

SUPPLY FUNCTION EQUILIBRIUM IN A CONSTRAINED TRANSMISSION SYSTEM

ROBERT WILSON

ABSTRACT. This paper characterizes equilibrium in an auction market constrained by limited capacities of links in a transportation network. The formulation is adapted to a wholesale spot market for electricity managed by the operator of the transmission system.

This paper derives conditions that characterize equilibrium in an auction market of the kind conducted by system operators in the electricity industry. These are wholesale spot markets in which the participants are suppliers (generators) and demanders (utilities and other load-serving entities). Because these participants are spatially distributed, the operator's allocation of production and consumption of electrical energy is constrained by the capacities of links in the transmission system. Moreover, storage is infeasible, supply must continually match demand, and both net demand and transmission capacities are affected by random shocks. Therefore, participants submit notional supply or demand functions in advance and then in each contingency the operator uses these functions to determine an optimal allocation.¹

Financial settlements in such markets use locational marginal pricing, also called nodal pricing. The operator chooses the allocation in each contingency to maximize the apparent gain from trade subject to the feasibility constraints imposed by limited transmission capacities. (This is the gain from trade "as bid" since the operator treats each supply function as though it reflects the actual marginal cost of production.) This optimization results in a vector $\lambda = (p, \mu)$ of Lagrange multipliers on the energy and capacity constraints. Then a supplier j for which each unit of energy output uses u_{ij} units of capacity on transmission link i is paid its "nodal" price $p_j = p - \sum_i \mu_i u_{ij}$. Thus, if j submitted the supply function $s_j(\cdot)$ then it is assigned to produce $s_j(p_j)$ units of energy and for this output it is paid $p_j s_j(p_j)$.

The formulation is established in Section 1. Sections 2 and 3 characterize a firm's optimal bidding strategy. In Section 2 this is done for a firm located at a single node in the network. Section 3 addresses the case that a firm controls supply resources located at multiple nodes in the network. Section 4 then applies these results to characterize an equilibrium among the firms. This is an ordinary Nash equilibrium and thus includes "rational expectations" in the sense that each firm anticipates correctly the bidding strategies of all other firms. This is a plausible approximation in wholesale electricity markets because the participants and their costs are known to all, as is the probability distribution of exogenous factors (random variations of demand and of transmission and generator capacities), and the market repeats every hour of every day. Since a few systems settle transactions using a pay-as-bid rule, Section 5 characterizes equilibrium bidding strategies

Date: 15 April 2005; corrected April 30. This is a draft — comments welcome!

I am indebted to Pär Holmberg for comments on previous drafts.

¹Similar markets are used in other industries, such as gas transmission, but in these industries storage is an important factor that is ignored here. In practice, there are additional constraints that are not addressed by our formulation, such as requirements for reserves (to sustain voltage and to protect against cascading failures of equipment) and dynamic constraints (e.g., "ramp rate" limits on the rate-of-change of generators' outputs). There are also additional financial aspects (e.g., the operator charges a network management fee, typically in the range of 1%-3% of the energy price, including costs of reserves) and fixed costs of starting-up and operating a generator that we ignore.

when this rule is used. Section 6 concludes by showing that an alternative formulation enables a supplier to estimate its optimal supply function using market data.

1. FORMULATION

There are several firms (either generators/suppliers or utilities/demanders) indexed by $j = 1, \dots, n$. For notational simplicity we treat demand as negative supply; thus, if firm j is a demander with demand function D_j then its supply function is $s_j = -D_j$. Except in Section 3, each firm is assumed to inject (or extract) power at a single node in the network.

There are also several capacity constraints indexed by $i = 1, \dots, m$. We assume throughout that the power transfer distribution factors [PTDFs] of the transmission network are fixed and known to all participants. Thus, let u_{ij} be the usage of link i required by a unit injection of energy by firm j at its node in the network. There is also a total energy constraint indexed by $i = 0$. For simplicity we ignore thermal losses, so each $u_{0j} = -1$; that is, a unit output of supplier j relaxes the energy constraint by one unit.² Let $a_{ij} = -u_{ij}$ and denote the matrix of all such factors by $A = (a_{ij})$. Since the column of j 's distribution factors is A_j , firm j 's nodal price is $p_j = \lambda A_j$ when the market clearing price of energy is $p = \lambda_0$ and the marginal value of enlarging the capacity of link i is λ_i .

Before the spot market opens, each firm submits a supply function s_j indicating its offered supply $s_j(p_j)$ at the nodal price p_j for injection at its node. We assume that each firm's supply function takes account of its own local capacity constraints; e.g., if supplier j can supply at most K_j then necessarily $s_j(p_j) \leq K_j$. We assume further that each supplier must offer its entire capacity; that is, $s_j(p_j) = K_j$ at every price p_j above a sufficiently high price p_j^* . This reflects the "must-offer" obligation that operators impose on suppliers to comply with regulatory mandates.

When the spot market opens the operator knows the realization of the shocks affecting the vector b of net energy demand and the capacities of transmission links. Therefore, it chooses a vector $\lambda = (\lambda_i)_{i=0,1,\dots,m}$ of energy and capacity prices that are the marginal values (Lagrange multipliers) of relaxing the constraints. The optimality conditions for the operator's allocation decision reduce to the requirement that the vector λ of marginal prices must satisfy the feasibility constraints:

$$\sum_j A_j s_j(\lambda A_j) \leq b, \quad \text{and} \quad \lambda \geq 0,$$

and the complementarity condition:

$$\lambda_i > 0 \quad \text{only if} \quad \sum_j a_{ij} s_j(\lambda A_j) = b_i.$$

We assume throughout that the number of firms (and their own capacities) is sufficiently large that the operator can satisfy the feasibility constraints; in particular, the number n of firms exceeds the number of binding constraints in every likely contingency. We also assume that the energy constraint is always binding; that is, $\sum_j s_j(p_j) = b_0$.

²The PTDFs are derived from the linear approximation of Kirchhoff's Laws obtained by assuming that phase angles are zero; cf. Chao and Peck (1996) and Chao, Peck, Oren, and Wilson (2000). In the engineering literature this is called the direct current approximation of an alternating current system. The PTDFs depend only on the topology of the network and the impedances of the transmission links. Some systems use this approximation as a standard operating procedure. In principle, thermal losses introduce quadratic terms but many systems rely on linear approximations, which is consistent with our formulation. Note that if u_{ij} is positive for an injection then it is negative for an extraction; also, if it is positive for flow along link i in one direction then it is negative for flow in the opposite direction.

Assuming that firm j is a supplier, let $C_j(q_j)$ and $c_j(q_j) = C'_j(q_j)$ be its total and marginal cost if it supplies output q_j . (For a demander the analogs are the negatives of its total and marginal value of consumption.) We assume that c_j is nonnegative, nondecreasing, and differentiable. Because our formulation is intended to model a spot market, it is important to recognize the role of firms' forward contracts on their financial positions and productive capabilities. Thus, each firm's cost function and available capacity should be interpreted as net of its forward contracts. Similarly, the operator's constraints are net of the aggregate flows implied by bilateral contracts; indeed, this is the treatment of forward contracts in most operators' spot markets.

We study an equilibrium among firms' supply functions that are differentiable except where some firm's own capacity constraint becomes binding. In previous work, Holmberg (2004, 2005a,b) has shown in a more restrictive formulation that piecewise-differentiability is implied by fundamental considerations. For this paper we rely on the presumption that his analysis can be extended to the more general formulation used here, and therefore we use the techniques of the calculus of variations (Elsgolc, 1962).

1.1. The Induced Distribution of Nodal Prices. To analyze a firm's bidding problem, we first derive the probability distribution of λ and hence the nodal price p_j of each firm j . For this we assume that all firms know the probability distribution F of the vector b of shocks realized in the spot market. Further, we assume that F has a density function f that is differentiable with support that is a convex full-dimensional subset of \mathbb{R}^{m+1} . (Actually, it need not be that the total probability is 1, since in those systems where each supplier submits a single supply function for each day, F and f can be interpreted as including the daily cycle of variation over the day.) As mentioned, each supply function s_j is piecewise-differentiable.

Given firms' supply functions, each realization b of the shock vector determines the subset I of constraints that are binding in the operator's optimization. Let b_I be the sub-vector of shocks for the binding constraints, let A_I be the corresponding submatrix, and let F_I and f_I be the marginal distribution and density functions of this sub-vector. Then from the operator's feasibility constraint

$$\sum_j A_{Ij} s_j(\lambda_I A_{Ij}) = b_I$$

one derives the Jacobian matrix $\lambda'_I(b_I)$ of partial derivatives of λ_I with respect to the components of b_I via implicit differentiation to obtain the relation

$$\left[\sum_j s'_j(\lambda_I A_{Ij}) A_{Ij} \cdot A_{Ij}^T \right] \cdot \lambda'_I(b_I) = \text{Id}_I,$$

which holds on each open set in the domain of shocks for which I is the set of binding constraints and the supply functions are differentiable. In this formula, Id_I is the identity matrix and the square matrix $A_{Ij} \cdot A_{Ij}^T$ is the outer product of the column vector A_{Ij} with its transpose, the row vector A_{Ij}^T . Therefore, the probability density $f_I(b_I)$ at the shock b_I induces the probability density $g_I(\lambda_I)$ at λ_I ,

$$g_I(\lambda_I) = f_I \left(\sum_j A_{Ij} s_j(\lambda_I A_{Ij}) \right) \left| \sum_j s'_j(\lambda_I A_{Ij}) A_{Ij} \cdot A_{Ij}^T \right|,$$

or $g_I \equiv 0$ if there is zero probability that I is the set of binding constraints. In this formula $|B_I|$ denotes the determinant of the matrix

$$B_I = \sum_j s'_j (\lambda_I A_{Ij}) A_{Ij} \cdot A_{Ij}^T.$$

This determinant is a multilinear function of (s'_j) ; that is, it is a linear function of each s'_j separately. For example, if

$$A_I = \begin{pmatrix} 1 & 1 & 1 \\ 1/3 & -1/3 & 0 \end{pmatrix},$$

then $|B_I| = (4s'_1 s'_2 + s'_3)/9$. The general fact that this determinant is multilinear can be proved using the transformation introduced in Section 2.

Analogous formulas pertain to each subset I of binding constraints, and thus to the corresponding domain of shocks for which this subset is binding, and the induced domain of λ_I . These domains partition the space of shocks and the space of marginal values; however, they need not induce a partition of the space of nodal prices.

2. CHARACTERIZATION OF A FIRM'S OPTIMAL SUPPLY FUNCTION

In this section we consider the bidding problem of one firm j given the supply functions of other firms. The formulation in this section is complemented by an alternative formulation in Section 6. We use k to index those firms other than j , and ℓ to index all firms.

Because firm j is paid its nodal price, its realized profit contribution is

$$\Pi_j(p_j, q_j) = p_j q_j - C_j(q_j) \quad \text{where} \quad p_j = \lambda A_j \quad \text{and} \quad q_j = s_j(p_j).$$

However, both p_j and q_j are uncertain when it submits its supply function s_j , except for the known fact that $q_j = s_j(p_j)$. Therefore its objective is to maximize its expected profit, taking account of its potential to affect the induced probability distribution of its nodal price.

Because j 's nodal price is $p_j = \lambda_o + \sum_{i>0} \lambda_i a_{ij}$ and k 's nodal price is $p_k = \lambda_o + \sum_{i>0} \lambda_i a_{ik}$, it follows that $p_k = p_j + \sum_{i>0} \lambda_i [a_{ik} - a_{ij}]$. Similarly, the operator's feasibility constraint $\sum_{\ell} A_{\ell} s_{\ell} \leq b$ can be represented from j 's viewpoint by adding appropriate multiples of row $i = 0$ of A to the other rows (as in Gaussian elimination) to obtain the equivalent form $\sum_{\ell} A_{\ell}^j s_{\ell} = b^j$, where $a_{\circ j}^j = 1$ and $a_{i j}^j = 0$ for each $i > 0$, and for each other firm $a_{\circ k}^j = 1$ and $a_{i k}^j = a_{i k} - a_{i j}$ for each $i > 0$; also, $b_{\circ}^j = b_{\circ}$ and $b_i^j = b_i - a_{i j} b_{\circ}$. This transformation has no effect on the determinant $|B_I|$ in the previous formula for $g_I(\lambda_I)$, but it alters the interpretation of λ_I . In particular, $p_j = \lambda_o$ and each $p_k = p_j + \sum_{i>0} \lambda_i a_{i k}^j$. Thus, from j 's perspective it is only the other suppliers that are charged for scarce transmission capacity. This transformation is equivalent to designating node j as a "trading hub" in the parlance of power markets.

In the remainder of this section we let $\lambda = (p, \mu)$ where $p = p_j$. We use f_I^j to denote the induced marginal density of b_I^j when I is the set of binding constraints. Let P be the interval that is the support of p , and let $M(I, p)$ be the support of μ_I given I and p . The following lemma will be useful later. We use the shorthand notation $d\mu_I \equiv \prod_{i \in I} d\mu_i$.

Lemma 2.1.

$$\left(\frac{\partial}{\partial s_j} - \frac{d}{dp} \frac{\partial}{\partial s'_j} \right) \sum_I \int_{M(I,p)} f_I^j \left(\sum_{\ell} A_{I\ell}^j s_{\ell}(\lambda_I A_{I\ell}^j) \right) \left| \sum_{\ell} s'_{\ell}(\lambda_I A_{I\ell}^j) A_{I\ell}^j \cdot A_{I\ell}^{jT} \right| d\mu_I = 0.$$

Proof. The Euler condition for maximizing

$$\int_P \left(\sum_I \int_{M(I,p)} f_I^j \left(\sum_{\ell} A_{I\ell}^j s_{\ell}(\lambda_I A_{I\ell}^j) \right) \left| \sum_{\ell} s'_{\ell}(\lambda_I A_{I\ell}^j) A_{I\ell}^j \cdot A_{I\ell}^{jT} \right| d\mu_I \right) dp,$$

by choosing s_j is the equation in the lemma. But this objective function is just the total probability $\int f^j(b^j) db^j = 1$, which is unaffected by firm j 's supply function s_j . Therefore, the Euler condition is satisfied identically. \square

We suppose that when choosing its supply function s_j , firm j 's objective is to maximize the expectation of its profit contribution $\Pi_j(p, q_j)$ given the supply functions of all other firms. This expectation is

$$E\{\Pi_j(p, s_j(p))\} = \int_P [ps_j(p) - C_j(s_j(p))] \sum_I \int_{M(I,p)} f_I^j \left(\sum_{\ell} A_{I\ell}^j s_{\ell}(\lambda_I A_{I\ell}^j) \right) \left| \sum_{\ell} s'_{\ell}(\lambda_I A_{I\ell}^j) A_{I\ell}^j \cdot A_{I\ell}^{jT} \right| d\mu_I dp.$$

In this formula each determinant can be written as

$$|B_I^j| \equiv \left| \sum_{\ell} s'_{\ell}(\lambda_I A_{I\ell}^j) A_{I\ell}^j \cdot A_{I\ell}^{jT} \right| = D_I^j + \delta_I^j s'_j(p).$$

Specifically,

$$D_I^j = \left| \sum_k s'_k(\lambda_I A_{Ik}^j) A_{Ik}^j \cdot A_{Ik}^{jT} \right|,$$

and δ_I^j is the cofactor of the element $\sum_{\ell} s'_{\ell}$ in position $(0, 0)$ of B_I^j ; viz.,

$$\delta_I^j = \left| \sum_k s'_k(\lambda_I A_{Ik}^j) A_{I^{\circ}k}^j \cdot A_{I^{\circ}k}^{jT} \right|,$$

where $I^{\circ} = I \setminus \{0\}$. As with $|B_I^j|$, each of these determinants is multilinear. Note that D_I^j is the same as $|B_I^j|$ except that firm j is omitted, and δ_I^j is the same as D_I^j except that the row and column of the energy constraint are omitted. In the special case that no transmission constraints are binding,

$$I = \{0\}, \quad D_I^j = \sum_k s'_k(p), \quad \text{and} \quad \delta_I^j \equiv 1.$$

Theorem 2.2. *The Euler condition for optimality of firm j 's supply function is*

$$s_j(p) = [p - c_j(s_j(p))] E \left\{ \sum_I D_I^j \mid p \ \& \ \sum_{\ell} s_{\ell}(p + \mu A_{\ell}^j) = b_{\circ} \right\}.$$

Proof. Write the integrand of j 's expected profit $E\{\Pi_j(p, s_j(p))\}$ as

$$G_j(p, s_j(p), s'_j(p)) = [ps_j(p) - C_j(s_j(p))] \sum_I \int_{M(I,p)} f_I^j \left(\sum_{\ell} A_{I\ell}^j s_{\ell}(\lambda_I A_{I\ell}^j) \right) \left| \sum_{\ell} s'_{\ell}(\lambda_I A_{I\ell}^j) A_{I\ell}^j \cdot A_{I\ell}^{jT} \right| d\mu_I.$$

Then for prices at which j 's capacity constraint is not binding (i.e., $s_j(p) < K_j$) and all supply functions are differentiable, the Euler condition is (omitting arguments of functions)

$$\begin{aligned}
0 &= \left(\frac{\partial}{\partial s_j} - \frac{d}{dp} \frac{\partial}{\partial s'_j} \right) G_j(p, s_j(p), s'_j(p)) \\
&= [p - c_j] \sum_I \int_{M(I,p)} f_I^j [D_I^j + \delta_I^j s'_j] d\mu_I + [ps_j + C_j] \sum_I \int_{M(I,p)} f_I^j [D_I^j + \delta_I^j s'_j] d\mu_I \\
&\quad - \left(s_j \sum_I \int_{M(I,p)} f_I^j \delta_I^j d\mu_I + [p - c_j] s'_j \sum_I \int_{M(I,p)} f_I^j \delta_I^j d\mu_I + [ps_j - C_j] \frac{d}{dp} \int_{M(I,p)} \sum_I f_I \delta_I^j d\mu_I \right) \\
&= -s_j \sum_I \int_{M(I,p)} f_I^j \delta_I^j d\mu_I + [p - c_j] \sum_I \int_{M(I,p)} f_I^j D_I^j d\mu_I \\
&\quad + [ps_j - C_j] \left(\frac{\partial}{\partial s_j} - \frac{d}{dp} \frac{\partial}{\partial s'_j} \right) \sum_I \int_{M(I,p)} f_I^j [D_I^j + \delta_I^j s'_j] d\mu_I \\
&= -s_j \sum_I \int_{M(I,p)} f_I^j \delta_I^j d\mu_I + [p - c_j] \sum_I \int_{M(I,p)} f_I^j D_I^j d\mu_I,
\end{aligned}$$

where the last equality applies Lemma 2.1. Therefore,

$$s_j = [p - c_j] \frac{\sum_I \int_{M(I,p)} f_I^j D_I^j d\mu_I}{\sum_I \int_{M(I,p)} f_I^j \delta_I^j d\mu_I}.$$

Now for each set I of binding constraints

$$\int_{M(I,p)} f_I^j \delta_I^j d\mu_I = \int_{M(I,p)} f_I^j(b_\circ, b_{I^\circ}) db_{I^\circ}$$

where $b_\circ \equiv \sum_\ell s_\ell$; that is, $\sum_I \int_{M(I,p)} f_I^j \delta_I^j d\mu_I$ is the marginal probability of the net energy shock b_\circ inferred from the requirement that it is met by the total supply $\sum_\ell s_\ell$. Hence at the price p ,

$$\frac{\sum_I \int_{M(I,p)} f_I^j D_I^j d\mu_I}{\sum_I \int_{M(I,p)} f_I^j \delta_I^j d\mu_I} = E\left\{ \sum_I D_I^j \mid p \ \& \ \sum_\ell s_\ell = b_\circ \right\}.$$

Note that this expectation is over all the variables $(\mu_i)_{i \in I^\circ}$ and hence also I via their shared dependence on b^j , but conditional on j 's nodal price p and all the supply functions. This completes the proof. \square

In this proof we use f_I^j throughout for clarity, but the superscript j is actually extraneous since in fact the probability density f_I of b_I is equivalent to the probability density f_I^j of b_I^j since $b_\circ^j = b_\circ$ and $b_i^j = b_i - A_{ij} b_\circ$ for $i > 0$.

If there are no transmission constraints then Theorem 2.2 specializes to the Euler condition

$$s_j(p) = [p - c_j(s_j(p))] \sum_k s'_k(p)$$

derived in previous studies of supply function equilibria; e.g., Klemperer and Meyer (1989) and Holmberg (2004, 2005ab). In this special case the Euler condition is independent of the probability distribution of shocks.

2.1. Transversality Conditions. The determination of an optimal supply function is completed by invoking transversality conditions. Assume that $s_j(p) = 0$ for those prices $p < p_*$ and $s_j(p) = K_j$ for those prices $p > p^*$. In particular, $s'_j(p) = 0$ in these two intervals of prices. Then the formula for j 's expected profit contribution is

$$E\{\Pi_j(p, s_j(p))\} = \int_{p_*}^{p^*} G_j(p, s_j(p), s'_j(p)) dp + \int_{p^*}^{\bar{p}} [pK_j - C_j(K_j)] \sum_I \int_{M(I, p)} \hat{f}_I^j \left(\sum_{\ell} A_{I\ell}^j s_{\ell}(\lambda_I A_{I\ell}^j) \right) \hat{D}_I d\mu_I dp.$$

This formula assumes that $p^* < \bar{p}$ if the operator imposes a bid cap \bar{p} . Also, we have written the second integral using \hat{f}_I^j and \hat{D}_I in case the sets of other suppliers that have not exhausted their capacities differs at prices below and above p^* .

Assume that there is a positive probability that firm j 's capacity will be exhausted; that is, its nodal price might exceed p^* .³ Then the transversality condition at p^* is

$$\begin{aligned} 0 &= \left(G_j - s'_j \frac{\partial}{\partial s'_j} G_j \right) - [p^* K_j - C_j(K_j)] \sum_I \int_{M(I, p^*)} \hat{f}_I^j \hat{D}_I d\mu_I \\ &= [p^* K_j - C_j(K_j)] \sum_I \int_{M(I, p^*)} [f_I^j D_I - \hat{f}_I^j \hat{D}_I] d\mu_I. \end{aligned}$$

If $\hat{f}_I^j \hat{D}_I = f_I^j D_I$ for every I then this condition is satisfied identically. However, Holmberg (2004) studies the case that the suppliers are symmetric and therefore all exhaust their capacities at the same price, which in his formulation is the price cap \bar{p} because demand is completely inelastic. Holmberg (2005a) also studies an asymmetric formulation with constant marginal costs and no transmission constraints. In this case, the suppliers can be ordered by the energy prices at which they exhaust their capacities; e.g., at sufficiently high prices one is a monopolist for the residual demand, over the next interval of lower prices two suppliers are duopolists for the residual demand, etc.

Supplier j can choose both its minimum price p_* and its quantity $s_j(p_*)$ offered at that price. Therefore there are two transversality conditions at p_* . The transversality condition for the optimal choice of p_* is

$$\begin{aligned} 0 &= G_j - s'_j \frac{\partial}{\partial s'_j} G_j \\ &= [p_* s_j - C_j(s_j)] \sum_I \int_{M(I, p_*)} f_I^j D_I d\mu_I. \end{aligned}$$

and the transversality condition for the optimal choice of $s_j(p_*)$ is

$$\begin{aligned} 0 &= \frac{\partial}{\partial s'_j} G_j \\ &= [p_* s_j - C_j(s_j)] \sum_I \int_{M(I, p_*)} f_I^j \delta_I d\mu_I. \end{aligned}$$

Either of these two conditions implies that

$$p_* s_j(p_*) = C_j(s_j(p_*)).$$

³As emphasized by Holmberg (2004), this implies that the operator has some scheme for rationing or curtailing excess demand when all supply capacity is exhausted. We ignore this aspect in cases where some strategic bidders are demanders.

That is, at the lowest nodal price for which j offers a positive supply, the profit contribution from the optimal supply is nil. If the supplier incurs no fixed cost then $p_* = c_j(0)$ and $s_j(p_*) = 0$. However, in the usual case the situation is more complicated. The supplier incurs fixed setup and operating costs and therefore its average cost $C_j(q)/q$ declines from an initial value of $+\infty$ as q increases from zero. Also, a generator typically has a minimum safe operating rate, say q_* , which requires that $s_j(p_*) \geq q_*$. Further, if there is a bid cap \bar{p} then it requires that $p_* \leq \bar{p}$. In practice, most operators circumvent these difficulties by imposing a must-bid obligation over the full range $[q_*, K_j]$ and then compensating by paying to a supplier any portion of its fixed costs not recovered from operating revenues. Therefore, for the analysis in the next section we assume that unrecoverable fixed costs are nil, and thus that $p_* = c_j(q_*)$ and $s_j(p_*) = q_*$.

2.2. Monotonicity Constraint. In practice an operator's software is typically designed to accept only a supply function that is nondecreasing and described by a limited number of linear segments. Baldick and Hogan (2002) emphasize the potential relevance of these constraints. Here we mention briefly the amendment to the Euler condition that ensures monotonicity.

To ensure that the constraint $s'_j(p) \geq 0$ is satisfied one uses a Lagrange multiplier, say $\varphi_j(p)$, and adjoins the term $\varphi_j(p)s'_j(p)$ to the integrand. Then the Euler condition is

$$\begin{aligned} 0 &= \left(\frac{\partial}{\partial s_j} - \frac{d}{dp} \frac{\partial}{\partial s'_j} \right) [G_j(p, s_j(p), s'_j(p)) + \varphi_j(p)s'_j(p)] \\ &= \left(\frac{\partial}{\partial s_j} - \frac{d}{dp} \frac{\partial}{\partial s'_j} \right) G_j(p, s_j(p), s'_j(p)) - \varphi'_j(p) \end{aligned}$$

Since $\varphi_j(p) = 0$ for each p where $s'_j(p) > 0$, this condition is equivalent to the requirement that for each interval (p_1, p_2) where the unconstrained Euler condition implies $s'_j < 0$ one selects a wider interval $[p_1^o, p_2^o] \supset (p_1, p_2)$ such that $s_j(p_1^o) = s_j(p_2^o)$ and over which

$$\int_{p_1^o}^{p_2^o} \left(\frac{\partial}{\partial s_j} - \frac{d}{dp} \frac{\partial}{\partial s'_j} \right) G_j(p, s_j(p), s'_j(p)) dp = 0$$

and then the optimal constrained supply function is $s_j(p) = s_j(p_1^o)$ for every price $p \in [p_1^o, p_2^o]$. This procedure is called "ironing" of the unconstrained supply function to ensure optimality subject to the monotonicity constraint. The theory is well developed and routinely applied in the theory of nonlinear pricing; e.g., Wilson (1993).

3. MULTIPLE SUPPLY FUNCTIONS

In this section we extend the characterization in Section 2 to the case that a firm controls supply resources located at multiple nodes in the transmission network. In this case we identify a firm with the set J of nodes at which its resources are located. Its realized profit contribution is therefore assumed to be

$$\Pi_J = \sum_{j \in J} p_j q_j - C_j(q_j) \quad \text{where} \quad p_j = \lambda A_j \quad \text{and} \quad q_j = s_j(p_j),$$

depending on the realized prices $(p_j)_{j \in J}$ at its nodes. Firm J 's bidding strategy thus specifies the collection $(s_j)_{j \in J}$ of supply functions that it submits to the operator. These supply functions are chosen to maximize its expected profit contribution $E\{\sum_{j \in J} p_j q_j - C_j(q_j)\}$, taking the supply functions of all other firms as given. The characterization therefore involves an Euler condition for each of J 's supply functions, together

with the associated transversality conditions. The derivation of the transversality conditions is similar to the one in subsection 2.1 so we focus on the Euler conditions. The key difference is that now the firm takes account of the combined effect of all its supply functions on all its nodal prices.

In the following we use $j \in J$ to indicate one of J 's nodes, h to index all of its nodes ($h \in J$), k to index nodes of other firms, and ℓ to index all nodes.

Theorem 3.1. *The Euler condition for optimality of firm J 's supply function s_j is*

$$s_j(p_j) = E\{[p_j - c_j(s_j(p_j))] \sum_I D_I^j - \sum_{h \neq j} [s_h(p_h) + (p_h - c_h(s_h(p_h)))s'_h(p_h)] \mid p_j \text{ \& } \sum_\ell s_\ell = b_o\}.$$

Proof. To obtain the representation from the perspective of node j as a trading hub, we use the same transformation as in Section 2; e.g., $p_j = p$ and $p_\ell = p + \mu A_\ell^j$, where $A_{i\ell}^j = A_{i\ell} - A_{ij}$ for $i > 0$. The integrand of J 's objective function is therefore (omitting arguments of functions)

$$G_J = \sum_I \int_{M(I,p)} \sum_h [p_h s_h - C_h] f_I^j [D_I^j + \delta_I^j s'_j] d\mu_I.$$

Because Lemma 2.1 remains valid, we omit the terms that, as in the proof of Theorem 2.2, cancel out in the following Euler condition for optimality of s_j . For prices at which j 's capacity constraint is not binding (i.e., $s_j(p) < K_j$) and all supply functions are differentiable, the Euler condition is

$$\begin{aligned} 0 &= \left(\frac{\partial}{\partial s_j} - \frac{d}{dp} \frac{\partial}{\partial s'_j} \right) G_J \\ &= [p - c_j] \sum_I \int_{M(I,p)} f_I^j [D_I^j + \delta_I^j s'_j] d\mu_I \\ &\quad - \left(\sum_I \int_{M(I,p)} [\sum_h s_h + \sum_h [p_h - c_h] s'_h] f_I^j \delta_I^j d\mu_I \right). \end{aligned}$$

Therefore,

$$s_j \sum_I \int_{M(I,p)} f_I^j \delta_I^j d\mu_I + \sum_I \int_{M(I,p)} \sum_{h \neq j} [s_h + (p_h - c_h) s'_h] f_I^j \delta_I^j d\mu_I = [p - c_j] \sum_I \int_{M(I,p)} f_I^j D_I^j d\mu_I.$$

This can be written as the conditional expectation stated in the theorem, as previously in the proof of Theorem 2.2. \square

Note that the effect of multiple ownership is summarized by the additional term

$$\sum_I \int_{M(I,p_j)} \sum_{h \neq j} [s_h + (p_h - c_h) s'_h] f_I^j \delta_I^j d\mu_I$$

in the Euler condition for supply function s_j at its nodal price p_j . Its effect is to reduce J 's total supply at each price, reflecting its greater market power when it controls supplies at multiple nodes.

If there are no transmission constraints then Theorem 3.1 specializes to the Euler condition

$$\sum_h s_h(p) = [p - c_j(s_j(p))] \sum_{\ell \neq j} s'_\ell(p) - \sum_{h \neq j} [p - c_h(s_h(p))] s'_h(p),$$

or if all the marginal costs are the same, say $c_j = c_h = c_J$, then firm J cares only about the aggregate supply that it offers at each price, so

$$\sum_{h \in J} s_h(p) = [p - c_J] \sum_{k \notin J} s'_k(p),$$

as obtained in previous studies of supply function equilibria.

4. SUPPLY FUNCTION EQUILIBRIUM

Theorems 2.2 and 3.1 establish that for each supplier an optimal supply function must satisfy a first-order differential equation that depends on all the other supply functions, including those of demanders. This differential equation holds on each domain of differentiability of the supply functions. Presuming continuity of supply functions at the boundaries of these domains, altogether these domains are linked together to yield a complete system of equations (for his formulation, Holmberg (2005a) proves continuity at such boundaries and provides an explicit example).

Under fairly general conditions, this collection of \bar{n} differential equations for the $\bar{n} < n$ suppliers that bid strategically has a solution for each specification of \bar{n} constants of integration. These constants of integration are provided by the suppliers' transversality conditions at the minimum prices at which their offered supplies are positive. Thus, given the initial condition that $p_{j*} = c_j(q_{j*})$ and $s_j(p_{j*}) = q_{j*}$ for each supplier j , the differential equations specified in Theorem 2.2 characterize an equilibrium collection of supply functions, and analogously in the case of Theorem 3.1. Due to the nonlinearities in the differential equations, however, there is no assurance that the equilibrium is unique.

When there are no transmission constraints the computation of an equilibrium is straightforward since the suppliers' output trajectories evolve together as the energy price increases. Computation of the equilibrium is considerably more complicated when there are transmission constraints. This is evident in Theorem 2.2 since firm j 's optimal supply $s_j(p_j)$ at its nodal price p_j depends on the probability distribution of the slopes $(s'_k(p_k))_{k \neq j}$ of other firms' supply functions at their nodal prices, which can differ over a wide range depending on which transmission constraints are binding. This feature implies that techniques more sophisticated than ordinary numerical integration are required. For example, if the computation is done by discretizing the differential equations then one obtains a set of simultaneous nonlinear equations whose solution approximates an equilibrium.

Some special cases are more amenable to solution; e.g., if the transmission system is "radial" then the market can be divided into zones, each with a single zonal energy price that is the nodal price for every firm located within the zone; that is, the zonal price fully summarizes all the effects of binding constraints on transmission into and out of the zone.

5. PAY-AS-BID SETTLEMENTS

In a few markets suppliers are paid their actual bids rather than market-clearing prices. In this section we adapt the previous analysis to characterize equilibrium when settlements are pay-as-bid, but for simplicity we address only the case that each supplier is located at a single node. Since settlements of this kind are used mainly when demand is perfectly inelastic, we assume this feature here by supposing that the effects of demand are included in the vector b .

In this case the operator minimizes its total cost $\sum_j P_j(p_j, s_j(p_j))$ of energy procurements subject to the feasibility constraint that $\sum_j A_j s_j(p_j) \leq b$, with equality required for the energy constraint $i = 0$. Using pay-as-bid settlements, the payment to firm j is

$$\begin{aligned} P_j(p_j, s_j(p_j)) &= p_* q_* + \int_{q_*}^{s_j(p_j)} s_j^{-1}(q) dq \\ &= p_j s_j(p_j) - \int_0^{p_j} s_j(\pi) d\pi, \end{aligned}$$

where p_* is the price at which the firm offers its minimum supply q_* . This payment differs from the settlement $p_j s_j(p_j)$ at market-clearing prices by the “rebate” $\int_0^{p_j} s_j(\pi) d\pi$. Assuming that each $s'_j > 0$, the implications of the operator’s optimality condition are essentially the same as before; namely, for each supplier j its nodal price is $p_j = \lambda A_j$, where λ is the vector of Lagrange multipliers for the constraints, but now this nodal price is paid only for the firm’s marginal unit of supply. Let F_j be the marginal probability distribution of j ’s nodal price p_j given the supply functions; i.e., $F_j(p) \equiv \text{Prob}\{\lambda A_j \leq p\}$, or using the transformation that makes j ’s node a trading hub, $F_j(p) \equiv \text{Prob}\{\lambda_o \leq p\}$ since $\lambda A_j^i = \lambda_o$ after the transformation. We use below the properties that

$$\begin{aligned} f_j(p) &\equiv \frac{d}{dp} F_j(p) = \sum_I \int_{M(I,p)} f_I^j(\sum_\ell A_{I\ell}^j s_\ell(p + \mu_I A_{I\ell}^j)) [D_I^j + \delta_I^j s'_j(p)] d\mu_I \\ \frac{\partial}{\partial s_j(p)} F_j(p) &= \sum_I \int_{M(I,p)} f_I^j(\sum_\ell A_{I\ell}^j s_\ell(p + \mu_I A_{I\ell}^j)) \delta_I^j d\mu_I \\ \frac{\partial}{\partial s'_j(p)} F_j(p) &= 0. \end{aligned}$$

Firm j ’s realized profit contribution is

$$\Pi_j(p_j, q_j) = P_j(p_j, q_j) - C_j(q_j) \quad \text{where} \quad q_j = s_j(p_j) \quad \text{and} \quad p_j = \lambda A_j.$$

Assume that $P_j(p_*, s_j(p_*)) = C_j(q_*)$ at j ’s minimum price p_* and minimum supply $q_* = s_j(p_*)$. Let p^* be the price at which j exhausts its capacity K_j . Then the expectation of its profit contribution is

$$\begin{aligned} E\{\Pi_j\} &= \int_{p_*}^{p^*} [P_j(p, s_j(p)) - C_j(s_j(p))] dF_j(p) \\ &= \int_{p_*}^{p^*} [ps_j(p) - \int_0^p s_j(\pi) d\pi - C_j(s_j(p))] dF_j(p) \\ &= \int_{p_*}^{p^*} \{[ps_j(p) - C_j(s_j(p))]f_j(p) - s_j(p)[1 - F_j(p)]\} dp \\ &= \int_{p_*}^{p^*} \{G_j(p, s_j(p), s'_j(p)) - s_j(p)[1 - F_j(p)]\} dp, \end{aligned}$$

where the third equality uses integration by parts, and G_j is the integrand of j ’s expected profit using market-clearing settlements as in Section 2.

Theorem 5.1. *With pay-as-bid settlements the Euler condition for optimality of j ’s supply function is*

$$1 - F_j(p) = [p - c_j(s_j(p))] \sum_I \int_{M(I,p)} f_I^j(\sum_\ell A_{I\ell}^j s_\ell(p_j + \mu_I A_{I\ell}^j)) D_I^j d\mu_I.$$

Proof. The Euler condition is the same as the one obtained in the proof of Theorem 2.2 except for the effect of the rebate term $s_j(p)[1 - F_j(p)]$. Specifically,

$$\begin{aligned}
0 &= \left(\frac{\partial}{\partial s_j} - \frac{d}{dp} \frac{\partial}{\partial s'_j} \right) \left(G_j(p, s_j(p), s'_j(p)) - s_j(p)[1 - F_j(p)] \right) \\
&= -s_j \sum_I \int_{M(I,p)} f_I^j \delta_I^j d\mu_I + [p - c_j] \sum_I \int_{M(I,p)} f_I^j D_I^j d\mu_I - \left(\frac{\partial}{\partial s_j} - \frac{d}{dp} \frac{\partial}{\partial s'_j} \right) s_j [1 - F_j] \\
&= -s_j \sum_I \int_{M(I,p)} f_I^j \delta_I^j d\mu_I + [p - c_j] \sum_I \int_{M(I,p)} f_I^j D_I^j d\mu_I - [1 - F_j] + s_j \frac{\partial}{\partial s_j} F_j \\
&= -[1 - F_j] + [p - c_j] \sum_I \int_{M(I,p)} f_I^j D_I^j d\mu_I,
\end{aligned}$$

as stated in the theorem. \square

The formula for j 's expected profit expresses the fact that it obtains the profit margin $p - c_j(q)$ on its q -th unit of output whenever its nodal price exceeds the price p at which it offered that unit. Correspondingly, the Euler condition merely says that for the q -th unit, where $q = s_j(p)$, a slightly higher price $p + dp$ obtains the gain on the left side, represented by the probability that j 's nodal price exceeds $p + dp$, balanced on the right side by the prospect that j loses the profit margin $p - c_j(q)$ on the sale of the q -th unit times the probability that j 's nodal price is between p and $p + dp$.

If there are no binding transmission constraints then the Euler condition specializes to the condition

$$1 - F_j(p) = [p - c_j(s_j(p))] f_j(p) \sum_{\ell \neq j} s'_\ell(p)$$

obtained by Holmberg (2005c).

The analysis of the transversality conditions at the firm's lower and upper limits p_* and p^* of its offered prices is essentially same as in subsection 2.1.

For comparisons between pay-as-bid and market-clearing settlements, it is sometimes useful to say that a pay-as-bid supply function \hat{s} is "revenue equivalent" to a market-clearing supply function s if, for each output quantity q ,

$$s^{-1}(q)q = \hat{s}^{-1}(q_*)q_* + \int_{q_*}^q \hat{s}^{-1}(x) dx,$$

or equivalently

$$s(p) + [p - \hat{s}^{-1}(s(p))]s'(p) = 0$$

assuming they obtain the same revenue at q_* . It might be thought that revenue equivalence maps an optimal supply function under one settlement rule into an optimal supply function under the other settlement rule. This conjecture is reinforced by the essential equivalence of the operator's optimization under the two settlement rules, and the similar roles of nodal prices, although interpreted as marginal prices when settlements are pay-as-bid but as average prices when settlements use market-clearing prices. But this conjecture is generally false; cf. Hästö and Holmberg (2005) and Holmberg (2005c). The explanation is evident by observing that the games are not strategically equivalent because a bidder's financial incentives differ in the two cases. In particular, a supplier's gain from raising the marginal price at its node is less than

the gain from raising the average price, even after adjusting the supply function to the altered settlement rule.

The debate between proponents of settlements based on market-clearing prices and pay-as-bid schemes has a long history, most prominently in the context of auctions of Treasury securities. In recent years the U.S. Treasury and several other central banks converted to settlements based on market-clearing prices. Amid the 2000-2001 crisis in California's wholesale electricity market, a panel convened to study the matter opted to continue relying on market-clearing prices; cf. Kahn, Cramton, Porter, and Tabors (2001). On the other hand, in 2001 the U.K. adopted pay-as-bid settlements. Holmberg (2005c) resolves this long-standing debate via an explicit model, albeit for a limited class of probability distributions.

6. CONCLUSION

At least since the seminal study by Green and Newbery (1992), it has been recognized that supply function equilibrium is the appropriate model for firms' bidding strategies in wholesale spot markets for electricity. Some countries encounter transmission congestion rarely, but in the U.S. most regional systems are tightly constrained by limits on transmission capacity. Moreover, most of these systems now use nodal pricing (implementations of nodal pricing in California and Texas are presently incomplete) and it is endorsed by the Federal Energy Regulatory Commission. A chief impediment to studies of supply function equilibrium has been the absence of a mathematical characterization of the necessary conditions for an equilibrium when transmission constraints might be binding, especially when nodal prices are used for settlements. This paper tries to fill that gap. On a technical note, a methodological contribution of this paper is to indicate the usefulness of techniques from the calculus of variations to characterize supply function equilibria.

Nevertheless, the results presented here are not especially encouraging. Unlike the equilibrium conditions when there is no congestion, the conditions in the general case depend on the probability distribution of random shocks to demand and transmission capacity, and the equations to be solved are highly nonlinear. This presents a challenging computational problem, but it also raises a conceptual problem. If the conditions for an equilibrium are so complicated as to impede academic and policy studies, then perhaps it is implausible to suppose that firms' bidding strategies approximate an equilibrium. But there is an alternative viewpoint. This paper takes the joint probability distribution of shocks to energy demand and transmission capacity as the primitive. From a firm's viewpoint, however, for its own optimization it suffices to use the joint probability distribution of marginal values (λ) or nodal prices ($p = \lambda A$), which it can estimate directly from market data.

To see this, observe that if $F_o^j(p)$ is the marginal distribution function of firm j 's nodal price p then its expected profit contribution is

$$E\{\Pi_j\} = \int_{p_*}^{\infty} [ps_j(p) - C_j(s_j(p))] dF_o^j(p) = \int_{p_*}^{\infty} [s_j(p) + [p - c_j(s_j(p))]s'_j(p)][1 - F_o^j(p)] dp,$$

where (as in Section 5) the second equality is obtained via integration by parts, assuming $p_*s_j(p_*) = C_j(s_j(p_*))$. Using this formulation, the Euler condition is (with some abuse of notation)

$$s_j(p) \frac{\partial F_o^j(p)}{\partial s_j(p)} = [p - c_j(s_j(p))] E \left\{ \sum_{k \neq j} \frac{\partial F^j(p, \mu)}{\partial s_k(p + \mu A_k^j)} s'_k(p + \mu A_k^j) \right\},$$

where the expectation is taken over the vector μ of Lagrange multipliers affecting other firms' nodal prices due to transmission congestion. (This is just another way of writing the condition in Theorem 2.2.) Thus, for firm j it suffices to estimate the marginal effect of incremental supply on the marginal probability distribution of its nodal price, and to observe the average effect of the slopes of other firms' supply functions. Since wholesale spot markets for electricity are repeated continually, some experimentation can complement observed market data to provide the requisite estimates.

Therefore, the seeming complexity of the equilibrium conditions in Sections 2-5 should be interpreted as a consequence of deriving the distribution of nodal prices from more primitive assumptions, whereas firms care only about the end result of this derivation, which can be estimated directly from experience.⁴

⁴For a survey of efficient computational methods for estimating the covariance matrix of the time series of nodal prices, see Vandenberghe and Boyd (1996).

References

- Ross Baldick and William Hogan, "Capacity Constrained Supply Function Equilibrium Models of Electricity Markets: Stability, Nondecreasing Constraints, and Function Space Iterations," Paper PWP-089, University of California Energy Institute, Berkeley, 2002.
- Peter Hästö and Pär Holmberg, "Some Inequalities Related to the Analysis of Electricity Auctions," Department of Mathematical Sciences of the Norwegian University of Science and Technology (Trondheim, Norway) and Department of Economics of Uppsala University (Uppsala, Sweden), January 2005.
- Hung-po Chao and Stephen Peck, "A Market Mechanism for Electric Power Transmission," *Journal of Regulatory Economics*, 10: 25-60, 1996.
- Hung-po Chao, Stephen Peck, Shmuel Oren, and Robert Wilson, "Flow-Based Transmission Rights and Congestion Management," *Electricity Journal*, October 2000, pp. 38-58.
- L. Elsgolc, *Calculus of Variations*. London: Pergamon Press Ltd., 1962.
- Richard Green and David Newbery, "Competition in the British Electricity Spot Market," *Journal of Political Economy*, 100: 929-953, 1992.
- Pär Holmberg, "Unique Supply Function Equilibrium with Capacity Constraints," (Working Paper 2004:20, November 2004); "Asymmetric Supply Function Equilibrium with Constant Marginal Costs," (March 2005 preliminary version); and "Numerical Calculation of an Asymmetric Supply Function Equilibrium with Capacity Constraints," (Working Paper 2005:12); "Comparing Supply Function Equilibria of Pay-as-Bid and Uniform-Price Auctions (April 2005 preliminary version). Department of Economics, University of Uppsala, Sweden.
- Alfred Kahn, Peter Cramton, Robert Porter, and Richard Tabors, "Pricing in the California Power Exchange Electricity Market: Should California Switch from Uniform Pricing to Pay-as-Bid Pricing?," Blue Ribbon Panel Report, California Power Exchange, January 2001. Available at <http://www.cramton.umd.edu/papers2000-2004/kahn-cramton-porter-tabors-blue-ribbon-panel-report-to-calpx.pdf>
- Paul Klemperer and Margaret Meyer, "Supply Function Equilibrium in Oligopoly under Uncertainty," *Econometrica*, 57: 1243-1277, 1989.
- L. Vandenberghe and Stephen Boyd, "Semidefinite Programming," *SIAM Review*, 38: 49-95, 1996.
- Robert Wilson, *Nonlinear Pricing*. London: Oxford University Press, 1993.

STANFORD BUSINESS SCHOOL, STANFORD, CA 94305-5015, USA.

E-mail address: `rwilson@stanford.edu`