



**Copenhagen
Business School**
HANDELSHØJSKOLEN

Department of Economics

Copenhagen Business School

Working paper 2-2014

**Restructuring the Electricity
Industry: Vertical Structure and the
Risk of Rent Extraction**

Anette Boom and Stefan Buehler

Restructuring the Electricity Industry: Vertical Structure and the Risk of Rent Extraction

Anette Boom*
Copenhagen Business School

Stefan Buehler†
University of St. Gallen

March 14, 2014

Abstract

We study the role of vertical structure in determining generating capacities and retail prices in the electricity industry. Allowing for uncertain demand, we compare three market configurations: (i) integrated monopoly, (ii) integrated duopoly with wholesale trade, and (iii) separated duopoly with wholesale trade. We find that equilibrium capacities and retail prices are such that welfare is highest (lowest) under separated (integrated) duopoly. The driving force behind this result is the risk of rent extraction faced by competing integrated generators on the wholesale market. Our analysis suggests that vertical structure plays an important role in determining generating capacities and retail prices.

Keywords: Electricity, Investments, Generating Capacities, Vertical Integration, Monopoly and Competition.

JEL-Classification: D42, D43, D44, L11, L12, L13

*Corresponding author: Copenhagen Business School, Department of Economics, Porcelænshaven 16 A, DK-2000 Frederiksberg, e-mail: ab.eco@cbs.dk

†University of St. Gallen, Department of Economics, FGN-HSG, Varnbuelstr. 19, CH-9000 St. Gallen, e-mail: stefan.buehler@unisg.ch

1 Introduction

Electricity markets around the world have been restructured in an effort to improve their performance. In several countries, legislators have allowed competition into statutory integrated monopoly and implemented regulations such as vertical unbundling to safeguard entrants and consumers from potentially harmful strategic behavior by integrated generators.¹ Yet, no consensus seems to have emerged as to which market configuration works best. A particular concern is that allowing competition into electricity markets might undermine investments in generating capacity (see, e.g., Joskow (2006), and Joskow and Tirole (2006)).

In this paper, we study the role of vertical market structure in determining investments in electricity generating capacity, retail prices, and welfare. We compare three market configurations that vary with respect to vertical market structure and the extent to which firms compete at the wholesale and the retail level: (i) integrated monopoly, (ii) integrated duopoly with wholesale trade, and (iii) separated duopoly with wholesale trade. Throughout the analysis, we allow for uncertain demand at the retail level.

A key feature of our analysis is that both capacity decisions and retail prices are determined before the state of retail demand is known. After retail demand is realized, the wholesale price for electricity is determined in a uniform-price auction due to von der Fehr and Harbord (1997) and (1993), and deliveries and payments are exchanged. This setting implies that wholesale prices react to changes in retail prices, whereas retail prices cannot react to changes in wholesale prices. The timing reflects a peculiarity of the electricity industry: capacity decisions are made under uncertainty about future demand, just as retail delivery contracts are signed before the state of demand is realized. The wholesale market then attempts to balance supply and demand based on available capacity and effective retail demand. Our main analysis will assume that a blackout occurs if the balancing act in the wholesale market fails. In an extension, we will also consider the case where

¹In the UK, for example, the industry was vertically separated into three generating firms, the National Grid company, and 12 regional distribution companies by the Electricity Act in 1989. However, some regional distribution companies later re-integrated vertically into generation (Newbery, 1999, 2005). The Californian restructuring bill from 1996 also forced the regulated utilities to sell lots of their generation facilities (Borenstein, 2002). The European Union ruled in its Directive 2003/54/EC concerning common rules for the internal market in electricity adopted on 26 June 2003 that electricity generating firms which are integrated into transmission and distribution have to be functionally disintegrated.

blackouts may be avoided by the rationing of retail demand.²

Our main results are the following. *First*, aggregate generating capacity is highest under integrated duopoly and lowest under integrated monopoly. The separated duopoly yields an intermediate level of generating capacity. The driving force behind this result is the rent extraction risk faced by an integrated duopoly generator: if individual capacity turns out to be too small to serve own retail demand, an integrated generator must buy electricity from its competitor in the wholesale market, thereby fully dissipating its rent. To avoid this outcome, an integrated duopoly generator will not only choose a large generating capacity, but also set a high retail price. Vertical separation eliminates this risk of rent extraction, as electricity generators are not committed to serve an uncertain demand at a pre-determined retail price. As a result, retail prices are lower and demand is higher than under vertical integration. In effect, vertical separation reduces the investment-enhancing effect of introducing competition into statutory monopoly, but does not fully eliminate it.

Second, equilibrium retail prices are lowest under separated duopoly and highest under integrated duopoly. The integrated monopoly yields an intermediate level of retail prices. Intuitively, the result follows again from the risk of rent extraction, which induces an integrated generator to charge a higher retail price. This result supports the notion that allowing competition into the electricity industry (alone) does not necessarily reduce retail prices.

Third, the combined effects of restructuring on investments in generating capacities and retail prices are such that social welfare is highest under separated duopoly and lowest under integrated duopoly. The integrated monopoly yields an intermediate level of social welfare. To understand the intuition for this result, note that irrespective of market configuration capacity is always large enough to satisfy retail demand at the relevant retail price (i.e., blackouts do not occur in equilibrium). This implies that low capacity investments do not have an adverse effect on welfare per se. Increasing capacity investments has two effects. (i) it raises generation costs without improving supply security; (ii) it increases the maximum level of retail demand that can be served without causing blackouts. To benefit from higher capacities, retail prices must decrease, which requires a restructuring from integrated monopoly to separated duopoly. Therefore, the welfare ranking is essentially a reversed ranking of the price levels under the various industry configurations.

Allowing for rationing of the retail demand does not undermine our analy-

²See section 5.1.

sis. The equilibrium outcome for the integrated duopoly remains unaffected, as the risk of rent extraction induces firms to choose capacities which make rationing unnecessary. In the other market configuration, equilibrium capacities and retail prices are no longer distorted upwards to avoid blackouts. Instead, some rationing occurs for very high demand realization, leading to a lower supply security compared to the integrated duopoly. Yet, this higher supply security affects the welfare ranking of the integrated duopoly configuration only for very low capacity costs.

Our paper contributes to the extensive literature on the impact of demand uncertainty on capacity choices (Drèze and Sheshinski (1976), Gabszewicz and Poddar (1997), von der Fehr and Harbord (1997), Castro-Rodriguez *et al.* (2009), Boom (2002) and (2009), Borenstein and Holland (2005)), Murphy and Smeers (2005), and Grimm and Zoettl (2013)). The key difference to this literature is that we focus on the role of vertical market structure in determining capacity choices.

We also contribute to the much scarcer literature on the competitive effects of changing the electricity industry's vertical structure. Previous contributions to this strand of the literature focus either on the loss of vertical economies due to the separation of generation and distribution of electricity (see e.g. Kwoka *et al.* (2010) and Kwoka (2002)), or on the effect of forward contracts on the wholesale prices (see e.g. Bushnell (2007), Mansur (2007), Bushnell *et al.* (2008), de Frutos and Fabra (2012) and Bosco *et al.* (2012)). None of these papers studies the role of vertical structure in determining the investments in generating capacities.³ We identify a new effect, the risk of rent extraction associated with vertical integration, as an important determinant of investments in generating capacities.

The remainder of the paper is structured as follows. Section 2 introduces the analytical framework. Section 3 discusses the benchmark cases of social optimum, integrated monopoly, and integrated duopoly. Section 3.4 studies the equilibrium outcome under separated duopoly. Section 4 compares the various market configurations and derives the main results. Section 5 discusses some limitations and explores rationing at the retail level as an extension. Section 6 concludes.

³Some authors interpret the integration of electricity generators into the retail market as forward contracting in the tradition of Allaz and Villa (1993), abstracting from the fact that vertically integrated firms commit on retail prices rather than retail quantities. It is worth noting that their empirical observation of lower wholesale prices with vertical integration is in line with our model if firms successfully avoid rent extraction (the wholesale price is then low despite a high retail price).

2 Analytical Framework

In this section, we outline the analytical framework for the various market configurations considered below, building on Boom (2009) and Boom (2007).

2.1 Demand

Suppose that customers' surplus is given by

$$V(x, \varepsilon, r) = U(x, \varepsilon) - rx = x - \varepsilon - \frac{(x - \varepsilon)^2}{2} - rx, \quad (1)$$

where $x \geq 0$ is the amount of electricity consumed, $r \geq 0$ is the retail price per unit of electricity, and ε is a demand shock, uniformly distributed on the interval $[0, 1]$.⁴ Maximizing $V(x, \varepsilon, r)$ with respect to x yields the linear retail demand for electricity

$$x(r, \varepsilon) = \max\{1 + \varepsilon - r, 0\}. \quad (2)$$

If there is more than one retailer, consumers subscribe to the retailer offering the lowest retail price (electricity is a homogeneous good). If retail prices are identical, consumers choose each retailer with equal probability.

2.2 Supply

We will compare three market configurations that differ in the number of active firms and the vertical market structure and the social optimum:⁵

- (i) *social optimum*;
- (ii) *integrated monopoly*;
- (iii) *integrated duopoly* with wholesale trade: Two integrated firms can buy and sell electricity on the wholesale market and serve retail demand;
- (iv) *separated duopoly* with wholesale trade: Two separated generators sell to the wholesale market, and two separated retailers buy from the wholesale market to serve retail demand.

⁴This specification implies that a positive demand shock ($\varepsilon > 0$) is associated with a negative effect on consumer surplus. Our key results do not depend on this specification (see section 5.2 below for a discussion).

⁵We abstract from the chain of (vertically separated) monopolies, which does not give rise to a sensible market configuration in our setting.

For simplicity, we assume that the marginal cost of generating electricity is constant and normalized to zero. The total cost of electricity generator $i = A, B$ is then given by

$$C(k_i) = zk_i, \quad (3)$$

where z is the constant unit cost of capacity and k_i is the generating capacity installed by firm i .⁶ We assume that capacity cost z satisfies

$$0 \leq z < \frac{1}{2} \quad (4)$$

to ensure strictly positive capacity investments in all market configurations. For simplicity, we further assume that the marginal cost of selling electricity to consumers is constant and normalized to zero

2.3 Timing

The timing reflects some key features of the electricity industry. We first consider the duopoly configurations, which presume the following five stages:

- (1) In the first stage, generators $i = A, B$ decide on capacities k_i before retail demand is known. In the integrated duopoly, capacity decisions are taken simultaneously. In the separated duopoly, we consider both simultaneous and sequential capacity decisions (assuming that A moves before B).
- (2) In the second stage, retailers $\ell = C, D$ simultaneously set retail prices r_ℓ in the separated duopoly, whereas generators $i = A, B$ simultaneously set retail prices r_i in the integrated duopoly. Consumers buy from the firm with the lower retail price, or, if prices are identical, from each firm with equal probability.
- (3) In the third stage, the demand shock $\varepsilon \in [0, 1]$ is realized. Since retail prices are already set, retail demand is fixed henceforth.
- (4) In the fourth stage, generators bid prices p_A and p_B for their full capacity $k_i, i = A, B$ to an auctioneer. The auctioneer determines the market clearing wholesale price p (if such a price exists) and the amount of electricity generators may supply to the grid.

⁶Firm indices may be ignored if there is only one generator.

- (5) Finally, in the fifth stage, if supply and demand are balanced, deliveries and payments are exchanged. If supply and demand cannot be balanced, a blackout occurs and market exchange is interrupted.

The monopoly configuration and the social welfare benchmark reflect the timing in the duopoly scenarios as closely as possible. In particular, both the social planner and the integrated monopoly must choose their capacity and retail price before retail demand is known.

2.4 Wholesale Market

As the wholesale price is determined after the retail price, the timing is reversed relative to the standard literature on vertically related markets. This reversion reflects the peculiarity that retailers must specify the terms of delivery before retail demand is known and then buy electricity on behalf of their customers on the wholesale market. That is, the retail market clears in the long run, whereas the wholesale market clears in the short run. Even though the wholesale price cannot affect retail prices, it is an important determinant of investments in generating capacity, as it affects the returns on investment for electricity generators. To fix ideas, we assume that the wholesale price is determined by a *uniform-price auction* due to von der Fehr and Harbord (1997) and (1993).⁷ Uniform-price auctions were used for the Electricity Pool in England and Wales before the reform in 2001, and are still in use elsewhere (e.g., for the Nord Pool in Scandinavia, or the Spanish wholesale market).⁸

The uniform-price auction we employ requires each firm i to bid a price p_i at which it is willing to supply its total capacity.⁹ The auctioneer then attempts to balance supply and demand on the grid, arranging the bids in ascending order and determining the marginal bid which equates supply and demand.¹⁰ The price of the marginal bid is the spot market price paid to all generators

⁷An alternative approach, based on Klemperer and Meyer (1989), has been suggested by Green and Newbery (1992). They assume that firms bid differentiable supply functions, whereas von der Fehr and Harbord (1997) and (1993) assume that they bid step functions.

⁸See Bergman *et al.* (1999), Crampes and Fabra (2005) and Newbery (2005).

⁹That is, we abstract from the problem of strategic capacity withholding (see Crampes and Creti (2005), and Le Coq (2002)).

¹⁰For simplicity, we ignore transmission constraints, although they might interact with constraints in the generating capacity. See Wilson (2002) for insights into this problem and for the analysis of isolated transmission constraints Borenstein *et al.* (2000), Joskow and Tirole (2000) and Léautier (2001)

for each unit that is dispatched on the grid (irrespective of individual bids).¹¹ The capacity of suppliers bidding below the spot market price is dispatched completely, whereas the marginal supplier delivers the residual amount which equates supply and demand.¹²

Since neither retail demand nor generating capacities can react to changes in the wholesale price, the auctioneer may fail to find a wholesale price that equates supply and demand. In this case, a blackout occurs, and market exchange is interrupted.¹³ In section 5.1, we also consider the case where blackouts are avoided by rationing at the retail level.

3 Alternative Market Configurations

In this section, we characterize the equilibrium outcomes for the various market configurations, drawing in part from Boom (2009) and Boom (2007).

3.1 Social Optimum

The social planner's capacity choice can be derived from the maximization of the social welfare function with respect to the retail price r and the capacity level k , respectively. Social welfare is given by

$$W(r, k) = \begin{cases} \int_0^1 U(x(r, \varepsilon), \varepsilon) d\varepsilon - zk & \text{if } r \geq 2 - k, \\ \int_0^{k-1+r} U(x(r, \varepsilon), \varepsilon) d\varepsilon \\ \quad + \int_{k-1+r}^1 U(0, \varepsilon) - zk & \text{if } 2 - k > r \geq \max\{1 - k, 0\}, \\ \int_0^1 U(0, \varepsilon) - zk & \text{if } 1 - k > r \geq 0, \end{cases} \quad (5)$$

where $U(\cdot)$ is the consumer surplus and $x(r, \varepsilon)$ is retail demand (from equations (1) and (2), respectively). The first segment of $W(r, k)$ is relevant when the retail price r is sufficiently large for retail demand to be smaller than capacity, even for the highest possible demand shock $\varepsilon = 1$.¹⁴ The second

¹¹Note the difference to Kreps and Scheinkman (1983), where the undercutting firm receives its *own* price per unit sold even if its capacity is too low to serve the market, so that some customers have to pay the price of the competitor with the next higher price.

¹²In line with Wilson (2002) we consider an integrated system because participation in the auction is compulsory if a generating firm wants to sell electricity.

¹³Under separated duopoly, a blackout will also occur if the wholesale price is larger than the retail price ($p > r$). In this case, retailers must declare bankruptcy and exit the market, so that generators can no longer sell electricity.

¹⁴In this case, the condition $k \geq (1 + \varepsilon - r)$ becomes equal to $r \geq 2 - k$. The lower bound of the integral assures a positive demand.

segment is relevant when r is in an intermediate range, such that demand is smaller than capacity for the lowest possible demand shock $\varepsilon = 0$ but larger than capacity for the highest possible demand shock $\varepsilon = 1$.¹⁵ Finally, the third segment is relevant if the retail price is low enough for demand to exceed capacity even for the smallest possible demand shock $\varepsilon = 0$. For this setting, one can show that the optimal retail price satisfies $r = \max\{2 - k, 0\}$. Substituting this price into the social welfare function and maximizing with respect to total capacity k yields proposition 1.

Proposition 1 (social optimum) *The welfare maximizing retail price and generating capacity, respectively, are given by $r^s = z$ and $k^s = 2 - z$.*

Proof: Boom (2007), Appendix B. ■

Note that the retail price (generating capacity, respectively) is increasing (decreasing) in capacity costs. In addition, the optimal retail price is such that there are no blackouts.

3.2 Integrated Monopoly

The integrated monopoly chooses the retail price r^m and the capacity level k^m so as to maximize expected profits¹⁶

$$\pi(r, k) = \begin{cases} \int_{\max\{r-1, 0\}}^1 r(1 + \varepsilon - r)d\varepsilon - zk, & \text{if } r \geq 2 - k, \\ \int_{\max\{r-1, 0\}}^{k-1+r} r(1 + \varepsilon - r)d\varepsilon - zk, & \text{if } \max\{0, 1 - k\} \leq r < 2 - k, \\ -zk, & \text{if } r \leq \max\{0, 1 - k\}. \end{cases}$$

Maximizing w.r.t. the retail price yields $r = \max\{2 - k, 3/4\}$. Substituting back into expected profits and maximizing w.r.t. capacity yields proposition 2.

Proposition 2 (integrated monopoly) *The profit maximizing retail price and generating capacity, respectively, are given by $r^m = \frac{3}{4} + \frac{z}{2}$ and $k^m = \frac{5}{4} - \frac{z}{2}$.*

Proof: Boom (2007), Appendix A. ■

Again, the retail price (generating capacity) is increasing (decreasing) in capacity costs. The retail price is higher and capacity lower than in the social optimum, but again there are no blackouts.

¹⁵The former requires $r \geq \max\{0, 1 - k\}$, and the latter $r < 2 - k$. The upper bound of the integral assures that there is no blackout (i.e., capacity is sufficient to satisfy demand).

¹⁶The different segments of the monopoly's profit function can be explained in the same way as the segments of the social welfare function.

3.3 Integrated Duopoly

Integrated duopoly generators $i = A, B$ are active both in the retail market and the wholesale market. The wholesale price is determined by a uniform-price auction. When the auction is held, aggregate retail demand is fixed and given by $x(r^{id}, \varepsilon)$, where $r^{id} = \min\{r_A, r_B\}$ is the equilibrium retail price under integrated duopoly. Firm i 's retail demand, in turn, is

$$d_i(r_A, r_B, \varepsilon) = \begin{cases} x(r_i, \varepsilon) & \text{if } r_i < r_j, \\ \frac{1}{2}x(r_i, \varepsilon) & \text{if } r_i = r_j, \\ 0 & \text{if } r_i > r_j, \end{cases} \quad \text{with } i, j \in \{A, B\}, i \neq j. \quad (6)$$

3.3.1 Wholesale Market and the Rent Extraction Risk

If aggregate capacity is sufficient to satisfy demand, i.e. $k_A + k_B \geq x(r^{id}, \varepsilon)$, the wholesale price is given by

$$p(p_A, p_B) = p_i \begin{cases} \text{if } p_i < p_j & \text{and } k_i \geq x(r^{id}, \varepsilon) \text{ or} \\ \text{if } p_i \geq p_j & \text{and } k_j < x(r^{id}, \varepsilon) \leq k_A + k_B. \end{cases} \quad (7)$$

If aggregate capacity is insufficient to satisfy demand, the auctioneer cannot find a wholesale price which equates supply and demand, and a blackout occurs. The price bids p_A and p_B also determine the wholesale volume that generator i can sell, which is given by

$$y_i(p_A, p_B) = \begin{cases} \min\{k_i, x(r^{id}, \varepsilon)\} & \text{if } p_i < p_j, \\ \frac{\min\{k_i, x(r^{id}, \varepsilon)\}}{2} + \frac{\max\{0, x(r^{id}, \varepsilon) - k_j\}}{2} & \text{if } p_i = p_j, \\ \max\{0, x(r^{id}, \varepsilon) - k_j\} & \text{if } p_i > p_j, \end{cases} \quad (8)$$

with $i, j = A, B$ and $i \neq j$.

Using (7) and (8), integrated generator i 's revenues are given by

$$\pi_i(r_i, r_j) = r_i d_i(r_i, r_j, \varepsilon) + p(p_i, p_j) [y_i(p_i, p_j) - d_i(r_i, r_j, \varepsilon)]. \quad (9)$$

Equation (9) shows that an integrated generator earns the retail price r_i times retail demand d_i , plus the wholesale price p times the difference between the wholesale volume y_i and retail demand d_i . As a consequence, an integrated duopoly generator faces a *rent extraction risk*: if its retail demand exceeds own wholesale capacity, the integrated generator becomes a net payer in the wholesale market. A competitor who can fill the gap will bid the maximum

price for selling its capacity, such that the net payer will forgo its rent. As we will show below, this risk induces integrated duopoly generators to charge relatively high retail prices and make large capacity investments. The wholesale price is zero if both firms are able to cover their retail demand, whereas the market breaks down if aggregate capacity cannot cover total demand.¹⁷

3.3.2 Retail Market

In the retail market, an integrated duopolist has three pricing options: First, it can undercut its competitor and corner the market. This strategy generates revenues if the demand shock is such that retail demand is positive and the undercutting generator's capacity is sufficiently large to serve it. Second, it can match the price of its competitor, splitting aggregate retail demand in half. Expected revenues then depend on the competitors' relative capacities: for the smaller firm, revenues are as in the undercutting case, except that it serves only half of the demand. For the firm with the larger capacity, however, revenues are different, as it can appropriate the rival's rent if its capacity is sufficiently large to make up for a lack of capacity of the smaller firm. Third, it can surcharge its competitor, in which case it will not attract retail subscribers. It will nevertheless make revenues if the competitor cannot serve aggregate retail demand and own capacity is sufficiently large to cover the gap at the wholesale level. More formally, equilibrium retail prices satisfy the following proposition.

Proposition 3 (retail prices) *Depending on capacity levels (k_i, k_j) , $i, j = A, B$, $i \neq j$, there are the following Nash equilibria in retail prices.*

- (i) *With symmetric capacities $k_i = k_j = k < \sqrt{5/2}$, the pareto-dominant Nash equilibrium results in symmetric retail prices*

$$r^{id} = r_i = r_j = \begin{cases} 2 - \sqrt{2k} & \text{if } 0 \leq k < 1/\sqrt{2}, \\ \frac{1}{2}(3 - \sqrt{4k^2 - 1}) & \text{if } \frac{1}{\sqrt{2}} \leq k < 1/\sqrt{5/2}. \end{cases} \quad (10)$$

- (ii) *With capacities not too asymmetric, the unique Nash equilibrium results in symmetric retail prices $r^{id} = r_i = r_j = 0$.*

- (iii) *With asymmetric capacities and $k_j < \max\{(k_i - 1)/2, \frac{1}{8}\}$, the pareto-dominant Nash equilibrium results in $r^{id} = r_i = \max\{\frac{3}{4}, 2 - k_i\} < r_j$.*

¹⁷The wholesale prices are explicitly derived in lemma 1 of Boom (2009).

(iv) If $\frac{1}{8} \leq k_j < (k_i - 1)/2$, the pareto-dominant Nash equilibrium results in $r^{id} = r_i = r_j = 1 - 2k_j$.

(v) If $k_i \neq k_j$ and $k_i + k_j < 1$, the equilibria cannot be pareto ranked, but they are payoff-equivalent as both firms realize zero revenues.

Proof: See Boom (2007), Appendix C. ■

With symmetric and limited capacities (case (i)), there is a common retail price for which both firms are indifferent between undercutting, matching, and surcharging their competitor. This forms a pareto-dominant Nash equilibrium. With slightly asymmetric capacities (case (ii)), both firms want to undercut (surcharge) the rival if he sets a high (low) price. Only the low capacity firm wants to match in an intermediate retail price range where the high capacity firm directly switches from surcharging to undercutting. Therefore, the only remaining Nash equilibrium is the Bertrand outcome. For very asymmetric capacities (case (iii)), the large-capacity firm can charge the monopoly price on the expected demand without inducing the rival to match its price, because the small-capacity firm will not be able to serve even half of the demand. With slightly less asymmetric capacities, the larger firm sets a price that induces the smaller firm to match it, even though it cannot satisfy its demand (case (iv)). Finally if capacities are asymmetric and relatively small, firms can also generate a Bertrand outcome by pricing low enough to generate blackouts all the time (case (v)).

3.3.3 Capacities

Using the retail prices from proposition 3, the following capacity choices emerge. If the competitor's capacity is very low, an integrated duopolist can either choose a very large capacity and corner the market, or it can match the competitor's capacity to generate positive revenues (for any smaller own capacity, revenues are zero). For a slightly larger competitor capacity, cornering the market is no longer an option. Yet, positive revenues are still possible if own capacity is chosen much larger than the competitor's capacity. For a still larger capacity of the competitor, positive revenues from installing a higher capacity are no longer possible. Finally, for a very large competitor capacity, own revenues are independent of own capacity. Solving for the non-pareto-dominated Nash equilibria yields the following result.

Proposition 4 (capacity investments) *Depending on capacity costs z , there are the following pareto-dominant Nash equilibria in capacities.*

(i) For low capacity costs ($0 \leq z < 0.2118$), there is a unique equilibrium where firms choose capacity levels $k_A = k_B = \widehat{k}$, with

$$\widehat{k} = \operatorname{argmax}_k \left\{ \frac{1 - 4k^2 + 3\sqrt{4k^2 - 1}}{8} - zk \right\}.$$

(ii) For intermediate capacity costs ($0.2118 \leq z \leq 1/(2\sqrt{2})$), the non-pareto dominated equilibria are characterized either by both firms choosing \widehat{k} , or by one firm choosing the monopoly capacity k^m and the other firm choosing 0.

(iii) For high capacity costs ($1/(2\sqrt{2}) < z < \frac{1}{2}$), there are two equilibria with one firm choosing k^m and the other firm 0.

Proof: See Boom (2007), Appendix C. ■

Proposition 4 indicates that a unique pareto-dominant equilibrium exists only for low capacity costs (case (i)). For intermediate capacity costs (case (ii)), there are three non-dominated equilibria, one with symmetric capacities and two in which only one firm invests (choosing monopoly capacity). For high capacity costs (case (iii)) only the latter two exist.

3.4 Separated Duopoly

In this market configuration, two separated generators sell to the wholesale market, while two separated retailers buy from the wholesale market and serve retail demand. We employ backward induction to characterize the subgame perfect equilibrium of the game.

3.4.1 Wholesale Market

The wholesale price and the generators' wholesale volumes are determined using the same uniform-price auction as in the integrated duopoly configuration. That is, if aggregate capacity is sufficient to satisfy demand, $k_A + k_B \geq x(r^{sd}, \varepsilon)$, the wholesale price is given by equation (7), with r^{id} replaced by r^{sd} . Otherwise a blackout occurs. Similarly, wholesale volumes are given by equation (8), with r^{id} replaced by r^{sd} .

Using the adapted versions of (7) and (8), the profit of a separated duopoly generator $i = A, B$ is given by

$$\pi_i(p_A, p_B) = p(p_A, p_B)y_i(p_A, p_B).$$

Note that a separated duopoly generator does not face the risk of rent extraction, since it is not committed to serve any specific level of retail demand. Best-response bidding now requires each generator to either undercut its competitor, or to bid the maximum price $p_i = r^{sd}$ at which retailers break even. The next proposition characterizes the resulting Nash equilibria in price bids.

Proposition 5 (wholesale prices) *Depending on capacity levels (k_A, k_B) and the retail price r^{sd} , there are the following types of Nash equilibria in price bids:*

- (i) *If $k_A + k_B < x(r^{sd}, \varepsilon)$, any pair (p_A, p_B) forms a Nash equilibrium in price bids. No wholesale price can equate supply and demand, and a blackout occurs.*
- (ii) *If $k_i \geq x(r^{sd}, \varepsilon) > k_j$, with $i, j = A, B$ and $i \neq j$, the Nash equilibrium in pure strategies is characterised by $p_i = r^{sd}$ and $p_j < r^{sd}(x(r^{sd}, \varepsilon) - k_j)/k_i$. The resulting equilibrium wholesale price is $p^{sd} = r^{sd}$, and firms sell the quantities $y_i = x(r^{sd}, \varepsilon) - k_j$ and $y_j = k_j$.*
- (iii) *If $k_A + k_B \geq x(r^{sd}, \varepsilon) > \max\{k_A, k_B\}$, there are two types of Nash equilibria in pure strategies: one with $p_A = r^{sd}$ and $p_B < r^{sd}(x(r^{sd}, \varepsilon) - k_B)/k_A$, and another with $p_B = r^{sd}$ and $p_A < r^{sd}(x(r^{sd}, \varepsilon) - k_A)/k_B$. The wholesale price is the same ($p^{sd} = r^{sd}$) for both types of equilibria, but the quantities sold in equilibrium differ: in the former $y_A = x(r^{sd}, \varepsilon) - k_B$ and $y_B = k_B$, whereas in the latter $y_A = k_A$ and $y_B = x(r^{sd}, \varepsilon) - k_A$.*
- (iv) *If $\min\{k_A, k_B\} \geq x(r^{sd}, \varepsilon)$ the Nash equilibrium $p_A = p_B = 0$ is unique. The resulting equilibrium wholesale price is $p^{sd} = 0$, and firms sell the quantities $y_A = y_B = x(r^{sd}, \varepsilon)/2$.*

Proof: Follows from appendix A of Le Coq (2002) or the proofs of proposition 1-3 in Crampes and Creti (2005), using that marginal generating costs are constant and normalized to zero by assumption and that the maximum wholesale price with positive demand is $p = r^{sd}$. ■

Proposition 1 is illustrated in Figure 1. Area \mathcal{A} corresponds to case (i), where demand exceeds aggregate capacity, so that a blackout occurs. Areas \mathcal{B} and \mathcal{D} are associated with case (ii): In area \mathcal{B} , firm A (B , respectively) is the large (small) firm. In area \mathcal{D} , these roles are reversed. In both cases, the large firm bids the maximum price r^{sd} , whereas the small firm bids just low

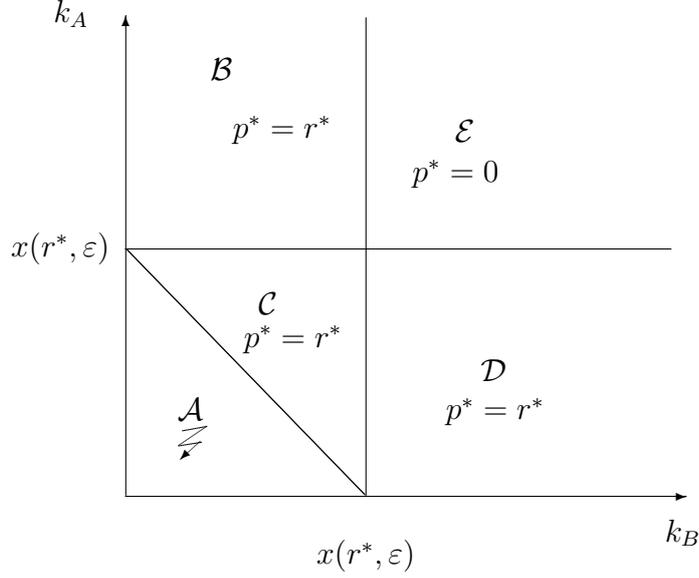


Figure 1: Prices on the Wholesale Market

enough to avoid undercutting by the large firm. In area \mathcal{C} , which corresponds to case (iii), the difference in installed capacities is smaller than in either \mathcal{B} or \mathcal{D} , and two types of equilibria are possible: Either the large or the small firm bids the maximum price, and the other firm bids low enough to avoid undercutting. In both cases the equilibrium wholesale price is $p^{sd} = r^{sd}$. Finally, area \mathcal{E} corresponds to case (iv). Here, each firm's capacity is sufficient to satisfy aggregate demand. Therefore, price bidding yields a Bertrand-type equilibrium.

Note that there are multiple pure-strategy Nash equilibria for cases (i)–(iii), as any lower bid that avoids undercutting and negative profits is admissible. These equilibria are pay-off equivalent for cases (i) and (ii), but not for case (iii), where the volume of dispatched electricity $y_i(p_A, p_B)$ depends on the type of equilibrium played. To deal with this multiplicity problem, we impose the following assumption:¹⁸

Assumption 1 *If capacities satisfy $k_A + k_B \geq x(r^{sd}, \varepsilon) > \max\{k_A, k_B\}$, generators coordinate on the Nash equilibrium where the large-capacity firm bids the maximum price and the small-capacity firm bids low enough to avoid undercutting by the large firm. If generators have equal capacities, they play each type of equilibrium with equal probability.*

¹⁸The assumption is equivalent to applying risk-dominance as a selection criterion. See Boom (2008) for a detailed discussion.

3.4.2 Retail Market

We first note that for retailers to obtain non-negative profits, the demand shock must satisfy $\varepsilon \in (\underline{\varepsilon}, \bar{\varepsilon})$, where $\underline{\varepsilon} \equiv r - 1$ is the critical value below which demand is zero, and $\bar{\varepsilon} \equiv \min\{k_A, k_B\} + r - 1$ is the maximum value for which generating capacities are large enough to avoid that generators extract rents from retailers.

With this in mind, and recalling that retailers compete à la Bertrand, the expected profits of retailer $\ell = C, D$ are given by

$$\pi_\ell(r_\ell, r_t) = \begin{cases} 0 & \text{if } r_\ell > r_t, \\ \frac{1}{2} \int_{\max\{0, \underline{\varepsilon}\}}^{\max\{0, \min\{\bar{\varepsilon}, 1\}\}} r_\ell(1 + \varepsilon - r_\ell) d\varepsilon & \text{if } r_\ell = r_t, \\ \int_{\max\{0, \underline{\varepsilon}\}}^{\max\{0, \min\{\bar{\varepsilon}, 1\}\}} r_\ell(1 + \varepsilon - r_\ell) d\varepsilon & \text{if } r_\ell < r_t, \end{cases} \quad (11)$$

with $\ell, t = C, D$, and $\ell \neq t$. Equation (11) indicates that retailers undercut each other until they reach zero profits. Therefore, the following Nash equilibrium in retail prices emerges.

Proposition 6 (retail prices) *Depending on the capacity levels (k_A, k_B) , there are the following Nash equilibria in retail prices.*

- (i) *If $\min\{k_A, k_B\} \geq 1$ there is a unique pure-strategy Nash equilibrium with $r_C = r_D = 0$.*
- (ii) *If $\min\{k_A, k_B\} < 1$ all pure-strategy Nash equilibria are characterised by $r_C \leq 1 - \min\{k_A, k_B\}$ and $r_D \leq 1 - \min\{k_A, k_B\}$.*

Proof: Suppose that $r_\ell > r_t$ with $\ell, t = C, D$ and $\ell \neq t$. This can only be an equilibrium if $r_t \leq 1 - \min\{k_A, k_B\}$ and $r_\ell \leq 1 - \min\{k_A, k_B\}$, because otherwise firm ℓ could increase its profits by undercutting and firm t by increasing its price. Suppose, alternatively, that $r_\ell = r_t$. Then either $r_\ell = r_t = 0$ if $\min\{k_A, k_B\} \geq 1$, or $r_\ell = r_t < 1 - \min\{k_A, k_B\}$ if $\min\{k_A, k_B\} < 1$, because otherwise each retailer could double its profit by undercutting. ■

Proposition (6) shows that, due to Bertrand competition, retailers cannot realize strictly positive profits, no matter whether the equilibrium is unique (case (i)) or not (case (ii)). In case (i) the generators' minimum capacity is not small enough to ensure that the rents are shifted from retailers to the generators for all demand realizations at a positive retail price, meaning

that the wholesale market equilibrium is for some small ε and all positive retail prices always located in area \mathcal{E} of figure 1. Therefore the retailers compete each other down to $r = 0$ in a Bertrand type manner. In case (ii) the minimum capacity of the generators is small enough that all rents are shifted from the retailers to the generators at a positive retail price even if the demand realization ε is close to zero. Thus, for the retail price range given in (ii) the wholesale equilibrium is for all ε located in either \mathcal{A} , \mathcal{B} , \mathcal{C} or \mathcal{D} meaning that the retailers realize zero profits before they compete the retail prices down to zero. To deal with the multiplicity problem in case (ii), we introduce the following assumption regarding equilibrium selection.

Assumption 2 *If $\min\{k_A, k_B\} < 1$, retailers choose the Nash equilibrium with $r_C = r_D = 1 - \min\{k_A, k_B\}$.*

Assumption 2 imposes that retailers select the equilibrium in which they choose the highest possible price which generates zero profits.

3.4.3 Capacity Investments

Separated duopoly generators $i = A, B$ must anticipate the impact of their capacity choices k_i on the retail market and the wholesale market. Due to Bertrand competition at the retail level, potential rents are shifted to generators. Provided that aggregate capacity exceeds retail demand, the wholesale price is given by

$$p^{sd} = r^{sd} = \max\{0, 1 - \min\{k_A, k_B\}\}. \quad (12)$$

With a strictly positive retail price r^{sd} , demand is characterized by $x(r^{sd}, \varepsilon) = 1 + \varepsilon - 1 + \min\{k_A, k_B\} = \varepsilon + \min\{k_A, k_B\}$. Therefore, generator i 's expected profits are given by

$$\Pi_i(k_i, k_j) = \begin{cases} \max\{0, 1 - k_j\} \int_0^{\min\{1, k_i\}} \varepsilon d\varepsilon - zk_i & \text{if } k_i > k_j, \\ \frac{\max\{0, 1 - k_j\}}{2} \left[\int_0^{\min\{1, k_i\}} \varepsilon d\varepsilon + \int_0^{\min\{1, k_j\}} k_i d\varepsilon \right] - zk_i & \text{if } k_i = k_j, \\ \max\{0, 1 - k_i\} \int_0^{\min\{1, k_j\}} k_i d\varepsilon - zk_i & \text{if } k_i < k_j, \end{cases} \quad (13)$$

with $i, j = A, B$ and $i \neq j$.¹⁹ To understand (13), suppose $\min\{k_A, k_B\} < 1$ and $x(r^{sd}, \varepsilon) \leq k_A + k_B$, and first consider the case where $k_i > k_j$. Firm i then bids high and serves residual demand $\max\{x(r^*, \varepsilon) - k_j, 0\} = \max\{1 + \varepsilon - 1 + k_j - k_j, 0\} = \varepsilon$. Next, consider the case where $k_i < k_j$: Firm i now bids low and delivers its total capacity up to the level of demand (i.e., $\min\{k_i, 1 + \varepsilon - 1 + k_i\} = k_i$). Finally, if capacities are identical ($k_i = k_j$), firm i bids high or low with probability one half each. As noted above, the condition $x(r^{sd}, \varepsilon) \leq k_A + k_B$ must hold, since generators cannot sell electricity in the event of a blackout. This condition is equivalent to $\varepsilon + \min\{k_A, k_B\} \leq k_A + k_B$ or $\varepsilon \leq \max\{k_A, k_B\}$, if $\min\{k_A, k_B\} < 1$, which explains the upper bound for integration in (13).

As we show in Appendix A, generator i 's best response is to choose a higher capacity than its competitor ($k_i = 1 > k_j$) if the competitor's capacity is relatively low, and to choose a lower capacity,

$$k_i = \max\{0, \min\{(k_j - z)/(2k_j), (1 - z)/2\}\},$$

if the rival's capacity is relatively high.²⁰ This is quite intuitive, since both residual demand and the wholesale price are large if the competitor's capacity is small. It therefore pays to install a large capacity. In contrast, if the competitor's capacity is large, it is more profitable to install a small capacity which is completely sold and supports a higher wholesale price. The next proposition summarizes the results for the case with simultaneous capacity choices.

Proposition 7 (simultaneous capacity choices) *With simultaneous capacity choices, the existence of a subgame-perfect Nash equilibrium (SPNE) in pure strategies is not guaranteed.*

- (i) *If $0 \leq z < 1/3$, there are two asymmetric SPNE in pure strategies with capacities $k_i^{sd} = 1$ and $k_j^{sd} = (1 - z)/2$, $i, j = A, B$ and $i \neq j$.*
- (ii) *If $1/3 \leq z < 1/2$, there is no SPNE in pure strategies.*

Proof: Solving the system of best-response functions (25) and (26) in Appendix A for equilibrium capacities yields results (i) and (ii). ■

¹⁹With $\min\{k_A, k_B\} \geq 1$, the wholesale price is zero and none of the generators will realize positive profits.

²⁰See (25) and (26) in Appendix A for a detailed description of firm i 's best response function.

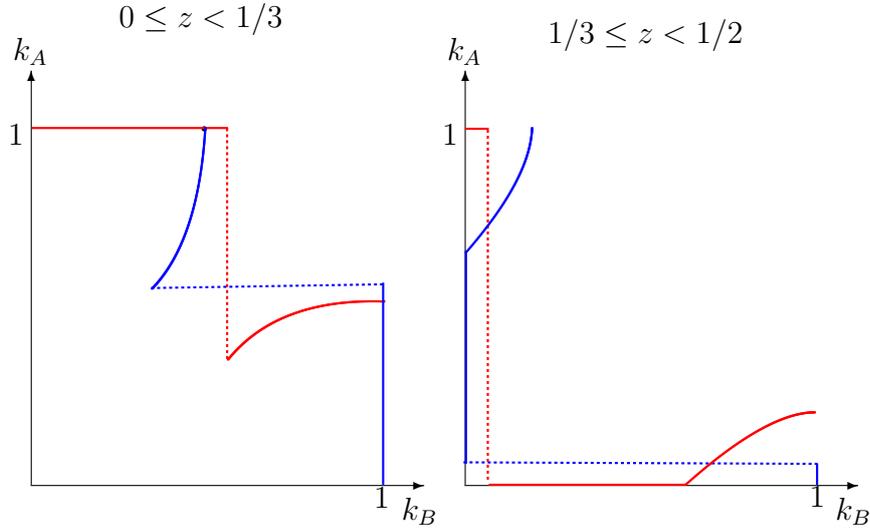


Figure 2: Best Responses in Capacities

Figure 2 illustrates that there exists no pure-strategy Nash equilibrium with simultaneous capacity choices and high capacity costs z . The next proposition shows that this non-existence problem disappears if separated generators choose capacities sequentially. For the sake of concreteness, it assumes that firm A moves first.

Proposition 8 (sequential capacity choices) *Suppose that firm A moves first.*

- (i) *If $0 \leq z < 1/3$, there is a unique SPNE in pure strategies where firm A chooses $k_A^{sd} = (1 - z)/2$ and firm B chooses $k_B^{sd} = 1$.*
- (ii) *If $1/3 \leq z < 1/2$, there is a unique subgame perfect Nash equilibrium in pure strategies where firm A chooses $k_A^{sd} = 1 - 2z$ and firm B chooses $k_B^{sd} = 1$.*

Proof: Substituting firm B 's best response function $k_B(k_A)$ into $\Pi_A(k_A, k_B)$ in Appendix A and maximizing with respect to k_A yields results (i) and (ii). ■

Note that the first mover, generator A , prefers to be the small-capacity firm: The small-capacity firm bids low and sells its total capacity, whereas the large-capacity firm, generator B , bids high, thereby determining the wholesale price, and serves only residual demand.

4 Ranking Market Configurations

We rank the various market configuration with respect to aggregate capacity, retail prices, and levels of social welfare. In doing so, we use the following notation. The first-best optimal level of aggregative capacity is given by k^s . Under integrated monopoly, capacity is denoted as k^m . Under integrated duopoly, aggregate capacity is $k^{id} = 2\hat{k}$, whereas under separated duopoly it is $k^{sd} = k_A^{sd} + k_B^{sd}$. We use similar notation to distinguish retail prices and levels of social welfare in the alternative market configurations.

Proposition 9 (ranking) *Suppose that capacity decisions are either taken sequentially by the separated generators or that $0 \leq z \leq 1/3$, and that integrated generators co-ordinate on the pareto-dominant competitive equilibrium. Then the ranking of market configurations in terms of*

(i) *aggregate capacity levels is*

$$k^s \geq k^{id} \geq k^{sd} \geq k^m; \quad (14)$$

(ii) *of retail prices is given by*

$$r^{id} \geq r^m \geq r^{sd} \geq r^s; \quad (15)$$

(iii) *and of welfare levels is given by*

$$W^s \geq W^{sd} \geq W^m \geq W^{id}. \quad (16)$$

Proof: (i) Follows from comparing Propositions 1, 2, 8 and 4. (ii) Follows from comparing Propositions 1, 2, 6 and 3. (iii) Since blackouts do not occur irrespective of market configuration, social welfare is given by

$$W(k) = \int_0^1 U(x(r, \varepsilon), \varepsilon) d\varepsilon - zk, \quad (17)$$

where k denotes total capacity. Substituting $U(x(r, \varepsilon), \varepsilon)$ from (1), $x(r, \varepsilon)$ from (2) and plugging in equilibrium values for r and k for each market configuration, yields the associated welfare levels. Comparing these welfare levels completes the proof. ■

Proposition 9 shows that, compared to the social optimum, capacity levels are inefficiently low and retail prices inefficiently high in all market configurations. We now want to discuss the intuition for the ranking of these configurations.

Let us first consider aggregate capacity levels. Result (i) shows that capacity levels are highest under integrated duopoly and lowest under integrated monopoly. The separated duopoly yields an intermediate level of aggregate capacity. To understand this result, consider the investment incentive of an integrated monopoly generator. Adding another integrated generator introduces competition both at the wholesale and the retail level. Since an integrated duopoly generator now faces the risk of rent extraction, it has an incentive to increase its investment relative to the integrated monopoly ($k^{id} > k^m$). Next, consider the impact of vertical separation on the investment incentives of duopoly generators. After vertical separation, generators trade with separated retailers (rather than themselves) on the wholesale market. Since they are no longer committed to serve any predetermined level of retail demand, generators do not face the risk of rent dissipation by their rival, and they thus install smaller capacities than integrated duopoly generators ($k^{id} > k^{sd}$). Result (i) indicates that vertically separating the duopoly eliminates the investment-enhancing effect of rent extraction but leaves a positive investment effect due to introducing competition ($k^{sd} > k^m$).

Result (ii) shows that retail prices are highest under integrated duopoly and lowest under separated duopoly (apart from the social optimum). The integrated monopoly yields an intermediate level of retail prices. The intuition for high retail prices under integrated duopoly parallels that for high aggregate capacity: Integrated duopoly generators face the risk of rent extraction and thus have an incentive to set a high retail price to keep demand low (in addition to making high investments to serve retail demand). This risk does not exist under integrated monopoly or separated duopoly. Also note that retail prices are lowest in the separated duopoly, where retail competition disciplines retail prices.

Finally, proposition 9 (iii) indicates that the combined effects of restructuring on capacity levels and prices are such that social welfare is highest under separated duopoly and lowest under integrated duopoly. The integrated monopoly yields an intermediate level of social welfare. To understand the intuition for the result, it is important to note that, irrespective of market configuration, aggregate installed capacity is always large enough to satisfy retail demand at the relevant equilibrium retail price, so that blackouts do not occur in equilibrium. This implies that raising capacity increases capacity costs rather than supply security. These increases in capacity costs must be weighed against the effects of changes in retail prices for the construction of the welfare ranking. Since both total capacity and retail prices are higher in the integrated duopoly than in the successive duopoly, welfare must be lower in the integrated duopoly. The welfare effect of changing from integrated

monopoly to separated duopoly is less obvious: Total capacity is higher, but retail prices are lower in the separated duopoly. Proposition 9 (iii) shows that the positive effect of lower retail prices dominates the negative effect of higher capacity costs, so that the separated duopoly performs better than the integrated monopoly.

5 Extensions and Limitations

So far, we have deliberately abstracted from a number of real-world issues to highlight the role of vertical structure in determining generating capacities and retail prices. In the following subsections, we consider various extensions and limitations of our analysis.²¹ First, we explore the case where blackouts may be avoided by the rationing of retail demand. Next, we consider alternative specifications of demand. Finally, we discuss alternative specifications of supply.

5.1 Rationing

In practice, system operators attempt to ration retail demand to avoid (or at least limit) blackouts. However, rationing is difficult to model as it is typically implemented in unsystematic ways and generally not in line with theoretical rationing models.²² By abstracting from the possibility of rationing, our main analysis has maximized the punishment of generators for providing insufficient capacity. We now consider the other extreme where demand can be rationed and generators are not punished for providing insufficient capacity.²³

Specifically, we assume that if demand exceeds aggregate supply on the wholesale market ($x(r, \varepsilon) > k_A + k_B$), blackouts may be avoided by the rationing of retail demand. Specifically, we suppose that rationing leads to a so-called “brownout”, that is, generators sell their total capacity to consumers.²⁴

²¹We are grateful to the referees for prompting us to study some of these extensions.

²²For instance, in case of excess demand due to an unplanned outage of facilities (e.g., a power station or a transmission line), consumers in the neighborhood are temporarily cut off from service to avoid a spread of the blackout.

²³In reality, unsystematic rationing is likely to provide at least some punishment, as valuable consumers might be cut off from service.

²⁴Note that in our model rationing is always efficient because we have a representative consumer model.

5.1.1 Social Optimum With Rationing

With rationing the social planner maximizes

$$W(r, k) = \begin{cases} \int_0^1 U(x(r, \varepsilon), \varepsilon) d\varepsilon - zk & \text{if } r \geq 2 - k, \\ \int_0^{k-1+r} U(x(r, \varepsilon), \varepsilon) d\varepsilon + \\ \int_{k-1+r}^1 U(k, \varepsilon) d\varepsilon - zk & \text{if } 2 - k > r \geq \max\{1 - k, 0\}, \\ \int_0^1 U(k, \varepsilon) d\varepsilon - zk & \text{if } 1 - k > r \geq 0 \end{cases} \quad (18)$$

with respect to r and k , respectively. The key difference to the welfare function given in (5) is that consumers now get served up to capacity if demand exceeds capacity. In this case consumers experience a brownout rather than a blackout.

Unsurprisingly, for any given capacity level, the optimal retail price with rationing is smaller than without rationing and effectively implies rationing for large demand shocks. The optimal capacity is also smaller than without rationing, inducing a retail price at the level of marginal cost ($r = 0$).²⁵

5.1.2 Integrated Monopoly With Rationing

The monopolist maximizes

$$\pi(r, k) = \begin{cases} \int_{\max\{r-1, 0\}}^1 r(1 + \varepsilon - r) d\varepsilon - zk, & \text{if } r \geq 2 - k, \\ \int_{\max\{r-1, 0\}}^{k-1+r} r(1 + \varepsilon - r) d\varepsilon \\ + \int_{\min\{k-1+r, 1\}}^1 rk d\varepsilon - zk, & \text{if } \max\{0, 1 - k\} \leq r < 2 - k, \\ \int_0^1 rk d\varepsilon - zk, & \text{if } r \leq \max\{0, 1 - k\} \end{cases} \quad (19)$$

with respect to r and k , respectively. Again, for any given capacity level, the profit maximizing retail price is smaller than without rationing. The chosen capacity level is also smaller, as insufficient capacity is no longer punished. In effect, the profit-maximizing retail price is smaller than without rationing, and rationing kicks in for large demand shocks.²⁶

²⁵See proposition 10 in Appendix B.1.

²⁶See proposition 11 in Appendix B.1 for further details.

5.1.3 Integrated Duopoly with Rationing

The integrated generator's retail demand is still given by equation (6). The wholesale price that results from the price bids p_A and p_B continues to be characterized by (7), with the change that the wholesale price is also determined in this way if demand exceeds aggregate capacity (i.e., $k_A+k_B < x(r, \varepsilon)$) with $r = \min\{r_i, r_j\}$. The capacity that is dispatched in the wholesale auction is now slightly different and given by

$$y_i^r(p_A, p_B) = \begin{cases} \min\{k_i, x(r, \varepsilon)\} & \text{if } p_i < p_j, \\ \frac{\min\{k_i, x(r, \varepsilon)\}}{2} + \frac{\min\{\max\{0, x(r, \varepsilon) - k_j\}, k_i\}}{2} & \text{if } p_i = p_j, \\ \min\{\max\{0, x(r, \varepsilon) - k_j\}, k_i\} & \text{if } p_i > p_j, \end{cases} \quad (20)$$

with $i, j = A, B$ and $i \neq j$. The profit function changes to

$$\begin{aligned} \pi_i(r_i, r_j) &= r_i d_i(r_i, r_j, \varepsilon) \cdot \min\left\{1, \frac{(k_i + k_j)}{x(r, \varepsilon)}\right\} + p(p_i, p_j) \left[y_i^r(p_i, p_j) \right. \\ &\quad \left. - d_i(r_i, r_j, \varepsilon) \cdot \min\left\{1, \frac{(k_i + k_j)}{x(r, \varepsilon)}\right\} \right]. \end{aligned} \quad (21)$$

It is obvious that the profits characterized in (21) are equivalent to the profits without rationing given in (9) if aggregate capacity is sufficient to serve retail demand ($k_A+k_B \geq x(r, \varepsilon)$). Moreover, there is still a risk of rent extraction in the wholesale market: Firms compete the wholesale price down to zero if each firm's individual capacity is sufficient to serve its (rationed) retail demand, as generators do not want to end up being a net payer in the wholesale market. If one of the firms nevertheless ends up being a net payer because of insufficient capacity, the other firm is bidding a wholesale price such that the net payer's profit is zero. The net payer, in turn, bids below the other firm's price to ensure that its payment is minimized.

Regarding retail prices, it is clear that an integrated generator cannot generate profits from undercutting if individual capacity is insufficient to cover the ensuing market demand.²⁷ Similarly, it is not profitable to match a larger rival's retail price. However, matching a smaller or equally sized competitor's retail price without being able to serve the ensuing retail demand does not necessarily reduce profit to zero. With identical capacities, each firm just sells its own capacity due to rationing, whereas with asymmetric capacities

²⁷See Appendix B.2 for a characterization of profits for undercutting, matching and overcharging the rival's retail price and the resulting best responses.

the larger firm acquires the market's total rent as soon as the smaller firm cannot serve its (rationed) retail demand. Analyzing the two firms' best responses in retail prices for all reasonable capacities $k_i, k_j \in [0, 2]$ reveals that, despite the changes in profits due to rationing, equilibrium retail prices remain essentially unaffected. Specifically, Proposition 3 (i)–(iv) remain valid, while (v) is no longer relevant as it describes blackout outcomes.

When deciding on their generating capacity, firms continue to compare profits when they match their rival's capacity (case (i) of proposition 3) with profits when they choose zero capacity or potentially (if the rival's capacity is small enough) with profits from much larger capacities than their rival (case (iii) and (iv) of proposition 3). If they go for any of the positive capacity options, the resulting retail price is always such that rationing does not kick in. Therefore, the firms' capacity choices are always the same as in the integrated duopoly case without rationing, and proposition 4 remains valid.

5.1.4 Separated Duopoly With Rationing

It is clear that, if rationing is not necessary, the wholesale price is still characterized by proposition 5, cases (ii) to (iv). If rationing is necessary ($x(r^{sd}, \varepsilon) > k_A + k_B$), then retailers demand electricity on behalf of their costumers on the wholesale market only if $p \leq r$. This leads to a Nash equilibrium with a wholesale price of $p = r^{sd}$ in case (i) of proposition 5, and each firm sells its full capacity. That is, in area \mathcal{A} of figure 1 the market no longer breaks down, but clears with the same wholesale price as in area \mathcal{B} , \mathcal{C} and \mathcal{D} . Note, though, that the firms sell different quantities than in \mathcal{B} , \mathcal{C} and \mathcal{D} , namely, given $k_i < k_j$, $y_i = k_i$ and $y_j = k_j$ instead of $y_i = k_i$ and $y_j = x(r^{sd}, \varepsilon) - k_i$. The slightly different wholesale prices have no impact on the retail price equilibrium described in proposition 6 or the resulting wholesale price given in (12). Nevertheless, generator i 's profit in terms of capacities changes, as blackouts are replaced by brownouts, and is now given by

$$\Pi_i(k_i, k_j) = \begin{cases} \max\{0, 1 - k_j\} \int_0^1 \min\{\varepsilon, k_i\} d\varepsilon - zk_i & \text{if } k_i > k_j, \\ \frac{\max\{0, 1 - k_j\}}{2} \left[\int_0^1 \min\{\varepsilon, k_i\} d\varepsilon + \int_0^1 k_i d\varepsilon \right] - zk_i & \text{if } k_i = k_j, \\ \max\{0, 1 - k_i\} \int_0^1 k_i d\varepsilon - zk_i & \text{if } k_i < k_j, \end{cases} \quad (22)$$

with $i, j = A, B$ and $i \neq j$ instead of (13). The separated generators' best response functions can again be derived from their profit maximization prob-

lems.²⁸ Analyzing these best response functions reveals that there exists no Nash equilibrium in pure strategies for simultaneous capacity choices. There is, however, a unique subgame perfect equilibrium for sequential capacity choices (with A choosing before B). In this setting, firm A chooses to be the small capacity producer and firm B invests in the larger capacity. Total capacities are again smaller than in the case without rationing, but the retail price is now higher. The higher retail price is provoked by the smaller minimal capacity in the market which determines the retail price.

5.1.5 Assessing the Role of Rationing

How does allowing for rationing affect our analysis? We have shown that rationing does not affect the equilibrium outcome under integrated duopoly. This is an important finding as it implies that the risk of rent extraction, which is the driving force in our main analysis, is robust to the introduction of rationing. In the other market configurations, the introduction of rationing tends to yield lower capacity investment, as capacity investments no longer need to be distorted upwards to avoid blackouts. With single decision makers (the social planner and the integrated monopoly) retail prices are also lower for the same reason. In the separated duopoly retail prices are, however, higher because the lower smaller capacity in the market shifts the rents from the retailers to the electricity generators already at a higher retail price. Moreover, with rationing brownouts may occur in equilibrium, whereas blackouts do not occur in equilibrium without rationing. Yet, the ranking of the market configurations provided in 9 remains largely unaffected.

To show this, we now consider aggregate capacity levels, retail prices, and welfare levels, in turn, across the different market configurations. To ensure existence of two active firms in both duopoly market configurations, we focus on capacity costs $0 \leq z \leq 1/(2\sqrt{2})$.

Aggregate Capacity Levels. For small capacity costs ($0 \leq z < 0.0267$), the ranking given in (14) remains unaffected. For larger capacity costs ($0.0276 \leq z \leq 1/(2\sqrt{2})$), there is one change in this ranking: Aggregate capacity in the integrated duopoly is now higher than the socially optimal level. Other than that, the ranking remains unaffected. Intuitively, the result follows from the fact that the risk of rent extraction induces integrated duopoly generators to choose capacities such that rationing does not occur in equilibrium. The logic of our above analysis is thus still correct.

²⁸See Appendix B.3 and proposition 12 for the detailed results.

Retail Prices. For the whole range of admissible capacity costs, the ranking given in (15) remains unaffected.

Welfare Levels. The key difference to the main analysis is that, with rationing, the integrated duopoly provides a higher supply security than the other market configuration, because the latter involve some rationing in equilibrium. Yet, this advantage of the integrated duopoly affects the ranking only for very low capacity costs ($0 \leq z < 0.05305$), in which case the integrated duopoly performs better than the monopoly but still worse than the separated duopoly. For higher capacity costs ($0.05305 \leq z < 1/(2\sqrt{2})$), the original welfare ranking given in (16) remains unaffected. That is, for most of the parameter space, introducing rationing does not affect the results of our main analysis.

5.2 Demand Specification

Consumer Surplus. In our main analysis, we have assumed that a positive demand shock ($\varepsilon > 0$) is associated with a negative effect on consumer surplus (see equation (1)). This is a natural assumption if the shock represents the extra demand for electricity due to extra cooling (heating, respectively) generated by, say, an extraordinarily hot (cold) day.²⁹ Let us now consider the case where a positive demand shock represents a boom that increases both demand and consumer surplus. In this case, one can construct the preferences such that demand is still represented by (2) and that the firms' decisions are not affected.³⁰ Of course, the social welfare benchmark is affected, but it can be shown that the ranking of market configurations given in Proposition (9) remains correct. Similarly, our focus on linear demand is not overly restrictive. Well-behaved, concave demand functions which are mainly shifted (rather than rotated) by demand shocks will give rise to similar results.

Distribution of Demand Shocks. The uniform distribution assumption for demand shocks is an important driving force of our results. Specifically, it imposes that extreme realizations are equally likely to occur as intermediate realizations. In such a setting, it is not surprising that firms' decisions are made such that, irrespective of market configuration, blackouts do not occur in equilibrium. For distributions which make extreme realizations less likely to occur (such as the normal), it appears more reasonable that firms allow for

²⁹Despite the lower benefit provided by energy, consumers demand more energy to adjust the room temperature accordingly.

³⁰Substitute for example $U(x, \varepsilon)$ in (1) by $\tilde{U}(x, \varepsilon) = x + \varepsilon - (x - \varepsilon)^2/2$ in (1).

blackouts for very large demand shocks that arise with very small probability. In such a setting, the incentive to make high capacity investments and charge high retail prices in the integrated duopoly may actually become a virtue as it leads to a higher supply security. That is, installing higher generating capacities may be valuable per se. Future research will have to establish how robust our results are to changes in the distribution of demand shocks.

Flexible Retail Prices. For simplicity, our analytical framework imposes that all consumers are faced with temporarily fixed retail prices. We are well aware that this must not necessarily be true in practice. For instance, large-demand consumers might have concluded contracts with electricity providers that offer options for real-time pricing, participate directly in the wholesale market, or own their own power stations. In the future, real-time pricing may even become viable for residential consumers, once smart metering is implemented on a large scale. Yet, our framework adequately reflects the fact that the vast majority of consumers is still faced with fixed retail prices (at least for the foreseeable future). This implies that integrated duopoly generators cannot avoid the risk of rent extraction by simply increasing real-time retail prices whenever necessary. More frequent retail price changes might moderate the distorted capacity and retail price outcomes of the integrated duopoly but do not solve the problem. Note that offering so-called *spot price products* is unlikely to change this: as long as these products charge consumers the average wholesale price on the basis of a fictitious monthly demand and not the real-time wholesale price, retail demand is effectively still fixed in the short run and cannot be easily reduced in the event of potential exploitation by a competitor in the wholesale market.³¹

5.3 Supply Specification

Electricity Generation Technologies. The above analysis focuses on the firms' overall incentives to invest in generating capacities and abstracts from alternative generation technologies (such as peak-load and base-load technologies with different levels of marginal and fixed costs). The paper therefore does not make any statements on the extent to which the generation portfolio varies (or is distorted) across the different market configurations. Having said this, one might conjecture that the structure of the generation portfolio interacts with the vertical structure via the amount of the firms' investments in capacities. For instance, if large (small) generators tend to

³¹See von der Fehr and Hansen (2010) for an empirical analysis of spot price products in Norway.

invest more in base-load (peak-load, respectively) technology, separated (integrated) duopoly generator will tend to invest more in peak-load (base-load) technology. Studying this issue more thoroughly is beyond the scope of this paper.

Payments for Installed Capacity. Our main analysis focuses on energy markets and abstracts from the possibility of payments for installed capacity. The reason is that, even though payments or artificial markets for installed capacities are important in the US and highly debated in the EU (Cramton and Ockenfels (2012)), their effect depend a lot on their particular design that varies a lot even in the US. It should be clear that introducing payments for installed capacities increases the generators' overall incentives to invest in capacities. It is less clear, though, how these extra incentives interact with vertical structure. Generally speaking, investments in generating capacity should be expected to increase in all market configurations. However, the relative size of the effect will depend on how the market for installed capacity is organized. For instance, if the payments for installed capacity decrease with the level of available aggregate capacity, the extra investment incentive will be smallest under integrated duopoly and highest in the integrated monopoly. The driving forces will then again be those that we described above.

6 Conclusion

Analyzing a stylized model of the electricity market, this paper has studied the role of vertical structure in determining generating capacities and retail prices. Our analysis identifies the risk of rent extraction associated with vertical integration as a key determinant of equilibrium capacity levels, retail prices, and welfare levels: An integrated generator whose capacity is too small to serve own retail demand must buy electricity in the wholesale market, thereby fully dissipating its rent. There is no such risk under vertical separation, as generators are not committed to serve an uncertain level of demand at a pre-determined retail price.

Our analysis supports the view that restructuring the electricity industry may lead to welfare gains. In addition, it highlights that introducing competition into the electricity industry (alone) does not necessarily reduce retail prices. The risk of rent extraction associated with vertical integration may prevent firms from reducing their retail prices and induce them to make large costly capacity investments. These insights are robust to the introduction of demand rationing at the retail level.

There is ample scope for future research. First, it would be interesting to allow for endogenous (and possibly asymmetric) vertical integration decisions, as suggested by Buehler and Schmutzler (2005) and Buehler and Schmutzler (2008). Doing so would further enrich our understanding of the firms' strategic investment decisions. Second, the discrimination of non-integrated competitors has rarely been considered in the context of electricity. Third, it would be useful to study models with different mechanisms determining wholesale prices and with more than two competitors to better understand the robustness of the risk of rent dissipation.

Acknowledgements

We are grateful to the editor, Michael Crew, and two anonymous referees for many helpful comments and suggestions. We further thank Gregor Zöttl, Nicholas Shunda, Chloé Le Coq, as well as seminar audiences at Copenhagen University, the University of Groningen, the University of Utrecht, the DIW Berlin, and numerous conferences and workshops for useful discussions. We gratefully acknowledge financial support from the Swiss National Science Foundation through grants PP0011-114754 and PP00P1-135143 and from the Social Science Section of the Danish Council of Independent Research through the *Risky Power* grant.

Appendix

A Firm i 's Best Response in Capacity in a Separated Duopoly.

For $k_j \geq 1$ firm i 's profit function (13) translates into

$$\Pi_i(k_i, k_j) = \begin{cases} -zk_i & \text{if } k_i \geq 1, \\ (1 - k_i)k_i - zk_i & \text{if } 0 \leq k_i \leq 1. \end{cases} \quad (23)$$

If $0 \leq k_j < 1$ holds, firm i 's profit function becomes

$$\Pi_i(k_i, k_j) = \begin{cases} \frac{1-k_j}{2} - zk_i & \text{if } k_i \geq 1, \\ \frac{(1-k_j)k_i^2}{2} - zk_i & \text{if } k_j < k_i \leq 1, \\ \frac{1}{2} \left[\frac{(1-k_j)k_i^2}{2} + (1-k_i)k_i k_j \right] - zk_i & \text{if } k_i = k_j, \\ (1-k_i)k_i k_j - zk_i & \text{if } 0 \leq k_i < k_j. \end{cases} \quad (24)$$

The best response of firm i which is derived from maximizing (23) or (24), respectively, with respect to k_i yields

$$k_i(k_j) = \begin{cases} \frac{1-z}{2} & \text{if } k_j \geq 1, \\ \frac{k_j-z}{2k_j} & \text{if } \frac{1-z-\sqrt{1-2z-2z^2}}{3} \leq k_j \leq 1, \\ 1 & \text{if } 0 \leq k_j \leq \frac{1-z-\sqrt{1-2z-2z^2}}{3}, \end{cases} \quad (25)$$

for $0 \leq z \leq 1/3$. If $1/3 < z \leq 1/2$ holds, the maximization of (23) and (24) with respect to k_i results in

$$k_i(k_j) = \begin{cases} \frac{1-z}{2} & \text{if } k_j \geq 1, \\ \frac{k_j-z}{2k_j} & \text{if } z \leq k_j \leq 1, \\ 0 & \text{if } 1-2z \leq k_j \leq z, \\ 1 & \text{if } 0 \leq k_j \leq 1-2z. \end{cases} \quad (26)$$

B Rationing

B.1 The Social Planner's and the Monopolist's Capacity Choice with Rationing

Proposition 10 (social optimum with rationing) *The welfare maximizing retail price and capacity are given by $r^{sr} = 0$ and $k^{sr} = 2 - \sqrt{2z}$.*

Proof: Taking the first derivative of (18) with respect to r setting it equal to zero and solving for r yields $r = \max\{0, 1 - k\}$ as the optimal retail price for a given capacity for the social planner. Substituting this into (18) and solving the first order condition with respect to k yields the optimal capacity given in the proposition if $0 \leq z \leq 1/2$. ■

Proposition 11 (integrated monopoly with rationing) *The profit maximizing capacity is given by*

$$k^{mr} = \frac{21 - 2z}{12} - \frac{(1 - i\sqrt{3})(153 + 252z + 4z^2)}{24g(z)} + \frac{(1 + i\sqrt{3})g(z)}{24}$$

with i being the imaginary number³² and

$$g(z) \equiv (1269 + 5778z + 1188z^2 - 8z^3 + 12\sqrt{3}\sqrt{-4563 - 7020z + 16137z^2 - 7454z^3 + 1272z^4 - 72z^5})^{\frac{1}{3}}.$$

The retail price is given by

$$r^{mr} = \frac{1}{3}[2(1 - k^{mr}) + \sqrt{1 + 4k^{mr} + (k^{mr})^2}].$$

Proof: Taking the first derivative of (19) with respect to r , setting it equal to zero and solving for r yields $r = 1/3[2(1 - k) + \sqrt{1 + 4k + k^2}]$ as the profit maximizing retail price for a given capacity. Substituting this into (19) and solving the first order condition of profit maximization with respect to k yields the profit maximizing capacity given in the proposition. ■

B.2 The Integrated Duopoly with Rationing.

When the firms choose their retail prices their profits depend on whether they undercut, match or overcharge their rival's retail price. Undercutting yields the following expected profit net of capacity costs.

$$\underline{\pi}_i(r_i, r_j) = \begin{cases} \int_{\max\{0, r_i - 1\}}^{\min\{1, k_i + r_i - 1\}} r_i x(r_i, \varepsilon) d\varepsilon & \text{if } 1 - k_i \leq r_i < r_j, \\ 0 & \text{if } 0 \leq r_i < \min\{1 - k_i, r_i\}, \end{cases} \quad (27)$$

that does not differ from the expected profit without rationing. Rationing does not play a role because, if firm i undercuts firm j on the retail market, firm j does not have any retail demand, and firm i supplies all consumers. If firm i can however not serve its retail demand, firm j can exploit this on the wholesale market and shift all the rents to firm j . This happens already for smaller demand shocks than those for which the system operator would start to ration demand.

³²Note that for the relevant range of z the profit maximizing capacity $k^{mr} \in [5/4, 0]$ and is a rational number that monotonously decreases in z .

If firm i matches its rival's retail price, then its expected profit net of capacity costs depends on whether $k_i < k_j$, $k_i = k_j$ or $k_i > k_j$ holds. With $k_i < k_j$, the structure of its expected profit does not differ from the situation without rationing and is given by

$$\pi_i |_{r_i=r_j} = \begin{cases} \int_{\max\{0, r_j-1\}}^{\min\{1, 2k_i+r_j-1\}} \frac{r_j x(r_j, \varepsilon)}{2} d\varepsilon & \text{if } r_j \geq 1 - 2k_i, \\ 0 & \text{if } 0 \leq r_j < 1 - 2k_i. \end{cases} \quad (28)$$

Due to the same logic than before nothing changes due to rationing.

However, with $k_i = k_j$, different from the no rationing case, the firm can now continue to sell up to its capacity even if it cannot satisfy its own retail demand and its expected profit net of capacity costs changes to

$$\pi_i |_{r_i=r_j} = \int_{\max\{0, r_j-1\}}^1 r_j \max \left\{ \frac{x(r_j, \varepsilon)}{2}, k_i \right\} d\varepsilon. \quad (29)$$

This happens because the rival cannot exploit the firm on the wholesale market when the firm's retail demand exceeds its capacity because this happens for the other firm at exactly the same demand shock.

If $k_i > k_j$ then firm i , when matching firm j 's retail price, can appropriate all its rival's rents in the wholesale market, if firm j is not able to serve its (rationed) retail demand because firm i is always able to serve its (rationed) retail demand. Its expected profit net of capacity costs is then

$$\pi_i |_{r_i=r_j} = \begin{cases} \int_{\max\{0, r_j-1\}}^1 \frac{r_j x(r_j, \varepsilon)}{2} d\varepsilon & \text{if } r_j \geq 2 - 2k_j, \\ \int_{\max\{0, r_j-1\}}^{2k_j+r_j-1} \frac{r_j x(r_j, \varepsilon)}{2} d\varepsilon + \int_{2k_j+r_j-1}^1 r_j \min\{x(r_j, \varepsilon), k_i + k_j\} d\varepsilon & \text{if } 1 - 2k_j < r_j < 2 - 2k_j, \\ \int_0^1 r_j \min\{x(r_j, \varepsilon), k_i + k_j\} d\varepsilon & \text{if } 0 < r_j < 1 - 2k_j. \end{cases} \quad (30)$$

If firm i charges a higher retail price than its rival it can only earn positive revenues exactly in the case where its rival cannot serve the (rationed) retail demand of its retail customers at its price r_j resulting in firm i appropriating all the rents via the wholesale auction. Thus, the expected profit net of capacity costs is:

$$\bar{\pi}_i(r_i, r_j) = \begin{cases} 0 & \text{if } r_i > r_j \geq 2 - k_j, \\ \int_{\max\{0, r_j-1+k_j\}}^1 r_j \min\{x(r_j, \varepsilon), k_i + k_j\} d\varepsilon & \text{if } 0 \leq r_j < \min\{r_i, 2 - k_j\}. \end{cases} \quad (31)$$

When solving for the two firms best responses in retail prices it turns out that they are close to identical with the ones derived in Boom (2007). They are given by If firm i undercuts its rival's retail price, its best response from below is then

$$\underline{r}_i(r_j) = \begin{cases} \max \left\{ 2 - k_i, \frac{3}{4} \right\} & \text{if } r_j > \max \left\{ 2 - k_i, \frac{3}{4} \right\} \\ r_j - \mu & \text{if } 0 \leq r_j \leq \max \left\{ 2 - k_i, \frac{3}{4} \right\} \end{cases} \quad (32)$$

with $\mu \rightarrow 0$ being the smallest unit in which retail prices can be announced. If firm i sets $r_i > r_j$, then it is indifferent between all prices that satisfy this restriction, because its profit, given in (31), does not depend on the level of r_i .

The overall best response is determined by the comparison of $\underline{\pi}_i(\underline{r}_i(r_j), r_j)$, derived from (27) and (32), with $\pi_i(r_i, r_j) \big|_{r_i=r_j}$ from (28), (29) or (30), respectively, and with $\bar{\pi}_i(r_i, r_j)$ defined in (31).

Suppose that $k_i > k_j$. Then the overall best response of firm i for $k_i > k_j \geq \sqrt{5 - k_j^2}$ is given by $r_i(r_j) = \underline{r}_i(r_j)$ from (32). For $\min \left\{ k_i, \sqrt{5 - k_i^2} \right\} > k_j \geq 0$ the overall best response is

$$r_i(r_j) \begin{cases} = \max \left\{ 2 - k_i, \frac{3}{4} \right\} & \text{if } r_j > \max \left\{ 2 - k_i, \frac{3}{4} \right\}, \\ = r_j - \mu & \text{if } \hat{r} < r_j \leq \max \left\{ 2 - k_i, \frac{3}{4} \right\}, \\ > r_j & \text{if } \max\{0, 1 - 2k_j\} < r_j \leq \min \left\{ \hat{r}, \max \left\{ 2 - k_i, \frac{3}{4} \right\} \right\}, \\ \geq r_j & \text{if } 0 < r_j \leq \min \left\{ 1 - 2k_h, \max \left\{ 2 - k_j, \frac{3}{4} \right\} \right\}. \end{cases} \quad (33)$$

with

$$\hat{r} = \begin{cases} \frac{3 - \sqrt{2(k_j^2 + k_i^2) - 1}}{2} & \text{if } k_i \geq k_j > \min\{\sqrt{1 - k_i^2}, k_i - 1\}, \\ 1 - k_j & \text{if } 0 \leq k_j \leq k_i - 1, \\ 2 - \sqrt{k_j^2 + k_i^2} & \text{if } 0 \leq k_j \leq \min\{\sqrt{1 - k_i^2}, k_i\}. \end{cases} \quad (34)$$

Firm j 's overall best response in retail prices is the equivalent to $\underline{r}_i(r_j)$ with firm j always undercutting firm i if $k_i > k_j \geq \sqrt{\frac{5}{2}}$. For $\min\{k_i, \sqrt{\frac{5}{2}}\} > k_j \geq$

$\max \{0, \frac{k_i-1}{2}\}$ firm j's best response in retail prices is given by

$$r_j(r_i) \begin{cases} = \max \{2 - k_j, \frac{3}{4}\} & \text{if } r_i > \max \{2 - k_j, \frac{3}{4}\}, \\ = r_i - \mu & \text{if } \max \{2 - k_j, \frac{3}{4}\} \geq r_i > r'_i, \\ = r_i & \text{if } r'_i \geq r_i > r''_i, \\ > r_i & \text{if } r''_i \geq r_i > 0, \\ \geq 0 & \text{if } r_i = 0, \end{cases} \quad (35)$$

where the critical prices r'_i and r''_i for the rival with the larger capacity is defined as:

$$r'_i = \max \left\{ \frac{3 - \sqrt{4k_j^2 - 1}}{2}, 2 - \sqrt{2k_j} \right\}, \quad (36)$$

and

$$r''_i = \max \left\{ \frac{3 - \sqrt{4k_i^2 - 1}}{2}, \frac{5 - \sqrt{12k_j^2 + 6k_i^2 - 2}}{3}, 0 \right\} \quad (37)$$

for $k_i > 1$,

$$r''_i = \begin{cases} \max \left\{ \frac{3 - \sqrt{4k_i^2 - 1}}{2}, \frac{5 - \sqrt{12k_j^2 + 6k_i^2 - 2}}{3} \right\} & \text{if } k_i \geq k_j \geq \sqrt{\frac{1-k_i^2}{2}}, \\ 2 - \sqrt{2k_j^2 + k_i^2} & \text{if } \sqrt{\frac{1-k_i^2}{2}} > k_j \geq 0 \end{cases} \quad (38)$$

for $\frac{1}{\sqrt{2}} < k_i \leq 1$ and

$$r''_i = \max \left\{ 2 - \sqrt{2k_i}, 2 - \sqrt{2k_j^2 + k_i^2} \right\} \quad (39)$$

for $0 \leq k_i < \frac{1}{\sqrt{2}}$.

For $\frac{k_i-1}{2} > k_j > 0$ firm j's best response in retail prices is

$$r_j(r_i) \begin{cases} = 2 - k_j & \text{if } r_i > 2 - k_j, \\ = r_i - \mu & \text{if } 2 - k_j \geq r_i > r'_i, \\ = r_i & \text{if } r'_i \geq r_i \geq 1 - 2k_j, \\ \geq 0 & \text{if } 1 - 2k_j > r_i \geq 2 - k_i, \\ > r_i & \text{if } 2 - k_i > r_i > 0, \\ \geq 0 & \text{if } r_i = 0, \end{cases} \quad (40)$$

where r'_j is defined in (36).

Now suppose that $k_A = k_B = k$, then each firm has the same best response in retail prices. If $k > \sqrt{\frac{5}{2}}$, then both firms' overall best response is given by $\underline{r}_i(r_j)$ from equation (32). If $0 \leq k < \sqrt{\frac{5}{2}}$, then each firm's best response in retail prices is given by:

$$r_i(r_j) = \begin{cases} = \max \left\{ 2 - k, \frac{3}{4} \right\} & \text{if } r_j > \max \left\{ 2 - k, \frac{3}{4} \right\}, \\ = r_j - \mu & \text{if } \max \left\{ 2 - k, \frac{3}{4} \right\} \geq r_j > \hat{r} \\ \geq r_j & \text{if } r_j = \hat{r} \\ > r_j & \text{if } \hat{r} > r_h > 0, \\ \geq 0 & \text{if } r_j = 0, \end{cases} \quad (41)$$

with \hat{r} from (34) with $k_i = k_j = k$. Then again there is always a Nash equilibrium with $r_A = r_B = 0$. Since the Nash equilibria in retail prices depend on when these best response functions change from undercutting to matching and then to overbidding, and since the retail prices at which this happens do not change, the possible Nash equilibria in retail prices are also close to the same as in proposition 3 with the exception of case (v) where now the only possible Nash equilibrium in retail prices implies $r_i = r_j = 0$.

B.3 The Separated Duopoly with Rationing

Firm i 's best response with $i = A, B$ can be derived from maximizing the profit function (22). It is given by

$$k_i = \begin{cases} 1 - \frac{z}{1-k_j} & \text{if } k_j \leq (3 - \sqrt{1 + 8z})/4, \\ \min \{k_j - \epsilon, \frac{1-z}{2}\} & \text{if } k_j > (3 - \sqrt{1 + 8z})/4, \end{cases} \quad (42)$$

with $\epsilon \rightarrow 0$. The best response function is downward sloping, jumps downward at the given threshold and finally increases up to the given limit.

Proposition 12 (separated duopoly with rationing) *If the two generators A and B simultaneously choose their capacities, a subgame perfect Nash equilibrium in pure strategies does not exist. If firm A chooses k_A before firm B chooses k_B , then a unique subgame perfect Nash equilibrium exists with $k_A = (3 - \sqrt{1 + 8z})/4$ and $k_B = (3 - \sqrt{1 + 8z})/2$ and a retail price of $r^{sdr} = (1 + \sqrt{1 + 8z})/4$.*

Proof: The non-existence of a subgame perfect Nash equilibrium in pure strategies with simultaneous capacity choice follows from the analysis of the two best response functions given in (42). The unique subgame perfect Nash equilibrium with sequential capacity choice is derived from the piecewise analysis of firm A's profit function after substituting firm B's best response. The retail price is still characterized by proposition 6. Taking into account Assumption 2 and substituting firm A's equilibrium capacity yields then the retail price for the sequential capacity choice case. ■

References

- Allaz, B. and J.-L. Villa (1993), 'Cournot Competition Forward Markets and Efficiency'. *Journal of Economic Theory*, 59(1), pp. 1–16.
- Bergman, L., G. Brunekreeft, C. Doyle, N.-H. M. von der Fehr, D. M. Newbery, M. Pollitt and P. Règebeau (eds.) (1999), *A European Market for Electricity?*, London and Stockholm: CEPR and SNS.
- Boom, A. (2002), 'Investments in Electricity Generation Capacity under Different Market Structures with Price Responsive Demand'. Diskussionsbeiträge des Fachbereichs Wirtschaftswissenschaft 2002/18, Freie Universität Berlin.

- Boom, A. (2007), 'Vertically Integrated Firms' Investments in Electricity Generating Capacities'. Discussion Paper 2007-14, Centre for Industrial Economics, University of Copenhagen.
- Boom, A. (2008), 'Equilibrium Selection with Risk Dominance in a Multiple-unit Unit Price Auction'. Working Paper 2-2008, Department of Economics, Copenhagen Business School.
- Boom, A. (2009), 'Vertically Integrated Firms' Investments in Electricity Generating Capacities'. *International Journal of Industrial Organization*, 27, pp. 544–551.
- Borenstein, S. (2002), 'The Trouble with Electricity Markets: Understanding California's Electricity Restructuring Disaster'. *Journal of Economic Perspectives*, 16, pp. 191–211.
- Borenstein, S. and S. P. Holland (2005), 'On the Efficiency of Competitive Electricity Markets with Time-invariant Retail Prices'. *The Rand Journal of Economics*, 36, pp. 469–493.
- Borenstein, S., J. Bushnell and S. Stoft (2000), 'The Competitive Effects of Transmission Capacity in a Deregulated Electricity Industry'. *The Rand Journal of Economics*, 31, pp. 294–325.
- Bosco, B., L. Parisio and M. Pelagatti (2012), 'Strategic Bidding in Vertically Integrated Power Markets with an Application to the Italian Electricity Auctions'. *Energy Economics*, 34, pp. 2046–2057.
- Buehler, S. and A. Schmutzler (2005), 'Asymmetric Vertical Integration'. *Advances in Theoretical Economics*, 5, p. Article 1, [http: www.bepress.com/bejte/advances/vol5/iss1/art1](http://www.bepress.com/bejte/advances/vol5/iss1/art1).
- Buehler, S. and A. Schmutzler (2008), 'Intimidating Competitors - Endogenous Vertical Integration and Downstream Investment in Successive Oligopoly,'. *International Journal of Industrial Organization*, 26, pp. 247–265.
- Bushnell, J. (2007), 'Oligopoly Equilibria in Electricity Contract Markets'. *Journal of Regulatory Economics*, 32., pp. 225–245.
- Bushnell, J. B., E. T. Mansur and C. Saravia (2008), 'Vertical Arrangements, Market Structure and Competition'. *American Economic Review*, 98(1), pp. 237–266.

- Castro-Rodriguez, F., P. L. Marín and G. Siotis (2009), ‘Capacity Choices in Liberalized Electricity Markets’. *Energy Policy*, pp. 2574–2581.
- Crampes, C. and A. Creti (2005), ‘Capacity Competition in Electricity Markets’. *Economia delle Fonti di Energia e dell’ Ambiente*, 48(2).
- Crampes, C. and N. Fabra (2005), ‘The Spanish Electricity Industry: Plus ça change...’. *The Energy Journal*, 26(European Energy Liberalisation Special Issue), pp. 127–153.
- Cramton, P. and A. Ockenfels (2012), ‘Economics and Design of Capacity Markets for the Power Sector’. *Zeitschrift für Energiewirtschaft*, 36, pp. 113–134.
- Drèze, J. H. and E. Sheshinski (1976), ‘Demand Fluctuation, Capacity Utilization and Costs’. *The American Economic Review*, 66, pp. 731–744.
- von der Fehr, N.-H. and P. V. Hansen (2010), ‘Electricity Retailing in Norway’. *The Energy Journal*, 31(1), pp. 25–45.
- von der Fehr, N.-H. M. and D. C. Harbord (1993), ‘Spot Market Competition in the UK Electricity Industry’. *The Economic Journal*, 103, pp. 531–546.
- von der Fehr, N.-H. M. and D. C. Harbord (1997), ‘Capacity Investment and Competition in Decentralised Electricity Markets’. Memorandum 27, Department of Economics, University of Oslo.
- de Frutos, M. and N. Fabra (2012), ‘How to Allocate Forward Contracts: The Case of Electricity Markets’. *European Economic Review*, 56, pp. 451–469.
- Gabszewicz, J. J. and S. Poddar (1997), ‘Demand Fluctuations and Capacity Utilization under Duopoly’. *Economic Theory*, 10, pp. 131–146.
- Green, R. J. and D. M. Newbery (1992), ‘Competition in the British Electricity Spot Market’. *Journal of Political Economy*, 100, pp. 929–953.
- Grimm, V. and G. Zoettl (2013), ‘Investment Incentives and Electricity Spot Market Competition’. *Journal of Economics & Management Strategy*, 22(4), pp. 832–851.
- Joskow, P. L. (2006), ‘Competitive Electricity Markets and Investment in New Generating Capacity’. In: Helm, D. (ed.), ‘The New Energy Paradigm’, Oxford and New York: Oxford University Press.

- Joskow, P. L. and J. Tirole (2000), ‘Transmission Rights and Market Power on Electric Power Networks’. *The Rand Journal of Economics*, 31, pp. 450–487.
- Joskow, P. L. and J. Tirole (2006), ‘Retail Electricity Competition’. *The Rand Journal of Economics*, 37, pp. 799–815.
- Klemperer, P. D. and M. A. Meyer (1989), ‘Supply Function Equilibria in Oligopoly under Uncertainty’. *Econometrica*, 57, pp. 1243–1277.
- Kreps, D. M. and J. A. Scheinkman (1983), ‘Quantity Precommitment and Bertrand Competition Yield Cournot Outcomes’. *The Bell Journal of Economics*, 14, pp. 326–337.
- Kwoka, J. (2002), ‘Vertical Economies in Electric Power: Evidence on Integration and its Alternatives’. *International Journal of Industrial Organization*, 20, pp. 653–671.
- Kwoka, J., M. Pollitt and S. Sergici (2010), ‘Divestiture Policy and Operating Efficiency in U.S. Electric Power Distribution’. *Journal of Regulatory Economics*, 38, pp. 86–109.
- Léautier, T.-O. (2001), ‘Transmission Constraints and Imperfect Markets for Power’. *Journal of Regulatory Economics*, 19, pp. 27–54.
- Mansur, E. T. (2007), ‘Upstream Competition and Vertical Integration in Electricity Markets’. *Journal of Law and Economics*, 50, pp. 125–156.
- Le Coq, C. (2002), ‘Strategic Use of Available Capacity in the Electricity Spot Market’. SSE/EFI Working Paper Series in Economics and Finance 496, Stockholm School of Economics.
- Murphy, F. H. and Y. Smeers (2005), ‘Generation Capacity Expansion in Imperfectly Competitive Restructured Electricity Markets’. *Operations Research*, 53, pp. 646–661.
- Newbery, D. (1999), ‘The UK Experience: Privatization with Market Power’. In: Bergman, L., G. Brunekreeft, C. Doyle, N.-H. M. von der Fehr, D. M. Newbery, M. Pollitt and P. Régibeau (eds.), ‘A European Market for Electricity’, pp. 89–115, London and Stockholm: CEPR and SNS.
- Newbery, D. (2005), ‘Electricity Liberalisation in Britain: The Quest for a Satisfactory Wholesale Market.’ *The Energy Journal*, 26(European Energy Liberalisation Special Issue), pp. 43–70.

Wilson, R. (2002), 'Architecture of Power Markets'. *Econometrica*, 70, pp. 1299–1340.