Price and Environment in Electricity Restructuring

Donald N. Dewees
Department of Economics
Faculty of Law
University of Toronto

August 23, 2001

An earlier version of this paper was presented at the Law and Economics Workshop at the Faculty of Law of the University of Toronto on March 19, 2001. The comments of the workshop participants are gratefully acknowledged.
Comments are welcome; please email to dewees@chass.utoronto.ca .

Table of Contents

Price and Environment in Electricity Restructuring

1. Introduction

1.1 Motives for restructuring

Electricity restructuring involves taking power away from monopolists and regulators and giving it to markets. Traditionally the planning of the electrical system and its operation have been the responsibility of regulated integrated utilities that generated, transmitted and distributed the electricity. Restructuring involves at a minimum eliminating the statutory monopoly on generation and allowing competition in generation as well as requiring the transmission owner to transmit power from generators to their customers. It may go further and establish a competitive wholesale power pool that sets an hourly (more or less) spot price. It may require integrated utilities to divest some of their assets; owners of the transmission and distribution networks, which remain natural monopolies, may be precluded from engaging in competitive activities such as generation and retailing. At the retail level, the distribution utility may retain a monopoly over the supply of electricity to customers in its service area, in which case it will purchase power on their behalf under contract and on the spot market, overseen by a regulator. Alternatively retail competition, also called retail access, may be mandated, allowing competitive retailers to arrange price terms with customers, paying the distributor for delivering the power.

The wave of electricity system restructuring during the last decade has been stimulated by several forces. The development of new generation technology, the combined cycle gas turbine (CCGT) has greatly reduced the minimum efficient scale of a generating plant, allowing large electricity consumers to install their own generation and facilitating generation by non-utility generators. The development of new metering technology and the information technology infrastructure to communicate prices and consumption at low cost has facilitated competition in metering and providing load-management services to consumers. While these technologies facilitate restructuring, however, Joskow (1997, p. 123) argues that the primary driver for restructuring has been the high prices charged by regulated utilities in the 1990's as they recovered the cost of investments in expensive facilities, when falling natural gas prices and increased efficiency of CCGT plants reduced the cost of new generation. Large customers rebelled at paying 6 or 7 cents per kilowatt-hour (kWh) to regulated utilities when the marginal cost of wholesale electricity was 2.5 cents/kWh and long run marginal cost was 3-4 cents/kWh. (Joskow, 1997, p. 126.) The solution was to introduce competition in generation, dealing with the high costs of incumbent utilities through an agreement on the amount of “stranded debt” that would be recouped, and who would pay for it. This seemed to be the logical next step after successfully eliminating rate regulation and introducing competition for railroads, trucking, airlines, natural gas, and long distance telephone rates from the late 1960's to the 1990's in the United States.

1.2 Experience with restructuring

By early 2001, the experience with restructuring was somewhat mixed. In England and Wales, generation costs fell dramatically after restructuring began in 1988, but this was the result mostly of declining coal and gas prices and only partly from increased labour productivity and improvements in the performance of nuclear plants. Restructuring coincided with the North Sea gas bubble, so there was a rash of CCGT construction. However two firms own most mid-merit (price-setting) plants (Wolfram, 1998, 706), and there was no retail competition for small customers until at least 1998, so the reductions in generation costs raised profits of generators much more than they lowered consumer prices. In Victoria, Australia, a monopoly generator was split into five one-plant firms, followed by similar divestment in four other states. By 2001, there were 17 generators competing in a national market, leading to improved plant utilization, better reliability, lower costs and reduced prices to consumers. In New Zealand a monopoly generator was split into two principal firms and a market power mitigation agreement was imposed. Consumer prices for electricity fell about 20% between 1992 and 2000 and new capacity has been constructed.

Alberta introduced the *Electrical Utilities Act*[1] in 1995 which adopted a mandatory power pool and encouraged market participants to enter into side contracts, contracts for differences, to hedge against price variations. Apparently few did. Despite steady economic growth, there was no new investment in generation capacity pending the resolution of some market design issues. Spot prices on the Power Pool of Alberta rose as a supply crunch developed, and by late 1999 average prices had doubled from their 1996 values. In 2000 they rose still further. Consumers, many of whom were paying spot prices, were furious at the price increases and accused the Alberta government of bungling restructuring. (Nikiforuk, 2001.) In 2001, new generation plants are being constructed, fuelled by both coal and gas.

Utilities in eastern Pennsylvania, New Jersey, Delaware and eastern Maryland have for decades coordinated electricity generation and transmission through PJM. In 1997 and 1998, PJM initiated a regional bid-based electricity market and established an independent system operator handling about 8% of US electrical power. The incumbent utilities remain in place, but customers can choose their supplier and generation is competitive. PJM adopted a nodal price system in which prices between nodes on the transmission network could vary when congestion limited flows between the nodes. The market seems to be working smoothly.

The California power market opened in March, 1998, after eight years of economic expansion during which no new generation was construction by the regulated public utilities because of environmental and community opposition, although municipal utilities and non-utility generators installed some capacity. After initial price drops, wholesale prices have skyrocketed, propelled in part by air pollution emission limits. A competition transition charge (CTC) equal to the difference (positive or negative) between the wholesale price and a retail price cap agreed to by the major stakeholders repaid some stranded debt when wholesale prices were low, but cost distributors $10 billion in one year when wholesale prices rose in 2000. The distributors' financial distress has led out-of-state suppliers to refuse to sell without secure payment. Consumer prices are being raised, the state is guaranteeing wholesale price payments, wholesale price caps have been imposed and environmental standards are being suspended to unleash all available generation. (Gallon, 2001.) Some of the municipal utilities, which serve over 20 percent of residential customers, opted out of restructuring for the time being, and those that are exporters of power (including Los Angeles) are little impacted except for enjoying some refreshing revenues. San Diego completed its retirement of stranded debt during 2000, so its customers were not covered by the price cap after that time, exposing them to spot prices and releasing their

[1] R.S.A., 2001, c. E-5.5.

wrath upon the California government which many accuse of bungling restructuring.

1.3 Questions

The brief histories above show that some restructuring appears to have brought broad benefits while others have brought broad misery. They may raise many questions, two of which are addressed here:

1. With costless metering, no transactions costs and no risk aversion, efficiency would call for all customers to pay a price equal to the marginal cost of generation every hour of every day. But sophisticated meters are costly, marginal cost pricing imposes substantial financial risks on both generators and customers, demand is relatively inelastic, and fixed prices are much easier to understand and compare than variable prices. Given these competing considerations, what would efficient retail price plans look like? Will markets find an efficient combination of spot prices and risk management? If not, is this the result of some artificial institutional barrier or is it an efficient response to information costs, transactions costs and inelastic demands?
2. What effect do environmental regulations have on the performance of restructured electricity markets? Do particular types of environmental regulations raise electricity prices artificially increased at times of peak demand? How can the design of environmental regulations create incentives for efficient control of pollution without artificially constraining electricity output?

To address these questions this paper will describe the central elements of electricity systems and restructured electricity markets. The analysis will draw on some of the experience with selected restructurings to date. Readers familiar with the electricity sector may wish to skip section 2, which provides background information on this industry.

2. Understanding Electricity Basics

Electricity cannot be stored; it must be generated exactly when it is consumed. It can be transported long distances, but long distance transmission capacity is limited. These facts cause electricity markets to differ substantially from markets for many commodities. Understanding the behaviour of electricity markets requires some understanding of the shape of electricity demand and supply functions including the time variations of demand, of average and marginal cost pricing, of investment in generation capacity, and of retail competition. We will consider these issues in a hypothetical electrical service area, such as a large metropolitan area.

2.1 Supply and Demand Elasticity

A typical service area will have three types of generation facilities: baseload, mid-merit (medium cost), and peaking. Baseload plants have high capital cost and low operating cost; they can be nuclear, run-of-the-river hydroelectric, or coal-fired thermal plants. Mid-merit plants have moderate capital and operating costs; they are usually coal or oil-fired steam turbine plants, or combined cycle gas turbines (CCGT).[2] Peaking plants have low capital cost and high operating

[2] CCGT involves a gas turbine, like a jet engine, which turns a generator. The exhaust from the turbine is fed into a boiler which makes steam that powers a steam turbine and a second generator. While thermal plants often require as much as 10,000 Btu to make a kilowatt-hour of

cost, and often rely on a simple cycle gas turbine or storage hydroelectric power. The aggregate marginal cost curve for a typical fleet of plants has costs that rise slowly over a range of output but rise rapidly as capacity is approached. See SMC in Figure 1. The marginal cost curve becomes vertical at the capacity of the system. If some of the capacity becomes unavailable, SMC could shift up or left, for example to SMC′. The mid-merit plants set the system marginal cost much of the time; in a competitive market, they usually set the price. Kahn (1988, p. 121) shows the system marginal cost for PG&E in California around $42/Mwh in a summer night, and over $60 during the day.

Empirical studies have found price elasticities of demand for electricity to be relatively small in the short run and near unity in the long run. Baughman, Joskow and Kamat (1979, pp. 52, 70) estimated own-price demand elasticities for the residential and commercial sectors at -0.19 in the short run (one year) and -1.0 in the long run (20 years), while the corresponding elasticities in the industrial sector were -0.11 in the short run and -1.28 in the long run. Their study, however, examined annual average electricity rates. Ham, Mountain and Chan (1997, 132-137) studied the response of small commercial customers (less than 50 kW peak demand) to very short run price changes, looking at peak/off-peak time-of-use pricing in Ontario and found own-price elasticities of -0.13 in the winter and -0.11 in the summer. There was significant substitution between peak and off-peak consumption only when the relative price ratio was as high as 6:1 between the two periods, and then the peak-period price elasticities were on the order of -0.067 to -0.091. The responsiveness to price increases was greater for a 2-hour peak than for a four-hour peak.

The existing studies of demand elasticity were conducted when many consumers had few options for reducing consumption in the short run or for shifting it by a few hours from peak to off-peak times, and when electricity prices had been relatively stable. The elasticities should increase with the development of improved load management technology, the expansion of load management services by energy service companies, increased variability of prices, and improved customer information regarding hourly prices and how to respond to them. Still, we should expect the price elasticity to be higher for within-day price variations as some customers shift loads by up to 12 hours by storing heat or cold; lower for periods of a day to a year, since storage is not useful for such periods, and higher as the time horizon expands beyond a year as new capital equipment may be installed to conserve electricity or to switch to alternative fuels.

2.2 Average Cost Versus Marginal Cost Pricing: Contract Prices and Spot Prices

Most traditional electric utilities and most competitive electricity markets (but not California) dispatch plants in merit order, which means they must know marginal generation costs on an hourly (more or less) basis. Some restructured markets use those marginal costs or bids by generators to generate a wholesale spot price reflecting the price-setting bid for that hour. If market is competitive, this spot price should represent the short run marginal cost of generation. While analyses of some existing markets has revealed “gaming” that gives rise to prices greater than marginal cost (Wolfram, 1998), and theoretical analysis has shown incentives to manipulate prices (Newbery, 1998) we will assume a market structure in which these deviations are not large.

---

electricity, a new CCGT may require only 6,500 Btu/kWh.

Economists' enthusiasm for competitive markets arises from the use of this price to equate supply and demand in real time; at every moment (or hour anyway) generators and consumers are matching supply and demand at a market price. At the margin, the cost of generation should equal the value of consumption.

Because electricity cannot be stored, and demand varies over time, marginal costs vary, sometimes substantially. Figure 1 shows a set of demand curves that may represent demand at different times on a given system. D1 could represent night time demand in the spring or fall, when demand is at its lowest, while D3 could be weekday demand in the same season. D2 could be night time demand in the winter when heating loads are high, and D4 could be weekday winter demand. D5 could be summer weekday daytime demand in a hot climate. And D6 could represent an unusual peak demand. The monthly average spot price in the PJM control area in Pennsylvania was $16 per MWh in February, 1999 but more than $90 in July as the summer air conditioning demand reached its peak.[3] The SMC′ curve in Figure 1 represents supply when some generation capacity is unavailable because of breakdown or maintenance or when fuel costs for high-cost plants have increased, as we saw in 2000 when upward leaps in gas prices increased generation costs for many utilities, most visibly in California. Figure 1 shows that shifts in either supply or demand may substantially increase or decrease the short run marginal cost, and thus the efficient price, especially if the system is close to full capacity utilization. In a competitive market, the spot price will respond at once to these factors.

Charging customers the spot price is quite unlike the traditional practice of charging a regulated price fixed for a period of a year or more to all customers in a class: a fixed price/kWh at all times for small customers, a fixed price/kWh plus a demand charge for highest usage in the billing period for medium users, and perhaps more sophisticated peak-load pricing for largest customers. Interruptible contracts, available to large customers, do not change the price, but allow the utility to cut off the customer if a shortage of capacity emerges.

The spot price may be efficient but it can lead to variable electricity bills as was dramatically demonstrated in Alberta in the year 2000 when persistent high spot prices doubled or tripled normal electricity bills, with little relief expected until new capacity comes on line. If the utilities had still been under rate of return regulation, their cost-based prices would have barely risen in 2000. Whether the utility operates for 5% of the year on D5 or 10% of the year on D5 will only modestly affect total costs. However if the price for all electricity generated in an hour is set by the marginal cost of generation in that hour, spending 5% more time on D5 rather than D4 can substantially increase the generator's revenue and of course the customer's cost. With marginal cost pricing, small shifts in supply or demand can substantially affect the load-weighted marginal cost.

There are at least two general solutions to the volatility of electricity costs in competitive markets. One is for retailers or distributors to purchase a portfolio of electricity supply contracts covering most of their customer demand and to charge consumers a price that blends the cost of this portfolio and of the spot price for the remaining quantity. If 80% of the supply is purchased on contract, then the variability of the customer's cost will be reduced by 80% from the variability

---

[3] See monthly price data at www.pjmiso.com.

of the spot price.[4] Unfortunately the marginal price paid by the customer no longer represents the spot price but the blend of portfolio and spot price, losing much of the short-run efficiency of marginal cost pricing.[5]

The other solution is for retailers or distributors to enter into contracts for fixed quantities of electricity at fixed prices and to offer to consumers a price that is fixed for a specified quantity, with deviations from that quantity bought or sold at the spot price. Under this pricing arrangement, consumers face the spot price at the margin, but they buy and sell relatively little at that price so cost impact is modest. This combines the efficiency of the spot price with cost and revenue stability similar to that of rate-of-return regulation. Interestingly, in the fall of 2000, the California PUC guaranteed that residential customers of San Diego Gas & Electric would pay no more than $75 per month for electricity in 2001 for the first 500 kWh per month; above that usage they would pay market rates. (Levesque, 2000, p. 14-15.) This effectively gives those customers a one-sided contract for 500 kWh per month; they pay spot prices for any excess power consumption, but do not save at the spot rate if they use less.[6] This would provide a high degree of cost protection while providing a full incentive for conservation at times of high prices.

2.3 Generation Investment

Regulated utilities, with their monopoly on generation, have traditionally forecast future demand and planned investment in new capacity to meet it. Plants are financed in anticipation of regulatory approval for the rates to pay for them. Indeed, one criticism of regulated utilities has been over-investment in facilities to minimize risks of power shortages and consumer complaints. Investment planning is made more difficult by the long lead times to plan and construct generation facilities: as much as a decade to construct a nuclear plant and as long again for environmental and planning approvals; up to five years for coal plant construction and another five for approvals; perhaps two years for construction of standard combined cycle gas turbine plants. (Applied

---

[4] One of the complaints about the California restructuring is that the distributors failed to purchase power for their customers on long-term contracts and were therefore fully exposed to spot prices. Some observers allege that the distributors were prohibited by the California PUC from purchasing power at fixed prices for their customers, while the PUC argues that it was willing to entertain proposals from the utilities for risk management, but few proposals were submitted. (Stavros, 2000.)

[5] Many variations are possible. Borenstein (2001, p. 12) proposes a system in which the retailer would purchase 80% of the required power on fixed price contracts, but charge its interval metered customers in any month the hourly price less the profit or loss that the retailer made comparing his contract price with the average monthly price. This leaves daily price variations intact, but reduces by 80% the monthly variation in electricity cost. However the customer's hourly price still does not equal the spot price.

[6] Suppose that most customers normally consume 550 kWh per month; they would be exposed to the spot price for only 9% of their consumption, and would be able to save at the spot price rate until their consumption dropped below 500 kWh for the month.

Decision Analysis, 1983, p. S-2, 3-4.)

If the wholesale market is competitive, then private investors should be prepared to invest in new generation when the forecast wholesale price will bring a satisfactory rate of return on that investment. However there is debate as to the willingness of investors in competitive generation plants to face these uncertainties and long lead times unless they can sell most of the power under long-term contracts. In any event, such investments would likely attract risk premiums that would increase with the lead time involved and increase with the fraction of the load that could not be sold at a fixed price for a decade or two. While it has been argued that competition will lower the cost of new investment by forcing investors rather than customers to bear the costs of bad investments, this argument ignores the increased risk premium those investors will demand for bearing the risk. Thus capital costs for competitive generators must be substantially greater than for a regulated utility making the same investment.[7] The experience in Australia, New Zealand, and the UK shows that private investors will build new capacity under some market rules. Indeed, it has been argued that high prices have stimulated new construction in most markets, even California, and that in 2000-2001 75,000 MW will be added to the US grid, more than was added during all of the 1990's.[8] (Seiple, 2000; Rose, 2000.)

Investors need to know the rules so they can evaluate the investment. Uncertainty delays investment. Yet restructuring takes time, during which the rules are, necessarily, unclear. Investment may stop when restructuring is proposed seriously and not resume until after market rules are finally agreed, as in Alberta. If this takes five years and the economy is growing, a supply/demand crunch could emerge by market opening. The smoothest openings of competitive markets have coincided with excess supply, while the worst were those that opened with shortages, as in Alberta and California.

If the new capacity is not small relative to the market, the price must rise above the target price, since the new capacity will depress the price once it comes on line. Over time, competitive prices in a market in which capacity investment is not small relative to installed capacity may follow a sawtooth pattern, rising until a new plant comes on line, then falling, only to rise again as demand grows.

2.4 Retail Competition

---

[7] Gordon (1996) estimated the increased capital cost for conversion of Ontario Hydro from a crown corporation to a private corporation. However his analysis focussed on changes in the capital structure and competitive costs of capital without regard to changes in the risks inherent in the unbundling of generation, transmission and retailing activities, so actual increases should greatly exceed his estimates.

[8] "New power plants are being built in markets with regulated reserve requirements - like NEPOOL - and in markets with no reserve requirements - like Texas and Western states. They are being built in regions with independent system operators, or ISOs, and in regions without ISOs. New power plants are even being built in markets with significant regulatory risk - like California - or significant permitting and environmental hurdles - like the Northeast." (Seiple, 2000.)

Retail competition requires consumers to compare price offers from competing retailers and choose intelligently among them. Yet a competitive electricity bill will price separately generation, transmission, distribution, billing and metering and other costs. The customer may or may not be able to match a quoted price to an existing line on a bill.[9] Customers may not read or understand all of the terms of a contract that is offered.[10] They may have trouble comparing offers of fixed prices to offers tied to a price index or with time-varying prices. More generally, experience has shown that small consumers have difficulty evaluating competing offers, and in consequence do not switch unless they have some assurance of cost savings of 10% or more.

Because many customers will not choose a competitive retailer unless they are paying a price significantly above the competitive level, it is important that they pay a fair price if they do not choose a retailer. One possibility would be to pass the spot price through to default (also called standard supply) customers. This was recommended (narrowly, and after vigorous debate) by the Ontario Market Design Committee (1998, RM 4-3). It gives the small customer the full advantage of the competitive wholesale price, without the costly intervention of middlemen. This seems like a reasonable proposition if one can expect the wholesale price to be reasonably stable, as would be the case in the first few years of the Ontario market because of a wholesale revenue cap for the legacy monopoly generator.[11] However the California and Alberta experience in 1999 and 2000 shows that the risks of high spot prices may be larger than even opponents of spot price pass-through imagined, at least if the market is poorly designed or the opening date coincides with a supply crunch.

So, customers may demand protection from spot price volatility. However, the distribution utility is a regulated monopolist and risk-management is a competitive business. Once the distribution utility moves beyond simply passing the spot price through to customers a host of problems arise. If the distributor may lock customers into default supply, competition is seriously constrained. If they cannot be locked in, how does the distributor ensure that customers will not depart when the spot price is low and return when it is high? How can the distributor purchase a portfolio of supply contracts for a customer base that may shift with the vagaries of the market?

---

[9] In Ontario, retailers were active in the summer of 2000 before the market opened, at which time residential bills showed an electricity price that included transmission charges and stranded debt charges. Several professors of Economics signed offers, not realizing that the price on the offer was more than 1 cent greater than the current commodity price.

[10] Few Ontario residents are likely to have understood this clause in a recent residential offer:

"In addition to the Price, the Applicant agrees to pay the Wholesaler the cost of any electricity that must be purchased by the Wholesaler to satisfy the Applicant's billing load profile should it differ from the utility's billing load profile forecast and any incremental costs incurred by the Wholesaler as a result of a failure to deliver by any of the Wholesaler's suppliers."

[11] The MDC recommended that until its share of the generation market falls below 35% Ontario Power Generation must refund to customers all revenues in excess of 3.8 cents per kWh. This allows the market price to vary hourly and daily with market supply and demand conditions, but it removes any incentive for OPG to use its market power to manipulate the price above 3.8 cents.

Yet the distributor that does not match its default load with power purchases risks the financial ruin that has rained down upon several imperfectly hedged firms in commodity markets in recent years. It is not certain that conservative local wires companies whose main duty is to keep the wires in good repair and send out accurate bills on time will be proficient in managing complex risk management systems. Restructuring jurisdictions have struggled with the competing goals of stable consumer prices and low-risk distribution utilities which leave risk-management to competitive retailers without converging on any wholly satisfactory solution.

3. Will Restructuring Lead to Efficient Prices?

3.1 Efficient pricing

Basic price theory says that when firms are price-takers they will adjust their output until price equals marginal cost, while price-taking consumers will adjust their consumption until the marginal value of another unit just equals the price. These two conditions ensure that consumer surplus and producer surplus are maximized, at least when there are no externalities and the commodity is homogeneous. Therefore marginal cost pricing should achieve efficient electricity production and consumption in the absence of transactions costs. (Joskow and Schmalensee, 1983, p. 81.) If consumers face a price not equal to marginal cost, there is a welfare loss equal to the Harberger triangle, and a transfer equal to the rectangle. See Figure 2. So, one test of any electricity market is its ability to transmit efficient marginal cost prices to consumers.

There are, however, competing considerations. First, it is more costly to meter electricity use by the hour than by the month. Second, risk-averse customers will want protection from price volatility, not because of short-term price spikes but because of the risk of prices that stay high enough long enough to substantially raise their monthly bills. Third, the complexity of anything other than fixed prices imposes costs on consumers to understand the price system so they can compare competing offers and so that they can respond efficiently. These factors mean that efficient pricing must involve minimizing the sum of lost consumer and producer surplus from mis-pricing plus the cost of metering and billing, the cost of risk-bearing, and the cost of understanding the price system.[12]

We can consider the efficiency of electricity pricing over several time frames. The very short run would examine the transmission of hourly prices and evaluate the responses of customers to changes in those prices. The short run would examine prices over a period such as a month during which time customers unable to respond to hourly prices could still respond to prices determined on a monthly or billing period basis. The medium run would examine the same issue over a year, which would allow customers to make some investment in energy conservation or energy management equipment or practices. The long run would examine prices over a period of several years during which customers could invest in demand management technology or strategies, and generators could invest in new capacity.

The evaluation of electricity pricing plans would therefore require some assumption about

[12] Joskow and Schmalensee (1983, p. 81) suggest that as a general principle "prices should reflect marginal costs, taking appropriate account of metering costs and other contractual complexities."

the pattern of variations in marginal cost over some defined scenarios. The scenarios should include typical years in which capacity is adequate and prices follow predictable patterns as well as plausible years in which capacity is short and unusual but not implausible price patterns are observed. While blackouts are rare in North America they do occur occasionally, and brownouts are not uncommon. To the extent that the risk of these events can be predicted, and to the extent that they impose costs on consumers, conditions that could cause them to occur should be included in the scenarios, to determine the extent to which various pricing plans can reduce demand when supply is short and thereby reduce the costs associated with a brownout or blackout. Because service areas differ with respect to the portfolio of generation plants, such cost variations would have to be determined specifically for the service area in question. A model of consumer demand would be needed for each customer class that would predict the demand response to any price plan when implemented in any of the marginal cost scenarios. The model could calculate the welfare costs of any deviations in price from marginal cost. The model would also have to incorporate information about consumers' and generators' preferences regarding price risk so that it could calculate welfare losses arising from price fluctuations. Finally, some means would have to be found to assess the costs to consumers of evaluating alternative price plans as a function of their complexity and to predict the extent to which they will be able to evaluate these alternatives satisfactorily.

3.2 Regulated Prices

How does pricing by a regulated monopoly match the model of efficient pricing? Joskow and Schmalensee (1983, p. 88) state bluntly that wholesale and retail power prices "are currently not generally based on marginal cost pricing principles." Regulated rates for small commercial and residential consumers in the United States and Canada are usually fixed for a year or more at a time in the form of block declining rates. Reliance on a meter that measures only the cumulative kWh used between readings, which are taken at intervals ranging from one to three months, precludes changing the price to reflect short-run cost variations. Medium and large size industrial and commercial customers use a "demand" meter that records the kWh used and records the maximum rate of use, in kilowatts (kW), between readings. The customer pays both a kWh charge and a demand charge for peak usage. This is intended as a peak-load-pricing system, although the customer's peak may not coincide with the system peak. The rates for kWh and kW are fixed in advance for periods of a year of more. The largest industrial customers, often connected directly to the high-voltage transmission grid, generally use an interval meter that records their electricity consumption every ten minutes or so. They may pay a fixed price or a price that varies with time.

Some jurisdictions also offer time-of-use (TOU) prices, especially in California which required TOU pricing for large customers in the 1970's. A TOU meter is really two or three kWh meters together with power run through one or another meter at different times of day according to a fixed schedule. In Los Angeles, for example, the summer (May through September) peak period is weekdays, 11AM to 6 PM, the semi-peak is weekdays 6AM to 11 AM and 6PM to 10PM. All other times are off-peak. In the winter, the peak moves to weekday evenings 5PM to 8 PM. The TOU meter records cumulative kWh consumption in each period at the time of meter reading. The peak period price may be four times the off-peak price or even more. The combination of TOU metering with seasonal price changes allows considerable flexibility to tailor prices to **expected** marginal costs, but not to actual costs. However, in the pre-restructuring North America, only a small fraction of all customers were likely on TOU metering.

Regulated rates generally cover the average cost of all power generation over the period for which the rates are set, usually a year or more. Hourly, daily and seasonal variations in marginal costs are lost in fixed annual rates. They are partially lost in TOU rates. There is no reason for the average cost for the year to equal the load-weighted average of the marginal costs for the year or the marginal costs of new generation capacity. In short, the variability of system marginal costs over time means that rate of return regulation will fail to transmit to customers the true marginal cost of the power they consume at any time. Joskow (1997, p. 126-7) reports that in the mid-1990's, the regulated price of electricity generation in the US northeast and in California was in the range of 6 to 7 cents/kWh, while the marginal cost of wholesale electricity was about 2.5 cents/kWh and the long run marginal cost was between 3 and 4 cents. In 2000, the marginal cost of wholesale electricity in California seemed to reach several times the regulated price.

3.3 Pricing Alternatives

In wholesale markets with a competitive structure and spot price, the hourly spot price should reflect the marginal cost of the most expensive unit dispatched and thus the opportunity cost of electricity for the hour. To what extent will this price be passed on to customers?

To pay an hourly price, a customer must have “interval meter” that records the kWh used in every hour of every day and transmits the data to the electricity supplier from time to time. While advancing technology has reduced cost of such meters and their reading, they still cost more on an annual basis than a kWh meter. As the size of the customer’s consumption falls, the potential savings from hourly metering will become too small to justify the cost of the more expensive meter. As a practical matter, most restructured markets have not required small consumers (often defined as less than 50 kW peak demand), residential and commercial, to install interval meters, and in those markets where such a meter is a prerequisite to signing up with a competitive retailer that requirement has been regarded as a barrier to competition. Restructuring usually requires wholesale customers, including distributors and large industrial end users, to replace demand meters with interval meters, and over time the maximum demand that need not install interval metering is reduced.

Whether it is economically desirable to install an interval meter and charge hourly prices rather than using a kWh meter and charging monthly prices depends on the costs of installation and operation and the welfare gains that arise from adjusting demand to the varying price. If short-run demand was completely inelastic, the welfare losses on the demand side from charging a price not equal to marginal cost would remain, and the gains from interval metering would be zero. As a customer becomes more responsive to hourly prices, the welfare loss arising from not facing the hourly price and adjusting quantity accordingly increases. Suppose that the price elasticity of demand in the very short run is -0.2, the marginal cost is 3 cents off-peak and 6 cents on peak, and the small customer’s usage is 635 kWh per month on-peak and 365 kWh/month off peak at a uniform price of 4.2 cents. Changing this customer to a real-time pricing scheme would decrease consumption modestly during the peak and increase it off-peak, eliminating a welfare loss of about 1.6% of the total revenue for the month. See Figure 2. This is unlikely to be large enough to justify the cost of interval metering compared to kWh metering for the smallest customers such as individual households. However the deadweight loss increases as the square of the price deviation. The PJM data referred to above suggest that the marginal cost in July, 1999 was triple the average marginal cost for the year of almost $30/mWh; if the marginal cost was triple the price for a month, the deadweight loss could exceed 25% of the expenditure for the month. Thus the value of more accurate pricing may depend as much on the extent of large deviations between marginal cost and price as on small routine deviations. A full analysis of the relative merit of different pricing systems would require a description of the anticipated variability of hourly marginal costs and an estimate of the response of consumers to those prices and the welfare loss associated with each pricing plan.

An intermediate step between the kWh meter and the interval meter is of course the TOU meter. The meter is more expensive than the kWh meter, but less costly than an interval meter. It is usually read manually, so it does not require a modem and telephone hookup. It allows pricing that varies according to pre-set periods of the day and that may vary by season.

Let us assume that the customers for whom interval metering is not justified will use either kWh meters or TOU meters and that the meter will be read every two months. Can we find a pricing scheme better than to charge a fixed annual price equal to the expected average cost of generation, or even the expected marginal cost of generation (the expected spot price)? There may be substantial changes in the weighted marginal cost of generation (spot price) among these bimonthly periods that could be captured in a pricing system. In many service areas the weighted bi-monthly average spot price varies significantly throughout the year. It is higher when demand is high because of cooling demands in the summer (especially in the south) and heating demands in the winter (especially in the north). It is lower in the in spring and fall when heating and cooling demands are modest. In jurisdictions with substantial hydroelectric generation prices may be especially low when river flow is at its maximum. One could capture the normal seasonal variation in price for small customers simply by varying the kWh price seasonally. Philadelphia has higher prices during the summer (June-September) than during the winter. As a seasonal pricing system would cost little more to administer than annual pricing, we should expect to see it in use where there are significant and predictable seasonal variations in the marginal cost of generation and non-trivial consumer responses to those seasonal variations.

We could imagine a range of pricing systems from a price that is fixed for a multi-year period, say five years, to a price that varies hourly with the spot price. When the customer pays the spot price we assume that they can also hedge the price risk by purchasing a futures contract for a fixed amount of power at a fixed price, or contract for differences (CFD).[13] Table 1 shows a

[13] The retailer and customer who is paying the spot price can enter into a contract for a fixed quantity of electricity at a fixed price for a fixed period of time. The contract will be settled by payment by one party to the other of the difference between the spot price and the contract price multiplied by the contract quantity. This contract is an investment in the price of electricity. If the spot price exceeds the contract price, the customer wins. If the spot price falls below the contract price, the retailer wins. The customer is exposed to the spot price of electricity for marginal purchases because the payment under the contract is not affected by actual consumption.

set of possible pricing plans. The first is a price fixed for five years covering all of the customer's electricity requirements. In a competitive market, the retailer should set this price at the weighted average expected spot price, where the expected price in every hour is weighted by the customer's expected consumption in that hour, plus a premium to compensate the retailer for administrative costs and for bearing the risks of unforseen spot price fluctuations and of variations in the customer's consumption.[14]

The second price plan is a fixed price for one year. The risk premium should be lower than for the 5-year contract because the customer assumes the risk of year-to-year price fluctuations so the retailer makes his forecast over a much shorter period of time. The welfare loss is less since the price can adjust every year, but the customer's volatility is higher for the same reason.

The third price plan is a seasonal plan, set somewhat in advance of the beginning of the year, like plan 2. It differs in that the price varies between seasons in accordance with the expected seasonal variation in the spot price. This plan transmits predictable seasonal prices to the consumer, encouraging conservation in the high price periods. To the extent that seasonal variations in the spot price can be predicted before the start of the year, this plan can capture them. Consumer and producer surplus losses are somewhat reduced from plan three, but volatility is no greater since prices are known at the beginning of the year.

The fourth price plan is a TOU plan, set in advance like plan 2. The price in each period: peak, semi-peak, off-peak, is set in accordance with the expected variation in the spot price. This plan transmits predictable prices to the consumer, encouraging conservation in high-price periods. Like plan 2, it works as well as our ability to forecast expected prices, but its performance degrades if the spot prices vary considerably even within the specified periods.

Plans 2, 3 and 4 would not have helped California deal with the supply/demand imbalance during 2000 because the very high prices of the summer and fall were not anticipated late in 1999. To reduce the welfare losses associated with unexpected shortages and price spikes, and to actually reduce the peak prices through demand response would require a system in which the price was based on actual realized spot prices, or real-time pricing as it is sometimes called. Plan 5 would compute the customer's bimonthly bill based on the customer's usage and the weighted average spot price during the two months. In this system, customers would not know the price until the end of the billing period, although it would be possible to forecast the price at the beginning of every billing period and news reports and retailers could inform consumers of expected daily prices so they could adjust their consumption accordingly. Such a pricing system would probably achieve much of the demand response that could be achieved by hourly pricing, at least for small customers, yet it would not require expensive interval meters. It would expose the

---

If consumption exactly equals the contract amount, the customer's cost is just the contract cost. If consumption exceeds the contract amount, the customer pays the spot price for each additional kWh. If consumption falls short of the contract amount, the customer is reimbursed at the spot price for each kWh of shortfall in consumption.

[14] Borenstein (2001, p. 8) reports empirical analysis showing that on average purchasers under forward contracts will not pay less than the expected spot price.

customer to the full variability of the spot price. This is essentially the price plan recommended by Ontario's Market Design Committee for customers who do not choose competitive retailers. (MDC, 1998, p. 4-5.)

While price volatility has been modest in some competitive systems, the California and Alberta price leaps in 2000 have been unpopular with those who face the spot price without hedging. Moreover it is possible that competitive investment may lead to smaller capacity reserves than a monopolist would provide, which would cause demand to intersect the sharply rising portion of the supply curve more often, increasing the volatility of prices beyond what would be predicted using the monopoly data. This concern could be met by adding a price hedge for the consumer in the form of a contract for differences (CFD) for a period of one to five years, shown in plan 6. We assume that the customer's normal consumption is 600 kWh per month and that the CFD covers 500 kWh per month. The customer is therefore exposed to the annual price for his marginal consumption, but the volatility of his electricity costs are 83 percent hedged, so his cost variability is only 17 percent of the variability of plan 5. On the assumption that the customer responds to increasing prices by reducing her consumption, the cost variability is less than 17 percent.

Why might the assumed CFD cover less than the full anticipated demand? In principle, it could cover the average customer's expected demand. However if the customer should draw less power than the amount of the CFD, and if the annual price is less than the CFD price, the customer will pay an apparent penalty because s/he is effectively buying the excess of the contract consumption over actual consumption at the CFD price and selling at the lower spot price. Customers will be unhappy enough at discovering that their contract price is above the spot price without the further insult of having to pay the contract premium for any electricity they did not consume. I therefore assume that customer satisfaction will be improved by a hedge that covers less than normal consumption so that the customer always pays more when the spot price is high and less when the spot price is low.

Plan 7 requires the customer to install an interval meter and passes the spot price to the customer on an hourly basis. It differs from plan 5 only in that the customer pays for actual hourly usage multiplied by actual hourly spot prices rather than using a load profile to allocate the customer's bimonthly usage to the hourly spot prices. Price volatility is the same for both plans, since billing is bimonthly. Indeed, effective price volatility could be slightly less under plan 7 than under 5 because the customer could reduce consumption during the highest-price hours and thereby escape the effect of the worst price spikes under plan 7. As with plan 5, the customer under plan 7 might purchase a CFD for any period from one to five years to hedge the spot price volatility.

The rate plans in Table 1 might be offered to any customer of any size, although plans 7 and 8 are unlikely to be attractive to small customers. To be certain about the relative merits of the plans would require analysis of the type suggested in section 3.1 above. However intuition suggests that since plan 6 requires no more sophisticated metering than plan 1, greatly reduces the welfare losses from plan 1 and increases volatility only modestly, plan 6 might be preferred by customers too small for interval metering, but who have a significant ability to vary their consumption in response to price. Plan 8 should be preferred by customers large enough to justify an interval meter.

3.4 Pricing after Restructuring

In many restructured jurisdictions, the regulated default or standard service price plan is a fixed price for a year or so, representing plan 2, and in some cases price caps have effectively provided fixed prices promised for several years, like plan 1. Some jurisdictions have actually passed through the spot price to default customers, averaged over their billing period if they do not have an interval meter. This represents plan 5 for small customers or plan 7 for large customers. That any jurisdiction would use plans 5 and 7 for default supply is encouraging, although when prices rose in Alberta, the government stepped in with subsidies that converted spot prices to fixed prices for many consumers. In California, retail prices were capped at \$60/Mwh, so prices that could in theory fluctuate with the spot price in real time were fixed at that price until the stranded debt was paid off. In fact, the regulated default pricing in many restructured jurisdictions is not obviously very different from what existed before restructuring. This should not be surprising since the same public utility commission is still regulating the same distribution utility as before.

Retailers in competitive markets appear to offer small customers fixed price contracts for a period of a year or so. The solicitation of residential customers in Ontario in 2000/01 has offered fixed price contracts for one or five years, representing plans 1 and 2, although commercial solicitation has offered a five-year price that would split the difference between the price in the year 2000 and the retailer's cost of acquisition.[15] Borenstein (2001, pp. 10-11) suggests that real-time pricing for retail customers of all sizes is rare.

Large industrial customers can afford to hire experts to manage their consumption and to optimize their equipment and operations to minimize their power bills. Thus the welfare gains from exposing large customers to the spot price should be greater than for small customers. Unfortunately there is little information available on the contracts faced by large customers in competitive markets because most contracts are bilateral and confidential. Even in a jurisdiction in which large customers pay the spot price, they may enter into CFD contracts with generators or retailers that reduce the expenditure risk even if it exposes them to the spot price on the margin. Many participants in the MDC process in Ontario were hostile to the concept of CFD's and argued instead for "physical bilateral" contracts, the nature of which was unclear. If medium and large customers contract for power at a fixed price regardless of quantity then the gains from marginal cost pricing will be lost. On the other hand, reports from Alberta suggest that many

---

[15] Direct Energy Marketing Limited solicited residential customers in Toronto in 2000 with a price offer of 5.75 cents/hWh for electricity for five years, with a 10 percent discount in the first year and a penalty for cancellation equal to 1.5 cents per kWh for the customer's forecast consumption over the remainder of the five years. Toronto Hydro Energy Services Inc. solicited residential customers with a price offer of 5.65 cents/kWh for electricity for one year, allowing cancellation with 30-days notice. Direct Energy Marketing limited solicited commercial customers in Toronto with an offer for five years equal to the local distributor's 2000 price less half the difference between that price and Direct Energy's acquisition cost plus a cost pass-through of up to one cent for costs of aggregation, risk management, etc. paid to an affiliate of Direct Energy. Both Direct Energy offers provide that the customer bears the risk of additional costs caused by the default of a supplier to Direct Energy.

customers who paid spot prices did not purchase CFDs and were therefore on plans 5 or 7, exposed to the full force of the price increases of 1999 and 2000. In California, the distribution utilities generally did not purchase hedges against changes in the wholesale price, leading to the fiscal immolation of 2000.

The research agenda that this suggests is to survey the pricing arrangements offered to default customers in restructured electricity markets, along with the contractual arrangements offered by retailers in those markets. Those plans, along with any of the hypothetical plans from Table 1, could then be evaluated by the model suggested in Section 3.1 above to determine the extent to which the market behaviour maximizes welfare calculated by the model assuming low transactions costs. If there are substantial divergences, further investigation could be undertaken to determine whether there are institutional barriers to offering seemingly more efficient plans. In addition, it might be possible to determine whether the price plans offered to default customers achieve the right balance between maximizing welfare and avoiding competitive activity by the distribution utility.

Pricing in electricity markets might be compared to pricing in markets for other commodities or services that cannot be stored. The default long distance telephone prices are fixed peak load rates, set for a year or more, but differing by time of day and week. This corresponds roughly to plan 4 in Table 1. Competitive suppliers of long distance telephone service offer plans some of which are flat rate at all times, and others of which vary by time of day and week, mirroring the default rate. None are responsive to actual conditions at all, despite precise metering of every call and the fact that there is some congestion on the long distance system at some times. Commuter railroads often offer off-peak fare discounts during mid-day and weekends, again with no responsiveness to real-time demand conditions. Airline fares effectively vary by time of day, day of week and season, as the proportion of low cost fares varies with demand; it is easier to find low cost seats on Wednesday at midday, or on Saturday than on Sunday evening, Friday evening or Monday morning. Natural gas prices for residential consumers can vary with market conditions, and as spot gas prices have risen during 2000 and 2001, many consumers have faced increased residential bills as distributors pass through their increased gas costs.

In summary, marginal electricity cost varies hourly, seasonally, and yearly. Regulated prices generally suppress all of this variability. Competitive markets have generally produced price plans similar to those that existed before restructuring, with customers rarely exposed to the spot price on the margin. Where customers are exposed to the spot price, there has been little hedging of the price risk. When prices are stable, these price plans may cause little welfare loss. In troubled times, however, large losses may arise from the failure to pass through marginal cost prices, and even larger losses may arise from passing through such prices in the absence of hedging. Quantitative analysis is necessary to determine whether existing markets have exhausted the gains from efficient pricing that could flow from restructuring.

3.5 The Role of Price Caps

California was not alone in imposing caps on retail prices when the market opened; several jurisdictions have capped prices or promised consumers that prices would fall. Economists have argued that the California price cap is inefficient and that the resulting problems were inevitable.

Why, then were they imposed? Joskow (2001, p. 7) says that when the price cap of $60/Mwh was set in 1996, it was expected that wholesale prices would average about $30/Mwh, so the cap was not expected to be binding. Indeed, in June, of 1998, just three months after the California market opened, the president of PG&E told a western economic conference that in addition to the 10 percent rate decrease for residential and small commercial customers that the state imposed effective January 1, 1998, "[b]y the end of 2001, we anticipate rates will decline at least another 10% to 20%." (Bilas, et al. 1999, p. 10.) But if prices will decline without caps, and caps will be damaging if they become effective, why do it?

Politically, the a price cap may act as a form of commitment in a restructuring package. If industry and government are in favour of restructuring, small consumers may be wary of it, worrying that they will not be well represented in the design process and may therefore lose out. If industry and government agree to price caps, this provides some assurance that the promised price reductions will indeed emerge. Indeed, it shifts the price risk of restructuring from small consumers to the electricity industry, with the exact locus of its residence depending on the design of the market.

The experience in California reminds us of the danger of promises that will be inefficient if they are effective. The distribution utilities in California were completely exposed to wholesale prices rising above the retail cap, and lost over $10 billion in one year as a result, impairing their credit and their ability to purchase power just when they needed it most. Moreover consumers were protected from high prices and thus not financially encouraged to conserve electricity, increasing the risk of brownouts and blackouts. This suggests a need to search for methods of commitment that will not be disastrous if they are called into action, or market designs that ensure that the commitment will not be called on. The alternative recommended by the Ontario Market Design Committee is a wholesale revenue cap imposed on the incumbent generator until the market structure becomes competitive. The revenue cap allows generation to be bid at marginal cost, but if average revenue exceeds the cap during a year, customers receive a rebate for the difference. This allows the hourly price to vary with marginal cost, but limits the total bill that consumers will pay over the course of the year. Another alternative, imposed in San Diego, is a price cap applied to a fixed quantity of electricity, with any further consumption paid at the market price. This encourages conservation, but protects those whose consumption is modest from price increases.

4. What Do Environmental Regulations Do to Restructured Markets?

It has been recognized that environmental regulations that were written specifically for integrated electric utilities may have to be amended to deal fairly and efficiently with a variety of types of generator (Dewees, 1996) and that demand side management programs that monopoly utilities have funded would likely not survive in a competitive market. These observations, while correct, turn out to have missed several important interactions between environmental regulation and the electricity sector which California, once again, has highlighted. These interactions occur in two distinct areas: environmental approvals of new generation plants and the regulation of emissions from new and existing plants.

4.1 Environmental Approval and Capacity Expansion

In most jurisdictions, the construction of new pollution sources requires environmental approval resulting in the issuance of an environmental permit along with other approvals under local zoning and/or planning laws. Any fossil-fuelled generation will emit some air pollutants, it may require water or cooling towers, and larger generation sites will require unsightly transmission lines for connection to the transmission grid, so all but the smallest generation plants will trigger an environmental approval process. The environmental approval may only require the issuance of a permit (in Ontario a Certificate of Approval under section 9 of the *Ontario Environmental Protection Act*), or it may require an environmental assessment of the proposed project, a detailed investigation that may require several years to complete. The zoning/planning approval may be straight forward if the proposed site is appropriately zoned, but it may be time consuming and costly if rezoning or amendment of (or exemption from) an official plan is required. Either approval process may allow or require notice to the affected community and a an opportunity for the community to comment or to question the proposal. In practice, either process may allow a project to be blocked if the political opposition is sufficient or if the environmental or planning problems are sufficiently serious.

The purpose of environmental and planning approvals is to allow consideration of the externalities associated with the project and an evaluation whether the externalities are warranted or whether they can be reduced and then justified. The approval process is designed to gather data and evaluate it, and to reject projects that do not meet statutory and regulatory requirements. In principle such processes may be part of an efficient regulatory regime, so that the rejection of some projects is consistent with the efficient control of externalities. However if the process is intended to balance costs and benefits of projects, however informally, to seek modification of projects where the benefit-cost ratio can be improved, and to reject projects where the costs exceed the benefits, the process must be designed so that it considers the relevant costs and benefits. In the case of electricity generation, the benefits of a project, in the form of lower electricity prices, are likely felt over the entire area of the uncongested grid which is often a very large area. The social costs are often much more local: particulate pollution may be of greatest concern within ten or twenty kilometres of the plant; the aesthetic concerns over power lines will arise in a narrow strip starting at the plant; NOx emissions may also be of local concern; while sulphur oxides and ozone arising from the NOx emissions may cause effects over many hundreds of kilometres. In general, those individuals bearing the costs and those reaping the benefits of the project will overlap only partially. This raises the risk that the approval process may not properly balance all costs and benefits, leading either to rejection of projects whose net benefits exceed the costs, or to allowing projects that benefit the general public but impose disproportionate costs on a politically ineffective minority. With possibilities of error in either direction, only a detailed examination of a particular process or a particular project can determine whether the balance is being properly weighed.

It has been suggested that California is in its present crisis, and would have been without restructuring, because environmental opposition blocked major new generation projects by the integrated utilities from 1990 to 2000. Even if true, this would not prove that there is a defect in the approvals processes in California. Perhaps the environmental harm from all proposed projects is so great that none should have been built. The test for this proposition is whether the harm anticipated from the construction of these projects would exceed the benefits of avoiding high prices. The recent public outrage over price increases that fell far short of covering marginal costs during 2000 suggests but does not prove that if they had known how high prices would go, Californians might have approved some of the rejected or postponed projects. Designing a process that would force decision-makers to balance the relevant costs and benefits would be a daunting task. However a first step to internalizing the externalities in the current California mess might be to suggest that the current retail price caps would not apply in any area that rejected a proposal for new generation capacity that met standard environmental requirements. While this would not accurately match costs and benefits, in a second-best world it might at least focus debate on the inherent trade-offs.

More generally, any jurisdiction might evaluate its approval procedures to test whether they are designed to identify the major costs and benefits of the project, to whomever they accrue, and to balance them in a reasonable fashion, reflecting the values of the society. A process that achieves this goal and does so in a reasonably expeditious manner ought to allow the construction of new generation in that is welfare-improving. If the losses are concentrated on a particular population, one might seek means of compensating the losers - perhaps a statutory tax reduction or tax subsidy payment for properties within a specified distance of a high-voltage transmission line or within the plume of greatest pollution from the stack.

What is the effect of an approvals process that slows or limits capacity expansion on the price of electricity? In a regulated monopoly, such restrictions reduce reserve capacity and may ultimately create power shortages, brownouts, or even blackouts. There is, however no significant effect on price in the short run, although in the longer run new capacity will be more expensive because of the cost of surmounting the approval hurdles. In competitive market, the reduction of reserve capacity will mean that demand curves intersect higher portions of the marginal cost curve, driving the spot price higher. This will create windfall gains for existing generators and losses for unhedged consumers. A competitive market will dramatically reveal the worsening supply situation with price spikes of increasing severity and duration. The extent of this revelation on the revenue of generators and the expenses of consumers will depend on their contractual arrangements and the degree of price hedging, but few are likely to be completely unaffected.

4.2 Air Emission Regulations

Air emission regulations can have a significant effect on the average cost and the marginal cost of operating a thermal electricity generation plant. Equally important, the magnitude of these effects depends on the form of the regulation as well as its extent. Moreover a competitive market will yield price changes in response to a given cost impact from environmental regulation that is much greater than the price change that would emerge under rate regulation. The effects of environmental regulations on product prices were not given much attention prior to the price explosion in California in 2000, except in astringent theoretical explorations (Baumol and Oates, 1998), but the revelation that limits on the emissions of nitrogen oxides (NOx) may have raised the price of electricity in August, 2000 by as much as $50 per MWh (Joskow and Kahn, 2001, p. 16) compels attention to this issue now.

Three specific design issues warrant attention. The first is the new/old distinction - the common practice of imposing more stringent emission limitations on new sources than are imposed on old sources. The second is the use of activity-based regulation, in which the amount of pollution allowed is proportional to the activity level of the polluter. The third is the use of emissions trading and the difference between two types of emissions trading - allowance-based systems, also known as cap-and-trade, and emission reduction credit, or ERC, systems.

4.2.1 The New/Old Distinction

Most air pollution regulation, in the US and in Canada distinguishes between new sources and existing sources. We adopt new source performance standards (NSPS) which must be met by any new source constructed after the date of adoption of the NSPS, while existing sources are not required to meet the NSPS and may only be required to meed some much less demanding standard or may not be regulated at all. One effect of more stringent new source regulation is to raise the cost of increasing capacity in the industry, thus raising product prices. A second effect is to delay the retirement of old sources. The owner of an old source has, in effect, a permit to pollute at a high rate, and gives up this right if it rebuilds or closes down the source. If the operating cost of an old dirty source is less than that of a new clean source, the imposition of NSPS will encourage owners to extend the life of old plants. (Hartman, Bozdogan, and Nadkarni 1979; Koch and Leone 1979.) In the electricity case the US NSPS for SO2 from utility boilers was first set at 1.2 lbs/mmBtu in 1970, yet a quarter of a century later 263 generating units still emitted more than 2.5 lbs/mmBtu. (Ellerman, *et al*. 1997, 12.)

Under rate of return regulation, a utility is allowed to recover the actual costs of operating its plant. NSPS may mean that new plant is more costly than old plant, but the utility can only set a price equal to the average of the old and new. In a competitive market, if new sources are more expensive than old, they will not be built until the market price will cover the full costs of the new source. Thus stringent NSPS should cause prices in a competitive electricity market in the long run to be greater than those in an otherwise identical market subject to rate of return regulation.

4.2.2 Activity-Based Regulation

In the United States, the new source performance standards (NSPS) for coal-fired electric power plants first enacted in 1970 limit sulfur dioxide emissions to 520 nanograms of pollutant per Joule (ng/J) of heat input, equal to 1.2 pounds of SO2 per million British thermal units of fuel heat input (lbs/mmBtu) with an added requirement that at least 70 or 90 percent of the SO2 be removed from the stack gases.[16] The NSPS for nitrogen oxides limits emissions to values ranging from 86 ng/J to 340 ng/J (0.2 to 0.8 lbs/mmBtu) depending on fuel type, with the highest limits allocated to coal combustion.[17] Canadian federal guidelines for thermal power generation

---

[16] 40CFR60.43a; Title 40 – Protection of the Environment, Chapter I – Environmental Protection Agency, Part 60 – Standards of performance for New Stationary Sources, Subpart Da – Standards of Performance for Electric Utility Steam Generating Units for Which Construction is Commenced After September 18, 1978, Sec. 60.43a Standard for sulfur dioxide.

[17] 40CFR60.44a: Title 40 – Protection of the Environment, Chapter I – Environmental Protection Agency, Part 60 – Standards of performance for New Stationary Sources, Subpart Da

emissions limit the discharge of nitrogen oxides, particulate matter, and sulfur dioxide in proportion to heat input.[18] Ontario's guidelines for stationary combustion turbines limit nitrogen oxides and sulfur dioxide in proportion to the power output of the turbine.[19] The common threads in all of these regulations are that the allowable pollution discharge depends on the amount of fuel burned and the type of fuel burned. A source that burns coal may discharge more air pollution per unit of heat in the fuel than a source that burns oil, while gas-fuelled generation is allowed the least pollution discharge.

The obvious effect of activity-based regulation is that an increase in economic activity leads to an increase in total pollution discharge. Every increase in the utilization of existing generating facilities increases total air pollution. Every new source commissioned increases the total pollution discharge, unless the construction of new sources is constrained by non-attainment of local air quality in the US. If the emission limit was chosen such that the marginal cost of pollution control was just equal to the marginal benefit of pollution control at some assumed level of economic activity, it is unlikely that those conditions would still be fulfilled as the economy grew. More likely, as economic activity grows over time, and as technological progress reduces the cost of pollution control over time, the efficient degree of control would become more strict.

A less obvious effect of activity-based regulation is that it fails to achieve the environmental result at least cost. Thomas (1980) and Helfand (1991) demonstrated that these activity-based regulations are less efficient than regulations that limit total discharge directly. The intuition in the argument is that efficiency requires that the polluter experience, at the margin, a marginal cost of abatement just equal to the harm caused by a marginal unit of discharge, and the polluter's product price should include the marginal private cost of production (including the cost of pollution control) **plus** the marginal social harm caused by pollution discharge. (Baumol and Oates, 1988, ch. 4.) If total discharge is limited absolutely, then to the polluter each unit of discharge acquires an opportunity cost, which, in an efficient system, should be just equal the marginal harm, so that the polluter's marginal cost of production includes the shadow price of the environmental discharge. On the other hand, when an increase in output increases allowable emissions, an expansion of output faces no such opportunity cost, since emissions may rise in proportion to output. The polluter's private marginal cost of production is less than the social cost, the price based on marginal cost is too low, and the good will be under-priced and over-produced.

This problem is compounded when, as is usually the case, the allowed discharge varies by type of fuel. Allowing more pollution discharge for coal-fired power plants than for gas-fired

---

– Standards of Performance for Electric Utility Steam Generating Units for Which Construction is Commenced After September 18, 1978, Sec. 60.44a Standard for nitrogen oxides.

[18] *Thermal Power Generation Emissions National Guidelines for New Stationary Sources* P.C. 1990-333. The SO2 limit for coal-fired power plants is 258 ng/J (0.6 lbs/mmBtu) or 90 percent removal, whichever is greater.

[19] Ontario Ministry of the Environment, "Guidelines for Emission Limits for Stationary Combustion Turbines" (Toronto, MOE, March, 1994.)

plants effectively subsidizes coal plants.

The inefficiency caused by activity-based environmental regulation is similar under rate-of-return regulation and under competition because in both cases there is no actual expenditure for added pollution control attendant on increased production and thus effectively no shadow price on use of the environment. In both cases, social costs will be understated and the electricity will be under-priced. In both cases, generators may be induced to use dirtier fuels.

4.2.3 Emissions Trading

Emissions trading allows firms to trade among themselves the right to pollute allocated by a regulatory regime. There are two basic types of emissions trading: allowance-based (cap-and-trade), and emission reduction credits (ERCs). Both are in use in the United States. An ERC is generated by reducing emissions below a regulated amount or historic baseline that is generally defined as an emission rate proportional to an activity level in the polluting plant. An allowance system starts with a maximum emission rate for the industry, usually expressed in tonnes per unit of time, and then allocates portions of this maximum to individual plants. The crucial difference is that with allowances the total allowed pollution emissions for the industry cannot increase with economic activity in the industry while with ERCs emissions may increase in proportion to industrial activity. As it turns out, these two emissions trading systems have substantially different effects on pollution emissions and on product prices.

Generally emissions trading is added to existing emission regulations rather than repealing those regulations. The continued existence of the old regulations helps to explain why trading has not saved as much as initially expected. Still, emissions trading can reduce the inefficiency of the new/old distinction inherent in most regulatory systems, if the emissions trading is used to substantially reduce total emissions. If substantial emission reductions are mandated using emissions trading, then the previous regulations will become less binding, and as the trading limit is reduced, it should become the dominant force for pollution control, reducing the inefficiency of the new/old distinction.

Emission Reduction Credits

Emission reduction credit trading was formalized in the EPA's 1986 "Emissions Trading Policy Statement" and the attached "Emissions Trading: Technical Issues Document."[20] This policy requires that emission reductions must be surplus, enforceable, permanent and quantifiable to qualify as ERCs. A permanent reduction is usually achieved by a change in technology that reduces emissions per unit of activity but it may arise from reduced activity so long as the reduction is made permanent by a revision of the source's permit. This means that a temporary reduction in output arising from reduced demand for the product cannot create ERCs. There are many limits on trading, the most important of which is that ERCs may not be used to meet or avoid new source performance standards. The result of this policy is that an electric utility could generate ERCs by installing pollution control equipment, permanently changing to a cleaner fuel, or permanently reducing output, including closing down (Technical Issues, I.A.1.c.(2)), but not by temporarily reducing output or substituting the output of clean generating station for that of a

---

[20] Federal Register, 51:233, December 4, 1986, pp. 43829-43859.

pollution-intensive station.

A somewhat different ERC program, Pilot Emission Reduction Trading (PERT), was developed in Canada by Ontario Hydro and other industrial, government and environmental representatives. Under PERT, ERCs can be created by "a specific and identifiable action or undertaking which is not a mere change in activity level, (e.g. due to typical business fluctuations)."[21] The quantity of ERCs created equals the activity level during the "creation period" measured in product output, such as Mwh generated, or input, such as fuel consumed, multiplied by the difference between the baseline emission rate and the creation period emission rate. PERT does not require that output reductions must be reflected in a change in the source's permit. Moreover, it implicitly rejects ERCs arising from shutdown of the source, since creation period activity would be zero.[22]

Because ERCs are based on activity-based regulations, they share all of the inefficiencies of those regulations. Even though firms can sell or buy ERCs, which creates an opportunity cost for pollution discharge, the linkage of pollution to activity means that the opportunity cost is too low. Moreover the subsidy to dirty fuels remains. Compared to an economically efficient regime, the firms will tend to burn too much dirty fuel and will charge a price that fails to fully reflect the marginal environmental harm from pollution discharge. (Dewees, 2001.) And, of course, pollution emissions rise and fall with activity, growing over time as activity grows.

Cap and Trade - Allowances

The cap-and-trade or allowance-based system is a very different form of emissions trading. Here the agency or ministry sets the total allowable discharge for the region and industry, shares of that total are distributed to the firms based on an allocation system that is independent of ongoing firm activity, and the firms can trade these allowances among themselves. Each year or period the firm must report its emissions and retire valid allowances equal to the emissions. Typically allowances may be banked from one year to the next.

The most prominent allowance trading program was established in 1990 by Title IV of the *1990 Clean Air Act Amendments*.[23] Title IV lists the power plants and the quantity of SO2 allowances each should receive, based on 1985-87 fuel consumption and emission rates of 2.5 lbs/mmBtu for Phase I and 1.2 lbs/mmBtu for Phase II. By the late 1990s, trading in the allowances created by this program had reached substantial levels, and prices had converged and stabilized. Trading appears to have saved considerable costs of pollution control. (Ellerman *et al*. 1997; Smith, Platt and Ellerman 1998.)

More recently, the RECLAIM program has used emission trading to control NO_x and SO2 in the Los Angeles basin. RECLAIM is an allowance-based trading system in which firms are given allowances based on the product of activity levels in 1989-1992 multiplied by emission factors for the current year. (Klier, Mattoon, and Prager 1997.) Pre-existing legislation provided

---

[21] Draft Rules for Emission Reduction Trading in Ontario 1996, section 2.4.1.

[22] Draft Rules, s. 2.10.

[23] 42 U.S. Code s. 7651.

for reductions in these emission factors over time.

Under an allowance based system, every ton of pollution discharged has an opportunity cost because it could be banked for the future or sold to another firm. The quantity of allowances that a firm receives is independent of its current activity, so pollution is not subsidized by output. The quantity of allowances is also independent of the type of fuel burned, so there is no subsidy to dirty fuels. Dewees (2001) shows that an allowance-based system in a competitive market achieves the efficient product price, assuming that the total allowed emissions are chosen at the efficient level. The marginal cost of generation under an allowance-based system is higher than under an ERC system because increased output consumes valuable allowances.

Under rate-of-return regulation, firms that consume mostly allowances that they were given for free (every allowance-based system distributes the allowances for free rather than auctioning them as economists recommend) will spend little to purchase allowances and will therefore fail to reflect the opportunity cost of allowances in its regulated rate. Thus in a jurisdiction that uses allowances to limit pollution discharge, a change from regulation to competition in generation should lead to an increase in the price of electricity to reflect the opportunity cost of the allowances.

Allowance Price Variability

The environmental stability of an allowance-based system can cause problems, as the California experience has showed. Suppose that you impose an emissions cap, based on equating marginal benefits of abatement with marginal costs and assuming a particular demand for the product and cost of abatement. If demand grows more than expected, firms will be forced to increase their degree of abatement, which could be quite expensive in the short run. The price of allowances could rise dramatically unless relieved by a stock of allowances in the bank. Whether this is desirable depends on the shape of the damage function. Suppose that the harm caused by pollution rose very steeply in the vicinity of the emission cap. In this case, increased industry activity causing sharp increases in abatement costs and product prices would be efficient because little increase in pollution could be tolerated. On the other hand, if marginal harm is constant over a range of emission rates, having an allowance price that fluctuated with economic conditions would be inefficient, failing to match marginal abatement costs to marginal benefits.

Again, California provides an instructive example. Studies of the harm caused by NOx emissions find values ranging from $220 to $9500 per ton (1992 US$) with an average of $2800. (Matthews and Lave, 2000, Table 1.) Let us assume that the California NOx limits were chosen on the assumption of damages of $2,000 per ton or $1 per pound. This is in line with estimates of the cost of NOx control ranging from $1,500 to $2,500 per ton. (Farrell, Carter and Raufer, 1999, p. 118.) Joskow and Kahn (2001, p. 15) report that in early 2000, NOx allowances in the RECLAIM area were selling for $1-2 per pound, equal to $2,000 to $4,000 per ton. They also report that the mid-merit gas-fired California generation plants emitted up to one pound of NOx per Mwh, so that trading NOx allowances added perhaps $1 per Mwh to the price of electricity. During the summer, the price of these allowances rose to $10 in June and to $35 by late August. While I would not want to minimize the disutility of smog in the Los Angeles basin, one could legitimately wonder if air quality was so damaging in August that it was worth $35 per pound, or $70,000 per short ton to limit it. It seems more likely that the price had in fact soared far above any amount anticipated by the designers of RECLAIM and far above any reasonable estimate of marginal harm.

No similar problem has arisen with the Title IV SO2 trading program for two reasons. First, because the SO2 problem is a long-distance long-term problem, trading covers the entire continental United States. While continental trading is inconsistent with some efficiency criteria, since the continent is not a perfectly mixed airshed, the size of the market prevents regional shifts in demand from greatly affecting prices, while banking provides further stability. But the area of southern California encompassed by RECLAIM is a relatively small economy and the electricity generation discharges a substantial fraction of its NOx. When the LA economy booms, demand for NOx emissions will rise. Second, banking is allowed from one year to the next under Title IV, but not in RECLAIM. With seasonal limits on NOx emissions and no banking substantial price volatility is not unlikely.

Does this mean that allowance-based trading is inappropriate for regional pollution control? I would not like to abandon it given its superior efficiency characteristics over ERC trading. However the California experience reminds us that the design of a regulatory program must pay careful attention to the true objective function and must allow for contingencies. Solutions to the problem have already been suggested in principle. A quarter of a century ago, Roberts and Spence (1976) proposed the use of a combination of quantity limits and an effluent charge for exceeding the limits.[24] Suppose that the emission limit has been set appropriately and that $1 per pound is a reasonable reflection of marginal harm most of the year. Suppose further that we think that marginal harm would double if total emissions increased by 25% and would rise by an equal amount again if emissions increased by another 25%. In that case, rather than keeping the statutory limit as an absolute limit, we should be prepared to expand the limit if allowance prices rise excessively. If the price rises to $2 per pound, we should be prepared to sell an additional 25% of allowances at $2 each. We should be prepared to sell a further 25% at $3 if the price should rise so high. By offering to sell allowances in limited quantities at increasing prices, we can try to match the shadow price of emissions to the marginal harm. In the California case there seems to be lots of room between $1 per pound and the $35 per pound price that was reached last August for a compromise relief system. If California had designed NOx emission relief at $10/lb of NOx, it could have cut perhaps $750 million from electricity bills in August, 2000 alone.[25]

---

[24] The principle has been proposed more recently regarding the design of global warming policies. Anticipating that countries may fail to meet their Kyoto emissions commitments, Ray Kopp, Richard Morgenstern and William Pizer proposed that countries that fell short could either buy permits on the market or pay a predetermined price per tonne of emissions. See http://www.weathervane.rff.org/features/parisconf0721/KMP-RFF-CIRED.pdf .

[25] Joskow and Kahn (2001, p. 20) estimate that with existing controls, the difference in electricity price between $10 NOx and $35 NOx would be about $35. Assuming an hourly load averaging 30,000 MW, this yields monthly consumption of 21.6 million Mwh.

5. Conclusions

Marginal costs of electricity generation have always been quite variable in many jurisdictions: hourly, seasonally, and annually, but these variations have been rendered invisible by regulated average cost pricing. In a regulated monopoly, a shortage of supply could lead to brownouts or blackouts, implying a very high shadow price of power, but it would cause no price increase. There would consequently be little incentive to conserve or to prepare to conserve. Restructuring renders that cost variability visible. As we see its magnitude, we should be more, not less, interested in restructuring and in passing those price signals to customers so they can respond appropriately.

The welfare losses associated with the blunt price signals of regulated monopoly prices may be small during quiet times when capacity reserves are ample, but they may be large in a supply crunch when demand approaches capacity and high cost generation must be run, or when costly blackouts actually occur. Welfare losses may be large when consumers who are paying low prices demand capacity increases that involve costs per kWh substantially greater than the prevailing price.

Electricity restructuring must create more volatile prices at the margin or it fails to do part of the job of passing meaningful price signals to consumers. But risk-averse customers do not like highly variable bills and generators do not like highly variable revenues. The design of competitive markets must balance the desire for efficient prices and the desire for cost and revenue stability. There may be substantial welfare gains from reaching the appropriate balance. It has not been reached in markets where most customers are on either fixed prices or unhedged and volatile spot prices.

Marginal cost pricing with modest bill volatility may be achieved by passing the spot price to customers combined with a hedge in the form of a fixed price for a fixed quantity. This can utilize a contract for differences or other mechanisms. The challenge is to embed these elements into a reasonable system of retail competition and default or standard supply that the customer can understand well enough to choose appropriately.

Electricity markets are not natural markets where we can tell the participants to act in their self-interest and rely on natural forces to achieve efficient trades. The complexity of electrical systems provides myriad ways for market participants to impose externalities on each other, so an efficient market requires myriad rules to minimize the externalities. It is not yet obvious what constitutes the best set of market rules. California is not an example of how a good competitive market works; rather it is a warning that poor design can involve failing to plan for plausible contingencies and that poor design can be enormously costly.

Environmental regulations and limitations can significantly affect the marginal cost of generation. Competitive markets highlight the effects of those regulations and may require more refinement in their design. Restructuring allows the transmission of these opportunity costs to consumers while regulated monopoly pricing does so much less effectively. With restructuring, the benefits of efficient market-based environmental regulations are increased, but the costs of inefficient regulations, including poorly-designed market-based regulations, are also increased.

The California experience in 2000 reveals potential inefficiencies in emissions trading. An allowance-based emissions trading system is only efficient if the allowed emission rate balances our concern for environmental harm and our concern for abatement costs. If we use regional trading systems; if we prohibit banking of emissions from one season to another; if activity in the polluting industry varies greatly from one season to another; then allowance prices may vary widely from one season to another and industry output may be constrained. Constraining the output of electricity can be very costly. Some relief should therefore be considered in the form of additional allowances sold at increasing prices. This flexibility will reduce spikes in the price of allowances or of electricity arising from environmental limits, while incorporating environmental damage in competitive electricity prices. Electricity prices will signal the attractiveness of environmentally friendly renewable energy, and reductions in emissions per kilowatt-hour of generation will be encouraged.

What will competition do to prices? Wholesale competition should lead to more efficient generation, higher capital costs for new generation, investment in new capacity only when the price will earn a reasonable rate of return on that new capital, and prices that are set by the marginal cost of generation. Since new generation is likely to be gas-fired CCGT, the price of natural gas will be an important determinant of competitive electricity prices in many areas. As environmental regulations become more strict, including increasing limits on NOx emissions from all thermal plants and limits on particulate and heavy metal emissions from coal-fired plants, there will be upward pressures on costs and prices. With all these factors, it is not clear that competition will reduce electricity prices below historic prices. It may not reduce prices below those that would have arisen in the absence of markets if marginal costs happen to rise after market opening. Competition promises to reduce some costs by increasing efficiency, but it promises to increase other costs by increasing risks and reflecting costs in prices. It does not promise lower prices; it only promises prices that better reflect true social and opportunity costs.

References

Applied Decision Analysis, 1983, "Analysis of Power Plant Construction Lead Times, Vol. 1, Analysis and Results," Final Report, EPRI, EA-2880.

Baughman, Martin L., Paul L. Joskow, Dilip P. Kamar, 1979, *Electric Power in the United States: Models and Policy Analysis,* Cambridge: MIT Press.

Baumol, William J., and Wallace E. Oates. 1988. *The Theory of Environmental Policy*. Cambridge: Cambridge University Press.

Bilas, Richard A., Kenneth L. Lay, Gordon R. Smith, Michal C. Moore, Robert J. Michaels, 1999, "Power Markets: Restructuring Electricity" in *Contemporary Economic Policy* 17:1, pp. 1-19. [Transcript of WEA meeting of June, 1998]

Borenstein, Severin, 2001, "The Trouble with Electricity Markets (and Some Solutions)," U.C. Energy Institute, Berkeley, POWER WP 081, Jan. 2001.

Borenstein, Severin, James Bushnell, and Steven Stoft, 2000, "The Competitive Effects of Transmission Capacity in a Deregulated Electricity Industry," *Rand J. of Econ.* 31:2, 294-325.

Dewees, Donald, 1996, "Hydro Restructuring and the Regulation of Conventional Pollutants," in Ron Daniels, ed., *Ontario Hydro at the Millennium* Kingston: McGill-Queen's University Press, pp. 169-197.

Dewees, Donald, 2001, "Emissions Trading: ERC's or Allowances?" forthcoming, *Land Economics* 77(4), November.

Ellerman, A.Denny, Richard Schmalensee, Paul L. Joskow, Juan Pablo Montero, and Elizabeth M. Bailey. 1997. *Emissions Trading under the US Acid Rain Program: Evaluation of Compliance Costs and Allowance Market Performance*. Cambridge, Mass.: MIT Center for Energy and Environmental Policy Research.

Farrell, Alex, Robert Carter, and Roger Raufer, 1999. "The NOx Budget: market-based control of tropospheric ozone in the northeastern United States," *Resource and Energy Economics* 21, pp. 103-124.

Gallon, Gary, 2001, "The Gallon Newsletter," Feb. 7, 2001.

Gordon, Myron J., 1996, "Comparative Cost of Financing Ontario Hydro as a Crown corporation and a Private Corporation," in Ron Daniels, ed., *Ontario Hydro at the Millennium* Kingston: McGill-Queen's University Press, pp. 231-255.

Ham, John C., Dean C. Mountain, M.W. Luke Chan, 1997, “Time-of-use Prices and Electricity Demand Allowing for Selection Bias in Experimental Data,” *Rand J. of Econ.* 28:0, S113-S141.

Hartman, Raymond S., Kirkor Bozdogan, and Ravindra M. Nadkarni. 1979. "The Economic Impacts of Environmental Regulations on the U.S. Copper Industry." *Bell J. of Econ.* 10 (2): 589.

Hazlett, Thomas W., 2001, “Dim Bulbs,” AEI Brookings, *Policy Matters* 01-03, http://www.aei.brookings.org/publications/policy/policy_01_03.asp

Helfand, Gloria E. 1991. “Standards versus Standards: The Effects of Different Pollution Restrictions.” *Am. Econ. Rev.* 81 (3): 622-34.

Joskow, Paul L., 1997, “Restructuring, Competition and Regulatory Reform in the U.S. Electricity Sector,” *Journal of Economic Perspectives* 11:3, Summer, 1997, pp. 119-38.

Joskow, Paul L. and Edward Kahn, 2001, “A Quantitative Analysis of Pricing Behaviour in California’s Wholesale Electricity Market During Summer 2000,” AEI-Brookings Joint Center for Regulatory Studies, WP 01-01.

Joskow, Paul L. and Richard Schmalensee, 1983, *Markets for Power: An Analysis of Electric Utility Deregulation,* Cambridge: MIT Press.

Kahn, Edward, 1988, *Electric Utility Planning and Regulation* Washington D.C.: American Council for an Energy-Efficient Future.

Koch, C.J., and R. A. Leone. 1979. "The Clean Water Act: Unexpected Impacts on Industry." *Harvard Environmental Law Rev*. 3: 84-111.

Levesque, Carl J., 2000, “California Price Wars,” *Public Utilities Fortnightly*, Oct 1, 2000, p. 14-15.

Linden, Henry R., 2000, “Fossil Fuels and Energy Policy: Understanding the new Natural Gas Economy,” *Public Utilities Fortnightly*, Nov. 15, 2000, p. 56.

Newbery, David M., 1998, “Competition, Contracts, and Entry in the Electricity Spot Market,” *Rand J. of Econ.* 29:4, 726-49.

Nikiforuk, Andrew, 2001, “Power Trip,” *Report on Business Magazine* March, 2001, p. 41.

Market Design Committee. 1998. *Second Interim Report of the Market Design Committee.* Toronto: Market Design Committee.

Mathews, H. Scott and Lester B. Lave, 2000, "Applications of Environmental Valuation for Determining Externality Costs," *Environ. Sci. and Technology* 34:8, pp. 1390-1395.

Roberts, Marc J. and Michael Spence, 1976, "Effluent Charges and Licenses under Uncertainty," *J. Public Econ.* 5, pp. 193-208.

Rose, Judah, 2000, "Price Spike Reality: Debunking the Myth of Failed Markets," *Public Utilities Fortnightly* Nov. 1, pp. 52-56.

Seiple, Chris, 2000, "Oversupply Ahead," *Public Utilities Fortnightly* Oct. 15, pp. 10-12.

Stavros, Richard, 2000, "Risk Management: Where Utilities Still Fear to Tread," *Public Utilities Fortnightly* Oct. 15, pp. 40-47.

Thomas, Vinod. 1980. "Welfare Cost of Pollution Control," *J. Environ. Econ. and Management*. 7 (2): 90-102

Wolfram, Catherine D., 1998, "Strategic Bidding in a Multi-unit Auction: An Empirical Analysis of Bids to Supply Electricity in England and Wales," *Rand J. of Econ.* 29:4, 703-725.

Table 1
Rate Plans, Efficiency, Volatility

Plan[26]	Description	Welfare Loss	Volatility
1. 5-year fixed	Expected average spot price, 5 years	max	0
2. 1-year fixed	Annual expected average spot price	med-high	medium
3. 1-year seasonal fixed	Set annually at expected prices to capture seasonal variations	medium	medium
4. 1-year TOU fixed	Set annually to reflect prices during 2-3 specific times of day.	med-low	medium
5. Bimonthly spot price pass-through	Set bimonthly ex post at actual average spot	low	high
6. – with CFD	Same as 5, but with CFD for 1-5 years	low	17% of high
7. Spot price	Hourly prices, interval meter	0	high
8. – with CFD	Same as 7, but with CFD for 1-5 years	0	17% of high

TOU time of use metering in which price can vary by time of day.
CFD contract for differences. Assumes that the CFD covers 5/6 of usual consumption.

[26] Assumed normal monthly consumption = 600 kWh. CFD is for 500 kWh per month. Meter reading assumed to be bimonthly. All plans require only kWh meter except #7, 8 which require interval meter that records hourly consumption.

Figure 1
Generation Cost and Demand

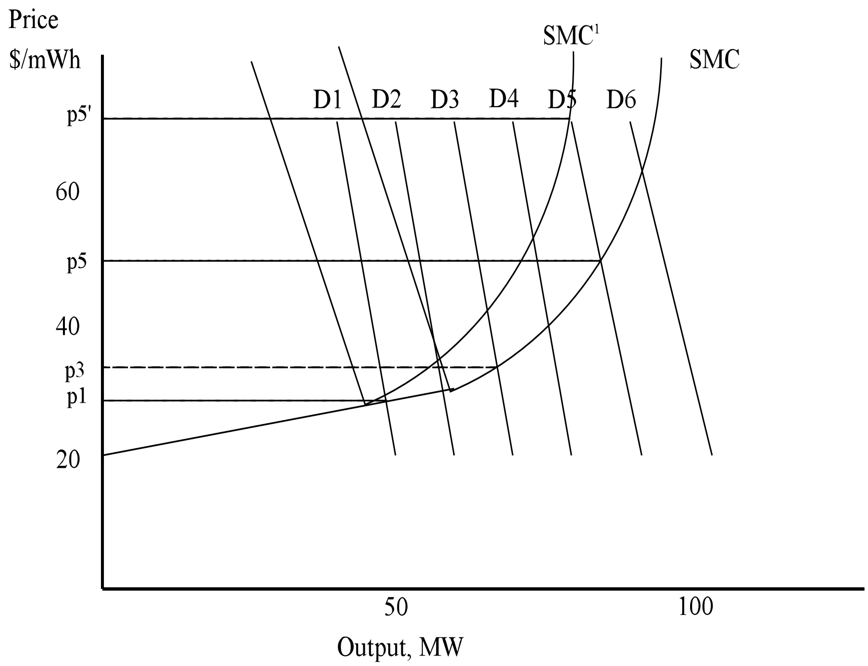
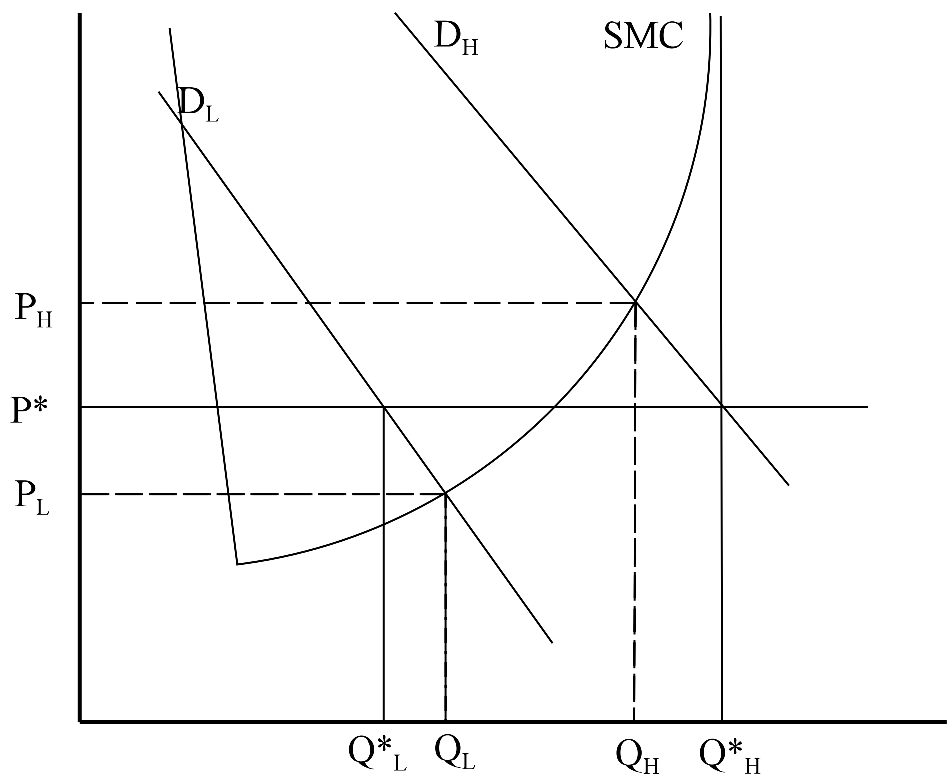


Figure 2
Lost Consumer and Producer Surplus from Mis-Pricing



P* is the average-cost price set by the utility.

Off-peak demand DL is Q^*_L at P*.
Off-peak demand DL is Q_L at P_L.

Peak demand DH is Q^*_H at P*.
Peak demand DH is Q_H at P_H.

Welfare loss depends on elasticity of supply and demand.