Supply Function Equilibrium with Asymmetric Capacities and Constant Marginal Costs

Pär Holmberg*

This paper analytically derives a Supply Function Equilibrium (SFE) of a real-time electricity market with multiple firms and asymmetric production capacities. There is a unique SFE, which is piece-wise symmetric when firms have identical constant marginal costs. It is believed that some of the properties of the derived SFE are valid for real-time markets in general. Firms' capacity constraints bind at different prices (i). Still, firms with non-binding capacity constraints have smooth residual demand (ii). Approximating an asymmetric real-time market with a symmetric one, tends to overestimate mark-ups for small positive imbalances and underestimate mark-ups for large positive imbalances (iii).

1. INTRODUCTION

Electric energy is expensive to store compared to its production cost. As a result, stored energy is negligible in most power systems. Thus power consumption and production have to be roughly in balance at all times. Most of the electric power is sold long before delivery. However, because neither consumption nor production is fully predictable, adjustments need to be made in real-time to maintain balance. The real-time market, also called a balancing market, is an important component in this process. The market functions as an auction — often a uniform-price auction — in which power producers can offer additional power relative to their contracted position. Both increments, balancing up/up-regulation, and decrements, balancing down/down-regulation, can be offered. The paper focuses on balancing-up bids, but corresponding results can be readily derived for balancing-down bids. A bid consists of a non-decreasing supply function and is submitted before the start of the delivery period. The demand in the real-time market is given by the system imbalance, and it is not known when bids are submitted, as unexpected demand shocks as well as generator and transmission outages may occur during the delivery period. The period is typically an hour, as in California, Pennsylvania-New Jersey-Maryland (PJM), and the Nordic countries, or half an hour as in Britain.

The Supply Function Equilibrium (SFE) with uncertain demand was introduced by Klemperer and Meyer (1989). Later on Green and Newbery (1992) and Bolle (1992) observed that the set-up of the model is similar to the organization of most electricity markets, and SFE is now an established model of bidding behavior in electricity auctions. In the non-cooperative Nash equilibrium of the static game, each producer commits to the bid that maximizes his expected profit given the bids of competitors. Klemperer and Meyer (1989) show that all smooth supply function equilibria are characterized by a differential equation, which in this paper is called the KM first-order condition.

---

* Department of Economics, Uppsala University, P.O. Box 513, SE-751 20 Uppsala, Sweden, +46 18 471 76 35; email: par.holmberg@nek.uu.se. I would like to thank my supervisor Nils Gottfries and my co-supervisor Chuan-Zhong Li for valuable comments, discussions and guidance. Comments by Börje Johansson, David Newbery, Andreas Westermark, and anonymous referees, mail correspondence with Robert Wilson, and suggestions from seminar participants at Uppsala University in October 2004 are also much appreciated. I am grateful to Meredith Beechey for proof-reading the paper. The Norwegian Water Resources and Energy Directorate (NVE) is acknowledged for providing data for electric power producers in Norway. The work has been financially supported by the Swedish Energy Agency and the Ministry of Industry, Employment and Communication.

The assumption of symmetric producers is convenient as it allows straightforward calculation of SFE for general cost functions, as shown by Rudkevich et al. (1998), Anderson and Philpott (2002), and Holmberg (2004). However, firms in electric power markets are typically asymmetric. In order to assess efficient antitrust policy and merger control, models that can analyze asymmetric markets are important. To get analytic results, the models of asymmetric markets tend to be greatly simplified, but are anyway useful as they improve the qualitative understanding of bidding in electricity auctions.

The case of linear SFE for asymmetric firms with linear marginal costs was analyzed by Green (1996). Baldick et al. (2004) extended this concept to piece-wise linear SFE that can be used to analyze asymmetric firms with asymmetric intercepts. However, both linear and piece-wise linear SFE are problematic in the presence of capacity constraints, an important feature of electricity markets.

For asymmetric capacities, SFE have only been analytically derived for firms with identical constant marginal costs. For this case, Newbery (1991) and Genc and Reynolds (2004) derive piece-wise symmetric and symmetric SFE[1], respectively. This paper extends their asymmetric duopoly models to multiple asymmetric producers. For two firms the unique equilibrium is similar to the piece-wise symmetric equilibrium presented by Newbery, but as in the analysis by Genc and Reynolds (2004), Newbery's linear demand has been replaced by perfectly inelastic demand and a price cap. This paper makes contributions beyond the work by Newbery (1991) in that it is thoroughly proven that the outlined equilibrium exists and that there are no other possible equilibria. Uniqueness occurs when maximum demand is so high that the up-regulation capacities bind for all firms, except possibly the largest. Similarly, minimum demand should be so low that the down-regulation capacities bind for all firms, except possibly the largest. As argued by Holmberg (2004), these assumptions are in theory satisfied by all real-time markets. Still it is too early to rule out other "equilibria". In practice, the probability of getting power shortages during a delivery period is very small, of the order 10^{-4}-10^{-6}, and even lower during off-peak periods. This may result in a very long learning period before the market finds the unique SFE. In addition, the risk may be so small that it is not considered at all by the bidders, i.e. they are not perfectly rational. Further, a regulator's implicit threat of reregulation might restrain the bids more than the price cap. These are all empirical questions, which are not further addressed here.

The unique SFE has the following properties for balancing-up bids. The supply functions of any two producers are identical and satisfy the KM first-order condition until the capacity constraint of the smaller firm binds. To compensate the smaller firm's kink at this price, all firms with non-binding capacity constraints have kinks in their supply functions. This ensures that firms with non-binding capacity constraints face a smooth residual demand. The capacity constraint of the second largest firm binds when the price reaches the price cap. Thereafter, the largest firm offers its remaining up-regulation capacity at a price equal to the price cap.

The piece-wise symmetric nature of the equilibrium is a specific result of piece-wise symmetric costs, and this result is only applicable to the few electricity markets that are dominated by one production technology. With this exception, the derived properties of the equilibrium: that firm's capacities bind at different prices, that the remaining firms (with non-binding capacity

---

[1] If maximum demand is above twice the capacity of the smallest firm, Genc and Reynolds (2004) suggest an asymmetric smooth equilibrium, which is not piece-wise symmetric. This equilibrium might exist if minimum demand is above some positive level. Such an assumption is reasonable for forward and day-ahead markets, but not for balancing and real-time markets, where the imbalance is often close to zero.

constraints) become more elastic at these prices, and that one firm sells its remaining capacity as a monopolist, is believed to be a general result if maximum demand is sufficiently high and if minimum demand is sufficiently low, and if an equilibrium exists under those circumstances. Further, it is likely that one can generally rule out perfectly elastic supply segments (except at the price cap)[2]. Moreover, it is unlikely that there are SFE where firms with non-binding capacity constraints have perfectly inelastic supply segments also when marginal costs are non-constant. In Holmberg (2005) the conjectured properties are used to numerically calculate SFE for markets where firms have both asymmetric costs and asymmetric capacities. This is a task that has been problematic in the past, see e.g. Baldick and Hogan (2002). Lastly, it is believed that when comparing an asymmetric market with a symmetric market that has the same average competition, e.g. with the same Herfindahl-Hirschman index, the asymmetric market is more competitive for small demand outcomes — when most asymmetric firms have non-binding capacity constraints — and less competitive for large demand outcomes, when most asymmetric firms have binding capacity constraints.

The structure of the paper is as follows. Notation and assumptions are introduced in Section 2 and the unique SFE is derived in Section 3. In Section 4, the unique SFE is numerically illustrated for the case of three asymmetric producers. In Section 5, the unique SFE is calculated for 153 firms in the Norwegian real-time market and Section 6 concludes.

2. NOTATION AND ASSUMPTIONS

The analysis in this paper is similar to that in Holmberg (2004). However, symmetric producers with strictly convex cost functions are replaced by producers with identical constant marginal costs c and asymmetric capacities. In the derivation of the equilibrium, the analysis is confined to real-time and balancing markets with positive imbalances but corresponding results can be readily derived for negative imbalances as illustrated in Section 5. The contract/forward positions are assumed to be exogenously given.

There are $N \geq 2$ producers, all of whom have different production capacities. The bid of a producer is assumed to be valid for one delivery period only. In the original work by Klemperer and Meyer (1989), analysis was confined to twice continuously differentiable supply functions. Here, as in Holmberg (2004), the set of admissible bids is extended to allow for kinks as well as perfectly inelastic and perfectly elastic segments. Perfectly elastic segments imply that there are several possible supplies at some prices. Thus in the general case the bid is not a supply function, but a supply correspondence. However, the correspondence can still be represented by a piece-wise smooth supply function, which is assumed to be left continuous in this paper.[3] Let $S_i(p)$ be the supply function of an arbitrary producer i. It is then understood that this firm is prepared to sell any supply in the range $[S_i(p), S_i(p+)]$ at the price p. Aggregate supply of his competitors is denoted $S_{-i}(p)$ and total aggregate supply is denoted $S(p)$. All supply functions are assumed to be non-decreasing, as this is required by most electricity auctions.

Let $\bar{\varepsilon}_i$ be the up-regulation capacity of producer i. It is given by the firm's production capacity that can be regulated on short notice and that has not

[2] Unless there is one firm with a lower marginal cost than its competitors at zero supply, see Anderson and Hu (2006).
[3] The supply correspondence can also be represented by a right-continuous supply function as in Genc and Reynolds (2004). Whether the supply function is assumed to be left- or right continuous does not influence the SFE.

been contracted in advance. Without loss of generality, firms can be ordered according to their up-regulation capacity, i.e. $\bar{\varepsilon}_1 < \bar{\varepsilon}_2 < \ldots < \bar{\varepsilon}_N$. Total up-regulation capacity is designated by $\bar{\varepsilon}$, i.e. $\bar{\varepsilon} = \sum_{i=1}^{N} \bar{\varepsilon}_i$. Denote consumers additional demand (relative to their contracts) by ε and the additional demand's probability density function by $f(\varepsilon)$. The density function is continuously differentiable and has a convex support set that includes zero demand. If in addition, it is assumed that up-regulation capacities of all producers, but possibly the largest, bind with a positive probability, then one gets a unique equilibrium.

The price cap $\bar{p}$ is important in a model with perfectly inelastic demand, as the equilibrium price would be infinite without it. In practice, the ISO uses forced disconnection of consumers to ensure that demand is zero above the price cap. Thus the market price equals the price cap when $\varepsilon > \bar{\varepsilon}$. Demand rationing can increase social welfare as consumers who do not switch off their equipment when electricity prices become extremely high do not necessarily have a high marginal benefit of power. Instead, they may not have the ability to switch off their consumption in real-time or they do not face the real-time price, e.g. household consumers with fixed retail rates. Thus at some sufficiently high price, which is often called Value of Lost Load (VOLL), social welfare is maximized by rationing demand.

The market design is such that when total supply is partially perfectly inelastic and coincides with perfectly inelastic demand, then the best price for the ISO is chosen, i.e. the lowest price in case of up-regulation.

Let $q_i(p,\varepsilon)$ be the residual demand that firm i faces for $p < \bar{p}$. As long as its competitors' aggregate supply is not perfectly elastic at p, firm i's residual demand is given by

$$q_i(p,\varepsilon) = \varepsilon - S_{-i}(p) \quad \text{if } p < \bar{p}. \tag{1}$$

If more than one producer has a supply function with a perfectly elastic segment at some price p_0, supply rationing at this price is necessary for some demand outcomes. This possibility is addressed in Holmberg (2004).

3. THE UNIQUE ASYMMETRIC SFE

Klemperer and Meyer (1989) derive the KM first-order condition, which is a differential equation satisfied by all smooth, i.e. twice continuously differentiable, supply function equilibria. In the case when supply functions are piece-wise smooth, Holmberg (2004) uses optimal control theory to show that in equilibrium the KM condition must be satisfied by all price-responsive supply functions in any price interval in which all supply functions are smooth. In Section 3.1, the system of differential equations of an asymmetric market with identical and constant marginal costs is solved. The solution contains some undetermined integration constants but is nonetheless useful when SFE candidates with binding slope constraints are ruled out in Sections 3.2 through 3.4. These potential SFE do not necessarily fulfill the KM first-order condition; supply functions may have perfectly elastic segments even if the price cap does not bind or perfectly inelastic segments even if the capacity constraint does not bind. It takes a long proof, which involves the seven propositions illustrated in Figure 1, to rule out these irregularities.

Section 3.2, combines Proposition 1 and 5 in Holmberg (2004) to prove that no up-regulation power is offered below marginal cost or withheld from the auction (Proposition i). The latter means that all up-regulation capacity is offered at a price lower than or equal to the price cap. It is further shown that no firm will have a price-responsive supply in a price interval for which residual demand is perfectly inelastic (Proposition ii).

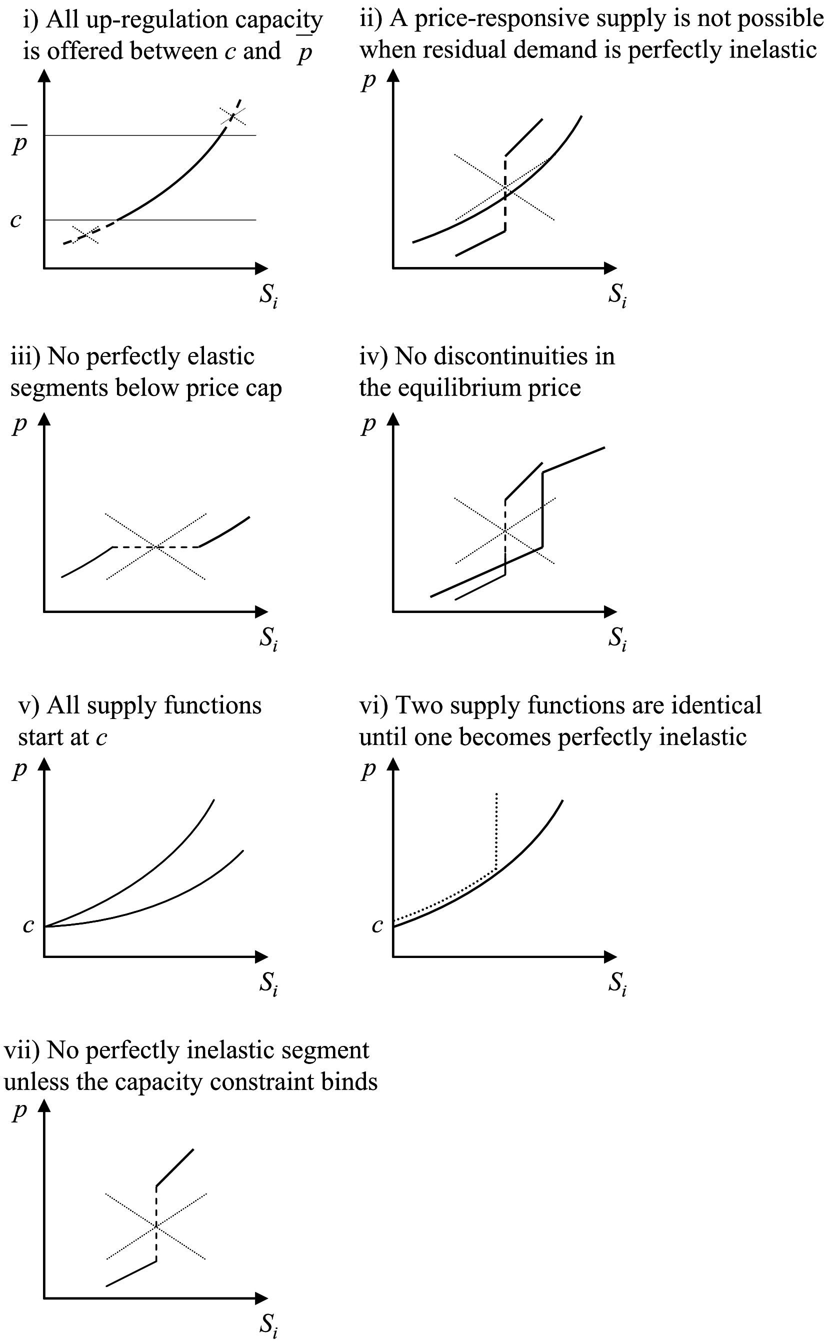

Figure 1. Graphical illustration of Propositions i-vii, the necessary properties of a supply function equilibrium with constant marginal costs when imbalances are positive.

Perfectly elastic segments below the price cap (Proposition iii) and discontinuities in the equilibrium price (Proposition iv) are ruled out for positive imbalances. These results follow from the proofs of Proposition 2 and 3 in Holmberg (2004), and Proposition 4 in Holmberg (2004), respectively. In addition, Proposition iii implies that at most one firm can have a perfectly elastic segment at the price cap.

Section 3.3 shows that all producers will offer their first unit of power at marginal cost (Proposition v), which provides an initial condition for the system of differential equations given by the KM first-order condition. Thus, in equilibrium, the supply function of each firm must fulfill the KM first-order condition from marginal cost up to the price at which either the up-regulation capacity or the price cap bind. Note that because of a singularity in the KM first-order condition when the price equals the marginal cost, the initial condition cannot be used to pin down the unique SFE.

Section 3.4 proves that any two supply functions of an equilibrium are identical up to the price at which one supply function becomes perfectly inelastic (Proposition vi). Using this result it is possible to show that firms with non-binding capacity constraints do not have perfectly inelastic segments in equilibrium (Proposition vii).

The integration constants in the solution of the first-order condition are determined in Section 3.5. Any two producers will have the same supply function until the capacity constraint of the smaller firm binds (see Proposition vi and Proposition vii). At least two firms must have non-binding capacity constraints up to the price cap (see Proposition ii and iv) and only one firm can have a perfectly elastic segment at the price cap (see Proposition iii). Thus the end-condition is that the up-regulation capacity of the second largest firm must start to bind at the price cap. The remaining capacity of the largest firm is sold with perfectly elastic supply at the price cap. Only one SFE candidate satisfies both the KM first-order condition and the end-condition.

Section 3.6 verifies that the unique candidate is indeed an equilibrium. If the competitors of firm *i* follow the SFE candidate, then the profit of firm *i* is globally maximized for every demand outcome if it also follows the SFE candidate. This is a sufficient condition for a SFE.

3.1. The first-order condition for smooth parts of the supply functions

Given the bids of competitors, each producer submits his best supply function out of the class of admissible supply functions. Now consider a price interval $[p_-, p_+]$, in which all supply functions $S_i(p)$ of an equilibrium are twice continuously differentiable, i.e. no firm has kinks or perfectly elastic supply in the interval. Denote equilibrium supply functions by $\widehat{S}_i(p)$. Assume that there are $M \geq 2$ firms with price-responsive supply in the whole interval. Let ***E*** be a set of all firms belonging to this group. The price interval is chosen such that firms not belonging to ***E*** have $\widehat{S}_j'(p) = 0 \quad \forall p \in [p_-, p_+]$. With the optimal control formulation used in Holmberg (2004), it can be shown that within the price interval, the supply function of each producer in ***E*** satisfy the KM first-order condition[4]

$$\widehat{S}_i(p) - \widehat{S}_{-i}'(p)(p - c) = 0. \tag{2}$$

Thus the supply functions of firms belonging to ***E*** are given by a system of differential equations. The system can be solved with a two-step approach. First, the total supply function of the *M* producers, $\widehat{S}_e(p)$, is calculated. Second, individual supply functions can be derived. We start with the first step. The

[4] The first-order condition derived by Klemperer and Meyer (1989) is more general, as it allows for general cost functions and elastic demand.

supply function of each firm in the set ***E*** follows a differential equation as in (2). Summing these equations yields

$$\sum_{i\in E}\hat{S}_i(p)-\sum_{i\in E}\hat{S}_{-i}'(p)(p-c)=\hat{S}_e(p)-(M-1)(p-c)\hat{S}_e'(p)=0.$$

Note that the assumption that firms' marginal costs are identical is crucial in this step, as the identity is only valid as long as *p-C'* is identical for all firms in ***E***. The differential equation is separable and has the following solution,

$$S_e=\beta(p-c)^{1/(M-1)}, \tag{3}$$

where $\beta > 0$ is an integration constant. Now, we can proceed with the second step. For any producer *i* belonging to the set ***E***, (2) can be written in the following form:

$$\hat{S}_i(p)-(p-c)\left(\hat{S}_e'(p)-\hat{S}_i'(p)\right)=0.$$

Rewriting it by means of (3) yields

$$\hat{S}_i(p)+\hat{S}_i'(p)(p-c)=\frac{\beta}{M-1}(p-c)^{1/(M-1)}.$$

Thus, it follows from the product rule of differentiation that

$$\frac{d}{dp}\left\{\hat{S}_i(p)(p-c)\right\}=\frac{\beta}{M-1}(p-c)^{1/(M-1)}.$$

Integrating both sides yields

$$\hat{S}_i(p)=\frac{\beta}{M}(p-c)^{1/(M-1)}+\frac{\gamma_i}{p-c}. \tag{4}$$

For M=2, this result becomes similar to results by Genc and Reynolds (2004). Note that all of the M firms with price-responsive supply have the same β in the interval $[p_-,p_+]$. On the other hand, they have individual specific constants γ_i. Nevertheless, the individual solutions in (4) must add up to the aggregate solution in (3). Thus

$$\sum_{i=1}^{M}\gamma_i=0. \tag{5}$$

It follows from (3) that the slope of the aggregate supply function in the interval $[p_-,p_+]$ is positive and given by

$$\hat{S}_e'(p)=\frac{\beta}{M-1}(p-c)^{(2-M)/(M-1)}.$$

Thus

$$p'(\varepsilon)=\frac{1}{\frac{d\varepsilon}{dp}}=\frac{1}{\hat{S}_e'(p)}=\frac{M-1}{\beta}(p-c)^{(M-2)/(M-1)}. \tag{6}$$

3.2. Basic properties of the supply function equilibrium

Proposition i: In equilibrium, no up-regulation capacity is offered below marginal cost or withheld. The latter means that all up-regulation capacity is offered at a price equal to or below the price cap.

Proof: See proof of Proposition 1 and 5 in Holmberg (2004).

It is always better to offer up-regulation capacity at the marginal cost than below the marginal cost and it is always better to offer up-regulation capacity at the price cap than to withhold power.

Proposition ii: In equilibrium, there is no price interval (p_-, p_+), in which a producer with price-responsive supply faces perfectly inelastic residual demand.

Proof: Assume that firm *i* has a price-responsive supply in (p_-, p_+), while its residual demand is perfectly inelastic in this interval. Without losing accepted supply, all units previously offered in the range (p_-, p_+) could be offered at a price arbitrarily close to, but still below, p_+. Thus firm *i* would deviate. □

Proposition iii: In equilibrium, no firm has a perfectly elastic segment below the price cap in its balancing-up bid. At most one firm can have a perfectly elastic segment at the price cap. Isolated points with perfectly elastic supply are only possible at marginal cost *c*.[5]

Proof: See Appendix.

The intuition of the proof is the same as for the Bertrand game in which producers undercut each others' perfectly elastic bids as long as price exceeds marginal cost. Thus a perfectly elastic segment in a balancing-up bid is only possible if it is the highest bid in the market, and if competitors capacity bind at this price. If the perfectly elastic segment is not at the price cap, the firm could increase the price without lost demand, as its residual demand is perfectly inelastic at this price level. The Bertrand equilibrium, and other equilibria with perfectly elastic segments at marginal cost, can be ruled out as competitors' up-regulation capacity is not enough to meet maximum demand.

It follows from Proposition iii that $S_i'(p) = \infty$ can be ruled out for every $p \in (c, \overline{p})$. This leads to the following corollary.

Corollary i: For every $p \in (c, \overline{p})$ one can, in equilibrium, find a sufficiently large p_- and a sufficiently low p_+ such that all supply functions $S_i(\cdot)$ are twice continuously differentiable in the intervals $[p_-, p]$ and $[p, p_+]$. Furthermore, all supply functions are continuous at *p*.

Proposition iv: In equilibrium $p'(\varepsilon)$ is finite for $\varepsilon \in (0, \overline{\varepsilon})$. From the lowest bid to the price cap, there must be at least two firms with price-responsive supply in each price interval.

Proof: See Appendix.

If there was a discontinuity in the equilibrium price, then a bid just below the discontinuity can be significantly increased without reducing the probability that the bid is accepted. It follows from Proposition ii that one firm will not have price-responsive supply if its residual demand is perfectly inelastic. Thus there must be at least two firms with price-responsive supply in each price interval.

---

[5] Mathematical clarification: Let the set $\mathbb{Q}$ consist of all points in the (p,S_i) space for which $S_i'(p) = \infty$. If there is any isolated point in $\mathbb{Q}$, it must be the point $(c,0)$.

3.3. Every firm offers its first unit of power at marginal cost

In Lemma i (below) it is shown that at least two producers will offer their first unit of power at $p=c$. This result is employed in Proposition v to prove that all producers must offer their first unit of power at $p=c$. This corresponds to the result for symmetric SFE, shown by Klemperer and Meyer (1989) and Holmberg (2004), where the lowest bid also equals the marginal cost. The intuition behind this result is that the bid of the first unit is only price-setting when demand is very small. As long as there are demand outcomes for which the first unit is not accepted, lowering the lowest bid increases the profit for these demand outcomes without hurting the profit for higher demand outcomes. Hence, the first unit is sold as if under Bertrand competition. If minimum demand is larger than zero, which is normally the case in day-ahead markets, the lowest bid is never price-setting and it may differ from marginal cost. As is shown by Genc and Reynolds (2004), this extra freedom introduces additional equilibria.

Lemma i: In equilibrium, the first unit of power is offered at c.

Proof: Denote the lowest offer in the total supply by p^*. Assume that $p^*>c$. According to Proposition iii, at most one producer has a supply function with a perfectly elastic segment at the price cap. Thus $c< p^* < \overline{p}$. There are three additional implications from Section 3.2. First, perfectly elastic segments at p^* can be excluded, according to Proposition iii. Second, Corollary i implies that a price p_+ can be found such that all producers have twice continuously differentiable supply functions in the interval $\left[p^*, p_+\right]$. Third, according to Proposition iv, at least two firms have price-responsive supply in the whole interval if p_+ is sufficiently small. The aggregate supply of the firms with price-responsive supply is given by (3). Thus $p^*>c$ can be excluded as it would imply $S_e\left(p^*\right) > 0$. It follows from Proposition i that $p^* \geq c$. Thus $p^*=c$. □

Proposition v: In equilibrium, there must be some $p_+>c$ such that all producers have price-responsive supply functions in the interval $\left[c, p_+\right]$.

Proof: Assume that there is a potential equilibrium in which producer i offers his first unit at the price $p^*>c$. Denote the aggregate supply and the equilibrium price of the potential equilibrium by $S^A(p)$ and $p^A(\varepsilon)$. It follows from Lemma i that $p^A(0)=c$. Now consider the following deviation. Producer i reduces the price for an infinitesimally small unit of power so that his first unit of power is offered at the price $p \in \left(c, p^*\right)$. For each unit of deviated power, the deviation leads to the following marginal change in expected profit:

$$\int_{S^A(p)}^{S^A\left(p^*\right)} \left[p^A(\varepsilon)-c\right] f(\varepsilon) d\varepsilon. \tag{7}$$

It follows from Proposition iv that $S^{A'}(p) > 0$ for $p \in \left[p, p^*\right]$. Thus $S^A\left(p^*\right) - S^A(p) > 0$ and the deviation is profitable. Accordingly, equilibria where producer i offers its first unit of power at a price $p^*>c$ can be eliminated. Supply functions with a perfectly elastic segment at c are excluded by Proposition iii. Thus the first unit of power of producer i must be offered with a price-responsive supply function in some interval $[c, p_+]$, and this is true for all producers. □

3.4. Each producer has a price-responsive supply, unless his capacity constraint binds

In this section, it is shown that a firm has perfectly inelastic segments in its supply function only when its capacity constraint is binding. A similar result is proven in Theorem 12 by Baldick and Hogan (2002) for the more general case with strictly elastic demand and general cost functions. On the other hand, Theorem 12 only considers the case when residual demand is smooth in the price range of the perfectly inelastic segment.

It is somewhat involved to intuitively explain why perfectly inelastic segments are only possible when the firm's capacity constraint is binding. Assume that no producer with a non-binding capacity constraint has perfectly inelastic supply below p_L. Assume further that supply of producer *i* and possibly some of the firm's competitors becomes perfectly inelastic just above $p_L < \overline{p}$. According to Corollary i and Proposition iv, there must be at least two producers with price-responsive supply functions that follow the KM first-order condition just above p_L. Denote this set of price-responsive supply functions by ***S***. It follows from the KM first-order condition in (2) that any supply function in ***S*** must face a continuous derivative of its residual demand at the price p_L. Thus to compensate for the switch to perfectly inelastic supply by producer *i,* the elasticity of supply functions in ***S*** must increase discontinuously at p_L. Likewise, this increases the elasticity of the residual demand of producer *i* discontinuously at p_L. As a result, the producer wants to sell more units just above p_L. Accordingly, he deviates unless his capacity constraint binds.

To accomplish the proof, it is first shown that any two firms have identical supply functions up to the price at which either has a perfectly inelastic segment.

Proposition vi: In equilibrium, any two producers have identical supply functions in the interval $[c, p]$, where $p < \overline{p}$, if neither of them have supply functions with perfectly inelastic segments in this interval.

Proof: Without loss of generality, denote the two producers by 1 and 2. According to Proposition iii, neither of the two producers has a supply function with perfectly elastic segments in the interval $[c, p]$. As it is also assumed that neither has perfectly inelastic segments in the interval, the supply functions of both producers follow the KM first-order condition in the whole interval. The number of competitors with price-responsive supply may change in the interval, but the supply functions of firm 1 and 2 continue to be piece-wise solutions of the type in (4) over the whole interval $[c, p]$. Proposition v implies that the two firms have the same initial condition $S_i(c) = 0$ and according to Corollary i both supply functions are continuous. Thus firm 1 and 2 will have identical γ_i over the whole interval $[c, p]$. □

Proposition vii: There is no equilibrium for which the supply function of a producer is perfectly inelastic in an interval $[p_L, p_U]$, where $c \le p_L < p_U \le \overline{p}$, unless his capacity constraint is binding.

Proof: Denote the potential equilibrium by the superscript *B*. Let ***I*** be a set with firms that have supply functions with perfectly inelastic segments before their capacity constraints bind. Define p_L by the following: no firm in ***I*** has a perfectly inelastic segment below p_L, but at least one firm in ***I*** starts being perfectly inelastic at p_L. Assume that producer *i* is one of these firms and that its supply is perfectly inelastic for $p \in [p_L, p_U]$. Proposition vi implies that all producers with a non-binding capacity constraint must have identical supply functions in the interval $[c, p_L]$.

According to Proposition iv, there must, for every price $p \in [p_L, p_U]$, be at least two producers — not necessarily the same over the whole interval — with price-responsive supply. Thus for a subinterval $[p_I, p_{II}] \subseteq [p_L, p_U]$, where a firm $j \neq i$ has a price-responsive supply function, its supply function must satisfy (2), i.e.

$$S_j^B(p) - S_{-j}^{B'}(p)(p-c) = 0 \;\; \forall\, p \in [p_I, p_{II}].$$

Because producer *i* has perfectly inelastic supply in the interval $[p_L, p_U]$, it follows that $S_{-i}^{B'}(p) > S_{-j}^{B'}(p)$ for $\forall\, p \in (p_I, p_{II})$. Furthermore, as $S_j^B(p_L) = S_i^B(p_L)$, $S_j^B(p) > S_i^B(p)$ for $\forall\, p \in (p_I, p_{II})$. Thus

$$S_i^B(p) - S_{-i}^{B'}(p)(p-c) < 0 \;\; \forall\, p \in (p_I, p_{II}). \tag{8}$$

Now consider the following deviation. Producer *i* decreases the price for an infinitesimally small unit of power previously offered at the price p_U and offers it at p_L instead. The supply function is unchanged above p_U and below p_L. Per marginal unit of deviated power, the deviation leads to the following marginal change in firm *i*'s expected profit:

$$\Delta E(\pi_i) = \int_{S^B(p_L)}^{S^B(p_U)} \left\{ \left[p^B(\varepsilon) - c\right] - S_i^B(p_L)\, p^{B'}(\varepsilon) \right\} f(\varepsilon)\, d\varepsilon. \tag{9}$$

The first term is due to increased sales and the second term due to the reduced price in the demand interval. As the supply of producer *i* is perfectly inelastic in the interval under consideration, (9) can be rewritten in the following way:

$$\Delta E(\pi_i) = \int_{S^B(p_L)}^{S^B(p_U)} \left\{ \left[p^B(\varepsilon) - c\right] - \frac{S_i^B(p_L)}{S_{-i}^{B'}(p_L)} \right\} f(\varepsilon)\, d\varepsilon.$$

There is a producer $j \neq i$ with a price-responsive supply at each $p \in [p_L, p_U]$. Hence, it follows from (8) that $\Delta E(\pi_i) > 0$, and a profitable deviation exists. In equilibrium, firm *i* cannot have a perfectly inelastic segment in the interval $[p_L, p_U]$ unless its capacity constraint binds. □

Proposition vii rules out perfectly inelastic segments in the supply of a firm unless its capacity constraint binds so the following corollary can be concluded by means of Propositions iii, v and vi.

Corollary ii: The supply of each firm is (i) price-responsive, (ii) has no perfectly elastic segments, (iii) follows the KM first-order condition in (2), and (iv) is identical to the supply of the largest firm from marginal cost up to the price at which either its capacity constraint or the price cap binds.

3.5. A unique SFE candidate that satisfies the necessary conditions

Let p_i denote the price at which the capacity constraint of firm *i* starts to bind. Recall that $\overline{\varepsilon}_1 < \overline{\varepsilon}_2 < \ldots < \overline{\varepsilon}_N$. Thus Propositions iii and iv and Corollary ii together imply the following:

Corollary iii: The supply function of the largest firm has a perfectly elastic segment at the price cap and $c < p_1 < \ldots < p_{N-1} = \overline{p}$.

An intuitive explanation of the result that the capacity constraints of small firms bind at lower prices is that small firms have less market-power and lower mark-ups for any percentage of their capacity, including full capacity.

Consider the first price-interval $[c, p_1]$, where all producers have non-binding capacity constraints. As all have identical supply functions within the interval, it follows from (4) and (5) that all firms have $\gamma_i = 0$ in the interval. Thus (4) yields

$$S_j(p) = \frac{\beta_1 (p-c)^{1/(N-1)}}{N} \qquad \text{for } j=1, 2, \ldots, N. \tag{10}$$

The subscript 1 on β is used to indicate that the constant is valid for the first price interval.

Note that as long as supply functions are continuous (Corollary i) and the number of firms with non-binding capacity constraints is constant, it follows from the KM first-order condition that slopes are continuous, i.e. there are no kinks. In the next price-interval $[p_1, p_2]$, there are *N*-1 remaining producers with non-binding capacity constraints. Following the line of argument used for the first interval, one can conclude that $\gamma_i = 0$ also in this interval. With *M*=*N*-1 it follows from (4) that

$$S_j(p) = \frac{\beta_2 (p-c)^{1/(N-2)}}{N-1}, \text{ if } p \in [p_1, p_2] \text{ and } j = 2 \ldots N.$$

Analogously, the solution for the price interval $[p_{n-1}, p_n]$ is

$$S_j(p) = \frac{\beta_n (p-c)^{1/(N-n)}}{N-n+1}, \text{ if } p \in [p_{n-1}, p_n] \text{ and } j = n \ldots N, \tag{11}$$

where $n = 1 \ldots N-1$ and $p_0 = c$. The latter is relevant for *n*=1. Combining the end-condition $p_{N-1} = \overline{p}$ with (11) yields

$$\beta_{N-1} = \frac{2\overline{\varepsilon}_{N-1}}{\overline{p}-c}. \tag{12}$$

Thus β_{N-1} can be uniquely determined. To avoid discontinuities in the supply functions — which would violate Corollary i — the following relations must be fulfilled at the boundary between two price intervals:

$$S_j(p_n) = \frac{\beta_n (p_n-c)^{1/(N-n)}}{N-n+1} = \frac{\beta_{n+1} (p_n-c)^{1/(N-n-1)}}{N-n}.$$

Thus

$$p_n = \left(\frac{(N-n)\overline{\varepsilon}_n}{\beta_{n+1}} \right)^{N-n-1} + c \tag{13}$$

and

$$\beta_n = \frac{(N-n+1)\beta_{n+1}(p_n - c)^{1/(N-n-1)-1/(N-n)}}{N-n}. \tag{14}$$

Accordingly, starting with (12), all β_n can be uniquely determined by iterative use of (13) and (14). Thus there is only one candidate that satisfies the necessary conditions for a SFE.

3.6. The only remaining equilibrium candidate is a SFE

SFE is a Nash equilibrium and no firm has a profitable deviation from the equilibrium given that competitors follow the equilibrium strategy. Consider an arbitrary producer *i*. Assume that its competitors follow the only remaining SFE candidate defined by (11) to (14). Their total supply is denoted by $\breve{S}_{-i}(p)$. This section will prove that firm *i* has no profitable deviation from its potential equilibrium bid $\breve{S}_i(p)$. The first-order condition in (2) guarantees that the profit of firm *i*, $\pi_i(\varepsilon, p)$, has an extremum at the equilibrium price for any demand outcome, but we cannot yet be sure that the extremum is a maximum for each demand outcome. This property is proven in this section and it is sufficient for a SFE.

It turns out that given $\breve{S}_{-i}(p)$, $\pi_i(\varepsilon, p)$ is globally concave in the price for each demand outcome, which is reminiscent to the result by Klemperer and Meyer (1989). One difference, however, is that the profit function is non-smooth; it has a kink at the price p_i, at which the capacity constraint of firm *i* binds. Further, due to the capacity constraint and the price cap, firm *i* can only set the market price in a range $[\tilde{p}_i(\varepsilon), \overline{p}]$. The lower boundary of the price interval is given by $\breve{S}_{-i}(\tilde{p}_i(\varepsilon)) = \varepsilon - \overline{\varepsilon}_i$, and arises when $\varepsilon > \overline{\varepsilon}_i$, because of firm *i*'s capacity constraint. For $\varepsilon \le \overline{\varepsilon}_i$, one can set $\tilde{p}_i(\varepsilon) = c$, which is the lowest bid of the competitors, and it follows from Proposition i that it is never profitable to bid below the marginal cost. As is shown in Figure 2, the type of maximum — interior maximum or boundary maximum — in the profit function depends on the demand outcome and the firm. The three outlined cases are analyzed separately.

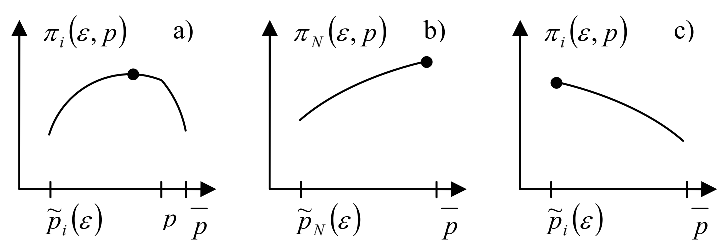

Figure 2. When $\varepsilon \le \breve{S}(p_i)$, $\pi_i(\varepsilon, p)$ has an interior max (a). The profit function of the largest firm, $\pi_N(\varepsilon, p)$, has a boundary max at $\overline{p}$ when $\varepsilon \in (\breve{S}(\overline{p}), \overline{\varepsilon}]$ (b). Smaller firms have a boundary max at $\tilde{p}_i(\varepsilon)$ when $\breve{S}(p_i) < \varepsilon < \overline{\varepsilon}$ (c).

There are obviously no profitable deviations when $\varepsilon > \overline{\varepsilon}$, then firms sell all of their capacity at the maximum price. Now consider the case $\varepsilon \le \breve{S}(p_i)$ when the market price of the equilibrium candidate does not exceed p_i, which also implies that $\widetilde{p}_i(\varepsilon) \le p_i$. Residual demand is given by (1). Hence, for given demand and price, the profit of producer *i* is

$$\pi_i(\varepsilon, p) = \underbrace{\left[\varepsilon - \breve{S}_{-i}(p)\right]}_{S_i}(p-c), \quad \text{if } p \in \left[\widetilde{p}_i(\varepsilon), \overline{p}\right). \tag{15}$$

Thus

$$\frac{\partial \pi_i(\varepsilon, p)}{\partial p} = \underbrace{\left[\varepsilon - \breve{S}_{-i}(p)\right]}_{S_i} - (p-c)\breve{S}_{-i}{}'(p), \text{if } p \in \left[\widetilde{p}_i(\varepsilon), \overline{p}\right). \tag{16}$$

With left- or right-hand derivatives, the result is also valid at p_i, where residual demand has a kink. With left-hand derivatives the result is valid for $p = \overline{p}$ as well. As in Klemperer and Meyer (1989), concavity of the profit function can be proven by means of the assumption that competitors' bids follow the first-order condition in (2). For the largest firm this is true from the marginal cost and up to the price cap.

$$\breve{S}_N(p) - \breve{S}_{-N}{}'(p)(p-c) = 0, \text{if } p \in \left[c, \overline{p}\right). \tag{17}$$

The equilibrium candidate is piece-wise symmetric, i.e. $\breve{S}_N(p) = \breve{S}_i(p)$ and $\breve{S}_{-N}{}'(p) = \breve{S}_{-i}{}'(p)$ if $p \in [c, p_i]$. The latter follows from the KM first-order condition. Moreover, $\breve{S}_i{}'(p) = 0$ if $p \in \left(p_i, \overline{p}\right]$. Thus subtracting (17) from (16) yields:

$$\frac{\partial \pi_i(p,\varepsilon)}{\partial p} = \begin{cases} \underbrace{\left[\varepsilon - \breve{S}_{-i}(p)\right]}_{S_i} - \breve{S}_N(p), \text{if } p \in \left[\widetilde{p}_i(\varepsilon), p_i\right) \\ \underbrace{\left[\varepsilon - \breve{S}_{-i}(p)\right]}_{S_i} - \breve{S}_N(p) - (p-c)\breve{S}_N{}'(p), \text{if } p \in \left(p_i, \overline{p}\right]. \end{cases} \tag{18}$$

As supply functions are piece-wise symmetric, it follows from (17) that

$$(p-c)\breve{S}'_N(p) = \frac{\breve{S}_N(p)}{\breve{M}-1},$$

where $\breve{M} \ge 2$ is the number of producers with non-binding capacity constraints at the price *p* in the potential equilibrium. Note that $\breve{M}$ decreases in the interval $\left(p_i, \overline{p}\right]$. Using the relation above, (18) can be written

$$\frac{\partial \pi_i(p,\varepsilon)}{\partial p} = \begin{cases} \underbrace{\left[\varepsilon - \breve{S}_{-i}(p)\right]}_{S_i} - \breve{S}_N(p), \text{if } p \in \left[\widetilde{p}_i(\varepsilon), p_i\right) \\ \underbrace{\left[\varepsilon - \breve{S}_{-i}(p)\right]}_{S_i} - \breve{S}_N(p) - \dfrac{\breve{S}_N(p)}{\breve{M}-1}, \text{if } p \in \left(p_i, \overline{p}\right]. \end{cases} \tag{19}$$

$\breve{S}_{-i}(p)$ and $\breve{S}_N(p)$ are both monotonically increasing in *p* and $\breve{M}$ is decreasing in *p*. Accordingly, $\frac{\partial \pi_i(p,\varepsilon)}{\partial p}$ is monotonically decreasing in *p* within the interval $\left[\widetilde{p}_i(\varepsilon), \overline{p}\right]$ and the profit function is concave in *p*. Thus the price *p*, at which $\frac{\partial \pi_i(p,\varepsilon)}{\partial p} = 0$, globally maximizes the profit for the demand outcome *ε*.

From (19) it can be deduced that producer *i* achieves the optimal price by following the piece-wise symmetric equilibrium candidate $\breve{S}_i(p)$, for which $\breve{S}_i(p) = \breve{S}_N(p)$ when $p \le p_i$. Thus there are no profitable deviations when $\varepsilon \le \breve{S}(p_i)$.

The largest firm requires special attention, as its bid is constrained by the price cap for the demand outcomes $\varepsilon \in \left(\breve{S}(\overline{p}), \overline{\varepsilon}\right]$ when all competitors' capacity constraints are binding.[6] For the largest firm (19) can be written

$$\frac{\partial \pi_N(p,\varepsilon)}{\partial p} = \underbrace{\left[\varepsilon - \breve{S}_{-N}(p)\right]}_{S_N} - \breve{S}_N(p), \text{if } p \in \left[\tilde{p}_N(\varepsilon), \overline{p}\right].$$

Also in this case $\frac{\partial \pi_N(p,\varepsilon)}{\partial p}$ is monotonically decreasing in *p* within the interval $\left[\tilde{p}_N(\varepsilon), \overline{p}\right]$ and the profit function is concave in *p*. Further, $\frac{\partial \pi_{NI}(\overline{p},\varepsilon)}{\partial p} > 0$ (the left derivative) as $\varepsilon - \breve{S}_{-N}(\overline{p}) - \breve{S}_N(\overline{p}) > 0$ when $\varepsilon \in \left(\breve{S}(\overline{p}), \overline{\varepsilon}\right]$. However, profit cannot be raised as price cannot be increased beyond the price cap. Therefore, producer *N* cannot do better than to follow $\breve{S}_N(p)$.

Now consider the remaining case $\breve{S}(p_i) < \varepsilon < \overline{\varepsilon}$ and *i*<*N*. For these demand outcomes, firm *i* offers all of its capacity below the market price in the potential equilibrium. The firm cannot push down the market price even if it lowers its bid. Thus $\tilde{p}_i(\varepsilon) > p_i$. In this case (19) can be written

$$\frac{\partial \pi_i(p,\varepsilon)}{\partial p} = \underbrace{\left[\varepsilon - \breve{S}_{-i}(p)\right]}_{S_i} - \breve{S}_N(p) - \frac{\breve{S}_N(p)}{\breve{M}-1}, \text{if } p \in \left(\tilde{p}_i(\varepsilon), \overline{p}\right] \qquad (20)$$

Once again $\frac{\partial \pi_i(p,\varepsilon)}{\partial p}$ is monotonically decreasing in *p* within the interval $\left[\tilde{p}_i(\varepsilon), \overline{p}\right]$ and the profit function is concave in *p*. As firm *i* cannot push down the market price when $\breve{S}(p_i) < \varepsilon < \overline{\varepsilon}$, competitors with non-binding capacity constraints, e.g. firm *N*, will always sell more power than firm *i*. Accordingly, it follows from (20) that $\frac{\partial \pi_i(p,\varepsilon)}{\partial p} < 0$. Thus producer *i* maximizes profit by minimizing the market price, which occurs if it offers all its capacity at or below $\tilde{p}_i(\varepsilon)$, as in the potential equilibrium, i.e. there are no profitable deviations when $\breve{S}(p_i) < \varepsilon < \overline{\varepsilon}$

The conclusion is that given $\breve{S}_{-i}(p)$, where *i*=1,2,...,*N*, the unique equilibrium candidate $\breve{S}_i(p)$ globally maximises the profit of firm *i* for every demand outcome *ε*. Thus the unique equilibrium candidate is a SFE.

---

[6] Recall that supply functions are left continuous. Thus $\breve{S}(\overline{p})$ does not include the largest firm's perfectly elastic supply at the price cap.

4. EXAMPLE 1 — THREE ASYMMETRIC PRODUCERS

Consider a market with three producers. The producers have identical and constant marginal cost c. Assume that the producers have $1/6$, $1/3$ and $1/2$ of total up-regulation capacity $\bar{\varepsilon}$ and that $\bar{p} = 3c$. Order the producers according to their up-regulation capacity so that producer 1 has the smallest capacity.

Firms 2 and 3 are symmetric up to the price cap, at which the capacity constraint of firm 2 starts to bind, i.e. $S_2(\bar{p}) = S_3(\bar{p}) = \frac{\bar{\varepsilon}}{3}$. The remaining capacity of firm 3 is sold at the price cap. Thus

$$p = \bar{p} = 3c, \text{ when } \frac{\bar{\varepsilon}}{3} \leq S_3 \leq \frac{\bar{\varepsilon}}{2}. \tag{21}$$

β_2 can be calculated from (12):

$$\beta_2 = \frac{2\bar{\varepsilon}/3}{2c} = \frac{\bar{\varepsilon}}{3c}.$$

Thus according to (11),

$$S_2(p) = S_3(p) = \frac{\bar{\varepsilon}(p-c)}{6c}, \text{ if } p \in [p_1, \bar{p}), \tag{22}$$

where p_1 is the price at which the capacity constraint of firm 1 binds. By means of (13) it can be shown that $p_1 = 2c$. Now, (14) can be used to calculate β_1:

$$\beta_1 = \frac{3\beta_2 c^{1/2}}{2} = \frac{\bar{\varepsilon}}{2c^{1/2}}.$$

According to (11)

$$S_1(p) = S_2(p) = S_3(p) = \frac{\bar{\varepsilon}}{6}\left(\frac{p-c}{c}\right)^{1/2}, \text{ if } p \in [c, p_1]. \tag{23}$$

The unique SFE, which is characterized by equations (21) to (23), is presented in Figure 3. All supply functions are symmetric up to the price p_1, at which point the capacity constraint of the smallest firm binds. Above this price, firms 2 and 3 have symmetric supply functions up to $\bar{p} = 3c$, at which point the capacity constraint of producer 2 binds. The remaining supply of producer 3 is offered with perfect elasticity at $p = \bar{p}$. The equilibrium is piece-wise symmetric like the SFE derived by Newbery (1991) for a duopoly facing linear demand.

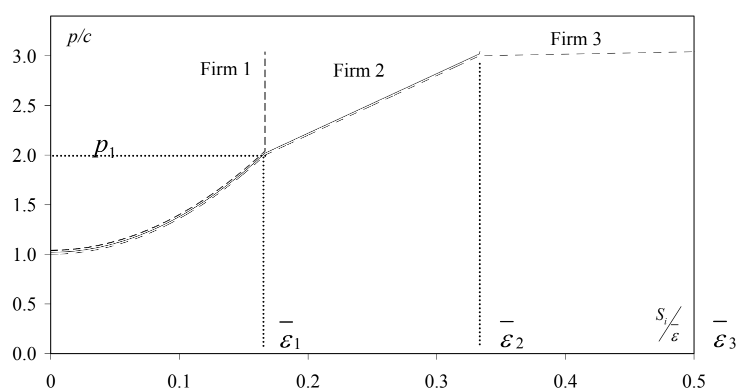

Figure 3. The unique supply function equilibrium is piece-wise symmetric if firms have identical constant marginal costs c.

In Figure 3 we can note that the supply functions of all producers become perfectly elastic at the point where the price approaches c in the limit. By differentiating (10), it is straightforward to verify that this is a general result for identical and constant marginal costs if N>2. The supply functions of firm 2 and 3 have kinks at p_1. It may seem counterintuitive that their elasticities increase at this price when the number of competitive bidders decreases. However, the increased elasticity is the only way to compensate the decreased elasticity in the supply of firm 1 and to ensure that the elasticity of the residual demand of both producer 2 and 3 is continuous at p_1. Figure 4 presents competitors' total supply functions for each of the producers. It is evident that firm 2 and 3 have a smooth residual demand from the marginal cost and up to the price cap. The elasticity of the residual demand of firm 1 increases discontinuously at p_1. This gives the firm incentives to be more competitive above this price, i.e. increase its supply. However, because of its capacity constraint this deviation is not possible. The residual demand of both firm 1 and 2 becomes perfectly elastic at the price cap. Once again deviations are not possible, because of their capacity constraints.

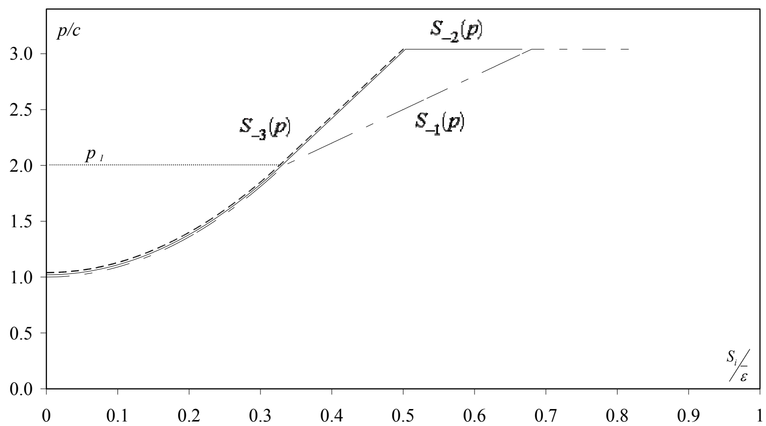

Figure 4. Competitors' total supply functions for each of the firms.

5. EXAMPLE 2 — THE NORWEGIAN REAL-TIME MARKET

The simple model is mainly useful for a qualitative understanding of bidding in electricity markets. However, it can also be used for quantitative calculations for markets dominated by one technology. More than 99 percent of electric power production in Norway is hydroelectric. The Norwegian real-time market is a uniform-price auction with a price cap at 50 000 NOK/MWh (≈6000 €/MWh). Firms can make real-time adjustments relative to their contracted position without making formal bids to the real-time market, but such adjustments are similar to making a very competitive bid that is always accepted, as the real-time price is the unit price for all deviations. During the 1990s, the Nordic countries developed a common day-ahead market. However, the real-time markets were formally separated until the beginning of the 2000s, and in this simple example it is assumed that Norway has a real-time market of its own. In the two calculations, the forward positions are taken as given. First, it is assumed that no capacity is sold forward for the delivery period under study and then that all firms sell half of their capacity forward for the delivery period under study.

The marginal cost of hydropower is very small, roughly c=50 NOK/MWh (≈6 €/MWh). But water is a limited resource. Thus hydropower bids are often driven by opportunity cost, the revenue from selling a unit of hydropower at a later day/hour. To avoid this complication, we consider an hour in the late spring when the alternative to power production is to spill water. The calculation of the SFE is based on the installed capacity of the 153 largest hydropower producers in Norway. The remaining firms and non-hydroelectric power are not included in the sample. The 10 largest producers in Norway and their share of the installed capacity are listed in Table 1.

Table 1: *Share of installed hydroelectric capacity of Norway's 10 largest power producers*

Company	Share of installed capacity	HHI
Statkraft	31.7%	1001.986
E-CO	7.3%	53.02816
Lyse	5.6%	31.04265
BKK	5.6%	30.92214
Norsk Hydro	4.8%	23.3467
Agder Energi	4.3%	18.37795
Skagerak Kraft	3.8%	14.52085
Otra Kraft	3.1%	9.856047
Trondheim Energiverk	2.7%	7.246727
Nord-Trøndelag Elverk	2.0%	4.20108
Rest (143 firms)	29.1%	27.2
Total	100%	1222

Source: Data provided by the *Norwegian Water Resources and Energy Directorate (NVE*).

The Herfindahl-Hirschman index (HHI) is a commonly used measure of market concentration and is given by the sum of market shares (in percentages) squared, see e.g. Tirole (2003). The Norwegian electric power market has HHI=1222 which corresponds to slightly less than 8 symmetric firms (HHI=1250). Thus it is interesting to compare the supply function equilibria of these two cases. By means of the formulae in Section 3.5, it is straightforward to calculate the supply of all asymmetric firms and the 8 symmetric firms. In the symmetric case, the capacity constraint of all firms will start to bind at the price cap, as shown in Holmberg (2004).

In the first calculation all power is assumed to be sold in real-time. In the asymmetric model nearly 20 firms have non-binding capacity constraints up to 40 percent of market capacity. Thus, as can be seen in Figure 5, there is almost perfect competition in this demand range, the mark-up is less than one percent. For the last 50 percent of the capacity, mark-ups are excessive, more than 200 percent. For small levels of demand, the asymmetric market is more competitive than the symmetric market, because many asymmetric firms have non-binding capacity constraints. On the other hand, the symmetric market is more competitive for large levels of demand when most asymmetric firms have binding capacity constraints. At 55 percent of market capacity, the number of firms with non-binding capacity constraints equals the number of symmetric firms and the mark-ups in the two models are also roughly the same.

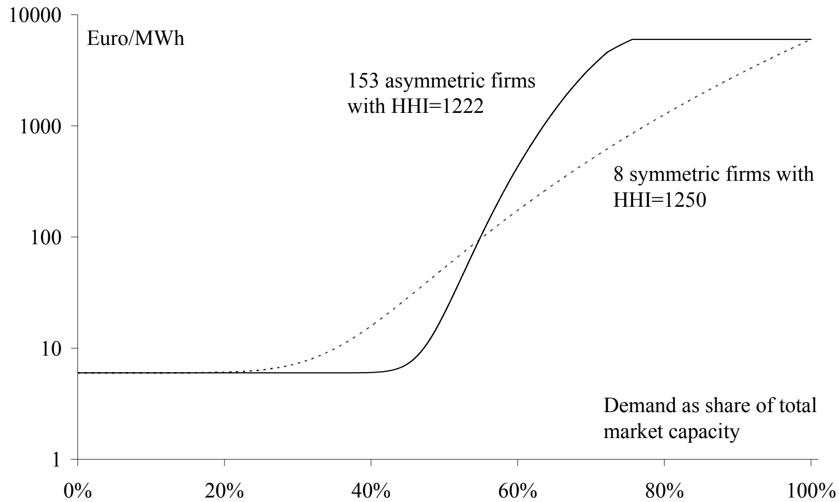

Figure 5. The unique supply function equilibrium of the real-time market in Norway compared to a case with 8 symmetric firms.

In the late spring, the typical demand in Norway is roughly 50 percent of market capacity. The result in Figure 5 would then imply a mark-up of more than 200%. To get a more realistic calculation, producers are now assumed to sell half of their capacity in the forward market. This means that producers will also offer decrements, balancing-down bids, in case realized demand turns out to be below half market capacity. Then the real-time market is a sales auction, i.e. the system operator sells power back to producers. Firms' alternative to not buying back power is to pay the marginal cost to meet their contracted production. Thus to make money on their balancing-down bids, firms offer to buy back power below marginal cost. The firm with the most down-regulation capacity, i.e. the firm with most market-power, will have the largest mark-downs in its balancing-down bid. In particular, in case of extreme demand outcomes that exhaust competitors' down-regulation capacity, the largest firm will offer to buy back power at the lowest possible price, the price floor $\underline{p}$, as shown in Figure 6. The equilibrium of the sales auction is anti-symmetric to a corresponding procurement auction. Let $\underline{\varepsilon}$ represent the total down-regulation capacity in the market. Imagine a fictive market for which $c^* = -c$, $\overline{p}^* = -\underline{p}$ and $\overline{\varepsilon}^* = -\underline{\varepsilon}$. Then the balancing-down equilibrium is anti-symmetric to the balancing-up equilibrium of the fictive market, i.e. $S_j(p) = -S_j^*(-p)$.

It would be natural to set the price floor equal to zero. However, in the calculation example it is set to 0.1 NOK (0.012 €/MWh), because of the logarithmic scale in Figure 6. In the new example, the real-time price is very close to the marginal cost, unless there is a large demand shock in the market that would require a production increase larger than 20% of market capacity or a decrease larger than 30% of market capacity. A close look at the data reveals that the asymmetric market is more competitive than the symmetric market for decrements below 27% of market capacity. Thus as in the previous calculation, the asymmetric market is more competitive for small imbalances and less competitive for large imbalances.

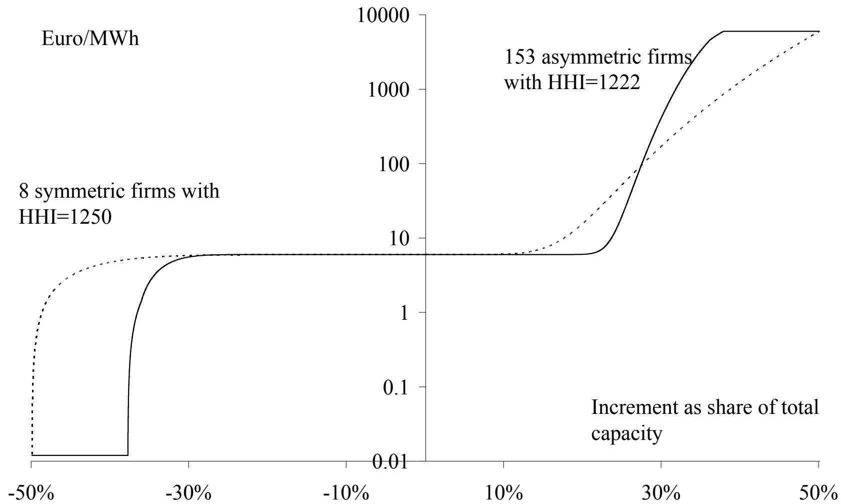

Figure 6. The supply function equilibrium of the real-time market when half of firms' capacities are sold in the forward market.

6. CONCLUSIONS

In the real-time market, also called balancing market, firms can offer increased production, up-regulation/balancing up, and decreased production, down-regulation/balancing down. The system operator accepts balancing-up bids when additional power production is needed to meet current demand. Similarly, the system operator accepts balancing-down bids when current power production exceeds current demand. The paper focuses on the up-regulation case, but corresponding results can readily be derived for the down-regulation case.

A Supply Function Equilibrium (SFE) is derived for a balancing market with multiple electric power producers having identical and constant marginal costs, but asymmetric capacities. The equilibrium is unique if maximum demand is sufficiently large and minimum demand sufficiently low. The up-regulation capacity of all firms, except possibly the largest, is assumed to bind at maximum demand. Similarly, minimum demand is so low that the down-regulation capacities bind for all firms, except possibly the largest. The unique equilibrium is piece-wise symmetric; two arbitrary producers have the same supply function until the capacity constraint of the smaller firm binds. The constraint of the producer with the second largest up-regulation capacity starts to bind at the price cap, while the capacity constraints of smaller firms bind below the price cap. The largest firm offers its remaining up-regulation capacity as a perfectly elastic supply at the price cap.

At the price for which the capacity constraint of a small firm starts to bind, the elasticity of the supply of larger firms will increase discontinuously. This ensures that firms with non-binding capacity constraints have smooth residual demand. The unique equilibrium has supply functions with kinks as well as perfectly elastic and inelastic segments. This implies that, in markets with asymmetric firms, one cannot limit attention to smooth supply functions, as is often the case in the SFE literature.

The equilibrium will only be piece-wise symmetric if costs are. However, in other aspects the qualitative properties of the asymmetric SFE are conjectured to hold for a more general class of cost functions. With the reservation that a valid SFE — with non-decreasing supply functions — must exist, this understanding can be used to numerically calculate valid SFE for asymmetric firms with increasing marginal costs as in Holmberg (2005).

Compared to a symmetric SFE with the same Herfindahl-Hirschman Index (HHI), the asymmetric SFE is much more competitive for small demand outcomes when the capacity constraints of few asymmetric firms bind. On the other hand, the asymmetric SFE is less competitive for large demand outcomes when the capacity constraints of most asymmetric firms bind. Intuitively, this result should also hold for increasing cost functions. As it is complicated to calculate asymmetric SFE for general cost functions, this result is useful when interpreting approximate symmetric SFE models of real electricity markets.

Even if uniqueness of the equilibrium depends on the support of the probability distribution of demand, the equilibrium itself is independent of this distribution; it is a supply function equilibrium for any distribution of demand. Thus using the notion introduced by Anderson and Hu (2006), one could call it a strong supply function equilibrium.

REFERENCES

Anderson, Edward. J., and Andrew B. Philpott (2002). "Using supply functions for offering generation into an electricity market", *Operations Research*, 50(3): 477-489.

Anderson, Edward. J., and Xinmin Hu (2006). "Finding Supply Function Equilibria with Asymmetric Firms", Workshop on Supply Function Equilibrium Modelling of Electricity Markets, April 2006, Auckland, New Zealand.

Baldick, Ross, Ryan Grant and Edward Kahn (2004).,"Theory and Application of Linear Supply Function Equilibrium in Electricity Markets", *Journal of Regulatory Economics*, 25(2): 143 – 167.

Baldick, Ross, and William Hogan (2002). *Capacity constrained supply function equilibrium models for electricity markets: Stability, non-decreasing constraints, and function space iterations*, POWER Paper PWP-089, University of California Energy Institute.

Bolle, Friedel (1992). "Supply function equilibria and the danger of tacit collusion. The case of spot markets for electricity", *Energy Economics,* 14: 94-102.

Genc, Talat, and Stanley Reynolds (2004). *Supply Function Equilibria with Pivotal Electricity Suppliers*, Eller College Working Paper No.1001-04, University of Arizona.

Green, Richard, and David Newbery (1992). "Competition in the British Electricity Spot Market", *Journal of Political Economy,* 100(5): 929-953.

Green, Richard (1996). "Increasing Competition in the British Electricity Spot Market", *Journal of Industrial Economics*, XLIV(2): 205-16.

Holmberg, Pär (2004). *Unique supply function equilibrium with capacity constraints*, Working Paper 2004:20, Department of Economics, Uppsala University, forthcoming, Energy Economics.

Holmberg, Pär (2005). *Numerical calculation of asymmetric supply function equilibrium with capacity constraints*, Working Paper 2005:12, Department of Economics, Uppsala University.

Klemperer, Paul, and Margaret A. Meyer (1989). "Supply function equilibria in an oligopoly under uncertainty", *Econometrica,* 57: 1243-1277.

Newbery, David (1991). "Capacity-constrained Supply Function Equilibria: Competition and Entry in the Electricity Spot Market", Unpublished manuscript, Cambridge: University of Cambridge, Department of Applied Economics.

Rudkevich, Aleksandr, Max Duckworth and Richard Rosen (1998). "Modeling electricity pricing in a deregulated generation industry: The potential for oligopoly pricing in poolco", *The Energy Journal*, 19(3): 19-48.
Rudkevich, Aleksandr (2005), "On the Supply Function Equilibrium and its Applications in Electricity Markets", forthcoming, *Decision Support Systems*.
Tirole, Jean (2003). *The Theory of Industrial Organization*, Cambridge, MA: MIT Press.

APPENDIX

Proof of Proposition iii. The proposition can be divided into three claims:

a) For every $p>c$, one can find a sufficiently high $p_-<p$ such that all supply functions are twice continuously differentiable in the interval $[p_-, p]$. Either no firm or at least two firms have price-responsive supply in the interval.

b) In equilibrium, two or more firms cannot have supply functions with perfectly elastic segments at the same price $p^* \in (c, \overline{p}]$.

c) In equilibrium, no producer has a supply function with a perfectly elastic segment at the marginal cost c.

d) In equilibrium, one firm cannot have a perfectly elastic segment at $p^* \in (c, \overline{p})$.

Lemma A below proves a technicality that will be useful when proving the claims.

Lemma A: An equilibrium with a smooth transition to a perfectly inelastic or perfectly elastic aggregate supply is not possible if the price p^* is above marginal cost c. In the perfectly inelastic case, the result applies also for $p^* \geq c$. In the perfectly elastic case, the result is valid both for individual firms and in aggregate.

Proof: Consider a smooth transition to p^* from the left. According to the assumed properties of the supply functions, a sufficiently large p_- can be chosen such that all supply functions are twice continuously differentiable in the range $[p_-, p^*)$. In addition, p_- can be chosen such that each supply function is either monotonically increasing in the whole range or perfectly inelastic in the whole range. If all supply functions are perfectly inelastic in the interval $[p_-, p^*)$, smooth transitions are not possible. Thus according to Proposition ii there must be $M \geq 2$ producers with monotonically increasing supply functions in the range $p \in [p_-, p^*)$.

It follows from (6) that one can find numbers $\underline{m} > 0$ and $\overline{m} < \infty$ such that $\underline{m} \leq p'(\varepsilon) \leq \overline{m}$, for every $p(\varepsilon) \in [p_-, p^*)$ if $p_->c$. For the upper boundary this is true also when $p_- \geq c$. Thus smooth transitions from the left to a perfectly elastic aggregate supply are not possible if $p^* > c$. This is valid for individual producers as well, as $p'(\varepsilon)=0$ if the marginal bidder has a perfectly elastic supply. Similarly, smooth transitions from the left to a perfectly inelastic aggregate supply are not possible if $p^* \geq c$.

Analogous proofs can be performed to rule out smooth transitions from the right. □

Proof of a) Given the left continuous and piece-wise smooth properties of supply functions, the proof follows directly from Lemma A and Proposition ii.

Proof of b) Given claim a), the proof is almost identical to the proofs of Proposition 2 and 3 in Holmberg (2004).

Proof of c) Assume that in equilibrium, producer *i* offers $S_i^Q(c+)=\lim_{p\to c+} S_i^Q(p)>0$ units of power with perfectly elastic supply at the constant marginal cost *c.* The aggregate supply of his competitors at this price is denoted by $S_{-i}^Q(c+)\geq 0$. Thus producer *i* may be the only firm with a perfectly elastic segment at *c*.

Assume first that aggregate supply is price-responsive just above *c.* Then claim a) implies that $M\geq 2$ firms are price-responsive just above *c.* The assumed properties of the supply functions ensure that a sufficiently low p_+ can be chosen so that in equilibrium, all supply functions are twice continuously differentiable in the interval (c, p_+), and the same $M\geq 2$ producers have a price-responsive supply in the price interval. The other *N-M* producers have a perfectly inelastic supply over the whole interval. The supply functions of the *M* producers are given by (2). The aggregate supply of this group is denoted by $S_e^Q(p)$. It follows from (3) that

$$S_e^Q(p)=\beta(p-c)^{1/(M-1)} . \qquad (24)$$

Thus $S_e^Q(p)$ approaches zero as the price approaches *c.* Accordingly producer *i* and any other producers with a perfectly elastic segment at *c* cannot belong to the group with price-responsive supply in the interval (c, p_+). Hence, their supply is perfectly inelastic in this interval.

Now consider the following marginal deviation of producer *i.* The price of an infinitesimally small unit, previously offered at *c,* is increased to the price $p^* \in (c, p_+]$. The marginal change in expected profit is given by

$$\Delta E(\pi_i)=\int_{S^Q(c+)}^{S^Q(p^*)}\left\{S_i^Q(c+)p^{Q'}(\varepsilon)-\left[p^Q(\varepsilon)-c\right]\right\}f(\varepsilon)d\varepsilon.$$

The first term reflects an increased price and the second term reflects reduced sales in the demand interval. By means of (6), the integral can be written as

$$\Delta E(\pi_i)=\int_{S^Q(c+)}^{S^Q(p^*)}\left\{S_i^Q(c+)\frac{M-1}{\beta}\left(p^Q(\varepsilon)-c\right)^{(M-2)/(M-1)}-\left[p^Q(\varepsilon)-c\right]\right\}f(\varepsilon)d\varepsilon =$$
$$=\int_{S^Q(c+)}^{S^Q(p^*)}\left[p^Q(\varepsilon)-c\right]\left\{S_i^Q(c+)\frac{M-1}{\beta}\left(p^Q(\varepsilon)-c\right)^{(M-2)/(M-1)-1}-1\right\}f(\varepsilon)d\varepsilon.$$

The value $\left(p^Q(\varepsilon)-c\right)^{(M-2)/(M-1)-1}$ can be made arbitrarily large for sufficiently small $p^Q>c$. Thus $\Delta E(\pi_i)>0$ for a sufficiently small p^*. Accordingly, there are profitable deviations, and equilibria where $S_i^Q(c+)>0$ can be ruled out if aggregate supply is price-responsive just above *c.*

Next consider the case where total supply is perfectly inelastic just above c. Smooth transitions to a perfectly inelastic supply are ruled out by Lemma A. Thus if p_+ is chosen sufficiently small, the bids of all producers are now perfectly inelastic over the whole price range $(c, p_+]$. Denote by ε_0 the demand outcome for which the last unit offered at the price c is sold. For the assumed equilibrium there is no contribution to expected profit from demands below ε_0. Now consider the following unilateral deviation of producer i. Increase the price for the $S_i^Q(c+)$ units to $p^* \in (c, p_+)$. This action increases the expected profit of producer i for demand outcomes below ε_0. Supply above p_+ is not affected, nor is the contribution to expected profit from demands above ε_0. As a result, the deviation increases expected profit of producer i. Accordingly, equilibria where a producer offers $S_i^Q(c+) > 0$ units of power at the constant marginal cost c can also be excluded if aggregate supply is perfectly inelastic just above c. □

Proof of d) Denote the potential equilibrium by the superscript W. It is assumed that $p^W(\varepsilon) = p^* < \overline{p}$ if and only if $\varepsilon \in [\varepsilon', \varepsilon'']$. Let firm i be the producer with the perfectly elastic segment.

All firms cannot have perfectly inelastic supply just above p^*, as, in that case, firm i would deviate. The price of the perfectly elastic segment can be increased without losing sales. Furthermore, smooth transitions to an aggregate perfectly inelastic supply are excluded by Lemma A. Thus it follows from claim a) that at least two producers have a price-responsive supply just above p^*, i.e. $S_{-ir}^{W\,\prime}(p^*) > 0$ (the right-hand derivative). However, any competitor $j \neq i$ with a price-responsive bid just above p^* would find it profitable to deviate. He can slightly reduce the price of his units offered just above p^* and instead offer them just below p^*. The marginal change in expected profit of producer j from deviating by one, infinitesimally small unit is

$$\int_{\varepsilon'}^{\varepsilon''} (p^* - c) f(\varepsilon) d\varepsilon.$$

The deviation is profitable because $p^* > c$. □

Proof of Proposition iv. Given Corollary i, discontinuities in the equilibrium price can be ruled out with a proof analogous to the proof of Proposition 4 in Holmberg (2004). This also ascertains that the highest bid must equal the price cap. Isolated points with perfectly inelastic aggregate supply are ruled out by Lemma A. Accordingly, it follows from Proposition ii that there must be at least two firms with price-responsive supply in each price interval, from the lowest bid to the price cap. □