Impact of virtual bidding on financial and economic efficiency of wholesale electricity markets

Wenyuan Tang[1,2*] Ram Rajagopal[2] Kameshwar Poolla[1] Pravin Varaiya[1]

Abstract

Virtual bidding is a financial mechanism that allows participants to speculate on the price spread between day-ahead and real-time wholesale electricity markets. We employ statistical learning methods to show that virtual bidding improves financial efficiency by narrowing the room for arbitrage. We then formulate a two-settlement market model and use it to analyze market data. We find that virtual bidding improves economic efficiency by reducing the system generation cost. Notwithstanding the empirical evidence of its effectiveness, virtual bidding has not fully eliminated market inefficiency, and we propose strategies that extract profits beyond those of the actual virtual bids. Lastly, we conduct a game-theoretic analysis based on the two-settlement model to develop the theory of virtual bidding. The analysis leads to the interpretation of spread as a measure of the average forecast accuracy of the market participants, and implies that introducing more qualified virtual bidders into the market can improve market efficiency.

The restructured electricity markets in the United States feature a two-settlement system at the wholesale level, consisting of a day-ahead (DA) forward market and a real-time (RT) spot market. Systematic nonzero spreads, defined as the differences between DA locational marginal prices (LMPs) and RT LMPs, are routinely observed, and signify market inefficiency. Virtual bidding was introduced as a financial mechanism to allow market participants including outside financial entities to speculate on the spread, without physically consuming or producing power. While the statistics of spreads across different regional markets and over different time horizons has been studied [1–11], the impact of virtual bidding on market efficiency has not been explored until recently. It is acknowledged that virtual bidding has improved price convergence [12]. A statistical framework is developed in [13] to test the existence of profitable bidding strategies. In [14], a hidden Markov model is proposed to characterize the stochastic process of the spread and to solve for optimal bidding strategies.

The aforementioned work focuses on price convergence, which we refer to as *financial efficiency*. Following a statistical learning approach, we propose a couple of measures that test the financial efficiency of the market, before and after the implementation of virtual bidding, and we design bidding strategies that outperform the actual virtual bids. Thus our findings agree with prior work that virtual bidding has reduced but not eliminated arbitrage opportunities. More fundamentally, we investigate the impact of virtual bidding on *economic efficiency*, which refers to generation cost minimization. To that end, we

[1]Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, CA 94720, USA.
[2]Department of Civil and Environmental Engineering, Stanford University, Stanford, CA 94305, USA.
[*]e-mail: tangwenyuan@berkeley.edu

formulate a two-settlement market model that captures the flexibility of generation, and addresses various phenomena observed from market data, in contrast with existing models [15–17]. Our proposed model provides a theoretical framework for understanding the interactions between DA and RT markets, and suggests a methodology for estimating the system generation cost. We then provide empirical evidence of the effectiveness of virtual bidding. Lastly, we conduct a game-theoretic analysis based on the two-settlement model, and develop the theory of virtual bidding. Our study interprets the spread as a measure of the average forecast accuracy of the market participants, and implies that introducing more qualified virtual bidders into the market can improve market efficiency.

Exploring market data

Our work is based on market data from the California independent system operator (CAISO) and PJM Interconnection (PJM). CAISO implemented the DA market in April 2009, and the virtual bidding mechanism in February 2011. This allows us to compare price performance before and after the implementation of virtual bidding. In particular we use the DA and RT LMPs at the CAISO NP15 trading hub for 2010 and 2012. PJM implemented the DA market and the virtual bidding mechanism concurrently in June 2000. We use PJM regional level data from the years 2012–2015, including DA and RT LMPs, DA and RT load, forecast load, and virtual bids. The additional data from PJM allows more sophisticated statistical analysis.

Summary statistics of DA and RT LMPs are shown in Fig. 1. While the DA and RT LMPs are very close on average, the standard deviation of the RT LMP is considerably larger than that of the DA LMP, indicating higher volatility in the RT market. Fig. 2a plots the time series of the spread for an off-peak hour (3–4 AM) and an on-peak hour (5–6 PM). The spread is large positive when the price spike occurs in the DA market, and large negative when the price spike occurs in the RT market. The distribution of the spread is shown in Fig. 2b. While the mean is close to zero, the distribution is heavy-tailed. It is also left-skewed, since price spikes are more likely to occur in the RT market than in the DA market. In general, the spread is difficult to forecast, and the characterization of its statistics is challenging. For example, it can be shown that the recent data from PJM does not support the classic Bessembinder and Lemmon model [15], which states that the spread is negatively related to the variance of the RT LMP, and positively related to the skewness of the RT LMP.

There are two types of virtual bids: incremental (INC) bids and decremental (DEC) bids. An INC bid is submitted as a supply bid in the DA market. It opens a short position in the DA market, with the obligation to be closed out in the RT market. The holder is paid the DA LMP and pays the RT LMP per cleared MWh. On the other hand, a DEC bid is submitted as a demand bid in the DA market. It opens a long position in the DA market, with the obligation to be closed out in the RT market. The holder pays the DA LMP and is paid the RT LMP per cleared MWh. Given the virtual bids and the LMPs, we obtain the cleared INC and DEC volumes, and then the net INC volumes as their difference. When the net INC volume and the spread have the same sign, the virtual bids make a net profit on aggregate, as shown in Fig. 3. The net INC volume is predominantly negative, which suggests that the virtual bidders speculate on negative spreads most of the time. The reason for this may be the left-skewness of the distribution of the spread: when the spread is large in magnitude, it is more likely to be negative than positive.

To take a closer look at the profit pattern of the virtual bids, we divide the 8760 hours in a year (or

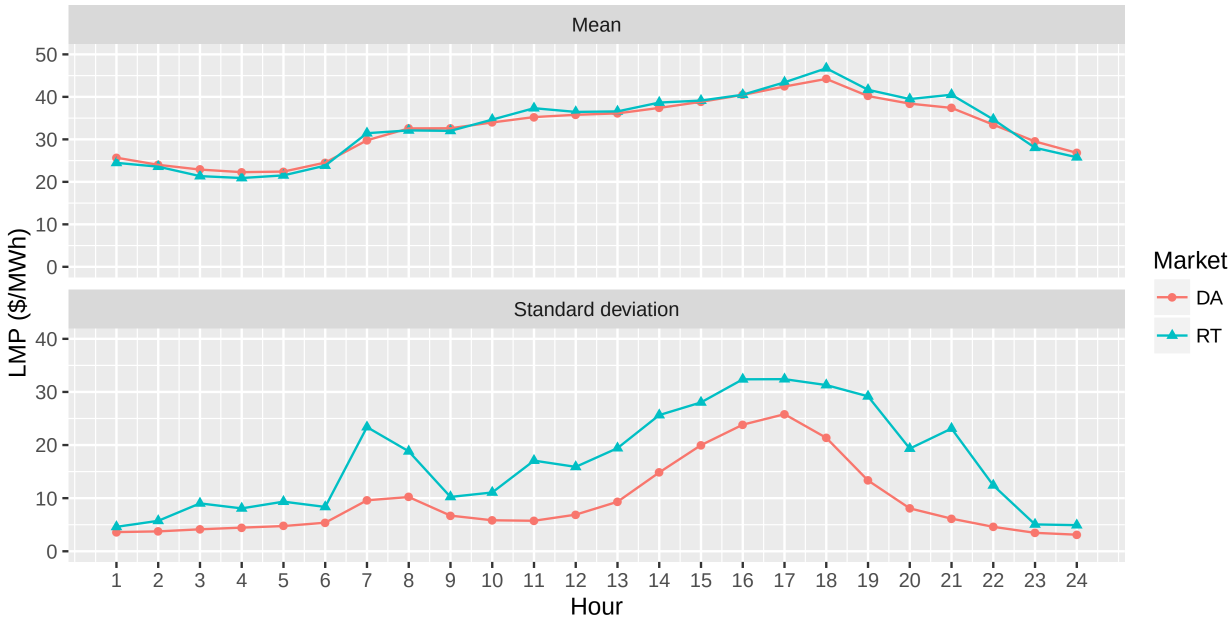
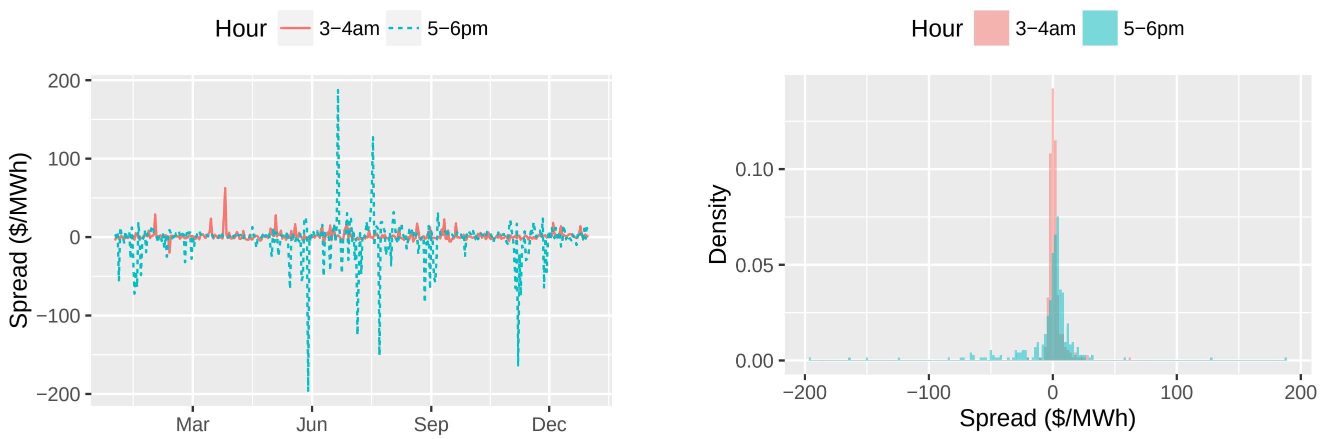


Figure 1: *Statistics of DA and RT LMPs, PJM, 2012.* While the mean of the RT LMP is close to that of the DA LMP, the standard deviation of the RT LMP is considerably larger.



(a) (b)

Figure 2: *Statistics of spread, PJM, 2012.* (a) The time series of the spread. (b) The histogram of the spread. The distribution of the spread is heavy-tailed and left-skewed.

8784 hours in a leap year) into two groups: normal hours, during which the spreads are between the 1st and the 99th percentiles; and abnormal hours otherwise. We examine the cumulative profits for all hours, normal hours, and abnormal hours, as shown in Fig. 4. The curves change rapidly, implying the high risk of virtual bidding, which arises from the heavy tails of the distribution of the spread. More interestingly, the virtual bids make small or even negative profits during the normal hours, but large profits during the abnormal hours. This suggests that the virtual bidders mainly speculate on extreme events, when the spreads are large in magnitude. We use the Sharpe ratio to evaluate the performance of the virtual bids, and find that the four-year Sharpe ratios for all hours, normal hours, and abnormal hours are 1.79, -1.29, and 2.56, respectively. As a benchmark, the Sharpe ratio of the S&P 500 index during the same period is

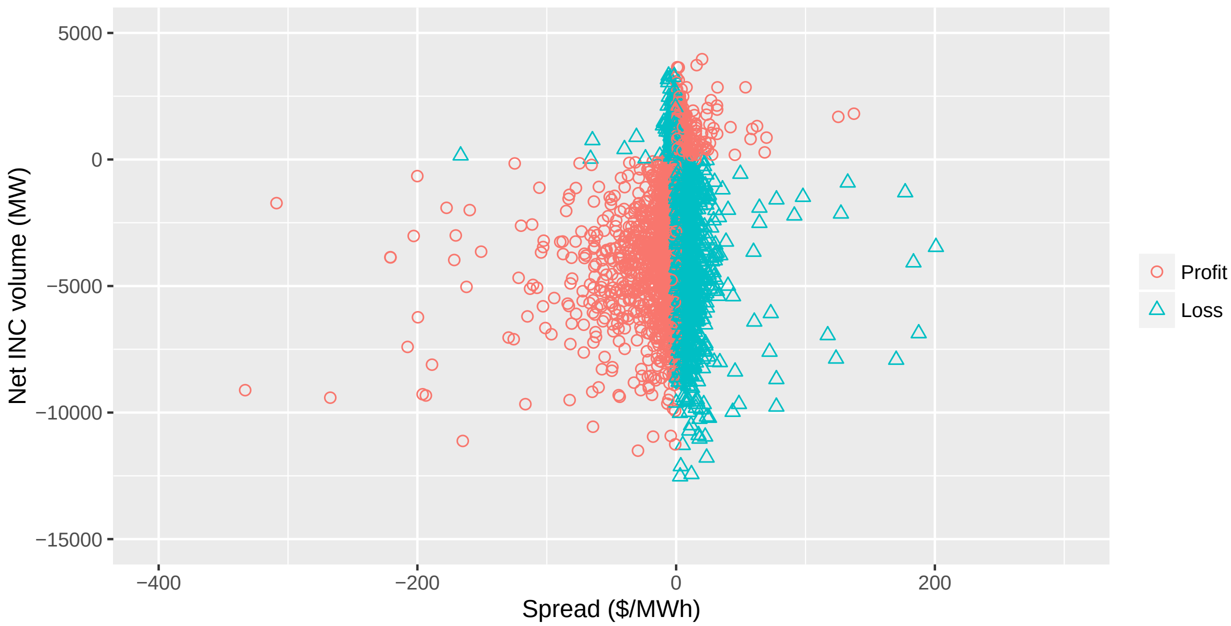


Figure 3: *Hourly net INC volume versus spread, PJM, 2012.* The net INC volume is predominantly negative, which suggests that the virtual bidders speculate on negative spreads most of the time.

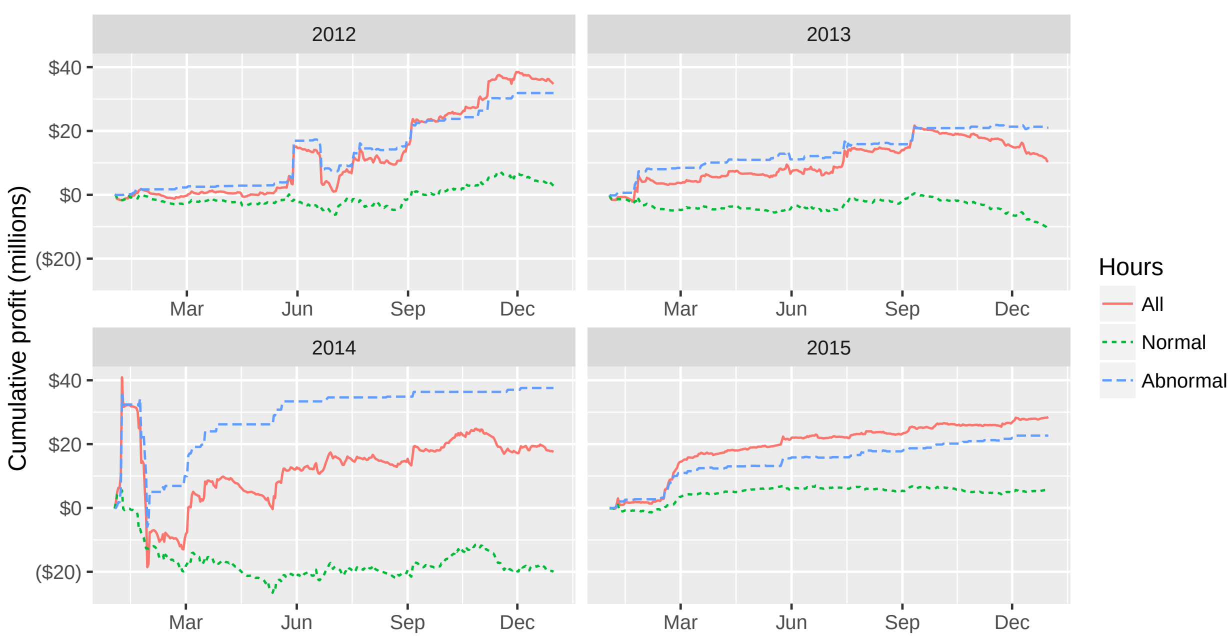
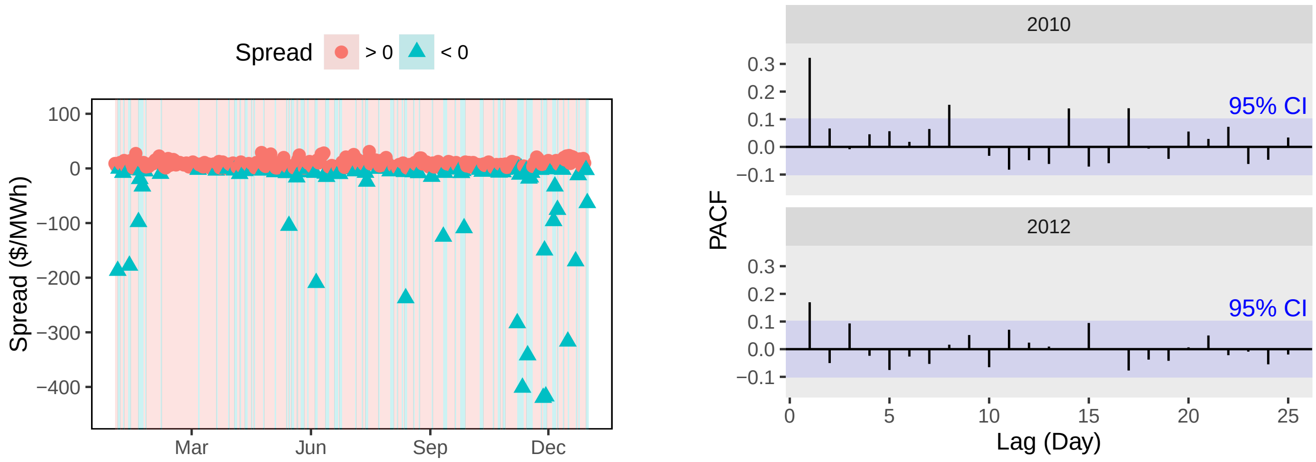


Figure 4: *Annual cumulative profit of virtual bids, PJM.* The virtual bids yield high return along with high risk. More interestingly, the virtual bids make small or even negative profits during the normal hours, but large profits during the abnormal hours.

1.68. While the Sharpe ratios are comparable, the virtual bids yield high return along with high risk.



(a) (b)

Figure 5: *Measures for testing the room for arbitrage.* (a) Runs test on the sequence of the sign of the hourly spread, CAISO, 2010, 5–6pm. A run is a segment of consecutive positive or negative spreads. The runs are highlighted by alternate color bars. Given the number of positive spreads (which is 268) and the number of negative spreads (which is 97), and under the null hypothesis that the sequence is independent and identically distributed, the number of runs (as the test statistic) is normally distributed, with the mean 143 and the standard deviation 7.4. The actual number of runs is 110, with a p-value less than 0.001, which strongly rejects the null hypothesis. (b) Partial autocorrelation of the daily average spread, CAISO. Overall, the sequence in 2012 is less autocorrelated than that in 2010. For 2012, the first lag is the only lag for which the partial autocorrelation is statistically significant at the 5% level.

Profitable bidding strategies and financial efficiency

Financial efficiency concerns whether there is opportunity for arbitrage. In practice, the mean of the spread is typically nonzero, so there is a question as to how to quantify the level of inefficiency. In fact, calculations from CAISO data show that the mean of the spread is closer to zero in 2012 than in 2010. While this is plausible evidence of the effectiveness of virtual bidding, it is mystifying that the variance of the spread remains high. As an alternative approach, we propose several measures and bidding strategies to gauge arbitrage opportunity.

The first measure is based on the Wald-Wolfowitz runs test, which tests the hypothesis that the elements of a two-valued sequence are independent and identically distributed. We do the runs test on the sequence of the sign of the spread, and take the p-value as the measure, as shown in Fig. 5a. A greater p-value indicates that the sequence of the signs is more random, which leaves less room for arbitrage. As we can see from Table 1, for most hours, the p-values for 2012 are greater than those for 2010, indicating that virtual bidding has improved financial efficiency.

The second measure is based on the partial autocorrelation of the sequence of the daily average spread shown in Fig. 5b. The sequence in 2012 is less autocorrelated than that in 2010—another indication that virtual bidding reduces arbitrage opportunity.

Inspired by the fact that the sequence of the daily average spread has a strong lag-1 (day) correlation, we propose a benchmark bidding strategy that is only based on the up-to-date price information. For operating day t, bids have to be submitted by noon of day $t-1$. Hence, we add up the spreads for the afternoon of day $t-2$ and the spreads for the morning of day $t-1$, and take the sign of the sum.

Table 1: Results from runs test, CAISO

Hour	2010	2012	Hour	2010	2012
1	**0.016**	0.331	13	**0.000**	0.074
2	**0.026**	**0.028**	14	**0.000**	0.123
3	0.161	**0.020**	15	**0.002**	**0.002**
4	**0.002**	0.625	16	**0.000**	**0.002**
5	0.063	0.352	17	**0.002**	**0.002**
6	**0.050**	0.998	18	**0.000**	**0.001**
7	**0.000**	0.561	19	**0.000**	**0.033**
8	**0.021**	0.730	20	**0.000**	**0.000**
9	**0.016**	0.501	21	**0.006**	0.323
10	0.524	0.066	22	**0.000**	**0.000**
11	**0.000**	**0.001**	23	**0.000**	**0.005**
12	**0.000**	**0.001**	24	**0.002**	0.703

The test results are given in terms of p-values. Those which are less than or equal to 0.05 are highlighted in bold. For most hours, the p-values for 2012 are greater than those for 2010.

If it is positive, we trade 1 MW INC for each hour on day t; and 1 MW DEC otherwise. Under the lag-1.5 algorithm (where the name is self-explanatory), the average profit is \$1.35/MWh for 2010, and \$1.08/MWh for 2012; and the Sharpe ratio is 1.93 for 2010, and 1.35 for 2012. As expected, it is more difficult to speculate after the implementation of virtual bidding, which suggests that virtual bidding has improved financial efficiency.

Besides the price data, we can also utilize other data such as load to design more sophisticated bidding strategies. In particular, we employ support vector machines (SVMs), specified in Methods, to forecast the sign of the spread, and then trade like in the lag-1.5 algorithm. Fig. 6 compares the profit patterns between the actual virtual bids and the SVM algorithm. While the traded volumes are in different scales, the use of the Sharpe ratio makes their performance comparable. Overall, the SVM algorithm outperforms the actual virtual bids. It is worth mentioning that the actual virtual bids profit in normal hours and lose in abnormal hours, whereas the SVM algorithm does quite the opposite. Our explanation is that virtual bidders may speculate on extreme events, while machine learning based algorithms aim to capture the key patterns in the data. Nevertheless, there is still room for arbitrage: even a simple machine learning algorithm can extract profits beyond those of the actual virtual bids.

Two-settlement market model and economic efficiency

To investigate the impact of virtual bidding on economic efficiency, we formulate a two-settlement market model. In the DA market, the supply consists of the physical supply, which is the DA generation, and the financial supply, which is the cleared INC volume; the demand consists of the physical demand, which is the DA load, and the financial demand, which is the cleared DEC volume. In the RT market, there is

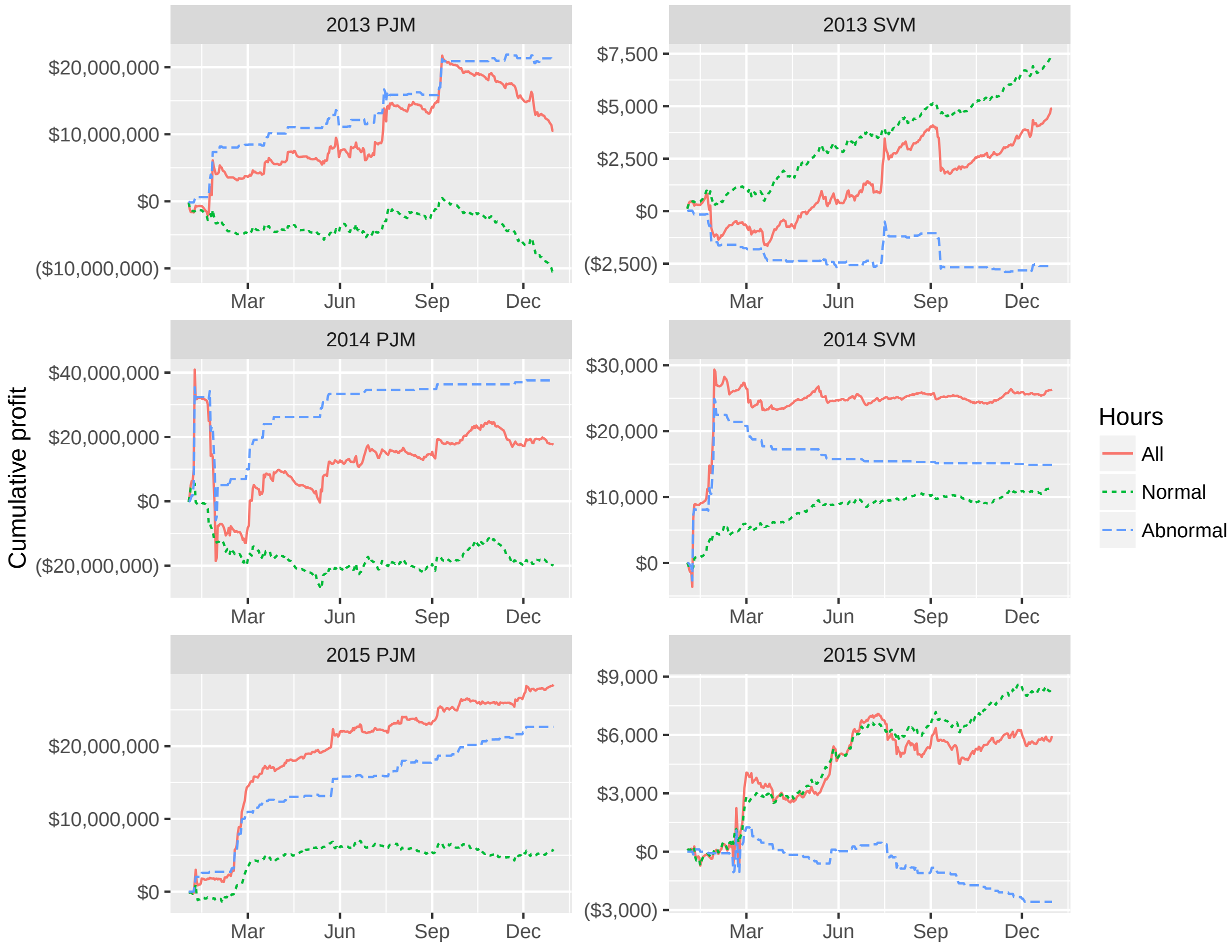


Figure 6: *Profit comparison between the actual virtual bids and the SVM algorithm, PJM.* For the actual virtual bids, the three-year Sharpe ratios for all hours, normal hours, and abnormal hours are 1.19, -1.60, and 1.97, respectively. For the SVM algorithm, the three-year Sharpe ratios for all hours, normal hours, and abnormal hours are 2.17, 6.05, and 0.63, respectively.

only the physical supply, which is the RT generation, and the physical demand, which is the RT load. The supply and demand should match in both markets.

The DA cost curve, which specifies the DA LMP as a function of the DA generation, is formed by all the DA bids, including generation, load, and virtual bids. The RT cost curve, which specifies the RT LMP as a function of the RT generation, is formed in a more complex manner: it depends on the DA dispatch, the bids of the generators scheduled in the DA market, and the newly submitted generation bids in the RT market. When fitted by linear regression as shown in Fig. 7, the DA cost curve is a better fit, and the RT cost curve is steeper, both of which are consistent with the fact that the RT market is more volatile.

The higher volatility of the RT market has several causes, including: (i) The RT market runs on a much shorter time scale, the operation of which is subject to instantaneous realizations of various sources of uncertainty; (ii) The DA unit commitment and economic dispatch impose operational limits on the scheduled generators; (iii) In addition to the generators scheduled in the DA market, only fast-start generators are dispatchable in the RT market; (iv) Intra-hour ramping constraints have to be taken care of in the operation of the RT market; etc. In general, not all the DA generation is dispatchable in RT.

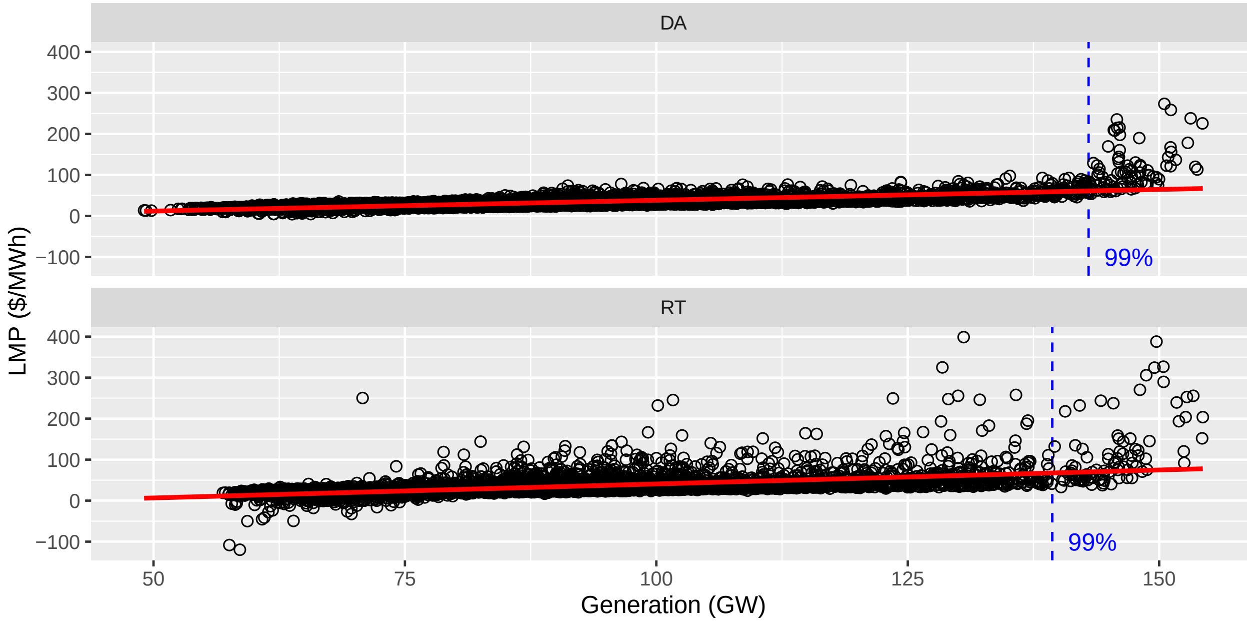


Figure 7: *LMP versus generation, PJM, 2012.* When fitted by linear regression, the DA cost curve is a better fit, and the RT cost curve is steeper.

As an abstraction, we classify generation as inflexible or flexible. Inflexible generation, once scheduled in DA, cannot be displaced in RT. Flexible generation can be either displaced (if scheduled in DA) or newly dispatched in RT (if not scheduled in DA). Note that a generation unit can have both inflexible and flexible segments. For example, the inflexible segment reflects the minimum output and the flexible segment reflects the difference between minimum and maximum output.

Since the DA dispatch dictates the schedule of the inflexible generation, the RT cost curve depends on the DA dispatch, which is the key point in the proposed model, described in more detail in Methods. The model specifies the system generation cost as the sum of the inflexible and the flexible generation costs, and aligns financial efficiency and economic efficiency: the generation cost is minimized if and only if the spread is zero. When there is no uncertainty in the RT cost curve, the spread is zero if and only if the DA generation equals the RT generation. In general, the market makes imperfect forecasts, in which case the spread can be viewed as a measure of the market's forecast accuracy, as in equation (5).

The fact that the market does not make perfect forecasts leads to possible efficiency loss. Virtual bids may help improve the efficiency. While virtual bids do not affect the RT generation, they affect the DA generation and therefore the DA LMP, the RT cost curve, the RT LMP, and the generation cost. We show that virtual bids are more effective in enhancing efficiency as the proportion of inflexible generation increases. Furthermore, the more competitive the virtual bids, the more likely that the resulting savings in generation cost outweighs payments to virtual bidders.

Based on the proposed two-settlement market model, we develop a methodology for empirical estimation, described in Methods, in order to investigate the effectiveness of virtual bidding. First, we estimate the parameters of the cost curves. Table 2 reports the results from the linear regression for the cost curves. All the coefficients and the intercepts are statistically significant at the 5% level. The R^2 for the DA cost curve is much greater than RT, indicating the higher volatility of the RT market. The proportion of

Table 2: Results from regression for the cost curves, PJM

Year	$\hat{p} = \hat{a}\hat{x}$		$p = ax + b$			$\eta = 1 - (\hat{a}/a)$
	$\hat{a}$ ($\times 10^{-3}$)	R^2	a ($\times 10^{-3}$)	b	R^2	
2012	0.37	0.93	0.68	−27.5	0.30	45.4%
2013	0.42	0.91	0.76	−31.8	0.34	44.6%
2014	0.56	0.55	1.72	−108.2	0.18	67.3%
2015	0.41	0.78	0.86	−44.7	0.28	53.0%

The DA cost curve takes the form $\hat{p}(\hat{x}) = \hat{a}\hat{x}$, where $\hat{x}$ is the DA generation, $\hat{p}$ is the DA LMP, and $\hat{a}$ is the parameter. The RT cost curve takes the form $p(x) = ax + b$, where x is the RT generation, p is the RT LMP, and a and b are the parameters. The proportion of inflexible generation η can be expressed in terms of $\hat{a}$ and a. Please refer to Methods for details.

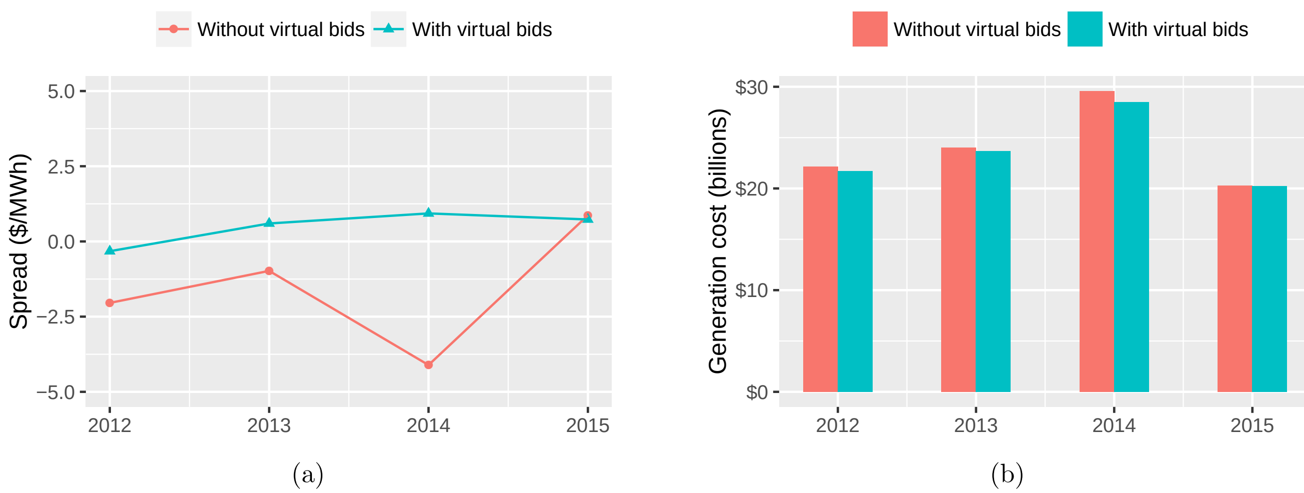


Figure 8: *Empirical evidence of the effectiveness of virtual bidding, PJM.* (a) For every year, the average spread (market data) with the virtual bids is closer to zero than the average spread (estimated) without the virtual bids. (b) For every year, the generation cost (estimated) with the virtual bids is smaller than the generation cost (estimated) without the virtual bids. The average saving is about 2%.

inflexible generation for 2014 is considerably higher, possibly due to the extreme weather events triggered by the polar vortex in January 2014. Next, we estimate the generation cost with and without the virtual bids, and the DA and the RT LMPs without the virtual bids. Estimates of the spread are shown in Fig. 8a. It can be seen that virtual bids drive the spread toward zero, thus improving financial efficiency. Estimates of the generation costs are shown in Fig. 8b. It can be seen that the virtual bids reduce the generation cost by an average of 2%, thus improving economic efficiency.

Game theory of virtual bidding

To investigate the strategic behavior of virtual bidders, we formulate a game based on the two-settlement market model, and develop the theory of virtual bidding. Through an equilibrium analysis, we provide a theoretical formula of spread, and discuss the practical implications of the results.

The detailed specification of the game is given in Methods. In this game, each virtual bidder is characterized by its own forecasts about the RT load and the disturbance of the RT cost curve, and its strategy which is the net INC volume traded. The strategy profile determines the DA dispatch and the DA LMP, which then gives the forecast RT LMP of each virtual bidder. The objective of each virtual bidder is to maximize its own forecast profit, given the strategies of the other virtual bidders.

We derive a closed form of the unique Nash equilibrium. The spread realized in equilibrium, as in equation (15), can be interpreted as a measure of the average forecast accuracy of the market and all the virtual bidders. Our further theoretical analysis sheds light on the virtual bidding mechanism. First, virtual bids can reduce the magnitude of the spread, when the average forecast of the virtual bidders is better than that of the market. Second, when the average forecast of the virtual bidders tends to be perfect under the law of large numbers, the spread converges to zero. This is analogous to the Cournot theorem in economic theory, which states that as the number of firms goes to infinity, the price converges to the marginal cost. Third, a virtual bidder makes a positive profit if and only if its participation drives the spread toward zero. Indeed, the mechanism provides the proper incentives: a virtual bidder can make more profits by improving its forecast accuracy, which is also favorable to the market.

Discussion

Based on market data, we formulate a two-settlement market model, which in turn provides a methodology for estimating the system generation cost. Aligning financial efficiency and economic efficiency, the proposed model lays a theoretical framework for understanding the interactions between DA and RT markets. The empirical evidence suggests the effectiveness of virtual bidding. First, the room for arbitrage has been narrowed after the implementation of virtual bidding in CAISO. Second, the estimation shows that the virtual bids in PJM improve the price convergence and save about 2% of the generation cost. On the other hand, there is still room for improving the market efficiency, since standard machine learning algorithms can extract profits beyond those of the actual virtual bids. Furthermore, we develop the game theory of virtual bidding, which conveys the implication that introducing more qualified virtual bidders into the market can improve the price convergence and therefore the market efficiency.

Methods

Support vector machines

For the delivery day t, the input of a virtual bidding strategy can only include information up to hour 12 on day $t-1$. To design practical bidding strategies, we employ support vector machines (SVMs), which have been shown to perform well in a variety of settings.

In the proposed SVM algorithm, the input variables include the spread, the DA load, the RT load, and the forecast load of day $t-7$, $t-6$, $\ldots$, $t-2$. The output variable is the sign of the daily spread of day t. We use the 365 or 366 samples in each year as the training data, and those of the following year as the test data. We choose the radial basis function kernel, with $\gamma = 1/24$ and cost $= 1$.

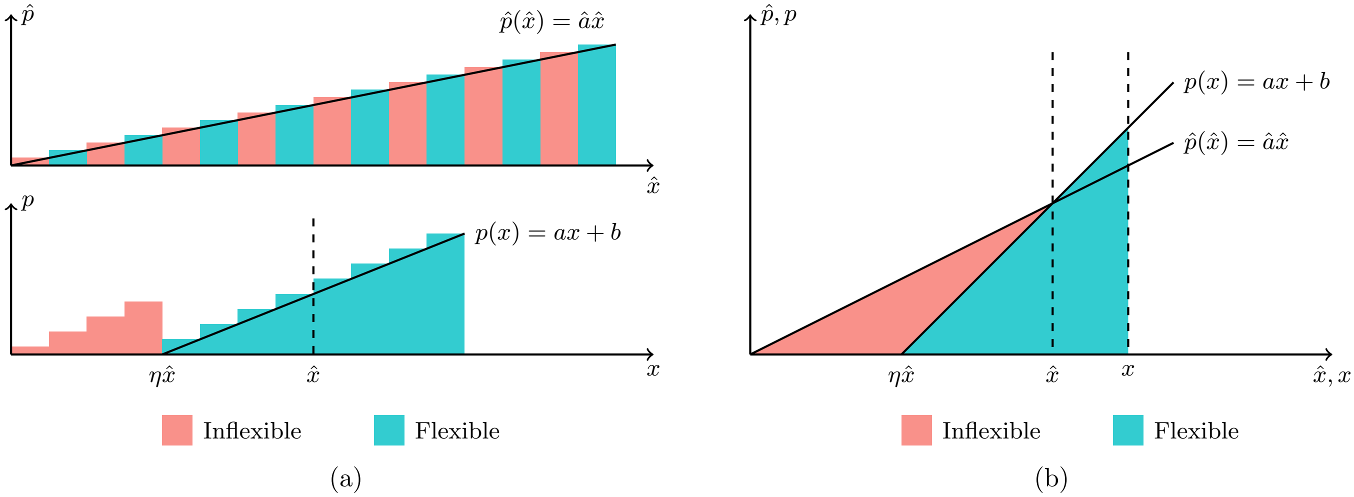


Figure 9: Two-settlement market model. (a) The RT cost curve depends on the DA cost curve and the DA generation. The RT cost curve starts from the inflexible generation $\eta\hat{x}$ with a zero LMP, and is steeper than the DA cost curve since only flexible generation is dispatchable in RT. (b) A compact representation of the DA and RT cost curves. The generation cost is the sum of the inflexible and the flexible generation cost.

Two-settlement market model

For the DA market, the supply consists of the DA generation $\hat{x}$, and the cleared INC volume v_+; the demand consists of the DA load $\hat{y}$, and the cleared DEC volume v_-. Let $v = v_+ - v_-$ be the net INC volume. We have $\hat{x} + v = \hat{y}$. For the RT market, the financial positions are closed out; the RT generation x constitutes the supply, and the RT load y constitutes the demand. We must have $x = y$.

Let the DA cost curve take the form $\hat{p}(\hat{x}) = \hat{a}\hat{x}$, where $\hat{p}$ is the DA LMP, and $\hat{a} > 0$ is a given parameter. Assume that inflexible generation is uniformly distributed, with a given proportion η such that $0 < \eta < 1$. Let the RT cost curve take the form $p(x) = ax + b$, where p is the RT LMP. The RT cost curve depends on the DA cost curve and the DA generation, as illustrated in Fig. 9a. To explore such dependence, consider the following scenarios: when the RT demand turns out to be the inflexible generation $\eta\hat{x}$, the RT LMP should be zero, i.e., $p(\eta\hat{x}) = 0$; when the RT demand turns out to be the DA generation $\hat{x}$, the RT LMP should be the same as the DA LMP, i.e., $p(\hat{x}) = \hat{p}(\hat{x})$. Solving for a and b, we obtain

$$a = \hat{a}/(1 - \eta), \quad b = (\hat{a} - a)\hat{x}. \tag{1}$$

A compact representation of the DA and RT cost curves is shown in Fig. 9b. Summing the inflexible and flexible generation cost, the generation cost can be expressed as a function of the DA generation:

$$c(\hat{x}) = \eta \int_0^{\hat{x}} \hat{a}z\,\mathrm{d}z + \int_{\eta\hat{x}}^{x} (az + b)\,\mathrm{d}z, \tag{2}$$

where the inflexible generation cost is directly proportional to the DA generation cost with proportionality constant η, under the assumption that inflexible generation is uniformly distributed.

In reality, there is uncertainty in the RT cost curve. We capture this phenomenon by adding an independent disturbance δ to b. Moreover, let L be the RT load, so that $x = L$. If L and δ are exactly

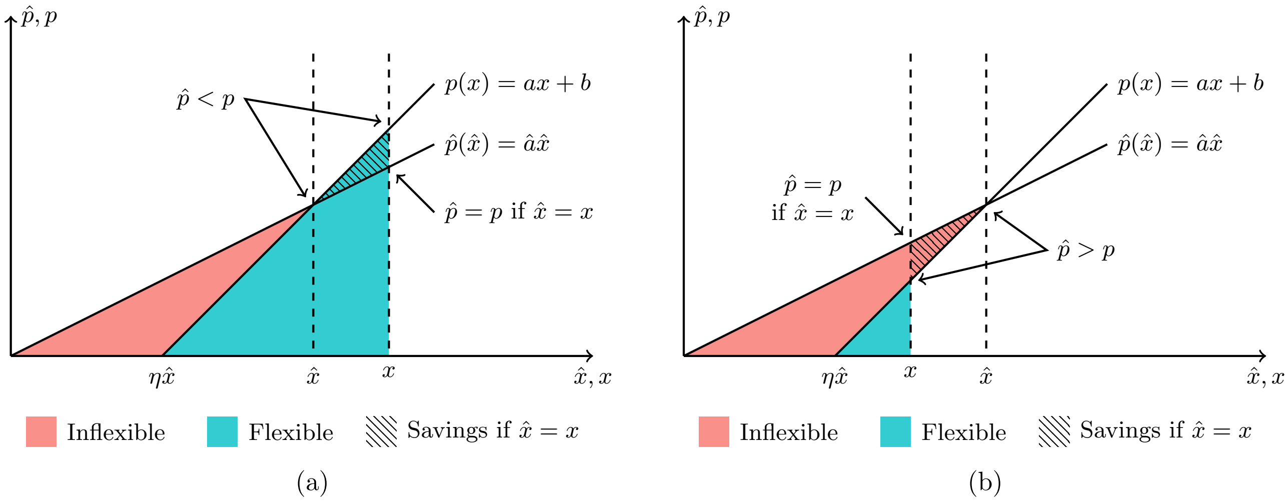


Figure 10: Equivalence between generation cost minimization, price convergence, and dispatch convergence, when $\delta = 0$. (a) When the DA generation $\hat{x}$ is less than the RT generation x, the DA LMP $\hat{p}$ is lower than the RT LMP p. If we had $\hat{x} = x$, the DA LMP $\hat{p}$ would equal the RT LMP p, and the generation cost would be minimized, with the savings highlighted in the figure. (b) When the DA generation $\hat{x}$ is greater than the RT generation x, the DA LMP $\hat{p}$ is higher than the RT LMP p. If we had $\hat{x} = x$, the DA LMP $\hat{p}$ would equal the RT LMP p, and the generation cost would be minimized, with the savings highlighted in the figure.

known, we can solve for the optimal DA generation that minimizes the generation cost (2):

$$\hat{x}^* = L + \delta/a. \tag{3}$$

Generation cost minimization is closely related to price convergence and dispatch convergence.

Proposition 1. *The generation cost is minimized if and only if the spread is zero. When the disturbance of the RT cost curve is zero, the spread is zero if and only if the DA generation equals the RT generation.*

Therefore, financial efficiency and economic efficiency are aligned in our model. This main result is illustrated in Fig. 10, where we establish the equivalence between generation cost minimization, price convergence, and dispatch convergence, when the disturbance of the RT cost curve is zero.

Impact of virtual bids on market outcomes

In reality, the market makes imperfect forecasts. Let L_0 and δ_0 be the market's forecasts about L and δ, respectively. Minimizing the market's forecast generation cost, the DA generation is given by

$$\hat{x}_0 = L_0 + \delta_0/a. \tag{4}$$

The resulting spread can be viewed as a measure of the market's forecast accuracy:

$$s_0 = a(L_0 - L) + (\delta_0 - \delta). \tag{5}$$

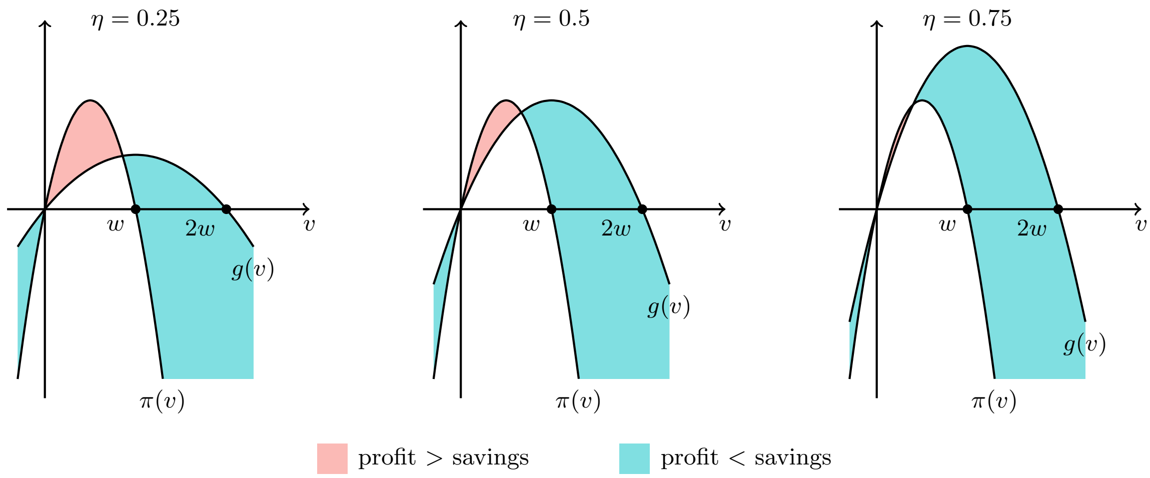


Figure 11: Comparison between the profit of virtual bids and the resulting savings of generation cost. Virtual bids are more effective as the proportion of inflexible generation increases. The more competitive the virtual bids, the more likely that the resulting savings outweigh the payments to them.

With the net INC volume v, the DA generation is $\hat{x}(v) = \hat{x}_0 - v$, and the spread can be calculated as

$$s(v) = -av + aw, \tag{6}$$

where $w = \hat{x}_0 - \hat{x}^*$ is defined as the net INC volume leading to a zero spread. The profit of the virtual bids is then given by

$$\pi(v) = -av^2 + awv, \tag{7}$$

and the savings of the generation cost due to the virtual bids is given by

$$g(v) = -\eta a v^2/2 + \eta awv. \tag{8}$$

The comparison between the profit and the savings is illustrated in Fig. 11.

Empirical estimation

We fit the cost curves by linear regression for each year, and estimate the yearly proportion of inflexible generation as $\eta = 1 - (\hat{a}/a)$. Then we make hourly estimates, taking η as fixed throughout a year:

1. The slope of the DA cost curve: $\hat{a} = \hat{p}/(\hat{y} - v)$;
2. The slope of the RT cost curve: $a = \hat{a}/(1 - \eta)$;
3. The intercept of the RT cost curve: $b = p - ay$;
4. The disturbance of the RT cost curve: $\delta = b - (\hat{a} - a)(\hat{y} - v)$;
5. The generation cost with the virtual bids:

$$c = \eta \int_0^{\hat{y}-v} \hat{a} z \mathrm{d}z + \int_{\eta(\hat{y}-v)}^{y} (az + b)\mathrm{d}z; \tag{9}$$

6. The DA LMP without the virtual bids: $\hat{p}_0 = \hat{a}\hat{y}$;

7. The RT LMP without the virtual bids: $p_0 = ay + (\hat{a} - a)\hat{y} + \delta$;

8. The generation cost without the virtual bids:

$$c_0 = \eta \int_0^{\hat{y}} \hat{a}z\mathrm{d}z + \int_{\eta\hat{y}}^{y} (az + (\hat{a} - a)\hat{y} + \delta)\mathrm{d}z. \tag{10}$$

Game-theoretic analysis

Consider N virtual bidders, indexed by $i = 1, \ldots, N$. Let L_i and δ_i be virtual bidder i's forecasts about L and δ, respectively. Let v_i be the net INC traded by virtual bidder i. Let $v = (v_1, \ldots, v_N)$ and $v_{-i} = (v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_N)$. Given the strategy profile v, the DA LMP is given by

$$\hat{p}(v) = \hat{a}(\hat{x}_0 - v). \tag{11}$$

Virtual bidder i's forecast RT LMP is given by

$$p_i(v) = aL_i + (\hat{a} - a)(\hat{x}_0 - v) + \delta_i. \tag{12}$$

Virtual bidder i's forecast profit, or utility, is then calculated as

$$\pi_i(v_i, v_{-i}) = -av_i^2 + \left(-a\sum_{j\neq i} v_j + a(L_0 - L_i) + (\delta_0 - \delta_i)\right) v_i. \tag{13}$$

In a Nash equilibrium $v^* = (v_1^*, \ldots, v_N^*)$, each v_i^* maximizes virtual bidder i's utility, given the others' strategies. It can be shown that the Nash equilibrium is unique, given by

$$v_i^* = \frac{aL_0 + \delta_0 + \sum_{j\neq i}(aL_j + \delta_j) - N(aL_i + \delta_i)}{(N+1)a}. \tag{14}$$

Furthermore, the realized spread in equilibrium can be calculated as

$$s^* = \frac{s_0 + \sum_i (a(L_i - L) + (\delta_i - \delta))}{N+1}, \tag{15}$$

where s_0 is the spread without the virtual bids, as in equation (5).

The following results convey the practical implications of the virtual bidding mechanism, as discussed in the main text.

Proposition 2. *If*

$$\frac{1}{N}\left|\sum_i (a(L_i - L) + (\delta_i - \delta))\right| < s_0 = a(L_0 - L) + (\delta_0 - \delta),$$

then

$$|s^*| < |s_0|.$$

Proposition 3 (Cournot theorem for virtual bidding)**.** *If*

$$\frac{1}{N}\left|\sum_i (a(L_i - L) + (\delta_i - \delta))\right| \to 0 \text{ as } N \to \infty,$$

then

$$|s^*| \to 0 \text{ as } N \to \infty.$$

Proposition 4. *Let s^*_{-i} be the realized spread in equilibrium without the participation of virtual bidder i:*

$$s^*_{-i} = \frac{s_0 + \sum_{j \neq i}(a(L_j - L) + (\delta_j - \delta))}{N}.$$

Then $s^* v^*_i > 0$ if and only if

$$\begin{cases} 0 < s^* < s^*_{-i}, & s^*_{-i} > 0, \\ s^*_{-i} < s^* < 0, & s^*_{-i} < 0. \end{cases} \tag{16}$$

References

[1] Celeste Saravia. Speculative trading and market performance: the effect of arbitrageurs on efficiency and market power in the New York electricity market. *Center for the Study of Energy Markets*, 2003.

[2] Francis A Longstaff and Ashley W Wang. Electricity forward prices: a high-frequency empirical analysis. *The Journal of Finance*, 59(4):1877–1900, 2004.

[3] Lester Hadsell and Hany A Shawky. Electricity price volatility and the marginal cost of congestion: an empirical study of peak hours on the NYISO market, 2001-2004. *The Energy Journal*, pages 157–179, 2006.

[4] Lester Hadsell and Hany A Shawky. One-day forward premiums and the impact of virtual bidding on the New York wholesale electricity market using hourly data. *Journal of Futures Markets*, 27(11):1107–1125, 2007.

[5] Lester Hadsell. Day-ahead premiums on the New England ISO. *The Electricity Journal*, 21(4):51–57, 2008.

[6] Severin Borenstein, James Bushnell, Christopher R Knittel, and Catherine Wolfram. Inefficiencies and market power in financial arbitrage: a study of California's electricity markets. *The Journal of Industrial Economics*, 56(2):347–378, 2008.

[7] Stratford Douglas and Julia Popova. Storage and the electricity forward premium. *Energy Economics*, 30(4):1712–1727, 2008.

[8] Craig Pirrong and Martin Jermakyan. The price of power: the valuation of power and weather derivatives. *Journal of Banking & Finance*, 32(12):2520–2529, 2008.

[9] Nicholas Bowden, Su Hu, and James Payne. Day-ahead premiums on the Midwest ISO. *The Electricity Journal*, 22(2):64–73, 2009.

[10] Lester Hadsell. Inefficiency in deregulated wholesale electricity markets: the case of the New England ISO. *Applied Economics*, 43(5):515–525, 2011.

[11] Erik Haugom and Carl J Ullrich. Market efficiency and risk premia in short-term forward prices. *Energy Economics*, 34(6):1931–1941, 2012.

[12] CK Woo, J Zarnikau, E Cutter, ST Ho, and HY Leung. Virtual bidding, wind generation and California's day-ahead electricity forward premium. *The Electricity Journal*, 28(1):29–48, 2015.

[13] Akshaya Jha and Frank A Wolak. Testing for market efficiency with transaction costs: an application to financial trading in wholesale electricity markets. *Working paper*, 2015.

[14] Ruoyang Li, Alva J Svoboda, and Shmuel S Oren. Efficiency impact of convergence bidding in the California electricity market. *Journal of Regulatory Economics*, 48(3):245–284, 2015.

[15] Hendrik Bessembinder and Michael L Lemmon. Equilibrium pricing and optimal hedging in electricity forward markets. *The Journal of Finance*, 57(3):1347–1382, 2002.

[16] SX Zhang, CY Chung, KP Wong, and H Chen. Analyzing two-settlement electricity market equilibrium by coevolutionary computation approach. *Power Systems, IEEE Transactions on*, 24(3):1155–1164, 2009.

[17] Daron Acemoglu, Ali Kakhbod, and Asuman Ozdaglar. Competition in electricity markets with renewable energy sources. *Working paper*, 2015.

Author contributions

All authors have contributed equally to all aspects of the work.

Additional information

Correspondence and requests for materials should be addressed to W.T.

Competing interests

The authors declare no competing financial interests.