OPERATIONS RESEARCH
Vol. 56, No. 2, March–April 2008, pp. 369–382
ISSN 0030-364X | EISSN 1526-5463 | 08 | 5602 | 0369

DOI 10.1287/opre.1070.0421




# Supply Function Equilibrium in a Constrained Transmission System

**Robert Wilson**
Stanford Business School, Stanford, California 94305, rwilson@stanford.edu

This paper characterizes a supply function equilibrium in an auction market constrained by limited capacities of links in a transportation network and limited input/output capacities of participants. The formulation is adapted to a wholesale spot market for electricity managed by the operator of the transmission system. The results are derived using the calculus of variations to obtain the Euler conditions and the transversality conditions that characterize a Nash equilibrium in an auction in which bids are as supply functions, and quantities and payments are based either on nodal prices or pay-as-bid.

*Subject classifications*: industries: electric; games/group decisions: bidding/auctions; networks/graphs: applications.
*Area of review*: Environment, Energy, and Natural Resources.
*History*: Received March 2006; revisions received October 2006, January 2007; accepted January 2007. Published online in *Articles in Advance* November 9, 2007.

## 1. Introduction

This paper derives conditions that characterize equilibria in an auction market with multilateral trading 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, other load-serving entities, and large consumers). 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]

Green and Newbery (1992) first proposed that an appropriate model for studying wholesale markets for electricity is a formulation as a game in which firms' strategies are their choices of supply functions, and the predicted outcome is a Nash equilibrium of this game, i.e., a supply function equilibrium as originally defined by Klemperer and Meyer (1989). One interpretation of the game among suppliers and the system operator is that it is a multileader Stackelberg game in which suppliers are the leaders, albeit in competition with each other, and the follower is the system operator, acting as the agent of demanders. For general discussions of the role of strategic behavior in electricity networks, see Anderson (2004), Berry et al. (1999), Bolle (2001), Ferrero et al. (1997), Hogan (1997), Rudkevich (2003), Stoft (1999), and Younes and Ilic (1997).

In many electricity markets, financial settlements rely on locational marginal pricing, also called nodal pricing. In each contingency the system operator chooses the allocation to maximize the apparent gain from trade subject to the feasibility constraints imposed by limited transmission capacities. (The apparent gain is the gain from trade "as bid" because 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 balance ($p$) and transmission capacity constraints ($\mu$). The settlement rules in systems that use locational marginal pricing prescribe that 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 firm $j$ 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 $p_j s_j(p_j)$.

### 1.1. Prior Studies of Equilibrium in Power Markets

This paper differs from prior work mainly in that transmission constraints are included. We use the calculus of variations to characterize optimal supply functions, as suggested by the formulation in Smeers and Wei (1997). The characterizations obtained are essentially the same as those obtained by Holmberg (2005) using more general methods for models without transmission constraints. Our formulation differs from Day et al. (2002), Anderson and Philpott (2002a, b), and Anderson and Xu (2002, 2005), who allow each participant to conjecture others' supply functions, and Hobbs and Rijkers (2004) and Hobbs et al. (2004), who allow conjectures about effects on transmission prices without requiring these conjectures to form an equilibrium. Most previous studies of electricity markets assume competition in the style of either Bertrand or Cournot; e.g., Cunningham et al. (2002), Oren (1997), Wei and Smeers

(1999), and Willems (2002). These formulations amount to specifying a supply function whose price elasticity is either infinite or zero. Those studies that allow less extreme versions still restrict the allowed form of the supply functions, and use mathematical programming methods to solve for equilibria. The usual restriction is that a supply function must be linear with only one strategic variable, either the slope or the intercept; e.g., Baldick et al. (2004) and Niu et al. (2005). Representative of early work are Ferrero et al. (1997) and Berry et al. (1999), who use a grid search of the strategy space to search for an equilibrium in a model with two suppliers. Hobbs et al. (2000) formulate a mathematical program subject to equilibrium constraints (MPEC) for each generator, and iterate among the generators to search for an equilibrium. Xu et al. (2004) derive the first-order conditions for each generator's (nonconvex) MPEC, and then collect them into a single equilibrium problem with equilibrium constraints (EPEC). MPEC and EPEC formulations have become the standard formulations for applied work that requires computational optimization of detailed models of generation and transmission.

### 1.2. Outline

The formulation is established in §2. To introduce the main concepts and methods in a simple context, §3 sketches a few preliminary properties of a "radial" transmission network. A radial network is one with no loops, i.e., there is a unique path in the network between every two nodes. In this case, the effect of congestion on a transmission link is simply to separate the system into "zones" at either end of the link, with each zone having its own price for energy. Radial networks are thus trivial for the purpose of this paper, but (as suggested by the referees and editors) some elementary results are presented to fix ideas.

The remainder of this paper considers the general case in which the topology of the network is arbitrary. After a preliminary characterization of the distribution of nodal prices induced by the distribution of shocks in §4, in §§5 and 6 we characterize a firm's optimal bidding strategy. In §5, this is done for a firm located at a single node in the network, and then §6 addresses the case that a firm controls resources located at multiple nodes in the network and submits multiple supply functions. These results are applied in §7 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. Because a few systems settle transactions using a pay-as-bid rule, §8 characterizes equilibrium bidding strategies when this settlement rule is used. Section 9 concludes by sketching an alternative formulation that suggests that from market data, a supplier can obtain estimates sufficient to calculate its optimal supply function.

## 2. Formulation

There are several firms (either generators/suppliers or utilities/demanders) indexed by $j = 1, \ldots, n$. For notational simplicity, we treat demand as negative supply; thus, if firm $j$ is a demander with valuation $V_j$ and demand function $D_j$, then its supply function is $s_j = -D_j$ and its cost is $C_j = -V_j$. In some contexts, we assume (realistically) that not all demanders act strategically. Except in §6, each firm is assumed to inject (or extract) power at a single node in the network.

There are also several transmission capacity constraints indexed by $i = 1, \ldots, m$. 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 due to electrical resistance, so each $u_{\circ j} = -1$; 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 $A = (a_{ij})$. Because 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 $p = \lambda_\circ$ and the marginal value of enlarging the capacity of link $i$ is $\lambda_i$.

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. (In §5.2, we address the constraint that the supply function must be monotonic.) We assume that each firm's supply function takes account of its own local generation capacity constraints; e.g., if supplier $j$ can supply at most $K_j$, then necessarily $s_j(p_j) \leqslant K_j$. We assume further that each supplier must offer its entire generation 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 $\lambda = (\lambda_i)_{i=\circ,1,\ldots,m}$ 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 $\lambda$ of marginal prices must satisfy the feasibility constraints

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

and the complementarity conditions

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

We assume throughout that the number of firms (and their own generation capacities) is sufficiently large so 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, $\sum_j s_j(p_j) = b_\circ$.

The intended interpretation of the shocks to net energy demand and the capacities of transmission links included in the vector $b$ is that they are the aggregate of terms added to the firms' supply and demand functions and to the system transmission constraints as predicted initially. They are insensitive (inelastic) to prices because they occur after the submission of supply functions, and they occur on a timeframe too short to allow adjustments to the firms' submissions. For example, they could be due to real-time deviations from predicted energy demands, and to real-time deviations from predicted transmission capacities due perhaps to loop flows from outside the system, variations in voltage and reactive energy, or tripped lines.

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 (2005) has shown in a more restrictive formulation, including constant marginal costs, that continuity and piecewise-differentiability are implied by fundamental considerations. Anderson and Xu (2006) obtain similar results (with a notable exception) without assuming constant marginal costs. For this paper, we rely on the presumption that their analyses can be extended to the more general formulation used here, and therefore we use the techniques of the calculus of variations. Elsgolc (1962) includes all the techniques used here.

## 3. Supply Function Equilibrium in a Radial Network

To introduce some of the main ideas, we first study a radial transmission network, i.e., one with no loops. In a radial system, each binding transmission constraint separates the spot market into separate zonal markets on either side of the constrained link. Typically, all those suppliers $j$ located on one side of the binding constraint $i > 0$ have positive PTDFs ($a_{ij} > 0$) and those on the other side have negative PTDFs. The separation of markets implies that each zone has an induced net demand for energy. That is, for a fixed realization of the shock $b$, if a zone is constrained by a fixed set of exports and/or imports obtained over congested transmission links, then its net demand (or shock to net demand) for energy is its original demand plus prescribed exports and minus prescribed imports. These exports and imports are determined solely from the fact that $I$ is the set of binding constraints when the shock is $b$. The following simple example illustrates market separation.

**Example of Market Separation in a Radial Network.** Consider an example in which a single transmission line $i = 1$ connects two zones $j = 1, 2$ in which the aggregate supply and demand functions are $S_1, D_1$ in zone 1 and $S_2, D_2$ in zone 2 for injections (by suppliers) and withdrawals (by demanders) at their nodes. Let $L$ be the capacity of this link in the direction from zone 1 to zone 2, which is the only source of congestion. The energy balancing constraint and the transmission capacity constraint require that the excess of supply in zone 1 is sufficient to cover the deficiency of supply in zone 2, and that the link's capacity is exhausted; hence

$$S_1(p_1) = D_1(p_1) + L,$$
$$S_2(p_2) = D_2(p_2) - L.$$

Thus, the price in each zone is determined to balance supply and demand in that zone net of exports ($L$ from zone 1) or imports ($L$ into zone 2). One obtains this same result from the specification of the coefficients $a_{\circ 1} = a_{\circ 2} = 1$, $b_\circ = 0$ for energy, and $a_{11} = 1$, $a_{12} = 0$, $b_1 = L$ for transmission. The latter supposes PTDFs are measured by assuming that injections flow to a demand hub in zone 2, whereas if the hub is in zone 1, then the coefficients are $a_{11} = 0$, $a_{12} = -1$ and the result is the same. The difference between the two zonal prices for energy is the shadow price on the transmission constraint.

In general, on the domain $B(I)$ of shocks $b$ for which the set of binding constraints is $I$, in each induced zonal market $i$ there is an induced shock $b_\circ^i$ to that zone's demand for energy. Most important for the following is that this derived zonal energy constraint is the only binding constraint within the zone.

For a shock $b \in B(I)$ for which the set of binding constraints is $I$, let $i$ index the induced set of zonal markets (or the entire market if there is no transmission congestion), and let $J(i)$ be the subset of firms whose nodes are in zone $i$. Within zone $i$, the only binding constraint is the induced net energy constraint, $\sum_{j\in J(i)} s_j(p_i) = b_\circ^i$, where $p_i$ is the zonal price of energy; that is, $p_j = p_i$ for all $j \in J(i)$. Therefore, the induced probability density of the zonal price is

$$f_i(p_i) = f_\circ^i\left(\sum_{j\in J(i)} s_j(p_i)\right)\left|\sum_{j\in J(i)} s_j'(p_i)\right|,$$

where $f_\circ^i$ is the induced density of $b_\circ^i$, and $\sum_{j\in J(i)} s_j'(p_i)$ is a scalar that we can assume to be positive without significant loss of generality.

The partition of the space of shocks into the sets $B(I)$ corresponding to different sets $I$ of binding constraints induces in some cases a partition of the space of nodal prices. For example, it might be that the firms are located at the two ends of a single transmission line. In such a situation, it is conceivable that a firm $j$ knows that its nodal price $p_j$ lies (a) in the range \$30–\$50 when there is no congestion, (b) below \$30 when exports are limited by transmission congestion, and (c) above \$50 when imports are limited by congestion. This case is obviously unrealistic in practice, but its analysis is simple and it provides building blocks for the general analysis that begins in §5.

For a given set $I$ of binding constraints, consider the bidding problem of a supplier $j \in J(i)$, where $i$ is the zone induced by $I$ that includes $j$'s injection node. Over the domain $B(I)$ of shocks, its expected profit from sales in zone $i$ is

$$\Pi_j(I) = \int_{P_i(I)} [p_i s_j(p_i) - C_j(s_j(p_i))] \cdot f_\circ^i\left(\sum_{l\in J(i)} s_l(p_i)\right)\left[\sum_{l\in J(i)} s_l'(p_i)\right] dp_i,$$

times the probability of $B(I)$, where $P_i(I)$ is the domain of prices in zone $i$ induced by shocks in $B(I)$. From this expected profit, one derives the Euler condition for an optimal supply function:[3]

$$s_j(p_i) = [p_i - c_j(s_j(p_i))] \sum_{l\in J(i),\, l\neq j} s_l'(p_i).$$

This necessary condition has the familiar interpretation that firm $j$ chooses $(p_i, s_j(p_i))$ along the residual demand curve induced by others' supply functions so as to maximize its expected profit. Note that in this simple case, the Euler condition does not depend on the probability distribution of shocks.

**Example of an Optimal Supply Function.** Suppose that there are $n > 1$ suppliers in zone $i$ and all suppliers have the same constant marginal cost $c$ up to the same capacity $K$, and demand is nonstrategic and inelastic, i.e., $s_j' = 0$ if $j$ is a demander. Then, the Euler condition implies that every supplier's optimal supply function has the form

$$s_j(p_i) = C[p_i - c]^\alpha,$$

where $\alpha = 1/[n-1]$ and $C$ is a constant of integration. Holmberg (2005) shows that if there is a price cap, then $C$ is determined uniquely by the transversality condition at the price cap if firms' capacities are not exhausted. See §5.1 for the transversality conditions where a firm's capacity is exhausted.

The Euler condition must be satisfied for each zonal price $p_i$ having a positive induced density. The index $l$ in the formulas above might include some firms (e.g., demanders) who are price takers (i.e., do not act strategically) and therefore their supply functions are fixed and not subject to a Euler condition; therefore, let $\hat{J}(i)$ be the set of suppliers in zone $i$ who act strategically in choosing their supply functions. Taken together for all firms $j \in \hat{J}(i)$, these Euler conditions define a set of $|\hat{J}(i)|$ ordinary differential equations for an equal number of supply functions whose solution is essentially unique for each specification of $|\hat{J}(i)|$ constants of integration. In §5.1, we show that these constants of integration are determined by the associated transversality conditions when the firms have limited capacities and there is no price cap.

### 3.1. An Example with Two Zones

In this subsection, we formulate a simple example in which two zones $i = 1, 2$ are connected by a single transmission line with limited capacity, and demand in each zone is affected by an additive shock. The capacity of the transmission line between the two zones is $L$ in both directions. To avoid trivial cases, assume that in each zone the aggregate generation capacity exceeds the transmission capacity plus demand at the monopoly price. For simplicity, assume further that demanders are nonstrategic and each supplier $j$ has constant marginal cost $c_j \geqslant 0$ up to its generation capacity $K_j$.

Demand in zone $i$ is $D_i(p_i) + \varepsilon_i$ when the price in zone $i$ is $p_i$. The two shocks $(\varepsilon_1, \varepsilon_2)$ are jointly distributed with a differentiable and positive density on $\mathbb{R}^2$. Define $\varepsilon = \varepsilon_2 + \varepsilon_1$ and $\delta = \varepsilon_2 - \varepsilon_1$. Let $s_j$ be the supply function submitted by supplier $j$, and let $S_i$ be the aggregate supply in zone $i$. Assume that each demand function $D_i$ is smooth, nonincreasing, and nonnegative, and each supply function $s_j$ is continuous, piecewise-smooth, nondecreasing, and nonnegative.

If there is no transmission congestion, then prices are the same in both zones, say $p_i = p$, and the energy balance constraint requires $\sum_i D_i(p) + \varepsilon = \sum_i S_i(p)$. Because the market clearing price $p$ in this case is less than $r$ if and only if there is excess supply at price $r$, i.e., $\sum_i D_i(r) + \varepsilon < \sum_i S_i(r)$, the cumulative distribution function of the clearing price is

$$H(r) \equiv \text{Prob}\{p \leqslant r\} = \text{Prob}\left\{\varepsilon \leqslant \sum_i [S_i(r) - D_i(r)]\right\} = F\left(\sum_i [S_i(r) - D_i(r)]\right),$$

where the induced marginal density and distribution functions of $\varepsilon$ are $f$ and $F$. Therefore, the induced density of $H$ at the price $p$ is

$$h(p) = f\left(\sum_i [S_i(p) - D_i(p)]\right)\left|\sum_i [S_i'(p) - D_i'(p)]\right|.$$

To enable mathematical symmetry in the subsequent derivations, assume that the PTDFs are +1 and −1. The amount of transmission, say from zone 1 to zone 2, is

$$T(p,\delta)=\Delta D(p)-\Delta S(p)+\delta,$$

where $p=(p_1,p_2)$ and $\Delta D(p)=D_2(p_2)-D_1(p_1)$ and $\Delta S(p)=S_2(p_2)-S_1(p_1)$. If there is congested transmission from zone 1 to zone 2, then $T(p,\delta)=L$, and the two zonal prices satisfy

$$D_1(p_1)+L/2+\varepsilon_1=S_1(p_1)\quad\text{and}$$
$$D_2(p_2)-L/2+\varepsilon_2=S_2(p_2),$$

where typically $p_1<p_2$.

### 3.2. Perfectly Correlated Shocks

To illustrate some aspects of the methods invoked later, we begin by sketching the case in which $\varepsilon$ and $\delta$ are perfectly correlated. Without loss of generality, assume that $\varepsilon_1=0$. In this case, the ranges of shocks $\varepsilon\equiv\varepsilon_2$ for which there is or is not congestion can be translated into corresponding intervals of prices. It suffices to consider only situations in which there might be congestion from zone 1 to zone 2 because the analysis of congestion in the other direction is analogous. We use the notational convention that $p_1=p$ and $p_2=p+q$, where $q\geqslant 0$ is the "congestion charge." Let $p^*$ be the price in zone 2 below which there is no congestion from zone 1 to zone 2.

Over any interval of zone 2 prices below $p^*$ an equilibrium of the firms' supply functions requires that each satisfies the Euler condition for an optimal reply to the aggregate of the others' supply functions. For a firm $j$, the relevant term of its expected profit is

$$\Pi_j=\int_{p_*}^{p^*}[p-c_j]s_j(p)\,dF\left(\sum_i[S_i(p)-D_i(p)]\right).$$

As above, the Euler condition for an optimal reply is

$$s_j(p)=[p-c_j]\left[\sum_{l\neq j}s_l'(p)-\sum_i D_i'(p)\right].$$

The solution of these differential equations depends on constants of integration, which we derive from each firm's transversality condition at the price $p^*$ that connects the two regimes with and without congestion of transmission from zone 1 to zone 2. For a firm $j$ in zone 2, the relevant terms of its expected profit from prices above and below $p^*$ can be written as

$$\Pi_j=\int_0^{p^*}[p-c_j]s_j(p)\,dF\left(\sum_i[S_i(p)-D_i(p)]\right)$$
$$+\int_{p^*}^{P_j}[p-c_j]\hat{s}_j(p)\,dF(\hat{S}_2(p)-D_2(p)-L/2),$$

where $P_j>p^*$ is the price at which firm $j$ exhausts its capacity (i.e., $\hat{s}_j(P_j)=K_j$), and $s_j$ and $\hat{s}_j$ are firm $j$'s supply functions over the intervals of prices below and above $p^*$. Invoking Holmberg's (2005) and Anderson and Xu's (2006) general argument that a firm's optimal supply function must be continuous, each firm $j$'s transversality condition at $p^*$ is subject to the constraint that $s_j(p^*)=\hat{s}_j(p^*)$, and collectively for all firms

$$S_2(p^*)=D_2(p^*)+\varepsilon^*-L/2=\hat{S}_2(p^*)$$

for the shock $\varepsilon^*$ that induces the price $p^*$.

To simplify here, suppose that firm $j$ is the only supplier in zone 2, i.e., $s_j\equiv S_2$. Then, the transversality condition is[4]

$$s_j'(p^*)=\hat{s}_j'(p^*).$$

This is a classical "smooth pasting" condition, as in Dixit and Pindyck (1994).[5] It requires that firm $j$'s optimal supply function transitions smoothly at $p^*$ from its supply function $s_j$ when there is no congestion to its supply function $\hat{s}_j$ when there is congestion into zone 2. Thus, from firm $j$'s optimal supply function $\hat{s}_j$ when transmission into zone 2 is fully congested, one computes its derivative $\hat{s}_j'(p^*)$ at $p^*$ and uses that to derive one of the constants of integration that determines $s_j$ at prices below $p^*$. The analysis of the transversality condition for a firm in zone 1, at the price $p_*$ at which there is congestion from zone 2 to zone 1, is analogous. Again, the result is a smooth pasting condition that requires that its supply function is differentiable at $p_*$. This construction determines the constants of integration for solution of the firms' Euler conditions in the range of prices between $p_*$ and $p^*$, where they compete in the absence of congestion. In sum, the chief implication of this simple case is that there must be a smooth transition from the regime of system-wide competition to the regime of within-zone competition.

### 3.3. Imperfectly Correlated Shocks

In a more general case, $\varepsilon$ and $\delta$ are imperfectly correlated. In this case, transmission can be congested because $\delta$ is large, regardless of whether $\varepsilon$ is large or small. In choosing its supply function, therefore, a firm must account for the fact that any price might arise with or without congestion of the transmission line. Here we illustrate the general method by sketching the effect of possible congestion from zone 1 to zone 2, ignoring for simplicity the prospect of congestion in the opposite direction. To simplify further, we assume that there is a single supplier $j=1$ in zone 1 and a single supplier $j=2$ in zone 2.

Define

$$F(\bar{\varepsilon},\bar{\delta})\equiv\text{Prob}\{\varepsilon\leqslant\bar{\varepsilon}\ \&\ \delta\leqslant\bar{\delta}\}\quad\text{and}$$
$$G(\bar{\varepsilon}_1,\bar{\delta})\equiv\text{Prob}\{\varepsilon_1\leqslant\bar{\varepsilon}_1\ \&\ \delta\geqslant\bar{\delta}\}.$$

Assume that the distribution function of the price $p$ received by firm 1 can be written as

$$H(r)\equiv\text{Prob}\{p\leqslant r\}$$
$$=F(\Sigma(r),\Delta(r))+G(S_1(r)-D_1(r),\Delta(r)),$$

where

$$\Sigma(r) \equiv \sum_i [S_i(r) - D_i(r)] \quad \text{and}$$

$$\Delta(r) = \Delta S(r,r) - \Delta D(r,r) + L.$$

In this formulation, the firm term $F(\cdot)$ is the probability (induced by the firms' supply functions) that there is no congestion because $\delta$ is sufficiently small and the system-wide price $p$ is less than $r$ because $\varepsilon$ is sufficiently small. The second term $G(\cdot)$ is the probability that there is congestion because $\delta$ is sufficiently large and the price in zone 1 is less than $r$ because $\varepsilon_1$ is sufficiently small. The probability density at the price $p$ is therefore the total derivative of $H$ at $p$.

Firm 1's expected profit (ignoring congestion in the opposite direction) is

$$\Pi_1 = \int_{c_1}^{P_1} [p - c_1] S_1(p)\, dH(p),$$

where $P_1$ is the price at which firm 1 exhausts its capacity $K_1$. Firm 1's optimal reply to firm 2's supply function requires that its supply function satisfies the Euler condition, and the transversality condition at $P_1$. The similar conditions for firm 2 then yield a pair of differential equations. The two transversality conditions at the prices where they exhaust their capacities then determine a particular solution of these differential equations.

Denote the partial derivatives of $F$, $G$ with respect to their first and second arguments by $F_1$, $F_2$, $G_1$, $G_2$, and let $E = F_1 - F_2 + G_1 - G_2$. Then, the Euler condition for firm 1 is $S_1(p) = [p - c_1] R_1(p)$, where (omitting arguments)

$$R_1 = \{F_1 \cdot [S_2' - D_1' - D_2'] - F_2 \cdot [-S_2' - D_1' + D_2'] + G_1 \cdot [-D_1'] - G_2 \cdot [-S_2' - D_1' + D_2']\}/E.$$

The Euler condition for firm 2 is analogous.

To determine the two constants of integration, suppose that, say, it is firm 1 that exhausts its capacity $K_1$ at a higher price than firm 2 does (e.g., firm 1 owns the peaking units); that is, $P_1 > P_2$. Then, for prices in the interval $P_2 < p \leqslant P_1$, it knows that it is a monopolist. Hence, firm 1 uses monopoly pricing in this interval, obtained using the special case of its Euler condition in which the formula for $R_1$ has $S_2' = 0$ and its transversality condition at $P_1$ is simply the condition that $(P_1, K_1)$ are its monopoly price and quantity (taking account of the shocks $(\varepsilon, \delta)$ affecting demands). Given this determination of $S_1$ in the interval $P_2 < p \leqslant P_1$, firm 1's transversality condition at $P_2$ is a smooth pasting condition requiring that its supply function transits smoothly from prices below $P_2$ (where it competes with firm 2) to those above. Firm 1's smooth pasting condition therefore determines firm 2's optimal supply at $P_2$ because $(P_2, K_2)$ must be its optimal price and quantity in reply to its (expected) residual demand given firm 1's supply function $S_1$ (and continuity of its derivative $S_1'$ below and above $P_2$). Thus, these two transversality conditions provide the requisite two conditions that determine a particular solution of the firms' pair of Euler conditions. A similar construction applies when it is firm 2 that exhausts its capacity at a higher price $P_2 > P_1$.

The foregoing illustrates Holmberg's (2005) general proposition (albeit, assuming constant marginal costs, which Anderson and Xu 2006 show is immaterial to his argument) that the firms' optimal supply functions are continuous, and that transversality conditions imply continuous derivatives across the boundaries of the domains $P_i(I)$ of prices associated with different sets $I$ of binding transmission constraints; furthermore, that the transversality conditions at the firms' capacity constraints ultimately determine exactly the firms' optimal supply functions in a supply function equilibrium.

As mentioned, the paper by Holmberg (2005) studies the solutions of these systems of differential equations, assuming constant marginal costs, using more general and powerful methods than the calculus of variations used here; and these characterizations are extended by Anderson and Xu (2006) to variable marginal costs. An important implication of Holmberg's results is that in a system without transmission constraints, the firms' capacity constraints (or if demand is completely inelastic, a price cap) determines a unique equilibrium. For the more general models studied in subsequent sections, it seems likely that Holmberg's methods might again establish that a unique equilibrium is determined from transversality conditions at the firms' capacity constraints, or at a price cap if demand is completely inelastic.

## 4. The Induced Distribution of Nodal Prices

For the remainder of this paper, we consider a transmission system with an arbitrary topology, which might be non-radial and hence includes loops. Henceforth, we use the general formulation in §2.

To analyze a firm's bidding problem, we first derive the probability distribution of the vector $\lambda$ of shadow prices 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.[6] Further, we assume that $F$ has a density function $f$ that is differentiable with support that is a convex full-dimensional subset of $\mathbb{R}^{m+1}$. 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 the 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 subvector 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 subvector. 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 $\lambda_I$ with respect to the components of $b_I$ via implicit differentiation to obtain the relation

$$\left[\sum_j s_j'(\lambda_I A_{Ij}) A_{Ij} \cdot A_{Ij}^T\right] \cdot \lambda_I'(b_I) = \mathrm{Id}_I,$$

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, $\mathrm{Id}_I$ 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 $\lambda_I$,

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

or $g_I \equiv 0$ if there is zero probability that $I$ is the set of binding constraints. The last term in this formula is the determinant $|B_I|$ of the Jacobian 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 $|B_I| = (4s_1's_2' + (s_1' + s_2')s_3')/9$. The general fact that this determinant is multilinear can be proved using the transformation introduced in §5.

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 $\lambda_I$. 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.

## 5. 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 §9. 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

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

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 $p_j = \lambda_\circ + \sum_{i>0} \lambda_i a_{ij}$ and $k$'s nodal price is $p_k = \lambda_\circ + \sum_{i>0} \lambda_i a_{ik}$, it follows that $p_k = p_j + \sum_{i>0} \lambda_i [a_{ik} - a_{ij}]$. Similarly, the operator's feasibility constraint $\sum_l A_l s_l \leqslant b$ 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 $\sum_l A_l^j s_l = b^j$, where $a_{\circ j}^j = 1$ and $a_{ij}^j = 0$ for each $i > 0$, and for each other firm $a_{\circ k}^j = 1$ and $a_{ik}^j = a_{ik} - a_{ij}$ for each $i > 0$; also, $b_\circ^j = b_\circ$ and $b_i^j = b_i - a_{ij} b_\circ$. This transformation has no effect on the determinant $|B_I|$ in the previous formula for $g_I(\lambda_I)$, but it alters the interpretation and value of $\lambda_I$. In particular, $p_j = \lambda_\circ$ and each $p_k = p_j + \sum_{i>0} \lambda_i a_{ik}^j$. 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 $\mu_I$ 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 5.1.

$$\left(\frac{\partial}{\partial s_j} - \frac{d}{dp}\frac{\partial}{\partial s_j'}\right) \sum_I \int_{M(I,p)} f_I^j\left(\sum_l A_{Il}^j s_l(\lambda_I A_{Il}^j)\right) \cdot \left|\sum_l s_l'(\lambda_I A_{Il}^j) A_{Il}^j \cdot A_{Il}^{jT}\right| d\mu_I = 0$$

*for any supply functions of the other bidders.*

Proof. The Euler condition for maximizing

$$\int_P \left(\sum_I \int_{M(I,p)} f_I^j\left(\sum_l A_{Il}^j s_l(\lambda_I A_{Il}^j)\right) \cdot \left|\sum_l s_l'(\lambda_I A_{Il}^j) A_{Il}^j \cdot A_{Il}^{jT}\right| d\mu_I\right) dp$$

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. □

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

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

Because $a^j_{0j} = 1$ and $a^j_{ij} = 0$, in this formula each determinant can be expanded into determinants of cofactors, and hence can be written as

$$|B^j_I| \equiv \left|\sum_l s'_l(\lambda_I A^j_{Il}) A^j_{Il} \cdot A^{jT}_{Il}\right| = D^j_I + \delta^j_I s'_j(p).$$

Specifically,

$$D^j_I = \left|\sum_k s'_k(\lambda_I A^j_{Ik}) A^j_{Ik} \cdot A^{jT}_{Ik}\right|,$$

and $\delta^j_I$ is the cofactor of the element $\sum_l s'_j$ in position $(0, 0)$ of $B^j_I$; viz.,

$$\delta^j_I = \left|\sum_k s'_k(\lambda_I A^j_{Ik}) A^j_{I^\circ k} \cdot A^{jT}_{I^\circ k}\right|,$$

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

$$I = \{0\}, \quad D^j_I = \sum_k s'_k(p), \quad \text{and} \quad \delta^j_I \equiv 1.$$

Theorem 5.2. *The Euler condition for optimality of firm j's supply function is*

$$s_j(p) = [p - c_j(s_j(p))] E\left\{\sum_I D^j_I \,\middle|\, p \ \text{ and } \sum_l s_l(p + \mu A^j_l) = b_\circ\right\}.$$

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

$$G_j(p, s_j(p), s'_j(p)) = [ps_j(p) - C_j(s_j(p))] \sum_I \int_{M(I,p)} f^j_I \left(\sum_l A^j_{Il} s_l(\lambda_I A^j_{Il})\right) \cdot \left|\sum_l s'_l(\lambda_I A^j_{Il}) A^j_{Il} \cdot A^{jT}_{Il}\right| d\mu_I.$$

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)

$$\begin{aligned}
0 &= \left(\frac{\partial}{\partial s_j} - \frac{d}{dp}\frac{\partial}{\partial s'_j}\right) G_j(p, s_j(p), s'_j(p)) \\
&= [p - c_j] \sum_I \int_{M(I,p)} f^j_I [D^j_I + \delta^j_I s'_j]\, d\mu_I \\
&\quad + [ps_j + C_j] \sum_I \int_{M(I,p)} f^{j\prime}_\circ [D^j_I + \delta^j_I s'_j]\, d\mu_I \\
&\quad - \Bigg(s_j \sum_I \int_{M(I,p)} f^j_I \delta^j_I\, d\mu_I + [p - c_j] s'_j \sum_I \int_{M(I,p)} f^j_I \delta^j_I\, d\mu_I \\
&\qquad + [ps_j - C_j] \frac{d}{dp} \int_{M(I,p)} \sum_I f_I \delta^j_I\, d\mu_I\Bigg) \\
&= -s_j \sum_I \int_{M(I,p)} f^j_I \delta^j_I\, d\mu_I + [p - c_j] \sum_I \int_{M(I,p)} f^j_I D^j_I\, d\mu_I \\
&\quad + [ps_j - C_j]\left(\frac{\partial}{\partial s_j} - \frac{d}{dp}\frac{\partial}{\partial s'_j}\right) \sum_I \int_{M(I,p)} f^j_I [D^j_I + \delta^j_I s'_j]\, d\mu_I \\
&= -s_j \sum_I \int_{M(I,p)} f^j_I \delta^j_I\, d\mu_I + [p - c_j] \sum_I \int_{M(I,p)} f^j_I D^j_I\, d\mu_I,
\end{aligned}$$

where the last equality applies Lemma 5.1. Therefore,

$$s_j = [p - c_j] \frac{\sum_I \int_{M(I,p)} f^j_I D^j_I\, d\mu_I}{\sum_I \int_{M(I,p)} f^j_I \delta^j_I\, d\mu_I}.$$

Now for each set $I$ of binding constraints

$$\int_{M(I,p)} f^j_I \delta^j_I\, d\mu_I = \int_{M(I,p)} f^j_I(b_\circ, b_{I^\circ})\, db_{I^\circ},$$

where $b_\circ \equiv \sum_l s_l$; that is, $\sum_I \int_{M(I,p)} f^j_I \delta^j_I\, d\mu_I$ is the marginal probability of the net energy shock $b_\circ$ inferred from the requirement that it is met by the total supply $\sum_l s_l$. Hence, at the price $p$,

$$\frac{\sum_I \int_{M(I,p)} f^j_I D^j_I\, d\mu_I}{\sum_I \int_{M(I,p)} f^j_I \delta^j_I\, d\mu_I} = E\left\{\sum_I D^j_I \,\middle|\, p \text{ and } \sum_l s_l = b_\circ\right\}.$$

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 it is conditional on $j$'s nodal price $p$ and all the supply functions. This completes the proof. □

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

If there are no transmission constraints, then Theorem 5.2 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; see, e.g., Klemperer and Meyer (1989) and Holmberg (2005). In this special case, the Euler condition is independent of the probability distribution of shocks.

If one writes the optimality condition in the abbreviated form $s_j = [p - c_j] \bar{D}^j / \bar{\delta}^j$, and then rewrites it as

$$\frac{p - c_j}{p} = \frac{1}{\eta(p)}, \quad \text{where } \eta(p) = \frac{p}{s_j} \frac{\bar{D}^j}{\bar{\delta}^j},$$

then it states the usual condition that the firm's supply at each price $p$ is an optimal reply to the conditional expectation given $p$ of its demand net of other firms' supplies; viz., its percentage profit margin (the so-called Lerner index) is the reciprocal of the conditional expectation $\eta(p)$ of the elasticity of its residual demand at the price $p$.

### 5.1. 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

$$
\begin{aligned}
&E\{\Pi_j(p,s_j(p)\}\\
&\quad=\int_{p_*}^{p^*} G_j(p,s_j(p),s_j'(p))\,dp+\int_{p^*}^{\bar p}[pK_j-C_j(K_j)]\\
&\qquad\cdot\sum_I\int_{M(I,p)}\hat f_I^j\Big(\sum_l A_{Il}^j s_l(\lambda_I A_{Il}^j)\Big)\hat D_I\,d\mu_I\,dp.
\end{aligned}
$$

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 differ 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 $p^*$.[7] Then, the transversality condition at $p^*$ is

$$
\begin{aligned}
0&=\Big(G_j-s_j'\frac{\partial}{\partial s_j'}G_j\Big)-[p^*K_j-C_j(K_j)]\sum_I\int_{M(I,p^*)}\hat f_I^j\hat D_I\,d\mu_I\\
&=[p^*K_j-C_j(K_j)]\sum_I\int_{M(I,p^*)}[f_I^jD_I-\hat f_I^j\hat D_I]\,d\mu_I,
\end{aligned}
$$

provided $p^*<\bar p$. If $\hat f_I^j\hat D_I=f_I^jD_I$ for every $I$, then this condition is satisfied identically. However, Holmberg (2005) 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 (2005) 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

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

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

$$p_*=C_j(s_j(p_*))/s_j(p_*),$$

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 $s_j(p_*)=0$ at $p_*=c_j(0)$ suffices. 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 (and if there is a bid cap $\bar p$, then it requires that $p_*\leqslant\bar p$). Also, a generator typically has a minimum safe operating rate, say $q_*>0$, which requires that $s_j(p_*)\geqslant q_*$. 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.[8] Therefore, for the analysis in the next section, we assume that unrecoverable fixed costs are nil, and thus that $s_j(p_*)=q_*$ at $p_*=C_j(q_*)/q_*$.

### 5.2. Monotonicity Constraint

In practice, an operator's software typically is 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 $s_j'(p)\geqslant 0$ 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

$$
\begin{aligned}
0&=\Big(\frac{\partial}{\partial s_j}-\frac{d}{dp}\frac{\partial}{\partial s_j'}\Big)[G_j(p,s_j(p),s_j'(p))+\varphi_j(p)s_j'(p)]\\
&=\Big(\frac{\partial}{\partial s_j}-\frac{d}{dp}\frac{\partial}{\partial s_j'}\Big)G_j(p,s_j(p),s_j'(p))-\varphi_j'(p).
\end{aligned}
$$

Because $\varphi_j(p)=0$ 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 $[p_1^\circ,p_2^\circ]\supset(p_1,p_2)$ such that $s_j(p_1^\circ)=s_j(p_2^\circ)$, and over which

$$\int_{p_1^\circ}^{p_2^\circ}\Big(\frac{\partial}{\partial s_j}-\frac{d}{dp}\frac{\partial}{\partial s_j'}\Big)G_j(p,s_j(p),s_j'(p))\,dp=0,$$

and then the optimal constrained supply function is $s_j(p)=s_j(p_1^\circ)$ for every price $p\in[p_1^\circ,p_2^\circ]$. 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; see, e.g., Wilson (1993).

## 6. Multiple Supply Functions

In this section, we extend the characterization in §5 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

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

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 §5.1, 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 6.1. *The Euler condition for optimality of firm $J$'s supply function $s_j$ is*

$$s_j(p_j) = E\Bigg\{[p_j - c_j(s_j(p_j))]\sum_I D_I^j - \sum_{h\neq j}[s_h(p_h) + (p_h - c_h(s_h(p_h))s_h'(p_h))] \,\Bigg|\, p_j \text{ and } \sum_l s_l = b_\circ\Bigg\}.$$

Proof. To obtain the representation from the perspective of node $j$ as a trading hub, we use the same transformation as in §5; e.g., $p_j = p$ and $p_l = p + \mu A_l^j$, where $A_{il}^j = A_{il} - A_{ij}$ 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 5.1 remains valid, we omit the terms that, as in the proof of Theorem 5.2, 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

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

Therefore,

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

This can be written as the conditional expectation stated in the theorem, as previously in the proof of Theorem 5.2. □

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. In principle, therefore, from the firm's observed supply functions one can infer what they would have been without common ownership of its supply resources.

If there are no transmission constraints, then Theorem 6.1 specializes to the Euler condition

$$\sum_h s_h(p) = [p - c_j(s_j(p))]\sum_{l\neq j} s_l'(p) - \sum_{h\neq j}[p - c_h(s_h(p))] s_h'(p),$$

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.

## 7. Supply Function Equilibrium

Theorems 5.2 and 6.1 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 2005 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 5.2 characterize an equilibrium collection of supply functions, and analogously in the case of Theorem 6.1. Due to the nonlinearities in the differential equations, however, there is no assurance that there is a unique equilibrium.

When there are no transmission constraints, the computation of an equilibrium is relatively straightforward because 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 5.2 because 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 a solution. For example, 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.

## 8. 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. Because 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 $\sum_j A_j s_j(p_j) \leqslant b$, with equality required for the energy constraint $i = 0$. Using pay-as-bid settlements, the payment to firm $j$ is

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

where $p_*$ is the price at which the firm offers its minimum supply $q_* = s_j(p_*)$. 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 $\lambda$ 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., $F_j(p) \equiv \text{Prob}\{\lambda A_j \leqslant p\}$, or using the transformation that makes $j$'s node a trading hub, $F_j(p) \equiv \text{Prob}\{\lambda_\circ \leqslant p\}$ because $\lambda A_j^j = \lambda_\circ$ after the transformation. We use below the properties that

$$f_j(p) \equiv \frac{d}{dp}F_j(p) = \sum_I \int_{M(I,p)} f_I^j\Big(\sum_l A_{Il}^j s_l(p + \mu_I A_{Il}^j)\Big) \cdot [D_I^j + \delta_I^j s'_j(p)]\,d\mu_I,$$

$$\frac{\partial}{\partial s_j(p)}F_j(p) = \sum_I \int_{M(I,p)} f_I^j\Big(\sum_l A_{Il}^j s_l(p + \mu_I A_{Il}^j)\Big)\delta_I^j\,d\mu_I,$$

$$\frac{\partial}{\partial s'_j(p)}F_j(p) = 0.$$

Firm $j$'s realized profit contribution is

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

Let $p^*$ be the price at which $j$ exhausts its capacity $K_j$. Then, the expectation of its profit contribution is

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

where the third equality uses integration by parts, and the fourth uses the same integrand $G_j$ of $j$'s expected profit using market-clearing settlements as in §5.

Theorem 8.1. *With pay-as-bid settlements, the Euler condition for optimality of $j$'s supply function is*

$$1 - F_j(p) = [p - c_j(s_j(p))] \cdot \sum_I \int_{M(I,p)} f_I^j\Big(\sum_l A_{Il}^j s_l(p_j + \mu_I A_{Il}^j)\Big) D_I^j\,d\mu_I.$$

Proof. The Euler condition is the same as the one obtained in the proof of Theorem 5.2, except for the effect of the rebate term $s_j(p)[1 - F_j(p)]$. Specifically,

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

$$= -s_j \sum_I \int_{M(I,p)} f_I^j \delta_I^j \, d\mu_I + [p - c_j] \sum_I \int_{M(I,p)} f_I^j D_I^j \, d\mu_I$$
$$- [1 - F_j] + s_j \frac{\partial}{\partial s_j} F_j$$
$$= -[1 - F_j] + [p - c_j] \sum_I \int_{M(I,p)} f_I^j D_I^j \, d\mu_I,$$

as stated in the theorem. □

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

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

obtained by Holmberg (2005).

The derivation of the transversality conditions at the firm's lower and upper limits $p_*$ and $p^*$ of its offered prices parallels §2.1. The transversality condition for the optimal choice of its nodal price $p^*$ at which $j$ exhausts its capacity $K_j$ is

$$1 - F_j(p^*) = [p^* - C_j(K_j)/K_j] \cdot \sum_I \int_{M(I,p^*)} f_I^j \left( \sum_l A_{Il}^j s_l(p^* + \mu_I A_{Il}^j) \right) D_I^j \, d\mu_I.$$

Assuming $p_*$ must be nonnegative and $s_j(p_*) \geqslant q_* > 0$, the transversality conditions for the optimal choices of $p_*$ and $s_j(p_*)$ are

$$0 \geqslant [p_* s_j(p_*) - C_j(s_j(p_*))] \cdot \sum_I \int_{M(I,p_*)} f_I^j \left( \sum_l A_{Il}^j s_l(p_* + \mu_I A_{Il}^j) \right) D_I^j \, d\mu_I - s_j(p_*)[1 - F_j(p_*)],$$

$$0 \geqslant [p_* s_j(p_*) - C_j(s_j(p_*))] \cdot \sum_I \int_{M(I,p_*)} f_I^j \left( \sum_l A_{Il}^j s_l(p_* + \mu_I A_{Il}^j) \right) \delta_I^j \, d\mu_I,$$

with equality required if $p_* > 0$ and $s_j(p_*) > q_*$, respectively. Thus, $s_j(p_*) = q_*$ and $p_* = C_j(q_*)/q_*$.

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 (2005). 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 et al. (2001). On the other hand, in 2001 the United Kingdom adopted pay-as-bid settlements. Hästö and Holmberg (2005) resolve this long-standing debate via an explicit model that, for a limited class of probability distributions, implies that pay-as-bid settlements yield lower average prices.

## 9. 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 United States most regional systems are tightly constrained by limits on transmission capacity. Moreover, most of these systems now use nodal pricing, 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 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. However, 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 ($\lambda$) or nodal prices ($p = \lambda A$), which it can estimate directly from market data.

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

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

where, as in §8, 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)

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

where the expectation is taken over the vector $\mu$ of Lagrange multipliers affecting other firms' nodal prices due to transmission congestion. (This is just another way of writing the condition in Theorem 5.2.) 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. Because 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 §§5–8 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.[9]

## Endnotes

1. Similar 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.
2. The PTDFs are derived from the linear approximation of Kirchhoff's laws obtained by assuming that the difference between phase angles at any two buses connected by transmission links is small or zero; cf. Chao and Peck (1996) and Chao et al. (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, resistance 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.
3. If the integrand is $I(p, S, S')$, then the Euler condition is $I_S = d[I_{S'}]/dp$. When $I = [pS - C(S)]f(S+X)[S'+X']$, the Euler condition reduces to $S = [p-c]X'$, where $c = C'$, after canceling the factor $f'(S+X)$ provided it is nonzero.
4. Elsgolc (1962, p. 69). For an objective of the form $\int_a^b I(p, S, S')\, dp$ and constraint $S(b) = \phi(b)$, the variation $db$ produces the variation $(I + [\phi' - S']I_{S'})db$ at $p = b$. The effect of variation of $a$ is similar. Here we apply these variations to the upper limit $p^*$ of the first integral and the lower limit of the second integral in the formula above for firm $j$'s expected profit $\Pi_j$.
5. The "subtlety" of the smooth pasting condition explains why, when there is congestion, attempts to compute supply equilibria by iterative procedures often fail to converge.
6. Anderson and Philpott (2003) develop methods for estimating "market distribution functions" that might be adapted to estimate $F$ from market data. Another method is considered briefly in §9.
7. As emphasized by Holmberg (2005), 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.
8. Some operators restrict this compensation to generators instructed to start to provide reserves or voltage support.
9. For a survey of efficient computational methods for estimating the covariance matrix of the time series of nodal prices, see Vandenberghe and Boyd (1996).

## Acknowledgments

The author is indebted to Pär Holmberg, the associate editor, and three referees for comments on previous drafts, and to the Electric Power Research Institute for financial assistance.

## References

Anderson, E., A. Philpott. 2002a. Optimal offer construction in electricity markets. *Math. Oper. Res.* **27** 82–100.

Anderson, E., A. Philpott. 2002b. Using supply functions for offering generation into an electricity market. *Oper. Res.* **50** 477–489.

Anderson, E., A. Philpott. 2003. Estimation of electricity market distribution functions. *Ann. Oper. Res.* **121** 21–32.

Anderson, E., H. Xu. 2002. Necessary and sufficient conditions for optimal offers in electricity markets. *SIAM J. Control Optim.* **41** 1212–1228.

Anderson, E., H. Xu. 2005. Supply function equilibrium in electricity spot markets with contracts and price caps. *J. Optim. Theory Appl.* **124** 257–283.

Anderson, E., H. Xu. 2006. Finding supply function equilibria with asymmetric firms. Working paper, Australian Graduate School of Management, Sydney, Australia.

Anderson, S. 2004. Analyzing strategic interaction in multi-settlement electricity markets: A closed-loop supply function equilibrium model. Ph.D. thesis, Harvard University, Cambridge, MA.

Baldick, R., W. Hogan. 2002. 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, CA.

Baldick, R., R. Grant, E. Kahn. 2004. Theory and application of linear supply function equilibrium in electricity markets. *J. Regulatory Econom.* **25** 143–167.

Berry, C., B. Hobbs, W. Meroney, R. O'Neill, W. Stewart. 1999. Understanding how market power can arise in network competition: A game theoretic approach. *Utilities Policy* **8** 139–158.

Bolle, F. 2001. Competition with supply and demand functions. *Energy Econom.* **23** 253–277.

Chao, H., S. Peck. 1996. A market mechanism for electric power transmission. *J. Regulatory Econom.* **10** 25–60.

Chao, H., S. Peck, S. Oren, R. Wilson. 2000. Flow-based transmission rights and congestion management. *Electricity J.* **13**(October) 38–58.

Cunningham, L., R. Baldick, M. Baughman. 2002. An empirical study of applied game theory: Transmission constrained Cournot behavior. *IEEE Trans. Power Systems* **17** 166–172.

Day, C., B. Hobbs, J. Pang. 2002. Oligopolistic competition in power networks: A conjectured supply function approach. *IEEE Trans. Power Systems* **17** 597–607.

Dixit, A., R. Pindyck. 1994. *Investment under Uncertainty*. Princeton University Press, Princeton, NJ.

Elsgolc, L. 1962. *Calculus of Variations*. Pergamon Press, London, UK.

Ferrero, R., S. Shahidehpour, V. Ramesh. 1997. Transaction analysis in deregulated power systems using game theory. *IEEE Trans. Power Systems* **12** 1340–1347.

Green, R., D. Newbery. 1992. Competition in the British electricity spot market. *J. Political Econom.* **100** 929–953.

Hästö, P., P. Holmberg. 2005. Some inequalities related to the analysis of electricity auctions. Working paper, Department of Mathematical Sciences, Norwegian University of Science and Technology, Trondheim, Norway and Department of Economics, Uppsala University, Uppsala, Sweden.

Hobbs, B., F. Rijkers. 2004. Strategic generation with conjectured transmission price responses in a mixed transmission system I: Formulation. *IEEE Trans. Power Systems* **19** 707–717.

Hobbs, B., C. Metzler, J. Pang. 2000. Strategic gaming analysis for electric power networks: An MPEC approach. *IEEE Trans. Power Systems* **15** 38–45.

Hobbs, B., F. Rijkers, A. Wals. 2004. Strategic generation with conjectured transmission price responses in a mixed transmission pricing system II: Application. *IEEE Trans. Power Systems* **19** 872–879.

Hogan, W. 1997. A market power model with strategic interaction in electricity networks. *Energy J.* **18** 107–141.

Holmberg, P. 2005. Modelling bidding behavior in electricity auctions: Supply function equilibria with uncertain demand and capacity constraints. Economic Studies 87, University of Uppsala, Uppsala, Sweden.

Kahn, A., P. Cramton, R. Porter, R. Tabors. 2001. 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, Folsom, CA.

Klemperer, P., M. Meyer. 1989. Supply function equilibrium in oligopoly under uncertainty. *Econometrica* **57** 1243–1277.

Niu, H., R. Baldick, G. Zhu. 2005. Supply function equilibrium bidding strategies with fixed forward contracts. *IEEE Trans. Power Systems* **20** 1859–1867.

Oren, S. 1997. Economic inefficiency of passive transmission rights in congested electricity systems with competitive generation. *Energy J.* **18** 63–83.

Rudkevich, A. 2003. Supply function equilibrium: Theory and applications. *Proc. 36th Hawaii Internat. Conf. System Sci.*, IEEE, Parsippany, NJ.

Smeers, Y., J. Wei. 1997. Spatially oligopolistic model with opportunity cost pricing for transmission capacity reservations—A variational inequality approach. CORE D.P. 9717, Université Catholique de Louvain, Louvain, Belgium.

Stoft, S. 1999. Using game theory to study market power in simple networks. H. Singh, ed. *Game Theory Tutorial*. IEEE Power Engineering Society, Parsippany, NJ, 33–40.

Vandenberghe, L., S. Boyd. 1996. Semidefinite programming. *SIAM Rev.* **38** 49–95.

Wei, J., Y. Smeers. 1999. Spatial oligopolistic electricity markets with Cournot generators and regulated transmission. *Oper. Res.* **47** 102–112.

Willems, B. 2002. Modeling Cournot competition in an electricity market with transmission constraints. *Energy J.* **23** 95–125.

Wilson, R. 1993. *Nonlinear Pricing*. Oxford University Press, New York.

Xu, H., D. Ralph, E. Ralph, P. Bardsley, M. Ferrisk. 2004. Electricity generation with looped transmission networks: Bidding to an ISO. Working paper, Australian Graduate School of Management, Sydney, Australia.

Younes, Z., M. Ilic. 1997. Generation strategies for gaming transmission constraints: Will the deregulated electric power market be an oligopoly? *Proc. IEEE Hawaii Internat. Conf. System Sci.*, IEEE, Parsippany, NJ, 112–120.