Electricity Markets, Reliability and the Environment: Smartening-up the Grid*

Richard E. Schuler
Dept. of Economics and School of Civil & Environmental Engineering
Cornell University
Ithaca, NY 14853
Res1@cornell.edu

Abstract
The opportunities and challenges of implementing a smart electric supply system are discussed, including how to integrate the local public goods of electric system reliability, and increasingly of environmental compliance, in an efficient way. The full potential of the smart grid cannot be attained, however, without the wide-spread availability of real-time pricing for electricity so that all parties, particularly customers, have an incentive to play “smart”. The incentives for behavioral, technological and environmental innovations resulting from a smart grid are discussed, if implemented in a market context, as are institutional and political inhibitions for doing so under a regulatory regime.

1. Introduction
First we had “Smart Markets” for electricity; now we’re being pushed toward the “Smart Grid”. While in the end, it’s highly desirable to be smart, if by smart we mean being able to produce more of what’s valued at a lower cost without precipitating a rebellion by some significant group in society, getting there is not easy. That’s because the current way of providing electricity combines private and public goods in complicated ways, and the institutions that have evolved to supply these services have been loathe to wake up individual customers to the potential power in their purses. In fact many suppliers, particularly those with regulated prices, and some government officials are deathly afraid of granting that power to consumers since it may challenge severely their comfort and security. With ever-increasing numbers of public concerns being imposed on the electric supply industry, from local to global environmental improvements, from renewable to sustainable use of resources, all while maintaining public health and safety, visual amenities, low cost supplies, plus the essential requirement of any power system for a modern society - that it be utterly reliable - compound the daunting challenges facing a smarter grid. And yet, the only hope for achieving these multi-faceted objectives is to motivate the greatest level of intelligent behavior by all participants in the system - suppliers, users and regulators, alike.

2. Today’s Industry and Smart Markets
This industry was dragged (not quite kicking and screaming) into the twentieth century near its end by a restructuring, in some regions of the country, which was designed to arrange wholesale supplies through market-like mechanisms. As emphasized over twenty years ago by Hobbs and Schuler [1] wholesale markets were most likely to be effective in allocating electricity efficiently in those regions of the country with a high spatial density of demand and a grid configuration of transmission lines (like in the northeast). They also estimated that markets would be less effective in sparsely-settled regions with radial transmission systems like vast rural areas in the west. The difference arises primarily because the higher transportation cost in these latter circumstances limits the number of competitive suppliers who are available to reach each customer economically. In fact in the 1930’s it was the formidable cost of constructing transmission and

* This work was sponsored by the Power Systems Engineering Research Center (PSERC) and the Consortium for Electricity Reliability Technology Solutions (CERTS). The support provided by PSERC’s industrial members, by the National Science Foundation’s Industry/University Cooperative Research Program and by the United States Department of Energy is greatly appreciated.
distribution lines, per customer, in those sparsely settled regions that led to government’s underwriting large portions of the cost of extending electricity supplies to those locations.

In today’s debates about the effectiveness of wholesale markets, many of the opponents come from those same less densely settled regions of the country or where public power is prevalent, and they may still be right for their portions of the country, but they are dead wrong about the other regions where wholesale markets are flourishing. The problem is that any discussion about the efficacy of market-based supplies of electricity quickly deteriorates, like arguments over politics or religion, without acknowledging that different solutions may be best for different locations. The nation would benefit from a uniform nationwide solution to its supply structure only if electricity were a pure public good whose shared impact spanned the continent. But electricity is not a pure public good as demonstrated by the fact that its use is metered, and each customer is charged in some relation to their individual consumption, in all regions of the country. What do differ by region are the cost characteristics of supply, and that suggests that different prices and/or institutional structures for supply should be abundant.

What has become a public good aspect of electricity supply is our utter reliance on its instant availability at the flip of a switch. As pointed out by Mount, Schuler and Schulze [2], it is the protection against unanticipated outages that is a public good, but it is a local public good in the sense that only the immediate neighbors who experience the same unplanned outage when a local line or network fails that no longer have the proper incentive to reveal truthfully their desire for reliability through market mechanisms. That is because any single neighbor who pays for upgrades and redundancy will automatically have that improvement extended to nearby neighbors who are served from the same lines. In anticipation of their neighbors stepping forth, many therefore will understake the value of reliability to them in order to “free ride” on their neighbors’ beneficence. This is why FERC has established an Economic Reliability Organization to set reliability standards that are in the collective public interest and to ensure that they are enforced. And note that an interruptible rate incentive is not the same thing, although it can relieve stress on the system in anticipated tight situations. The difference under interruptible rates is that the participant is warned ahead of time and is usually given an option of not complying if the reduction of service seems too inconvenient. By comparison the cost of unanticipated interruptions is usually much greater for most customers, and it is signaling the willingness-to-pay to avoid this type of unplanned outage that takes on public-good, free-rider characteristics.

And it is because of these local public good aspects of the grid, combined with several other unique features - - the facts that 1) electricity can’t be stockpiled for a rainy day, therefore 2) parallel paths and duplicate generation facilities must be available real-time as the only means of providing reliable service and 3) electricity flows according to the laws of physics, not the rules of commerce - - requires that market-based transactions must be “optimized” in order to clear the market reliably. Yes, bid curves are arranged in descending order and offer curves in an ascending sequence just as they are in any other market exchange, but which offers are accepted and which bids are satisfied are subject to all of the physical and reliability constraints that exist on the system. This is why Vernon Smith [3] dubbed these electricity markets as “smart”: someone needs to optimize the dispatch (market-clearing) to serve the customers and to satisfy the constraints at the lowest cost to the buyer. And given the complexities of the system and the needed speed of clearing, a computerized optimization algorithm is essential.

3. Add Environmental Concerns and Sustainability

In addition to reliability there is another public good whose regulation affects and is affected by the reliable, least-cost dispatch of the power grid: the quality of the environment, both nearby and increasingly, afar. In most cases, separate regulatory agencies set and monitor compliance with electricity system reliability and with regional environmental quality; yet both public goods are delivered as a result of the operation of a single system, the power grid. With increasing public awareness of environmental costs, including their effects on public health and amenities, there is increasing pressure to strengthen environmental standards further in many locations.

Fortunately the increasing use of price by the environmental community, either through cap-and-trade mechanisms or the levy of effluent fees, would seem to offer a nice complement to existing market-based allocation mechanisms for the wholesale supply of electricity. And, even in areas with regulatory allocations of electricity supplies, the addition of market-price-based environmental adders to the cost of generation can reflect the differential external (non-marketed) environmental costs of
supplying power from different sources; the externalities are internalized! Thus power plants burning high sulfur coal without scrubbing will have those social costs of pollution added to their modest fuel costs in determining whether or not they are economical to dispatch.

But, an important distinction needs to be made between the emission of greenhouse gases, as compared to particulates and sulfur and nitrogen oxides. Eventually, the adverse effects of greenhouse gas emissions are experienced by everyone on the planet; their reduction is a pure public good in the economist’s jargon, and the emissions from each source add-up to what is received collectively by all. By comparison, the other pollutants affect different people in different locations differently, depending on the way the wind is blowing, other atmospheric conditions and topography and geography. The reduction of these “criteria” pollutants (SOx, NOx and particulates) is a local public good (just like electric reliability) since everyone in the same neighborhood receives the same benefit; what’s different from greenhouse gases is that different neighborhoods can receive different annual average levels of these criteria pollutants unless the generation from each unit is managed to equalize the impact across locales (Note, this is exactly what’s done to equalize electric supply reliability.). So the “transport paths” by which these criteria pollutants are distributed from generators who emit them to the customers who may experience adverse impacts is important to consider. But, those transport paths for pollutants are usually very different than the flow paths over the network by which electricity is distributed from generators to customers.

The problem that arises in managing these criteria pollutants is that the agency setting the emissions caps (or effluent fees) is different from the entity that is obligated to dispatch the power in order to maintain reliability, and because the electricity and effluents travel different routes and usually land in different places (plus those transport paths will change, depending upon atmospheric conditions in the case of pollutants and upon changing demand patterns, particularly when lines become congested, in the case of reliability). It is easy to imagine the consequent tugging and pulling between the respective oversight agencies, as emphasized in a recent paper by Mitarotonda [4]. As an example, suppose the NYISO dispatcher increases generation by a 10 MWs in Newburgh, NY to maintain reliability in New York City, but as a result a coastal breeze from the south dumps more criteria pollutants on Albany that is experiencing severe ozone conditions. As a result, the environmental regulators in NY state might raise the emissions prices imposed on the Newburgh generator dramatically, causing the generation reserves needed for NYC to be transferred to an oil-fired generator across the Hudson River in New Jersey, which dumps more particulates and NOx on both NYC and northern New Jersey, which causes a different air-shed monitor to wake-up and reduce allowed emissions, which has an impact on electricity dispatchers in PJM - - and so on. In addition to a much smarter grid and an integration of environmental costs in real time with the other costs of supplying electricity from different sources, solving this combined energy and environmental resource allocation problem at the least cost to society requires a degree of collaboration and cooperation between environmental and electricity supply oversight agencies rarely seen before. Nevertheless, it is important that if the lowest cost way of supplying reliable electricity to New York City is to ramp up a New Jersey generator that dumps its emissions in northern New Jersey, that the cost of those environmental impacts on New Jersey be faced by New York City’s electricity customers (Presumably, NYS’s ISO and its Department of Environmental Conservation will consider the impact of increased generation in NJ on NYC’s air quality.) .

But there are signs of hope for cooperation. In many jurisdictions, the subsidies paid at the instigation of environmental agencies to the developers of renewable resources have been in the form of a lump sum or a flat per MWh subsidy applied to the locationally different prices charged and paid by the electricity system dispatcher to generators that use renewable resources. By comparison some developers of renewable resources have requested guaranteed minimum payments for their generation, regardless of their location, but this type of guarantee would provide no incentive to the renewable generators to locate (in cases when they have some discretion) where their supplies are of the greatest value to the entire system (the end-use customers). By comparison, flat per MWh subsidies maintain the locational differences in prices that arise from electrical system congestion, and therefore result in a lower overall cost of electricity supply while meeting society’s objective of moving toward renewable sources of generation.

The other aspect of most renewable sources of energy that are used to generate electricity, like wind and solar, is that frequently their peak availability does not coincide with the periods of greatest customer demand for electricity (the key exception being the effect of solar load on air-conditioning demand). So in addition to properly differentiated prices for electricity by location, there needs to be a
fine-grained differentiation of prices by time-of-day so that potential suppliers of these renewable sources optimize the location and amount of generating capacity they install with respect to their buyers’ demand. But buyers too need to see and pay for their electricity by location and time-of-day if renewable resources are to be fully developed economically. That’s because, given a choice between paying a high price for gas-fired generation, or next to nothing for electricity when the wind is blowing or the sun is shining, some customers may be willing to switch their usage patterns. It’s all a matter of economics and the first requirement is that all parties be informed of the consequences of their actions.

4. Enter the Smart Grid

Like the term, sustainability, the smart grid means something different to nearly everyone who utters the phrase. To some, it implies the application of the latest sensors and computerized algorithms to nearly every aspect of the operation of the grid so that everything is on a special type of auto-pilot, able to anticipate and respond to difficulties and imbalances automatically and more-rapidly than any human operator could hope to respond. To others the wave of change extends down to electric distribution so that that part of the system is also designed and operated with the same degree of sophistication and redundancy as is the bulk power network (This type of distribution-level design does exist to some extent in major urban areas that have underground distribution network systems.). The natural extension of these automated systems at all levels of supply is the micro-grid with distributed generation, combined lighting, heat, air-conditioning and power that is managed with its own optimization routines while accounting for the economic interface with the external transmission system. And the external bulk power system must anticipate and account for the sophisticated actions and responses of that black-box-like micro-grid.

Evolving technology merely amplifies the potential value of a smart-grid that encompasses all of these perspectives. Consider the combination of wind-power, solar and the advent of the plug-in-hybrid vehicle. The latter, for the first time, would allow for economically-based distributed storage of electricity (since it’s justified by the much higher price of gasoline, rather than by the price of electricity) at the distribution level that could take advantage of renewable sources of generation. But to the extent that more and more of these technologies emerge at smaller scales and are implemented on the distribution system, achieving economical coordination with the customers’ use of electricity becomes ever-more important. Since the sun may not be shining or the wind blowing when a driver remembers to re-charge her electric vehicle’s batteries, the efficient development and use of generating resources - - particularly of less-polluting and/or renewable-resource-based generation - - cannot happen without nearly every customer having a sensor measuring their usage in real time. And accurate measurements are of far greater benefit in modifying peoples’ behavior efficiently when those users are charged for their consumption based upon time-differentiated costs of supply that include external environmental costs.

5. Smart-Grid = Every Customer with a Smart-Meter

Why have we been so reluctant to install smart-meters that sense and record electricity usage instantaneously? Perhaps because the next step would be to charge customers according to their usage patterns (and according to the costs, including environmental costs, that that consumption imposes on others). The popular presumption has been that customers don’t want to be bothered and that they would rebel at having real-time pricing imposed, but previously reported experimental work suggests otherwise.

Several years ago my colleagues and I at Cornell conducted a series of experiments [5], where we actually paid the customers (students) real money for the money they saved in their electricity purchases in a computer-simulated electricity market. They first tried their hand under traditional constant price tariffs under a range of weather conditions and simulated day-night differences in the value of using electricity. Next, they ran through the same sequence of purchase conditions where they paid real-time market-clearing prices (they also tried a third scenario similar to many demand response programs that are available throughout the country where a pre-specified saving could be reaped if purchases were reduced in pre-announced stress periods). The experimental results overwhelmingly favored real time pricing; price spikes were reduced in most peak periods, as were suppliers’ profits. In economic jargon, the industry was more efficient. But in addition, before the experiments were begun, the participants were asked which pricing system they thought they would prefer. Beforehand, two-third said they’d choose the constant fixed price, but after trying all three systems, two-third selected real time
pricing! This reversal in preferences is statistically significant by any test, and it refutes the popular political wisdom that customers won’t like it.

But the benefits of letting all customers into the electricity markets don’t end there. In further simulations of the effects of these alternative demand systems on a 30-bus electrical system with six generators, it was estimated that maximum transmission line capacities could be reduced by an average of seven percent, and peak generation requirements would be reduced by five to ten percent [6]. In fact, in these experiments, total electricity consumption increased slightly under real-time pricing, as compared to constant fixed prices, since night-time usage expanded (because of much lower prices) more than high-priced daytime usage declined. This is because of the greater inelasticity of customer demand for daytime use.

In the long run, these two effects of real-time pricing would translate into lower capital costs per MWh sold, but these further benefits weren’t included in the estimates of increased overall customer value. Nor were the potential benefits from the incentives that real-time pricing might provide for the proper use of emerging technological advances like wind and solar generation, micro-grids and plug-in-hybrid vehicles accounted for. Just imagine the economies resulting from storing electricity in your plug-in hybrid car that was generated by low-cost base-load and wind units at night, compared to recharging during the day when expensive gas-fired peaking units may be the marginal generator. Without real-time pricing, the incentive is smaller for customers to make the extra investment in plug-in hybrid cars.

Furthermore, full two-sided markets for electricity where every customer has an opportunity to react to prices the way suppliers can, should make the system easier to operate, as detailed by Schuler [7]. As an example, operating evidence from Australia [8] where the cap on wholesale prices is ten times higher than in all jurisdictions in the United States has led many load serving entities (LSEs) and large customers to install automatic load-shedding devices that are triggered by low frequency, since there is an inverse correlation between system capacity stress (high prices) and frequency. The large LSEs are mandated to install these devices by the system operator, but many large customers seek to avoid price spikes. As a result, following a sudden 3100 MW loss of capacity over a twenty second interval (1800 MW in the first five seconds) on August 15, 2004, the system began to stabilize and resume a path back to normal frequency within less than twenty-five seconds from the time of the initial generator trips. Here the operators’ activity was certainly enhanced by the automatic reduction of load; it’s like additional damping was introduced into the electrical system.

Why has real time pricing not caught on? Why hasn’t it been actively promoted by regulated utilities and many public power agencies (there are exceptions)? The answer may lie in the expected level of a supplier’s (utility’s) profit margins. Even with decreasing costs, profit relationships as a function of price (the difference between the TR and TC curves in Figure 1) can be shown to be concave in prices as depicted by the dashed-line, smaller parabola. If under real time prices $P_1$ and $P_0$ in peak versus low-load periods, the supplier were to earn profits $\pi_1$ and $\pi_0$, respectively, then their average profit would be $\overline{\pi}$. If however, the supplier is allowed to charge the same average price in both periods, with the same concave profit function the firm would charge $\overline{P}$ and earn $\pi(\overline{P})$ which is greater than the result under real time pricing. This perspective is analogous to the Friedman and Savage [9] analysis of choice under uncertainty that attempts to explain why individuals are willing to pay a premium for insurance rather than bear the risk of uncertainty.

One possible explanation for the persistence of averaged, bundled prices for electricity customers is that they may have been advantageous to suppliers by providing a greater profit-cushion in the face of varying demand. A second reason may be that by masking the truly high cost of providing electricity in some periods, the manufacturers (and developers) of storage devices and/or means of providing electricity only in brief spurts might be deterred from entering the industry. Thus averaged retail prices may reduce

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{figure1.png}
\caption{Profits under varying vs. average prices charged to buyers}
\end{figure}

the potential for a great deal of supply-side competition at the retail level.
In addition, elected officials, regulators, and therefore regulated utilities do not like headlines about high prices, no matter if those $1.00/kWh prices only persist for one hour out of twenty-four, and the prices were at record lows for twelve of the remaining hours. By comparison, it’s those price spikes in a market environment that set the creative juices of entrepreneurs into motion developing new storage technologies, smart controls on appliances and home-heating and air-conditioning units and/or lower cost methods for generating at the peaks. But so long as public outrage focuses primarily on the isolated high price, and not the consequent declining average price, regulators and politicians are certain to pander to the instincts of their constituents, rather than try to educate them about the realities.

6. Comparisons with Other System-Service Industries

Airlines, highways and banking are other services that are highly reliant on a coordinated network for the efficient and reliable (predictable) flow of their services; yet curiously, all three have a long way to go to match the reliability of electricity supply. Air travel has become a nightmare, as has commuting by car during rush hour in most urban areas throughout the country, and the nation is still suffering from a lingering shortage of liquidity following the financial collapse in Fall 2008, despite the action of the Federal Reserve to pump trillions of dollars of additional funds into the banks’ coffers. It’s like trying to serve New York City’s electrical demand by building a large amount of new generation upstate, but not increasing the capacity of any of the connecting transmission lines.

What’s curious is that these other, network-related industries do provide reasonably accurate market signals to their customers with highly differentiated prices for use at different times and places (the exceptions are the lack of highway congestion pricing in most areas, and the lack of penalties for airlines who do not meet their schedules). What these industries haven’t sorted-out is the effective management of their local-public good, network-related problems. They simply aren’t utilizing the existing network capability as efficiently as they might, and instead they seek to develop more capacity (more roads, airplanes, etc.). These industries could learn much by studying the electricity supply system and by moving toward its equivalent of the smart grid.

But what the electricity industry has been missing is to include all of its customers in the “smartness” game. How can we have a smart grid without smart customers? Telecommunications seems to be the one network industry that has gotten both aspects right, and the truly smart grid would be a marriage of these two industries. Would the I-phone have come about if all telephone customers had to be served by one regulated utility and paid a single flat monthly fee? Imagine levying a charge of $599 for a new telephone (the initial asking price for an I-phone) under a regulated regime (note that two years later the price is under $100). Electricity customers simply don’t have that type of choice, and because they don’t, the imagination and creativity of individuals and businesses to provide new products and services haven’t emerged in the electric supply industry as they have in others. No wonder the best and brightest students are frequently attracted elsewhere. But, if we can open electricity markets to retail customers, not only might we achieve the full promise of the smart grid, the way may have been paved to transform the electric supply industry into something totally unimagined. By integrating environmental services, it might even become a pathway toward a sustainable society - - all at the touch of an I-phone.

7. References


