Demand response: the sequel to DSM?

Fereidoon Sioshansi, Menlo Energy Economics

1. SYNOPSIS

With the introduction of competition in electricity markets, time has arrived for the demand side to take a more active role in balancing supply and demand.

2. ABSTRACT

For decades, US regulators and policy makers coerced monopoly utilities to deliver a host of **energy efficiency** and **load management** services to their captive customers. It was an uphill battle. Now, in restructured electricity markets, the incentives to engage in conservation and load management are driven by market pull, not regulatory push. This paper describes the business proposition behind the concept of **demand response (DR)**. The basic idea is to provide economic incentives for customers with elastic and flexible demand to curtail usage so that other customers with more pressing needs can continue to be served. The idea is ripe for applications in regions where generation supplies are short and the delivery infrastructure cannot serve the peak demand.

3. INTRODUCTION

Following the 1973 energy crisis, both state and federal regulators in the US started to encourage regulated utilities to implement a wide range of **energy conservation** and **demand-side management** (DSM) programs to their captive customers. Although these programs resulted in significant savings in both capacity and energy, there was an inherent problem. Regulators were attempting to push utilities to sell fewer kilowatt hours, while investor-owned utilities (IOUs) had a tendency to do the opposite. This dichotomy between *negawatts* and megawatts is described in Sioshansi (1996).

To overcome this problem, regulators devised convoluted schemes to make negawatts more palatable to the IOUs through incentives, shared-savings, and so on. But despite their best efforts, the DSM programs of the 1980s and early 1990s produced uneven results, sometimes achieved at relatively high costs, as documented in Joskow and Marron (1992). The administrative costs and the regulatory burden of overseeing the implementation, measurement and evaluations also added to the costs.

Due to restructuring, DSM programs gradually fell out of focus in a number of states since the mid 1990s. Moreover, the regulators now have very little leverage on many market players, notably independent generators, competitive retail suppliers, power traders and re-sellers. Nor is it clear who should be responsible to implement DSM programs (Sebold *et al*. 2000).

California and New York, two pioneering states, have instituted a public purpose fund on retail rates in CA and NY. They have also established new policy setting, administrative, and implementation agents to carry out the programs. In the case of California, the initial attempt to establish the California Board for Energy Efficiency (CBEE) turned out to be a colossal mistake. With restructuring now sweeping the country, other states are grappling with similar issues – namely how to fund, administer, and implement DSM in the restructured environment (Kushler, 2000).

4. THE DAWN OF A NEW AGE: MARKET-DRIVEN DSM

With restructuring now spreading across the country, regulated utilities are becoming open-access *conduits* who essentially rent their distribution wires to competing **energy service providers** (ESPs). Moreover, in many cases, the

utility distribution companies (UDCs) have been forced to sell their generation assets. In these cases, they serve the needs of their remaining customers by buying from a competitive wholesale auction.

This functional unbundling has fragmented what used to be a vertically integrated supply chain. Consequently, any savings derived from implementing energy conservation or peak-load reduction at the retail level have been disconnected from those who stand to gain from it, namely generation and/or distribution companies.

Other things have changed too. Historically, most customers bought electricity from monopoly utilities who offered them *stable prices* that were regulated and *low*. With deregulation, restructuring, and privatisation of the electric power sector in North America and elsewhere, electricity prices are becoming more *volatile*. Customers who were used to predictable monthly electricity bills will increasingly be exposed to unpredictable bills. These phenomena, combined with the competitive business pressures have created an atmosphere, which is ripe for demand-side aggregation and bidding or demand response.

What is demand-response?

Demand response refers to the ability to encourage some customers with elastic and/or low-value usage to *get off* congested networks during periods of high demand. By doing so, these customers allow others with highly inelastic and/or high-value usage to remain on the congested network and be served. Although somewhat similar to the concept of real-time pricing (RTP), DR’s main objective is distinctly different.

RTP programs are primarily focused on advising customers *a priori* on how to adjust their usage based on variable electricity prices. RTP programs encourage customers to:

- curtail load during high demand (and high-price) periods;
- increase load during low demand (and low-price) periods; and
- shift load from on-peak to off-peak periods.

DR programs, by contrast, are primarily focused on allowing customers with elastic and/or discretionary demand *to forego usage* during capacity constrained periods so that others with inelastic and/or critical usage can continue to be served. DR programs:

- maximise revenues (or yield) derived from capacity-constrained infrastructure; and
- encourage low-value added customers to forego their usage during peak demand periods so that higher-value customers can continue to receive service.

5. AIRLINE ANALOGY

To understand the DR concept and to contrast with RTP, it may be helpful to use an airline analogy. Although the two industries differ in many ways, airlines and the electric power industry have several common features. Both industries are:

- **network-based** in the sense that they have to invest in vast, capital-intensive, interconnected networks to deliver the required services;
- **capacity-constrained** in the sense that they experience bottlenecks during periods of high demand and have under-utilised capacity during all other times; and
- **high-fixed cost** in the sense that their marginal cost of serving an additional customer or additional unit of output is minimal relative to their fixed costs.

These common features mean that both industries must design their systems to serve a *finite capacity* (peak demand), which becomes over-subscribed during heavy demand periods. For example, airlines never have enough seats during heavy travel season around holidays or on popular routes during peak-demand periods. They can, of course, expand capacity to satisfy these short-duration peaks, but this would come at a high cost (e.g., more planes, more landing rights, more gates, more crew, more fuel, more maintenance, more finance costs, etc.).

However, adding capacity to meet peak demand would mean that they will have a higher number of empty seats (and fixed costs) during low-demand periods. In short, adding capacity indiscriminately to meet the peak load may result in financial ruin. The answer, therefore, is not increased capacity, but *high utilisation* of the finite available capacity.

Since deregulation, major airlines employ sophisticated yield management systems. Broadly speaking, these systems are designed to *maximise revenues* (hence profit) on finite available capacity. The systems do this by dynamically adjusting ticket prices. For example seats tend to be more expensive on popular routes and/or during peak-demand periods.

6. OVERBOOKING

Yield management programs also allow airlines to engage in another practice known as *over-booking*. Overbooking allows airlines to sell a few more tickets than the actual number of seats available. There are two objectives:

- First, to compensate for the passengers who have bought a ticket but do not show; and
- Second, and more importantly, to ration scarce capacity during heavy demand periods.

Occasionally, of course, everyone who has bought a ticket actually shows up at the gate. Since all ticket holders are entitled to get on board (notwithstanding the warning on overbooking on the tickets), a mechanism must be in place to get a few people *voluntarily* off the plane. Airlines conduct an auction to see which customers are willing to voluntarily forego their “right” to fly on the overbooked flight in exchange for a seat on a later fight and a travel voucher. They start the auction at low prices and raise the price until the desired number of seats are vacated. In practice, there are always a few individuals with elastic demand – i.e., people who are willing to take the next flight in exchange for a travel voucher. Incidentally, the travel vouchers are essentially worthless empty seat on a future flight, so the airlines suffer very little when they give away these rewards.

The scheme is quite clever. Full fare business travellers with pressing schedules get to fly on over-booked flights. Passengers with flexible schedules and on tight travel budget voluntarily vacate their seats. The airline makes a lot more money without greatly inconveniencing any one. It is win, win, win.

It is instructive to examine the similarities between the airlines’ practice of overbooking and the concept of DR. DR allows the electric power industry to generate more revenues from its over-subscribed network during peak demand periods. It can do this by luring customers with discretionary requirements to curtail their usage so that customers with essential needs can continue to be served.

Why is time ripe for DR

Three factors will drive the demand for DR as well as wide-spread use RTP and other load management programs in the future:

- Competitive pressures to generate more revenues from fixed networks;
- Increased wholesale price volatility in liberalised electricity markets; and
- Impracticality of investing in generation to serve infrequent, short duration peak loads.

Not surprisingly, a number of enterprising companies are entering this business to offer DR and other forms of proactive load management services to customers. For example, **Idaho Power Company** in Boise, Idaho has bought roughly 250MW of load back from its agricultural customers during the coming summer period at a cost of $150/MWh using an auction system. Most of these customers use electricity for irrigation pumping. The rational is simple economics. Like most utilities in the West, Idaho Power has to buy some energy from the wholesale spot market at exorbitant prices. Buying back demand from its own customers at $150/MWh is far cheaper than buying it from the spot market. The customers are happy, since getting paid $150/MWh for not pumping provides them with more income than they would from producing the crops.

Likewise, the **Bonneville Power Administration** (BPA) has proposed to shut down the aluminum smelters in the Pacific Northwest for 2 years. This would free up over 1,000 MW of capacity, which the agency does not have to buy from the high-priced wholesale spot market. Due to the current draught in the West, hydro capacity is strictly limited and BPA has to buy power to meet its commitments. Paying customers not to use electricity during the current energy crisis is far more economical and logical. BPA has pointed out that aluminium smelters cannot economically justify operations when price of electricity rises above $30/MWh, which is definitely the case for the next couple of years. So the agency has asked the smelters to remain *off line,* allowing the 1,000 MW of spare capacity to flow to customers who are willing to pay much higher prices for it.

These two examples, however, illustrate the lazy man's version to DR. In both cases, selected customers are paid to stay off line at all times when wholesale prices are known to be high. Such static DR schemes are extremely easy to implement, but they are inflexible and unintelligent. Most customers cannot afford to forego the use of electricity for extended periods of time. But they can adjust their usage or short periods of time given sufficient incentives. This is where more proactive, intelligent forms of DR programs come to play.

Sacramento Municipal Utility District (SMUD) in Sacramento, CA, for example, has a program that provides large industrial & commercial customers fairly high incentives for curtailing usage for a few hours, say 2-6 pm when peak demand usually occur in California. Although the details of these programs have not been made public, it is understood that customers who are willing to curtail partial or total usage on short notice will be offered a portion of the prevailing wholesale prices.

In this case, the motivation is more than saving money. By selectively curtailing the load of willing customers, SMUD can effectively avoid highly disruptive rolling blackouts, which would otherwise be necessary during emergency episodes. Rolling blackouts do not discriminate among customers with elastic or non-elastic load. DR is a much preferred alternative to rolling blackouts.

Other utilities as well as the California Independent System Operator (ISO), and the California Energy Commission (CEC) are also working to implement similar schemes in time for this summer's peak demand period. A number of private **energy service companies** (ESCOs) and specialty firms such as **Silicon Energy** and **Apogee Interactive** are also actively engaged. These services are sorely needed for this summer in California as well as in other regions of the US where capacity shortages and/or high wholesale prices are expected.

The market for DR services is expected to grow significantly in the next few years, particularly in capacity-constrained and/or deregulated markets. The main driver is the high cost of wholesale electricity during peak demand periods. Advancements in DR technology make the solutions not only feasible but affordable. Companies that are offering these types of services include:

- energy service companies (ESCOs) load and load management companies;
- load aggregators, energy re-sellers, scheduling co-ordinators;
- retail energy suppliers;
- independent systems operators (ISO);
- trading companies; and
- traditional distribution companies and/or their non-regulated subsidiaries.

Critical success factors for DR

Succeeding commercially as a DR service provider will require mastery of a few *technical* and *business* complexities include the following:

- **Volume –** Many customers' loads have to be aggregated and managed for the concept to become commercially viable.
- **Customisation** – Each customer has a unique load profile and service requirements, hence the need for individualised solutions within the broad business concept.
- **Real-time –** Prices and loads must be controlled and managed in real-time, or close to it, thus requiring dynamic and instantaneous two-way communications and feedback.

- **Co-ordination** – Sophisticated real-time communication and co-ordination is required between the DR service provider and customers on the one hand, and the independent system operator (ISO) or scheduling coordinator (SC) on the other.
- **Collaboration –** Collaboration with numerous distribution supply companies (DSCs), scheduling co-ordinators, metering and billing companies is necessary.
- **Monitoring** – Sophisticated metering and billing systems are required to measure what is consumed and how much load has been shed to bill customers accurately and in a timely manner.
- **Settlement** – Settlement and back office systems are required to manage the transactions efficiently, securely, and at low costs.

In short, DR service providers need powerful back office systems to effectively handle large volumes of business transactions among multiple players securely and in real time. More technical details may be found in Sioshansi and Vojdani (2001).

7. ELECTRICITY RESTRUCTURING HAS BROUGHT CAPACITY CONSTRAINTS TO FOCUS

Electricity supply industry (ESI) restructuring, which started in a handful of states beginning in 1998, has rapidly spread to 26 states and the District of Columbia. Over the next few years, more states may follow suite, even though California's recent problems have caused a few states to postpone the opening of the market. At the same time, the **Federal Energy Regulatory Commission** (FERC) has essentially turned the nations' transmission grid into an open-access regime. This allows traders, generators, customers, suppliers, and others to access the grid for purpose of shipping energy from one place to anther.

In some regions of the country, however, supply has not kept up with robust demand growth for electricity. Consequently, many regions of the US are short of capacity, particularly during the hot summer months when the use of air conditioning creates sharp demand peaks. This problem is exasperated by an even more acute lack of investment in transmission capacity. Not surprisingly, many local load centers are poorly connected. As a result, during critical demand periods, these load centers cannot be adequately served even though excess generation capacity may be available in a nearby region.

Other factors, including an unprecedented rise in natural gas prices (which fuels many electric power plants), increases in oil prices, and environmental factors have made matters worse. The effect of acute shortages in generation and transmission capacity combined with deregulation of wholesale electricity markets has resulted in massive increases in wholesale prices at major trading hubs in the US, as shown in Table 1.

Table 1. Wholesale price increases at major US trading hubs

TRANSMISSION INTERCONNECTION POINT	% CHANGE 1997-2000
Texas	+293
Louisiana-Mississippi-Arkansas	+216
Tennessee Valley Authority	+165
California-Oregon border	+162
New York-West	+138
Chicago area	+130
New England	+117
New York-East	+101
Upper Midwest	+ 99
Florida	+ 89
Mid Atlantic region	+ 80*

* Includes only years 1999-2000
Source: Federal Energy Regulatory Commission (FERC)

The net effect of these developments has been two pronounced trends – which are likely to be amplified in the next few years as more states open their markets to competition:

- wholesale electricity prices are likely to increase in many regions of the US – notably in areas with acute capacity shortages; and
- prices are likely to become more volatile and less predictable as more utilities rely on competitive wholesale markets to secure their supplies (as opposed to generating all their internal needs).

These concerns, as previously mentioned, have prompted a few states to postpone and/or modify the opening of their markets, at least until more workable solutions to competitive markets can be found.

Short term solutions?

In the short-run (e.g., next 2-3 years), there is little that can be done to address the fundamental supply shortages. For example, additional generation capacity – now in planning and under constructions – is expected to alleviate the supply shortages in the West. But it will take at least a couple of years to bring additional units into service. It will take even longer to fix transmission capacity bottlenecks. Additional supplies of natural gas can be brought to market, but that will take longer still. In the short-run, there are only a handful of feasible solutions:

- increased reliance on energy conservation to make existing supplies go farther; and
- increased reliance on schemes such as RTP and DR to allocate limited capacity during high demand periods.

Energy conservation refers to overall reduction in energy usage (at all hours). This can be achieved, for example, by replacing inefficient incandescent light bulbs with more efficient compact fluorescent lights (CFLs), or by replacing old, inefficient refrigerators with more efficient models. Energy conservation will play a growing role, no doubt, in the coming years. In California, for example, Governor Davis has proposed a $1 Billion budget to fund energy conservation efforts. Energy conservation helps by reducing demand at all hours – thus stretching existing capacity.

Schemes such as DR and RTP, however, offer more surgical and highly effective weapon against high prices – caused by sharp demand peaks and the necessity of keeping supply and demand in balance in real-time.

DR: How would it work?

DR may be offered to a wide variety of customers under a number of schemes. One possible scheme may be offered by a middle-man, be it a service aggregator or ESP. In principle, the customers can be large or small, industrial, commercial or residential. In practice, however, the concept works best with a small number of large customers, simply because more volume can be aggregated with fewer customers, making it more manageable.

The aggregator combines the load of a number of customers and works as the middle-man to provide a packaged service or solution. The scheme's main selling point to customers is that by participating in the program, they can potentially make money by reducing their energy costs during high demand periods. This is similar, in principle, to the passenger who gets off the plane in exchange for a travel voucher.

The DR service aggregator can make money in any number of ways. One possibility would be to charge a flat monthly service or membership fee, plus a percentage of any savings derived from each customer's participation in the scheme. To make the scheme work, however, the aggregator must offer a number of critical services in real-time and manage the aggregate load of its customers in such a way that does not inconvenience anyone involuntarily. To achieve this, the aggregator must encourage program participants to reduce consumption during periods when energy is scarce and expensive – i.e., peak demand periods.

Case Study: California

California policy makers restructured the state's electricity market by passing a legislation in 1996. They promised lower prices in an efficient, competitive and self-regulating market. The restructured market opened in April 1998 and worked relatively smoothly for the first two years. But starting in May of 2000, wholesale prices have risen (Figure 2) and have remained highly volatile. Moreover, California's dysfunctional market is now affecting prices in the entire Western region of the US, which is inter-connected. Consequently, high prices are not limited to California alone, but spill into the neighbouring states (Figure 3).

Figure 2. Wholesale electricity prices in California

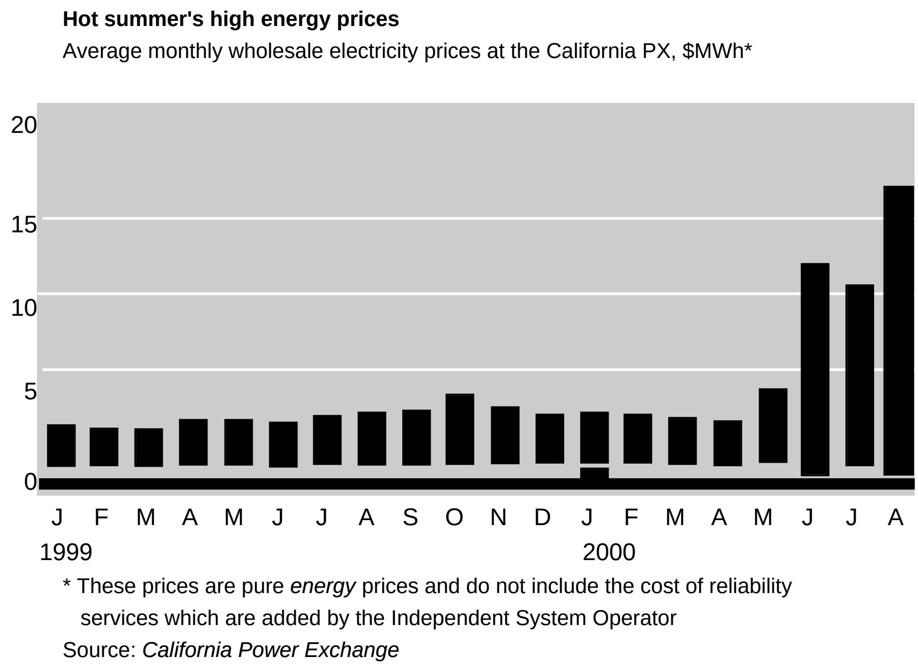
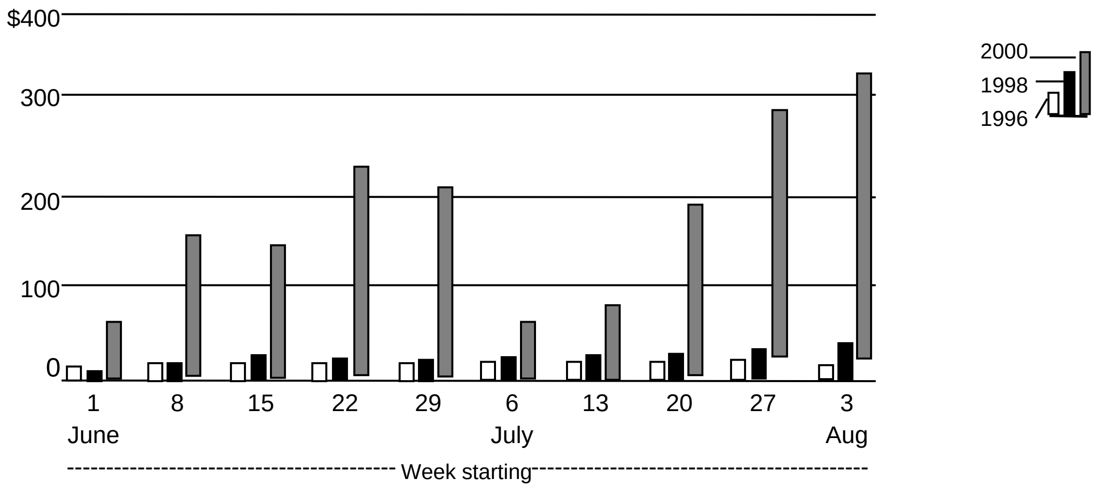


Figure 3. High prices are not limited to the California market



Source: Northwest Power Planning Council (NPPC)

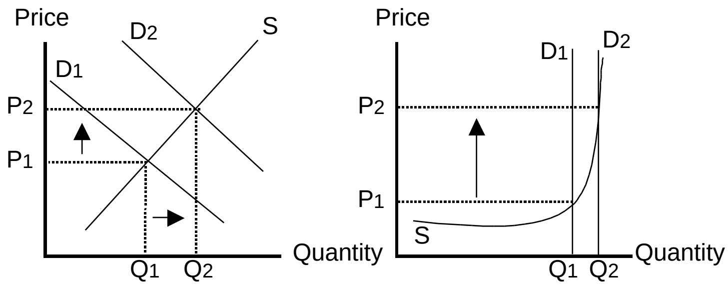
The experience of California provides a wealth of insights about the complexities of restructured electricity markets and what precautions may be taken to avoid a repetition of some of the things that went wrong. A number of fixes are being tried and are expected to help the market function as originally envisioned. These include several longer term solutions that will take several years to bear fruit. In the short-run, however, there are only few things that can be done to avert disaster: energy conservation and load management options such as RTO and DR.

8. ONLY HALF OF A MARKET

One of the fundamental flaws of the California market, which is common to many other newly restructured markets around the world, is that it is virtually a half market. Only the supply side of the equation sets the market clearing price (MCP). Customer demand is treated as a *given*, because the customer's end energy prices are fixed and thus not fluctuating with the whole sale price. It thus is not allowed to respond, or react, to fluctuations in price, except to a minimal degree. This major drawback is the result of a faulty assumption that demand is *in-elastic*, i.e., it does not respond to high prices (EEnergy Informer, 2000)

Consequently, customers, by and large, do not reduce consumption even during periods when prices are extremely high (Figure 4). The net result is that consumers continue to use electricity when it is exceedingly expensive to produce goods and services that are barely worth producing. From an economic (as well as purely logical) perspective, this is highly inefficient. According to economic theory, firms use inputs only when their use is justified by the market value of the output.

Figure 4. Supply and demand curves in *normal* markets vs. in capacity constrained markets with perfectly inelastic demand



For example, it would not make any sense for an aluminium smelter to produce aluminium when electricity costs 30 cents/kWh. That is why several large smelters in the Pacific Northwest have shut down operations in recent months, figuring that they can make more money by selling their energy to power hungry California market. This is the so-called concept of economic efficiency – that electrons should flow to where they can extract the highest price.

The DR concept would allow, in principle, all customers, not just big aluminium smelter, to decide when it is cost effective to use energy – and when it makes sense to reduce or shift usage, or shut down all operations. Presently, for most customers (i.e., everyone other than big steel mills and like) the concept would only work if a service provider would take care of the chores. The savings may not be worth the bother for many small to medium-sized companies. And here lies an enormous business opportunity for DR service providers.

9. SUMMARY AND CONCLUSIONS

This paper has described the business context for a new type of service. DR offers a logical, commercially viable, and sustainable business proposition in newly restructured, capacity-constrained markets, such as those in many regions of North America. In states like California (and neighbouring interconnected states), the concept is in high demand, and will remain so for the foreseeable future. As competition spreads to other parts of the US (and elsewhere), the demand for DR is expected to become universal.

The challenges for DR service providers are many – but these are not insurmountable. A lot of energy and ingenuity is at work to address these technical issues, and companies are hard at work to develop integrated systems and solutions. The opportunities to save costs and improve the efficiency of capacity-constrained networks are too large to suggest otherwise.

10. REFERENCES

Sioshansi, F. P., 1996, The Rise and Fall of DSM: Why Did it Sizzle and How did it Fizzle, **EEnergy Informer** publications, Menlo Park, CA

Joskow, P. and D. Marron, 1992, What Does a Negawatt Really Cost? **The Energy Journal**

Sebold, F., *et al.*, 2000 A Framework for Planning and Assessing Publicly Funded Energy Efficiency, AESP Conference, New Orleans, LA

Kushler, M, 2000, What's Up With Public Benefits? A Review and Initial Assessment

of Public Benefits Policies and Programs in Restructured States, AESP Conference, New Orleans, LA

Sioshansi, F. P., 2001, California's Dysfunctional Electricity Market: Policy Lessons on Market Restructuring, forthcoming in **The Energy Policy**

EEnergy Informer, October 2000, Is There a Cure for California's Wobbly Market?

Sioshansi, F. P. and A. Vojdani, Why the time is Ripe for Demand Response, forthcoming in **The Electricity Journal**, June 2001

11. END NOTE

***** For additional information contact the author via e-mail fpsioshansi@aol.com