Convergence Bidding and the Enforcement of Day-Ahead Commitments in Electricity Markets

James Bushnell and Frank Wolak
June 1, 2006

Summary

Operators of electricity markets continue to struggle with the distinction between “physical” and “financial” transactions. The interpretation of this distinction can have implications for the reliability and efficiency of operations, as well as the vulnerability of market prices to the strategic actions of market participants. At a basic level, transactions in real-time markets are by definition physical, and all transactions made in advance of the real-time market are financial. Historically, problems have arisen in wholesale electricity markets that have ignored these distinctions by assuming that short-term forward market transactions are physical.

This paper first explains the distinction between physical transactions and financial transactions in wholesale electricity markets. We then describe the consequences of the failure of the California market design to recognize this distinction. We present several examples of how the strategic behavior of market participants can impact the relationship between day-ahead or hour-ahead prices and real-time prices and describe what factors contribute to the profitability of these strategic actions. Because day-ahead or hour-ahead forward market schedules that are substantially different from real-time output and consumption levels increases the likelihood of system failures, mechanisms have been implemented in most markets to limit the magnitude of purely financial transactions that take place between the close of the short-term (day-ahead or hour-ahead) forward market and the real-time market. The two sets of responses to these reliability problems are: (1) tight restrictions (or penalties) on the ability to adjust day-ahead commitments, or (2) the introduction of explicitly financial trading, such as convergence bidding, that allows for adjustments in financial positions without disrupting system reliability. The paper closes with a discussion of these two responses and our reasons for preferring the second approach.

I. Background: Characterizing Electricity Market Transactions

Before discussing the specific implementation of various transactions in electricity markets it is useful to define the different types of transactions that are seen in electricity and other commodity markets. A spot market transaction is generally defined as a payment for the immediate exchange of a good. Such a definition applies to everything from grocery store purchases to sales in real-time electricity markets. A forward transaction is an agreement or immediate payment for a promise for the delivery of the commodity at pre-specified date in the future. In most cases forward transactions are specifically tailored to the needs of the buyer and seller and negotiated bilaterally (or over-the-counter).
Forward market prices are heavily influenced by expectations about what the underlying spot price will be at the date of delivery of the forward contract. This is because a spot transaction at the delivery date is an alternative to the contract forward trade. A seller would prefer to wait to sell in a spot market, if the forward price is below its expectation of the spot market at the delivery date. Buyers are in the opposite position. If buyers and/or sellers are risk-averse, perceptions about the amount of spot price risk a market participant faces also influences the relationship between the forward market price and the expected spot price at the delivery date of the forward contract.

In general, if there are a sufficient number of risk-neutral market participants on both sides of the market, the forward price should equal the expectation of what spot prices will be at delivery date of the forward contract. Therefore, at any give point in time, the spot price will almost certainly not be the same as the corresponding forward price. However, if risk is relatively balanced amongst buyers and sellers, forward prices can be expected to be equal to spot prices on average.

A futures contract is a specific form of forward contract transaction, one that is standardized to meet specific definitions of quantity, delivery time and location. Futures are traded under standardized terms and conditions on organized exchanges, which assume the role of counter-party to every trade. In other words, the exchange assumes the risk of non-payment or non-delivery, freeing individual traders from the burden of verifying the reliability of their trading partners.

In the electricity context, real-time (or balancing) markets run by Independent System Operators (ISOs) and their equivalents are spot markets. ISOs must maintain a constant balance of supply and demand in real time to ensure the reliability of the transmission grid. If an ISO detects that there is likely to be too little or too much power being injected into a network, it will immediately buy power from a set of standing offers (bids) made by generators, or in effect sell back power to those generators by instructing them to reduce their output. All electricity markets have mechanisms for maintaining system balance in real-time. Some of these involve explicit market mechanisms; others are hybrids of market and regulatory mechanisms that penalize net energy purchases or sales between the close of the short-term forward market and real-time system operation.

Virtually all ISOs around the world run formal or informal day-ahead markets. The California ISO will have a formal day-ahead market once its Market Redesign and Technology Update (MRTU) process is completed. In day-ahead markets, buyers and sellers offer to supply or bid to purchase electricity that will be consumed at a specific time and location the following day. Day-ahead markets are considered beneficial because many sources of electric power supply need advance instructions in order to efficiently and reliably adjust their production levels. If major adjustments to overall supply were limited to the balancing market, some units would be limited or completely unable to respond to market conditions. Alternatively, there would be serious costs consequences if firms were expected to rigidly adhere to their forward positions, with no adjustment for unforeseen events (such as weather and fuel price changes).
Day-ahead markets are forward transactions. A seller or buyer is making a commitment to consume or supply energy the following day. In this sense, these transactions are purely financial because no physical exchange of goods is made at the time of the transaction. However, because of the need for advance notice to bring on generation units, there is a both a reliability and economic cost associated with large deviations between a forward market financial commitments and actual real-time production and consumption. This concern has led some system operators and regulators to argue that day-ahead and hour-ahead market should be physical despite the fact that the opportunity exists to undo any commitment made in these markets in the real-time market.

II. Financial versus Physical Transactions in California’s Markets

Since their inception, California’s short-term wholesale electricity markets have been officially considered to be “physical” markets. For the reasons discussed above, this definition was at odds with the reality of market operation. The only electricity market that is purely physical is the ISO’s real-time energy market. This market involves transactions for the immediate supply or consumption of energy.

Until early 2001, the California Power Exchange (PX) operated a day-ahead forward market commitment to supply or consume energy. Because there is no way to practically match the electrons transacted to the parties involved, a failure to meet a commitment made in the PX had to be reconciled in the balancing market. Firms could in practice reverse their PX positions in the imbalance energy market without paying additional penalties beyond the replacement costs for fulfilling their PX commitment out of the ISO’s imbalance market. For this reason, the PX was effectively a purely financial market.

The matter of discouraging ‘financial’ trades can be distilled down to how one penalizes a failure to meet forward commitments, or how one prevents certain parties from making forward transactions in the first place. The sum of such policies in California was a haphazard application of a physical standard. Technically, only firms with actual generation plants were to submit advanced schedules to the ISO. The ISO however, had no ability, then or now, to distinguish between schedules that will be carried out in real time or those that the market participant plans to undo in the real-time market.

This is particularly, the case for the interties into the California market. The only thing the ISO is able to see in the day-ahead market is a schedule at an intertie (i.e. a commitment to inject or withdraw a certain amount of power from a point that connects the CAISO control area with a neighboring system). It has no way of knowing what generation unit located outside of California will supply this energy. The restriction against financially selling energy in the forward market could (and still can) be circumvented by simply scheduling more energy in the day-ahead market from a generation unit or intertie than the market participant intends to supply in real-time. A restriction on buying energy without physical demand can be circumvented by
purchasing more energy than the market participant intends to consume in real-time. Until recently, the ISO made no attempts to restrict financial purchases from forward markets and did not penalize for the resulting ‘over-scheduling’ of demand.

The process of market redesign and the planned addition of a day-ahead market have reopened the question of how to deal with financial transactions. The historical practice with regards to purely financial transactions was therefore to officially forbid them, but to tolerate them in practice. One of the biggest difficulties in the past with enforcing a ban on financial transactions was that all but the extremely large or idiosyncratic financial trades are impossible to distinguish from physical trades. A purely financial transaction appears as a deviation in the balancing market. But so do transactions by buyers with erroneous load forecasts or suppliers that suffer from forced outages. It is important to accommodate these latter, ‘legitimate,’ imbalances.

Excessive penalties for such deviations would increase the costs of participating in the real-time market without increasing reliability. Usually ISOs allow for deviations within some ‘reasonable’ bound with no penalty or they apply a graduated schedule of imbalance penalties, with small imbalances paying little or no penalty. In the case of the England and Wales New Electricity Trading Arrangements (NETA) market, there was a deliberate decision to make the cost of deviations from final schedules extremely expensive to discourage transactions in the NETA balancing mechanism.

II.1. Consequences of Inconsistent Policies in California

As noted above, a number of markets, including California, have taken a “don’t ask don’t tell” approach to the question of physical versus financial transactions. Although short-term forward markets are officially considered physical, there are limited explicit penalties for deviating from day-ahead commitments. Thus firms could buy or sell power in the day-ahead market and reverse their position (i.e. sell or buy back their day-ahead commitment) by simply not performing according to their schedule. A firm’s costs or compensation would be based upon the imbalance energy price.

For example, a firm could bid to buy power in the day-ahead market and then simply not consume it. That firm would pay the day-ahead price for its purchase and earn back the imbalance energy price because its non-consumption can be viewed as selling power back to the market in real-time. In this way a firm could “arbitrage” expected price differences in between the two markets. Such transactions are not necessarily harmful and can be beneficial to markets. However, under some circumstances such trades can create problems.

Large volumes of transactions in the imbalance market that are not supported by actual physical supply or demand can create reliability problems for the ISO if the ISO is expecting physical resources to be available. Imbalance energy markets are designed to handle relatively small volumes of physical transactions. When large-scale changes to system conditions arise because physical conditions on the grid turn out to be quite different from those implied by the day-ahead market, the ISO is forced to call upon
generation and possibly even load to perform in ways that are extremely costly and potentially unstable. Relatively few resources are nimble enough to respond to large-scale changes with little advance warning. This is why most systems have a day-ahead market in the first place.

Second, markets are more vulnerable to strategic behavior of market participants under a regime where financial trades are tolerated only if supported by a fiction of physical resources. This is because a limited number of players are able to undertake such transactions. Only firms that can credibly represent their trades as physical can attempt to undertake such pseudo-physical trades. By restricting the number of players, it becomes much easier for any single market participant to influence market prices through their unilateral actions.

One concern about limiting the number of market participants able undertake purely financial transactions is that it can distort day-ahead prices relative to the true real-time price. Under ideal conditions, day-ahead and imbalance energy markets would display a relationship like that illustrated in Figure 1. Supply is offered into the day-ahead market along the supply curve \( S \), and demand is bid along the demand curve \( D \). Both supply and demand are bid in a way that results in a market clearing price consistent with what the real-time price would be in expectation. In the real-time imbalance market, an unexpected change in demand would result in movement along a similar supply curve, yielding a price that could in fact be higher or lower than the day-ahead price. Under these conditions, any difference between day-ahead and real-time prices is the result of random, unexpected changes to market conditions that arise after the day-ahead market closes.

If there are restrictions on the number or type of market participants in short-term forward markets, a large buyer in the day-ahead market may find it profitable to reduce its day-ahead purchases and thereby lower the price in the day-ahead market (see Figure 2). If there is no response by any other market participants (such as suppliers) this would

![Figure 1: 100% Scheduling of Demand in Forward Market](image-url)
lower overall procurement cost because only the portion of real-time demand not purchased in the day-ahead market is pays the “unbiased” market clearing price $E(P_s)$. The rest is purchased at the lower forward price $P_f$.

While such strategies could successfully distort prices for a time, there is reason to believe that a significant price difference between the day-ahead and forward markets would not persist. One natural reaction would be for suppliers to raise their offer prices into the day-ahead market, thereby counteracting the reduced purchases by demand. Alternatively, other firms may attempt to arbitrage such a price difference by bidding in fictional demand into the day-ahead market and implicitly selling it back into the real-time market.\(^2\)

![Figure 2: Strategic under-purchasing](image)

Such a procurement strategy raises several concerns. Reliability could be reduced due to the increased reliance on the real-time market. If the strategy succeeds in reducing overall procurement costs and the market is competitive on the supply side, it can eventually lead to problems with supply investment, as lower-cost generation units would lose revenues if they sell in the day-ahead market.

While these outcomes are clearly a concern, several aspects of the current California market greatly mitigate the incentive for load-serving entities (LSEs) to execute such a strategy. Most important, if close to 100% of the LSE’s expected real-time demand has been hedged in long-term forward financial contracts negotiated in advance of the day-ahead market, LSEs will have little incentive to influence the day-

---

\(^1\) It is worth noting that a similar result can arise from an implementation of a “soft” bid cap in a market. Under a soft cap, firms with offers above a certain level would earn a higher price, but the bulk of the market would earn only the capped level. If actual incremental production costs rise above the cap level, then buyers are effectively buying only the residual (above cap) power at the “true” market-clearing price. The bulk of their purchases may be held to a price below this level by the cap.

\(^2\) This last strategy is effectively what was described as the “fat-boy” strategy employed by Enron.
ahead price through this sort of behavior. Just as forward contracts mitigate the incentives of suppliers to exert their market power in short-term markets, long-term contracts limit the incentives of LSEs to do the same. Second, current resource adequacy mandates have must-offer requirements on generation units that also limit the incentives of LSEs to bid to reduce the quantity of energy scheduled in the day-ahead market. Third, the addition of the residual unit commitment (RUC) process under MRTU will further penalize load that consistently purchases from the imbalance energy market, as the cost of procuring RUC capacity will fall primarily upon the load with schedules below their actual consumption.

III. Scheduling and Deviations under MRTU

Unlike the period from 1998-2000, the California ISO’s market redesign (both the current Phase Ib design and the forthcoming MRTU) provides for penalties for uninstructed deviations by suppliers. Although there is still some uncertainty as to whether these penalties will be imposed, the ISO plans to impose penalties on suppliers who produce less than 97% or more than 103% of their combined scheduled and instructed production levels at each individual plant. The ISO will allow for substantial deviations from forward schedules that are based upon declared physical problems with the generation facility. Of course this creates a serious verification problem. Many have questioned the veracity of declared generation outages during the period June 2000 to June 2001, although meaningful penalties have only been applied in the extreme cases where explicit documentation of withholding has been produced.

The output capability of a given generation unit at any given moment is in large part a judgment call on the part of the unit’s operator. It will be very difficult for the ISO or other authorities to second-guess such judgments ex-post. In the absence of virtual bidding, however, there will be a powerful financial incentive for firms to test the limits of the ISO’s verification of declared unit capabilities.

Like over-generation, over-procurement (or under-consumption) will also be potentially subject to penalties from the ISO. The ISO’s oversight and investigations protocol calls for it to investigate instances where a firm consumes less than 95% of its scheduled power purchases. If such behavior is judged to be inconsistent with legitimate sources of demand uncertainty (i.e. an unexpected drop in temperature), it could be subject to a $10,000 fine. As with the verification of generation capabilities, there is an increased burden placed on the ISO to determine the intent of various scheduling practices. It is not clear at this time what standards the ISO might apply, or the impacts of such standards on the behavior of firms.

There will not be explicit penalties for consuming “too much” power in the imbalance energy market. The management of the ISO is now satisfied that the introduction of residual unit-commitment (RUC) as part of the MRTU, eliminates the concern that under-procurement creates a serious reliability problem. In essence the ISO will dispatch units according to its own load forecasts rather than according to final schedules and charging all LSEs that are “under-scheduled” relative to their real-time
consumption for the resulting RUC costs. In this sense, the RUC costs will become a tax on real-time consumption. At this time, it is difficult to predict how significant these charges will be. It is important to note that resources that have been contracted for under resource adequacy requirements will not be eligible to earn RUC payments. If the RUC process impacts mostly RA units, RUC costs on LSEs may not be very significant.

III.1. Scheduling Requirements

One proposed solution to the concern that LSEs may attempt to reduce day-ahead energy prices is a 95% day-ahead scheduling requirement. Such requirement would in theory prevent a LSE from lowering its day-ahead purchase quantity far below its expected actual demand. However, the practical enforcement of such a restriction is still problematic. In essence, any policy imposing restrictions on forward transactions must balance the need for accurate monitoring and meaningful enforcement with the potential severity of the restrictions on legitimate commercial practices. If the restrictions necessary to fully eliminate financial transactions impose serious restrictions upon market participants, the costs must be weighed alongside the benefits of such restrictions.

A core difficulty in enforcing a scheduling requirement is the need to distinguish between an intentional reduction in purchases and an error, or difference of opinion in a forecast. The approach currently employed by the ISO requires LSEs to schedule to within at least 95% of their own load forecast. However, this can create a situation where compliance is relatively straightforward, but the scheduling requirement is effectively meaningless. The LSE can always say that its load forecast was less than or equal to its actual day-ahead schedule divided by 0.95. This would guarantee compliance with the requirement. Over time, the ISO may detect a systematic error in a forecast. Therefore such a policy would require a standard of accuracy from LSE forecast that would almost certainly allow for some room for error.

Given these practical considerations, however, it is not clear how effective such a policy would actually be at limiting the ability of LSE’s to strategically lower day-ahead prices. Even if the ISO required 100% accuracy on average in forecasts over some period, such as six months, LSEs could still benefit from strategically missing their forecasts during key high price periods, and compensating by missing in the opposite direction during other periods. Even a 5% reduction in demand during some high price periods, where the supply curve might be very steep, could result in a non-trivial reduction in price. Thus, if eliminating the potential for such strategies were the primary concern, even more severe restrictions may be necessary. However, these more severe restrictions carry more serious cost implications.

For example, the ISO could require LSEs to schedule to within 95% of the ISO’s load forecast for that LSE. However, this requirement would rightly raise concerns among LSEs that they would be required to pay the real-time market costs associated with errors in the ISO’s load forecast. Alternatively, the ISO could measure compliance by comparing a LSEs actual load or generation output relative to its day-ahead schedule. For generation this creates a circumstance where the generation unit owner would be
reluctant to respond to a real-time dispatch instruction because this could make their actual output multiplied by 0.95 exceed their day-ahead schedule. For LSEs, validation of compliance with the 95% percent scheduling requirement using their actual consumption compared to their day-ahead schedule would provide strong incentives for LSEs to implement demand response programs.

There are also market efficiency consequences of a 95 percent scheduling requirement. California is increasingly import-dependent. Many of these imports only become available at the close of the day-ahead market, or one hour before the imbalance market is run. This is why the CAISO currently allows for adjustments to schedules one-hour in advance of real-time. A 95% day-ahead scheduling requirement could raise wholesale energy purchase costs because it would limit the magnitude of purchases California LSEs are allowed to make between the close of the day-ahead market and real-time market.

Thus, penalties severe enough to discourage financial trading may raise costs significantly. The UK market, which provides a penalty for any deviations from day-ahead schedule on an hour-by-hour and plant-by-plant basis, provides a lesson in this regard. Plant operators in the UK market have adopted a practice of holding some of a plant’s capacity in reserve in order to better ensure hitting its day-ahead scheduling target. A system that leads each plant to provide its own operating reserve works against the benefits of pooling operations and sharing reserves over larger systems.

**IV. Limiting Financial and System Reliability Concerns**

In the past, a policy that treated all transactions as if they were physical created potential reliability problems. Apparent predictable price differences between markets created a strong incentive for firms to attempt to arbitrage those price differences. Since explicit arbitrage trades were not officially allowed, such trades were executed under a façade of a physical transaction. However, by representing that these arbitrage trades were in fact commitments to physical generate or consume power, reliability was potentially threatened. System operators need sufficient resources available at various locations in the control area at various time horizons in advance of real-time operation. One role of the day-ahead market and hour-ahead market is to ensure that these reliability requirements are met. For market efficiency and reliability reasons, it is desirable that system operators see an accurate picture of the physical condition of network resources in advance of real-time.

A large number of trades that imply the availability of physical resources that are not in fact available can threaten reliability. One way to limit such a problem would be to seriously penalize deviations from positions implied by forward trades. However, as described above, to be fully effective such limitations may have to be quite severe. These severe penalties would likely catch many legitimate deviations along with any attempts at purely financial trades.
Another approach for dealing with the reliability problem is to allow explicitly financial transactions between short-term forward markets and the real-time market. This approach redirects the pressure created by profitable trading opportunities into an outlet in which such trades no longer threaten system reliability. These kinds of trades, which have been referred to as *virtual*, or *convergence* bidding, would be explicitly identified as having no specific relation to physical resources. System operators can then focus on the physical trades when evaluating the reliability of the network, as the incentive to misrepresent an arbitrage trade as physical would be removed. Further, since allowing such trades will decrease the likelihood of persistent predictable differences between day-ahead and real-time prices, the incentive to hold back physical resources for the real-time market will be reduced. Both markets will present roughly the same opportunities for buyers and sellers.

To the extent that there are remaining concerns with large players taking large financial positions in the day-ahead market, limits on the size of financial positions have merit in the context of electricity markets. This is an important contrast to the existing policy, which effectively grants a franchise on making such trades to a limited number of physical market participants. The existing policy just bestows an additional level of market power on those firms; market power in financial trading. Financial transactions from these firms would continue to be treated as physical by the ISO, again creating potential reliability problems. We instead suggest that *all* financial trades be explicitly identified as such by accommodating such trades within a set of market protocols, rather than forcing such transactions to appear as physical trades that fail to perform.

One of the concerns raised by some stakeholders is that accommodating financial trades would legitimize and facilitate market manipulation. However, the accommodation of financial trades can be one of the most effective means of deterring behavior that might be termed market manipulation. The conventional view among economists is that in a market with more participants it is more difficult for any one to move the market price through their unilateral actions than in market with fewer participants. If financial transactions must be identified in advance, then it is possible for the ISO to punish substantial deviations from forward market schedules from physical resources such as generation units and interties into California without adverse market efficiency consequences.

**IV. Summary**

When products for the delivery of the same good to the same time and location are traded in different markets, there are at times profitable opportunities to trade on the prices in the different markets. In other words, if one market has a predictably lower price for the same product than another market, firms will want to buy low and sell high until such differences vanish. In electricity markets, this trading pressure has coincided with rhetorical restrictions that trades have to be among “physical” participants. The combination of profitable trading opportunities and a restriction to physical trades has led
some firms to engage in trades that they represented as backed by physical assets, when in fact they were not.

Another form of trade of potential concern is the attempt to influence prices by reducing purchases in day-ahead markets. Large buyers could in theory reduce day-ahead prices by shifting demand out of the day-ahead market. However, it is unlikely such price differences would persist as one would expect supply to also shift from the day-ahead to the imbalance energy market. While such a reaction would restore the relationship between forward and expected imbalance energy prices, it would also result in higher volumes in the real-time market. Like more conventional strategies for exercising market power, the incentive to influence price by reducing purchases in the day-ahead market will be greatly reduced if most of the energy purchased by LSEs is bought under forward contracts, rather than in the day-ahead or real-time market.

Two alternative approaches have been proposed for dealing with these problems. The first would attempt to apply meaningful enforcement to the restriction that trades must be backed by physical resources. Penalties great enough to offset any gains from financial trades would be sufficient to discourage such trades. However, in order to be effective, such penalties would have to be applied with performance standards strict enough that they would also likely have a negative impact on legitimate operations and market opportunities. The second alternative is to allow explicit arbitrage transactions that would be identified to operators as trades that are not backed by physical resources. Firm specific position limits for such transactions could be used to mitigate concerns that some firms could artificially influence prices with such transactions.