Quasi-Stochastic Electricity Markets

Jacob Mays*

School of Civil and Environmental Engineering, Cornell University, Ithaca, NY 14850

October 8, 2019

Abstract

Operators of wholesale electricity markets clear sequential, deterministic auctions that attempt to represent an underlying dynamic, stochastic reality, with the additional complicating factor of non-convex production costs. In order to help resolve price formation issues that arise in this context, market designers have proposed operating reserve demand curves that inject an element of stochasticity into deterministic market clearing formulations, altering the procurement of reserves and therefore the pricing of both reserves and energy. The construction of these curves relies on contentious administrative parameters that lack strong theoretical justification. This paper proposes instead to construct curves based on reserve valuations implicit in non-market reliability processes performed by system operators. The proposed strategy promotes greater consistency between commitment decisions and eventual prices, reducing the need for out-of-market uplift payments or enhanced pricing schemes to address non-convexity.

Keywords: Electricity market design, operating reserves, stochastic competitive equilibrium

1 Introduction

In most commodity markets, we conceive of production decisions as being driven by prices. In real-time electricity markets, the need to tightly manage frequency, voltage, and line flows precludes such a decentralized control mechanism. Prices are instead formed by the system operator after the optimal production schedule has been determined. In this context, the engineering challenge is to commit units and dispatch the system in a way that maximizes gains from trade while ensuring reliability. The pricing challenge is to distribute that surplus in a manner consistent with the commitment and dispatch decisions that have been made. Ideally, the prices produced should meet the goals of efficiency, transparency, and simplicity. Unfortunately, electricity markets struggle to achieve all three.

This paper connects three aspects of the price formation problem, each of which has caused significant disagreements among market designers, monitors, regulators, and participants. The

*jacobmays@cornell.edu.
first aspect is that the real-time market is a deterministic representation of an underlying stochastic reality. It is well understood that taking uncertainty into account can improve decisions, whether they are made algorithmically or by human operators exercising engineering judgment. It is equally well understood that if the reason for a decision is not explicitly represented in the market clearing engine, it will not be reflected in prices. Accordingly, while debates about stochastic market clearing have largely been confined to academic outlets [3, 16, 17, 24, 33, 38, 39, 40], the effects of a failure to explicitly represent uncertainty in market models are already felt. The second aspect is the treatment of start-up and no-load costs, which lead to non-convex generator production functions. System operators have proposed a variety of schemes to incorporate these avoidable fixed costs in prices [18]. These schemes, however, can lead to a number of non-intuitive market outcomes [34]. To date, there are no universally agreed-upon metrics with which we can compare the performance of the various schemes on offer. The third aspect is the valuation of reserves through an operating reserve demand curve (ORDC). While system operators have long penalized a shortage of reserves, only recently have they begun to consider the incremental value of holding reserves above a fixed requirement [13, 30]. Some reject the premise altogether: in response to a recent proposal from PJM, for instance, the Independent Market Monitor claims that “ORDCs are an attempt to directly manage energy prices at levels that exceed LMP on a consistent basis,” and instead recommends that PJM “focus on fundamental improvements to energy market efficiency” [22].

An underlying source of disagreement about ORDCs is their reliance on administrative parameters having unclear connection to the traditional goals of market efficiency. The goal of this paper is to put ORDCs on firmer theoretical ground by connecting them to outcomes we would expect in efficient stochastic markets. In doing so, we illustrate why existing proposals may fail to solve the fundamental issues with price formation and sketch a new approach that may better address some of these issues. Deterministic models result in a clearing price under the expected operating conditions, while stochastic models calculate clearing prices under a range of potential operating conditions. ORDCs in principle allow construction of a deterministic model that yields the expected price (or revenue) from the stochastic model. Absent high-quality real-time information on uncertainty distributions, the expected price in the stochastic model is difficult to calculate. However, operator decisions to commit units in non-market reliability processes imply an estimated price in future intervals. We propose to tune ORDCs to match these implied prices, effectively tying
ORDCs to operator decisions rather than relying on administrative parameters. Implementing an ORDC that underestimates this expected price will tend to result in units failing to cover their avoidable fixed costs. An ORDC that overestimates this value, on the other hand, may lead to increased self-commitment of inefficient units or to lost opportunity costs for non-committed units. One implication of the analysis is that if ORDCs are successful, the use of make-whole payments and enhanced pricing to address non-convexity may no longer be justified.

Section 2 provides background on recent pricing debates in centrally committed US markets. We present the ideal case of stochastic competitive equilibrium in Section 3 and introduce an example in Section 4 to help guide the remainder of the discussion. Section 5 identifies pricing issues caused by using deterministic instead of stochastic models for market clearing, and Section 6 addresses the potential of current ORDC formulations to resolve these issues. We outline our proposed approach in Section 7 and conclude in Section 8.

2 Price Formation

A curious feature of US wholesale electricity markets is that system operators, on behalf of the demand side of the market, regularly commit to buying energy at above-market prices. Many believe that the net effect of these commitments is to suppress energy market prices, allowing the demand side to procure the balance of its energy at slightly lower cost. Were loads entering into these contracts themselves, such a pattern could be seen as buyer-side market manipulation. Instead, uplift payments are widely agreed to be a necessary component of electricity markets, going by the names of bid cost recovery, revenue sufficiency guarantees, and make-whole payments in different regions. In the final accounting, load does not benefit from this price suppression: insufficient energy market revenue leads to poor incentives for investment, contributing to the need for supplementary payments for capacity [15].

Explanations for side payments start from the observation that wholesale electricity markets lack uniform clearing prices due to the non-convex production costs characteristic of most generation technologies [12, 27]. The total revenue earned by a generator may be insufficient to cover the avoidable fixed costs associated with starting up and operating at its minimum output. Despite this condition, the decision to commit the unit can still be efficient in the sense of maximizing expected surplus. To encourage their participation in the market despite this risk of shortfall, system oper-
ators have authorized a variety of uplift payments that guarantee cost recovery [6]. Dissatisfaction with the discriminatory and non-transparent nature of these side payments, as well as concern about potential long-term consequences of price suppression, has led many to propose alternate price formation scheme to reduce or eliminate these side payments [18]. Strategies that have been or soon will be implemented include variants of Fast-Start Pricing in ISO-NE [14], PJM [9], and NYISO [8], as well as Extended Locational Marginal Pricing in MISO [37]. Throughout this paper, we use the basic LMP pricing scheme.

Non-convexity itself, however, is only a partial explanation for uplift. In the aggregate, evidence in [21] suggests that an efficient long-term capacity mix is supported by a strict implementation of LMP, without any side payments or enhanced price formation schemes. Complicating this result, it has now been observed in several test systems [4, 21, 35] that the LMP can be quite sensitive to the particular commitment solution chosen by the optimization solver. Any bias toward holding additional units online can therefore result in lower prices on average. While difficult to model in a rigorous way, upward load adjustments are made on a regular basis (see Attachment E within [30] for a description of these actions). Simulations on the PJM system for 2018 result in uplift of $30M after re-running the unit commitment software using actual demand, versus the $200M actually incurred in 2018 [23]. A direct comparison would be unfair, since perfect knowledge of future demand is impossible. Nevertheless, the difference demonstrates the degree to which uplift is driven not purely by non-convexity but also by the exigencies of decision making under uncertainty.

Independent System Operators (ISOs) and Regional Transmission Organizations (RTOs) in the United States fulfill a dual role as reliability coordinators and market operators for the transmission grid. This paper suggests that the uplift gap between the $30M attributable directly to non-convexity and the $200M actually observed represents a failure to fully reflect the needs of the reliability function in pricing models. Price formation is driven by a three-step process involving commitment, dispatch, and pricing [34]. In the commitment stage, operators solve a series of mixed-integer programs (MIPs) to determine the optimal start-up and shut-down schedule of units on the system. Importantly, the day-ahead market is just one of these MIPs. It is accompanied by a series of reliability processes in which the bid-in load is replaced by an operator demand forecast. The participation of virtual bidders imparts some features of stochastic models to the day-ahead market, allowing prices to converge between day-ahead and real-time [20]. This feature of day-
ahead markets is lacking in subsequent reliability processes, making it more likely for operators to commit units that will not be profitable in real time.

Apart from handling non-convexity, a separate trend in price formation addresses the value of maintaining reserves beyond the required minimum on the system. In contrast to typical forward markets, which in the electricity setting would estimate the expected incremental cost over a range of possible operating conditions, current “real-time” auctions calculate the incremental cost under the expected operating conditions. The ORDC construct attempts to rectify this issue by inserting the expected cost of deploying reserves in the objective function of the pricing problem described above [13, 30]. While not strictly necessary, ORDCs could also be included in commitment and dispatch models.

In theory, two key inputs govern the shape of current ORDCs. The first is the cost of recourse actions that will be taken if the system falls below a mandated level of contingency reserves. In ERCOT, which has no capacity market, this is the value of lost load, currently assessed at $9,000/MWh since the assumed recourse action is involuntary load shedding. PJM, in which the capacity market is intended to virtually guarantee the availability of supply, proposes a lower value of $2,000/MWh. The second is the probability that the assumed recourse action will take place given a certain level of reserves. Determining these probabilities is a challenging problem. The basic strategy in ERCOT and PJM is to use the distribution of historical forecast errors over a chosen lookahead period. The length of this lookahead period is a largely unjustified parameter in both markets. The most natural choice is a lookahead period equal to the frequency at which electricity is repriced. The argument in [30] appeals to the future value of unused reserves at the end of a pricing interval to defend the choice of a much longer lookahead period. In theory this is not required, since any residual value of reserves should be incorporated in the new price of electricity at the time of repricing. However, including uncertainty over a longer horizon might be warranted if there are intertemporal constraints omitted from current single-period dispatch and pricing models. Incorporating the residual value could therefore be interpreted as a proxy for a better representation of, e.g., the opportunity cost of energy-limited resources. A more explicit representation of these constraints, such as that embedded in the ramping product implemented in MISO [36], would offer a stronger justification for cross-temporal pricing effects.

Reserve penalty factors and probabilities are also present in the general formulation in this pa-
per, perhaps suggesting that the resulting reserve prices will be no less administratively determined than current proposals. The presence of these parameters, however, belies an important philosophical difference. The approach here enforces a clear separation between the reliability and market functions fulfilled by system operators: the role of the reliability function is to maximize expected surplus, and the role of the pricing function is to produce prices consistent with the reliability function. This contrasts with current ORDC proposals, in which the price-setting function can overrule the reliability function, potentially leading to inconsistent prices or inefficient commitments. That said, to the extent penalty factors and probabilities are used to make reliability decisions, they will be expressed in prices. This paper does not address the problem of determining these parameters, instead focusing on the problem of producing consistent prices.

In practice, the difficulty of linking reserve levels in one period to the probability of recourse actions in future periods makes it a weak foundation for pricing logic. It is therefore necessary to seek other ways to validate choices made in the construction of administrative ORDCs. In ERCOT, the choice of metric is clear: since the primary goal of the ORDC is to guarantee resource adequacy, it is natural to assess it according to the reserve margin that it can support in equilibrium [26]. The resulting curve has only loose connection to its theoretical mooring in the loss of load probability (see, e.g., the administrative parameters described in [5]). In other markets resource adequacy is guaranteed by other means, necessitating a different metric. In the following sections, we propose such a metric, bid cost recovery in expectation, by connecting ORDCs with the ideal of stochastic competitive equilibrium.

3 Stochastic Ideal

An underlying assumption of the paper is that implementation of a fully stochastic market design will be impossible, at least in the foreseeable future. That said, the ideal of stochastic competitive equilibrium provides a target that we might hope to approximate with ORDCs. Here we develop a simple model of ISO/RTO operations to demonstrate this ideal, as well as the key property of bid cost recovery in expectation. Variants of these results appear in several related works [3, 16, 38, 39, 40]; for completeness, as well as to introduce subsequent results, we present them here as well.

While the results here serve as a theoretical foundation, the remainder of the article is structured
such that readers who are more interested in policy implications will be able to follow the core arguments of the paper without reference to this section.

3.1 Unit Commitment

Significant progress has been made on large-scale unit commitment models that incorporate uncertainty using stochastic [2, 10] or robust [1, 19] optimization, with performance substantially enhanced by tight formulations of the underlying technical constraints [25, 28]. If uncertainty can be appropriately characterized, these models offer significant performance improvements over current deterministic implementations. As a benchmark for the best possible decision process, we formulate a stochastic program with perfect knowledge of the distribution of the random variables. Throughout the paper, commitment is assumed to be a non-market reliability process, in which no virtual bidders participate and from which no financial positions result.

The commitment problem is to determine the commitment status of each generator $g \in \mathcal{G}$ in each time period $t \in \mathcal{T}$, given a finite number of dispatch scenarios $s \in \mathcal{S}$ characterizing uncertain demand and generator availability. If a generator is brought online, it must remain on for a minimum up time $UT_g$. Starting generator $g \in \mathcal{G}$ incurs a start-up cost $C_{SU}^g$, while keeping it online for one time period entails a no-load cost $C_{NL}^g$. For each load $l \in \mathcal{L}$, demand served up to the quantity $D_{slt}^+$ is valued at $V_{D}^l$. The system is willing to spend $V^R$ for reserves up to a quantity of $R^+$; stated differently, $V^R$ is the cost of recourse actions that will be taken to maintain security in the event of a reserve shortfall. At time $t$, the conditional probability of scenario $s$ is $\phi(s|\mathcal{F}_t)$, where $\mathcal{F}_t$ represents all information available immediately before the beginning of period $t$. Any scenarios with $\phi(s|\mathcal{F}_t) = 0$ are removed from the model. We assume initial unit commitment decisions occur before any new information is available, and denote the unconditional probability of scenario $s \in \mathcal{S}$ as $\phi_s$. Binary variables for start-up $v_{gt}$, commitment status $u_{gt}$, and shutdown $q_{gt}$ for generator $g \in \mathcal{G}$ in time period $t \in \mathcal{T}$ are shared across all scenarios, while power generation $p_{sgt}$ and provision of reserves $r_{sgt}$ for each generator $g \in \mathcal{G}$ and cleared demand $d_{slt}$ for each load $l \in \mathcal{L}$ can vary by scenario. The variable $o_{st}$ represents the total quantity of reserves procured by the system operator in scenario $s \in \mathcal{S}$ and time period $t \in \mathcal{T}$. Online generators are constrained to operate above a minimum operating level $P_{sgt}^-$, while supply of energy and reserves cannot exceed maximum output $P_{sgt}^+$. Provision of reserves is assumed to be costless; incorporating offer prices
would require small changes. Inclusion of minimum down times and ramping constraints would similarly require only small changes to the model. We assume a single reserve product and do not include network or security constraints.

As indicated by the subscript \( s \), uncertainty arises from the minimum and maximum availability of generators, represented by the random variables \( P^{-\text{sgt}} \) and \( P^{+\text{sgt}} \), as well as demand \( D^{\text{slt}} \). In the case of typical wind and solar generators, \( P^{-\text{sgt}} = 0 \) while \( P^{+\text{sgt}} \) can fluctuate between 0 and the rated capacity. For thermal generators, \( P^{-\text{sgt}} = P^{+\text{sgt}} = 0 \) represents a failure. The stochastic unit commitment problem can be stated as

\[
\text{(SUC)}: \quad \max_{u,v,q,p,r,d} \sum_{s \in S} \sum_{l \in L} \sum_{t \in T} \phi_s V^D d_{slt} + \sum_{s \in S} \sum_{t \in T} \phi_s V^R o_{st} - \sum_{s \in S} \sum_{g \in G} \sum_{t \in T} \phi_s C^E g p_{sgt} - \sum_{g \in G} \sum_{t \in T} (C^{NL} g u_{gt} + C^{SU} g v_{gt}) \tag{1a}
\]

subject to

\[
\phi_s \left( \sum_{l \in L} d_{slt} - \sum_{g \in G} p_{sgt} \right) = 0 \quad \forall s \in S, t \in T \quad (\lambda_{st}) \tag{1b}
\]

\[
\phi_s \left( o_{st} - \sum_{g \in G} r_{gst} \right) = 0 \quad \forall s \in S, t \in T \quad (\mu_{st}) \tag{1c}
\]

\[
d_{slt} \leq D^{+\text{slt}} \quad \forall s \in S, l \in L, t \in T \quad (\alpha^{D}_{slt}) \tag{1d}
\]

\[
o_{st} \leq R^t \quad \forall s \in S, t \in T \quad (\alpha^{R}_{st}) \tag{1e}
\]

\[
P^{-\text{sgt}} u_{gt} \leq p_{sgt} \quad \forall s \in S, g \in G, t \in T \quad (\alpha_{sgt}^{P+}) \tag{1f}
\]

\[
p_{sgt} + r_{sgt} \leq P^{+\text{sgt}} u_{gt} \quad \forall s \in S, g \in G, t \in T \quad (\alpha_{sgt}^{P+}) \tag{1g}
\]

\[
v_{gt} + u_{g,t-1} - u_{gt} - q_{gt} = 0 \quad \forall g \in G, t \in T \quad (\gamma_{gt}) \tag{1h}
\]

\[
\sum_{t' = \max(1,t-UT_g+1)}^t v_{gt'} - u_{gt} \leq 0 \quad \forall g \in G, t \in T \quad (\delta_{gt}) \tag{1i}
\]

Constraints in this MIP are used in several subsequent linear programs. For convenience, we include
Greek letters to the right of each constraint denoting dual variables that will be used in these linear programs. The objective function in (1a) calculates the value of serving load and supplying reserves minus the three-part cost of meeting that load. Power balance is enforced in constraint (1b), while constraint (1c) calculates the total reserves supplied. These equations are scaled by probability $\phi_s$ in order to produce unscaled prices $\lambda_{st}$ and $\mu_{st}$ in each scenario. Combined with the objective function, constraint (1d) models load curtailment, while constraints (1e) and (1f) model a reserve shortfall. Constraints (1g) and (1h) represent the minimum and maximum output of generators, which depends on their commitment status $u_{gt}$. Equation (1i) and (1j) ensure consistency between start-up, shut-down, and commitment variables given generator minimum run times. In the first time period, the variable $u_{g,t-1}$ in Eq. (1h) is replaced by the current commitment status of each generator; similar substitutions can be made in Eq. (1i) if generators have not already met their minimum run time.

### 3.2 Relaxation and Duality

To establish formal results, we will instead work with a convex relaxation of model (SUC). The relaxed stochastic unit commitment problem is stated as

\[
(RSUC) : \begin{align*}
\text{maximize} & \quad (1a) \\
\text{subject to} & \quad (1b)-(1l) \quad \text{for } u,v,q,p,r,d,o \\
& \quad u_{gt} \leq 1 \forall g \in \mathcal{G}, t \in \mathcal{T} \quad (\beta^ON) \\
& \quad v_{gt} \leq 1 \forall g \in \mathcal{G}, t \in \mathcal{T} \quad (\beta^SU) \\
& \quad u_{gt} \leq 1 \forall g \in \mathcal{G}, t \in \mathcal{T} \quad (\beta^SD) \\
& \quad u_{gt}, v_{gt}, q_{gt} \geq 0 \forall g \in \mathcal{G}, t \in \mathcal{T}. \quad (2f)
\end{align*}
\]

The dual to model (RSUC) is stated as follows:

\[
(DRSUC) : \begin{align*}
\text{minimize} & \quad \sum_{g \in \mathcal{G}} \sum_{t \in \mathcal{T}} (\beta^ON + \beta^SU + \beta^SD) \\
& \quad + \sum_{s \in \mathcal{S}} \sum_{l \in \mathcal{L}} \sum_{t \in \mathcal{T}} \alpha^{D^+}_{slt} D_{slt}^+ \\
& \quad - \sum_{s \in \mathcal{S}} \sum_{t \in \mathcal{T}} \alpha^R_{st} R_t^+ \\
\text{subject to} & \quad \forall s \in \mathcal{S}, g \in \mathcal{G}, t \in \mathcal{T} \quad (3a)
\end{align*}
\]

\[
-\phi_s \lambda_s - \alpha^P_{sgt} + \alpha^P_{sgt} \geq -\phi_s C_g^{EN} \quad \forall s \in \mathcal{S}, g \in \mathcal{G}, t \in \mathcal{T} \quad (p_{sgt}) \quad (3b)
\]
\[
\begin{align*}
-\phi_s \mu_{st} + \alpha^{P+}_{sgt} & \geq 0 \quad \forall s \in S, g \in G, t \in T \quad (r_{sgt}) \\
\phi_s \mu_{st} + \alpha^R_{st} & \geq \phi_s V^R \quad \forall s \in S, t \in T \quad (o_{st}) \\
\phi_s \lambda_{st} + \alpha^{D+}_{slt} & \geq \phi_s V^D \quad \forall s \in S, l \in L, t \in T \quad (d_{slt}) \\
\sum_{s \in S} (P^-_{sgt} \alpha^P_{sgt} - P^+_{sgt} \alpha^P_{sgt}) & \geq 0 \\
+ & \sum_{s \in S} \sum_{t \in T} \phi_s p_{sgt} (\lambda_{st} - C^E_{g}) \\
& + \sum_{s \in S} \sum_{t \in T} \phi_s r_{sgt} \mu_{st} \\
& - \sum_{t \in T} (C^N_{g} u_{gt} + C^S_{g} v_{gt}) \\
\gamma_{gt} + & \sum_{t' = t} \alpha^D_{slt} \geq 0 \quad \forall s \in S, l \in L, t \in T \\
\alpha^P_{sgt}, \alpha^P_{sgt} & \geq 0 \quad \forall s \in S, g \in G, t \in T \\
\alpha^R_{st} & \geq 0 \quad \forall s \in S, t \in T \\
\beta^O_{g}, \beta^S_{gt}, \beta^D_{gt}, \gamma_{gt} & \geq 0 \quad \forall g \in G, t \in T.
\end{align*}
\]

3.3 Market Participants

Model (RSUC) represents the problem faced by the wholesale market operator. A decentralized perspective focuses on the problems faced by the market participants. Given prices \( \lambda_{st} \) and \( \mu_{st} \) for energy and reserves, the relaxed problem faced by generator \( g \in G \) is stated as

\[
(RGEN)_{g} : \\
\begin{align*}
\text{maximize} & \quad \sum_{s \in S} \sum_{t \in T} \phi_s p_{sgt} (\lambda_{st} - C^E_{g}) \\
& + \sum_{s \in S} \sum_{t \in T} \phi_s r_{sgt} \mu_{st} \\
& - \sum_{t \in T} (C^N_{g} u_{gt} + C^S_{g} v_{gt}) \\
\text{subject to} & \quad P^-_{sgt} u_{gt} \leq p_{sgt} \quad \forall s \in S, t \in T \\
& \quad p_{sgt} + r_{sgt} \leq P^+_{sgt} u_{gt} \quad \forall s \in S, g \in G, t \in T \\
& \quad v_{gt} + u_{g,t-1} - u_{gt} - q_{gt} = 0 \quad \forall t \in T \\
& \quad \sum_{t' = \max(1,t-UT_g+1)} v_{gt'} - u_{gt} \leq 0 \quad \forall t \in T \\
& \quad u_{gt} \leq 1 \quad \forall t \in T \\
& \quad v_{gt} \leq 1 \quad \forall t \in T \\
& \quad u_{gt} \leq 1 \quad \forall t \in T \\
& \quad p_{sgt}, r_{sgt} \geq 0 \quad \forall s \in S, t \in T.
\end{align*}
\]
Generators choose a production schedule that maximizes expected profits while following all of the generator-specific constraints described in model \((RSUC)\), but have no responsibility for system-level constraints. The simpler problem for loads is stated as follows:

\[
(LOAD)_l : \quad \begin{align*}
\text{maximize} & \quad \sum_{s \in \mathcal{S}} \sum_{t \in \mathcal{T}} \phi_s d_{slt} (V_{lt}^D - \lambda_{st}) \\
\text{subject to} & \quad d_{slt} \leq D_{slt}^+ \quad \forall s \in \mathcal{S}, t \in \mathcal{T} \quad (\alpha_{slt}^D) \\
& \quad d_{slt} \geq 0 \quad \forall s \in \mathcal{S}, t \in \mathcal{T} \quad (\beta_{slt}) 
\end{align*}
\]

### 3.4 Stochastic Competitive Equilibrium

Using these problem definitions, we define a stochastic competitive equilibrium for the market.

**Definition 1.** A stochastic competitive equilibrium is a set of prices \((\lambda_{st}, \mu_{st}) \forall s \in \mathcal{S}, t \in \mathcal{T}\), generator schedules \((u_{gt}, v_{gt}, q_{gt}) \forall g \in \mathcal{G}, t \in \mathcal{T}\) and production quantities \((p_{sgt}, r_{sgt}) \forall s \in \mathcal{S}, g \in \mathcal{G}, t \in \mathcal{T}\), load consumption quantities \(d_{slt} \forall s \in \mathcal{S}, l \in \mathcal{L}, t \in \mathcal{T}\), and operator reserve procurements \(o_{st} \forall s \in \mathcal{S}, t \in \mathcal{T}\) such that:

1. the generator schedules and production quantities solve \((RGEN)_g \forall g \in \mathcal{G}\) and the load consumption quantities solve \((LOAD)_l \forall l \in \mathcal{L}\) at the given prices, and

2. the market clears in every scenario, i.e.,

\[
\sum_{l \in \mathcal{L}} d_{slt} = \sum_{g \in \mathcal{G}} p_{sgt} \quad \forall s \in \mathcal{S}, t \in \mathcal{T}
\]

and

\[
o_{st} = \sum_{g \in \mathcal{G}} r_{gst} \quad \forall s \in \mathcal{S}, t \in \mathcal{T}.
\]

In other words, a stochastic competitive equilibrium occurs when all market participants maximize their individual expected surplus and system constraints are satisfied. We now show that this is the case for an optimal solution to the system operator’s problem.

**Theorem 1.** Assume all market participants \(g \in \mathcal{G}\) and \(l \in \mathcal{L}\) agree on probabilities \(\phi_s \forall s \in \mathcal{S}\). Let \((u^*, v^*, q^*, p^*, r^*, d^*, o^*)\) be an optimal solution to \((RSUC)\) and let \((\lambda^*, \mu^*, \alpha^*, \beta^*, \gamma^*, \delta^*)\) be an optimal solution to \((DRSUC)\). Then the prices \((\lambda_{st}^*, \mu_{st}^*) \forall s \in \mathcal{S}, t \in \mathcal{T}\), generator schedules \((u_{gt}^*, v_{gt}^*, q_{gt}^*) \forall g \in \mathcal{G}, t \in \mathcal{T}\), generator production quantities \((p_{sgt}^*, r_{sgt}^*) \forall s \in \mathcal{S}, g \in \mathcal{G}, t \in \mathcal{T}\),
load consumption quantities \( d^*_s, l \in L, t \in T \), and operator reserve procurements \( o^*_s, s \in S, t \in T \) represent a stochastic competitive equilibrium.

**Proof.** The theorem follows from duality theory. For each generator \( g \in G \), the dual problem to \((RGEN)_g\) can be stated as follows:

\[
(DRGEN)_g: \quad \min_{\alpha_g, \beta_g, \gamma_g, \delta_g} \sum_{t \in T} (\beta_{gt}^{ON} + \beta_{gt}^{SU} + \beta_{gt}^{SD})
\]

subject to

\[
\begin{align*}
-\alpha_g P^- s &+ \alpha_g P^+ s \geq \phi_s (\lambda_{st} - C^{EN}_g) \quad \forall s \in S, t \in T \quad (p_{sgt}) \\
\alpha_g P^+ s &\geq \phi_s \mu_{st} \quad \forall s \in S, t \in T \quad (r_{sgt}) \\
\sum_{s \in S} (P^- s \alpha_{sgt}^P - P^+ s \alpha_{sgt}^P) &+ \gamma_{gt+1} - \gamma_{gt} + \beta_{gt}^{ON} - \delta_{gt} \geq -C^{NL}_g \quad \forall t \in T \quad (u_{gt}) \\
\gamma_{gt} + \sum_{t' = t}^{\min(t + T_g - 1)} &\delta_{gt} + \beta_{gt}^{SU} \geq -C^{SU}_g \quad \forall t \in T \quad (v_{gt}) \\
-\gamma_{gt} + \beta_{gt}^{SD} &\geq 0 \quad \forall t \in T \quad (q_{gt}) \\
\alpha_{sgt}^P, \alpha_{sgt}^P &\geq 0 \quad \forall s \in S, t \in T \quad (o_{sgt}) \\
\beta_{gt}^{ON}, \beta_{gt}^{SU}, \beta_{gt}^{SD}, \gamma_{gt} &\geq 0 \quad \forall t \in T. \quad (g_{ht})
\end{align*}
\]

Similarly, for each load \( l \in L \), the dual problem to \((LOAD)_l\) can be stated as follows:

\[
(DLOAD)_l: \quad \min_{\alpha_l} \sum_{s \in S} \sum_{t \in T} \alpha^D_{slt} D^+_{slt} \quad (7a)
\]

subject to

\[
\begin{align*}
\alpha^D_{slt} &\geq \phi_s (V^D_l - \lambda_{st}) \quad \forall s \in S, t \in T \quad (d_{slt}) \\
\alpha^D_{slt} &\geq 0 \quad \forall s \in S, t \in T. \quad (7c)
\end{align*}
\]

Since \((u^*, v^*, q^*, p^*, r^*, d^*, o^*)\) is an optimal solution to \((RSUC)\) and \((\lambda^*, \mu^*, \alpha^*, \beta^*, \gamma^*, \delta^*)\) is an optimal solution to \((DRSUC)\), the complementary slackness conditions hold. While we omit these conditions for brevity, the variables corresponding to each constraint in \((RSUC)\), \((DRSUC)\), \((RGEN)_g\), \((DRGEN)_g\), \((LOAD)_l\), and \((DLOAD)_l\) are shown. It can be readily seen that the complementary slackness conditions for \((RSUC)\) and \((DRSUC)\) are a superset of those for the individual market participant problems. Accordingly, the generator schedules \((u^*_g, v^*_g, q^*_g)\) and production quantities \((p^*_g, r^*_g)\) solve \((RGEN)_g \forall g \in G\), while the load consumption quantities \(d^*_l\) solve \((LOAD)_l \forall l \in L\) given prices \((\lambda^*, \mu^*)\). Furthermore, since \((u^*, v^*, q^*, p^*, r^*, d^*, o^*)\) is an optimal so-
solution to \((RSUC)\), the market clears. Thus, the prices \((\lambda^*_s, \mu^*_s) \forall s \in S, t \in T\), generator schedules \((u^*_{gt}, v^*_{gt}, q^*_{gt}) \forall g \in G, t \in T\) and production quantities \((p^*_{sgt}, r^*_{sgt}) \forall s \in S, g \in G, t \in T\), load consumption quantities \(d^*_{slt} \forall s \in S, l \in L, t \in T\), and operator reserve procurements \(o^*_{st} \forall s \in S, t \in T\) represent a stochastic competitive equilibrium.

A key assumption for stochastic competitive equilibrium is that market participants agree on the probability \(\phi_s\) that each scenario \(s \in S\) will arise. We return to the consequences of this assumption in Section 7.

3.5 Bid Cost Recovery in Expectation

One implication of Theorem 1 is the property of bid cost recovery in expectation, i.e., a guarantee that the expected profit of every generator is at least zero in the optimal solution to the relaxed problem. Scenario-specific profit for generator \(g \in G\) can be calculated as

\[
\pi_{sg} = \sum_{t \in T} \left( p_{sgt}(\lambda^*_s - C^E_{g}) + r_{sgt}\mu^*_s - C^NL_{gt}u_{gt} - C^SU_{gt} v_{gt} \right)
\]

(8)

Corollary 1. Let \((u^*, v^*, q^*, p^*, r^*, d^*, o^*)\) be an optimal solution to \((RSUC)\) and \((\lambda^*, \mu^*, \alpha^*, \beta^*, \gamma^*, \delta^*)\) be an optimal solution to \((DRSUC)\). Then, given probabilities \(\phi_s \forall s \in S\), expected profit \(\mathbb{E}[\pi_{sg}] \geq 0 \forall g \in G\).

Proof. By Theorem 1, expected profit \(\mathbb{E}[\pi_{sg}]\) can be calculated as the objective function value for \((RGEN)_{g}\). By strong duality, this value can also be calculated through the dual problem \((DRGEN)_{g}\) as

\[
\mathbb{E}[\pi_{sg}] = \sum_{t \in T} \left( \beta^O_{gt} + \beta^SU_{gt} + \beta^SD_{gt} \right).
\]

(9)

Since \(\beta^O_{gt}, \beta^SU_{gt}, \beta^SD_{gt} \geq 0 \forall g \in G, t \in T\), we can conclude that \(\mathbb{E}[\pi_{sg}] \geq 0\).

3.6 Profitability of Marginal Generators

A second implication of Theorem 1 is that expected profit is exactly zero for what we might consider marginal generators.

Corollary 2. Let \((u^*, v^*, q^*, p^*, r^*, d^*, o^*)\) be an optimal solution to \((RSUC)\) and \((\lambda^*, \mu^*, \alpha^*, \beta^*, \gamma^*, \delta^*)\) be an optimal solution to \((DRSUC)\). For some generator \(g \in G\), let \(u_{gt}^*, v_{gt}^*, q_{gt}^* < 1 \forall t \in T\). Then
expected profit \( \mathbb{E}[\pi_{sg}] = 0 \).

Proof. In the optimal solution, \( u^*_g, v^*_g, q^*_g < 1 \forall t \in \mathcal{T} \). By complementary slackness, we know that \( \beta^{ON}_{gt}(1-u_{gt}) = 0 \), \( \beta^{SU}_{gt}(1-v_{gt}) = 0 \), and \( \beta^{SD}_{gt}(1-q_{gt}) = 0 \) \( \forall t \in \mathcal{T} \). Therefore, \( \beta^{ON}_{gt} = 0 \), \( \beta^{SU}_{gt} = 0 \), and \( \beta^{SD}_{gt} = 0 \) \( \forall t \in \mathcal{T} \). As in Corollary 1, expected profit for generator \( g \) can then be calculated as

\[
\mathbb{E}[\pi_{sg}] = \sum_{t \in \mathcal{T}} (\beta^{ON}_{gt} + \beta^{SU}_{gt} + \beta^{SD}_{gt}) = 0.
\] (10)

We note that the generators considered in Corollary 2 are not necessarily marginal in the sense of providing the incremental unit of energy at a given point of time. Instead, their entrance into the optimal commitment solution is marginal in the relaxed problem. In the binary context, these generators cannot be partially on or off. With this non-convexity, expected profits will not precisely equal zero. There is no theoretical guidance predicting whether profits will increase or decrease when moving to the optimal integer solution. Absent such guidance, the expected value of these shifts in profit may also be zero, informally preserving the property of bid cost recovery in expectation.

4 Example System

In this section, we develop a simple example with a single load, single time period, and single source of uncertainty to illustrate the pricing issues caused by disconnects between commitment and dispatch in the presence of uncertainty. The lone source of uncertainty is a wind generator. The availability of wind is represented by the random variable \( W \), which is uniformly distributed between 0 and 100, i.e., \( W \sim \mathcal{U}(0,100) \). Where appropriate, we also use \( W \) to indicate a realization of this random variable. For ease of exposition, we construct a simplified version of the stochastic unit commitment problem in model (1). Along these lines, we assume wind is costless and does not provide reserves, remove it from the set of generators \( \mathcal{G} \), and define a new decision variable \( w \) indicating the quantity of wind dispatched. We retain variable and parameter definitions but drop the unneeded subscripts \( s, l, \) and \( t \) from model (1). Demand \( D^+ = 200 \). The reserve requirement \( R^+ \) is set at \( 20 - \varepsilon \), where \( \varepsilon \) is a small positive number; this quantity is perturbed to avoid multiple
optimal dual solutions, but we use $R^+ = 20$ in calculations of probabilities and revenues. The unit penalty for a reserve shortfall $V^R = 950$, while the value of demand $V^D = 10,000$. In addition to wind, the system is served by 101 thermal generators. The first has capacity $P^+_g = 120$ and can operate flexibly, i.e., $P^-_g = 0$. Its no-load cost $C^{NL}_g = 0$, while its energy cost $C^{EN}_g = 50$. The remaining 100 generators are all block loaded units of size 1; i.e., $P^-_g = P^+_g = 1$. The units are conveniently arranged by no-load costs increasing from 51 to 150 in increments of 1, while energy costs $C^{EN}_g$ for each are zero. There are no start-up costs, i.e., $C^{SU}_g = 0 \forall g \in G$, and minimum up times are omitted. We refer to the largest generator as generator 0, and the remaining 100 in order according their cost (e.g., the generator with $C^{NL}_g = 70$ is referred to as generator 20).

The stochastic unit commitment problem for this example can be stated as

$$\max_{u: \ u_g \in \{0,1\}} - \sum_{g \in G} C^{NL}_g u_g + \mathbb{E}[H(u; W)],$$

where $u = (u_g)_{g \in G}$ and $\mathbb{E}[H(u; W)]$ represents the expected surplus from dispatching the system given uncertain wind availability. The subproblem for a specific realization of wind $W$ is stated as follows:

$$H(u; W) = \max_{p, r, d, o, w} V^D d + V^R o - \sum_{g \in G} C^{EN}_g p_g$$

subject to

$$d - w - \sum_{g \in G} p_g = 0 \quad (12b)$$
$$o - \sum_{g \in G} r_g = 0 \quad (12c)$$
$$d \leq D^+ \quad (12d)$$
$$o \leq R^+ \quad (12e)$$
$$w \leq W \quad (12f)$$
$$P^-_g u_g \leq p_g \quad \forall g \in G \quad (12g)$$
$$p_g + r_g \leq P^+_g u_g \quad \forall g \in G \quad (12h)$$
$$d, o, w \geq 0 \quad (12i)$$
$$p_g, r_g \geq 0 \quad \forall g \in G \quad (12j)$$

4.1 Optimal Solution

The system is constructed such that it can be solved exactly with a simple rule. Since $C^{NL}_0 = 0$, generator 0 will be committed. Moreover, only generator 0 will supply reserves, since all remaining
thermal units will be producing energy equal to their maximum output and it will always be optimal to dispatch wind at its full availability. We commit the remaining units in order until the no-load cost associated with model (11) exceeds the expected improvement of the dispatch in model (12). Let $u^n$ indicate the solution to model (11) in which the cheapest $n$ block-loaded units are committed. Then, we can commit units until

$$C_{n+1}^{NL} \geq \mathbb{E}[H(u^{n+1}; W)] - \mathbb{E}[H(u^n; W)].$$

(13)

The optimal solution is to commit 90 block loaded units. This leads to a ten percent chance of a reserve shortfall, which will occur if wind availability turns out to be less than 10, but no chance that load will need to be curtailed. Though non-convexity and the increasing cost of the block-loaded units complicates the analysis, the example is similar to the classic newsvendor problem. If committing generator 90 at a cost of $140 allows us to avoid a unit of reserve shortfall, the system has a benefit of $1000: $50 from reducing the energy output of generator 0 plus $950 from supplying an additional unit of reserve. If there is no shortfall, however, the system only saves $50 in production costs. Because the overage cost of $90 is much smaller than the underage cost of $860, the optimal solution is to commit far more units than would be required if we saw average wind availability of 50.

4.2 Dispatch and Pricing

As the actual operating interval approaches, the system operator solves the economic dispatch problem with the commitment decisions fixed. In principle, the uncertainty in the system may evolve between the time of commitment and the time of dispatch. Since the generators in our example system have no constraints on their flexibility apart from their commitment status, the distinction does not matter for dispatch. However, the timing of uncertainty realization has important implications for pricing.

We first consider the ideal case, in which uncertainty is completely resolved before the time of dispatch. In practice, this would require clearing markets in real time at a frequency far higher than the current five minutes. In this case, the optimal dispatch is found by solving a version of model (12) with the realized value of $W$ and the chosen commitment solution. Emphasizing that
instead of a subproblem this is a separate, standalone problem with commitment variables fixed to the solution \( u = \hat{u} \), we formulate the economic dispatch problem as

\[
(ED - VERT) : \begin{align*}
\text{maximize} & \quad p, r, d, o, w \\
\text{subject to} & \quad d - w - \sum_{g \in G} p_g = 0 \\
& \quad o - \sum_{g \in G} r_g = 0 \\
& \quad (12d) - (12j) \\
& \quad u_g = \hat{u}_g \quad \forall g \in G. 
\end{align*}
\] (14a)

The name \((ED - VERT)\) is chosen in light of the vertical demand curve for reserves, which we will later contrast with the sloped demand curve created by an ORDC. We write the outputs of this model as functions of the random variable, further conditioned by the commitment solution \( \hat{u} \) chosen prior to dispatch. The LMP \( \lambda(W; \hat{u}) \) is the dual variable associated with the power balance constraint, Eq. \((14b)\), while the price of reserves \( \mu(W; \hat{u}) \) is given by the dual variable associated with the reserve constraint, Eq. \((14c)\). Assume the operator has chosen the optimal commitment solution, i.e., \( \hat{u} = u^{90} \). As shown in Table 1 there are two possible pricing outcomes: if wind \( W < 10\), the reserve shortfall leads to prices \( \lambda(W; u^{90}) = \$1,000 \) for power and \( \mu(W; u^{90}) = \$950 \) for reserves. If wind \( W \geq 10\), generator 0 sets the price at \( \lambda(W; u^{90}) = \$50 \).

**Table 1: Dispatch and prices given optimal commitment**

<table>
<thead>
<tr>
<th>Range</th>
<th>Probability</th>
<th>Wind Availability</th>
<th>( p^*_g(W; u^{90}) )</th>
<th>( r^*_g(W; u^{90}) )</th>
<th>( \lambda(W; u^{90}) )</th>
<th>( \mu(W; u^{90}) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.1</td>
<td>( 0 \leq W &lt; 10 )</td>
<td>110 - W</td>
<td>10 + W</td>
<td>$1,000</td>
<td>$950</td>
</tr>
<tr>
<td>2</td>
<td>0.9</td>
<td>( 10 \leq W \leq 100 )</td>
<td>110 - W</td>
<td>20 - \varepsilon</td>
<td>$50</td>
<td>$0</td>
</tr>
</tbody>
</table>

The profitability of a given generator is similarly a function of the random variable \( W \). In this single-period example, we can compute profitability \( \pi_g(W; \hat{u}) \) as

\[
\pi_g(W; \hat{u}) = p^*_g(W; \hat{u})\lambda(W; \hat{u}) + r^*_g(W; \hat{u})\mu(W; \hat{u}) - C_g^{EN} p^*_g(W; \hat{u}) - C_g^{NL}\hat{u}_g, 
\] (15)

In the optimal solution \( u^{90} \), consider the profitability of the most expensive committed unit, generator 90. It incurs a no-load cost of \$140. Since it is block loaded, we know \( p^*_{90}(W; u^{90}) = 1 \) and \( r^*_{90}(W; u^{90}) = 0 \) for any value of \( W \). If it receives only the LMP, generator 90 will therefore earn a profit of \$1,000 - \$140 = \$860\ ten percent of the time, while it loses \$140 - \$50 = \$90\ ninety
percent of the time. On average, this means that it earns a profit of \(0.1 \cdot \$860 - 0.9 \cdot \$90 = \$5\).

US markets guarantee bid cost recovery, ensuring that no generator loses money over some defined interval (e.g., each block for which it is committed). In the example, generator 90 would be entitled to a make-whole payment ninety percent of the time, i.e., whenever \(W \geq 10\), to cover the difference between its no-load cost and the revenue it earns from selling power. Including such payments, generator 90 would break even in the ninety percent of cases where it previously lost, leading its expected profit to increase to \$86. In the stochastic case, these payments have the effect of socializing generator 90’s losses and privatizing its gains. Accordingly, if the stochastic ideal could be achieved, it would be preferable to relax the standard of bid cost recovery to the less stringent bid cost recovery in expectation. This goal can be stated

\[
E[\pi_g(W; \hat{u})] \geq 0 \quad \forall g \in G. \tag{16}
\]

We showed in Section 3 that the property of bid cost recovery in expectation holds for an optimal solution to a convex relaxation of the stochastic unit commitment problem, while the expected profits are precisely zero for marginal generators. While this property cannot be guaranteed in general due to non-convexity, we suggest that the proximity of generator 90’s expected profit to zero is not an accident. Neglecting commitment costs, profits from the sale of energy and reserves under the basic LMP scheme give an estimate of a unit’s contribution to the dispatch, i.e., the right-hand side of the rule in Eq. (13). Thus, it is reasonable to suggest that even in the non-convex case, the most expensive committed generator in the optimal commitment solution will come close to zero profits in expectation.

5 Two Defects of Deterministic Pricing

While significant progress is being made in stochastic commitment and dispatch, the optimization routines currently used by system operators are deterministic. These deterministic routines give rise to two mechanisms by which operators may commit units for which the property of bid cost recovery in expectation does not hold. First, in the commitment process, operators regularly adjust the load forecast upwards to account for uncertainty that has not been explicitly represented. Second, at the time of dispatch, operators use the best available point forecast for grid conditions.
during the dispatch interval. In this section, we discuss the consequences these workarounds have for pricing.

5.1 Unit Commitment with Biasing

For the deterministic unit commitment, we modify the stochastic model (11) in two ways. First, we substitute the random variable $W$ with its expected value $\bar{W}$. Second, we add a load biasing term $b$ to the demand. Reusing constraints from the subproblem in model (12), the deterministic unit commitment with load biasing can be stated as

$$
\begin{align*}
\text{maximize} & \quad V^D d + V^R o - \sum_{g \in G}(C_{NL} u_g + C_{EN} p_g) \\
\text{subject to} & \quad (12c)-(12e), (12g)-(12j) \\
& \quad d + b - w - \sum_{g \in G} p_g = 0 \\
& \quad w \leq \bar{W} \\
& \quad u_g \in \{0, 1\} \quad \forall g \in G.
\end{align*}
$$

Equipped with a deterministic model, the challenge for human operators is to convert uncertainty into an appropriate bias $b_t$ in each time period $t$. In the single-period example system, expected wind $\bar{W} = 50$, while the optimal solution to the true stochastic commitment problem is to commit the cheapest 90 block-loaded units. Solving a bias-free model would lead to poor results: with 50 units of wind available along with 120 units of capacity from generator 0, the remaining demand for energy and reserves can be satisfied with 50 block-loaded units. Ideally, the operator will instead choose $b = 40$, which will induce the solver to commit 90 block-loaded units.

5.2 Stochastic Pricing with Suboptimal Commitments

While the optimal bias can be determined in the example system, there is no general method to calculate appropriate biases for more complex systems. As such, the choice of bias is left to the engineering judgment of system operators. It bears emphasizing that, despite any negative connotations associated with the word “bias,” some amount of biasing is necessary to achieve a good commitment solution with a deterministic model. However, conservatism on the part of operators could lead to excess biasing. This choice can have important consequences for the prices observed in the system.
We first consider the ideal pricing situation from Section 4.2, in which uncertainty resolves before the time of dispatch and prices are calculated through model \((ED - VERT)\). The energy price \(\lambda(W; u^{90})\) takes values of $1,000 and $50 depending on the availability of wind. The relative probabilities of the two prices, however, depends on the commitment solution. Table 2 shows the decline in average price, \(E[\lambda(W; \hat{u})]\), as additional units are committed. At the optimal commitment solution, \(u^{90}\), there is a ninety percent chance of avoiding a reserve shortfall. As discussed in Section 3, this leads to bid cost recovery in expectation for generator 90. In the solution \(u^{91}\), generator 91 becomes the most expensive unit. Bid cost recovery in expectation no longer holds: the average price is $135.50, lower than this unit’s $141 no-load cost. The situation deteriorates with additional commitments. In the most conservative case, in which the operator chooses bias \(b = 50\) and avoids any chance of a reserve shortfall, the expected price drops to \(\lambda(W; u^{100}) = 50\).

**Table 2: Dependence of reserve shortfall and prices on commitment**

<table>
<thead>
<tr>
<th>Commitment Solution</th>
<th>Probability of Reserve Shortfall</th>
<th>Average Price</th>
</tr>
</thead>
<tbody>
<tr>
<td>(u^{90})</td>
<td>0.10</td>
<td>$145.00</td>
</tr>
<tr>
<td>(u^{91})</td>
<td>0.09</td>
<td>$135.50</td>
</tr>
<tr>
<td>(u^{92})</td>
<td>0.08</td>
<td>$126.00</td>
</tr>
<tr>
<td>(u^{93})</td>
<td>0.07</td>
<td>$116.50</td>
</tr>
<tr>
<td>(u^{94})</td>
<td>0.06</td>
<td>$107.00</td>
</tr>
<tr>
<td>(u^{95})</td>
<td>0.05</td>
<td>$97.50</td>
</tr>
<tr>
<td>(u^{96})</td>
<td>0.04</td>
<td>$88.00</td>
</tr>
<tr>
<td>(u^{97})</td>
<td>0.03</td>
<td>$78.50</td>
</tr>
<tr>
<td>(u^{98})</td>
<td>0.02</td>
<td>$69.00</td>
</tr>
<tr>
<td>(u^{99})</td>
<td>0.01</td>
<td>$59.50</td>
</tr>
<tr>
<td>(u^{100})</td>
<td>0.00</td>
<td>$50.00</td>
</tr>
</tbody>
</table>

5.3 Ex-Ante Deterministic Pricing

A separate pricing issue can arise when uncertainty is not completely resolved before the pricing model is run. In practice this is necessarily the case, since pricing intervals are at least five minutes long in real-world systems. Interpreting our example system differently, we can imagine that wind output will fluctuate between 0 and 100 over the course of the interval. Accounting for every fluctuation would require continuous-time pricing of electricity [29]. Instead, prices are set by running a single instance of model \((ED - VERT)\) using the expected wind output \(\bar{W} = 50\). With this output reserves are plentiful, and the price \(\lambda(\bar{W}; u^{90}) = 50\). In other words, the price reflects marginal cost under the expected operating conditions, rather than the expected price under a
range of operating conditions. Even under the optimal commitment $u^{90}$, this price fails to satisfy the property of bid cost recovery in expectation.

6 Quasi-Stochastic Improvements

The pricing problems identified above lead to a difference between the expected price in the stochastic ideal with the optimal commitment, $E[\lambda(W; u^*)]$, and the single price seen by market participants with the actual commitment, $\lambda(W; \hat{u})$. We propose that the goal of ORDCs should be to bridge the gap between these two prices. This conception generalizes the ORDC beyond its original focus on times of scarcity [13], during which the difference $E[\lambda(W; u^*)] - \lambda(W; \hat{u})$ is likely greatest. This generalization helps explain why an ORDC might be active even at times when no chance of scarcity exists. In this section we present current ORDC formulations, demonstrating that they achieve our goal under ideal circumstances, and then discuss three issues with price formation they may fail to address.

6.1 Probability-Driven ORDC

The example system presents the best possible circumstances for construction of an ORDC along the lines of [13]: we have complete knowledge of the probability distribution for wind availability, are able to identify the optimal commitment solution $u^{90}$, can easily determine that the marginal unit has a variable cost of $50, and have a constant recourse cost of $950. Ninety percent of the time, holding additional reserves would have no value. The remaining ten percent of the time, additional reserves would mitigate the system’s reserve shortfall, providing value of $950. Before uncertainty is resolved and assuming risk neutrality, therefore, holding additional reserves has marginal value $E[\mu(W; u^{90})] = $95. Measured against the expected wind W, this commitment solution holds 60 units of reserves, much more than the required 20. The uniform distribution of W simplifies the calculation of expected value at different levels of reserve: with only 20 units of reserve online, there is a 50 percent chance of shortage, so the expected value of additional reserves is $475. With 70 units of reserve, there is no chance of a shortfall, so the expected value is $0. Between these reserve levels, expected value falls linearly.

The economic dispatch problem in model (ED – VERT) utilizes a vertical demand curve for reserves, assigning a value of 0 to any reserves above the level $R^+$. The sloped ORDC formulation
replaces this with a function reflecting the declining expected value of reserves. In keeping with the formulation in [13], implementation in ERCOT, and proposal for PJM, we define the ORDC for our example with respect to the expected wind $W$ and ignore the negative net load changes that would come with wind above this level. As in model $(ED - VERT)$, we formulate the dispatch and pricing problem as a single-period problem. Instead of the linear coefficient $V^R$ up to the quantity $R^+$, we express the total value of these reserves is expressed through the concave, non-decreasing function $V^O(o)$, which in this example is linear up to the quantity $o = 20$, quadratic in the range $20 \leq o < 70$, and constant for $o \geq 70$. The economic dispatch with sloped ORDC can be stated as

\[
(ED - ORDC) : \maximize_{p, r, d, o, w} \quad V^D d + V^O(o) - \sum_{g \in G} C^E_{g} p
\]

\[
\text{subject to} \quad (12d), (12g)-(12j)
\]

\[
d - w - \sum_{g \in G} p_g = 0 \quad (18b)
\]

\[
o - \sum_{g \in G} r_g = 0 \quad (18c)
\]

\[
w \leq W. \quad (18d)
\]

In the deterministic pricing setting discussed in Section 5.3, prices were set using the vertical demand curve of model $(ED - VERT)$ and the expected wind output $W$. Model $(ED - ORDC)$ similarly uses $W$, but incorporates the uncertain $W$ in the function $V^O(o)$. Under this model, the LMP $\lambda^O(\hat{u})$ and reserve price $\mu^O(\hat{u})$, calculated as the dual variables to Eqs. (18b) and (18c), are still functions of the commitment solution, but no longer functions of the random variable. Assuming the optimal commitment solution, $u^{90}$, solving model (18) leads to wind clearing at 50 units, each of the 90 block-loaded units providing 1 unit of energy, and generator 0 meeting the remaining 60 units of demand. However, instead of being limited to 20 units, all 60 units of generator 0’s remaining capacity now clear as reserves. Since the marginal value of reserves is $95 at this level, the reserve price $\mu^O(u^{90}) = $95, while $\lambda^O(u^{90}) = $145. In other words, the ORDC recovers the expected price under the stochastic ideal, $\mathbb{E}[\lambda(W; u^*)]$, despite being a deterministic model.
6.2 Stochasticity without Shortage

In the example system, limiting uncertainty to a one-dimensional random variable enables straightforward calculation of the probability of reserve shortage. In real-world systems, a wide array of transmission contingencies, load fluctuations, and generator failures can impact the cost of operating the system and thus the prices that would arise in the stochastic ideal. Combinations of these events could lead to a wide range of prices in a stochastic ideal setting, in which markets could be cleared in real time as conditions changed. This stochasticity in system cost is present at all times, even when the probability of lost load is negligible.

With their focus on reserve shortages, current formulations in ERCOT and PJM do not capture this more general form of stochasticity, instead assuming a constant recourse cost. By itself, this could mean that both underestimate the value of reserves. However, both markets potentially overestimate the value of reserves by incorporating forecast errors over much longer timescales than the five minutes for which prices are valid. Without an external metric to validate ORDC parameters, it is difficult to assess whether the net effect is to under- or overvalue reserves.

6.3 Robustness to Commitment Solution

With perfect knowledge, the ORDC in model (18) solves the pricing issue identified in Section 5.3, namely, the need for the system operator to set prices before uncertainty is realized. It does not, however, address the issue of suboptimal commitments raised in Section 5.2. If the operator in the example commits all 100 block-loaded units, the system registers 70 units of reserve and the LMP drops to $\lambda^O(u^{100}) = $50. In other words, the clearing price of reserves is not determined solely by the cost of the recourse action and the probability it will take place; an additional factor is the bias chosen by the operators. In recognition of this dependence, ERCOT utilizes a Real-Time On-Line Reliability Deployment Price Adder, which attempts to counteract the pricing impacts of operator actions [32]. The example system illustrates a potential difficulty of calculating this adder. As discussed in Section 5.1, reliability deployments at a level of 40 units restore the stochastic optimal commitment solution. It is only bias above this level that creates a problem for pricing in need of correction. Differentiating between commitments that do and do not warrant correction is not straightforward.

One approach to help resolve inconsistencies between the commitment solution and prices is
to use an ORDC in the commitment problem itself. It is likely that incorporating an ORDC would represent an improvement over the deterministic status quo, mitigating the need for operator biasing. However, since stochastic and robust optimization techniques are typically seen as more promising ways to handle uncertainty at the commitment stage, it is unlikely that the need for intervention would be eliminated. This paper takes the view that system operators should adapt pricing to the reliability function, rather than letting the needs of the market dictate reliability processes. A more subtle issue with this approach arises from the fact that forecast uncertainty at the time of commitment is different from forecast uncertainty at the time of dispatch. Time consistency between prices in the commitment and dispatch models is therefore not guaranteed if the same ORDC is used in both.

6.4 Flexibility Compensation

While an ideal ORDC may be able to bridge the gap between the expected price under uncertain conditions and the deterministic price in the expected conditions, it does not recover all the properties of the stochastic ideal. In the example, consider the revenue earned by the wind generator. In the stochastic setting, prices and wind output are negatively correlated: since reserve shortages only occur when $0 \leq W < 10$, average wind generation in these high price times is only 5. The average wind output when prices are lower is 55. Expected revenue for the wind generator is thus $5 \cdot $1,000 \cdot .1 + 55 \cdot $50 \cdot .9 = $2,975. In the ORDC setting, by contrast, wind earns $145 per unit regardless of its metered output, giving an expected revenue of $50 \cdot $145 = $7,250.

While the example uses wind, in real-world systems many generators fail to meet their desired dispatch point. In recognition of physical limitations, most systems assess penalties for deviations only outside a wide tolerance (e.g., 8 percent in MISO [31]). Mismatches between cleared and metered output can result from these physical limitations as well as “price chasing” behavior on the part of generators. Recognizing that less frequent settlement effectively under-compensates flexible resources that are able to consistently meet their desired dispatch point, US markets have moved to five-minute settlement intervals [7].
7 ORDC Construction

Given the challenges faced by current formulations, this section considers how we might construct ORDCs to bridge the gap between the expected price that would arise in a stochastic model, \( \mathbb{E}[\lambda(W; u^*)] \), and the price in the single scenario chosen by system operators for the deterministic dispatch, \( \lambda(W; \hat{u}) \). In doing so, we hope to approximate the outcomes that might be achieved in a stochastic competitive equilibrium, in which system operators maximize expected surplus and market participants maximize expected profits. We begin with the case where system operators are able to implement stochastic models for background reliability processes (e.g., the PJM IT-SCED and ISO-NE Real-Time Unit Commitment), but are unable to implement a fully stochastic market design. We then consider the problem faced by current system operators using deterministic commitment models. Lastly, we return to the subject of bid cost recovery in expectation, highlighting three implications of the analysis for future market design.

7.1 Stochastic Reliability Models

If the system operator runs a stochastic model along the lines of model (11) for commitment, it can reasonably be assumed that the ex ante optimal commitment solution \( u^* \) will be chosen. The expected price \( \mathbb{E}[\lambda(W; u^*)] \) in any period can then be calculated using the dual variables in the dispatch subproblem, model (12). After the system operator chooses a particular scenario for use in the deterministic pricing model, it is straightforward to find the marginal cost of energy and the quantity of reserves available given that choice of scenario. In the example system, \( \mathbb{E}[\lambda(W; u^*)] = $145 \) and we chose the scenario \( W = 50 \). This enables us to calculate a marginal energy cost of $50 under the vertical reserves demand curve in model \( ED - VERT \) and to determine that a total of 60 units of reserves are online. When constructing the ORDC, this means we should choose a curve such that \( V^O(60) = $95 \). Inserting this valuation in model (18) yields a reserve clearing price of \( \mu^O(u^*) = $95 \) and an energy price of \( \lambda^O(u^*) = $145 \).

Here we formalize this methodology, which extends naturally to the dynamic, multi-period setting, by returning to the more extensive model developed in Section 3. Prior to the beginning of each pricing interval, we assume the operator has access to information \( \mathcal{F}_t \), from which they can calculate the probabilities \( \phi(s|\mathcal{F}_t) \). The operator chooses random variable realizations from a
single scenario $s \in S$ in order to calculate prices with a deterministic, single-period model. Our proposed method recovers the expected energy price $\sum_{s \in S} \phi(s|F_t)\lambda_{st}$ through construction of an ORDC specific to the operator scenario choice, time period, and available information. The ORDC as constructed takes the form of a piecewise constant value function broken into segments $c \in \mathcal{C}$.

Reserves up to a quantity $R_{ct}^+(s, F_t) \geq 0$ in segment $c \in \mathcal{C}$ are valued at $V_{ct}^o(s, F_t) \geq 0$, while the total quantity supplied in that segment is expressed as $o_{sc}$. While these parameters will depend on information $F_t$ and choice of scenario $s$, we omit this notation from the formulation. We assume generator commitments are fixed at $u_{gt}^*$, as determined by a stochastic unit commitment performed prior to the dispatch interval. The economic dispatch problem with operating reserve demand curve for time period $t$ given scenario choice $s$ is stated as

$$(ED - ORDC)_{st} :$$

maximize $\sum_{l \in \mathcal{L}} V_l D_{sl} + \sum_{c \in \mathcal{C}} V_{ct}^O o_{sc} - \sum_{g \in \mathcal{G}} C_g^E p_{sgt}$

subject to

$\sum_{l \in \mathcal{L}} d_{slt} - \sum_{g \in \mathcal{G}} p_{sgt} = 0$ (19b)

$\sum_{c \in \mathcal{C}} o_{sc} - \sum_{g \in \mathcal{G}} r_{gst} = 0$ (19c)

$d_{slt} \leq D_{slt}^-$ $\forall l \in \mathcal{L}$ (19d)

$o_{sc} \leq R_{ct}^+$ $\forall c \in \mathcal{C}$ (19e)

$P_{sgt}^- u_{gt}^* \leq p_{sgt}$ $g \in \mathcal{G}$ (19f)

$p_{sgt} + r_{gst} \leq P_{sgt}^+ u_{gt}^*$ $g \in \mathcal{G}$ (19g)

$d_{slt} \geq 0$ $\forall l \in \mathcal{L}$ (19h)

$o_{sc} \geq 0$ (19i)

$p_{sgt}, r_{gst} \geq 0$ $g \in \mathcal{G}$.

With this problem defined, an ORDC can be constructed on the basis of Theorem 2.

Theorem 2. Suppose that for a chosen scenario $s \in S$, we are able to identify generator $g' \in \mathcal{G}$ as a marginal supplier of energy, i.e., $P_{sgt}^- u_{gt}^* < p_{sgt}^* < P_{sgt}^+ u_{gt}^*$. Further suppose that we are able to determine a segment $c' \in \mathcal{C}$ for which $0 < o_{sc'} < R_{ct}^+$, and let the value of reserves $V_{ct}^O = \sum_{s \in S} \phi(s|F_t)\lambda_{st} - C_g^E$ in this segment. Then the clearing price of energy $\lambda_{st}^O = \sum_{s \in S} \phi(s|F_t)\lambda_{st}$. 

26
Proof. In the dual to \((ED - ORDC)_{st}\), the constraint related to the primal variable \(o_{st}\) is

\[
\mu_{st}^O + \alpha_{st}^R \geq V_{ct}^O.
\] (20)

Since in the optimal solution \(o_{st}^* > 0\), by complementary slackness this inequality is binding. Further, since \(o_{st}^* < R_{ct}^+\), we know that \(\alpha_{st}^R = 0\). Therefore,

\[
\mu_{st}^O = V_{ct}^O = \phi(s|F_t)\lambda_{st} - C_{g'}^{EN}.
\] (21)

Now we consider the dual constraint related to the primal variable \(r_{sg't}\), i.e.,

\[
-\mu_{st}^O + \alpha_{sg't}^{P^+} \geq 0.
\] (22)

Since reserves are costless to provide and \(V_{ct}^O \geq 0\), we can always choose an optimal solution in which generator \(g'\) supplies reserve up to its capacity, i.e., \(p_{sg't} + r_{sg't} = P_{sg't}^+ u_{gt}^*\). Since \(p_{sg't}^* < P_{sg't}^+ u_{gt}^*\), this implies that \(r_{sg't} > 0\), and thus \(\mu_{st}^O = \alpha_{sg't}^{P^+}\). Lastly, the dual constraint related to the primal variable \(p_{sg't}\) gives

\[
-\lambda_{st}^O - \alpha_{sg't}^{P^-} + \alpha_{sg't}^{P^+} \geq -C_{g'}^{EN}.
\] (23)

Since \(P_{sg't}^- u_{gt}^* < p_{sg't}^*\), by complementary slackness this inequality is binding and \(\alpha_{sg't}^{P^-} = 0\). Substituting \(\phi(s|F_t)\lambda_{st} - C_{g'}^{EN} = \mu_{st}^O = \alpha_{sg't}^{P^+}\) gives

\[
\lambda_{st}^O = \sum_{s \in S} \phi(s|F_t)\lambda_{st}
\] (24)

In this construction, the marginal value of reserves \(V_{ct}^O\) reflects a weighted average of the cost of all recourse actions included in the stochastic model, including the reserve penalty factor \(V^R\). In principle, constructing an ORDC in this way allows reconstruction of the expected price in a deterministic model. The assumption that we can identify the marginal supplier of energy may seem somewhat limiting. However, we note that the same assumption is made in the ORDC construction in [13]. In practice, a reasonable proxy for the marginal cost of energy is likely to be available based
on the results of recent dispatch periods. The assumption that we can identify which segment of the ORDC will be active is much weaker, since the total quantity of reserves available is tracked in real time by system operators.

A complication in practice, but not in theory, is that commitment decisions do not occur at the same time as dispatch and pricing decisions, and accordingly are made with different information available. Suppose that at time $t_1$, the operator makes a decision to commit a generator for a future time $t_2$. We note that by the tower property,

$$\mathbb{E}\left[\mathbb{E}\left[\lambda_{st}\mid F_{t_2}\right] \mid F_{t_1}\right] = \mathbb{E}\left[\lambda_{st}\mid F_{t_1}\right].$$

(25)

This states that the expected price in interval $t_2$, given the information the operator learns between $t_1$ and $t_2$, is equal to the expected price when it was committed.

## 7.2 Deterministic Reliability Models

A potentially greater challenge is to calculate a marginal value of reserves consistent with operator commitments even if those commitments are made with a deterministic model. If the operator instead uses a deterministic model with biasing along the lines of model (17), we do not have access to the optimal commitment solution or expected price. We propose to address the first issue by assuming that the choice of bias made by operators is in fact optimal. To ensure consistency between operator actions and prices, the simplest choice is to trust the engineering judgment of the operators. As operators gain access to increasingly sophisticated algorithmic decision support tools, this assumption is likely to become increasingly valid over time. The second issue can be addressed by invoking the result shown in Section 3.6 in the convex case: at the time of commitment, the expected profit for a marginal generator in the commitment solution should be precisely zero.

Suppose in the example that the operator chooses a bias of $b = 42$. The most expensive generator committed in the optimal solution, generator 92, has a no-load cost of $142 and will earn all its revenues from the sale of energy, which it produces at a cost of $0. The fact that operators committed the unit through their choice of bias reveals their preference for holding an additional unit of capacity online at a cost of $142. When constructing an ORDC, we should therefore choose a curve such that $V^O(62) = 92$, leading to reserve price $\mu^O(u^{92}) = 92$ and energy price $\lambda^O(u^{92}) = 142$. 28
More generally, the goal of bid cost recovery in expectation for the most committed unit offers a lower bound on expected prices. Accordingly, such a strategy could be successful as long as at least one unit in the system remains uncommitted. Translating this deterministic methodology to the dynamic, multi-period setting, however, is less straightforward than in the stochastic case.

7.3 Implications for Future Market Design

The proposed method recovers some version of the expected price, whether that is calculated explicitly in a stochastic model or implicitly in a deterministic one. In the example system, this leads to bid cost recovery: the price produced is sufficiently high that all committed units cover their avoidable fixed costs. However, three complicating factors prevent more general guarantees on bid cost recovery in expectation. Each of these factors has implications for future market design improvements beyond the ORDCs discussed in this paper.

The first is the flexibility effect discussed in Section 6.4. A generator whose production is positively correlated with prices may achieve bid cost recovery in expectation under the stochastic ideal, but not when the uncertainty in prices is removed. Moreover, as described the methodology targets the expected price, rather than the total expected cost to operate the system. Since the generation side of the market has historically been significantly more flexible than the demand side, it may be preferable to target a price that is weighted by the cleared load in the scenario. In other words, instead of the average price $\mathbb{E}[\lambda(W; u^*)]$, the market operator could target the average generator revenue $\mathbb{E}[d^*(W; u^*)\lambda(W; u^*)]$. This revenue could then be shifted to more flexible generators through a system of penalties and rewards for adherence to dispatch signals.

The second is non-convexity. Recognition of non-convexity in the deterministic context has led to the introduction of make-whole payments and enhanced pricing schemes, justified on the grounds that market participants following operator dispatch signals should not lose money. As illustrated by the example, however, the stochastic analysis offers a new interpretation of the challenge posed by non-convexity. In a stochastic setting, commitments may be profitable in expectation but result in losses under some realizations. In such cases it is difficult to justify make-whole payments or enhanced pricing, which amount to socializing losses and privatizing gains. If ORDCs effectively raise real-time prices to the level expected at the time of commitment, these efforts to address non-convexity may no longer be warranted.
The third is asymmetric beliefs. The model of stochastic competitive equilibrium in Theorem 1 relies on the assumption that all market participants share a common assessment of the probability that each scenario materializes. If this assumption does not hold, the market operator may believe that a unit will be profitable in expectation, but the market participant may not share that belief. Reconciling the two beliefs may be possible with the introduction of a broader set of market instruments [11]. For example, if in the example the system operator expected a price of $145 while generator 90 expected a price of $130, the system operator could buy a future at a price between $140 and $145 at the time of commitment. This approach introduces other challenges, since the operator would then be at risk of revenue shortfall. However, socializing net losses on these contracts may be preferable to maintaining current uplift payments.

8 Conclusion

This article explains the inconsistency between commitment, dispatch, and pricing observed in current markets as arising due to the use of deterministic models to represent a stochastic reality. A full stochastic model gives information on the expected cost to operate the system given a range of potential conditions, whereas a deterministic model estimates the cost given one set of assumed conditions. We propose that the goal of ORDCs should be to recover this expected cost, thus improving the quality of the price signal produced in real-time markets. This interpretation generalizes ORDCs, helping explain their relevance in markets where they are not required for resource adequacy. Restoring the expected price can be achieved with only modest changes to current market clearing mechanisms, making ORDCs promising for near-term implementation.

Whereas previous ORDC implementations rely primarily on administrative parameters, this paper proposes to construct ORDCs by tying them to the commitment decisions made by system operators. Since current proposals rely on administrative choices loosely connected to the probability of involuntary load shedding, they are unlikely to fully align with operator decisions. If these administrative parameters underestimate the value placed on reserves by operators, ORDCs will fail to achieve price consistency between commitment and dispatch. If they overestimate the value of reserves, they may instead induce inefficient self-commitment of units not needed for reliability.

The ORDC construction envisioned in Section 7.1 relies on information available in stochastic unit commitment models. Implementation of these models for non-market reliability processes
may be realistic in the near future. At present, however, operators in centrally committed markets rely on deterministic models with biasing. We sketch an approach in Section 7.2 to construct ORDCs on the basis of the more limited information available in these models. Starting from either a stochastic or deterministic model, translating our approach to dynamic, multi-period settings with network constraints and multiple reserve products presents several challenges. Instead of constructing ORDCs, a more immediate application of the analysis in this paper may be to assess the quality of curves generated by administrative choices. In other words, administrative ORDCs can be judged on the basis of achieving bid cost recovery in expectation while avoiding uneconomic self-commitments.

The analysis identifies three additional shortcomings of current markets connected to stochasticity. First, deterministic market models have the effect of undercompensating flexible resources. While a shift to more frequent settlement intervals helps alleviate the problem, this undercompensation could become more significant with increased contributions from variable renewable resources. Second, ORDCs weaken the justification for uplift payments. If ORDCs are successfully implemented, stakeholders may wish to reexamine current practices regarding these payments as well as efforts to implement enhanced pricing schemes. Third, the presence of uncertainty leads to the potential for asymmetric beliefs between market participants and the operator. These asymmetries are particularly important for energy-limited resources, the offers of which are contingent on self-assessed opportunity costs. Asymmetry could lead to suboptimal commitment and dispatch and may necessitate a broader set of instruments to facilitate market clearing. Accordingly, while ORDCs offer the potential to resolve one issue with deterministic market designs, further improvements are needed to fully reflect the stochastic reality.

**Acknowledgment**

The author thanks staff in the Business Architecture and Technology department at ISO New England. The views expressed in this paper are those of the author and do not represent the views of ISO New England.
References


