

## **CALIFORNIA: ENERGY CRISIS OR MARKET DESIGN CRISIS?**

Vernon Smith, Stephen Rassenti and Bart Wilson  
International Foundation for Research in Experimental Economics and Mercatus Center  
George Mason University

Electrical energy as a commodity has the uncommon feature that the typical consumer does not know how much he has purchased and what it will cost him until the end of the month when he gets his bill. If he should go to the trouble of looking at his meter at 3 p.m. on a hot summer day with his air conditioning on high, and his clothes dryer cranking away, he will note that the little monitoring wheels are spinning much faster than when he inspects it at 3 a.m. in the morning. This comparison would give him the barest of clues as to his time-of-day consumption rate. If the consumer lives in an apartment complex, her choice to move there may have been influenced by the manager's advertisement stating 'free utilities,' which means that there is but one meter for all the apartments in the entire complex. In both examples the time and seasonal variation in the cost of energy is averaged across all the hours of the month, and in the second the private cost is also artificially socialized across all consumers who dip into a common pool of available energy.

This is an incentive nightmare -- a disaster in waiting, caused by the failure to appropriately deregulate retail entry and prices along with wholesale entry and prices. The crisis in California is not one of energy; it is a crisis in bad market design.

Here is the crux of the so-called California energy crisis, signaled loud and clear by the earlier wholesale price spikes in the Midwest, South and East in the summer of 1998, continuing into 1999, 2000 and likely to persist in 2001: no significant portion of the time rate of demand for electricity is responsive to the extreme time rate of variation

in its supply cost. A porch light left burning in the daytime and people riding in elevators are afforded the same priority for energy. If retail utilities get some relief from the cost squeeze in 2001 it will be a consequence of previous high wholesale prices having attracted new peaking capacity; it will not be due to the political outcry, the price ceilings, or the threat of punishing suppliers by making them give back some of their “unjust riches.”

How can this be happening in the long-arrived electronic age of cheap sophisticated switching, sensing and control devices capable of selective programmed interruption of power flows to particular circuits and appliances, conditional upon almost anything: time of day, price, temperature? Responsive demand can be implemented voluntarily by contract between the customer and some energy provider/wholesaler, or directly by the consumer who buys a programmable load management device. In this paper, we are going to try to explain some of the technical problems in electricity, and why its delivery to the end user is trapped in yesterday’s technology and cultural mind-set, surrounded by unrealized opportunity for gains in investment efficiency, energy efficiency, and for a reduced environmental emissions load.

### **The Dynamics of Consumption and Supply Cost**

The electrical power industry has some key features that help to illuminate its current market design problems:

1. All local regions exhibit a natural daily cycle in consumption starting at its lowest level in the early morning hours and rising to a peak sometime in the mid to late afternoon, with peak levels of consumption two to three times the off-peak level.

2. Supply side investment has adapted to this cycle by providing three basic generator types:

- Large capital-intensive, efficient, but inflexible *base-load* generators that operate with very low energy costs. These are nuclear and fossil fuel units that are not designed to be ramped up and down with changes in the consumption load. They satisfy almost all of the off-peak daily consumption. Most people do not know that off-peak power sometimes fetches spot prices at or near zero. This was common in the California market, and is typical of spot markets all over the world.
- Higher energy cost, flexible *load following* generators kick in to supplement the base-load units as the daily load increases from its early morning low. These generators are designed to ramp up and down with the jiggles and weaves in consumption, and this essential ramping flexibility is why they cost more to supply a unit of power. Flexibility costs money.
- Before, during and following a consumption peak, still higher cost *peaking* generators are brought into service, if and as needed. These tend to be smaller quick-start generators that serve most of the time by waiting. They provide emergency reserves and peaking power. The energy cost of these generators is high when committed, but their readiness (capital and maintenance) cost is less than would be incurred with increased, but more often idle, load follower capacity. When peak load is small or modest, no peaking capacity may be needed.

3. Consequently, as the consumption load cycles each day from low to high, there is a corresponding, but larger amplitude cycle in the marginal cost of generation. Energy cost alone can vary by a factor of 6 or more from low to high demand. The purpose of deregulating wholesale markets is to allow wholesale prices to reflect these natural variations in marginal cost thereby enabling such prices to efficiently guide economic decisions. But without also deregulating entry and prices at the retail level, only producer, not consumer, decisions will be able to adapt to the pricing structure that emerges.

4. Note carefully that if it were not for the diurnal consumption cycle, and if total consumption were flat throughout the day, electricity would be much less costly. High generation costs are induced by the dynamic pattern of daily load variation. Similarly the cost of transmission and distribution is higher because of this pattern. *The capacity of all electrical facilities and their investment cost depends on peak, not average, consumption.*

*But the legacy of regulation is to average everything.* Only competition, induced by free choice among alternatives freely offered, can set the stage for a discovery process

that can undermine the natural tendency toward average cost thinking. Competition drives every hotel/motel manager to slash prices below average cost off-season and on weekends, and to increase prices above average cost in-season, during the week, on holidays, and during annual special events; they do this even if they are average cost thinkers. No one has to tell them, or teach them, and no regulator has to order them, to do this because they get immediate negative feedback if they depart from this pattern. Competition--the search for profit niches--is the reason why you can fly more cheaply on a 14-day advance nonrefundable airline ticket with a Saturday night stay-over, why you can sometimes fly more cheaply with a 605am departure than an 8am departure, and why the ticket may cost you 4 to 10 times more if you want to fly on the same day you purchase it. This competitive price structure fills the airplane seats on most of the flights. Before deregulation and the discovery of this price structure the flights were moving lots more empty seats on weekends. Competition discovers opportunity cost price structures by trial-and-error experimentation and imitation. Reason alone (without test bedding in the market) is unlikely to find simple and ingenious conditions, like the Saturday night stay-over rule for the economy rate, because thinking is not driven by opportunity cost considerations.

Most people will tell you that it is not 'fair' to pay more for power at 3 p.m. than at 3 a.m.; to pay more for a flight ticket if you do not stay over Saturday, more if you are leaving at 8am than at 6am, or because you are going home to see grandma on Thanksgiving; or to pay more for a hotel room in the peak demand season. But it is the other way around. *It is unfair for peak users not to be charged the full dynamic cost of their decision to consume on peak.* Any other policy will force off-peak users to

subsidize the peak users. If everybody pays the same price averaged over the cycle, then the off-peak users pay more than the cost occasioned by their consumption, the difference being an implicit tax. Similarly, the peak users pay less than the cost of their consumption, the difference being a hidden subsidy financed by the tax on off-peak consumption.

Conduct the following mental experiment in the airline and hotel cases: imagine that a flat average cost price is charged independent of day of week, holidays and season; then there would be a shortage of airplane seats and accommodation rooms at all peak demand times, more airplanes and more hotels would have to be built and this extra capacity would be idle at all other times. All customers able and willing to consume more off-peak, if they could save money, would be forced to help pay for the idle capacity.

Under retail electricity regulation, monopoly protection from competitive entry allowed flat time prices to be imposed by average cost thinking, operating under an adversarial political process that is not well suited to experimentation. This led to investment in enough capacity to meet peak demand at prices that artificially subsidize that peak demand, tax the off-peak demand, and hence distorts investment insofar as consumers would change their demand pattern if they had to pay the true cost at the time of consumption. Without competitive demand-responsive pricing at the end-user level, too much of the capacity is idle most of the time, and the mix of generator types is inefficient. *No one* can say what the capacity level and mix should be because it is *not* a planning problem; it is a market decision problem to be based on millions of bits of information not given to any one mind. It is planning by people who did not and could

not possess all such dispersed information, and driven by bad incentives, that caused the problem in the first place.

Some well-meaning officials have suggested that the California experience shows that the energy crisis is here to stay, and that we need more supplies of energy because you cannot get more for less. Shades of the bogus energy crisis of the 1970s! You *can* get more for less: this is precisely what the airline and hotel/motel industries do when they price seats and rooms to better fill them off-peak, and as a consequence utilize capacity and operations more efficiently over time. The market will determine whether we need more energy supplies after retail competition has filled in some of the idle capacity intervals.

Others have blamed the crisis on California market rules that limited the use of long term contracting for power. *But long term contracting is just another form of averaging the cost over time.* It seeks smooth prices without giving consumers an opportunity to adjust their consumption decisions to the time rate of costs they impose on the supply system. You must begin with a robust and well-structured two-sided spot market for power, and then allow people to engage in whatever financial arrangements derived from that spot market that best suits their circumstances. Its the spot market that allows intelligent long term contracts to be struck. But you have to begin with the horse, and follow him with the cart.

A related issue is the concern and emphasis that has gone into the development of electricity market derivatives--futures, options, etc. Just as long term contracting is not the crucial issue, neither are these instruments. All of them provide means for hedging, smoothing or averaging the risk of unpredictable wholesale price changes. They are fine

and desirable for dealing with any price volatility in the wholesale market that is left over after you get the retail markets deregulated. Derivatives are not a substitute for exposing retail customers to the realities of time-of-day variations in the real cost of producing power, letting them choose the technologies or contracts that realize the amount of conservation and demand responsiveness that suits their circumstances, and allowing the capacity of all components in the system to adjust to the smoother real time demand.

### **How Did We Get Here?**

The regulation of electrical utilities began about 85 years ago as, one-by-one, the States moved to grant each local area an exclusive franchise monopoly to serve its local electricity customers. For most of that history the industry had neither the technology nor the competitive motivation to implement demand responsive time-of-day pricing. If they had had this competitive motivation, then to the extent that demand is the mother of invention the technology would have been forthcoming sooner. The adversarial process for justifying prices and investment under regulation based on an average rate of return on investment created a cultural mind-set on the part of managers, regulators and customers. The cultural mind-set of management is that they face what they are fond of calling--so revealingly--their “must-serve” demand, one that they are “obligated to serve” at a fair cost-of-service average price. The regulators, who implemented this policy, share and unwittingly helped to create this mind-set, while for the retail customers it has become an entitlement. Protected from entry and consumer choice among alternatives, the managers never were guided, as were the hotel/motel managers, by the competitive search for a profitable return, to slash off-peak rates, raise peak rates, and adjust investment in capacity to consumer decisions to conserve their on-peak consumption.

The irony of it all is that in California the “obligation to serve” at a fixed price could not even be implemented at times of severe stress because the unresponsive demand exceeded energy supply, and the shortfall was met by rolling blackouts that stranded people in elevators. *There is no such thing, under regulation or deregulation, as a “must serve” demand whose satisfaction can always be guaranteed.* When peak demand strains supplies, because the reservoirs are low and/or the temperature is high, the wholesale price paid by the distributors in California (and earlier in other states) substantially exceeded the fixed regulated price at which the power was resold to customers, sometimes by a large factor. It’s called buying high and selling low, and if it persists, bankruptcy is inevitable! Thus, during the week of June 26 the wholesale spot price in California leaped as high as \$1.10 per kwh, but the power purchased was resold to customers at around \$0.10-0.11 per kwh. (Incidentally, \$1.10 is not the record, which apparently occurred earlier in the East when the spot market hit \$7.50). The way to make gobs of money (reduce these huge losses) in this environment is by selling *less* to customers by offering a discount if they consume less. But the idea of making more money by selling less power is a way of thinking that does not come naturally to an industry conditioned by the “obligation to serve,” under a regulatory process in which a decline in the average profit return below the approved regulated rate invites application to the regulatory commission for an increase in the average price charged customers. According to press reports California utilities lost \$8-9 billion trying to avoid blackouts. A miniscule fraction of this sum invested in voluntary time-of-day interruptible technologies would have disciplined prices and avoided the blackouts. We summarize the relevant behavioral evidence below.

## **How Much Interruptible Consumption Will Discipline Prices?**

Of course no one knows the answer to a question that will be determined by a few million market decisions, but we know from spot markets the world over that wholesale prices are very sensitive to small changes in peak demand, and very insensitive to changes in demand off-peak. Elsewhere ([Regulation](#), 2001), we have reported the results of laboratory market experiments using profit-motivated subjects who compete as wholesale electricity sellers and buyers. We compare prices and volume in a wholesale spot market under two alternative treatment conditions: (1) a one-sided market in which only sellers (five) actively bid to supply demand; (2) a two-sided market with demand side bidding in which buyers (four) bid to buy power simultaneously with sellers who bid to supply power. Demand each trading day moves through a cycle of four phases including a peak and an off-peak phase. A key design feature is that a relatively modest 16% of retail peak demand can be interrupted at discount price steps below the fixed average retail rate. Hence, 84% of peak demand is unresponsive so-called “must-serve” demand. Without demand side bidding, power is interrupted by automatic schedule when the market clearing supply side bid price rises above a discount price step. With demand side bidding the buyers are free to strategically underbid their resale demand schedule, including the discount steps, in an effort to discipline any attempt by sellers to raise the clearing price. In doing this, they deny load to the marginal generators attempting to increase the selling price.

Comparing data from eight independent experimental groups, four run under each of the two bidding conditions, the results are dramatic: demand side bidding yields lower

prices, and upward price spikes that jump by 50-100% under supply-side only bidding, are effectively eliminated. *There is no substitute for rules that empower buyers to defend their own interest through the competitive market process.* Because of the great sensitivity of prices to peaking generator supply, interruptible demand need be only a modest proportion of the peak demand.

### **How Do We Get From Here to There?**

The key policy need is *to remove restrictions on the entry of competing retail energy suppliers* by giving them the right to bypass the local distribution supplier who then becomes a protected monopoly supplier of wires services only, not also of energy. Why not a simple rule that passes the wholesale spot price through to all end users? Because consumers will vary in the extent to which they are willing to be interrupted, on what conditions, and at what prices. They need to be prepared, technologically and psychologically for supply cost pricing. If a company is cooking microchips, there is no way that it can afford interruption, but a household may be perfectly willing to run a clothes dryer when the prices are lower. Some may be willing to pay a market determined premium to satisfy all their demand; others may be willing to pay no such premium and be freely interrupted to always get the lowest available spot prices; still others will fall in between these extremes. We do not need a new one-size-fits-all decree from the top.

Some have argued that when consumers have been offered the opportunity to enroll in time-of-day pricing schemes offered by their local franchised monopoly, they have shown little interest. We suspect that this is because they were not offered inducements that corresponded more closely to the savings in supply cost. Judging from

the reaction of the customers of San Diego Electric, when for a short time their management passed through the full wholesale price, many customers would conserve at peak prices up to 10 times the average retail rate, which were a far higher multiple of the off-peak rate. If in fact consumers were indifferent to discounts and premiums, then of course there would be no reason not to pass through the wholesale rate, once this fact is established.

Opening up entry will make the full range of what are called distributed generation resources available. These are capital resources that are located at or near the end-use consumer; they include the switching, metering and control technologies that enable better utilization of energy over time, but also alternative power sources--micro-turbines, solar, fuel cell, etc. that bypass the transmission and distribution wires. Competing suppliers will find the various technological niches into which various end-use consumers will fit; uncontested planning will not find these niches, as the history of the world has made plain thousands of times over.

Remember when Ma Bell would not let you connect any telephone except her issue, and none but her service agents could have access to the telephone circuits in your house? You may even remember an earlier day when she only issued Ma-Bell-Black phones--no choice. The excuse for all this was that--*in your interest*--Ma had to protect the integrity and reliability of the network, and some regulatory economists of the day mumbled phrases like "network externalities." Welcome to the unpleasant memory of telephone regulation. Today, the same culture rules your friendly local power distribution utility, which, in pursuing his interest, not the consumers, has steadfastly opposed all deregulation.

To summarize, we need to deregulate both entry and prices in our highly regulated retail power markets, and allow technology to make consumer demand as responsive to prices as consumers choose to make it when exposed to the reality of time variable supply cost. To the extent that consumers choose to switch from higher cost peaking power to lower cost off-peak power, a change in the pattern of investment in generator types will be encouraged, including investment in distributed resources and technologies located nearer to the end-use consumer. Power losses in transmission and distribution will be decreased, peak transmission constraints will be reduced, and the more efficient pattern of consumption over time will lower carbon and nitrogen emissions. Until this policy is implemented, no one can say whether we have, or will ever have, an energy crisis, because the policy enables the market to provide more for less by filling in the gaps where we now have idle capacity.

### **Reference**

Rassenti, Stephen J., Vernon L. Smith, and Bart J. Wilson (2001). "Turning Off the Lights," Regulation, forthcoming.