Optimal Bidding and Contracting Strategies in the Deregulated Electric Power Market: Part I

D.-J. Wu
LeBow College of Business, Drexel University, Philadelphia, PA 19104
wudj@drexel.edu

Paul Kleindorfer
The Wharton School, University of Pennsylvania, Philadelphia, PA 19104,
kleindor@grace.wharton.upenn.edu

Jin E. Zhang
City University of Hong Kong, Kowloon, Hong Kong
efjzhang@cityu.edu.hk

Abstract

This paper models the interaction of long-term contracting and spot market transactions between one Genco and one or more Discos. The basic model proposed allows the Genco and Discos to negotiate bilateral electric power contracts and then, on the day, to sell or buy in an associated spot market. The type of bilateral contracts studied has a two-part tariff structure that is at the foundation of practical contracts used in industry. The first part is a reservation cost per unit of capacity, and the second an execution cost per unit of output when this capacity is actually used. We show that Discos’ optimal contracting strategies follow an index that combines the Genco’s reservation cost and execution cost. The Genco’s optimal strategy is to reveal its production cost but extract its margin from the Discos from the capacity reservation charge. The optimal reservation charge depends on the opportunity cost the Genco loses from the spot market as well as the inverse elasticity of the Discos’ contract demand.

1. Introduction

The energy sector has been going through a period of rapid restructuring from vertically integrated franchise operations to competitive structures in which production, transmission and distribution are separately priced and typically also provided by different Gencos. The resulting market structure has stimulated an explosive growth in financial and contractual innovations. Foremost among these is the development of bilateral contracts of the sort discussed here. In natural gas, contracts for production and delivery, which used to have a duration of years, are now being bid on a weekly and monthly basis [1]. In electric power, bilateral contracts are at the foundation of many of the evolving models of the restructured market place [4]. Contracting serves in these markets both the important role of price discovery as well as the obvious direct role of coordinating capacity commitments with anticipated demand. Such contracting takes place against the backstop supply source of the spot markets that provide output “on the day”. In natural gas and electric power, such spot markets are well developed, including various futures trading instruments and other derivatives to assist in price discovery. These exist through the New York Mercantile Exchange in the U. S. and through other exchanges throughout Europe, Latin America and Asia. Beyond these trading activities, which serve as fundamental drivers of overall supply chain efficiency, bilateral contracting among Gencos and large buyers in the energy market is fundamental to its overall functioning. Such bilateral contracts take many forms, but basically they commit a Genco to reserve a certain amount of capacity (e.g., 100 Megawatts of power) during a specific period (e.g., each weekday from 6 a.m. to 10 p.m. during a specific week) for a particular buyer’s use (e.g. a distribution company selling energy to final customers). Additional fees levied if the capacity is actually used (in the example noted, these would be levied per Megawatt-hours delivered). Such contracts are the foundation for a significant portion of the $600 Billion dollars in annual sales of the U.S. electric power industry (see ref. [2], for further details on contracting in electric power markets).

This paper considers a deregulated electric power market
in which there are a producing Genco, Discos and a Transmission Company. The Genco may either sign long-term contracts with Discos to provide up to a fixed quantity of output ex post, or sell its output in the spot market. Similarly, Discos may sign long-term contracts to cover their anticipated needs, or they may cover these needs through purchases in the spot market.

The non-storability of electric power gives rise to the so-called two-goods problem. The first of two goods is the availability of capacity itself, pre-committed to a specific buyer. The second good is the output actually delivered on the day to the buyer. The key issue we study here is how spot pricing and bilateral contracting in these markets are or should be linked. From the Genco’s perspective, pre-committing capacity at a fixed price to a particular buyer may exclude more profitable opportunities through the spot market on the day. The same is true for buyers signing such pre-commitment agreements. Thus, as we shall see, the key tradeoff in determining how such contracts should be priced and how much capacity should be committed to them by Gencos and Discos are the relative costs and risks of sourcing from the contract versus the spot market.

The rest of this paper is organized as follows. In Section 2, we consider the case of a single Genco and a single Disco, first neglecting transmission costs, then adding transmission costs. Last, in Section 3, we extend these results to the more general case of a single Genco facing many contracting opportunities with multiple Discos. We conclude in Section 4 with a number of suggestions for future research.

Note: Due to space limitation, this is an abbreviated version of the first [8] in a two-part series of papers dealing with this subject. Omitted proofs of lemmas, theorems, or corollaries can be found in the full Version of this paper [8]. For Part II, see [9].

2. Basic model

This section models procurement of electric power from two potential sources: long-term contracting and spot market purchases. There are two agents in this model: A Disco that needs to purchase electric power and a Genco with whom the Disco can sign long-term contracts. Other Gencos provide additional electric power on the day through a spot market. The sole source of uncertainty is the spot market price P_s. This situation is modeled as von Stackelberg Game with the Genco as the leader. The Genco selects his profit maximization contract prices (reservation cost s per unit of capacity and execution cost g per unit of output actually used) anticipating how the Disco will react. Given these costs [s, g], the Disco determines its optimal reservation or contract level Q. Contracts are signed before the market price P_s is known. After observing the spot price (on the day), the Disco determines how many units of the contracted goods to use and how many to purchase from the spot market.

We now get into the details of our model. The capacity for the Genco is K; The cost of the Genco is characterized by the parameter pair (b, β), where b is the short-run marginal cost of providing a unit at the Genco’s door, and β is unit capacity cost per period. The Genco is assumed to offer long-term contracts to the Disco in the form of a specified amount of reserved capacity Q and subject to a two-part tariff of the form [s, g]. In the form, s is the reservation cost per unit of capacity per period and g is execution cost, the price per unit the Genco charges to the Disco if the Disco exercises its contract with the Genco on the day.

We denote the Disco’s total demand, on the day, from both contract and spot purchases as D(P_s(ω), Q). We show below that, as this notation suggests, the Disco’s optimal consumption on the day will depend only on the spot market price P_s(ω) and the capacity reservation level Q. The spot market price at the state of the world ω is denoted P_s(ω) and is assumed to be uninfluenced by either the Genco or the Disco. The distribution of the spot market price is assumed to be common knowledge.

Given the Genco’s offer [s, g], the Disco decides how much to contract with the Genco, Q(s, g). On the day, when the state of the world ω is known, the Disco decides how much to use from this contract, denoted as q(P_s(ω)) and how much to purchase from the spot market, denoted as x(P_s(ω)). The Disco consumes up-to its contracted amount only if the spot market price is higher than its contracted price, i.e., its actual consumption q(P_s(ω), g, Q, D) satisfies

\[
q(P_s(ω), g, Q, D) = \min[Q(s, g), D(P_s(ω), Q)] \chi(P_s(ω) - g), \tag{1}
\]

where D = D(P_s(ω), Q(s, g)) is the total demand of the Disco, and \(\chi(\cdot)\) is the indicator function (which takes the value of 1 if its argument is positive and 0 else).

The Disco’s residual demand not covered by the long-term contract is covered in an open spot market. This open spot market is supported by the Genco as well as by other Gencos. Given a market price, the Genco will bid its output into the spot market if its short-run marginal cost (b) of supply is less than the price paid him (P_s(ω)). For notational simplicity, we will omit the state variable ω in the following sections.

2.1. The Disco’s problem

In this subsection, we define the Disco’s objective function, provide justifications of underlying assumptions, then derive the Disco’s optimal reservation level Q*, and finally give some examples.
Using the standard quasi-linear form of utility (in ref. [5], pp. 97), the Disco’s utility at \( P_s \) is given by

\[
V(D, Q, P_s) = U(D) - sQ - g q(P_s, g, Q, D) - P_s x,
\]

(2)

where the first term is its Willingness-To-Pay (WTP) at \( P_s \), evaluated at the realized demand \( D(P_s, Q) \), the second and the third term together are the payment for the goods delivered under the long-term contract, and the fourth term is the payment for goods \( x \) purchased in the spot market.

Concerning the form of the Disco’s “utility”, this is standard in both economics and marketing science [3, 5]. Note, in particular, that the WTP function \( U \) is the basis for normal demand theory since the solution, call it \( D_s(P_s) \), to maximizing \( U(D) - P_s D \) is simply \( U'(D) = P_s \) or \( D_s(P_s) = (U')^{-1}(P_s) \). Thus, the standard demand curve for the Disco is just \( D_s = (U')^{-1} \) where \( (U')^{-1} \) is the inverse of \( U' \). The standard tools of Marketing Science would be used to specify and estimate the WTP function \( U \). Since we are considering here dual sources of fulfillment (namely, spot or contract purchases), a more complicated solution for the demand equation results than the traditional single-source demand function \( D_s(P_s) \). Nonetheless, as we now show, the dual-source demand function of interest here is a transformation of \( D_s(P_s) \), so that the standard demand curve \( D_s(P_s) \) is the basic building block of the results presented here. Define \( x^+ = \max(x, 0) \).

Using (1), (2) can be rewritten as

\[
V(D, Q, P_s) = U(D) - P_s D + (P_s - g)^+ \min[D, Q] - sQ. \tag{3}
\]

The Disco’s problem can be solved in two stages. Starting at the end (the first-stage), once \( P_s \) is known and \( Q \) is set, the Disco solves for \( q \) and \( x \). Knowing these solutions, the Disco can solve for \( Q \) (the second-stage) by maximizing the expected value of \( V(D, Q, P_s) \) in (3). In keeping with decreasing marginal utility of consumption (note that this does not imply anything about risk preferences. The Disco is in fact risk neutral in our model since \( V(\cdot) \) is linear in money), we will assume that the Disco’s WTP \( U(x) \) is strictly concave and increasing so that

\[
U'(x) > 0, \quad U''(x) < 0, \quad \text{for } x \geq 0. \tag{4}
\]

In the first-stage, \( D(P_s, Q) \) is given as the solution to:

\[
\max_{D \geq 0} V(D, Q, P_s) = \max_{D \geq 0} [U(D) - P_s D + (P_s - g)^+ \min[D, Q] - sQ]. \tag{5}
\]

The following Lemma 1 characterizes the Disco’s demand curve when facing both a spot market as well as a contract market.

Lemma 1:

\[
D(P_s, Q) = \max[D_s(P_s), Q] = \max[(U')^{-1}(P_s), Q]. \tag{6}
\]

Lemma 1 indicates that when the Disco has dual sources for its procurement, its demand curve would change from downward sloping (single-source) to kinked (dual-source). The Disco’s first-stage decision, whether to use contract purchases or to use spot market purchases, depends entirely on the comparison of spot market price \( P_s \) on the day with the Disco’s marginal WTP \( U'(Q) \): If \( P_s < U'(Q) \), then \( x = 0 \), i.e., contract purchases dominate spot purchases.

A key observation that has been used in our proof of Lemma 1 in [8], is that the Disco will never contract for any greater amount than it would find reasonable to use at the contract execution price \( g \), so that, in fact, \( g \leq U'(Q) \) (or equivalently \( Q \leq D_s(g) \)) must hold for any reasonable \( Q \). Lemma 1 is only defined for such “reasonable” values of \( Q \). We discuss further below the implications of Lemma 1 for optimal \( q \) and \( x \). In any case, total demand is derived from the basic building block of the Disco’s standard demand curve \( D_s(P_s) \).

Lemma 1 gives the optimal consumption portfolio of \( q \) and \( x \). Given Lemma 1, (1) can be simplified as

\[
q = \begin{cases} 
Q, & P_s > g, \\
0, & \text{otherwise}
\end{cases}
\]

This indicates that the optimal decision on contract use is a rather simple bang-bang solution: If \( P_s > g \), use all reserved capacity \( Q \), otherwise do not use any reserved capacity. For spot market purchase, \( x \) can be simplified as

\[
x = \begin{cases} 
D_s(P_s), & P_s < g, \\
D_s(P_s) - Q, & g \leq P_s < U'(Q), \\
0, & P_s \geq U'(Q).
\end{cases}
\]

This indicates that the optimal purchase decision is determined by checking whether \( P_s \) fall in or out of the range of \([g, U'(Q)]\). If it is out of the range but \( P_s \geq U'(Q) \), purchase nothing from the spot market; else if \( P_s < g \), purchase only from the spot market. If \( P_s \) falls in the range, then fulfill only the residual demand from the spot market by purchasing \( D_s(P_s) - Q \).

The second-stage of the Disco’s decision is to find a reservation level \( Q \) in order to maximize its expected utility.
\[ EV(Q), \text{i.e.,} \]
\[ \max_{Q \geq 0} EV(Q) = \int V(D, Q, P_s) f(P_s) dP_s, \quad (7) \]
where \( D = D(P_s, Q) \) is given by (6), and \( f(P_s) \) is the probability density function of spot market price \( P_s \), assumed to be common knowledge among all market participants.

We now define the “effective price function” ("G" function for short; we will explain its meaning below) \( G(p) \) as
\[ G(p) = \int_0^p (1 - F(y)) dy = \int_0^p P_s f(P_s) dP_s + p(1 - F(p)) = E[\min(P_s, p)], \quad (8) \]
where \( F(y) \) is the cumulative distribution function of the spot market price \( P_s \). The second equality holds due to the consequence of integration by parts. Notice that,
\[ \lim_{p \to \infty} G(p) = \lim_{p \to \infty} \int_0^p (1 - F(y)) dy = \bar{P}_s, \quad (9) \]
and
\[ E(P_s - x)^+ + E[\min(P_s, x)] = \bar{P}_s, \quad (10) \]
The meaning of the “G” function is the following. By (9), \( G(p) \) equals the expected spot price as \( p \) approaches infinity. For finite \( p \), by (8), \( G(p) \) is just the expected value of \( \min(P_s, p) \). If a contract can be executed at price \( p \), then this represents the expected price per unit paid by a Disco, since the Disco will elect to use the spot price \( P_s \) when this is less than \( p \) and will use the contract when \( p \) is less than the spot price. Thus, the terminology for the \( G \) function as “effective price function”.

**Lemma 2:** \( EV(Q) \) is concave in \( Q, \forall (s, g) \).

**Proof:** This follows from \( \frac{\partial^2 EV}{\partial Q^2} = G'(U'(Q))U''(Q) = (1 - F(U'(Q)))U''(Q) < 0. \)

Denote \( G^{-1}(-) \) as the inverse function of \( G(-) \). Denote \( Q^*(s, g) \) as the optimal reservation level for the Disco.

**Theorem 1 (Disco’s optimal contracting policy):** When \( s + G(g) > G(U'(0)) \), \( Q^*(s, g) = 0. \) Otherwise, \( Q^*(s, g) \) is determined by any of the following equivalent identities:

(i) \( Q^*(s, g) = (U')^{-1}(G^{-1}(s + G(g))) \);

(ii) \( E[\min(P_s, U'(Q^*(s, g)))] = s + E[\min(P_s, g)] \);

(iii) \( E(P_s - U'(Q^*(s, g)))^+ = s + E(P_s - g)^+ \);

(iv) \( U'(Q^*) = s + g + \int_g^{U'(Q^*)} F(P_s) dP_s \);

(v) \( G(U'(Q^*)) = G(g) = s. \)

**Proof:** When \( s + G(g) > G(U'(0)) \), the optimal contracting strategy for the Disco solving (7) is \( Q^*(s, g) = 0. \) Now we consider the case of \( s + G(g) \leq G(U'(0)) \). Using (3) we write the expected utility \( EV(Q) \) as
\[ EV(Q) = -sQ + \int_g^{\infty} (P_s - g) \min(Q, D) f(P_s) dP_s + \int_0^{\infty} [U(D) - P_s D] f(P_s) dP_s. \quad (11) \]

From Lemma 1, (11) can be written as
\[ EV(D(Q), Q) = -sQ + \int_g^{\infty} (P_s - g) Q dF(P_s) + \int_0^{U'(Q)} [U(D_s(P_s)) - P_s D_s(P_s)] dF(P_s) + \int_{U'(Q)}^{\infty} [U(Q) - P_s Q] dF(P_s). \]

Then the first order condition gives,
\[ \frac{\partial EV}{\partial Q} = -s + \int_g^{\infty} (P_s - g) dF(P_s) \]
\[ + \int_{U'(Q)}^{\infty} (U'(Q) - P_s)dF(P_s) \]
\[ = -s - \int_0^g P_s dF(P_s) - g[1 - F(g)] \]
\[ + \int_0^{U'(Q)} P_s dF(P_s) + U'(Q)[1 - F(U'(Q))] \]
\[ = -s - G(g) + G(U'(Q)) = 0. \quad (12) \]

The first equality of (12) is due to (8), the second equality (FOC) gives identity (v) in Theorem 1. Identities (i), (ii) and (iv) are direct consequences of identity (v) by using the “G” function (8). Use (10), we obtain identity (iii). Hence the proof.

The intuition underlying Theorem 1 is as follows. If the combined effective price per unit of capacity \( s + G(g) \) exceeds the threshold level of \( G(U'(0)) \), the Disco will not contract for any capacity at all, but will rely entirely on the spot market for supply. The upper bound \( s + G(g) \) represents the sum of the cost per unit to reserve capacity \( s \) plus the expected cost of using this capacity on the day \( G(g) \). When this combined effective price exceeds the maximum effective price the Disco is prepared to pay for any unit of capacity (recall that \( U'(Q) \) is decreasing in \( Q \)), then it is better to rely on the spot market for all purchases.

Concerning the case of positive contracting, consider for example identity (iv). The left-hand side is the Disco’s
Proof: Taking the derivatives w.r.t. \( s \) and \( g \) in identity (\( v \)) in Theorem 1,

\[
\frac{\partial Q^*}{\partial s} = \frac{1}{U'(Q^*)[1 - F(U'(Q^*))]} < 0,
\]

\[
\frac{\partial Q^*}{\partial g} = \frac{1}{U'(Q^*)[1 - F(U'(Q^*))]} < 0.
\]

The above two equalities imply (13).

Corollary 1: \( Q^* \) is monotonically decreasing in \( s \) and \( g \).

\( \forall Q^* > 0, \)

\[
\frac{\partial Q^*}{\partial g} = (1 - F(g)) \frac{\partial Q^*}{\partial s}. \tag{13}
\]

Example 1 (Numerical): Suppose \( P_s \) follows the Erlang (Gamma) distribution \( f(x) = \frac{1}{(2)^x}xe^{-x} \), and \( U(x) = 10(1 - e^{-10x}) \). Figure 1 shows two optimal contracting strategies (for iso-quants \( Q = 1, 2 \)) for various offers from the Genco.

Example 2 (Analytical): Assume the spot market price is uniformly distributed between \( P_d \) and \( P_u \), and the Disco’s WTP has the form \( U(x) = \frac{\alpha}{\gamma}(1 - e^{-\gamma x}) \). Assume

\[
g < \min[\alpha, P_u - \sqrt{2s(P_u - P_d)}] \quad \text{and} \quad s < \alpha, \tag{14}
\]

then the Disco’s optimal contracting strategy is

\[
Q^*(s, g) = \frac{-1}{\gamma} \ln \frac{\alpha}{P_u - \sqrt{(P_u - g)^2 - 4\sqrt{3s\sigma}}}, \tag{15}
\]

where \( \sigma \), the variance of the spot market price, is \( \sigma = \frac{P_u - P_d}{2\sqrt{3}} \).

If either of the inequalities is violated in (14), then \( Q^*(s, g) = 0 \).

2.2. The Genco’s problem

Henceforth we will take the Disco’s contracting strategy \( Q(s, g) \) to be the optimal strategy \( Q^*(s, g) \) given in Theorem 1. When no confusion is likely, we will omit the superscript “***”. Given this contracting strategy, the Genco is assumed to determine its optimal bidding strategy \([s^*, g^*]\) that maximizes its expected profit. In this subsection, we first derive the Genco’s profit function, then we derive Genco’s
profit maximization strategies anticipating the Disco’s reaction, last, we derive the condition for the existence of such an equilibrium for which \( Q > 0 \) and give one example. The profit function \( \Pi(s, g; P_s, Q(s, g)) \) for the Genco at \( P_s \) from the long-term contract and from the spot market is given as

\[
\Pi(s, g; P_s, Q(s, g)) = sQ(s, g) + gq - (\beta K + bq) + (P_s - b)^+m(K - q).
\]  

(16)

The first two terms on the right hand side of equation (16) represent the Genco’s revenue from the contract, the third term is the Genco’s cost of supplying \( q \) units to the Disco, and the fourth term is the Genco’s profit from the spot market (where the Genco will only sell if \( P_s \geq b \)). We assume that on the day the Genco can only sell a percentage \((0 \leq m \leq 1)\) of its residual output in the spot market. The risk factor \( m \), which is assumed to be fixed here, provides the incentive for the Genco to sign a contract with the Disco.

Concerning the risk factor \( m \), this would in practice be estimated from historical data, leading in stable markets to rational expectations about its value, which typically will vary from one Genco to another. There are several reasons why this factor may be less than unity. The spot market may be “thin” or not well organized, making it difficult to identify Discos on the day. There may be difficulties in arranging transmission, because of transmission capacity constraints or thinness in the transmission market. This latter case is the situation, for example, in many energy markets in which the transmission is fraught with congestion and uncertainty. We will see below (Corollary 4) that when \( m = 1 \), there will be no long-term contracting in equilibrium; all transactions will occur through the spot market. Thus, as \( m \) goes to unity, long-term contracts disappear. While we do not model the dynamics of the contracting process, it is natural to think of \( m \) as decreasing as the spot market becomes more liquid and contracting becomes short term. These are exactly the changes noted in ref. [1] in the natural gas market. As the short-term contracting market picked up steam in the late 1980s, more and more of producers’ capacity was contracted for in the short-term market, eventually driving the long-term (multi-year) contract market transactions to near zero. Quarterly and monthly contracting is now the rule of the day, coupled with various short-term risk management instruments. Presumably the reason that such contracting is not further displaced by the (very liquid) spot market in natural gas is that producers are not certain that they can find Discos on the day for all their output (i.e., \( m < 1 \)), just as modeled here.

Alternative explanations of risk factor \( m \) could be that the Genco is risk-averse and hedges by selling some at the current prices and the Disco agrees to reserve this capacity by the same token. Gencos attempt to assure reasonable utilization of their capacity through bilateral contracts and Discos, who face uncertainty in the spot market, attempt to hedge against price volatility through the same bilateral contracts. Thus, risk-averse behavior of either party might be another argument.

The Genco’s decision problem is the following:

\[
\text{Max}_{s, g} \mathbb{E}\Pi(s, g)
\]

\[
= \int \Pi(s, g; P_s, Q(s, g)) f(P_s) dP_s
\]

\[
= [s + (1 - m\chi(g-b))(g-b)(1 - F(g)) - m(P_s - G(\max(g, b)))] Q + mK(P_s - G(b)) - \beta K.
\]  

(17)

Define \( \epsilon_s(Q) \) as the demand elasticity w.r.t. to the subscription charge \([5]\),

\[
\epsilon_s(Q) = \frac{s}{Q} \left| \frac{\partial Q}{\partial s} \right| = -\frac{s}{Q} \frac{\partial Q}{\partial s}.
\]

**Theorem 2 (Genco’s optimal bidding policy):** Assume \( QU''(Q) + 2U''(Q) < 0 \), then the optimal bidding policy for the Genco is:

\[
s^* = \begin{cases} 
G(U'(Q)) - G(b) = \frac{mE(P_s - b)^+}{1 - \alpha}, & Q < K, \\
G(U'(K)) - G(b), & Q \geq K.
\end{cases}
\]

It should be noted that the rather technical condition assumed in Theorem 2 (viz. \( QU''(Q) + 2U''(Q) < 0 \)) is a sufficient condition for the global optimality of the expected profit function reached at \((s^*, g^*)\). This condition is satisfied, for example, if the Disco’s normal demand curve \( D_s(p) = (U')^{-1}(p) \) is linear or concave (since \( D_s \) is concave if \( U'' \leq 0 \)). Theorem 2 goes through for the most common convex demand case of constant elasticity demand. This is the case that derives from the WTP function \( U(x) = ax^\alpha \), where \( a > 0 \) and \( 0 < \alpha < 1 \). It gives rise to a constant elasticity of demand for the Retailer’s demand function \( D_s(p) \). For this case our sufficient condition in Theorem 2 holds, even though the demand function, \( D_s(p) = (a p)^{-\alpha} \) is clearly convex. It should be noted that the sufficient condition is satisfied for the class of convex demand functions used in examples 1-3 as well.

The conclusions of Theorem 2 are derived in standard fashion from first-order conditions. The concavity of the demand function is a typical condition for assuring concavity of expected profits but weaker sufficient conditions for optimality of the first-order conditions likely exist (as they do for the traditional one-good case). Interestingly, we do not
Theorem 2 indicates that the optimal bidding strategy for the Genco is to use the reservation charge \( s \) to extract margin from the Disco, setting the execution cost \( g \) as low as possible consistent with recovering unit variable operating costs \( b \). The inverse elasticity term in the expression for \( s^* \) reflects the usual tradeoff between price (positive with increasing price) and quantity (negative with increasing price) effects in pricing any good (see, e.g., [5]). The optimal reservation charge depends on the unit opportunity cost \((mE(P_s - b)^+)\) the Genco loses from the spot market due to capacity commitment to the Disco as well as the inverse elasticity of the Discos’ contract demand (reservation level). The key here is that margin is collected through the reservation charge, essentially because increasing \( g \) above its minimum level \( b \) decreases the anticipated value of the contract to the Disco at a faster rate than increases in \( s \). This is so because the Disco knows that it will only execute the contract on the day in those states of the world where \( P_s \) exceeds \( g \) so increases in \( g \) increase both the Disco’s cost of execution as well as the utility of the contract, as a consequence of the Disco’s behavior, and because of (13), any execution charge \( g \) deviating from the unit variable operating costs \( b \) decreases the Genco’s profit. A rigorous analysis of this can be found in the proof of Theorem 2 in [8].

Denote \( Q \) as the Genco-Disco joint equilibrium.

**Corollary 2 (Genco-Disco joint equilibrium):** \( Q \) is implicitly determined by:

\[
QU''(Q)[1 - F(U'(Q))] + G(U'(Q)) - mP_s - (1 - m)G(b) = 0. \tag{18}
\]

**Proof:** It is a direct consequence of using Theorem 1 and Theorem 2.

Denote \( K^* \) as optimal Genco-Disco joint investment policy.

**Corollary 3 (Genco-Disco joint investment policy):** Assume \( \beta > m(P_s - G(b)) \), then

\[
K^* = (U')^{-1}(G^{-1}(\frac{\beta}{1 - \gamma(k^*)} + G(b))) = Q(s^*, g^*),
\]

where the optimal bidding policy for the Genco is,

\[
s^* = \frac{\beta}{1 - \gamma(k^*)}, \quad g^* = b.
\]

**Corollary 4 (Riskless spot market for the Genco):** If \( m = 1 \) then \( Q = 0 \).

Corollary 4 is a direct consequence of Theorem 3 since \( m = 1 \) violates the condition. Corollary 4 says that when the Genco incurs no risk on the spot market \((m = 1)\), i.e., it can sell all its residual output in the spot market, then the Genco will not sign any contract with the Disco \((Q = 0)\). The impact of the risk factor \( m \) is further elucidated in the following example.

**Example 3 (Risky spot market for the Genco):** Suppose the spot market price distribution and the Disco’s WTP function are the same as in Example 1. Assume \( b = 5, \beta = 0 \).
1. $m = 0$, the Genco faces an extremely risky spot market. Its profit relies entirely on the contract. From (18), we obtain $Q^* = 0.440$. The Genco charges $g^* = 5$, $s^* = 0.0336$. The maximum overall profit the Genco gets is $\Pi_{\max} = 0.0148$.

2. $m = 0.1$, the Genco faces a less risky spot market. Its profit comes from both the contract and the spot market. We obtain $Q^* = 0.424, g^* = 5, s^* = 0.0349, \Pi_{\max} = 0.0128$.

Figure 2. The sensitivity of $Q$ to $m$, $0 \leq m \leq \frac{G(U'(b)) - G(b)}{P_s - G(b)} = 0.988$.

Example 3 shows how the changing of $m$ will change the behavior of the Disco as well as the Genco. We see that, as $m$ increases, the Genco faces less risky spot market, hence he would demand more by charging a high reservation fee, to offset its opportunity loss from the spot market. As a consequence, the Disco would reserve/contract less from the Genco due to such an increasing in reservation cost. The Genco losses part of its profit from the contract market, however, it gains back from the spot market, hence he would demand more by charging a high reservation fee, to offset its opportunity loss from the spot market.

2.3. Impact of transmission costs

Now we introduce transmission costs and show that the previous results hold with positive transmission costs. In many sectors, including the energy sector, such costs can be a significant proportion of total landed cost, so their inclusion here is certainly warranted. We again assume the standard two-goods pricing structure, now for transmission. The first good is transmission capacity and the second is transmission use. In some sectors, the cost of transmission is bundled in with the Genco’s bid price; in others, it is separately negotiated. In any case, it should be intuitively clear that the impact of transmission cost will be simply to add fixed amounts (the transmission costs) to the Genco’s unbundled bid prices $[s, g]$. As we will see, this intuition is essentially correct.

Denote $h$ as the transmission price the Disco pays to the transmission company for shipping output from the Genco to the Disco’s door, $h_0$ the transmission price for shipping output from the spot market to the Disco’s door. Denote the transmission cost difference between these two different routes $h - h_0$ as $\Delta h$. In addition to the actual transmission usage charge ($h, h_0$), there is a fixed charge ($t$) from the transmission company to the Disco for reserving transmission capacity for the Disco.

The Disco’s optimal contracting will be affected by the introduction of transmission costs. Its revised demand curve is defined as $D(P_s + h_0, g + h, Q)$, and its revised optimal contracting is defined as $Q(s + t, g + h, h_0)$ accordingly. The specific analytical solutions are as follows.

The Disco’s revised total demand $D(P_s + h_0, g + h, Q)$ is given as the solution to:

$$\max_{D \geq 0} [U(D) - (P_s + h_0)D + (P_s - g - \Delta h)^+ \min[D, Q] - (s + t)Q].$$

Comparing (5) and (19), it is apparent that the effect of transmission costs $(h, h_0, t)$ is obtained by replacing the spot price $P_s$ by the price $P_s + h_0$, by replacing the reservation cost $s$ by the cost $s + t$, by replacing the Disco’s demand curve $D(P_s, Q)$ by $D(P_s + h_0, Q)$ and by replacing the Disco’s optimal contracted capacity $Q(s, g)$ by $Q(s + t, g + h, h_0)$ with $D(P_s, Q)$ as given in Lemma 1 and $Q(s, g)$ as given in Theorem 1. For Lemma 1, Theorem 1 and Corollary 1, the results are direct. For Theorem 2, the result of these substitutions is the following.

Theorem 2a (Optimal bidding with transmission costs): Assume $QU''(Q) + 2U''(Q) < 0$ and $\Delta h > 0$, then the optimal bidding policy for the Genco with transmission cost is the following:

$$s^* = \frac{m(\bar{P}_s - G(b + \frac{\Delta h}{1 - m}))}{1 - c(Q)}$, $g^* = b + \frac{m}{1 - m}\Delta h$.

Notice that, from Theorem 2a, $g^* = b$ is no longer opti-
nal when transmission costs are non-zero. $g^*$ depends on $m$ and on the transmission costs difference $\Delta h$. Specifically, $g^*$ increases as $m$ or $\Delta h$ goes up. The Genco increases its execution charge $g^*$ from its unit operating cost to cover the transmission costs difference between delivering from the Genco and from the spot market, weighted by a positive factor $\frac{m}{1-m}$ that characterizes its accessibility to the spot market.

3. Extensions to multiple Discos

This section extends our previous results to the more general case where there are multiple Discos purchasing from the same Genco, and using a radial transmission network. The assumption of a radial supply chain implies that no economies of scale or routing are reflected in the transmission cost structure.

Assume there are $J$ Discos (or customers) with demand functions $D_j(P_s, g + h_j, t_j, Q_j)$, $j = 1, \ldots, J$. Denote $h_j$ as the transmission price per unit from the Genco to Disco $j$. Denote $t_j$ as the transmission capacity reservation cost for Disco $j$. Assume the transmission network connecting the Genco and Discos is radial.

We assume that the Discos, the transmission company and the Genco exchange bids through an “electronic bulletin board” on which they post bids and offers until they have reached an equilibrium. The offers by the Genco are simply the Genco’s reservation charge $s$, contract execution cost $g$, and possibly state-dependent cost functions for the Gencos’ bids. The Transmission Company posts the transmission charges $h_j$ and $t_j$ (which are assumed arbitrary but fixed in this analysis) and the transmission capacity reservation cost $t_j$. Discos post their demands at the indicated contract prices and the market clears when no further adjustment occurs. We again assume that the Genco is a von Stackelberg leader, anticipating the demands of the Discos when it posts $[s, g]$. Each Disco determines its optimal contract $Q_j(s + t, g + h_j, h_{o_j})$ using revised version of Theorem 1 (where transmission costs are added) once they see the cost structure $(s, g, t_j, h_j, h_{o_j})$ on the electronic bulletin board. Based on our earlier results on the structure of Disco demand and contracting decisions, we obtain the following

Theorem 2b (Optimal bidding with multiple Discos): Assume the Transmission Company charges all Discos the same $(P_s = G(b + \frac{\Delta h}{1-m}))$ for $j = 1, \ldots, J$. Then the Genco’s optimal bidding policy when facing multiple Discos is,

$$s^* = \frac{m(P_s - G(b + \frac{\Delta h}{1-m}))}{1 - \frac{1}{v_j}}, \quad g^* = b + \frac{m}{1-m} \Delta h.$$

Practically, what Theorem 2b says is that the optimal bidding policy when the Genco facing multiple Discos is essentially the same as facing one Disco, treating all Discos as one aggregated Disco, by replacing $Q_j$ in Theorem 2a with aggregated demand $Y = \sum_{j=1}^{J} Q_j$.

From Theorem 2b, and revised identity (v) in Theorem 1 $G(U_j(Q_j) - h_{o_j}) - G(g + \Delta h) = s + t$, we note that if Discos have heterogeneous demand functions (i.e., different WTP functions $U_j$), then the transmission company can (if it has market power) benefit from this by charging a discriminating transmission price from the spot market to the Disco’s door, according to the following pricing rule:

$$h_{o_j} = U_j(Q_j) - C,$$

where $C = G^{-1}[s + t + G(b + \frac{\Delta h}{1-m})]$ is a constant that is independent of $j$. This rule says that the transmission company can charge higher transmission costs for high value Discos for delivery from the spot market to those Discos. As noted above, we assume transmission charges are fixed in this analysis. But the results show that if competition does not prevail in the transmission sector, regulation of transmission prices may well be required.

4. Summary

We have further generalized the above results to the case of multi-Genco, multi-Disco general network [9]. What we show there is that when Gencos accurately anticipate demands to their bids, then a dominant strategy for their contract execution price is for each Genco to truthfully reveal its variable cost, with subscription fees determined in interaction with Discos to trade off the risk of underutilized capacity against unit capacity costs. Discos’ optimal portfolios are shown to follow a merit order or greedy shopping rule, under which contracts are signed following a newsvendor strategy in order of a certain index of the capital intensity of Gencos’ bids. The results reported in this paper serve as the theoretical foundation for this overall program of research.

Besides the extensions noted to the multi-Genco, multi-Disco case, the results of this paper suggest a number of important directions for continuing this research. On the theoretical side, these results need to be generalized to allow for state-dependent Willingness-to-Pay functions by Discos and possibly state-dependent cost functions for the Gencos. In this regard, the foundational Lemma 1 continues to hold (at the corresponding state-dependent utility functions
but the conditions for optimal contracting policies are more complex.

A second theoretical extension concerns results for closed spot markets in which Gencos or Discos may have market power and therefore may have some influence on spot market price.

On the empirical side, it will be interesting to explore the increasing use of bilateral contracting in energy and service markets to exploit these results in understanding the structure of multi-tiered markets, especially those in which financial intermediaries play an important role.

Finally, while the above analysis employs a traditional analytical approach, it will be interesting to explore the use of electronic agents to develop near-optimal contracting and bidding strategies [6, 7] in more complex environments (e.g. those involving more complex contracting of service quality features than can be represented in an analytical model). The theoretical results presented here can provide a useful benchmark for this analysis, which would allow the exploration of learning and computational strategies in these more realistic environments.

References


