i.*

$f_i(y_i)$ - Probability density function over energy yield of resource type *i*

$P_i(y_i)$ - Cumulative probability over energy yield of resource type *i*

p - Market clearing price

r - Shortfall penalty rate as a wholesale price multiplier.

The aggregator's objective is to maximize expected profit by selecting the committed curtailment capacity k_i for each resource type and the offered quantity Q of curtailed energy, solving the optimization problem:

$$
\begin{aligned}
&Max_{Q,k_1,..k_N}\left[pQ-\sum\nolimits_{i=1}^{N}c_ik_i-r\cdot p\cdot E_{y_1,...y_n}\{Max[0,(Q-\sum\nolimits_{i=1}^{N}y_ik_i)]\}\right]\\
&Subject\ to:\\
&0\le k_i\le K_i,\quad i=1,....N
\end{aligned}
$$

The shortfall penalty rate may reflect an actual penalty imposed by the system operator or just be used as a control parameter to manage the shortfall probability.

The offer Q is supported by the separate offers Q_i corresponding to each resource category *i.* so that $Q=\sum\nolimits_{i=1}^{N}Q_i$. Then the shortfall is bounded above by:

$$
Max[0,(Q-\sum\nolimits_{i=1}^{N}y_ik_i)=Max[0,\sum\nolimits_{i=1}^{N}(Q_i-y_ik_i)]\le\sum\nolimits_{i=1}^{N}Max[0,(Q_i-y_ik_i)]
$$

We approximate the profit function maximization by maximizing a lower bound on the expected profit function which ignores the possibility that surplus energy curtailment in one category can offset the shortfall in another category and reduces the penalty. The approximate optimization problem is:

$$
\begin{aligned}
&Max_{Q_1,...,Q_N,k_1,..k_N}\left[\sum\nolimits_{i=1}^{N}(pQ_i-c_ik_i)-r\cdot\sum\nolimits_{i=1}^{N}E_{y_i}\{Max[0,(Q_i-y_ik_i]\}\right]\\
&Subject\ to:\\
&0\le k_i\le K_i,\quad i=1,....N
\end{aligned}
$$

The above approximation problem is separable and can be solved separately for each resource category whereas the total supply function is the sum of the individual supply functions for each resource category.

We therefore only need to solve the problem for a single resource and can suppress the category index as follows:

$$
\begin{aligned}
&Max_{Q,k}\left[pQ-ck-r\cdot p\cdot E_y\{Max[0,(Q-yk]\}\right]\\
&Subject\ to:\\
&0\le k\le K,
\end{aligned}
$$

The above optimization problem can be transformed by simplifying the penalty term in the objective function using integration by parts, resulting in.

$$
Max_{Q,k}\left[pQ-ck-rpk\int_0^{\frac{Q}{k}}P(y)dy\right]
$$

$$\mathit{Subject\ to}:$$
$$0 \leq k \leq K,$$

First order necessary conditions for an interior maximum with respect to Q yields

$$p[1 - rP(\tfrac{Q}{k})] = 0$$

Resulting in an expression for the optimal Q, $Q^* = kP^{-1}(\tfrac{1}{r})$

Substituting Q^* in the profit function and using integration by parts yields

$$\text{Expected profit} = pQ^* - ck - rpk\int_0^{Q^*/k} P(y)dy = pkP^{-1}(\tfrac{1}{r}) - ck - rpk\int_0^{P^{-1}(\frac{1}{r})} P(y)dy = -ck + rpk\int_0^{P^{-1}(\frac{1}{r})} yf(y)dy$$

The expected profit is linear in k so it is maximized at either 0 or K (Bang-Bang solution) depending on the sign of the coefficient. i.e., $k^* = K$, if $\dfrac{c}{p} \leq r\int_0^{P^{-1}(\frac{1}{r})} yf(y)dy$, $k^* = 0$ otherwise

To simplify the notation Let $G(r) = \int_0^{P^{-1}(\frac{1}{r})} yf(y)dy$

then the optimal supply quantity is $Q^* = K \cdot P^{-1}(\tfrac{1}{r})$, if $\dfrac{c}{p} \leq r \cdot G(r)$, and $Q^* = 0$ otherwise

The ***expected shortfall*** is given by $k^*\int_0^{\frac{Q^*}{k^*}} P(y)dy = K \cdot \int_0^{P^{-1}(\frac{1}{r})} P(y)dy$. And the ***probability of a shortfall*** is $\Pr[(Q^* - k^* y) > 0] = \Pr[K \cdot (P^{-1}(\tfrac{1}{r}) - y) > 0] = \int_0^{P^{-1}(\frac{1}{r})} f(y)dy = \tfrac{1}{r}$ if $c \leq p \cdot r \cdot G(r)$, and 0 otherwise. In other words, the penalty parameter r determines the constant shortfall probability achieved by the optimal policy.

For illustrative purpose consider the case where y is uniformly distributed between 0 and 1. Then $G(r) = \tfrac{1}{2r^2}$ so $Q^* = K / r$ if $c \leq \tfrac{p}{2r}$, and $Q^* = 0$ otherwise

The ***expected net profit*** is $K \cdot p \cdot Max\left[(\tfrac{1}{2r}) - \tfrac{c}{p}), 0\right]$ and the ***expected* energy *shortfall*** is $K \cdot \tfrac{1}{2r^2}$

The VPP supply function corresponding to the entire portfolio has a staircase shape as illustrated in Figure 4 below.

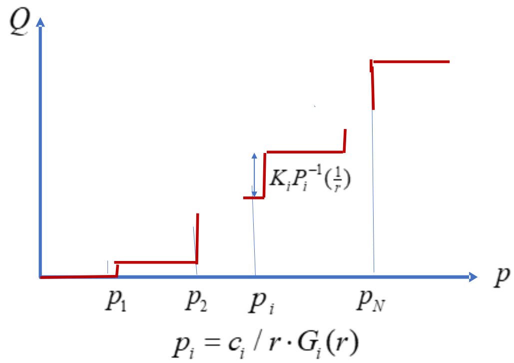


Figure 4: VPP Supply Function under optimized delivery risk

For uniform yield distributions $p_i = 2rc_i$ and the offered energy for resource *i is* K_i / r .

7. Conclusion

In this paper, we examine three cases with varying sophistication of metering technology 1) integrated resource planning with no meters, 2) ex ante pricing with traditional time-of-use meters and 3) priority service pricing, or nonlinear pricing of priority services with smart meters. Priority service Pareto dominates ex ante pricing. Priority service pricing facilitates demand-side management offering a menu of contingent contracts for distribution of uncertain supplies. Priory service prices are expectations of ex post spot prices for comparable services. Customers' selections reveal the benefit of capacity expansion. Priority service pricing can be implemented under alternative market organizations, e.g., in retail markets via buying compensatory insurance or in wholesale markets via selling demand reserves. Under conditions of supply uncertainty and zero marginal cost, priority service pricing achieves first-best ex post efficiency and assures revenue sufficiency for merchant investments. Priority service yields perfect service reliability which reflects not only efficient rationing but also customer self-selection obviating involuntary service curtailment, a defining signature of unreliable service. In essence, priority service transforms service reliability from a common public good to differentiated private products.

To address practical implementation challenges, we discuss a stochastic market auction platform that enables innovative demand-side management with priority service via an end-to-end business model, in ways that flexible demand devices (e.g., hot water heaters, air conditioners, energy storage etc.) on the customer end are aggregated into a “virtual power plant,” and an aggregator for distributed resources addresses the financial risks for curtailed energy on the supply side. In essence, priority service is implemented through curtailment options that can be employed by the service provider on the demand side. Furthermore, we address implementation issues concerning the contract design so that the benefits of demand-side flexibility harnessed through priority service mechanism can be monetized by a revenue management approach for financial risk management. The supply-side delivery risk can be controlled through contractual arrangements with customers that enables aggregator to form VPP offers in the wholesale market.

Acknowledgements

Hung-po Chao is affiliated with Energy Trading Analytics, Shmuel Oren with University of California, Berkeley, and Robert Wilson with Stanford University. The authors assume the sole responsibility for the views and any remaining errors in this paper. In actually writing, the first author had the initial responsibility for preparing the first version in 2019 when he was affiliated with PJM Interconnection. Subsequently, the revision has been subject to reviews and enhancements by the other authors. The authors are grateful to partial financial support of a grant from Department of Energy, Advanced Research Projects Agency-Energy PERFORM Program Office to Energy Trading Analytics, LLC. (DE-AR0001281).

The authors declare that they have no conflicts of interest regarding this work.

This article does not contain any studies with human or animal subjects performed by any of the authors.

References

1. Boiteux, M. (1949). “La Tarification des demandes en point: application de la theorie de la vente au cout marginal.” *Revue Generale de l’Electicité* 58 (August): 321-40; translated as “Peak Load pricing,” *Journal of Business* 33 (2) (1960): 157-79.
2. Chao, H. (1983). “Peak load pricing and capacity planning with demand and supply uncertainty.” *Bell Journal of Economics* 14(1), 179-190
3. Chao, H. (2012) “Competitive Electricity Markets with Consumer Subscription Service in a Smart Grid,” *Journal of Regulatory Economics*.
4. Chao, H. (2019) “Incentives for efficient pricing mechanism in markets with non-convexities,” *Journal of Regulatory Economics* https://doi.org/10.1007/s11149-019-09385-w
5. Chao, H. and R. Wilson (1987) "Priority Service: Pricing, Investment and Market Organization," *American Economic Review*, 77 (5): 899-916.
6. Chao, H., S. Oren, and R. Wilson (2008) “Reevaluation of Vertical Integration and Unbundling in Restructured Electricity Markets”, Chapter 1 in the book: Competitive Electricity Markets, Design, Implementation and Performance, edited by Fereidoon P. Sioshansi

7. Chao, H., S. Oren, S. Smith and R. Wilson (1986) "Multi-level Demand Subscription Pricing for Electric Power," *Energy Economics,* pp. 199-217
8. Chao, H., S. Oren, S. Smith and R. Wilson (1988) "Priority Service: Market Structures and Competition," *The Energy Journal*, Special Issue on Electricity Reliability, 77-104.
9. Crew, M., Fernando, C., and Kleindorfer, P. (1995). "The theory of peak-load pricing: A survey." *Journal of Regulatory Economics*, 8(3), 215–248.
10. EIA (2019) U.S. Energy Information Administration, *Electric Power Monthly*, April 2019 https://www.eia.gov/todayinenergy/detail.php?id=39992#:~:text=U.S.%20electricity%20generation%20from%20renewables%20surpassed%20coal%20in%20April&text=In%20April%202019%2C%20U.S.%20monthly,electricity%20generation%20to%20coal's%2020%25.
11. EPRI (1986) *Priority Service: Unbundling the Quality Attributes of Electric Power*, Electric Power Research Institute, Report EA-4851, co-authored by Chao, H., S. Oren, S. Smith and R. Wilson
12. EPRI (2004) *Integrated Engineering and Economic Operation of Power Systems*, Electric Power Research Institute Report, co-authors: Chao, H., S. Oren, A. Papalexopoulos, D. Sobajic and R. Wilson.
13. FERC (2020) Federal Energy Regulatory Commission (FERC), Docket No. RM18-9-000; Order No. 2222 (September 17)
14. Hogan, W. W. (2021) "Electricity Market Design and Zero-Marginal Cost Generation" Prepared for *Current Sustainable/Renewable Energy Reports*. https://scholar.harvard.edu/files/whogan/files/hogan_emd_102121.pdf
15. Gedra, T.W. and Varaiya P.P. (1993)" Markets and Pricing of Interruptible Electric Power", *IEEE Transactions on Power Systems*, 8(1):122-128
16. Joskow, P. (2019) "Challenges for Wholesale Electricity Markets with Intermittent Renewable Generation at Scale." *Oxford Review of Economic Policy*, 35 (2): 291-331
17. Joskow, P. L. and J. Tirole, (2007), "Reliability in Competitive Electricity Markets, *Rand Journal of Economics*," 68:159-170.
18. Kamat, R. and Shmuel S. Oren. (2002) "Exotic Options for Interruptible Electricity Supply Contracts," *Operations Research*, 50(5): 835-850
19. Oren, S. S., S. A. Smith and R. B. Wilson (1985), "Capacity Pricing," *Econometrica*, 53 (3): 545-567.
20. Ozer, O. and R. Phillips (edit), (2012) *The Oxford Handbook of Pricing Management*, Oxford University Press, Oxford
21. Weitzman, M. L. (1974) "Prices vs. Quantities" *Review of Economic Studies*, 41 (3): 477-491
22. Wilson, R. (1989) "Efficient and Competitive Rationing", *Econometrica*. 57 1-40.

23. Wilson, R. (2002) “Architecture of Power Markets," *Econometrica*, 70, 1299-1340