

Vertical Structure and the Risk of Rent Extraction in the Electricity Industry

Anette Boom* Stefan Buehler†

This draft: May 31, 2019

Abstract

This paper studies how competition and vertical structure jointly determine generating capacities, retail prices, and welfare in the electricity industry. Analyzing a model in which demand is uncertain and retailers must commit to retail prices before they buy electricity in the wholesale market, we show that welfare is highest if competition in generation and retailing is combined with vertical separation. Vertically integrated generators choose excessively high retail prices and capacities to avoid rent extraction in the wholesale market when their retail demand exceeds their capacity. Vertical separation eliminates the risk of rent extraction and yields lower retail prices.

Keywords: Electricity, Generating Capacities, Vertical Integration, Monopoly, 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, Institute of Economics (FGN-HSG), Varnbühlstr. 19, CH-9000 St. Gallen, e-mail: stefan.buehler@unisg.ch

1 Introduction

Electricity markets around the world have recently been reformed in an effort to improve their performance. In several countries, legislators have introduced competition into statutory monopoly and imposed mandatory unbundling on vertically integrated electricity generators.¹ While introducing competition has hardly been controversial, it is fair to say that the role of vertical structure in the electricity industry is not very well understood. Why should electricity generators be vertically separated from other layers of the industry, as suggested by legislators (see, e.g., European Commission (2010))? Does mandatory vertical separation undermine investments in generating capacity when generators compete (Joskow, 2006)?

In this paper, we employ a simple model of the electricity industry to study how competition and vertical structure jointly determine generating capacities, retail prices, and welfare. We consider three vertical layers—generation, wholesaling, and retailing—and examine four industry configurations that vary with respect to vertical structure and competition: social optimum, integrated monopoly, integrated duopoly, and separated duopoly. Throughout, we assume that retail demand is rationed if total retail demand exceeds aggregate capacity. Excess demand thus leads to “brownouts” (rather than “blackouts,” which would reflect a complete market breakdown) in the event of insufficient capacity.

Our key result is that—apart from the social optimum—welfare is highest (lowest)

¹In the UK, for instance, the industry was 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 into generation (Newbery, 1999, 2005). Similarly, the Californian restructuring bill from 1996 forced regulated utilities to divest generation facilities (Borenstein, 2002). The European Union ruled in its Directive 2003/54/EC on 26 June 2003 that electricity generating firms must be functionally disintegrated.

if competition is combined with vertical separation (vertical integration, respectively). The integrated monopoly yields an intermediate level of welfare. Our analysis thus supports the view that introducing competition into the electricity industry should be combined with vertical separation, even though investments in generating capacity will indeed be smaller than with vertical integration.

The driving force behind this result is the risk of rent extraction that is associated with vertical integration: An integrated generator whose capacity is insufficient to serve own retail demand at pre-determined prices must buy electricity in the wholesale market.² This drives up the wholesale price to a level at which the troubled generator must give up its rent. To avoid such rent extraction in the wholesale market, competing integrated generators not only make large capacity investments, but also set excessively high retail prices, thereby avoiding commitments to large retail sales. The combination of high capacities and low retail demand gives rise to inefficient capacity utilization. Vertical separation eliminates the risk of rent extraction because vertically separated generators do not need to serve an uncertain retail demand at retail prices that they have committed to. As a result, both capacity investments and prices are lower, and welfare is higher, even though consumers are rationed in high-demand states when demand exceeds capacity. Intuitively, the result follows because, from a welfare perspective, it is better to reduce demand in high-demand states only (by rationing) than to reduce demand in *all* states (by higher prices).

The risk of rent extraction emerges from the interplay of three ingredients of our

²It is well known that residential consumers typically buy electricity at pre-determined prices. According to Borenstein and Bushnell (2018), only 4% of US residential customers face time-varying electricity prices (and among the few customers on time varying electricity prices most face relatively static time-of-use tariffs).

analysis: (i) uncertain retail demand; (ii) retailers must commit to retail prices before the wholesale price is determined; and (iii) the wholesale price increases when demand exceeds capacity. In our view, these are natural ingredients of any model of the electricity industry that studies multiple vertical layers. It should be clear, though, that the risk of rent extraction is mitigated when there are more than two generators, in which case competition among generators may prevent the wholesale price from increasing even if one of the firms lacks sufficient capacity. However, if all generators must offer some of their capacity in order to satisfy aggregate demand, the risk of rent extraction persists.

Our paper contributes to two related strands of literature. First, we add 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), Borenstein and Holland (2005), Murphy and Smeers (2005), Grimm and Zoettl (2013) as well as Fabra *et al.* (2011)). The key difference to these papers is that we study the role of vertical structure for capacity choices when demand is uncertain.

Second, we add to the literature on the role of vertical structure in the electricity industry, which has either focused on the loss of economies of scope from (mandatory) vertical separation (see e.g. Kwoka *et al.* (2010) and Kwoka (2002)), or interpreted vertical integration into the retail market as forward contracts and analysed their effects on wholesale prices (see, e.g., Bushnell (2007), Mansur (2007), Bushnell *et al.* (2008), de Frutos and Fabra (2012) and Bosco *et al.* (2012)).³

³These papers build on Allaz and Villa (1993) who show that, if two Cournot competitors can sell quantities also in a forward market and arbitrageurs equalize the prices in the forward and the spot market, then the spot market price is lower than without the opportunity to sell forward. Mahenc and Salanié (2004) show that this result fails to hold if firms compete in prices rather than

None of these papers studies the role of vertical structure in determining generating capacities. Baldursson and von der Fehr (2007) analyze the effect of vertical integration on the performance of long-term and spot markets when spot market prices are uncertain and both independent retailers and electricity generators are risk averse. They find that vertical integration impairs market performance by increasing the gap between contract prices and expected spot prices. This effect disappears if agents are risk neutral as in our framework.

The paper closest to ours is Boom (2009), which uses a similar analytical framework. There are two important differences to this paper. First, we presume that excess demand leads to rationing and brownouts (rather than blackouts and complete market breakdown). In doing so, we eliminate extreme punishments from blackouts as the driving force behind investment and pricing decisions. Second, we introduce the separated duopoly configuration, which is indispensable for studying how competition and vertical structure jointly determine generating capacities, retail prices, and welfare.

The remainder of the paper is structured as follows. Section 2 introduces the model, and Section 3 characterizes the equilibrium outcomes in the various market configurations. Section 4 provides a ranking of the market configurations in terms of capacities, retail prices, and welfare, and discusses the role of rent extraction in determining this ranking. Finally, Section 5 offers conclusions and directions for future research.

quantities.

2 The Model

2.1 Vertical Structure

We consider a stripped-down model of the electricity industry with three vertical layers: generation, the wholesale market, and the retail market (see Figure 1). There is ex ante uncertainty about the level of retail demand, which is represented by the demand shock $\varepsilon \geq 0$. Before the demand uncertainty is resolved, generators A and B must choose capacities k_A and k_B , and retailers C and D must commit to retail prices r_C and r_D at which they are willing to serve demand. After the demand uncertainty is resolved and retailers have learned their demands $d_C(r_C, r_D, \varepsilon)$ and $d_D(r_C, r_D, \varepsilon)$, retailers buy electricity from generators on the wholesale market. An auction determines the wholesale price $p(p_A, p_B)$ and the quantities $y_i(p_A, p_B) \leq k_i, i = A, B$, which generators supply to the grid, where p_A, p_B denote the generators' wholesale bids. If necessary, the rationing of retail demand assures that total capacity covers aggregate demand, $d_C + d_D \leq k_A + k_B$.

We study how generating capacities, retail prices, and welfare are endogenously determined across the following market configurations: (i) social optimum, (ii) integrated monopoly, (iii) integrated duopoly, and (iv) separated duopoly. The wholesale market is superfluous in the benchmark configurations (i) and (ii), where capacities and retail prices are chosen so as to maximize welfare and industry profit, respectively. In configurations (iii) and (iv), however, the wholesale market plays a crucial role. To see this, suppose that two integrated supply chains, (A, C) and (B, D) , interact on the wholesale market. In addition, assume that, say, (A, C) cannot serve its own retail demand, $k_A < d_C$, while (B, D) has excess

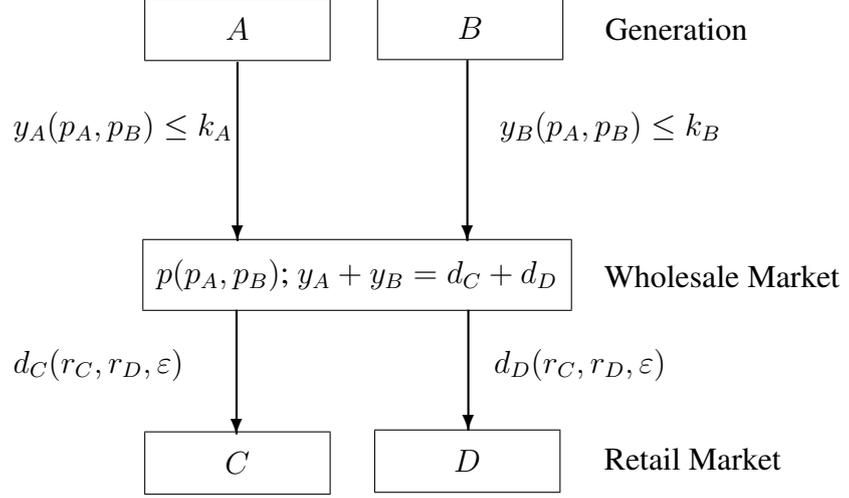


Figure 1: Three Vertical Layers of the Electricity Industry

capacity, $k_B > d_D$. Then, (B, D) has an incentive to drive up the wholesale price $p(p_A, p_B)$ in order to extract the rent of (A, C) when selling (part of) its excess capacity.

2.2 Demand and Supply

We assume that consumer 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 aggregate amount of electricity consumed, $r \in \{r_C, r_D\}$ is the retail price per unit of electricity, and the demand shock ε is uniformly distributed on the interval $[0, 1]$.⁴ Maximizing $V(x, \varepsilon, r)$ with respect to x yields the following

⁴This specification implies that the demand shock generates a negative effect on consumer surplus. Our key results do not rely on this assumption (see Boom and Buehler (2014)).

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 lower retail price. If retail prices are identical, consumers choose each retailer with equal probability.

The marginal cost of generating electricity is constant and normalized to zero. The total cost of electricity generator $i = A, B$ is given by $C(k_i) = zk_i$, where z is the constant unit cost of capacity and k_i is the generating capacity installed by firm i . We assume that capacity cost z satisfies $z \in [0, 1/(2\sqrt{2})]$ to ensure strictly positive equilibrium capacity investments by all generators in all market configurations. For simplicity, we further assume that the marginal cost of selling electricity to final consumers is constant and normalized to zero.

2.3 Timing

We assume the following sequence of events:

(1) Generators $i = A, B$ simultaneously choose their capacities k_i . In line with Fabra *et al.* (2011), firms subsequently decide whether or not they want to mothball part of their capacities, starting with the smaller-capacity firm (or a randomly chosen firm in case of identical capacities).⁵

(2) Retailers $\ell = C, D$ simultaneously set their retail prices r_ℓ , and consumers

⁵This assumption secures the existence of a pure-strategy subgame perfect Nash equilibrium in the separated duopoly configuration.

subscribe to the firm with the lower retail price.

- (3) The demand shock $\varepsilon \in [0, 1]$ is realized, and retailers learn their demand.
- (4) In the wholesale market each generator simultaneously bids a reservation price p_i for its full capacity k_i , $i = A, B$ in a uniform price auction, which is modelled as in von der Fehr and Harbord (1993) and (1997).⁶ Retailers bid their clients' fixed demand as long as the wholesale price ensures non-negative profits. The auctioneer determines the market-clearing wholesale price p and the quantities y_i generators can sell. If necessary, retail demand is rationed.
- (5) Finally, supply and demand are balanced, and deliveries and payments are exchanged.

Note that we require each firm to bid its total capacity in the wholesale market and therefore abstract from strategic capacity withholding (Le Coq, 2002; Crampes and Creti, 2005). Instead, we focus on the role of vertical market structure for capacity investments and retail prices. Whenever there are multiple equilibria, we assume that firms can coordinate on a Pareto dominant equilibrium.

3 Market Configurations

In this section, we characterize the equilibrium outcomes in the various market configurations. To simplify exposition, we use reduced-form notation wherever

⁶Several European wholesale markets are indeed organized as uniform price auctions (e.g., the Nord Pool in Scandinavia, or the Spanish wholesale market). Before the reform in 2001, the Electricity Pool in England and Wales was also organized as uniform price auction. See Bergman *et al.* (1999), Crampes and Fabra (2005) and Newbery (2005).

possible. All proofs are relegated to the Appendix.

3.1 Benchmark Cases

Our benchmark configurations are the social optimum and the integrated monopoly. In these configurations, price and capacity are chosen so as to maximize expected welfare and profit, respectively.

Social Optimum. The social planner chooses the generating capacity k and retail price r so as to maximize expected welfare,

$$\max_{k,r} \mathbb{E}[W(k, r, \varepsilon)] = \max_{k,r} \{ \mathbb{E}[U(\min\{x(r, \varepsilon), k\}, \varepsilon)] - zk \}, \quad (3)$$

where the expectation is taken over ε . We derive the following result.

Proposition 1 (social optimum) *The welfare-maximizing retail price and capacity are given by $r^s = 0$ and $k^s(z) = 2 - \sqrt{2z}$.*

Proposition 1 shows that the welfare-maximizing retail price is zero, and that rationing optimally occurs in high-demand states if capacity is costly ($z > 0$). Since the marginal cost of generating electricity is zero, restricting consumer demand is optimal only if capacity is binding. Put differently, the social planner prefers rationing in high-demand states to setting a strictly positive retail price, as the latter would imply that demand is restricted in all demand states.

Integrated Monopoly. The integrated monopoly chooses the generating capacity

k and retail price r so as to maximize expected profit

$$\max_{k,r} \{ \mathbb{E}[r \min\{x(r, \varepsilon), k\}] - zk \}. \quad (4)$$

The monopoly's profit-maximizing retail price $r^m(z)$ and capacity $k^m(z)$ are characterized by equations (14) and (15) in Appendix A.2, where we also derive the following results.

Proposition 2 (integrated monopoly) *The monopoly retail price and capacity satisfy $r^m(z) > r^s = 0$ and $k^m(z) < k^s(z)$, with $dr^m/dz \geq 0$ and $dk^m/dz \leq 0$, for any admissible z .*

The result states that the integrated monopoly charges a higher retail price than the social planner and invests less in generating capacity for any capacity cost z . Due to the strictly positive retail price, consumer demand is smaller than optimal for given capacity cost. Nevertheless, rationing still occurs in high-demand states. In addition, the profit-maximizing price (capacity) increases (decreases) with the cost of capacity. Next, we consider market configurations that combine competition with vertical integration and separation, respectively.

3.2 Integrated Duopoly

In this market configuration, integrated duopoly generators $i, j = A, B, i \neq j$, compete both in the retail and the wholesale market. We use backward induction to work out the equilibrium outcome.

Wholesale Market. Recall that the wholesale market is organized as a uniform price auction. Generating capacities (k_A, k_B) and aggregate retail demand $x_R(r^{id}, \varepsilon) = \min\{x(r^{id}, \varepsilon), k_A + k_B\}$ are predetermined, where $r^{id} = \min\{r_A, r_B\}$ denotes the retail price under integrated duopoly and the subscript R indicates that rationing occurs when demand exceeds capacity. Firm i 's predetermined retail demand is

$$d_i(r_i, r_j, \varepsilon) = \begin{cases} x_R(r^{id}, \varepsilon) & \text{if } r_i = r^{id} < r_j, \\ \frac{1}{2}x_R(r^{id}, \varepsilon) & \text{if } r_i = r_j = r^{id}, \\ 0 & \text{if } r_i > r_j = r^{id}. \end{cases} \quad (5)$$

Note that firm i 's retail demand is zero if it charges a higher price than firm j . The generators' wholesale price bids (p_i, p_j) determine both the wholesale price

$$p(p_i, p_j) = p_i \begin{cases} \text{if } p_i < p_j & \text{and } k_i \geq x_R(r^{id}, \varepsilon) \text{ or} \\ \text{if } p_i \geq p_j & \text{and } k_j < x_R(r^{id}, \varepsilon), \end{cases} \quad (6)$$

and the wholesale quantity

$$y_i(p_i, p_j) = \begin{cases} \min\{k_i, x(r, \varepsilon)\} & \text{if } p_i < p_j, \\ \frac{\min\{k_i, x(r, \varepsilon)\}}{2} + \frac{\min\{\max\{0, x(r, \varepsilon) - k_j\}, k_i\}}{2} & \text{if } p_i = p_j, \\ \min\{\max\{0, x(r, \varepsilon) - k_j\}, k_i\} & \text{if } p_i > p_j \end{cases} \quad (7)$$

that generator i can supply to the grid. Using equations (5) to (7), the profit of an integrated generator is given by

$$\pi_i(p_i, p_j) = r_i d_i(r_i, r_j, \varepsilon) + p(p_i, p_j) [y_i(p_i, p_j) - d_i(r_i, r_j, \varepsilon)] - z k_i. \quad (8)$$

Equation (8) shows that an integrated generator runs the *risk of rent extraction*: Generator i must buy electricity in the wholesale market if its retail demand turns out to exceed its capacity ($d_i(r_i, r_j, \varepsilon) > k_i \geq y_i$). Generator j can then drive up the wholesale price $p(p_i, p_j)$ to a level at which generator i 's rent is fully extracted, provided that it has spare capacity (i.e., $k_j > d_j(r_j, r_i, \varepsilon)$). If both generators can cover own retail demand, however, the wholesale price is zero. Note that the commitment to serve uncertain retail demand at predetermined retail prices is key for the risk of rent extraction. If retail prices were flexible, integrated generators could always increase their retail prices and thereby force their retail demand down to capacity.

Retail Market. An integrated generator can undercut, match, or surcharge its competitor's retail price (see equation (5)). Clearly, undercutting is not profitable if own capacity cannot cover the extra demand from the lower price for most demand states. Similarly, it is not profitable to match a larger rival's retail price because of earlier capacity exhaustion. However, it may be profitable to match a smaller or equally sized rival's retail price. With equal capacities, each firm sells its own capacity in case of rationing, whereas with asymmetric capacities the larger firm acquires the total market profit if the smaller firm cannot serve its own (rationed) retail demand.

With asymmetric capacities matching each other's retail prices can only be an equilibrium if both firms set $r_i = r_j = 0$, so that undercutting is not profitable. However, if the asymmetry in capacities is very large, there exists a retail price range in which the large firm can serve total demand even in the largest demand state, whereas the smaller firm cannot even serve half of total demand when it

matches the rival's retail price in the smallest demand state. The Pareto dominant retail price equilibrium is then either that the large firm sets the monopoly price and the small firm overbids, or that the small firm sets the monopoly price and the large firm overbids or matches. Either way, the large firm earns all the profits in the market. For symmetric, not too large capacities, matching each other's prices at a positive level where firms are just indifferent between underbidding, matching and overbidding is an equilibrium.⁷

Generation. Integrated generators choose capacities in order to maximize profits, anticipating the effects on the retail and wholesale market. As we show in Appendix A.3.3, their symmetric capacities are always large enough to avoid rationing. Our next result summarizes the market outcome for the integrated duopoly.

Proposition 3 (integrated duopoly) *There is a unique Pareto dominant subgame perfect Nash equilibrium in which integrated generators set the retail price*

$$r^{id} = \frac{1}{2} \left(3 - \sqrt{4(k^*)^2 - 1} \right)$$

and each invests in capacity k^ , such that total capacity satisfies $k^{id} = 2k^*$, with*

$$k^* = \arg \max_{k \in \left[\frac{1}{\sqrt{2}}, \min\left\{ \frac{\sqrt{13}}{4}, \bar{k} \right\} \right]} \left\{ \frac{1 - 4k^2 + 3\sqrt{4k^2 - 1}}{8} - zk \right\}, \quad (9)$$

and with \bar{k} being the largest capacity level at which matching the rival's capacity investment yields zero profit. Rationing does not occur in equilibrium.

Generating capacity decreases and the retail price increases if capacity costs z

⁷See Appendix A.3.2 for the derivation of the Pareto dominant retail price equilibrium.

increase. Rationing does not occur in equilibrium, because, as we will show later, integrated generators not only charge excessively high retail prices, but also make excessively large capacity investments in order to avoid rent extraction in the wholesale market.

3.3 Separated Duopoly

In this configuration, separated generators sell to the wholesale market, while separated retailers buy from the wholesale market and serve retail demand.

Wholesale Market. The wholesale price $p(p_A, p_B)$ and the generators' wholesale quantities (y_A, y_B) are determined by the uniform price auction discussed above, given predetermined capacities (k_A, k_B) and retail demand $x_R(r^{sd}, \varepsilon)$, where r^{sd} denotes the retail price in the separated duopoly market configuration. The profit of a separated duopoly generator $i = A, B$ is given by

$$\pi_i(p_A, p_B) = p(p_A, p_B)y_i(p_A, p_B) - zk_i. \quad (10)$$

Since vertically separated generators do not need to commit to serve uncertain retail demand at predetermined prices, there is no risk of rent extraction. Best-response bidding requires each generator to either undercut the competitor or to bid the maximum price $p_i = r^{sd}$ at which retailers break even (at higher wholesale prices, retailers are better off declaring bankruptcy and leaving the market). The undercutting generator serves total demand up to capacity, whereas the other generator serves residual demand.

The equilibrium wholesale price is illustrated in Figure 2. In areas \mathcal{A} (rationing),

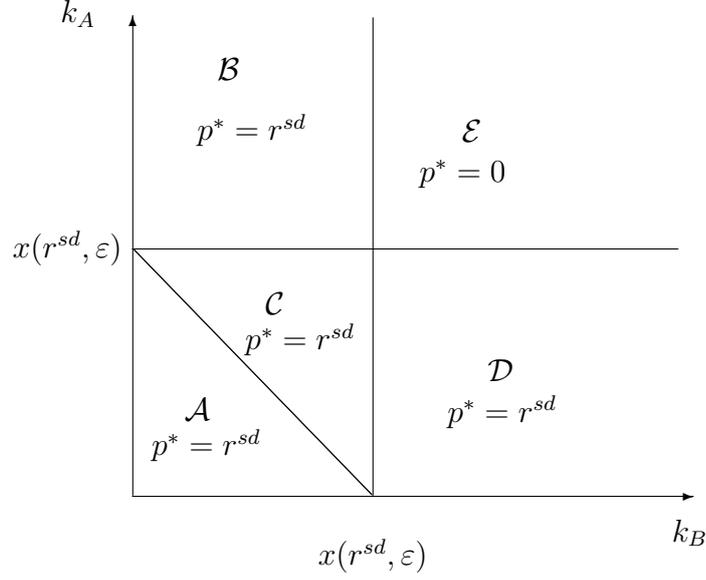


Figure 2: Wholesale prices in the separated duopoly

\mathcal{B} , \mathcal{C} , and \mathcal{D} , the wholesale price is the maximum price at which retailers break even. The larger generator bids this price, whereas the smaller bids sufficiently low, such that the larger does not want to undercut. In area \mathcal{C} there are also equilibria with the roles of the two firms reversed. In order to ensure a unique pay-off equivalent wholesale market equilibrium, we introduce an additional equilibrium selection criterion which resembles risk dominance (see our Assumption 1 in Appendix A.4.1). In area \mathcal{E} , the wholesale price is zero. Here, different from areas \mathcal{A} , \mathcal{B} , \mathcal{C} , and \mathcal{D} , each individual capacity is sufficient to ensure for each individual generator a residual demand of zero, which leads to Bertrand bidding.

Retail Market. Separated retailers $\ell, \iota = C, D, \ell \neq \iota$, choose their retail price in order to maximize expected profit, accounting for potential rationing,

$$\max_{r_\ell} \mathbb{E}_R[r_\ell d_\ell(r_\ell, r_\iota, \epsilon)]. \quad (11)$$

They undercut each other, with the outcome depending on predetermined levels of capacities. If $\min\{k_A, k_B\} \geq 1$, capacities are too large to induce a rent shift from retailers to generators for all demand realizations, and retail prices are competed down to $r_C = r_D = 0$. If $\min\{k_A, k_B\} < 1$, retailers earn zero profits at retail prices $r_C = r_D = 1 - \min\{k_A, k_B\} > 0$, and rents are shifted to generators for all demand states.⁸

Generation. Separated generators choose capacities in order to maximize expected profits (see equation (10)), accounting for the effects of investment on the retail and wholesale market, and factoring in potential rationing. In the Appendix, we show that there are two Pareto dominant asymmetric subgame perfect equilibria in which the smaller generator realizes a larger profit than the larger one. Proposition 4 summarizes the equilibrium outcome.

Proposition 4 (separated duopoly) *In the Pareto dominant asymmetric subgame perfect equilibrium, the retail price and total capacity are given by*

$$r^{sd}(z) = \frac{1 + \sqrt{1 + 8z}}{4} \text{ and } k^{sd}(z) = \frac{3(3 - \sqrt{1 + 8z})}{4}.$$

Note that at this equilibrium retail price and aggregate capacity, rationing occurs in some high-demand states.

⁸This retail price is Pareto dominant in the whole game, but the retailers in the retail stage are indifferent between all retail prices with $\min\{r_C, r_D\} \leq 1 - \min\{k_A, k_B\}$. See also the discussion in Appendix A.4.2.

4 Evaluating Market Configurations

We now evaluate the four market configurations in terms of equilibrium prices, capacities, and welfare. Using Propositions 1 to 4, we derive the following result.

Proposition 5 (ranking) *Consider the following market configurations: social optimum (s), integrated monopoly (m), integrated duopoly (id), and separated duopoly (sd). These market configurations compare as follows in terms of*

(i) *aggregate capacities:*

$$k^s \geq k^{id} > k^{sd} > k^m, \quad \text{if } 0 \leq z \leq 0.0276,$$

$$k^{id} > k^s \geq k^{sd} > k^m, \quad \text{if } z > 0.0276,$$

(ii) *retail prices:*

$$r^{id} \geq r^m > r^{sd} > r^s \quad \text{for all } z,$$

(iii) *welfare levels:*

$$W^s > W^{sd} > W^m \geq W^{id} \quad \text{for all } z.$$

Apart from the social optimum, welfare is highest (lowest) if competition in generation and retailing is combined with vertical separation (vertical integration, respectively). The result follows from the fact that both the retail price and aggregate generating capacity are highest under integrated duopoly (except for very low capacity cost, in which case capacity is higher in the social optimum). Consequently, capacity utilization is inefficient under integrated duopoly, and consumers

suffer from a high deadweight loss because of the excessively high retail price in all demand states.

Our analysis thus suggests that introducing competition into statutory monopoly should be combined with vertical separation. The driving force behind this finding is the risk of rent extraction under integrated duopoly, which induces competing generators to charge excessively high retail prices and make excessively large capacity investments.

5 Conclusion

This paper has studied how competition and vertical structure jointly determine generating capacities, retail prices, and welfare in the electricity industry. The key insight of our analysis is that welfare is highest (lowest) if competition is combined with vertical separation (integration). The driving force behind this result is the risk of rent extraction: A vertically integrated duopoly generator whose capacity is insufficient to serve own retail demand must buy electricity in the wholesale market at an increasing price. To avoid rent extraction, integrated duopoly generators choose excessively high capacities and retail prices, which leads to inefficient capacity utilization. Vertical separation eliminates the risk of rent extraction and allows for effective competition in electricity generation and retailing.

Our analysis suggests that the introduction of competition should be combined with vertical separation, as argued by several legislators. It should be clear, though, that the risk of rent extraction is mitigated when there are more than two generators. If these generators have sufficient spare capacities to make up for the lack

of capacity of an individual generator, competition in the wholesale market may prevent the wholesale price from increasing. In such a setting, vertical separation may not be needed to discipline retail prices and investments in electricity generating capacity. However, if all generators must offer some of their capacity in order to satisfy aggregate demand, the risk of rent extraction persists.

There is ample scope for future research. For instance, it would be interesting to analyze the role of vertical market structure when there is repeated interaction. In addition, one could study endogenous (and possibly asymmetric) vertical integration. We hope to address these issues in future research.

Acknowledgements

We thank Gregor Zöttl, Nicholas Shunda, Chloé Le Coq, the referees and the co-editor for helpful comments, as well as seminar audiences at Copenhagen University, the University of Groningen, the University of Utrecht, the DIW Berlin, Aarhus University, and at 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.

A Appendix

A.1 Proof of Proposition 1 (Social Welfare)

Spelling out the expected social welfare function in equation (3), we have

$$\mathbb{E}[W(k, r, \varepsilon)] = \begin{cases} \int_0^1 U(x(r, \varepsilon), \varepsilon) d\varepsilon - zk & \text{if } r \geq 2 - k, \\ \int_0^{k-1+r} U(x(r, \varepsilon), \varepsilon) d\varepsilon + \\ \int_{k-1+r}^1 U(k, \varepsilon) d\varepsilon - zk & \text{if } 2 - k > r \\ \geq \max\{1 - k, 0\}, \\ \int_0^1 U(k, \varepsilon) d\varepsilon - zk & \text{if } 1 - k > r \geq 0. \end{cases} \quad (12)$$

Maximizing (12) with respect to r yields $r = \max\{1 - k, 0\}$, which is the optimal retail price for given capacity. Substituting back into (12) and maximizing with respect to k yields $r^s = 0$ and $k^s = 2 - \sqrt{2z}$ for $0 \leq z \leq \frac{1}{2\sqrt{2z}}$, such that consumers are rationed in some high-demand states.

A.2 Proof of Proposition 2 (Integrated Monopoly)

Spelling out the expected profit function in (4), we have

$$\mathbb{E}[\pi(r, k, \varepsilon)] = \begin{cases} \int_{\max\{r-1, 0\}}^1 r(1 + \varepsilon - r) d\varepsilon \\ - zk, & \text{if } r \geq 2 - k, \\ \int_{\max\{r-1, 0\}}^{k-1+r} r(1 + \varepsilon - r) d\varepsilon \\ + \int_{\min\{k-1+r, 1\}}^1 rkd\varepsilon - zk, & \text{if } \max\{0, 1 - k\} \leq r \\ < 2 - k, \\ \int_0^1 rkd\varepsilon - zk, & \text{if } 0 \leq r \leq 1 - k \end{cases} \quad (13)$$

Maximizing (13) with respect to r and k yields

$$r^m = \frac{1}{12} \left[(3 + 2z) \left(1 + \frac{3 + 2z}{g(z)} \right) + g(z) \right] \text{ and} \quad (14)$$

$$k^m = \frac{1}{12z(g(z))^2} \{ 4h(z)(g(z) - 5z - 3) + 21z(g(z))^2 - 9zg(z) - 8z^4 - 4(g(z) - 81)z^3 - 2[(42 + g(z))g(z) - 81]z^2 - 27z \} \quad (15)$$

with $g(z) \equiv \sqrt[3]{27 + 54z - 180z^2 + 8z^3 + 12h(z)}$

and $h(z) \equiv \sqrt{3z^2 \{ 2z[4z(9 - z) - 27] - 27 \}}$.

for $0 \leq z \leq 1$. One can show that $r^m \leq \min\{2 - k^m, 1\}$, which implies that demand is always strictly positive and that rationing is necessary for some high-demand states. It is straightforward to show that $dr^m/dz \geq 0$ and $dk^m/dz \leq 0$.

A.3 Proof of Proposition 3 (Integrated Duopoly)

In order to derive a subgame perfect equilibrium we apply backward induction. Specifically, we start with the analysis of the final stage, the bidding behaviour of the two integrated firms in the wholesale market. Next, we analyze the second stage, the retail price competition between the two integrated firms. In this stage, firms anticipate the consequences that their respective retail positions have for the outcome in the wholesale market. Finally, we consider the capacity choices of the two firms, assuming that firms anticipate how their choices will influence the outcomes on the retail and on the wholesale market.

A.3.1 The Wholesale Market

Firm i 's profit function in terms of the bids on the wholesale market is given by:

$$\pi_i(p_i, p_j; r_i, r_j, \varepsilon) = \begin{cases} r^{id}d_i(r_i, r_j, \varepsilon) + p(p_i, p_j, \varepsilon) \\ \quad \cdot (\min\{k_i, x^R(r^{id}, \varepsilon)\} - d_i(r_i, r_j, \varepsilon)) & \text{if } p_i < p_j, \\ \\ r^{id}d_i(r_i, r_j, \varepsilon) + \frac{p(p_i, p_j, \varepsilon)}{2} \\ \quad \cdot [\min\{k_i, x^R(r^{id}, \varepsilon)\} - d_i(r_i, r_j, \varepsilon) \\ \quad + \max\{x^R(r^{id}, \varepsilon) - k_j, 0\} - d_i(r_i, r_j, \varepsilon)] & \text{if } p_j = p_i, \\ \\ r^{id}d_i(r_i, r_j, \varepsilon) + p(p_i, p_j, \varepsilon) \\ \quad \cdot (\max\{x^R(r^{id}, \varepsilon) - k_j, 0\} - d_i(r_i, r_j, \varepsilon)) & \text{if } p_i > p_j, \end{cases}$$

with $i, j = A, B$ and $i \neq j$. We need to distinguish two different situations: (i) both firms are able to serve the demand of their retail customers; (ii) only one firm is able to serve the demand of its retail customers.⁹

- (i) Both firms can serve their respective retail costumers if $k_i \geq d_i(r_i, r_j, \varepsilon)$ holds for $i, j = A, B$ and $i \neq j$. Thus, firm i becomes a net seller on the wholesale market if it underbids its rival because $\min\{k_i, x^R(r^{id}, \varepsilon)\} - d_i(r_i, r_j, \varepsilon) \geq 0$ holds. It becomes a net buyer each time it overbids because

⁹Note that $k_i < d_i(r_i, r_j)$ cannot hold for both firms because of rationing.

$\max\{x^R(r^{id}, \varepsilon) - k_j, 0\} - d_i(r_i, r_j, \varepsilon) \leq 0$ necessarily holds. Therefore, both firms want to undercut their rival and $p_i = p_j = 0 = p(p_i, p_j, \varepsilon)$ is the unique Nash equilibrium. Substituting this into the profit function $\pi_i(p_i, p_j; r_i, r_j, \varepsilon)$ yields $\pi_i(r_i, r_j, \varepsilon) = r^{id}d_i(r_i, r_j, \varepsilon)$, which is the profit they will anticipate in the earlier retail stage for all demand shocks ε for which both firms can satisfy their respective retail demand.

- (ii) Only firm j can meet its retail demand, $k_j \geq d_j(r_i, r_j, \varepsilon)$, while firm i cannot, $k_i < d_i(r_i, r_j, \varepsilon)$. Then firm i is always a net buyer and firm j a net seller in the auction irrespective of their price bids. Consider first the net buyer, firm i . Overbidding firm j at a given p_j may lead to a higher wholesale price, $p_i > p_j$, if firm j is not able to cover total demand, and will certainly lead to a higher quantity that firm i must buy from firm j because of

$$\begin{aligned} k_i - d_i(r_i, r_j, \varepsilon) &\geq \max\{x^R(r^{id}, \varepsilon) - k_j, 0\} - d_i(r_i, r_j, \varepsilon), \\ \Leftrightarrow k_i &\geq \max\{x^R(r^{id}, \varepsilon) - k_j, 0\}. \end{aligned}$$

Thus, in any Nash equilibrium the net buyer, firm i , always underbids the net seller, firm j , such that $p_i < p_j$.

Now consider the net seller, firm j . For a given bid p_i of the net buyer, firm i , firm j sells a lower quantity $x^R(r^{id}, \varepsilon) - k_i - d_j(r_i, r_j, \varepsilon)$ when overbidding firm i , compared to the quantity $\min\{k_j, x^R(r^{id}, \varepsilon)\} - d_j(r_i, r_j, \varepsilon)$ when underbidding. At the same time, the resulting auction price $p(p_i, p_j, \varepsilon) = p_j > p_i$ is certainly higher, when overbidding compared to underbidding. Firm j 's overbidding profit increases linearly in its bid $p_j > p_i$ and therefore its optimal overbidding price is the maximum price $p_j = \bar{p}_j(r_i, r_j, \varepsilon)$ that it can bid, such that the net buyer firm i realizes non-negative profits with

$$\bar{p}_j(r_i, r_j, \varepsilon) = \frac{r^{id}d_i(r_i, r_j, \varepsilon)}{d_i(r_i, r_j, \varepsilon) - k_i}.$$

With $p_j = \bar{p}_j(r_i, r_j, \varepsilon) > p_i$ firm j 's optimized overbidding profit is

$$\begin{aligned} \pi_j(p_i, p_j, r_i, r_j, \varepsilon) \Big|_{p_j = \bar{p}_j(r_i, r_j, \varepsilon) > p_i} &= r^{id}d_j(r_i, r_j, \varepsilon) \\ &\quad + \bar{p}_j(r_i, r_j, \varepsilon)[d_i(r_i, r_j, \varepsilon) - k_i] \\ &= r^{id}[d_j(r_i, r_j, \varepsilon) + d_i(r_i, r_j, \varepsilon)] \\ &= r^{id}x^R(r^{id}) \end{aligned}$$

However, firm j only prefers overbidding with $p_j = \bar{p}_j(r_i, r_j, \varepsilon) > p_i$ to underbidding if the optimal overbidding profit is not exceeded by the optimal underbidding profit. If both firms' capacities are needed to satisfy

market demand ($k_j < x^R(r^{id}, \varepsilon)$), firm j 's optimal underbidding price bid from below is any bid with $p_j < p_i$. Otherwise, if $k_i \geq x^R(r^{id}, \varepsilon)$, firm j 's underbidding profit is linearly increasing in p_j , and its optimal undercutting bid is $p_j = p_i - \mu$ with $\mu \rightarrow 0$. The auction price in both cases is $p(p_i, p_j, \varepsilon) = p_i$, such that firm j 's optimized underbidding profit is

$$\begin{aligned} \pi_j(p_i, p_j; r_i, r_j, \varepsilon)|_{p_j < p_i} &= r^{id} d_j(r_i, r_j, \varepsilon) \\ &\quad + p_i (\min\{k_j, x^R(r^{id}, \varepsilon)\} - d_j(r_i, r_j, \varepsilon)). \end{aligned}$$

Firm j prefers overbidding to underbidding or matching firm i 's price bid p_i if and only if

$$\pi_j(p_i, p_j; r_i, r_j, \varepsilon)|_{p_j < p_i} \leq \pi_j(p_i, p_j, r_i, r_j, \varepsilon)|_{p_j = \bar{p}_j(r_i, r_j, \varepsilon) > p_i},$$

which is equivalent to $p_i \leq \hat{p}_j(r_i, r_j, \varepsilon)$, with

$$\hat{p}_j(r_i, r_j, \varepsilon) = \frac{r^{id} d_i(r_i, r_j, \varepsilon)}{\min\{k_j, x^R(r^{id}, \varepsilon)\} - d_j(r_i, r_j, \varepsilon)} < \bar{p}_j(r_i, r_j, \varepsilon).$$

Thus, if $k_i < d_i(r_i, r_j, \varepsilon)$ for only one firm, there are multiple Nash equilibria with $p_i \leq \hat{p}_j(r_i, r_j, \varepsilon)$ and $p_j = \bar{p}_j(r_i, r_j, \varepsilon) > \hat{p}_j(r_i, r_j, \varepsilon)$. For all these Nash equilibria the auction price is unique with $p(p_j, p_h, \varepsilon) = \bar{p}_j(r_j, r_h, \varepsilon)$, and the firms' profits are $\pi_i(r_i, r_j, \varepsilon) = 0$ and $\pi_j(r_i, r_j, \varepsilon) = r^{id} x^R(r^{id}, \varepsilon)$.¹⁰

A.3.2 The Retail Market

The expected profits of the two firms, when they choose their retail prices and anticipate the wholesale price bidding, depends on whether they undercut, match or overcharge their rival's retail price.

Undercutting ($r_i < r_j$): The retail price is $r^{id} = r_i$ and firm i 's retail demand is $d_i(r_i, r_j, \varepsilon) = x^R(r_i, \varepsilon)$. The expected profit net of capacity costs from undercutting is

$$E\pi_i(r_i, r_j) = \begin{cases} \int_{\max\{0, r_i-1\}}^{\min\{1, k_i+r_i-1\}} r_i x(r_i, \varepsilon) d\varepsilon & \text{if } 1 - k_i \leq r_i < r_j, \\ 0 & \text{if } 0 \leq r_i < \min\{1 - k_i, r_j\}. \end{cases} \quad (16)$$

¹⁰The Nash equilibria in wholesale prices are the same as in Boom (2009) with the only exception that, due to the assumed rationing instead of a blackout when demand exceeds capacities, a pure Nash equilibrium in wholesale prices also exists in this case.

Rationing does not affect firm i 's expected undercutting profit $E\pi_i(r_i, r_j)$, because when it undercuts firm j 's retail price, firm i supplies all consumers in the retail market. However, in all demand states in which firm i lacks the capacity to serve its retail demand ($d_i(r_i, r_j, \varepsilon) = x^R(r_i, \varepsilon) > k_i$), firm j exploits this on the wholesale market by setting $p_j = \bar{p}(r_i, r_j, \varepsilon)$, thereby shifting all the rents to firm j . This happens for smaller demand shocks, namely at $\varepsilon = k_i + r_i - 1$, than the one for which the system operator would start to ration demand. The latter happens at $\varepsilon = k_i + k_j + r_i - 1$.

From analysing $E\pi_i(r_i, r_j)$ in (16) one can identify firm i 's optimal undercutting strategy. It is

$$\underline{r}_i(r_j) = \begin{cases} \max\{2 - k_i, \frac{3}{4}\} & \text{if } r_j > \max\{2 - k_i, \frac{3}{4}\}, \\ r_j - \mu & \text{if } 0 < r_j \leq \max\{2 - k_i, \frac{3}{4}\}, \end{cases} \quad (17)$$

with $\mu \rightarrow 0$ being the smallest unit in which retail prices can be announced.

Matching ($r_i = r_j$): The retail price is $r^{id} = r_i = r_j$ and firm i 's retail demand is $d_i(r_i, r_j, \varepsilon) = x^R(r_i, \varepsilon)/2$. In this case, firm i 's expected profit net of capacity costs depends on its capacity relative to its rival, that is, whether $k_i < k_j$, $k_i = k_j$ or $k_i > k_j$ holds.

With $k_i < k_j$, rationing does again not affect firm i 's profit for the reasons discussed above for the undercutting case. Rationing happens only in demand states $\varepsilon > k_i + k_j + r_j - 1$ where firm i 's rents will already be shifted to firm j in the wholesale market, which happens for all $\varepsilon > 2k_i + r_j - 1$. Firm i 's expected profit net of capacity costs is given by

$$E\pi_i \Big|_{\substack{r_i=r_j \\ k_i < k_j}} = \begin{cases} \int_{\max\{0, r_j-1\}}^{\min\{1, 2k_i+r_j-1\}} \frac{r_j x(r_j, \varepsilon)}{2} d\varepsilon & \text{if } r_j \geq 1 - 2k_i, \\ 0 & \text{if } 0 \leq r_j < 1 - 2k_i. \end{cases} \quad (18)$$

With $k_i = k_j$ rationing plays a role because the firm can now continue to sell up to its capacity even if it cannot satisfy its own retail demand. Here neither of the two firms can exploit its rival on the wholesale market, because, when the rival's retail demand exceeds its capacity, this happens for both firms at exactly the same demand shock. The expected profit net of capacity costs is therefore

$$E\pi_i \Big|_{\substack{r_i=r_j \\ k_i=k_j}} = \int_{\max\{0, r_j-1\}}^1 r_j \min\left\{\frac{x(r_j, \varepsilon)}{2}, k_i\right\} d\varepsilon. \quad (19)$$

With $k_i > k_j$, firm i , when matching firm j 's retail price, can appropriate all its rival's rents in the wholesale market for some demand shocks with $\varepsilon >$

$2k_j + r_j - 1$ in which firm j cannot serve its (rationed) retail demand. Due to its larger capacity firm i can always serve its (rationed) retail demand. Its expected profit net of capacity costs is

$$E\pi_i \Big|_{\substack{r_i=r_j \\ k_i > k_j}} = \begin{cases} \int_{\max\{0, r_j-1\}}^1 \frac{r_j x(r_j, \varepsilon)}{2} d\varepsilon & \text{if } r_j \geq 2 - 2k_j, \\ \int_{\max\{0, r_j-1\}}^{2k_j+r_j-1} \frac{r_j x(r_j, \varepsilon)}{2} d\varepsilon + \\ \int_{2k_j+r_j-1}^1 r_j x^R(r_j, \varepsilon) d\varepsilon & \text{if } 1 - 2k_j \leq r_j \\ & < 2 - 2k_j, \\ \int_0^1 r_j x^R(r_j, \varepsilon) d\varepsilon & \text{if } 0 \leq r_j \\ & < 1 - 2k_j. \end{cases} \quad (20)$$

Overcharging ($r_i > r_j$): The retail price is $r^{id} = r_j$ and firm i 's retail demand is $d_i(r_i, r_j, \varepsilon) = 0$. Here firm i can only earn positive revenues in those demand states ($\varepsilon > r_j - 1 + k_j$) in which its rival cannot serve the (rationed) retail demand of its retail customers, such that firm i appropriates all its rival's rents via the wholesale auction. Thus, the expected profit net of capacity costs from overcharging is:

$$E\bar{\pi}_i(r_i, r_j) = \begin{cases} 0 & \text{if } r_i > r_j \geq 2 - k_j, \\ \int_{\max\{0, r_j-1+k_j\}}^1 r_j x^R(r_j, \varepsilon) d\varepsilon & \text{if } 0 \leq r_j < \\ & \min\{r_i, 2 - k_j\}. \end{cases} \quad (21)$$

If firm i overcharges firm j , then it is indifferent between all retail prices satisfying $r_i > r_j$, because its profit does obviously not depend on the level of r_i .

Firm i 's overall best response in retail prices is determined by comparing the optimized undercutting expected profit $E\pi_i(\underline{r}_i(r_j), r_j)$, derived from (16) and (17), with the relevant expected profit from matching $E\pi_i(r_i, r_j) \Big|_{r_i=r_j}$, given either in (18), (19) or (20), and with the expected profit from overcharging $E\bar{\pi}_i(r_i, r_j)$ defined in (21). We need to distinguish again different cases depending on the capacities k_i and k_j of the two firms.

(i) Retail Market Competition with Both Firms Having Large Capacities

If both firms have relatively large capacities, $k_i \geq \sqrt{\frac{5}{2}}$ and $k_i \geq k_j > \sqrt{5 - k_i^2}$, each firm i 's best response in retail prices is $r_i(r_j) = \underline{r}_i(r_j)$ from (17), such that they always want to undercut each other. The unique Nash Equilibrium in retail prices is then

$$r_i = r_j = 0 = r^{id}.$$

(ii) *Retail Market Competition with Asymmetric, but not Large Capacities*

Here we assume throughout that firm i is the large and firm j is the small firm with $\min\{k_i, \sqrt{5 - k_i^2}\} > k_j$. We start with identifying the larger firm's best response function and can show that, depending on the smaller firm's retail price, the larger firm usually switches from undercutting to overcharging and then becomes indifferent between overcharging and matching when the small firm's retail price becomes smaller.

There is, however, one exception. When the two firms' capacities are extremely asymmetric, $0 < k_j < \min\{(k_i - 1)/2, 1/8\}$, the larger firm i 's best response in retail prices is

$$r_i(r_j) \begin{cases} = \max\{2 - k_i, \frac{3}{4}\} & \text{if } r_j > \max\{2 - k_i, \frac{3}{4}\}, \\ \geq r_j & \text{if } 0 \leq r_j \leq \max\{2 - k_i, \frac{3}{4}\}. \end{cases} \quad (22)$$

It switches directly from undercutting (with the profit-maximizing monopoly price) to becoming indifferent between overcharging and matching, because the smaller firm's capacity is so small that it can never (even for the lowest demand with $\varepsilon = 0$) serve its own retail demand, if it sets its retail price r_i , such that $r_j \leq r_i \leq \max\{2 - k_i, \frac{3}{4}\}$ holds. For all retail prices up to the monopoly price the large firm realizes all the revenues, whereas the small firm generates always zero profits.

However, the usual pattern already emerges for still very asymmetric capacities with $\min\{(k_i - 1)/2, 1/8\} \leq k_j < \min\{k_i - 1, \frac{1}{4}\}$. Here the large firm i 's best response is

$$r_i(r_j) \begin{cases} = \max\{2 - k_i, \frac{3}{4}\} & \text{if } r_j > \max\{2 - k_i, \frac{3}{4}\}, \\ > r_j & \text{if } 1 - 2k_j < r_j \leq \max\{2 - k_i, \frac{3}{4}\} \\ \geq r_j & \text{if } 0 \leq r_j \leq 1 - 2k_j. \end{cases} \quad (23)$$

The smaller firm can serve its own retail demand when matching the large firm's monopoly price at the lowest retail demand with $\varepsilon = 0$, but cannot serve it when undercutting the large firm's monopoly price with $r_j \leq r_i \leq \max\{2 - k_i, \frac{3}{4}\}$. Therefore, the large firm switches directly from undercutting with its monopoly price to overcharging if the smaller firm's retail price decreases, starting from firm i 's monopoly price. The large firm becomes indifferent between overcharging and matching at a lower retail price of the low capacity firm, when the low capacity firm can no longer serve its retail demand with matched retail prices even for the smallest demand $\varepsilon = 0$.

	$0 \leq k_i \leq 1$	$1 < k_i < 2$
$0 \leq k_j \leq \max \left\{ \min \{ \sqrt{1 - k_i^2}, k_i \}, k_i - 1 \right\}$	$2 - \sqrt{k_j^2 + k_i^2}$	$1 - k_j$
$\max \{ \sqrt{1 - k_i^2}, k_i - 1 \} \leq k_j \leq k_i$	$\frac{3 - \sqrt{2(k_j^2 + k_i^2) - 1}}{2}$	

Table 1: Definition of \hat{r} for Different k_i and k_j

Finally, for intermediately asymmetric capacities with $\min\{k_i - 1, \frac{1}{4}\} \leq k_j < \min\{k_i, \sqrt{5 - k_i^2}\}$ the larger firm's overall best response is

$$r_i(r_j) \begin{cases} = \underline{r}_i(r_j) & \text{if } \hat{r} < r_j, \\ > r_j & \text{if } \max\{0, 1 - 2k_j\} \leq r_j \leq \hat{r}, \\ \geq r_j & \text{if } 0 < r_j \leq 1 - 2k_j. \end{cases} \quad (24)$$

Note that the rival's retail price $r_j = \hat{r}$, where the large firm switches from undercutting to overcharging, depends on the capacities of both firms (see Table 1 for the exact definitions) and is lower than the large firm's monopoly price.

To derive a Nash equilibrium in retail prices, we now need to consider the smaller firm j 's best response in retail prices. We can distinguish two cases. For not too asymmetric capacities with $\min\{k_i, \sqrt{5 - k_i^2}\} > k_j \geq \max\{0, \frac{k_i - 1}{2}\}$, the smaller firm's overall best response in retail prices is given by

$$r_j(r_i) \begin{cases} = \underline{r}_j(r_i) & \text{if } r_i > r'_i, \\ = r_i & \text{if } r'_i \geq r_i > r''_i, \\ > r_i & \text{if } r''_i \geq r_i > 0, \\ \geq 0 & \text{if } r_i = 0, \end{cases} \quad (25)$$

where the critical prices

$$r'_i = \begin{cases} 2 - \sqrt{2}k_j & \text{if } 0 \leq k_j < \min\left\{\sqrt{\frac{1}{2}}, k_i\right\}, \\ \frac{3 - \sqrt{4k_j^2 - 1}}{2} & \text{if } \sqrt{\frac{1}{2}} \leq k_j < k_i, \end{cases} \quad (26)$$

	$0 \leq k_i < \frac{1}{\sqrt{2}}$	$\frac{1}{\sqrt{2}} \leq k_i < 1$	$1 \leq k_i < \sqrt{\frac{5}{2}}$	$\sqrt{\frac{5}{2}} \leq k_i \leq 2$
$0 \leq k_j < \min \left\{ \frac{k_i}{\sqrt{2}}, \sqrt{\frac{1-k_i^2}{2}} \right\}$	$2 - \sqrt{2k_j^2 + k_i^2}$		-	-
$\min \left\{ \sqrt{\frac{1-k_i^2}{2}}, \frac{k_i-1}{2} \right\} \leq k_j$ $< \min \left\{ \frac{\sqrt{2k_i^2 + \sqrt{4k_i^2 - 1}}}{2}, \frac{\sqrt{9-2k_i^2}}{2} \right\}$	-	$\frac{5 - \sqrt{12k_j^2 + 6k_i^2 - 2}}{3}$		
$\min \left\{ \frac{k_i}{\sqrt{2}}, \frac{\sqrt{2k_i^2 + \sqrt{4k_i^2 - 1}}}{2}, \frac{\sqrt{9-2k_i^2}}{2} \right\} \leq k_j < k_i$	$2 - \sqrt{2}k_i$	$\frac{3 - \sqrt{4k_i^2 - 1}}{2}$	0	

Table 2: Definition of r_i'' for Different k_i and k_j

and r_i'' for the rival with the larger capacity again depend on the capacities of the two firms (see Table 2 for the exact definition of r_i'').¹¹ For a sufficiently high retail price of its larger rival i the smaller firm j switches from undercutting the larger firm's retail price r_i to matching and then to overbidding it, if the larger firm reduces its retail price r_i .

However, for $\frac{k_i-1}{2} > k_j > 0$ the small firm j 's best response in retail prices changes to

$$r_j(r_i) \begin{cases} = \underline{r}_j(r_i) & \text{if } r_i > r_i', \\ = r_i & \text{if } r_i' \geq r_i \geq 1 - 2k_j, \\ \geq 0 & \text{if } 1 - 2k_j > r_i \geq 2 - k_i, \\ > r_i & \text{if } 2 - k_i > r_i > 0, \\ \geq 0 & \text{if } r_i = 0, \end{cases} \quad (27)$$

where $\underline{r}_j(r_i)$ is still firm j 's optimal undercutting strategy (the equivalent to (17))

¹¹The cells in Table 2 with entries “-” do not exist, that is, there are no capacities $k_i, k_j \in [0, 2]$ which fulfill simultaneously the conditions given in the vertical column and the horizontal row.

and where r'_i is still defined in (26). Different from the more symmetric case, the smaller firm here no longer directly switches from matching to overbidding when its larger rival's retail price decreases. Due to the very asymmetric capacities the smaller firm here is indifferent between matching, overbidding and undercutting for a particular retail price range of its rival. In this retail price range the smaller firm is neither able to ever serve its own retail demand while undercutting or only matching its larger rival (even with the smallest demand, $\varepsilon = 0$), nor can it ever acquire its larger rival's revenues in the wholesale market when overbidding the larger rival in the retail market (even with the largest demand $\varepsilon = 1$), because its larger rival is always able to serve its demand.

Now consider first the not too asymmetric case with $\min\{k_i, \sqrt{5 - k_i^2}\} > k_j > \frac{k_i - 1}{2}$. For this case one can prove that $r'_j > \min\{\hat{r}, \max\{2 - k_i, \frac{3}{4}\}\} > r''_j > 1 - 2k_j$, where r' is defined in equation (26), \hat{r} in Table 1 and r'' in Table 2. This means that for all retail prices where the smaller firm wants to undercut the larger firm, the larger firm wants to do the same. For all the retail prices where the smaller firm wants to match the larger firm, the larger firm either wants to undercut or to overcharge, but never wants to match, and, finally, for all retail prices where the smaller firm wants to overcharge the larger, the larger also wants to overcharge or is indifferent between matching and overcharging. Therefore, like in a usual Bertrand competition, the unique equilibrium in retail prices is given by $r_i = r_j = 0 = r^{id}$ for all $k_i > k_j > \frac{k_i - 1}{2}$. Depending on the exact level of capacities rationing might or might not take place in equilibrium.

However, this changes for extremely asymmetric capacities with $\frac{k_i - 1}{2} > k_j \geq 0$. Analyzing (27) together with the large rival's counterpart, which is either (22) or (23), reveals that for these extremely asymmetric capacities additional equilibria exist, in which the large firm i earns positive profits and the small firm j earns nothing. One can show that for all $\frac{k_i - 1}{2} > k_j \geq 0$ a whole range of retail price equilibria exist in which the small firm j undercuts or matches the large firm i 's retail price r_i with $\max\{r_j, 2 - k_i\} \leq r_i \leq \min\{1 - 2k_j, 3/4\}$ without being able to ever cover its own retail demand with its capacity. The retail price is then r_j . The small firm is indifferent between all these equilibria, whereas the large firm obviously prefers $r_i = r_j = 2 - k_i$ for $k_i \leq 5/4$ and $r_i = r_j = \min\{3/4, 1 - 2k_j\}$ for $k_i > 5/4$. For even more asymmetric capacities with $0 \leq k_j < \min\{(k_i - 1)/2, 1/8\}$ another retail price equilibrium occurs where the small firm j overcharges the large firm i with $r_i = \max\{2 - k_i, 3/4\} < r_j$. The retail price would be r_i .

Thus, there is a unique Pareto dominant retail price equilibrium for extremely

asymmetric capacities, $\frac{k_i-1}{2} > k_j \geq 0$, with

$$r_i = \begin{cases} 2 - k_i & \text{if } k_i \leq \frac{5}{4} \\ \frac{3}{4} & \text{if } k_i > \frac{5}{4} \end{cases} \text{ and } r_j \geq r_i \text{ if } k_j < \frac{1}{8} \text{ also holds,} \quad (28)$$

$$\text{and } r_i = 1 - 2k_j = r_j \text{ if } k_j \geq \frac{1}{8}. \quad (29)$$

Rationing never takes place in this Pareto dominant equilibrium.

(iii) *Retail Market Competition with Symmetric, but not Large Capacities*

Here we assume that $k_i = k_j = k < \sqrt{\frac{5}{2}}$ holds. Each firm has the same best response in retail prices, given by:

$$r_i(r_j) = \begin{cases} = \underline{r}_i(r_j) & \text{if } r_j > \hat{r}, \\ \geq r_j & \text{if } r_j = \hat{r}, \\ > r_j & \text{if } \hat{r} > r_h > 0, \\ \geq 0 & \text{if } r_j = 0, \end{cases} \quad (30)$$

with \hat{r} defined in Table 1, given $k_i = k_j = k$. For all rival's prices above $r_j = \hat{r}$, each firm wants to undercut, at the rival's price $r_j = \hat{r}$ each firm is indifferent between overcharging and matching the rival's price, whereas for all rival's retail prices below, with $\hat{r} > r_j > 0$, each firm wants to overcharge, and finally at $r_j = 0$ each firm is indifferent between overcharging and matching. Thus, obviously the Nash equilibrium with $r_i = r_j = 0$ and zero profits always exists.

For $k \leq \sqrt{\frac{5}{2}}$ there is, however, also another equilibrium with

$$r_i = r_j = \hat{r}|_{k_i=k_j=k} = \begin{cases} 2 - \sqrt{2}k & \text{if } 0 \leq k < \sqrt{\frac{1}{2}}, \\ \frac{1}{2} (3 - \sqrt{4k^2 - 1}) & \text{if } \sqrt{\frac{1}{2}} \leq k < \sqrt{\frac{5}{2}}. \end{cases} \quad (31)$$

In this retail price equilibrium rationing is never necessary because $2 - \hat{r} < 2k$ always holds. Since the expected profits of the two firms with $r_i = r_j = \hat{r}|_{k_i=k_j=k}$ is positive, this is the unique Pareto-dominant Nash equilibrium in retail price for $k=k_j = k < \sqrt{\frac{5}{2}}$.

The following Lemma summarizes all our results with respect to the retail prices.

Lemma 1 *Assuming that the two firms always act according to the Pareto optimal Nash equilibrium in case of multiple Nash equilibria in retail prices, the equilibrium in retail prices, anticipating the outcomes for different demand states on the wholesale market, depends on how large and how asymmetric the integrated firms' $i, j \in \{A, B\}$ and $i \neq j$, generating capacities are.*

- (i) *With intermediately asymmetric capacities, $k_i > k_j \geq \frac{k_i-1}{2}$, or with large symmetric capacities, $k_i = k_j = k \geq \sqrt{\frac{5}{2}}$, firms choose $r_i = r_j = 0 = r^{id}$ in equilibrium and realize an expected profit of zero, net of their capacity costs. Rationing may be necessary in equilibrium.*
- (ii) *If the firms' capacities are more extremely asymmetric, $\frac{k_i-1}{2} > k_j \geq 0$, they choose the retail prices, depending on the exact capacities, defined either in equation (28) or in (29), where the retail price paid by the consumers is in both cases $r^{id} = r_i$, the respective retail price demanded by the larger firm. The small firm's j expected profit net of capacity costs is always zero, whereas the large firm's i net expected profit is positive. Rationing is never necessary in this equilibrium.*
- (iii) *For not too large symmetric capacities with $k_i = k_j = k < \sqrt{\frac{5}{2}}$ both firms choose $r_i = r_j = \hat{r}|_{k_i=k_j}$ given in (31), and consumers pay $r^{id} = \hat{r}|_{k_i=k_j}$. Both firms realize positive expected profits net of their capacity costs and rationing is never necessary in equilibrium.*

A.3.3 The Integrated Firms' Investment Decision

When deciding on their electricity generating capacities, integrated firms anticipate the consequences of their capacity choices on the wholesale prices for the different demand scenarios and on retail price competition.

Let us assume for now that they stick to their original capacity investments k_i and k_j and do not mothball any capacities. Then, their expected profit $\Pi_i(k_i, k_j)$ is characterized in Table 3.¹² Since demand in our model is never larger than 2 we assume that $k_i, k_j \in [0, 2]$.

The capacity decisions of the two firms are determined in three steps. First both firms invest in capacity k_i and k_j and bear the respective capacity costs. Then the smaller firm, let us suppose it is firm i , due to $k_i \leq k_j$ after the investment

¹²As before in Table 2, a cell with entry “-” indicates that no combination of k_i and k_j exists which satisfies simultaneously the conditions given in both the description of the row and of the column.

	$0 \leq k_j < \frac{1}{8}$	$\frac{1}{8} \leq k_j < \frac{1}{2}$	$\frac{1}{2} \leq k_j < \sqrt{\frac{1}{2}}$	$\sqrt{\frac{1}{2}} \leq k_j < \sqrt{\frac{5}{2}}$	$\sqrt{\frac{5}{2}} \leq k_j \leq 2$
$k_i < k_j$	$-zk_i$				
$k_i = k_j$	$k_i^2 \left(1 - \frac{k_i}{\sqrt{2}}\right) - zk_i$			$\frac{1}{8} (1 - 4k_i^2 + 3\sqrt{4k_i^2 - 1}) - zk_i$	$-zk_i$
$k_j < k_i \leq 2k_j + 1$	$-zk_i$				
$2k_j + 1 < k_i \leq \frac{5}{4}$	$(2 - k_i) \times (k_i - \frac{1}{2}) - zk_i$	-			
$\max\{\frac{5}{4}, 2k_j + 1\} < k_i \leq 2$	$\frac{9}{16} - zk_i$	$(1 - 2k_j) \times (\frac{1}{2} + 2k_j) - zk_i$	-		

Table 3: Firm i 's expected Profit $\Pi_i(k_i, k_j)$ in Terms of k_i and k_j

decisions, has the opportunity to costlessly mothball part of its capacity, such that it has only $\hat{k}_i \leq k_i < k_j$ left to bid later into the wholesale market. Finally, the originally larger firm j can also reduce its capacity to $\hat{k}_j \leq k_j$.

- (i) Let us first consider the final step, the choice of firm j with $k_j \geq k_i \geq \hat{k}_i$. Investment costs are sunk at this stage. Therefore we analyze $\Pi_j(\hat{k}_i, \hat{k}_j) + zk_j$, implicitly characterized in Table 3, in order to derive firm j 's best response to firm i 's potentially adapted capacity $\hat{k}_i \leq k_i$ which depends also on its own originally invested capacity k_j . It is given by

$$\hat{k}_j(\hat{k}_i, k_j) = \begin{cases} k_j & \text{if } 0 \leq \hat{k}_i < \min\{0.45063, \frac{k_j-1}{2}\}, \\ \hat{k}_i & \text{if } \min\{0.45063, \frac{k_j-1}{2}\} \leq \hat{k}_i < \sqrt{\frac{5}{2}}, \\ k_j & \text{if } \sqrt{\frac{5}{2}} \leq \hat{k}_i \leq 2, \end{cases} \quad (32)$$

and implies that the larger firm j either reduces its capacity and matches the smaller firm's capacity \hat{k}_i , resulting in both firms having the same positive profit, or it sticks to its original investment k_j . The latter is beneficial if firm i has either such a small capacity \hat{k}_i that firm j can monopolize the market (with profits described in the south west of table 3), or if firm i has chosen such a large capacity $\hat{k}_i \geq \sqrt{\frac{5}{2}}$ that neither of the two firms can earn any profits.

- (ii) In the intermediate step firm i anticipates the best response $\hat{k}_j(\hat{k}_i, k_j)$ of firm j and knows therefore that it can never become the large supplier by mothballing capacity and that it can only generate positive profits if firm j later matches \hat{k}_i in the final step. Firm i chooses \hat{k}_i such that it maximizes $\Pi_i(\hat{k}_i, \hat{k}_j(\hat{k}_i, k_j)) + zk_i$ for $\hat{k}_i \leq k_i \leq k_j$. Its best response is

$$\hat{k}_i(k_i, k_j) = \begin{cases} k_i & \text{if } 0 \leq k_i \leq \min\{\frac{\sqrt{13}}{4}, k_j\} \\ \frac{\sqrt{13}}{4} & \text{if } \frac{\sqrt{13}}{4} \leq k_i \leq k_j. \end{cases}$$

- (iii) In the first step neither of the two firms has an incentive to invest in a level of capacity such that it later is going to mothball it in either stage 2 or stage 3 of the process. Therefore only two different types of equilibria are possible. Either one firm monopolizes the market with $k_j = \frac{5}{4} - \frac{z}{2}$ and the other firm does not invest at all, $k_i = 0$, or both firms match each others capacities with $0 \leq k_i = k_j \leq \frac{\sqrt{13}}{4}$.

Let us first consider the monopolization equilibrium. This can only exist as long as the small firm i which earns zero profits in this case cannot deviate

to a positive capacity $\frac{k_j-1}{2} = \frac{1}{8} - \frac{z}{4} < k_i \leq \min\{\frac{\sqrt{13}}{4}, \frac{5}{4} - \frac{z}{2}\}$, which would later, in step 3, induce the large firm to reduce its capacity to the level of firm i , such that firm i generates a positive profit with matched capacities. One can show that such a profitable deviation exists for firm i , as long as $0 \leq z < \frac{1}{2\sqrt{2}}$. In addition the monopolizing firm should not realize zero profits. From Table 3 one can see that for $(k_j, k_i) = (\frac{5}{4} - \frac{z}{2}, 0)$ firm j 's expected profit is

$$\Pi_j \left(\frac{5}{4} - \frac{z}{2}, 0 \right) = \frac{1}{16}(4(z-5)z + 9) \geq 0 \Leftrightarrow z \leq \frac{1}{2}.$$

Thus the monopolization equilibrium exists only for $\frac{1}{2\sqrt{2}} \leq z \leq \frac{1}{2}$ and not for the range of capacity costs assumed here with $0 \leq z < \frac{1}{2\sqrt{2}}$.

Now consider a potential matching equilibrium with $0 \leq k_i = k_j$. Note that matching equilibria cannot exist for $k_i = k_j > \frac{\sqrt{13}}{4}$ because then the firm, which is allowed to mothball part of its capacity first, always can do better by reducing k_i to $\hat{k}_i = \frac{\sqrt{13}}{4}$. This capacity level is then matched by the other firm j with $\hat{k}_j = \frac{\sqrt{13}}{4}$. In addition there should be neither an incentive to choose ex ante a capacity of zero in order to avoid negative profits in a matching equilibrium, nor should there be an incentive to ex ante invest in a capacity which would allow the respective firm to monopolize the market.

A matching equilibrium yields negative profits if both firms invest in the same level of capacity $k_i = k_j$ below \underline{k} which is the smallest capacity that, when matching it, just ensures zero profit with

$$\underline{k} \equiv \frac{1}{\sqrt{2}} \left(1 - \sqrt{1 - 2z\sqrt{2}} \right). \quad (33)$$

or if both firms invest in the same level of capacity $k_i = k_j$ above \bar{k} with

$$\bar{k} \equiv \left\{ k \in \left[\frac{1}{\sqrt{2}}, \sqrt{\frac{5}{2}} \right] \mid \frac{1 - 4k^2 + 3\sqrt{4k^2 - 1}}{8} - zk = 0 \right\}. \quad (34)$$

Note that $\underline{k} \leq \bar{k}$ for all $0 \leq z \leq 1/(2\sqrt{2})$. In addition, one can show that $\bar{k} > \frac{\sqrt{13}}{4}$ for all $0 \leq z < 9/(8\sqrt{13}) \approx 0.312$, such that \bar{k} is only relevant for $9/(8\sqrt{13}) \leq z \leq 1/(2\sqrt{2}) \approx 0.35355$ as a limit for how much integrated firms could potentially invest in a matching equilibrium.

The best deviating monopolization strategy from $k_i = k_j$ is to choose ex ante $k_j = \min\{2k_i + 1, 5/4 - z/2\}$. However, this is only more profitable

than matching the competitor's capacity k_i as long as $k_i \leq \tilde{k}$ with

$$\tilde{k} \equiv \left\{ k \in \left[0, \frac{1}{2} \right] \mid k^2 \left(1 - \frac{k}{\sqrt{2}} \right) - zk = \right. \quad (35)$$

$$\left. (1 - 2k) \left(\frac{1}{2} + 2k \right) - z(2k + 1) \right\}.$$

Note that $\tilde{k} < \min\{\bar{k}, \sqrt{13}/4\}$ for all $0 \leq z \leq 1/(2\sqrt{2})$ and that $\tilde{k} \leq \underline{k}$ for all $0.2484 \leq z \leq 1/(2\sqrt{2})$.

Thus, we can prove the following Lemma.

Lemma 2 *For $0 \leq z < 1/(2\sqrt{2})$ multiple subgame perfect equilibria exist. All these equilibria are characterized by both firms $i, j \in \{A, B\}$ choosing $k_i = k_j = k$ with $\max\{\tilde{k}, \underline{k}\} \leq k < \min\{\sqrt{13}/4, \bar{k}\}$. Rationing is never necessary in all these subgame perfect equilibria.*

Since we focus here on Pareto dominant subgame perfect Nash equilibria we can, of course, identify the equilibrium $k_i = k_j = k^*$ among the multiple matching equilibria with $0 \leq z < 1/(2\sqrt{2})$ which maximizes each firm's profit. It is given by (9) in proposition 3.

Corollary 1 *For $0 \leq z < \frac{1}{2\sqrt{2}}$ the unique Pareto dominant subgame perfect Nash equilibrium is that the two firms choose $k_A = k_B = k^*$ defined in (9). The resulting retail price is*

$$r^* = \frac{1}{2} \left(3 - \sqrt{4(k^*)^2 - 1} \right) \quad (36)$$

and rationing is never necessary in this equilibrium.

Note that the total capacity k^{id} coincides with $2k^*$ from equation (9) in Proposition 3, and that r^{id} is equivalent to r^* from equation (36) after substituting k^* . Explicitly solving equation (9) for k^* , taking into account that $k^{id} = 2k^*$ and substituting

k^* into (36) yields

$$r^{id} = \frac{3}{2} - \frac{1}{2} \left\{ \left[\frac{\sqrt{3}\sqrt{26 + 16z^2 + m(z)}}{12} - z + \left(\frac{13}{12} - \frac{2z^2}{3} - \frac{m(z)}{48} - \frac{5\sqrt{3}z}{2\sqrt{26 + 16z^2 + m(z)}} \right)^2 - 1 \right]^{\frac{1}{2}} \right\} \text{ and}$$

$$k^{id} = \left[\frac{13}{12} + \frac{2z^2}{3} - \frac{m(z)}{48} - \frac{5\sqrt{3}z}{2\sqrt{26 + 16z^2 + m(z)}} \right]^{\frac{1}{2}} - z + \frac{\sqrt{3}\sqrt{26 + 16z^2 + m(z)}}{12},$$

with $m(z) \equiv \frac{(13 - 16z^2)^2}{n(z)} + n(z)$,

$$n(z) \equiv \sqrt[3]{-2197 + 144q(z)z - 22992z^2 - 9984z^4 + 4096z^6},$$

and $q(z) \equiv \sqrt{3}\sqrt{2197 + 7440z^2 + 9984z^4 - 4096z^6}.$

A.4 Proposition 4 (Separated Duopoly With Rationing)

A.4.1 The Wholesale Market

Analyzing each firm's profit function $\pi_i(p_A, p_B)$ in equation (10) reveals that wholesale bidding game is the same as the one analyzed in appendix A of Le Coq (2002), or in the proofs of Proposition 1-3 in Crampes and Creti (2005), taking into account that marginal generating costs are constant and normalized to zero by assumption and that the maximum wholesale price with positive demand is $p = r^{sd}$. We summarize their results with regard to the Nash equilibria in the wholesale market in Lemma 3.

Lemma 3 (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 $(p_i, p_j) = (r^{sd}, p_j)$ with $p_j \leq r^{sd}$, $i, j \in \{A, B\}$ and $i \neq j$ forms a Nash equilibrium in price bids. The resulting equilibrium wholesale price is $p^{sd} = r^{sd}$, and each firm sells its total capacity, meaning $y_i(r^{sd}, p_j) = k_i$ and $y_j(r^{sd}, p_j) = k_j$.*
- (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 $y_A = y_B = x(r^{sd}, \varepsilon)/2$.

The retail demand is only rationed in case (i) where $x_R(r^{sd}, \varepsilon) = k_A + k_B$ holds.

The uniform price auction has multiple equilibria in all the cases (i)-(iii). These are pay-off equivalent in case (i) and (ii), but not in case (iii), where each firm prefers to be the low price bidder instead of the high price bidder, with $p_i = r^{sd}$ because both firms get the same uniform price with $p^{sd} = r^{sd}$, but only the low bidder sells its total capacity. In order to be able to derive the subgame perfect equilibrium of the whole game via backward induction we need an assumption on which equilibrium the two generators coordinate. Here we follow Fabra *et al.* (2011) and apply

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.*

Assumption 1 can also be motivated by applying risk-dominance as a selection criterion to a very similar game.^{13 14}

¹³The game is the same as ours in this stage here, but the price bids can only be made as a multiple of the smallest monetary unit γ which is small but not zero. See Boom (2008) for a detailed discussion of how risk dominance can be applied in such a game.

¹⁴In an empirical study of the New York capacity market which is organized as a uniform price multi unit auction Schwenen (2015) finds that in all but one of the 55 analyzed auctions, it was always the firm with the largest capacity which bid the highest price among all accepted bids.

A.4.2 The Retail Market

For retailers to obtain non-negative profits, the realized 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 such that demand is sufficiently small to avoid a wholesale price $p^{sd} = r^{sd}$. If $p^{sd} = r^{sd}$ holds, generators extract all the rents from the retailers (areas \mathcal{A} , \mathcal{B} , \mathcal{C} and \mathcal{D} in Figure 2).

Retailers compete à la Bertrand and the expected profits of retailer $\ell = C, D$ is given by

$$\pi_\ell(r_\ell, r_\iota) = \begin{cases} 0 & \text{if } r_\ell > r_\iota, \\ \frac{1}{2} \int_{\max\{0, \underline{\varepsilon}\}}^{\max\{0, \min\{\bar{\varepsilon}, 1\}\}} r_\ell(1 + \varepsilon - r_\ell) d\varepsilon & \text{if } r_\ell = r_\iota, \\ \int_{\max\{0, \underline{\varepsilon}\}}^{\max\{0, \min\{\bar{\varepsilon}, 1\}\}} r_\ell(1 + \varepsilon - r_\ell) d\varepsilon & \text{if } r_\ell < r_\iota, \end{cases} \quad (37)$$

with $\ell, \iota = C, D$, and $\ell \neq \iota$. Equation (37) indicates that retailers undercut each other until they reach zero profits. Therefore, the following Nash equilibrium or equilibria in retail prices emerge(s).

Lemma 4 (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 $r_D \leq 1 - \min\{k_A, k_B\}$.*

Its proof is equivalent to a standard proof for a Bertrand equilibrium.¹⁵ In case (i) retail competition results in a unique equilibrium, whereas in case (ii) we have again multiple equilibria.

In case (i) the generators' minimum capacity is too large and induces positive profits for the retailers for some small demand realizations if $r \geq 0$ holds. Therefore, retailers compete each other down to $r = 0$ in a Bertrand type manner.

¹⁵Suppose that $r_\ell > r_\iota$ with $\ell, \iota = C, D$ and $\ell \neq \iota$. This can only be an equilibrium if $r_\ell \leq 1 - \min\{k_A, k_B\}$, because otherwise firm ℓ could increase its profits by undercutting and firm ι by increasing its price. Suppose, alternatively, that $r_\ell = r_\iota$. Then either $r_\ell = r_\iota = 0$, if $\min\{k_A, k_B\} \geq 1$ holds, or $r_\ell = r_\iota \leq 1 - \min\{k_A, k_B\}$, if $\min\{k_A, k_B\} < 1$ holds. Otherwise each retailer could double its profit by slightly undercutting its rival.

In case (ii) the minimum capacity of the generators is small enough for all rents to be shifted from retailers to generators at some positive retail prices, even if the demand realization is very small ($\varepsilon = 0$). We can deal with the multiplicity problem in case (ii), if we apply Pareto dominance as an equilibrium selection criterion.

Corollary 2 *If $\min\{k_A, k_B\} < 1$ holds, the unique Pareto dominant Nash equilibrium is*

$$r_C = r_D = 1 - \min\{k_A, k_B\}. \quad (38)$$

Corollary 2 means that retailers select the equilibrium strategy in which they choose the highest possible price which generates zero profits. The equilibrium is Pareto dominant, although the retailers C and D in the retailing stage are indifferent between all prices described in case (ii). The generators A and B , however, clearly prefer the retailers to choose the retail price identified in Corollary 2.

A.4.3 The Investment Decisions of the Separated Generators

Anticipating the outcome of the wholesale market price bidding and the retailers' Bertrand competition on the retail market, the generator i 's profit in terms of capacities with $i, j = A, B$ and $i \neq j$, given that none of the two generators mothballs any capacity, is

$$\Pi_i(k_i, k_j) = \begin{cases} \max\{0, 1 - k_j\} \int_0^1 \min\{\varepsilon, k_i\} d\varepsilon - zk_i & \text{if } k_i > k_j, \\ \frac{\max\{0, 1 - k_j\}}{2} \left[\int_0^1 \min\{\varepsilon, k_i\} d\varepsilon + \int_0^1 k_i d\varepsilon \right] & \text{if } k_i = k_j, \\ -zk_i & \text{if } k_i < k_j, \\ \max\{0, 1 - k_j\} \int_0^1 k_i d\varepsilon - zk_i & \text{if } k_i < k_j. \end{cases} \quad (39)$$

Let us assume that firm i has originally invested more than firm j in the first stage of the three-stage investment process, meaning $k_i \leq k_j$ and consider the final stage of the investment process first.

- (i) In the final third stage of the investment process firm j decides on whether to mothball part of its capacity, after potential mothballing by firm i in the second stage. Thus, firm i 's capacity is \hat{k}_i with $\hat{k}_i \leq k_i \leq k_j$. In order to decide on whether to reduce its current capacity k_j to $\hat{k}_j \leq k_j$, firm j

maximizes $\hat{\Pi}_j = \Pi_j(\hat{k}_i, \hat{k}_j) + zk_j$, given by

$$\hat{\Pi}_j = \begin{cases} 0, & \text{if } 1 < \hat{k}_i \text{ and } 1 < \hat{k}_j, \\ \frac{1-\hat{k}_i}{2}, & \text{if } \hat{k}_i \leq 1 < \hat{k}_j, \\ (1-\hat{k}_i) \left[(1-\hat{k}_j)\hat{k}_j + \frac{\hat{k}_j^2}{2} \right], & \text{if } \hat{k}_i < \hat{k}_j \leq 1, \\ \frac{1-\hat{k}_i}{2} \left[(1-\hat{k}_j)\hat{k}_j + \frac{\hat{k}_j^2}{2} + \hat{k}_j \right], & \text{if } \hat{k}_i = \hat{k}_j \leq 1, \\ (1-\hat{k}_j)\hat{k}_j, & \text{if } \hat{k}_j < \hat{k}_i \text{ and } \hat{k}_j \leq 1. \end{cases} \quad (40)$$

Maximizing this profit function with respect to firm j 's capacity \hat{k}_j yields firm j 's best response in stage 3 of the investment process for a given capacity \hat{k}_i of the originally smaller firm i and for a given original capacity k_j of firm j . It is either

$$\hat{k}_j(\hat{k}_i, k_j) = \begin{cases} \hat{k}_i - \alpha \text{ with } \alpha \rightarrow 0 & \text{if } k_j < 1 - \sqrt{1 - 2\hat{k}_i} \\ \min\{k_j, 1\} & \text{if } k_j \geq 1 - \sqrt{1 - 2\hat{k}_i}, \end{cases} \quad (41)$$

and if $\hat{k}_i \leq \frac{1}{2}$, or it is

$$\hat{k}_j(\hat{k}_i, k_j) = \frac{1}{2}, \text{ if } \hat{k}_i > \frac{1}{2}. \quad (42)$$

If the originally smaller firm's capacity after stage 2 is small, then firm j prefers to become the small capacity provider via sufficient mothballing, if its own original capacity is also relatively small. Otherwise it stays larger. If the originally smaller firm's capacity after stage 2 of the investment process is larger, firm j always mothballs its large capacity such that it becomes the smaller firm.

- (ii) In stage 2 the originally smaller firm i with $k_i \leq k_j$ anticipates firm j 's best response in stage 3, when it itself can reduce its capacity k_i to $\hat{k}_i \leq k_i$. It knows that if $k_j \geq k^+(\hat{k}_i)$ holds with

$$k^+(\hat{k}_i) \equiv 1 - \sqrt{1 - 2\hat{k}_i}, \text{ and } \hat{k}_i \leq \frac{1}{2}$$

then firm j does not reduce its capacity k_j in stage 3 below the level of \hat{k}_i , but sticks to $\hat{k}_j = k_j$. If either $k_j < k^+(\hat{k}_i)$ holds or, alternatively $\hat{k}_i > \frac{1}{2}$, then firm j reduces its capacity in stage 3 below the capacity level of firm i and chooses $\hat{k}_j = \min\{\hat{k}_i - \alpha, \frac{1}{2}\}$ with $\alpha \rightarrow 0$.

Considering now $\hat{\Pi}_i = \Pi_i(\hat{k}_i, \hat{k}_j(\hat{k}_i, k_j)) + zk_i$ as implicitly defined by (40) and by firm j 's best response function (41) or (42) above, reveals that the small firm i in stage 2 has no interest in firm j becoming later in stage 3 the small firm. Firm i 's profit in stage 2, $\hat{\Pi}_i$, is maximized if firm i chooses in stage 2 of the investment process

$$\hat{k}_i = \min \left\{ k_i, \left(1 - \frac{k_j}{2} \right) k_j \right\} \text{ for all } k_i \leq k_j \leq 1$$

$$\text{and } \hat{k}_i = \min \left\{ k_i, \frac{1}{2} \right\} \text{ for all } k_i \leq k_j \text{ and } k_j > 1.$$

This capacity choice of firm i in stage 2 with $k_i \leq k_j$ triggers later a final capacity choice of firm j in stage 3 with $\hat{k}_j = \min\{k_j, 1\}$. Note that this is also true, if firms are symmetric with $k_i = k_j$ to begin with. Then the firm which can accidentally mothball part of its capacity first, will do so.

- (iii) In stage 1 of the investment process both generators simultaneously choose their capacity and anticipate how the investment process continues in stage 2 and 3. One can show that both generators choosing the same capacity is never an equilibrium because each firm has an incentive to deviate to $k_i = (1 - k_j/2)k_j < k_j$ ensuring that it always becomes the small generator, instead of playing this role only with a probability of 1/2. Both generators also anticipate that, finally, a generator, which chooses a larger capacity than its rival in stage 1, will always continue to be the larger capacity provider and might only mothball any capacity in stage 3, exceeding $k_j = 1$, if it invests into $k_j > 1$. Thus in equilibrium $1 \geq k_j > k_i$ must hold.

Therefore the optimal larger capacity in stage 1 must maximize

$$\bar{\Pi}_j(k_i, k_j) = (1 - k_i) \left[(1 - k_j)k_j + \frac{k_j^2}{2} \right] - zk_j$$

which yields the best response from above in stage 1, being

$$\bar{k}_j(k_i) = 1 - \frac{z}{1 - k_i} \tag{43}$$

as long as $\bar{k}_j(k_i) > k_i \Leftrightarrow k_i < 1 - \sqrt{z}$ holds.

The smaller firm i maximizes

$$\underline{\Pi}_i(k_i, k_j) = (1 - k_i)k_i - zk_i$$

$$\text{subject to } k_i \leq \left(1 - \frac{k_j}{2} \right) k_j \text{ and } \bar{\Pi}_j(k_i, k_j) \geq \underline{\Pi}_j(k_i, k_j).$$

Note that the second of the two constraints is equivalent to $k_i \leq \underline{k}_i(k_j)$ with

$$\begin{aligned} \underline{k}_i(k_j) &= \frac{1}{4} \left\{ 2 + (2 - k_j)k_j - 2z \right. \\ &\quad \left. - \sqrt{4z^2 - 4(2 - k_j)(k_j + 2)z + (2 - (2 - k_j)k_j)^2} \right\} \\ &\leq \left(1 - \frac{k_j}{2} \right) k_j \text{ for all } 0 \leq z < 1, \end{aligned} \quad (44)$$

meaning that $k_i \leq \underline{k}_i(k_j)$ is the only binding constraint.¹⁶ It ensures that a generator, which chooses to be the large generator in the first stage of the investment process does not want to deviate and become the smaller generator. In addition one can show that

$$\arg \max_{k_i} \Pi_i(k_i, k_j) = \frac{1 - z}{2} \geq \underline{k}_i(k_j),$$

for all $0 \leq z < 1$ and $0 \leq k_j \leq 1$. This means that the best response from below in stage 1 is $k_i = \underline{k}_i(k_j)$ defined in equation (44).

Solving $k_j = \bar{k}_j(k_i)$ and $k_i = \underline{k}_i(k_j)$ from equations (43) and (44) simultaneously for (k_i, k_j) , yields the two Nash equilibria identified in Lemma 5. With these capacities both firms choose best responses to each other's capacity and, by construction of k_i , the large generator j does not want to become the small generator and the small generator i is in no way tempted to choose $k_i > k_j$.

Lemma 5 *With two generators $i, j = A, B$ with $i \neq j$ which are separated from the two retailers $\ell, \iota = C, D$, and $\ell \neq \iota$, and with $0 \leq z \leq \frac{1}{2\sqrt{2}}$ there exist two subgame perfect Pareto dominant Nash equilibria in which assumption 1 is satisfied. In each of these two equilibria the two generators $i, j = A, B$ with $i \neq j$ invest in (k_i, k_j) such that*

$$k_i = \frac{3 - \sqrt{1 + 8z}}{4} \text{ and } k_j = \frac{3 - \sqrt{1 + 8z}}{2}. \quad (45)$$

The total capacity is $k^{sd} = k_i + k_j$, and can easily be derived from equation (45) in Lemma 5. It is explicitly given in Proposition 4 together with the retail price r^{sd} .

¹⁶This is intuitive. If the small capacity is small enough, such that the large generator does not want to become the small generator in stage 1 where it even can save capacity costs by deviating, then it is certainly small enough to prevent the large firm from becoming the small firm in the third stage where mothballing does not save any capacity costs.

The latter can be derived from substituting k_i from equation (45) into the highest retail price which generates zero profits for the retailers, given in equation (38) in Corollary 2. Rationing occurs for high demand states with $\varepsilon > k_j = \frac{3-\sqrt{1+8z}}{2}$. Obviously, capacities decrease and the retail price increases if capacity costs z increase.

A.5 Proof of Proposition 5 (Ranking)

The statements in (i) and (ii) directly follow from comparing the capacities and retail prices for the two benchmark cases and the two market configurations characterized in Propositions 1, 2, 3 and 4, given the range of investigated capacity costs $0 \leq z \leq \frac{1}{2\sqrt{2}}$.

For statement (iii) we need to compare welfare for the different cases. Welfare is given by the expected social welfare defined in equation (12), where $U(\cdot)$ is implicitly defined in equation (1), $x(r, \varepsilon)$ is defined in (2), and r refers to the retail price and k to the total capacity in the considered configuration. Substituting the relevant retail price and total capacity for the two benchmark cases, given in Proposition 1 and characterized in Proposition 2, and for the two market configurations, characterized in Proposition 3 and Proposition 4, yields the result, stated in (iii) in Proposition 5. The relevant capacities, retail prices and social welfare levels compared in Proposition 5 are illustrated in Figures 3, 4 and 5.

Note that the social welfare if no electricity is consumed coincides with

$$W_0 = \int_0^1 -\varepsilon - \frac{(-\varepsilon)^2}{2} d\varepsilon = -\frac{2}{3},$$

which can be derived from integrating $V(x, \varepsilon, r)$ from equation (1) for $x = 0$ over all demand states $\varepsilon \in [0, 1]$. Thus, it is no problem that social welfare for some configurations and capacity cost levels z is negative, because it always exceeds the social welfare level without any electricity consumption.

References

- Allaz, B. and J.-L. Villa (1993), ‘Cournot Competition Forward Markets and Efficiency’. *Journal of Economic Theory*, 59(1), pp. 1–16.
- Baldursson, F. M. and N.-H. von der Fehr (2007), ‘Vertical Integration and Long Term Contracts in Risky Markets’. Memorandum 01/2007, Department of Economics, University of Oslo.

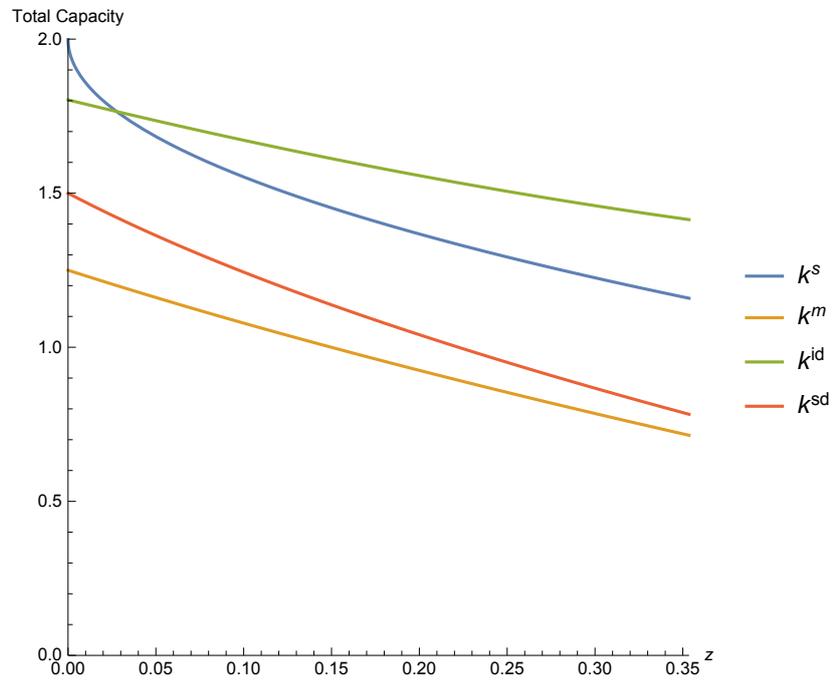


Figure 3: The Capacities in the Different Market Configurations

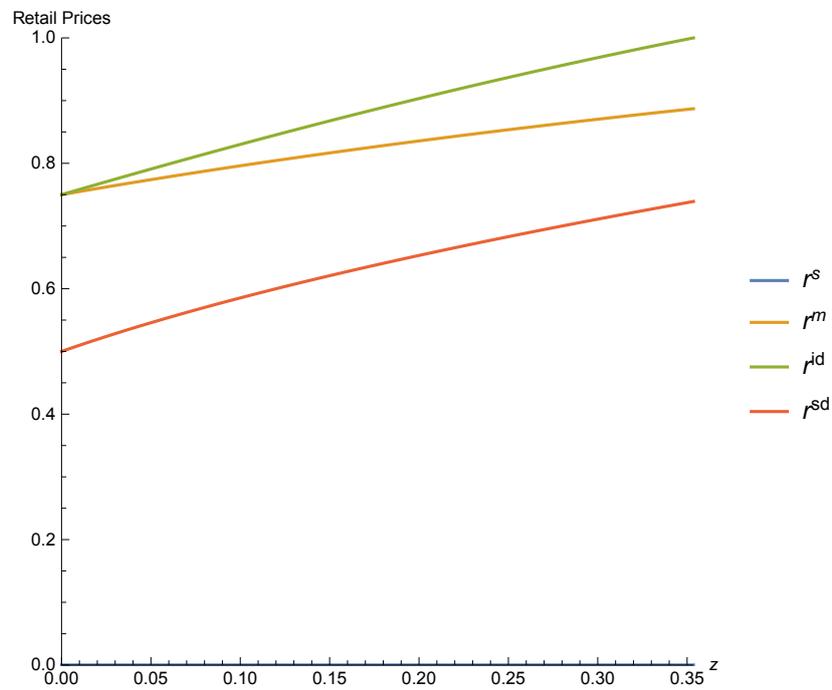


Figure 4: The Retail Prices in the Different Market Configurations

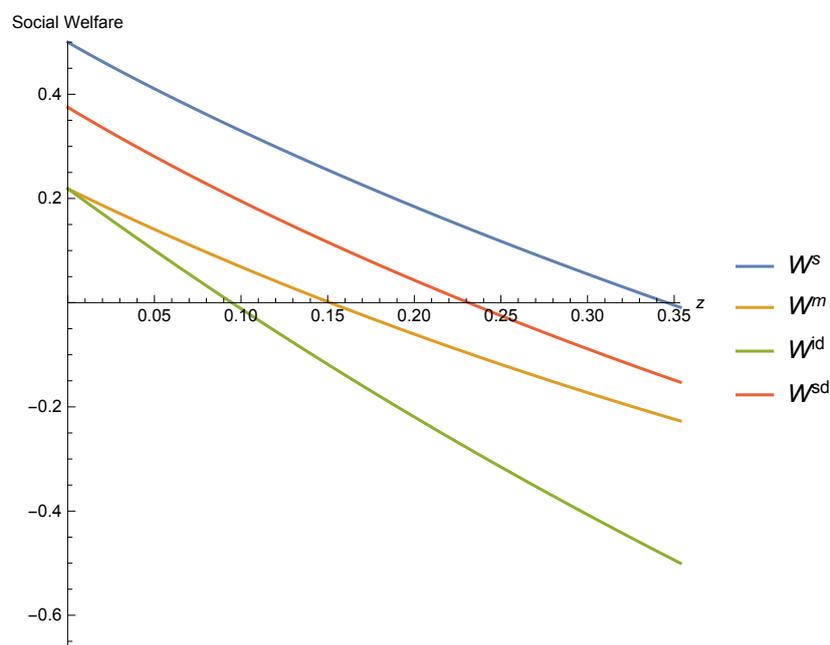


Figure 5: Social Welfare in the Different Market Configurations

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. (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.

Boom, A. and S. Buehler (2014), 'Restructuring the Electricity Industry: Vertical Structure and the Risk of Rent Extraction'. Discussion Paper 2014-37, School of Economics and Political Science, Department of Economics, University of St. Gallen.

Borenstein, S. (2002), 'The Trouble with Electricity Markets: Understanding Cal-

- ifornia's Electricity Restructuring Disaster'. *Journal of Economic Perspectives*, 16, pp. 191–211.
- Borenstein, S. and J. B. Bushnell (2018), 'Do Two Electricity Pricing Wrongs Make a Right? Cost Recovery, Externalities, and Efficiency'. Working Paper 24756, National Bureau of Economic Research, Cambridge, MA, revised September 2018.
- 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.
- 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.
- 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.
- Drèze, J. H. and E. Sheshinski (1976), 'Demand Fluctuation, Capacity Utilization and Costs'. *The American Economic Review*, 66, pp. 731–7442.
- European Commission (2010), 'The Unbundling Regime. Interpretative Note on Directives 2009/72/EC Concerning Common Rules for the Internal Market in Electricity and Directive 2009/73/EC Concerning Common Rules for the Internal Market in Natural Gas'. Commission staff working paper, Brussels.
- Fabra, N., N.-H. M. von der Fehr and M.-Á. de Frutos (2011), 'Market Design and Investment Incentives'. *The Economic Journal*, 121, pp. 1340–1360.

- 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.-Á. 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.
- 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.
- 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.
- Mahenc, P. and F. Salanié (2004), 'Softening Competition through Forward Trading'. *Journal of Economic Theory*, 116, pp. 282–293.
- 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.
- Schwenen, S. (2015), 'Strategic Bidding in Multi-unit Auctions with Capacity Constrained Bidders: The New York Capacity Market'. *The Rand Journal of Economics*, 46(4), pp. 730–750.