Efficiency Impact of Convergence Bidding in the California Electricity Market

Ruoyang Li*, Alva Svoboda**, Shmuel Oren*
*University of California at Berkeley
**Pacific Gas and Electric Co.

Presented at
Conference on the Economics of Energy and Climate Change
Toulouse, France
September 8-9, 2015

Scope

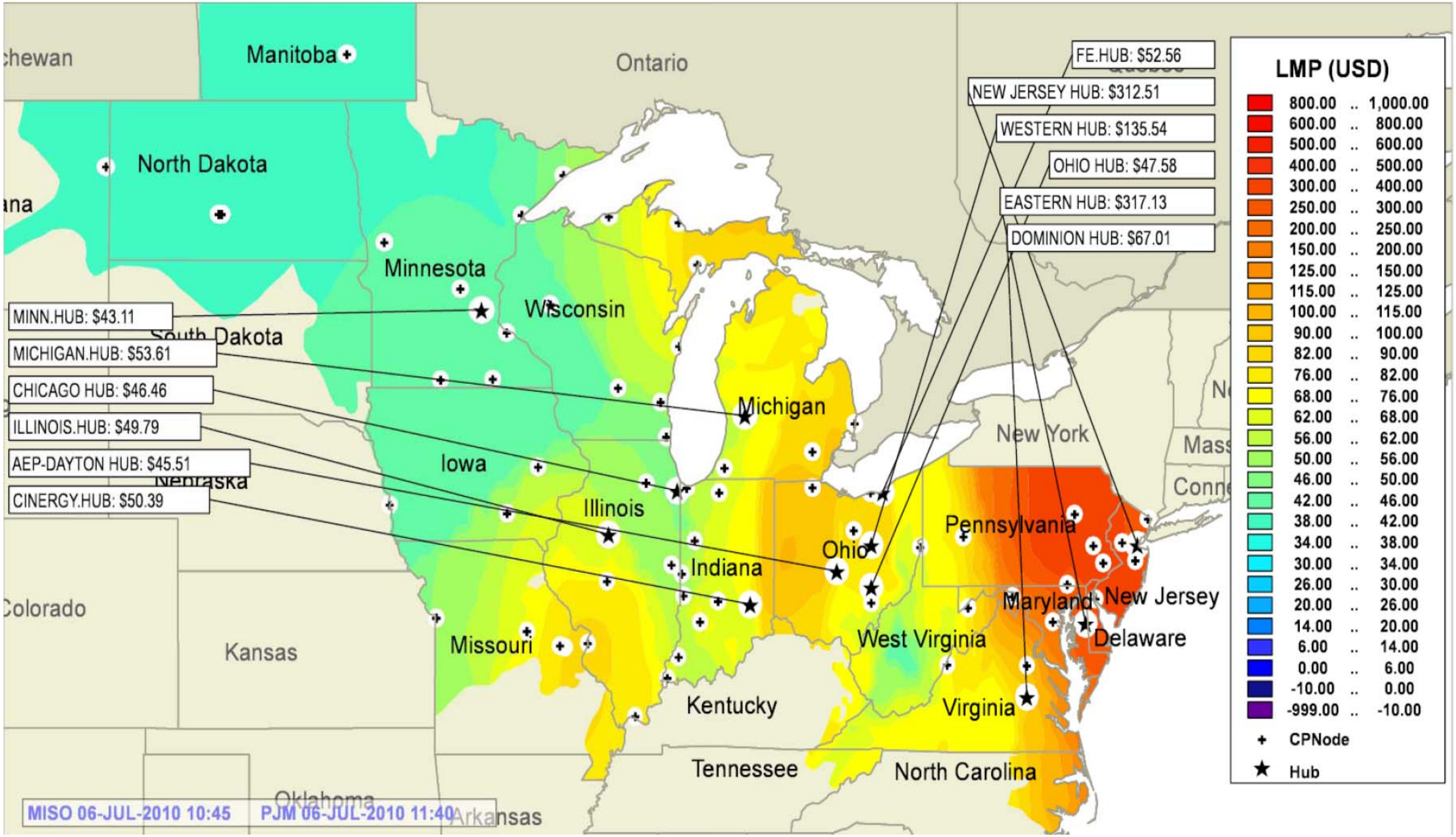
- The California Independent System Operator (CAISO) has implemented CB on February 1, 2011 under Federal Energy Regulatory Commission's (FERC) September 21, 2006 Market Redesign and Technology Upgrade (MRTU) Order.
- CB is a pure financial mechanism that allows market participants to arbitrage price differences between forward and spot electricity markets without physically obligation.
- The central question of this study is to address whether the CAISO's forward and spot electricity markets are efficient, and if not, to what extent CB improves market efficiency.

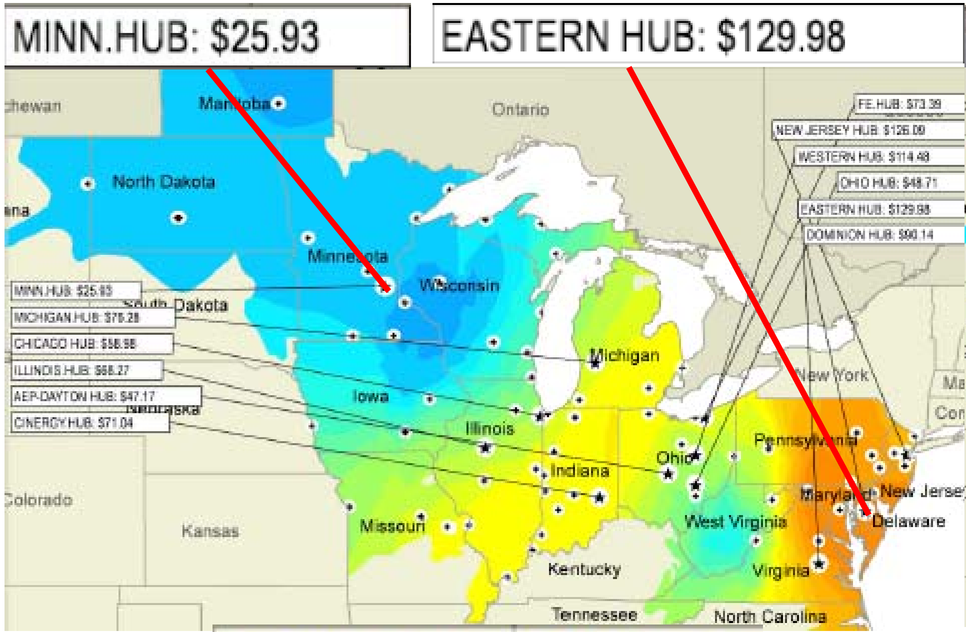
Pricing Mechanisms

- The prevailing mechanism for pricing electric energy in US electricity markets operated by Independent System Operators (ISO) is Locational Marginal Prices (LMP), defined as the incremental (least) cost of supplying a marginal MW of power to the specific location while meeting all security constraints
- LMP is the pricing mechanism proposed by Bohn, Caramanis, and Schweppe (1984) to internalize significant externalities arising from the presence of Kirchhoff's laws that governs the power flow in transmission systems.
- When the transmission lines are congested and the import of electricity from cheap producers are constrained, the ISO is forced to use some local but expensive producers for power generation in order to satisfy the demand.
- As a result, LMPs are high in the downstream areas of the congested transmission lines, and low in the upstream areas.

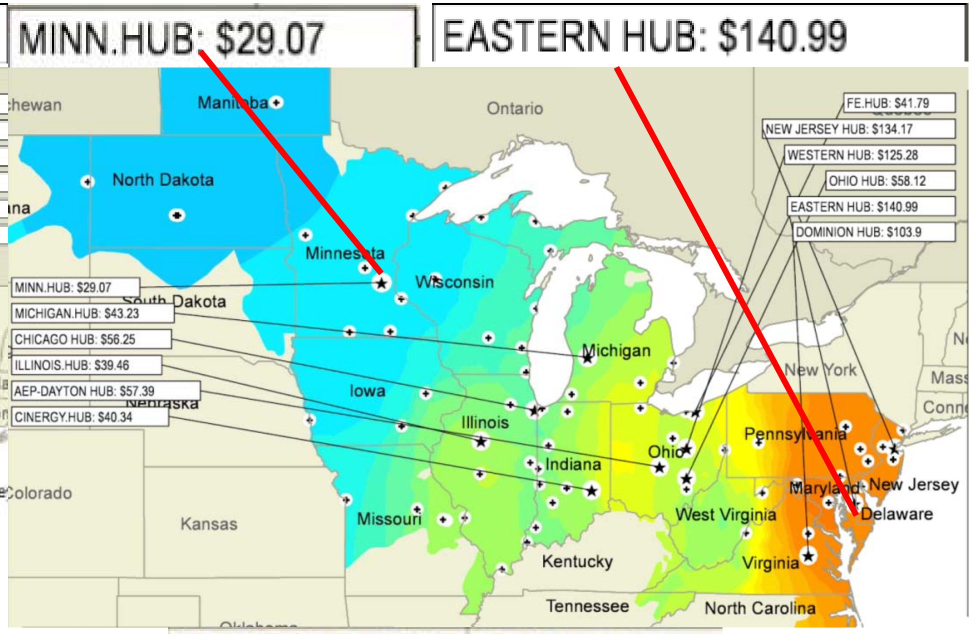
MidwestISO
Energizing the Heartland

pjm

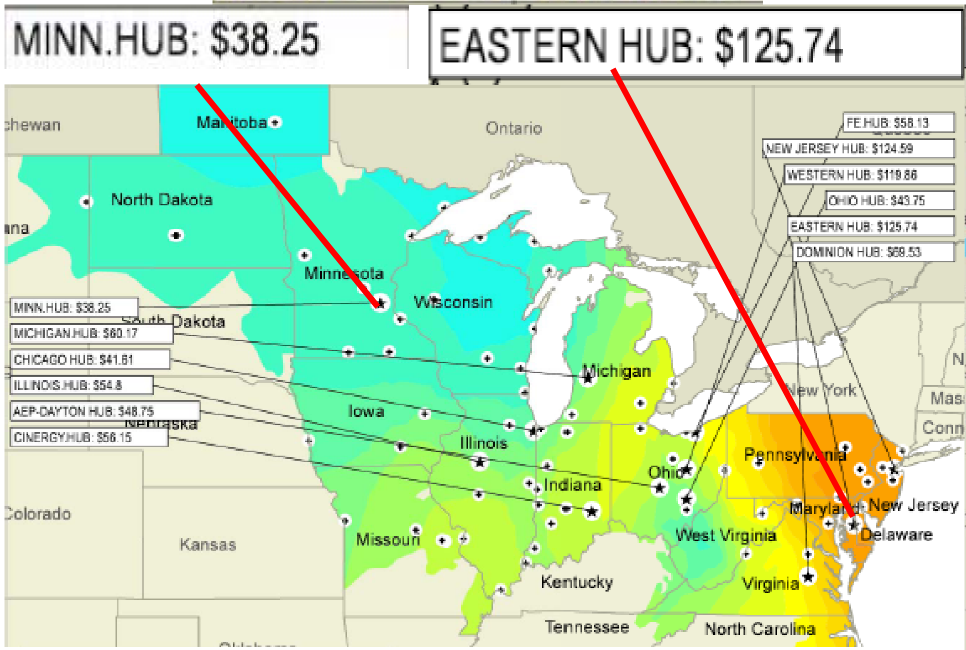




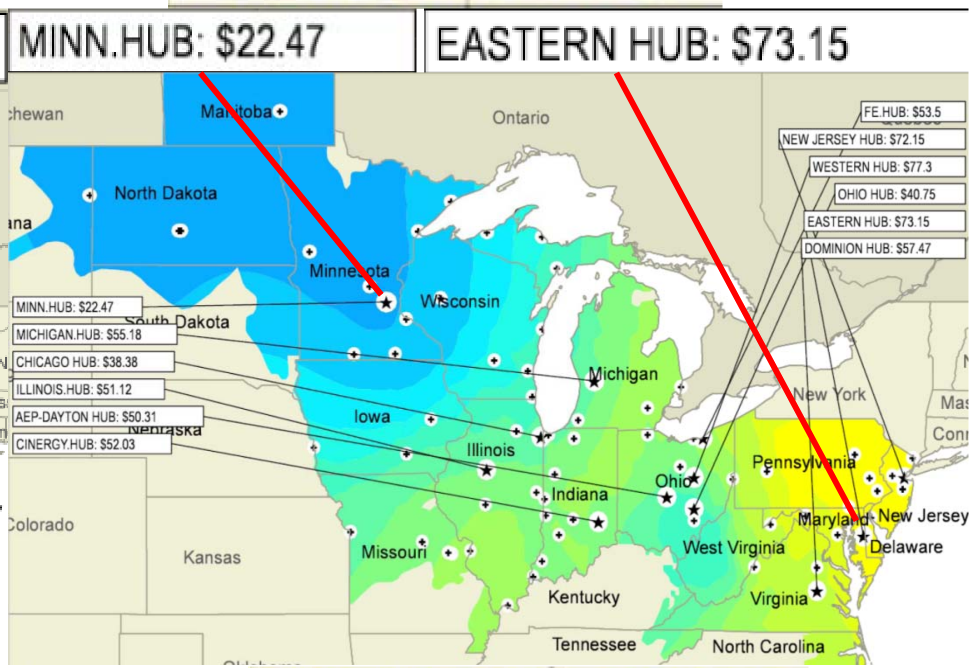
MISO 05-JUL-2010 17:35



MISO 05-JUL-2010 17:50

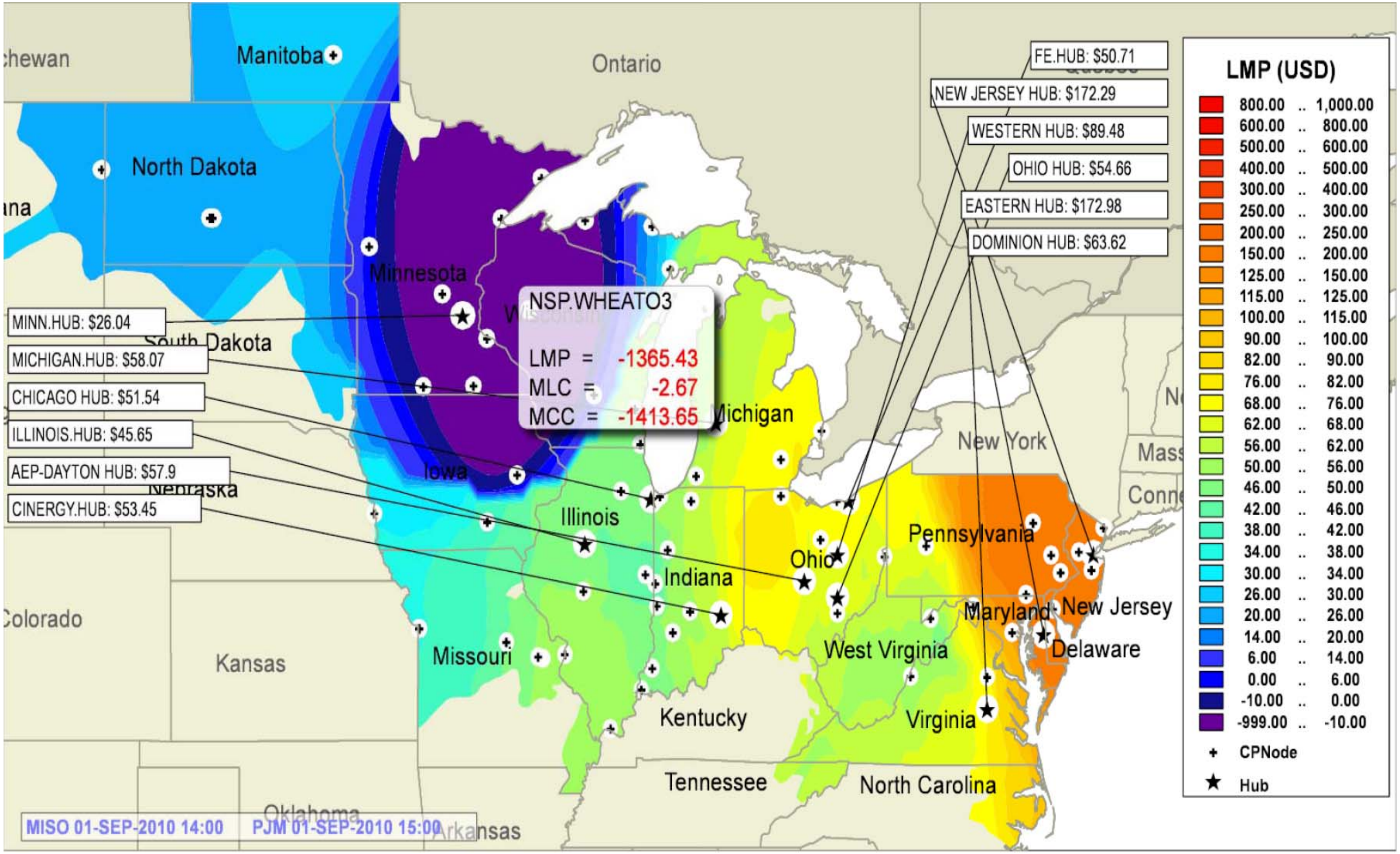


MISO 05-JUL-2010 17:45



MISO 05-JUL-2010 18:55

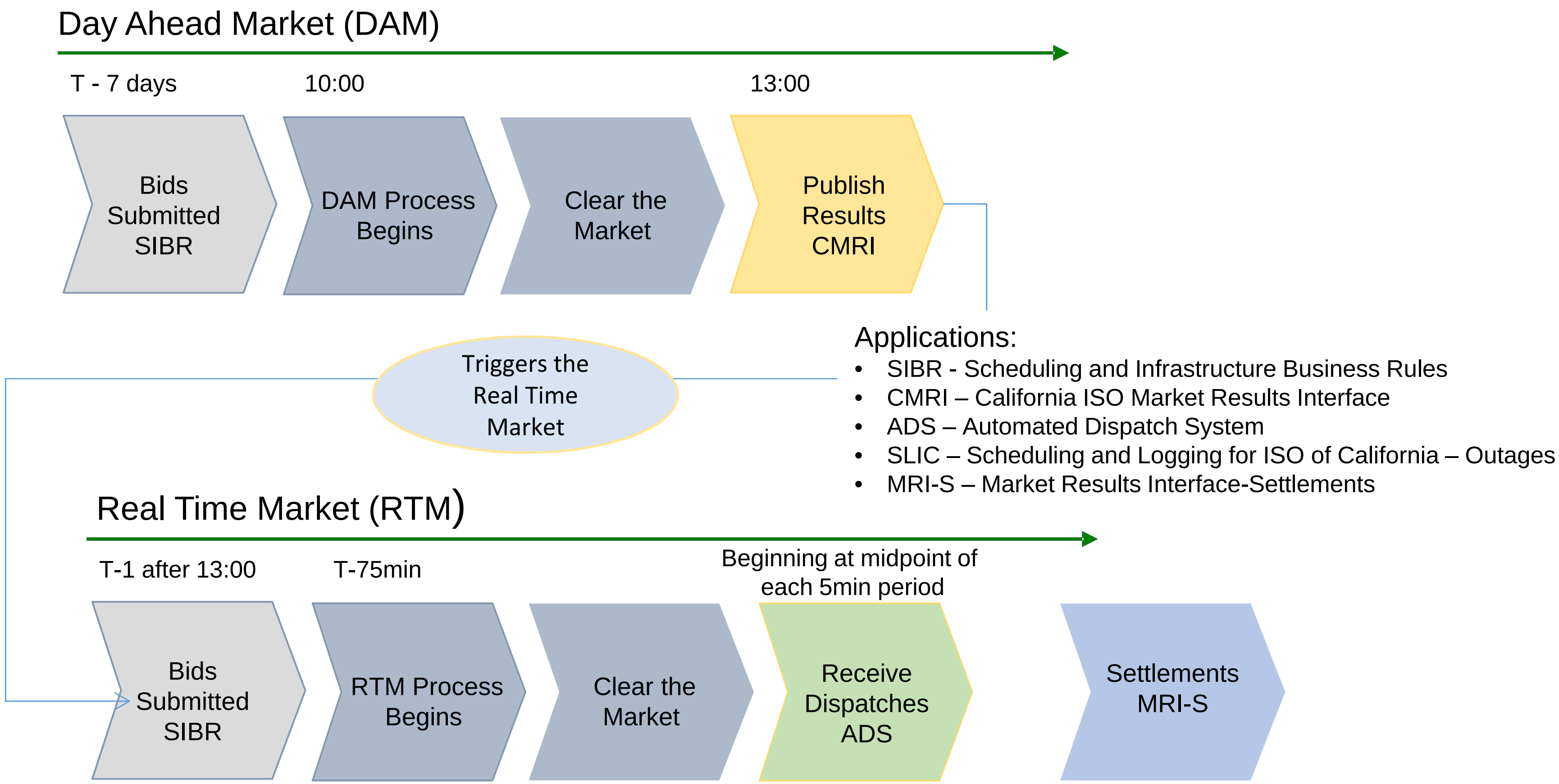
Over-generation, congestion and no storage capability can lead to negative prices



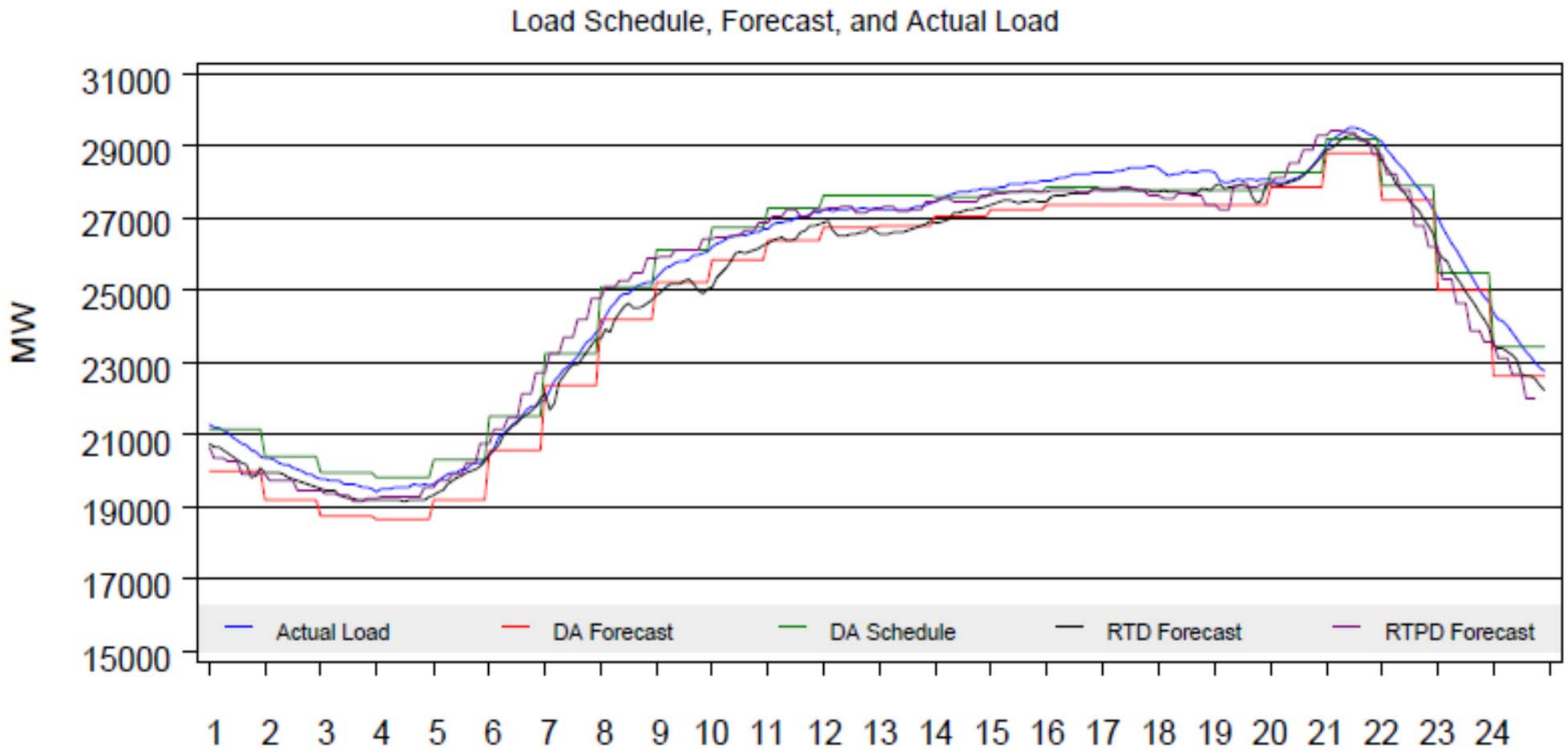
Two-settlements Electricity Markets

- The two-settlement electricity markets consist of two interrelated markets: day-ahead (DA) market, and real-time (RT) market.
- DA LMPs are generally considered more stable than RT LMPs.
- The DA market includes three sequential processes: market power mitigation and reliability requirement determination (MPM-RRD), integrated forward market (IFM), and residual unit commitment (RUC).
- In the RT market, the ISO runs the economic dispatch process every 5 minutes to rebalance the residual demand.
- If a resource does not cover its total cost including start-up and minimum load cost through its energy revenue at DA and RT LMPs, its shortfall is covered by an uplift payment which is allocated to market participants.

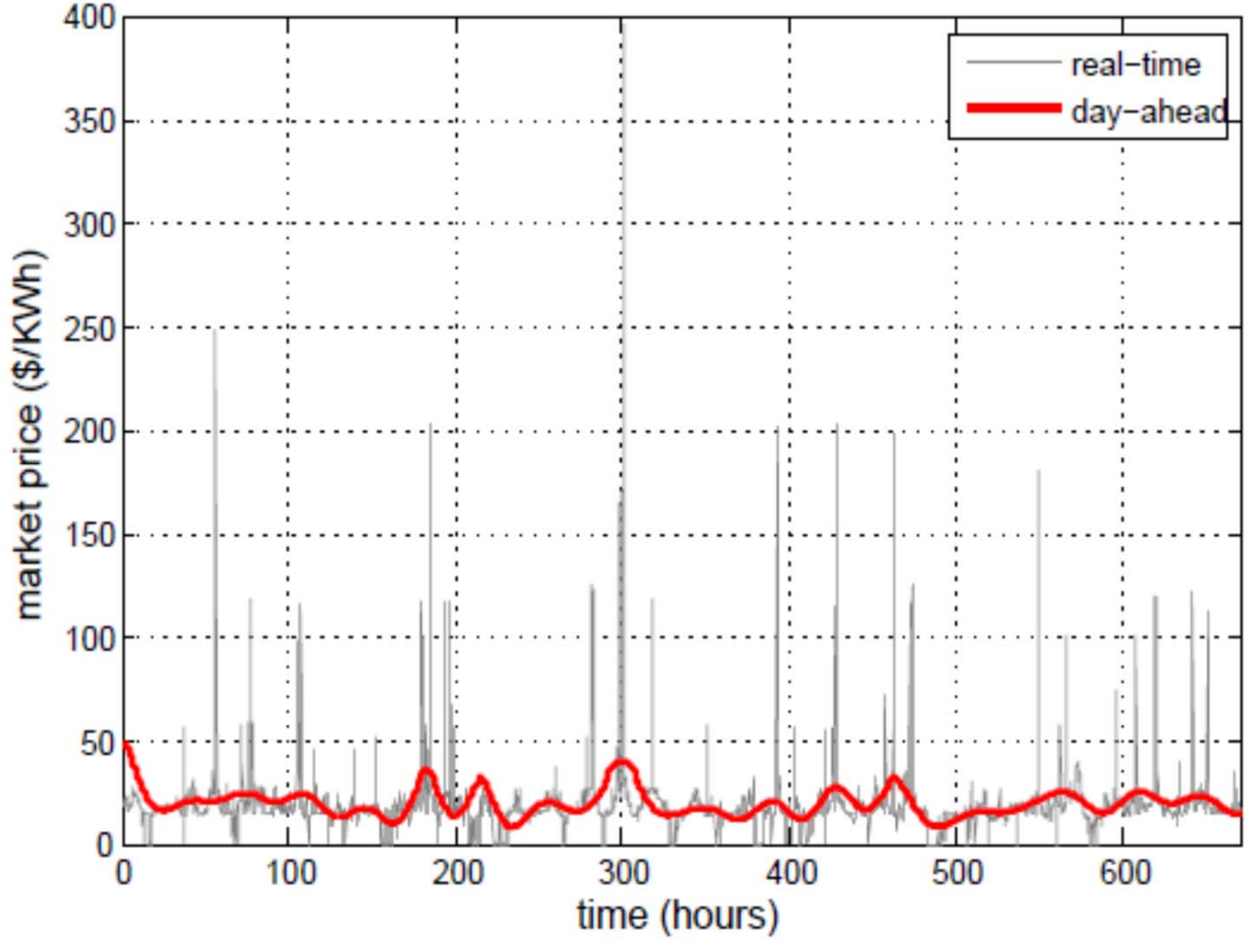
California ISO Market Timeline



Typical Daily Schedule



Example of DA-RT Price Spread



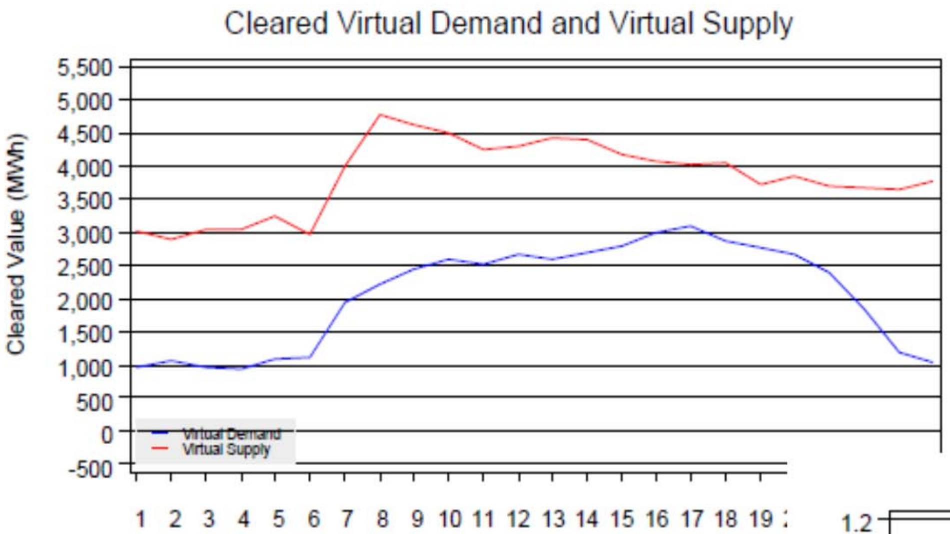
Convergence (Virtual) Bidding (CB)

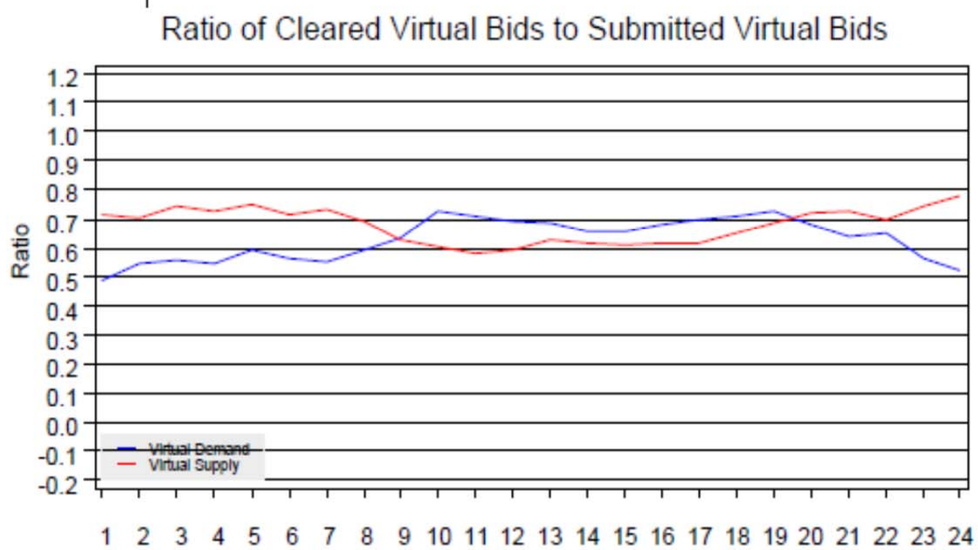
- The California Independent System Operator (CAISO) has implemented CB on February 1, 2011 under Federal Energy Regulatory Commission's (FERC) September 21, 2006 Market Redesign and Technology Upgrade (MRTU) Order.
- CB is a pure financial mechanism that allows market participants to arbitrage price differences between forward and spot electricity markets without physically obligation.
- CB also enables market participants executing physical trades to opt for RT prices instead of DA prices. It Also increases market liquidity by enabling participants with no assets to take positions arbitraging the DA-RT spread

Theoretical Benefits of CB

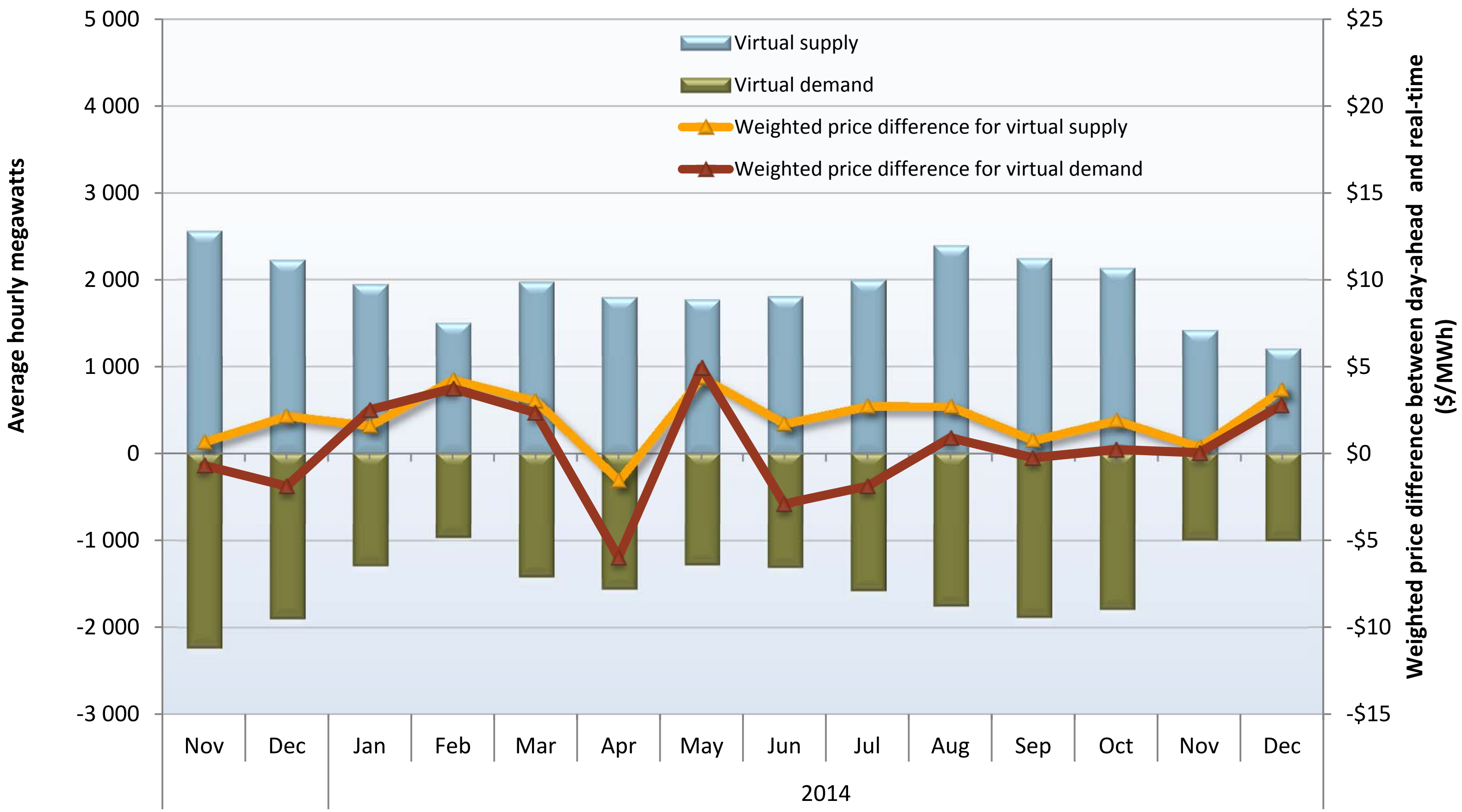
- Price convergence is regarded as a benefit to the DA and RT markets.
- It reduces the incentives for market participants to defer their physical resources to the RT market in expectation of favorable RT LMPs.
- The benefit of CB also comes from the fact that it relieves market participants from using physical resources to arbitrage price differences between the DA and RT markets, also called implicit virtual bidding in some literature.
- The central question of this study is to address whether the CAISO's forward and spot electricity markets are efficient, and if not, to what extent CB improves market efficiency.

Typical Submitted and Cleared CB Volumes





Convergence bidding volumes and weighted price differences Q4 2014



The Efficient Market Hypothesis

- The efficient market hypothesis first formalized by Samuelson (1965) and Fama (1970), asserts that at any given time asset prices should always reflect all available information, and change quickly to incorporate new information.
- Jensen (1978) defines market efficiency in terms of trading profitability – "a market is efficient with respect to [an] information set, if it is impossible to make economic profits by trading on the basis of [this] information set."
- Following this methodology, we empirically test for market efficiency by evaluating the performance of trading strategies based on market data in the CAISO electric power markets.

Our Approach

- Use the perspective of a trader attempting to construct an optimal portfolio of virtual hourly positions.
  - Construct and estimate a statistical time series model of the returns from historical data
  - Obtain an optimal portfolio (based on appropriate constraints on risk)
  - Evaluate the arbitrage profits of the optimal portfolio using in sample and out of sample data
  - Asses market efficiency based on profitable arbitrage opportunity.
- Alternative approach proposed by Jha and Wolak [2013] was based on evaluating the "implied trading cost" for which one cannot reject the null hyphotesis of "no arbitrage" (based on average return = uniform portfolio assumption)

Portfolio Optimization

- Let $P_t^{DA} \in \mathbb{R}^{24}$ and $P_t^{RT} \in \mathbb{R}^{24}$ denote DA and RT LMPs for one node on day t.
- DA-RT spreads can be expressed as $R_t = P_t^{DA} - P_t^{RT}$.
- The energy trader's optimization problem can be formulated as,

$$\text{(P0)} \quad \max_{x_t} \quad E\left[R_t^T x_t\right] - \tau \|x_t\|_1 \tag{1}$$

$$s.t. \quad C\|x_t\|_1 \leq W_0 \tag{2}$$

  where τ is the costs allocated to 1 MWh of virtual position, C is the reference price for 1 MWh of virtual position, and W_0 is the initial collateral.
- $x_t^{(i)} \geq 0$ denotes virtual supply, while $x_t^{(i)} < 0$ denotes virtual demand.

Portfolio Optimization (cont'd)

- Without loss of generality, we assume $W_0 = 1$.
- The collateral used to establish virtual positions in DA-RT spreads is $y_t = Cx_t$ and the costs associated with 1 dollar of collateral are $\tau^c = \frac{1}{C}\tau$.
- The returns on DA-RT spreads are then defined as $R_t^c = \frac{1}{C}R_t = \frac{1}{C}(P_t^{DA} - P_t^{RT})$.
- With these substitutions, (P0) is equivalent to (P1),

$$
\begin{aligned}
\text{(P1)} \quad \max_{y_t} \quad & E[R_t^{c\,T} y_t] - \tau^c \|y_t\|_1 && (3)\\
\text{s.t.} \quad & \|y_t\|_1 \leq 1, && (4)
\end{aligned}
$$

which is a portfolio optimization problem in the presence of linear transaction costs.

Portfolio Optimization with VaR Constraints

- VaR is a modern way of measuring the risk of a portfolio, based on computing probabilities of large losses of the portfolio (Rockafellar and Uryasev, 2000).
- Mathematically, $\mathrm{VaR}(z;\eta) = \inf\{\gamma | \mathrm{P}(z \le \gamma) \ge \eta\}$ is the level η-quantile of the random variable z. (P1) can be reformulated as a portfolio optimization problem $(\mathrm{VAR0}(\gamma,\eta))$ with VaR constraint (6),

$$
\begin{array}{llll}
(\mathrm{VAR0}(\gamma,\eta)) & \max_{y_t} & E[{R_t^c}^T y_t] - \tau^c \|y_t\|_1 & (5) \\
 & \text{s.t.} & \mathrm{VaR}(-{R_t^c}^T y_t;\eta) \le \gamma & (6) \\
 & & \|y_t\|_1 \le 1 & (7)
\end{array}
$$

where γ is the predetermined upper bound for the VaR of the portfolio.

Portfolio Optimization with VaR Constraints (cont'd)

- DA-RT spreads are negatively skewed in most of the hours, which cannot be modeled properly by a normal distribution.
- Without assuming normality, VaR cannot be written in a closed form, and there is no guarantee that VaR is convex.
- Nemirovski and Shapiro (2006) propose a computationally tractable approximation of the non-convex VaR constraint (6) with the Chebyshev bound,

$$
\begin{aligned}
&(\mathrm{VAR2}(\gamma,\eta)) \\
&\max_{y_t} \quad \mu_t^T y_t - \tau^c \|y_t\|_1 && (8)\\
&\text{s.t.} \quad -E[(R_t^{c\,T} y_t + \gamma)] + (\eta E[(R_t^{c\,T} y_t + \gamma)^2])^{\frac{1}{2}} \le 0 && (9)\\
&\qquad \|y\|_1 \le 1. && (10)
\end{aligned}
$$

Portfolio Optimization with CVaR Constraints

- Since VaR is incapable of addressing the distribution of losses beyond $\mathrm{VaR}(z;\eta)$, CVaR is introduced by Rockafellar and Uryasev (2000) as an alternative risk assessment technique to account for losses in the tail of the distribution.
- In this case, the optimization problem can be stated as follows,

$$
\begin{aligned}
(\mathrm{CVAR0}(\gamma,\eta)) \quad \max_{y_t} \quad & E[{R_t^c}^T y_t] - \tau^c \|y_t\|_1 && (11)\\
\text{s.t.} \quad & \mathrm{CVaR}(-{R_t^c}^T y_t;\eta) \le \gamma && (12)\\
& \|y_t\|_1 \le 1. && (13)
\end{aligned}
$$

Portfolio Optimization with CVaR Constraints (cont'd)

- VaR and CVaR can be characterized by function $g_\eta(z,\rho) = \rho + \frac{1}{1-\eta} E[(z-\rho)_+]$,

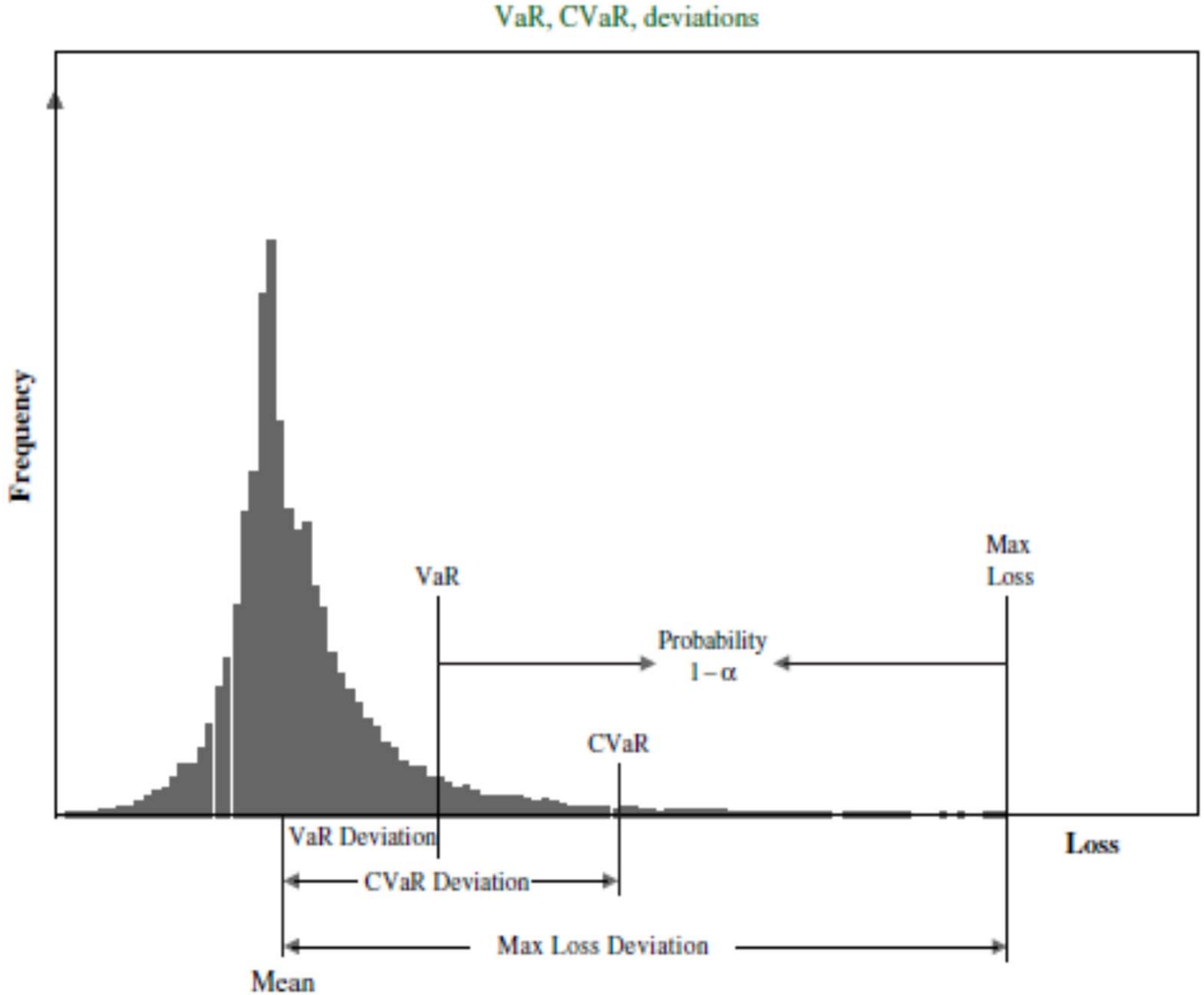
$$\mathrm{CVaR}(z,\eta) = \min_{\rho} g_\eta(z,\rho), \tag{14}$$

$$\mathrm{VaR}(z,\eta) = \arg\min_{\rho} g_\eta(z,\rho). \tag{15}$$

- Thus, by substituting $\mathrm{CVaR}(z,\eta)$, $(\mathrm{CVAR0}(\gamma,\eta))$ becomes,

$$
\begin{array}{lll}
(\mathrm{CVAR1}(\gamma,\eta)) & \max_{y_t} & E({R_t^c}^T y_t) - \tau^c |y_t|_1 \qquad (16) \\
 & s.t. & g_\eta(-{R_t^c}^T y_t, \rho) \le \gamma \qquad (17) \\
 & & |y_t|_1 \le 1. \qquad (18)
\end{array}
$$

Risk Measures: VaR vs. CVaR



Time Varying Forward Premium

- In electricity markets, the 24 hourly forward premia FP_t on day t take the form,

$$FP_t = E_{t-1}[P_t^{DA} - P_t^{RT}] = E_{t-1}[R_t]. \tag{19}$$

- There exists extensive literature on the time-varying property of the forward premium – a situation where the forward premium varies through time to reflect economic risk.
- The time-varying forward premium is observed and well documented in exchange rates and traditional commodity markets, including Fama (1984), Fama and French (1987), Bekaert and Hodrick (1993), Backus, Foresi, and Telmer (2001), and Baillie and Kilic (2006).
- Recently, there is a growing literature investigating the time-varying forward premium in electricity markets.

Time Varying Forward Premium (cont'd)

- Bessembinder and Lemmon (2002) model the forward market as a closed system, where the only participants are producers and consumers.
- The forward premium can be expressed as,

$$P_t^{DA} - E[P_t^{RT}] = \theta_1 Var[P_t^{RT}] - \theta_2 Skew[P_t^{RT}], \tag{20}$$

  where $\theta_1 \leq 0$, and $\theta_2 \leq 0$, implying that the forward premium is negatively related to the variance of RT LMPs, and positively related to the skewness of RT LMPs.
- To express forward premia in terms of DA-RT Spreads $R_t = P_t^{DA} - P_t^{RT}$, we can rewrite (20) as,

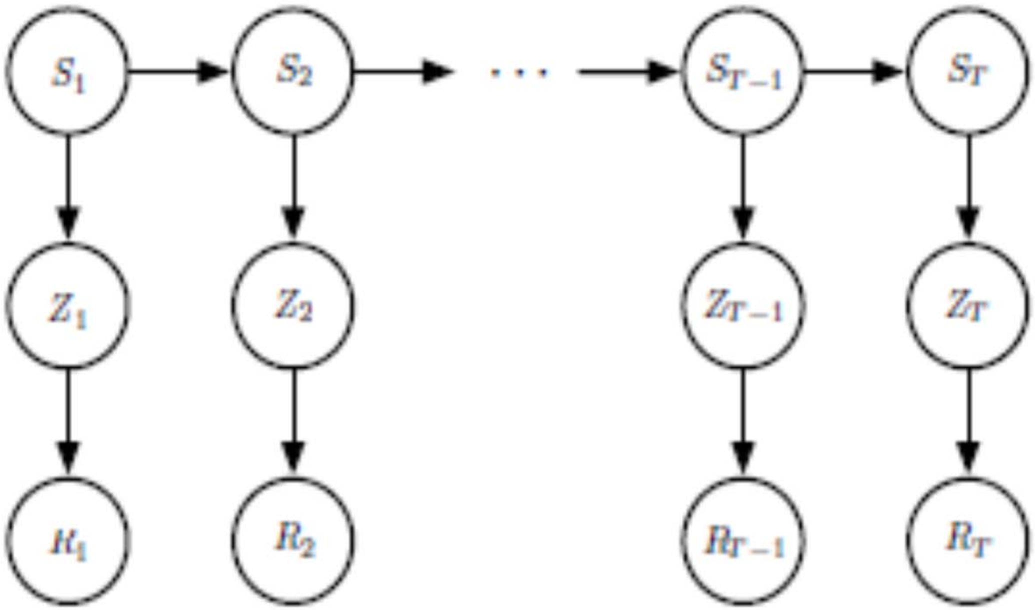
$$E[R_t] = \theta_1 Var[R_t] + \theta_2 Skew[R_t]. \tag{21}$$

Time Varying Forward Premium (cont'd)

- The existing literature extensively studies the time-varying forward premium in electricity markets by statistical models with observable state variables, namely the volatility and skewness of spot prices, the level of risk aversion, market structure, and demand and supply capacity in Shawky, Marathe, and Barrett (2003), Cartea and Villaplana (2008), and Benth, Cartea, and Kiesel (2008).
- The choice of state variables is largely predetermined and varies across different electricity markets, which limits the possibility to arrive at a generalization.

Model Description

- In the CB context, Hidden Markov Model (HMM) is a discrete-time stochastic process $\{S_t, R_t\}_{t=1}^{T}$ where the sequence of states $\{S_t\}_{t=1}^{T}$ is an unobserved Markov chain.
- In this study, we assume the conditional probability density function of R_t, given the occurrence of S_t, follows a Gaussian mixture distribution.

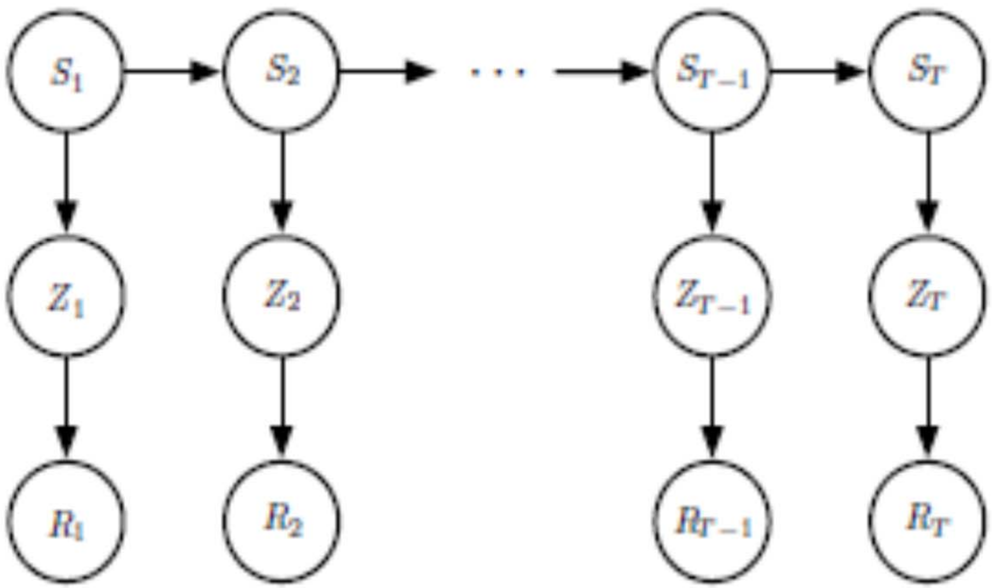


Model Description (cont'd)

- We assume $P(z_t|s_t) = c_{s_t,z_t}$, $P(s_{t+1}|s_t) = a_{s_t,s_{t+1}}$.
- The equation for the returns on DA-RT spreads R_t in cluster z_t is,

$$R_t = \mu_{z_t} + \Sigma_{z_t}^{\frac{1}{2}} \epsilon_t, \tag{22}$$

where μ_{z_t} denotes the conditional mean given the cluster z_t, Σ_{z_t} denotes the conditional covariance given the cluster z_t, and ϵ_t denotes the noise.



Data

- The data for this study consist of the historical DA and RT LMPs at the CAISO NP15 EZ Gen Hub before and after the implementation of CB.
- The costs τ are assumed to be \$0.085 for 1 MWh of cleared virtual position and the reference price C for 1 MWh of cleared virtual positionand is calculate by the 95th percentile value of the historical price differences between DA and RT LMPs.

Table : Seasonal Means of Post-CB DA-RT Spreads

Hour	November - January	February - April	May - July	August - October
1	1.84	1.18	2.25	1.95
2	1.98	2.71	1.63	1.61
3	1.47	4.25	2.02	1.68
4	2.93	3.46	3.19	2.78
5	1.51	1.44	1.79	0.55
6	-0.74	-1.73	1.85	0.18
7	-0.16	0.34	5.82	-0.29
8	2.32	-3.53	-3.70	2.58
9	-3.99	-0.60	2.44	-0.14
10	-1.44	-10.37	2.03	0.27
11	0.15	-11.31	-0.72	0.90
12	-0.70	-5.47	1.85	4.99
13	-2.25	-1.93	0.92	3.70
14	2.76	-0.13	-2.22	8.27
15	1.09	-2.11	-1.71	1.43
16	0.58	1.15	-8.49	-3.04
17	2.68	-0.98	-14.45	-9.92
18	3.55	2.02	-23.82	-3.18
19	0.35	2.46	-7.51	-8.79
20	3.32	-0.39	-7.79	-0.17
21	0.93	3.39	1.38	0.44
22	-1.85	1.27	-10.84	4.32
23	-0.75	0.43	-1.19	1.66
24	1.88	2.02	-2.47	3.61
Overall	0.73	-0.52	-2.41	0.64

Table : Seasonal Standard Deviations of Post-CB DA-RT Spreads

Hour	November - January	February - April	May - July	August - October
1	7.40	7.31	8.20	4.77
2	10.37	10.48	9.36	6.72
3	8.54	13.02	13.68	8.78
4	11.23	15.42	15.22	11.65
5	9.53	10.54	14.43	9.05
6	17.46	20.04	9.15	5.22
7	19.30	22.21	14.97	22.00
8	9.44	55.29	51.24	8.76
9	56.30	17.93	7.34	17.61
10	31.29	77.13	15.91	20.01
11	30.15	98.84	27.36	36.17
12	29.90	60.60	23.17	10.13
13	53.25	20.47	36.14	27.79
14	9.72	10.40	61.68	7.89
15	16.01	21.82	56.80	49.59
16	14.09	6.75	99.58	118.79
17	14.82	20.07	111.17	108.03
18	22.96	10.59	117.88	81.25
19	40.83	12.25	47.47	78.20
20	13.03	29.47	60.89	63.16
21	20.30	7.40	21.73	43.60
22	40.20	7.09	70.50	23.48
23	18.00	10.51	22.21	16.58
24	14.12	14.65	44.51	5.70
Overall	25.50	33.81	51.67	46.29

Optimal Numbers of States and Clusters

- We choose the number of states $M = 2$ to avoid the overfitting problem commonly encountered in learning a large state-space HMM.
- We choose the number of clusters $N = 3$, according to "elbow criterion".

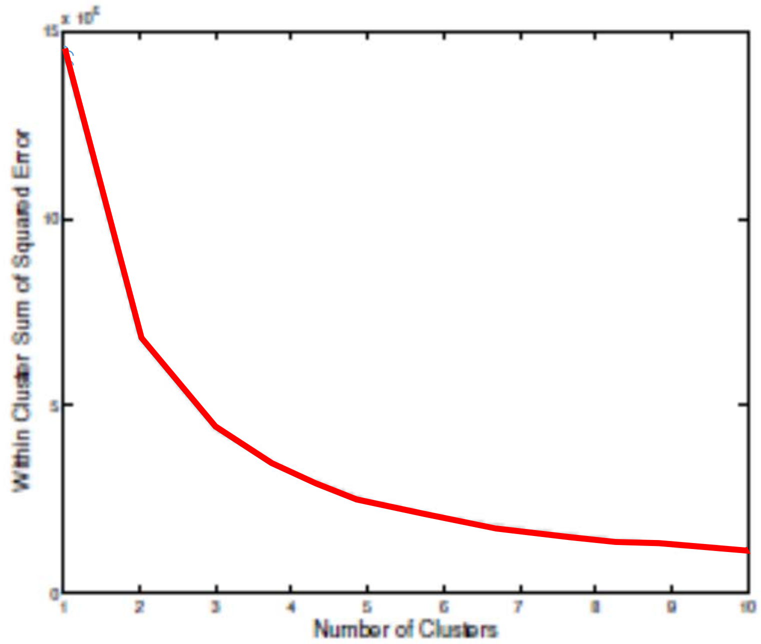


Figure : Post-CB Within-Cluster Sum of Squared Error

Transition probabilities of the Post-CB GMHMM

Table : Transition Probabilities of the Post-CB GMHMM

	State 1	State 2
State1	95.23%	4.77%
State2	4.00%	96.00%

Table : Cluster Probabilities of the Post-CB GMHMM

	Cluster 1	Cluster 2	Cluster 3
State1	89.65%	10.35%	0.00%
State2	56.84%	42.10%	1.05%

Post-CB Posterior State Probabilities

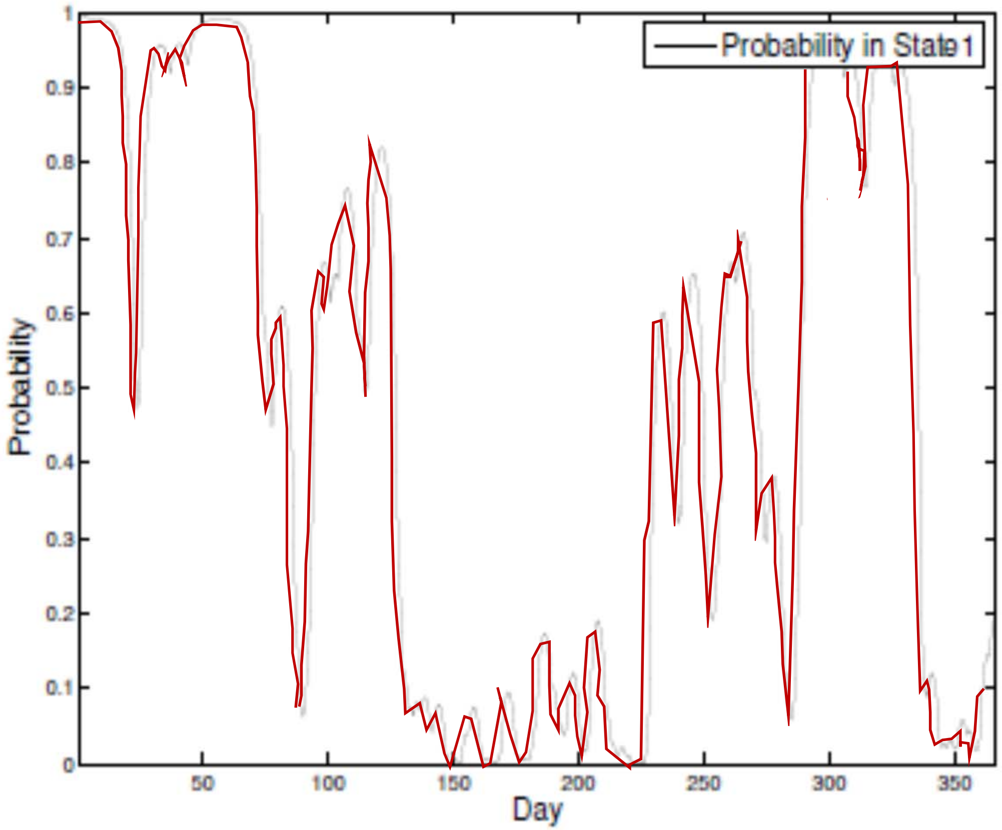


Figure : Post-CB Posterior State Probability $\lambda_1(t) = P(s_t = 1|r_1, \ldots, r_T)$

Marginal Distribution of Post-CB DA-RT Spread

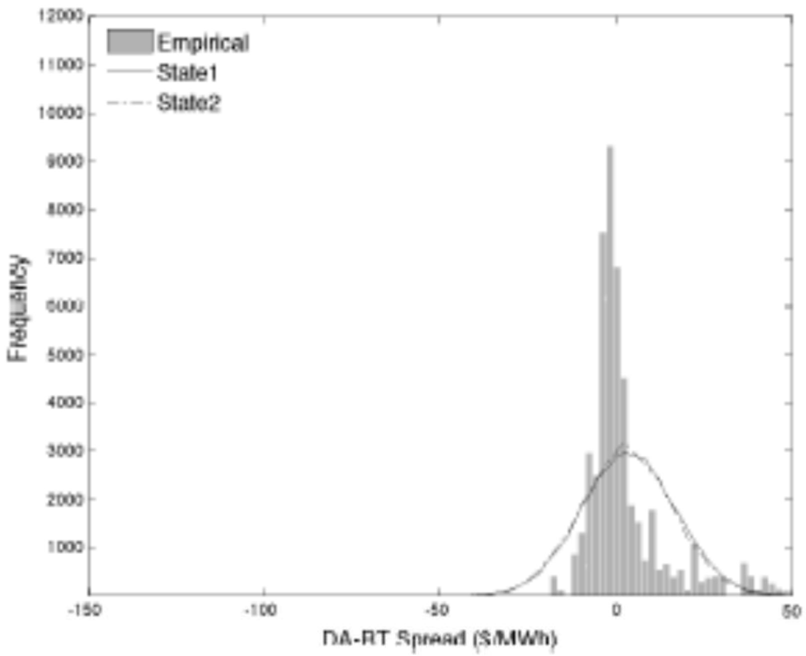


Figure : Marginal Distribution of Post-CB DA-RT Spreads for 3 a.m.

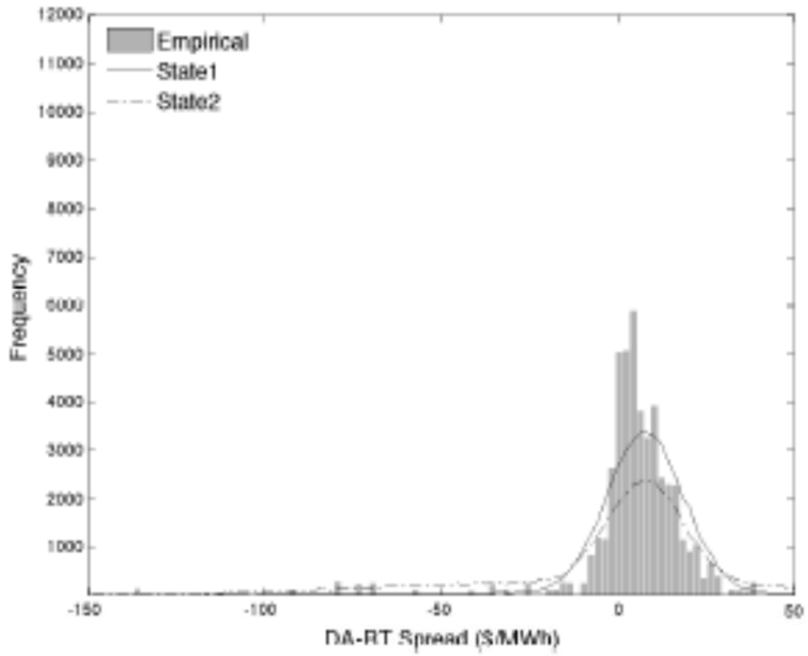


Figure : Marginal Distribution of Post-CB DA-RT Spreads for 3 p.m.

Regression Analysis

- To further test the implications of Bessembinder and Lemmon (2002), we regress the means for each of the 24 hours in the 2 states of the post-CB GMHMM on the corresponding variance and skewness measures in Table 6.
- The regression specification can be written in the form of (23),

$$StateMean_i = \theta_0 + \theta_1 StateVar_i + \theta_2 StateSkew_i + \epsilon_i. \tag{23}$$

Table 10 Pre-CB Regression Analysis

θ_0	θ_1	θ_2	t_{θ_0}	t_{θ_1}	t_{θ_2}	R Squared	DF
0.9050	-0.0032	-1.1511	1.2250	-16.1359	-2.0904	0.8732	45

Table 11 Post-CB Regression Analysis

θ_0	θ_1	θ_2	t_{θ_0}	t_{θ_1}	t_{θ_2}	R Squared	DF
1.8041	-0.0021	-1.0617	4.6235	-13.4369	-4.2229	0.8246	45

In Sample and Out of Sample Tests

- In the in-sample test, the performance is tested over the same dataset used to fit the model. Technically, this should yield the best possible results.
- In the out-of-sample test, the strategy is evaluated over a period which is different from the one the strategy is optimized on.

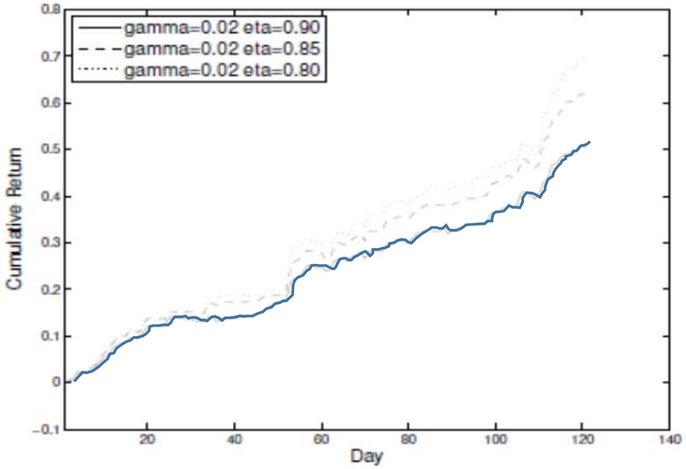


Fig. 6 Pre-CB In-Sample Performance under a VaR constraint

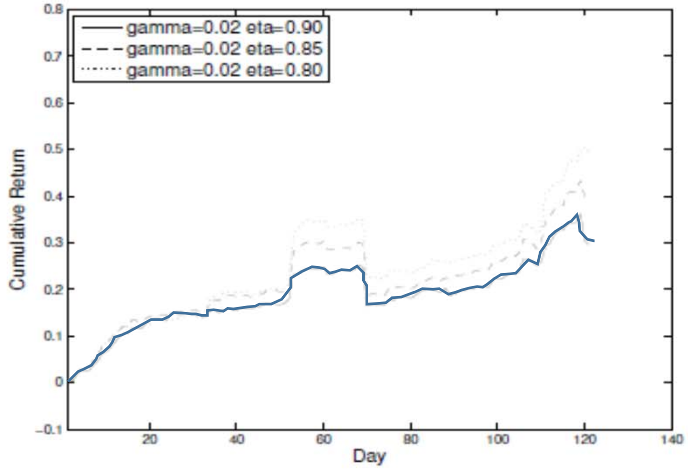


Fig. 7 Pre-CB Out-of-Sample Performance under a VaR constraint

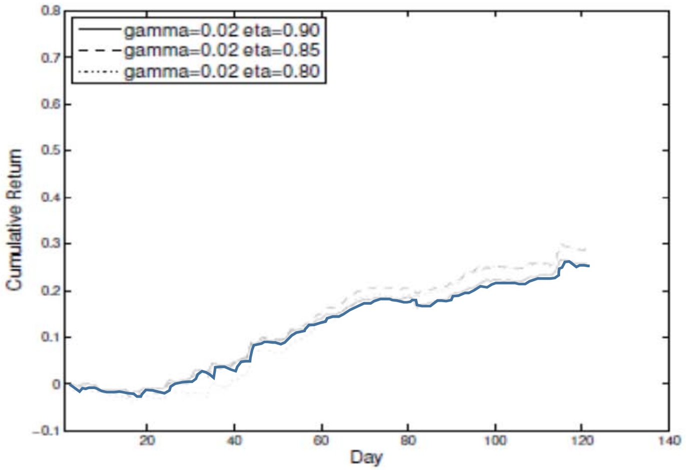


Fig. 8 Post-CB In-Sample Performance under a VaR constraint

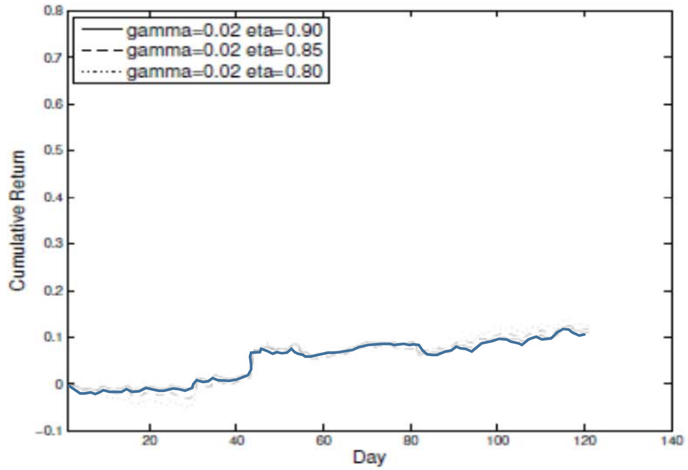


Fig. 9 Post-CB Out-of-Sample Performance under a VaR constraint

Performance Under a VaR constraint

Table : Performance under a VaR Constraint

Strategy Parameter	Expected Return	Standard Deviation	Sharpe	Max Drawdown
	Pre-CB In-Sample Performance			
$\gamma = 0.02$ $\eta = 0.90$	153.42%	14.60%	10.34	1.12%
$\gamma = 0.02$ $\eta = 0.85$	187.50%	19.54%	9.48	1.53%
$\gamma = 0.02$ $\eta = 0.80$	210.34%	24.71%	8.43	2.09%
	Pre-CB Out-of-Sample Performance			
$\gamma = 0.02$ $\eta = 0.90$	90.36%	24.50%	3.58	6.91%
$\gamma = 0.02$ $\eta = 0.85$	118.70%	29.42%	3.95	8.96%
$\gamma = 0.02$ $\eta = 0.80$	150.92%	35.15%	4.23	9.45%
	Post-CB In-Sample Performance			
$\gamma = 0.02$ $\eta = 0.90$	78.03%	11.10%	6.79	2.23%
$\gamma = 0.02$ $\eta = 0.85$	86.66%	13.08%	6.42	2.82%
$\gamma = 0.02$ $\eta = 0.80$	87.61%	15.21%	5.59	3.72%
	Post-CB Out-of-Sample Performance			
$\gamma = 0.02$ $\eta = 0.90$	33.47%	13.03%	2.35	2.73%
$\gamma = 0.02$ $\eta = 0.85$	35.11%	14.35%	2.25	3.19%
$\gamma = 0.02$ $\eta = 0.80$	38.03%	16.55%	2.13	5.08%

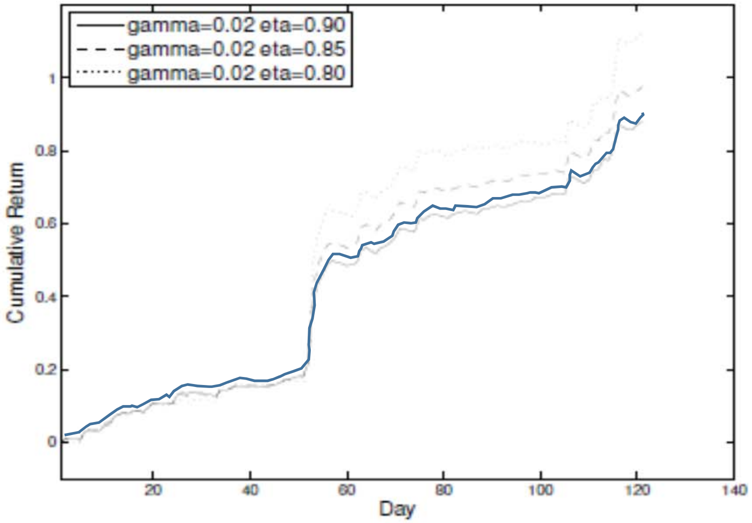


Fig. 10 Pre-CB In-Sample Performance under a CVaR constraint

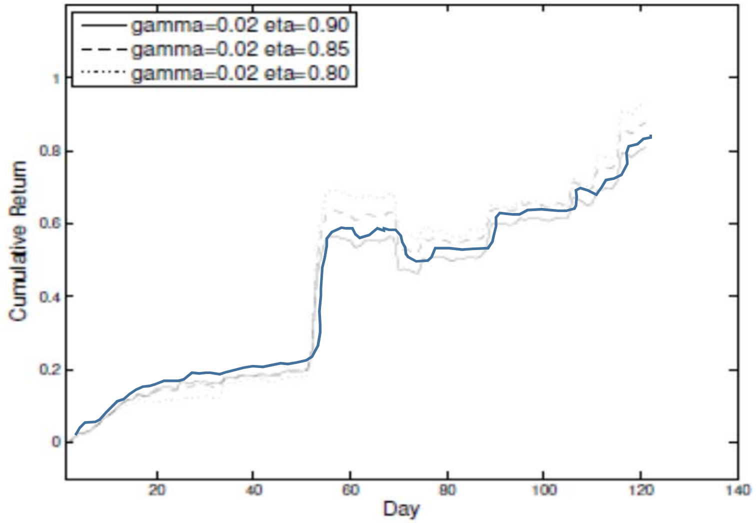


Fig. 11 Pre-CB Out-of-Sample Performance under a CVaR constraint

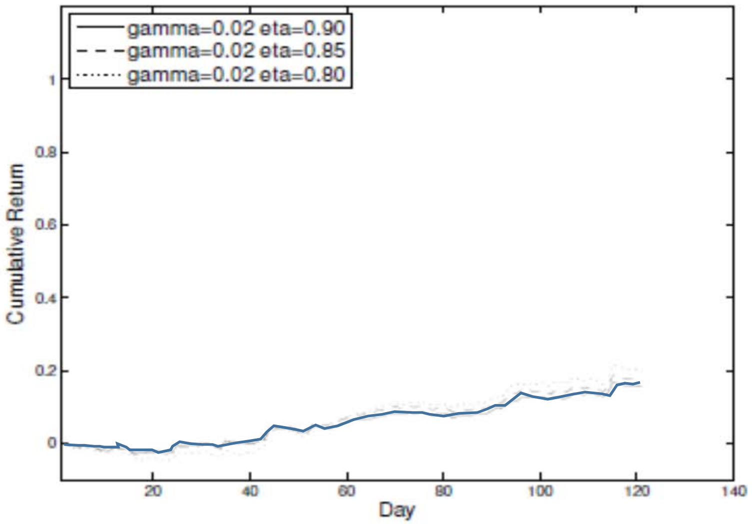


Fig. 12 Post-CB In-Sample Performance under a CVaR constraint

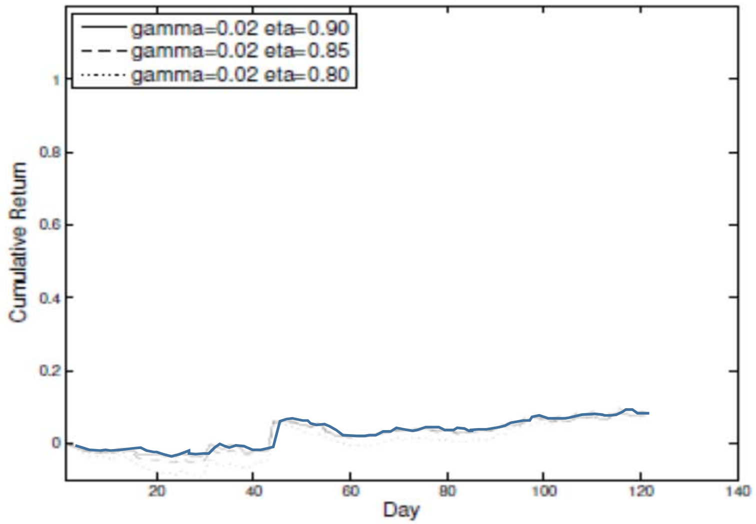


Fig. 13 Post-CB Out-of-Sample Performance under a CVaR constraint

Performance Under a CVaR Constraint

Table : Performance under a CVaR Constraint

Strategy Parameter	Expected Return	Standard Deviation	Sharpe	Max Drawdown
Pre-CB In-Sample Performance				
$\gamma = 0.02$ $\eta = 0.99$	267.77%	37.04%	7.18	1.19%
$\gamma = 0.02$ $\eta = 0.98$	296.75%	44.60%	6.61	1.14%
$\gamma = 0.02$ $\eta = 0.95$	341.83%	59.30%	5.74	1.42%
Pre-CB Out-of-Sample Performance				
$\gamma = 0.02$ $\eta = 0.99$	245.86%	47.78%	5.10	6.59%
$\gamma = 0.02$ $\eta = 0.98$	266.79%	55.65%	4.76	7.23%
$\gamma = 0.02$ $\eta = 0.95$	284.81%	66.01%	4.29	7.80%
Post-CB In-Sample Performance				
$\gamma = 0.02$ $\eta = 0.99$	47.35%	11.54%	3.86	2.75%
$\gamma = 0.02$ $\eta = 0.98$	52.93%	13.29%	3.77	3.16%
$\gamma = 0.02$ $\eta = 0.95$	60.82%	16.58%	3.50	4.60%
Post-CB Out-of-Sample Performance				
$\gamma = 0.02$ $\eta = 0.99$	22.58%	16.56%	1.19	4.02%
$\gamma = 0.02$ $\eta = 0.98$	25.55%	17.87%	1.27	5.29%
$\gamma = 0.02$ $\eta = 0.95$	22.18%	22.83%	0.84	9.01%

Conclusions

- Clearly, the deteriorated profitability in the post-CB period provides compelling evidence for the improved market efficiency, which demonstrates the benefit of CB.
- The profitability in the post-CB period, however, conveys empirical implications that can be interpreted differently, depending on the level of competition and the level of risk aversion of virtual traders.
- If virtual traders are risk-neutral and the competition among virtual traders is intense, the profitability in the post-CB period is convincing evidence against the fully efficient DA and RT markets.
- Otherwise, the profitability in the post-CB period might only rationally reflect the economic profit to incentivize the participation of risk averse virtual traders, which has nothing to do with market inefficiency and the mispricing of financial instruments.