

What's Wrong with Capacity Markets?

By

Frank A. Wolak

Department of Economics

Stanford University

Stanford, CA 94305-6072

June 23, 2004

1. Introduction

This paper describes why capacity markets do not enhance the efficiency or reliability of wholesale electricity markets. In the United States, capacity markets have been touted as an essential ingredient to a successful wholesale electricity market. However, this view does not appear to be held outside of the United States. The wholesale electricity markets in Australia, United Kingdom, and the Nordic countries of Denmark, Finland, Norway, and Sweden have been in existence for at least ten years and are generally acknowledged as the best-performing, yet none has a capacity market.

The superior performance of these wholesale markets is not surprising because capacity markets considerably complicate the task of providing the appropriate signals for the new investment necessary to meet current and future demand. In particular, capacity payments encourage over-investment and new generation capacity mix that more expensive that is necessary to meet an increase in annual electricity demand. Capacity market are also extremely susceptible to the exercise of unilateral market power, which implies that regulatory intervention is often needed to set the price paid for capacity.

The major rationale for capacity markets in the United States appears to be a holdover from the vertically integrated regulated regime when capacity payments compensated generation units for their capital costs, because the regulatory process typically reimbursed unit owners for their variable operating costs. In a wholesale market regime all generation unit owners have the opportunity to earn the market-clearing price which is typically above generation unit's average variable cost when the unit is operating. In this way, the generation unit earns a return to capital during each hour it produces electricity.

This paradigm for earning a return on capital from the difference between the market price and the firm's average variable cost of production has managed to provide the appropriate incentives for investment in new productive capacity all workably competitive industries. Given the international experience cited above, there is little reason to expect that it could not work in the wholesale electricity industry as well.

The remainder of this paper describes the major market inefficiencies introduced by capacity markets. It then outlines an alternative approach to ensuring capacity adequacy that is used in the successful wholesale markets from around the world. The paper closes with suggestions for implementing this approach in the California market.

2. Why Capacity Market Don't Benefit Consumers

Consumers want to be able to withdraw electricity from the network when they need it at the lowest cost possible. Although some consumers also care whether the electricity they consume is produced in an environmentally friendly manner, the physics of electricity transmission and distribution make this impossible. The best that a supplier can offer is that he will inject the same amount of energy from his generation units as the consumer withdraws from the network during that hour. Electricity flows along the path of least resistance, so it is impossible for a seller to deliver his electricity to a specific location in the transmission network or a specific buyer.

Consequently, it is unclear why electricity is so fundamentally different from other products that it requires paying suppliers for their generation units to exist. Consumers want cars, not automobile assembly plants. They want point-to-point air travel, not airplanes. They want a loaf of bread, not a bakery. In these markets we do not pay producers capacity payments for the fact that they own the facilities to provide the goods and services we desire. All of these industries are high fixed cost, relatively low marginal cost production processes, yet all of these firms earn their return on capital invested by selling the good that consumers want at a price above the variable cost of producing it. Clearly cars, air travel, and bread are in many ways essential commodities, yet we don't need a capacity payment to ensure that there is sufficient productive capacity for these products to meet our needs.

Conceding the argument that electricity is special because it is an essential part of the modern information economy, I will now address the question of whether a capacity payment is the least cost way to achieve a desired level of reliability now and in the future. The first step to answer this question is a clear statement of what is purchased with a capacity payment. The typical capacity market requires load-serving entities to procure available generation capacity that can be delivered to load. Available generation capacity is something that the independent system operator (ISO) certifies. It is usually the nameplate capacity of a generation unit deflated by the amount of time it was available to supply electricity over some historical time period. For example, if a 500 MW unit was available 90 percent of the hours of the past year, this unit might therefore have an available capacity rating of 450 MW. The demand for available capacity is created by the ISO requiring that all load-serving entities (LSEs) purchase a certain multiple of its monthly peak demand in available capacity. For example, the requirement could be to purchase 115 percent of each month's peak demand in available capacity in advance.

So what do consumers get from this available capacity purchased on their behalf by the LSE? Typically, resources that are sold in the capacity market are subject to certain restrictions on their bidding behavior. For example, they are required to bid their capacity into the market below the price cap and must be subject to a local market power mitigation mechanism. In addition, they are often required to curtail any exports they may have sold if the ISO needs them for additional energy in real time. Although these

restrictions on supply behavior are potentially valuable to consumers, it is not necessary to make capacity payments to generation unit owners in order to subject suppliers to these restrictions. In other markets around the world, many of these restrictions are imposed as a pre-condition to sell electricity in the wholesale market. Because the supplier has the right to receive a wholesale price in excess of their variable cost of production they have an obligation to be subject to these restrictions.

Another problem with capacity markets is that they are extremely susceptible to the exercise of unilateral market power for the simple reason that the demand for this product is extremely price inelastic. In the Eastern US markets, there have been numerous instances of the exercise of the enormous market power in these capacity markets. During the off-peak months of the year when no single supplier is pivotal in the capacity market, the price paid for capacity is very close to zero, which is the marginal cost of a supplier providing an additional megawatt (MW) of available capacity. During the peak and shoulder months when one or more suppliers are pivotal in the capacity market, there is no limit on the price a supplier can charge. For example, suppose a market has 10 suppliers each of which owns 1200 MW and the peak demand for the system during the peak month is 10,000 MW. Under these circumstances all suppliers know that the aggregate available capacity requirement of 11,500 MW ($=1.15 \times 10,000$ MW) cannot be met without some of their capacity. As consequence in all of the Eastern US markets, very stringent market power mitigation measures have had to be put in place. Consequently, capacity prices typically fluctuate from very close to zero to the regulatory price cap. It is difficult to see how these very volatile prices provide very useful signals about the need for new investment in generation capacity.

Instead, capacity payments look very much like subsidies to the construction of new generation capacity. Recall that they are payments made to generation unit owners for being available to produce electricity. The requirement to make these payments to owners of generation capacity can lead to over-investment in new generation capacity, which will reduce the spot price of electricity and increase the need for higher capacity payments in the future. These higher capacity payments could induce further investment, which would lead to lower energy prices. This unstable sequence of capacity and energy prices appears to be playing out in the PJM market at the moment. Energy prices are very low and suppliers are arguing in favor of higher capacity payments and other non-market-based compensation schemes to keep them in business.

Even if a wholesale electricity market has a capacity market, there is no way to compel generation unit owner provide electricity if they would prefer to withhold this capacity to drive up the spot price of energy. Recall the “sick day” problem that occurred with generation units during the period December 2000 to May 2001 when many units were “declared” unavailable to operate. A capacity market cannot prevent this from occurring, although it is true that if generation unit owners declared too many sick days this would significantly reduce the amount of available capacity they could sell in the future.

Finally, a capacity market makes even less sense for an import-dependent market like California than it does for the capacity sufficient markets in the eastern US. If a capacity market is going to be of use to the ISO in maintaining a reliable transmission network, the LSE must be able to point to an actual physical resource as its capacity resource. California relies on imports for close to 20 percent of its energy needs. Although it is possible to point to a specific resource outside of California that is supplying an LSE in California with available capacity, it is difficult to enforce many of the restrictions on these resources that apply to capacity resources located in California. For example, how would an export schedule be curtailed if the unit is located outside of California and the transmission line connecting this resource to the California ISO control area is down? Because generation resources located outside of California cannot be seen by the California ISO control area, it is difficult to see how these resources could sell a capacity product in California that is equivalent to the product sold by resources located in California.

3. Ensuring Adequate Energy at Least Cost to Consumers

So how do the successful markets described earlier ensure that sufficient generation capacity is available to meet current and future demand without a capacity market? They do it the same way as other industries that make products with high fixed cost and low marginal costs and significant lags between order and delivery of the product: through an active forward market. For example, jet aircraft and jet aircraft engines are typically sold on a forward market, far in advance of the delivery date. In particular, the buyer that first purchased the product may not be the buyer that ultimately takes delivery. These successful wholesale electricity markets described earlier focus on procuring in advance the product the consumers actually want—energy at a specific location in the transmission network.

These markets focus on contract adequacy, not generation adequacy. For example, if an LSE would like a certain pattern of energy at a location in the network, then it should purchase a financial forward contract for this pattern of energy cleared against the spot price at the location where the LSE withdraws energy from the transmission network, not a new generation facility to supply this energy. The LSE should focus on purchasing the product that is needed, a hedge against the spot price risk at the locations in the network where it withdraws energy, and rely on the financial incentives that will exist once these forward financial contracts has been purchased to supply the needed energy at least cost.

A major mistake made by the California Department of Water Resources (CDWR) in negotiating the forward contracts during the winter and spring of 2001 is that it focused on purchasing power plants instead of hedges against the spot price of energy at the locations where the three large LSEs withdrew energy from the network. This procurement strategy created the seller's choice forward contract problem that is stalling the movement toward a more efficient spot market design for California. In addition, this inappropriate focus on buying generation assets, instead of forward financial contracts cleared against spot prices at locations in the network where the LSEs withdraw power,

has created substantial congestion costs in the Miguel and Imperial Valley region and the significant reliability problems for the California ISO. Had these forward contracts been purely financial contracts cleared against spot price at locations where the three large LSEs actually withdrew energy from the transmission network, both the seller's choice problem and the Miguel-Imperial Valley congestion problem would not exist. California could more easily move toward a spot market design that enhances market efficiency and benefits California consumers with lower-priced wholesale electricity.

By focusing on contract adequacy rather than building generation facilities, California would have a portfolio of forward contracts that provide incentives for least cost production of electricity in the short and long term. Firms that have sold these forward financial contracts would have strong incentives to ensure that the spot prices at the locations in the California ISO control area where these financial contracts are cleared are as low as possible. That is because once a supplier has signed a fixed-price forward financial contract that clears against the spot price at a given location in the network, the supplier's revenue stream is fixed, so it has the strongest possible incentive to ensure that the cost of meeting this forward contract obligation in real-time is as low as possible. Consequently, it is extremely unlikely that suppliers that sold these forward financial contracts would have constructed generation units at locations in the network where their plants could not operate to the greatest extent possible because of transmission constraints.

An active forward financial market also has other hedging instruments besides pure swap contracts, where a supplier and an LSE agree to a fixed price at a location in transmission network for a fixed quantity of energy. Cap contracts are also very effective instruments for guarding against price spikes in the spot market and for funding the appropriate amount of peak generation capacity. For example, a supplier might sell an LSE a cap contract that says that if the price at a specific location exceeds the cap price the seller of the contract pays the buyer of the contract the difference between the spot price and the cap exercise price times the number of MWh of the cap contract sold. For example, suppose the cap exercise price is \$300/MWh and market price is \$400/MWh, then the payoff from the cap contract is $\$100/\text{MWh} = \$400/\text{MWh} - \$300/\text{MWh}$ times the number of MWh sold. If the spot price is less than \$300/MWh, then the buyer of the cap contract does not receive a payment.

Because the seller of a cap contract is providing insurance against price spikes, it must make payouts when the price exceeds the cap exercise price. This price spike insurance obligation implies that the buyer must make a fixed up-front payment to the seller in order for the seller to be willing to take on this obligation. This up-front payment can then be used by the seller of the cap contract to fund a generation unit that provides a physical hedge against price spikes at this location, such as peaking generation unit. The Australian electricity market has an active financial forward market where these types of cap contracts are traded. These contracts have been used to fund peaking generation capacity to provide the seller of the cap contract with a physical hedge against this insurance obligation.

One question often asked about a contract adequacy approach is whether sufficient resources will be built to meet demand if consumers only buy forward financial hedges against spot price risks at their location in the network. In this regard it is important to bear in mind that the incentives faced by the seller of the forward financial contract once this contract has been sold. The supplier has an obligation to insure that forward contract quantity of energy can be purchased at the agreed upon location in the spot market (or whatever market the forward contract clears against) at the agreed upon forward price. They bear all of the risk associated with higher spot prices at that location. In order to prudently hedge this risk, they have a very strong incentive to construct sufficient generation capacity to ensure that in virtually all states of nature the risk associated with guaranteeing the price of spot power at that location in the network is minimal. This means that if a supplier signs a forward contract guaranteeing the price for 500 MWh of energy for 24 hours a day and 7 days per week at a specific location in the network, it will construct or contract for more than 500 MWh of generation capacity to hedge this spot price risk. Building only a 500 MW facility to hedge this risk would be extremely imprudent and expose the supplier to significant risk, because if this 500 MW facility is unavailable to provide electricity, the supplier must effectively purchase the energy from the spot market at the price that prevails at the time. If this generation unit is unavailable, it is very likely that the spot price is extremely high.

4. Implementing Contract Adequacy in California

Implementing the contract adequacy approach in California is relatively straightforward in the core/non-core retail market model currently under consideration. For the three large LSEs, the California Public Utilities Commission (CPUC) could set a target average wholesale price of electricity for all core customer electricity sold by the each LSE. For example, the CPUC could say that the LSE is entitled to receive 6 cents/kwh in revenue recovery for each kwh of energy sold to its core customers. If the LSE's spot price risk management process results in average wholesale energy purchase costs below this 6 cents/kwh target, then LSE gets to keep X percent of the revenues as profits to be distributed to its shareholders. Similarly, if its average wholesale energy purchase costs are above this level, then the LSE must make its shareholder pay for X percent of these cost over-runs. The CPUC can determine the appropriate sharing fraction X.

The CPUC can also monitor the forward contracting strategies of the three LSEs to ensure that they meet certain minimal prudency standards. For example, the CPUC could require that at least 75% of the LSE's core demand be hedged in a swap contract clear one of the locations in the network the LSE withdraws energy one year in advance, 85% one month in advance and 95% one day in advance. These and other minimal hedging standards could be enforced through penalty mechanisms.

For non-core customers, CPUC should impose minimal requirements on how these entities procure their wholesale energy. The only requirement would be that the default wholesale price a non-core customer pays for all of their power is the real-time spot price, unless the customer signs a spot price hedging agreement with a supplier. In

addition, all non-core customers should be required to subject themselves to automatic curtailment if there is insufficient generation capacity bid into the ISO's real-time market. Wolak (2004) discusses this core/non-core model in more detail.

References

Wolak, Frank A. (2004) "Designing a Retail Electricity Market That Enhances Wholesale Competition" Testimony to California Public Utilities Commission, April 20, 2004, for Core/Non-Core Electric Market Structure Options: En Banc Hearing, available at (<http://www.stanford.edu/~wolak>)