Utilities Policy 18 (2010) 165–173

Contents lists available at ScienceDirect

Utilities Policy

journal homepage: www.elsevier.com/locate/jup

Three-part auctions versus self-commitment in day-ahead electricity markets

Ramteen Sioshansi [a], Shmuel Oren [b,*], Richard O'Neill [c]

[a] *The Ohio State University, United States*
[b] *University of California, Berkeley, CA 94720, United States*
[c] *Federal Energy Regulatory Commission, United States*

ARTICLE INFO

Article history:
Received 28 November 2009
Received in revised form
24 May 2010
Accepted 26 May 2010

Keywords:
Day-ahead electricity market
Centralized market
Unit commitment
Self-commitment
Combinatorial auction

ABSTRACT

We examine the economic consequences of a bid-based security-constrained centralized unit commitment paradigm based on three-part offers, which is the prevalent day-ahead market-clearing mechanism in restructured electricity markets in the United States. We then compare this approach with an energy-only auction with self-commitment (such as in Australia) addressing efficiency and pricing as well as the tradeoff between coordination losses and incentives to bid truthfully.

1. Introduction

The introduction of competition in the electricity supply industry has led to a number of important questions regarding the need for organized markets to efficiently and reliably coordinate the power system, and the desirable features and scope of those markets. Complicating electricity market design, power systems are subject to a number of 'network' constraints, in that these constraints depend on the actions of every market participant and each participant can impose an externality on others using the power system. Furthermore, generators' cost structures are non-convex due to startup and no-load cost components and generating units are constrained in the time it takes them to startup or shut-down, and the rate at which they can adjust their output. Thermal units typically have non-zero minimum generating constraints and 'forbidden zones,' in which they cannot operate stably, when they are online. Other types of generating units, such as combined-cycle gas turbines (CCGT) and cascaded watershed hydroelectric systems tend to have complex constraints restricting their operation. Due to the stochastic nature of demand fluctuations, generators must be able to adjust their real and reactive power outputs in real-time to ensure constant load balance. Other random contingencies such as transmission equipment failures or forced generator outages may also require generators to adjust their outputs within a short period of time to maintain system reliability. Thus, efficient and reliable operation of the system requires having a sufficient number of generators online and available to react to variations in load and other contingencies at least cost.

These complexities call into question the ability of decentralized markets, where suppliers respond autonomously to market signals, to efficiently and feasibly commit and dispatch units while respecting power system constraints. On the other hand, while a centralized market can, in theory, find the most efficient dispatch of the generators, the market designs suffer equity and incentive problems. Decentralized designs can overcome some of these issues but will suffer efficiency losses due to the loss of spatial and temporal coordination among resources. These design issues arise particularly in the context of determining the proper role for the system operator (SO) in making day-ahead unit commitment decisions.

As electricity markets in various countries have been restructured and have evolved, different approaches have been used with varying degrees of success. In the US, for example, the move towards standard market design has led to heavy reliance on open and transparent centralized markets where an SO operates central energy markets and has the authority to commit and schedule generators based on load forecasts and a multipart auction with offers specifying nonconvex cost components and unit-operating constraints (including startup and no-load costs, minimum load, energy offer curve, ramp rates, and generator-specific operating limits) [see Bowring (2006) for one example of such a market in

* Corresponding author.
E-mail addresses: sioshansi.1@osu.edu (R. Sioshansi), oren@ieor.berkeley.edu (S. Oren), Richard.oneill@ferc.gov (R. O'Neill).

0957-1787/$ – see front matter
doi:10.1016/j.jup.2010.05.005

the US]. The unit commitment in these systems is based on a centralized market-clearing mechanism using a security-constrained bid-based economic dispatch optimizer. Output levels of generators can then be adjusted at a second set of settlement prices in a real-time market. The British market, by contrast, started with a centralized market in the original Electricity Pool and moved to a more decentralized design under the New Electricity Trading Arrangements (NETA) and subsequent reforms under the British Electricity Trading and Transmission Agreements (BETTA), which were meant to overcome some of the problems experienced under the original centralized pool design [see Newbery (2006)]. The Australian National Electricity Market relies on self-commitment by generators who then submit energy-only offers into a single-settlement market [see Moran and Skinner (2008)]. Although these design differences are driven to a large extent by realities of the market such as asset ownership, generation mix, and system infrastructure, different 'philosophies' regarding the proper role of centralized markets have also played a role in determining the scope of any centralized markets.

This paper revisits the issue of dispatch efficiency raised by the design of markets based on centralized versus decentralized day-ahead dispatch and examines other economic implications of the two approaches attributable to the problem structure of unit commitment optimization and to the 'cost of anarchy' in self-commitment. Unfortunately, in addressing these issues one must recognize that 'the devil is in the details' and much of the discussion hinges on the computational complexity of the unit commitment problem. Until recently, the Lagrangian relaxation (LR) algorithm was the only practical means of solving a commercial-scale unit commitment. However, advances in optimization technology over the last decade have enabled the formulation and solution of such problems as mixed-integer programs (MIPs) using branch and bound (B&B) algorithms [see Streiffert et al. (2005), which discusses the use of MIP in PJM, the largest SO market in the United States]. One of the issues that has traditionally plagued the use of MIP in solving unit commitment problems has been the inability to provide a solution within a reasonable amount of time. However, an MIP-based formulation has significant advantages as it can represent complex units, such as CCGTs, pumped hydroelectric storage, and cascaded watershed hydrosystems, better than LR can. Furthermore, an MIP-based solution algorithm allows SOs to easily introduce new types of unit-operating and system constraints to the formulation of the problem and is less dependent on heuristics that must be tailored to specific resource characteristics. In contrast to LR methods, even if the B&B algorithm times-out before finding an optimum, one is still left with a primal-feasible solution and a bound on the optimality gap.[1] These intermediate solutions are often found within the same amount of time that an LR-based algorithm takes, and typically have optimality gaps of the same size or smaller than LR solutions. These overwhelming advantages and the tractability of MIP algorithms have led SOs, such as PJM, and the California ISO's Market Redesign and Technology Update, that started on April 1, 2009, to implement MIP-based solution methods as opposed to LR. Furthermore, the forthcoming Texas nodal market redesign features centralized commitment solved using MIP, and ISO New England (ISONE) is similarly exploring a switch from LR to MIP.

Due to the computational complexity of unit commitment problems and limited solution times, SOs that implement B&B-based algorithms do not solve their unit commitment problems to complete optimality or prove that the best solution found is optimal.[2] PJM, for instance, allows its MIP optimizer to run within a certain period of time or until the optimality gap is below some maximal threshold, and uses whatever intermediate integer-feasible solution the solver has found. An obvious issue raised in using MIP to solve the commitment is, therefore, how robust the solution is in terms of economic efficiency and fairness to market participants. Such issues have been studied in the context of an LR-based unit commitment by Johnson et al. (1997). Sioshansi et al. (2008a) reexamined the issue under an MIP formulation for a commercial-scale unit commitment problem based on an ISONE dataset and solved to optimality with a B&B algorithm.

We will review these results and discuss their policy implications with regards to the implementation of a centralized day-ahead market based on a security-constrained bid-based economic dispatch. We then compare these results to an alternative design based on self-commitment and an energy-only single-settlement auction wherein generators offer energy supply curves and are paid only for the energy they sell. We review simulation studies of such a decentralized approach conducted by Sioshansi et al. (2008b) using the same ISONE data employed for the centralized MIP-based approach. The decentralized market is modeled as a simultaneous Walrasian-type day-ahead auction for 24 h in which generators decide whether to commit their units (i.e., turn them on) and how much energy to offer in order to maximize their profits at the posted energy prices. The auctioneer will iteratively adjust the vector of hourly energy prices until the demand is served by the generators. We model our decentralized self-committed market after the multiround auction design proposed by Wilson (1997), although our specific market design differs slightly. The Wilson's (1997) proposal called for generators and loads to iteratively submit one-part offers, subject to some proposed activity rules,[3] specifying prices at which they would be willing to supply and consume energy until converging to an equilibrium set of prices, commitments, and dispatches. Our simulated market, instead, assumes that the auctioneer iteratively adjusts prices and the generators declare how much energy they are willing to provide at those set of prices. This market could also allow for loads to participate by declaring how much energy they would consume at the price vector given by the auctioneer, but because we assume loads to be fixed, the demand side does not participate in our market. It is worth noting that the Wilson's (1997) proposal was geared towards the original California market design, and while some of the elements were used, the design eventually settled on a simpler single-round auction. We do not advocate or refute this iterative auction design but rather employ it as an algorithm that will converge to a set of hourly energy prices that are consistent with individual rationality of the generators regarding their commitment and output decisions over a 24-h period. Our analysis focuses on the resulting energy prices, which should be independent of the process by which they are arrived at.

The comparison between the central unit commitment and decentralized market paradigms is based on competitive benchmarks assuming that generators do not behave strategically in

[1] In theory, the B&B algorithm may time out before finding an integer-feasible solution, in which case heuristics or an alternative solution method would have to be employed. Indeed, Streiffert et al. (2005) mention the concern when introducing MIP to the PJM market that the B&B algorithm could time out without finding a feasible solution. They overcame this issue by running the old LR-based algorithm in parallel as a backup. This was eventually phased-out due to the excellent performance of the MIP.

[2] Since LR-based algorithms almost invariably have a duality gap, they are not solved to optimality either.

[3] The purpose of the activity rules is to ensure early price discovery, fast convergence, and to prevent large generators and consumers from manipulating the auction by withholding themselves from the market until the final round of bidding.

manipulating their offers in the two markets and instead submit truthful offers. This comparison shows the extent of productive efficiency losses from a decentralized market and the distributional consequences of the two designs but does not capture incentive effects that will be discussed qualitatively. Sioshansi and Nicholson (2007) consider the incentive properties of the two market designs using a simplified symmetric duopoly model. They show that although generators generally have incentives to overstate their costs, offer caps can be designed to make the two markets cost equivalent. Sioshansi and Nicholson (2007) assume, however, that the SO is able to solve the commitment problem to optimality in the centralized market, and do not account for the incentive effects of suboptimal unit commitments.

The simulation results discussed in this paper are not new. They are pulled from the two previous papers by the authors. The main contribution here is a comparative examination of these results that highlights the tradeoffs between the centralized and decentralized approaches to day-ahead unit commitment and policy implication of such tradeoffs. We also examine the market design implications of nonconvexities and computational complexity of the auction clearing engine. To clarify some of these issues we produce a simple example that illustrates the causes for coordination losses in decentralized markets with nonconvex cost functions and the potential impact of such decentralization on energy prices faced by consumers.

2. The unit commitment problem, energy pricing, and properties of LR and B&B solutions

The unit commitment problem finds the least-cost commitment and dispatch of a set of generating units to meet expected load over a time horizon consisting of a fixed number of periods, typically, twenty-four single-hour periods. The problem can be formulated as an MIP in which the operating status of each unit (online or offline) in each planning period is characterized by a set of binary variables, and a set of continuous variables indicate the generating output of each unit in each planning period. In addition to a load-balance constraint, which ensures that expected demand is met in each period, unit commitment formulations will typically also have ancillary service requirements, upper- and lower-generating capacities for each unit, ramping constraints, minimum up and down times when units are started and stopped, startup costs (which can depend on the length of time that a unit has been offline), and transmission network constraints. The objective function to be minimized includes energy, no-load, and startup costs for each generating unit. Bidders in a centralized day-ahead market specify these costs as part of their offers in a multipart auction along with operating constraints that are included in the constraints.

Historically, solving a commercial-scale problem with hundreds of generating units was impractical using a B&B algorithm. As such, LR techniques were employed in which a Lagrangian dual is obtained by relaxing the load-balance, reserve, and any other 'coupling' constraints and penalizing violations in the objective function. When these constraints are relaxed, the problem can be decomposed into a set of problems for each generating unit, making the dual problem relatively simple to solve. The LR algorithm then works iteratively to try and find a set of energy and reserve capacity 'prices' (the objective function penalty coefficients), which incent an optimal commitment and dispatch of the units while satisfying the relaxed constraints. However, the solution resulting from solving the decomposed dual problem is typically not primal-feasible, due to nonconvexities, and additional processing of the LR solution and heuristics are needed to restore primal-feasibility. Hence the LR solution to the unit commitment problem is inherently approximate and, as demonstrated by Johnson et al. (1997), the objective function near the optimum is relatively 'flat' so that there are multiple feasible solutions that differ only slightly in terms of the respective values of the objective function. This observation has been confirmed by Sioshansi et al. (2008a) using a dataset from ISONE with 276 dispatchable generation units. Once a commitment and dispatch is obtained from the centralized unit commitment, generators are typically given linear energy payments based on the dual variables associated with the hourly load-balance constraints, which give the marginal cost of energy in each hour.

Regardless of whether the unit commitment is solved using LR, B&B, or another technique, these prices are often found either from the unit commitment problem itself or from an optimal power flow (OPF) problem. If they are based on the unit commitment solution, this is typically done by fixing the integer variables in the problem at their final value and resolving the resulting dispatch problem, which is continuous and will yield dual variables. If the unit commitment problem includes a load-flow model, then the resulting load-balance constraints at each network bus will yield a set of locational marginal prices (LMPs) at each bus in the network. Otherwise, the single system wide load-balance constraint will yield a single market-clearing marginal cost price (MCP). If the OPF problem is used for pricing, this will generally take the unit commitments as fixed and solve for the optimal power flow (oftentimes using a more complex ac load-flow model, as opposed to a dc approximation) for each planning period, ignoring intertemporal constraints. Again, the dual variables associated with the load-balance constraints at each bus are used to determine the LMPs.

Thus, the LMPs found by the unit commitment and OPF model will generally differ due to their different treatments of intertemporal and load-flow constraints. When there are no binding load-flow constraints, the OPF LMPs will simply be set by the highest marginal cost unit, which is not running at minimum load. The unit commitment prices, on the other hand, will reflect intertemporal cost shifting when price-setting units have binding intertemporal ramping constraints. To see this, consider the simplified dispatch problem in time period t:

$$\begin{aligned} &\min_{q_{g,t}} \sum_g C_g\left(q_{g,t}\right) \\ &\text{s.t } D_t = \sum_g q_{g,t} \\ &Q_g^- \le q_{g,t} \le Q_g^+, \end{aligned}$$

where C_g is generator g's cost function, D_t is the demand in hour t, $q_{g,t}$ is the amount of power provided by generator g, and Q_g^- and Q_g^+ are lower- and upper-bounds on generator g's output. If we let λ_t denote the Lagrange multiplier associated with the load-balance constraint and $\mu_{g,t}^-$ and $\mu_{g,t}^+$ the multipliers for the lower- and upper-bound constraints, then the first-order necessary condition (FONC) for an optimum implies:

$$\lambda_t = C_g'\left(q_{g,t}\right) + \mu_{g,t}^+ - \mu_{g,t}^- \text{ with } \mu_{g,t}^+,\ \mu_{g,t}^- \ge 0.$$

Under this formulation the cost of serving an incremental demand unit, which defines the MCP, is λ_t, whereas $\mu_{g,t}^-$ and $\mu_{g,t}^+$ are zero when the lower- and upper-bound constraints on generator g's output are not active. It follows that the MCP is set by the marginal cost of whichever generator is not operating at its lower- or upper-bound constraint, i.e, $\lambda_t = C_g'(q_{g,t})$. If ramping constraints are binding in the unit commitment, expensive peaking units may have to be started or dispatched at increased output levels during shoulder periods in order to meet demand at the peak period. Thus setting the price at the marginal cost of the most expensive operating unit unfairly penalizes consumption during shoulder periods and produces perverse incentives for load shifting that could alleviate ramp constraints. Such distortions could be corrected by setting energy prices based on the unit commitment solution, which explicitly accounts for the intertemporal ramp constraints. To illustrate this point we introduce intertemporal ramping constraints in the above simplified dispatch problem which becomes:

$$\begin{aligned}\underset{q_{g,t}}{\text{Min}} \sum_{g,t} C_g\left(q_{g,t}\right)\\ \text{s.t } D_t = \sum_{g} q_{g,t}\ \forall t\\ Q_g^- \le q_{g,t} \le Q_g^+\ \forall t\\ -R_g \le q_{g,t} - q_{g,t-1} \le R_g \quad \forall t,\end{aligned}$$

where R_g is generator g's ramping limit. If we let $\eta_{g,t}^-$ and $\eta_{g,t}^+$ denote the Lagrange multipliers associated with ramp-down and ramp-up constraints respectively, the FONC becomes:

$$\begin{aligned}\lambda_t &= C_g^{'}\left(q_{g,t}\right) + \mu_{g,t}^+ - \mu_{g,t}^- + \eta_{g,t}^+ - \eta_{g,t+1}^+ - \eta_{g,t}^- + \eta_{g,t+1}^-,\\ &\text{with } \mu_{g,t}^-, \mu_{g,t}^+, \eta_{g,t}^+, \eta_{g,t+1}^+, \eta_{g,t}^-, \eta_{g,t+1}^- \ge 0\end{aligned}$$

The Lagrange multipliers are zero if the corresponding constraint is non-binding, hence if the only binding constraint on the price-setting generator, g, is its ramp-up constraint in hour t, then the FONC in hours $t-1$ and t become:

$$\lambda_{t-1} = C_g^{'}\left(q_{g,t-1}\right) - \eta_{g,t}^+ \text{ and } \lambda_t = C_g^{'}\left(q_{g,t}\right) + \eta_{g,t}^+.$$

Because the multiplier, $\eta_{g,t}^+$, is non-negative these conditions imply that the MCP in hours $t-1$ will be subsidized by an increased MCP in hour t. Such an intertemporal subsidy would mitigate the higher prices imposed on period $t-1$ due to the ramp constraint activated by the demand in period t. It assigns the increased cost to those who cause it and it creates the correct incentives for load shifting from period t to period $t-1$. From generator g's perspective the above price adjustment is beneficial since it increases its revenue by $\eta_{g,t}^+ \cdot R_g$, which is the shifted price increment from period $t-1$ to period t times the increase in generator g's output level. For inframarginal generators operating at their upper-bound or generators operating at minimum load the adjustment is revenue neutral. A binding ramp-down constraint would have a similar but opposite effect on the MCPs.

In spite of the compelling argument in favor of accounting for ramp constraints in setting marginal cost prices, this is not done in practice and LMPs are set for each time interval with no consideration of ramp constraints (even when such constraints are enforced in determining the generators' output level). The primary reason for this approximation is the computational burden of accounting for intetemporal constraints in conjunction with a full network representation OPF. In our simulation, however, the two sets of prices are nearly identical so our analysis uses the unit commitment prices, which are produced as a byproduct of the unit commitment optimization.

Table I shows the value of the objective function and corresponding duality gaps for 11 near-optimal solutions produced by LR for the above dataset in comparison to the true MIP-optimal solution produced by B&B. These results are based on a simplified model of the ISONE commitment problem, which includes minimum up and down times, ramping constraints, hourly load-balance constraints, and a single type of load-based ancillary service requirement. To simplify the model, marginal generating costs are assumed to be constant, startup costs are not time-dependent, and transmission constraints are ignored. The different LR solutions are generated by adjusting the rate of convergence of the step-size sequence used in the subgradient algorithm underlying the LR. The troubling issue highlighted by Johnson et al. (1997) and reconfirmed by Sioshansi et al. (2008a) is the fact that near-optimal solutions may result in large deviations in surplus accrued by individual generators and in energy prices. The last column in Table I shows the number of units that have a different dispatch in each near-optimal solution compared to the MIP-optimal solution. While such deviations were inconsequential in a regulated monopoly setting they have significant economic implications in a deregulated market with dispersed ownership of generation units. Furthermore, the near-optimal solutions can result in negative surplus for some generators, which is confiscatory and not sustainable. This creates incentive problems since generators may attempt to manipulate their offers to shift the solution in their favor, which may result in a suboptimal commitment and dispatch. One could argue that an MIP B&B solver that determines the true optimal solution would resolve the ambiguity resulting from the approximate nature of the LR approach. However, the MIP-optimal solution could be confiscatory as well, since the presence of alternate near-optimal solutions and the consequential incentives to misrepresent offers are inherent in the problem structure and do not depend on the solution technique.

Table I
Comparison of LR solutions and MIP optimum, with linear energy payments only.

Solution	Total cost ($)	Optimality gap (%)	Units affected
MIP	8074022.55		
1	8181665.93	1.33	125
2	8212269.01	1.71	128
3	8202929.15	1.6	132
4	8171416.63	1.21	141
5	8125904.86	0.64	109
6	8220547.52	1.82	128
7	8211208.66	1.7	132
8	8124845.51	0.63	112
9	8180528.2	1.32	129
10	8189867.05	1.44	125
11	8116566.01	0.53	112

The confiscatory issue is resolved by all SOs employing day-ahead centralized unit commitment by means of 'make whole' payments, which ensure that a centrally dispatched unit will cover its cost in any 24-h period. These payments are computed as the difference between the cost incurred by each generator (which are calculated on the basis of the cost components in the generator's multipart bids in the auction) and the energy payments received from the market—if this difference is positive. If so, then the sum of energy and make-whole payments ensure that the generator recovers all of its costs. If not, then this implies that the generator has earned inframarginal rents from energy payments, which it keeps as surplus. These make-whole payments do not, however, make the hourly dispatch *ex-post* incentive-compatible in every hour. The make-whole payments will, however, smooth out the payoff differences to individual generators between alternative near-optimal solutions and the MIP optimum by truncating the payoff distribution at zero. Table II shows the effect of make-whole payments on the distribution of surplus deviations (normalized on a per MWh basis) between the near-optimal LR solutions and the MIP optimum. Because the surplus calculations are based on costs specified in generator offers, as opposed to actual costs, the surplus calculations are offer-based (i.e. generators' actual surplus may differ from our calculations, since generators may have misrepresented their costs).

Table II
Comparison of unit offer-based surplus deviations between LR near-optimal solutions and MIP optimum.

Solution	As-bid surplus ($/MWh)					
	Without make-whole payments			With make-whole payments		
	Mean	Max	cv	Mean	Max	cv
1	61.19	942.24	3.66	0.33	4.68	1.76
2	71.92	945.92	3.38	0.33	4.8	1.79
3	73.66	939.3	3.31	0.33	4.8	1.79
4	57.46	943.55	3.33	0.39	4.8	1.61
5	7.28	952.56	8.2	0.33	4.8	1.79
6	71.9	945.92	3.38	0.29	4.54	1.8
7	73.64	939.3	3.31	0.29	4.54	1.8
8	5.57	353.56	5	0.29	4.54	1.8
9	62.95	931.94	3.56	0.29	4.54	1.8
10	61.22	942.24	3.66	0.29	4.54	1.8
11	5.6	353.56	4.97	0.33	4.8	1.79

3. MIP implementation

Although recent computational and algorithmic advances make direct solution of the unit commitment by B&B tractable, SOs cannot currently solve their commitment problems to complete optimality within the allotted timeframe. Most SOs, upon receiving generation offers and other market data day-ahead, must return commitments and a schedule to market participants within a few hours. The formation of these schedules oftentimes requires solving multiple unit commitment, optimal power flow, and other optimization problems. As such, SOs that have implemented or are proposing to use MIP in their unit commitment set limits on the solution time and rely on the best integer-feasible solution found at the end of that time. Although SOs boast their ability to find feasible solutions with minuscule optimality gaps, if an SO is left to rely on an intermediate integer-feasible but suboptimal solution, the same issues of generator payoffs, energy pricing, and inequity of the resulting dispatch arise as with suboptimal LR commitments.

To illustrate the consequences of truncating the B&B process before completion, Table III summarizes the progression of the MIP optimizer in CPLEX 9.120 solving the simplified ISONE unit commitment problem mentioned above with the default settings. It should be noted that unlike commercial MIP-based unit commitment software packages, the formulation of the problem or the settings in CPLEX were not fine-tuned, nor were problem-specific cutting planes introduced to improve the solutions or solution times of the problem.[4] CPLEX finds 5 intermediate suboptimal integer-feasible solutions, all of which have smaller optimality gaps than the near-optimal LR solutions shown in Table I. Moreover, should the SO use one of the intermediate solutions but include a make-whole provision, the net offer-based surplus to each unit is identical to that under the MIP optimum in all but the first solution, with the largest deviation being $0.02/MWh. As in the case of the near-optimal LR solutions, energy prices corresponding to intermediate solutions can deviate substantially in some hours from those corresponding to the MIP optimum, as shown in Table IV. In the absence of make-whole payments, the profitability of some units can vary erratically among near-optimal intermediate solutions. For instance two identical units (in terms of stated costs and operating constraint parameters) receive identical commitments and dispatches in the MIP optimum (with negative surplus to each) but are given different commitments and dispatches in every intermediate MIP solution.

[4] The integrality and optimality gap tolerances were set to zero in order to ensure that the final solution given by CPLEX is indeed the MIP optimum.

Table III
Progression of integer-feasible solutions found by B&B.

Solution	Total cost ($)	Optimality gap (%)
1	8074400.39	0.0049275
2	8074045.7	0.0005345
3	8074020.25	0.0002192
4	8074014.91	0.0001531
5	8074003.06	0.0000063

Although the above results suggest that make-whole payments resolve the payoff instability issues when using near-optimal LR solutions or intermediate B&B solutions, the formulation upon which these observations are drawn is a simplification of any actual unit commitment solved by SOs and excludes many important details. Sioshansi et al. (2008a) present a summary of the progression of integer-feasible solutions found by CPLEX in solving ISONE's complete unit commitment problem, which includes virtual transactions, demand bids, time-dependent startup costs, stepped generation costs, multiple types of ancillary services, and a dc load-flow model. Due to inclusion of demand side bids the problem is formulated so as to maximize total surplus of energy traded. In the course of optimizing the model, 38 intermediate integer-feasible but suboptimal solutions were found, which are quite revealing. Solutions that are very close to the optimum in terms of the MIP duality gap result in different payoffs to individual generators. Moreover, no monotonicity is evident in the pattern of surplus deviations or energy prices. Intermediate solutions, regardless of how close to optimal, can result in significant differences in energy prices with some extreme cases showing 10% deviations from the MIP optimum when the objective function value is a millionth of a percent away from the optimum. These observations are based on simulations that assume truthful revelation of costs and constraint parameters. However, the above observations are likely to lead to misrepresentation and manipulation attempts by market participants that engage in this process on a daily basis. Such distortions call into question the efficiency justification for the security-constrained bid-based economic dispatch which underlines the day-ahead market design. To address this question one must consider the alternative approach based on self-commitment and a single-part energy-only auction-based dispatch.

4. One-part energy auctions with self-commitment

The decentralized market model we analyze assumes that energy is traded through an energy-only market. As suggested in the discussion above, due to nonconvexities, network externalities, and other complexities of power systems, a decentralized market can suffer from both efficiency losses and higher settlement costs, even under a competitive assumption. To see this, consider a very simple example in which a baseload coal generator and a combustion turbine (CT) must be committed and dispatched to serve a 1000 MW load in a single hour. Table V summarizes the characteristics of the two generators in the example. Assuming that the two generators behave competitively and truthfully reveal their generation costs and operating constraints, then the least-cost solution to the centralized market would be to commit the coal generator and have it generate 1000 MW. The energy price would

Table IV
Energy prices corresponding to Intermediate B&B solutions.

Hour	Energy price of solution ($/MWh)					
	1	2	3	4	5	Optimum
2	45.84	44.3	44.3	44.3	44.3	44.3
11	114.95	64.49	59.72	59.72	59.72	59.72
23	48.96	50.24	48.96	50.24	50.24	50.24

Table V
Generator characteristics in example.

Generator	Capacity (MW)	Startup cost ($)	Variable cost ($/MWh)
Coal	2000	75,000	10
CT	200	0	75

be set at $10/MWh, but this generator would have to be given a $75,000 make-whole payment so that it could recover its fixed startup cost. In a decentralized energy-only market (which we still assume to be perfectly competitive), by contrast, the energy price would have to be raised to $85/MWh or higher in order for it to be individually rational for the coal generator to startup and provide energy. At this price, however, the CT, which is assumed to behave competitively and offer its energy at the marginal cost of $75/MWh, would want to be fully dispatched at prices between $75/MWh and $85/MWh, however the offer quantity below $85/MWh is not sufficient to meet demand. At $85/MWh the aggregate competitive supply function has a discontinuity and the offer quantity jumps from 200 MW to 1200 MW, which will require the SO to use some rationing rule in its energy procurement. Efficiency losses will occur if the CT is dispatched in this energy-only market, which is a likely outcome given that it submits the lowest offer. Moreover, if the CT is dispatched at any level then the coal generator will produce less than 1000 MW and the energy price will have to be raised above $85/MWh, in order to ensure that the coal generator recovers it startup costs (otherwise it will withdraw from the market). It should be noted that if the energy price is $85/MWh, then the two market designs would yield the same total settlement costs (when make-whole payments are taken into account), but if the CT is dispatched and the energy price in the decentralized market is raised above $85/MWh to assure sufficient supply. However, this market outcome will be inefficient and more costly to consumers than a coordinated centralized market. In an extreme case in which the CT is dispatched to its capacity of 200 MW and the coal generator to 800 MW, the energy price would have to rise to at least $103.75 in order to ensure the coal generator recovers its costs, which would yield a 22% increase in settlement costs to consumers, and a 15% increase in commitment and dispatch costs, which would be the allocative efficiency losses.

Although this example is a gross oversimplification of any actual electricity market, it highlights the fact that due to nonconvexities, energy-only markets will be prone to efficiency losses and higher settlement costs than a centralized market, even under a price-taking assumption, which is consistent with our findings based on the ISONE dataset.

To compute a competitive benchmark using the ISONE dataset, the market is modeled as a competitive auction in which the auctioneer[5] announces a set of hourly energy prices and price-taking generators individually determine their hourly commitments and output levels to maximize profits and submit offers to the auctioneer indicating how many MWh they are willing to supply in each hour. The auctioneer then iteratively adjusts the hourly energy prices until it finds a set of prices that incent sufficient generation to serve the load. This iterative price-updating process is meant to mimic the Wilson's (1997) proposal for a self-committed market with two important differences. One is that loads are fixed in each hour as opposed to being price-elastic.[6] Thus the market is assumed not to accept demand bids but rather solicit sufficient generation at any price to serve a fixed hourly load. The other is that under Wilson's proposal, generators are assumed to submit offers consisting of quantity/price pairs. Because the model analyzed in this section assumes generators to behave competitively, generators are modeled as price-takers, that take the auction prices as fixed and decide their commitments and generation offers to maximize profits individually, as opposed to strategically adjusting their energy offers to raise energy prices. It should also be noted that a generator's decision to produce energy is independent of whether it is contracted, since we assume that it can always fulfill any contractual obligation through purchases from the market. If a generator is not contracted, then the production decision is based on sales revenue versus the cost of production. On the other hand, if a generator is contracted then market prices represent the opportunity cost of not producing and the generator's decision is driven by minimizing the cost of fulfilling its contractual obligation.

Although the model assumes that energy is traded through a centralized energy market, it can be thought of as solving for a competitive equilibrium of direct bilateral trade between generators and consumers *a la* a Walrasian auction model. The model further assumes that the auctioneer starts with a set of prices that incent sufficient generation to serve the load, and iteratively adjusts prices until finding a set of supporting minimal prices—which is a set of prices such that generators offer sufficient energy to serve the load, but would no longer do so if any of the energy prices were reduced. As the energy prices are dropped, higher-cost units will no longer find it profitable to commit themselves and the total quantity offered for generation will be driven towards the system load.

One difficulty with finding a set of supporting minimal prices is that the binary nature of the generators' commitment decisions means that a set of supporting minimal prices will generally not be market-clearing, meaning generators will offer more total generation than there is load to serve, yet reducing any of the energy prices will cause a unit to decommit itself, leaving insufficient energy to serve the load.[7] One solution is to assume that the auction uses some type of rationing rule to determine how the load is divided amongst generators willing to commit themselves. Our model assumes, instead, that because higher-cost generators drop out of the commitment as prices are iteratively reduced, then if there is excess generation offered and multiple units are competing for the same load, the one with the lowest average cost over the course of the day will prevail.

In modeling generators' profit-maximizing behavior, they are assumed to perfectly rationally expect the behavior of other generators and take into account the 'winner determination assumption' in making their own commitment decisions. This is to preclude the possibility that a unit may commit itself in expectation of being dispatched but finds that it doesn't, resulting in a net profit loss. This assumption is enforced algorithmically in the model formulation and simulation. Finally, the model assumes that each generator acts independently in making its commitment decisions, as opposed to making commitment decisions for portfolios of generators being owned by generating firms. This assumption is made because the dataset used does not have unit ownership information, although the technique and results would translate to a setting with generation asset portfolios. As indicated earlier, the above assumptions are not intended to represent an accurate behavioral model but rather specify an algorithm that will lead to a set of prices that will support individually rational decisions of generators regarding daily self-commitment and output while meeting the load in each hour.

[5] The auction can be thought of as being operated by the SO, or it can be a separate outside market.

[6] In order for our simulations of the two market designs to be comparable, we use the same underlying cost, load, and generator constraint data in the two sets of market simulations.

[7] This inability to find market-clearing prices is a direct consequence of the non-zero duality gap that remains if the SO unit commitment problem is solved using an LR algorithm.

Table VI
Cost and profit comparison of centrally- and self-committed market designs.

Market design	Energy payments	Make-whole payments	Total settlements	Commitment costs	Total unit profits
Central	$16,075,121	$0.00	$16,075,121	$5,758,201	$10,316,920
Self	$25,060,666		$25,060,666	$6,003,274	$19,057,392
%-Difference			55.90%	4.16%	84.72%

5. Comparative analysis based on market simulation

To examine the pros and cons of centrally- and self-committed markets we conduct simulations of the two approaches based on actual market data from an ISONE unit commitment problem in February of 2005, consisting of 276 dispatchable units. This dataset is used because of availability of the data, and is meant to be an illustrative example of the relative efficiency losses and settlement costs between the two market designs. The ISONE system covers approximately 6.5 million retail customers, includes more than 350 generators with 31,000 MW of installed capacity, an all-time peak load of 28,127 MWh, and $11 billion of annual energy trade.

For ease of analysis and discussion the unit commitment formulation used in the simulations is a simplification of a commercial model. It includes stepped marginal costs and fixed (not time-dependent) startup and no-load costs for each unit. Demand is price-inelastic, there are no network flow constraints, virtual bids, self-schedules, or ancillary service requirements. The computations assume that the centrally-committed market settles with a uniform hourly energy price—specifically the dual variable associated with the hourly load-balance constraints in the unit commitment problem. As noted before, because linear energy-only payments can be confiscatory, the analysis further assumes that the SO includes a make-whole provision, which pays each unit the difference between its total costs incurred in the commitment and dispatch (calculated on the basis of costs stated in its offer) and the total energy payments received over the course of the 24 h, if that difference is positive. These payments ensure that total net profits (on the basis of stated costs) are always non-negative.

The operating costs and constraint parameters used in the market simulations are those which were submitted by generators to ISONE. The competitive benchmark assumption takes these generator-offered parameters as reflecting actual costs and unit-operating constraints—thereby assuming away any incentive compatibility issues. The computation of the central commitment assumes generators will offer these actual cost and constraint parameters to the SO for use in its commitment problem, as opposed to strategically misstating them to increase profits. The computation of the self-commitment assumes that generators behave as price-takers and maximize profits with the same cost and constraint parameters.

Table VI compares the total settlements paid to generators, commitment costs, and profits of the generators in the simulations of the two market designs. Although the centrally-committed market is assumed to include a make-whole provision, the dataset and optimal central commitment is such that each generator receives sufficient inframarginal rents to recover all of its costs and no supplemental make-whole payments are required. Nonetheless, Fig. I shows that the set of supporting minimal prices found in the self-committed market far exceed the energy prices paid in the central unit commitment.

Indeed, a critical assumption underlying a centrally-committed market is that the SO can force cross subsidies of ‘losing’ hours by profits from other hours and has a means of preventing generators from making adjustments to their assigned schedules. Fig. II shows the resulting load imbalances which would occur in a centralized market if generators could individually adjust their outputs to maximize profits against the hourly energy prices—known as uninstructed deviations. While the scale of these deviations may seem small, it is important to note that they would have disastrous consequences in a power system, since any difference between power demand and supply will threaten system stability and may cause brown- or black-out conditions. Due to the potential for such deviations, SOs penalize such deviations in generation by requiring generators to buy or sell

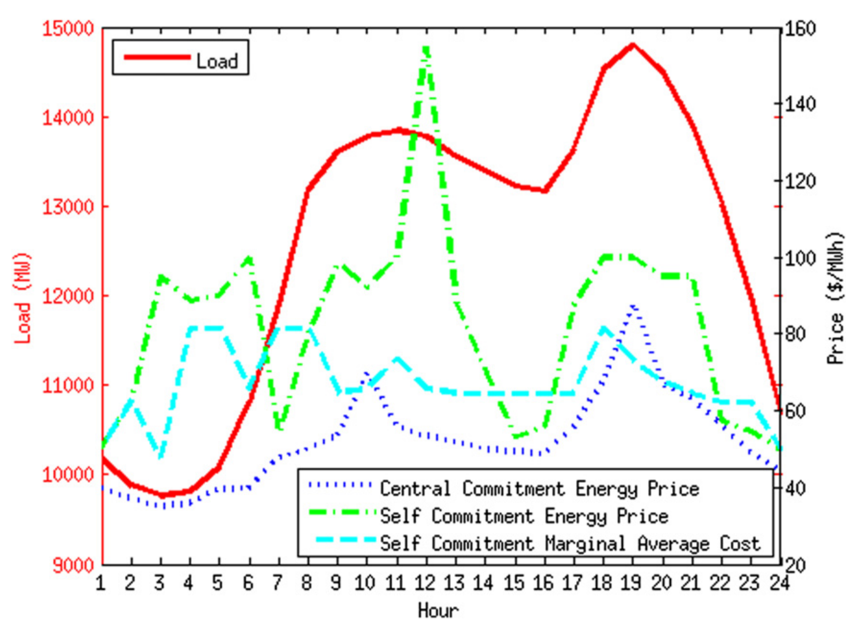

Fig. I. Energy prices under central unit commitment and self-commitment.

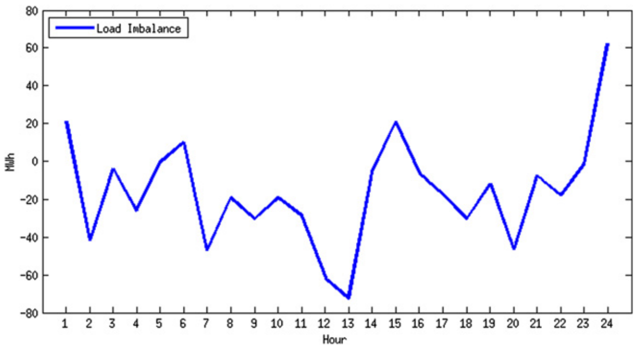

Fig. II. Load imbalance from profit-maximizing uninstructed deviations in centrally-committed market.

back their insufficient or excess generation at the locational marginal price (LMP)—oftentimes with an added deviation penalty—thereby removing any incentive for such deviations.

This enforcement mechanism can be problematic, however, in multiple-settlement systems in which the SO computes different sets of LMPs at different time intervals in real-time. Since there can be differences between the prices at which an uninstructed deviation is paid and penalized, a generator may be inclined to change its output if these price differences are predictable.

More importantly, the simulation demonstrates that a self-committed market requires higher energy prices than a centrally-committed one. Because the model assumes that demand is fixed and inelastic, these higher prices are simply a wealth transfer from consumers to generators, without any efficiency losses. In a market with demand response, the higher prices will generally result in allocative distortions. Indeed, if we return to the example described in Table V and suppose that demand is given by the price-elastic inverse demand function, $p(D) = 1010 - D$, we can show how the higher energy prices from centralized commitment will result in social and consumer welfare losses. With this inverse demand function, the welfare-maximizing solution of the centrally-committed market would be to still commit the coal generator and have it produce 1000 MW. This will yield a social surplus of \$425,000, a consumer surplus of \$500,000, and an energy price of \$10/MWh. Again, because the coal generator will not recover its startup costs a make-whole payment of \$75,000 will be required. Because consumer surplus is \$500,000, the cost of this make-whole payment could be imposed on consumers (which ensures that the market is revenue-adequate) in the form of a two-part tariff, that would not have any allocative efficiency losses.

In a self-committed market, by contrast, the energy price would have to rise to at least \$85/MWh in order for the coal generator to be willing to startup and provide energy, which would reduce demand. Moreover, because this \$85/MWh is greater than the marginal cost of the CT, it would want to startup and generate energy as well. The SO's rationing rule between the coal and CT generators in the self-committed market will determine the extent of productive and allocative efficiency losses, but as an upper-bound, we will consider a case in which the SO would favor the CT over the coal generator in its dispatch decision.[8] In such a case, because the CT would always offer itself at a lower price, it would be dispatched to produce 200 MW, and the coal generator would produce $(1010 - 200 - p) = (810 - p)$ MW, where p is the equilibrium energy price. The equilibrium energy price would then be a solution to the equation:

$$p = 10 + \frac{75000}{810 - p},$$

which ensures that the price is sufficiently high to recover all of the coal generator's costs. The solution to this equilibrium condition yields an energy price of about \$118.45, and a total demand of 891.55 MW.[9] Because of the higher price and lower demand, however, consumer surplus is reduced to \$397,431 (a 22% loss compared to centralized commitment), social surplus is \$406,120 (a 4% loss compared to centralized commitment), and generation costs rise by about 14%.

Table VI also shows that a self-committed market will generally suffer productive efficiency losses, as demonstrated by the more than 4% increase in total commitment and dispatch costs. These efficiency losses are not the result of units committing themselves under a self-committed market when they would not be committed under the central commitment. Rather these losses stem from the fact that a central commitment gives the most efficient coordination of generator dispatches, which are lost when generators dispatch themselves independently. Of the 276 units, 108 are committed in at least 1 h under the central unit commitment solution. Of these 108, 73 follow the same commitment and dispatch schedule under the self-committed market as under the central unit commitment, with some shuffling of generation amongst the remaining 35.

6. Conclusions

Clearly centrally- and self-committed markets present tradeoffs, which must be evaluated in addressing market design issues. Centrally-committed markets strive for the least-cost commitment and dispatch of generators by solving for a commitment that minimizes the SO's cost objective. However, as shown above and by Sioshansi et al. (2008a), this approach presents equity and incentive issues that also call into question the efficiency of the centralized solution. Although make-whole payments, which are made by most SOs that operate centrally-committed markets, reduce these equity and incentive issue, they do not completely eliminate them. As shown in Table II, generators can in some cases earn up to \$5/MWh of excess surplus from a suboptimal unit commitment solution. Given the fact that generators participate in these markets repeatedly, it is not inconceivable that they may learn to manipulate their bids in order to influence the resulting commitment and dispatch. Indeed, Newbery (2006) notes that because the original Electricity Pool market in Britain was operated using the same commitment and dispatch model that generators used under the monopoly regime, they were able to do this type of market manipulation.

Self-commitment, on the other hand, has been offered as a viable alternative, which addresses and reduces some of the issues with centrally-committed markets but suffers from loss of coordination amongst generators. This loss of coordination will result in some efficiency losses, mainly due to the nonconvexities in generator cost and operating constraints. Moreover, because generators' nonconvex cost components must be covered by energy payments, energy prices and total settlement costs will generally be higher than in a centrally-committed market. Thus, central- and self-commitment are two imperfect market models with inherent shortcomings since

[8] As we noted before, because the CT would offer its generation at \$75/MWh, whereas the coal generator would bid a higher price, it is not unreasonable to assume that the SO would favor the CT in its dispatch.

[9] A simple calculation will confirm that this price recovers all of the coal generator's costs.

centralized markets will be fraught with incentive problems and decentralized markets with coordination losses.[10]

On one hand an SO with broad economic authority can, in theory, determine the most efficient commitment to meet forecasted demand. However, centrally-committed markets, are not strategy-proof and are prone to incentive compatibility issues, meaning that generators can profitably manipulate their offers to increase profits. This has both been shown through simple examples, for example by Sioshansi and Nicholson (2007), and was one criticism of the original Electricity Pool in Britain. Proponents claim that a decentralized energy-only market, in which generators individually determine their commitments, can reduce the incentive issues of a central unit commitment while minimizing efficiency losses. The simulation of a competitive benchmark discussed above and conducted by Sioshansi et al. (2008a,b) provides an estimate of the productive efficiency losses from a self- as opposed to centrally-committed market design, under a competitive assumption. While these losses are relatively small, around 4% in the case examined here, this would nonetheless represent a significant welfare loss in absolute terms considering that SO markets typically trade energy worth billions of dollars on an annual basis. The efficiency loss in the market simulations would amount to an annual loss of nearly $90 million for the ISONE system, if the results are typical of most days.[11] Moreover, a self-committed market may also yield consumer surplus losses in the presence of demand response due to the higher energy costs that will generally be seen in self-committed markets. The potential for allocative efficiency losses with demand response may be an important consideration as these programs are slowly becoming more prevalent.

Finally, it is interesting to note that many centrally-committed markets also allow generators to self-schedule their generation, in which case they individually make commitment decisions and submit one-part energy-only bids. Thus, many markets operate as a hybrid of centrally- and self-committed markets. Our findings regarding energy prices in centrally- versus self-committed markets may suggest that generators in such markets would prefer self-committing, but in practice generators tend to favor using the centralized commitment. One possible reason for this behavior is that a generator that self-commits must recover all of its nonconvex costs through energy payments, and will have to roll these fixed costs into its one-part energy bid. Units that are centrally committed by the SO, on the other hand, are given make-whole payments to recover their nonconvex costs. These payments are not considered in the optimization underlying the centralized unit commitment. This different treatment of nonconvex costs can serve to make a self-committed generator less attractive in the SO's dispatch, since the cost of a self-committed unit will seem higher than an identical generator that participates in the centralized unit commitment.

References

Bowring, J., 2006. The PJM market. In: Sioshansi, F.P., Pfaffenberger, W. (Eds.), Electricity Market Reform: An International Experience. Elsevier.

Johnson, R.B., Oren, S.S., Svoboda, A.J., March 1997. Equity and efficiency of unit commitment in competitive electricity markets. Utilities Policy 6, 9–19.

Mas-Colell, A., Whinston, M.D., Green, J.R., 1995. Microeconomic Theory. Oxford University Press, New York, New York.

Moran, A., Skinner, B., 2008. Resource adequacy and infrastructure investment. In: Sioshansi, F.P. (Ed.), Competitive Electricity Markets: Design, Implementation & Performance. Elsevier.

Newbery, D.M., 2006. Electricity liberalization in Britain and the evolution of market design. In: Sioshansi, F.P., Pfaffenberger, W. (Eds.), Electricity Market Reform: An International Experience. Elsevier.

O'Neill, R.P., Sotkiewicz, P.M., Hobbs, B.F., Rothkopf, M.H., Stewart, W.R., 1 July 2005. Efficient market-clearing prices in markets with nonconvexities. European Journal of Operational Research 164, 269–285.

Sioshansi, R., Nicholson, E., November 4–7 2007. Equilibrium offers in unit commitment auctions with nonconvex costs. In: 2007 INFORMS Annual Meeting. Institute for Operations Research and Management Sciences.

Sioshansi, R., O'Neill, R.P., Oren, S.S., May 2008a. Economic consequences of alternative solution methods for centralized unit commitment in day-ahead electricity markets. IEEE Transactions on Power Systems 23, 344–352.

Sioshansi, R., Oren, S.S., O'Neill, R.P., 2008b. The cost of anarchy in self-commitment based electricity markets. In: Sioshansi, F.P. (Ed.), Competitive Electricity Markets: Design, Implementation & Performance. Elsevier.

Streiffert, D., Philbrick, R., Ott, A., 2005. A mixed integer programming solution for market clearing and reliability analysis. In: Power Engineering Society General Meeting, 2005. IEEE, San Francisco, CA.

Wilson, R.B., 1997. Activity rules for a power exchange. In: Proceedings of the 2nd Annual POWER Conference of the University of California Energy Institute. Berkeley, California.

[10] In theory, a Vickery-Clarke-Groves (VCG) mechanism (assuming the SO's unit commitment problem could be solved to complete optimality) would address both the incentive and efficiency issue. This approach has serious practical limitations, however, one of which being the inability to solve a unit commitment problem to optimality in a timely fashion. In fact, the VCG mechanism requires solving the unit commitment problem multiple times—once for each generation firm in the market. Moreover, VCG payments are discriminatory, complicated, and not budget-balanced, making the mechanism an unrealistic option. O'Neill et al. (2005) provide a more complete discussion on the application of VCG in the context of unit commitment and Mas-Colell et al. (1995) discuss mechanism design and the VCG auction more generally.

[11] This may likely be a lower-bound on the annual losses, since more energy would presumably be traded during summer peak periods, and our market data is taken from February.