

Zöttl, Gregor

**Conference Paper**

## On Optimal Scarcity Prices

Beiträge zur Jahrestagung des Vereins für Socialpolitik 2010: Ökonomie der Familie - Session: Regulation and Pricing Policies, No. F8-V3

**Provided in Cooperation with:**

Verein für Socialpolitik / German Economic Association

*Suggested Citation:* Zöttl, Gregor (2010) : On Optimal Scarcity Prices, Beiträge zur Jahrestagung des Vereins für Socialpolitik 2010: Ökonomie der Familie - Session: Regulation and Pricing Policies, No. F8-V3, Verein für Socialpolitik, Frankfurt a. M.

This Version is available at:

<https://hdl.handle.net/10419/37363>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

# On Optimal Scarcity Prices\*

GREGOR ZÖTTL<sup>†</sup>

University of Munich

March 1, 2010

## Abstract

This article contributes to the debate of *missing money* (e.g. Joskow(2007)) which has seriously questioned the desirability of caps on scarcity prices in markets with fluctuating demand by emphasizing their potentially negative impact on firms investment decisions in the long run. A prominent example are recently liberalized electricity markets, where competition authorities have imposed caps in order to mitigate the exercise of market power at the spot markets.

In order to shed light on the still incompletely explored impact of such caps in the long run we analyze investment of strategic firms in base-load and peak-load technologies in a market with fluctuating demand. We show that under imperfect competition appropriately chosen caps on scarcity prices not only are beneficial in the short run but also in the long run, since they lead to a strict increase of investment in the peak-load technologies, leaving investment in the base-load technologies unchanged. We furthermore characterize the optimal cap on scarcity prices.

**Keywords:** Investment Incentives, Fluctuating Demand, Scarcity Rents, Missing Money, Electricity Markets

**JEL classification:** D43, L13, D41, D42, D81.

---

\*We thank Rabah Amir, Claude d'Aspremont, Veronika Grimm, David Newbery, and Yves Smeers, as well as seminar participants at Berlin, Brussels, Cambridge, and Cologne for helpful comments and suggestions.

<sup>†</sup>University of Munich, Ludwig Strasse 28, München, Germany, gregor.zoetl@lrz.uni-muenchen.de

# 1 Introduction

Prices at electricity spot markets are highly volatile and typically range from 0\$/MWh to more than 1000\$/MWh<sup>1</sup>. As many experts argue, especially in hours of high demand and tight supply, electricity markets might very likely be subject to the exercise of market power.<sup>2</sup> Caps on those price spikes have thus been proposed, or are already used to mitigate market power at electricity wholesale markets during those critical hours.<sup>3</sup> In the words of a Federal Energy Regulatory Commissioner: "If you cap those prices, you eliminate any incentive to withhold . . . you may well sell into the market at a capped price as long as you're covering your running cost and making a reasonable profit".<sup>4</sup> In the short run, for fixed investment choice, such caps indeed are welfare enhancing. When taking into account firms' investment decisions, however, their desirability is much less clear cut and the questions involved are not yet well explored in the literature.

Recently, economists have seriously questioned the desirability of capping scarcity prices in markets with fluctuating demand, arguing that this considerably distorts firms' investment decisions. Their argument is clear cut, originating in the findings of the peak-load pricing literature<sup>5</sup>, which analyzes firms' investment incentives under fluctuating demand: High price spikes during hours of high demand and scarce supply do send accurate price signals for investment decisions. In equilibrium market prices will be such as to sustain optimal investment levels. Restricting scarcity prices distorts those price signals and lowers firms' investment incentives. In the literature this is often referred to as a problem of "missing money". As Paul Joskow (2007) summarizes in a recent contribution: "As a result, many economists assume that the primary source of the missing money problem must be the price caps and related market power mitigation procedures imposed by regulators. That is, that the efforts to mitigate market power have had the effect of suppressing energy prices too much, especially during scarcity conditions when prices should be very high." As a reaction to those considerations caps at several wholesale electricity markets have recently been substantially increased.<sup>6</sup>

---

<sup>1</sup>E.g. wholesale prices at the German market ranged from 0€/MWh to 2489€/MWh in 2006.

<sup>2</sup>Exercise of market power at electricity spot markets has received much attention, compare Borenstein et al. (2002) for California, Wolfram (1999) for the UK, or Schwarz and Lang (2006) for Germany.

<sup>3</sup>In the major electricity markets in the US, the PJM and the Californian market wholesale prices are capped at 1000\$/MWh and 2500\$/MWh respectively (compare FERC(2000) and FERC(2009)).

<sup>4</sup>See Walsh (2001).

<sup>5</sup>For a survey see Crew, Fernando, and Kleindorfer (1995).

<sup>6</sup>Caps in the Californian electricity market have now been adjusted from 250\$/MWh to 2500\$/MWh, see FERC(2000b) and FERC(2009).

It is the purpose of the present article to contribute to the discussion of "missing money" and suppressed scarcity rents. Up to date this discussion has only considered the case of perfectly competitive markets, we want to extend the focus also to strategically behaved firms.<sup>7</sup> This seems to be a natural consideration: If firms are suspected to behave strategically<sup>2</sup> when producing at the spot market,<sup>8</sup> then it seems natural to consider firms to also behave strategically when making their investment decisions. Notice furthermore that a contribution to this discussion necessarily has to take into account the fluctuating nature of spot market demand. Approximation by a single deterministic spot market would erase occurrence of price spikes and (by construction) eliminate the fundamental problem under study. That is, the well known results<sup>9</sup> obtained for a single spot market in the absence of fluctuations cannot contribute substantially to the discussion of "missing money", neither for the case of perfect competition nor for strategically behaved firms.

We thus choose to explicitly take into account fluctuating demand at the spot markets, giving rise to scarcity pricing and price spikes. Notice that under fluctuating demand firms choose to invest in a mix of different technologies. Base-load technologies (e.g. nuclear or coal plants), which allow to produce cheaply, have to run most of the time in order to recover their high investment cost. Peak-load technologies, which are cheap to build, are used in order to additionally serve more infrequent demand peaks. We then consider a cap on scarcity prices at the spot market and analyze its impact on strategic firms' investment decisions and on overall welfare. As we show, too low limits on scarcity prices lead to a reduction of investment, at least partially confirming the already known findings of the missing money discussion. In contrast to those results, however, appropriately chosen limits on scarcity prices do always lead to a strict increase of overall investment of strategic firms. More specifically, they have a strictly positive impact on investment in the peak-load technology, leaving investment in the base-load technologies unchanged. We are able to

---

<sup>7</sup>Indeed the peak-load pricing literature which analyzes investment incentives under demand fluctuation does not provide results for the case of strategically behaved firms, we thus extend those results to the case of imperfect competition, see section 3.

<sup>8</sup>Indeed the rationale for the presently imposed price caps exclusively is the mitigation of market power at the spot market, compare Joskow and Tirole (2007).

<sup>9</sup>For a single deterministic spot market, under perfect competition all price caps above total marginal cost are irrelevant, those below total marginal cost lead to a complete stoppage of investment and production. For the case of strategic investment decisions, independently of spot market competition, firms can fully determine the market outcome by choosing investment in a single technology. The entire market game can then be modeled as a static Cournot framework with the optimal price cap at marginal cost (compare for example Kreps and Scheinkmann (1983) and the subsequent contributions). Those results provide little insight on the problem of suppressed scarcity rents and missing money, however.

characterize those caps and finally also determine the optimal cap on scarcity prices.

That is, if firms are suspected to have market power in a market with fluctuating demand and a cap on scarcity prices is put in place to mitigate the exercise of market power at the spot market, then an appropriately chosen cap does not lead to a reduction of firms' investment incentives. Unlike suggested in the current discussion of missing money and suppressed scarcity rents, this implies that there is no trade-off to be solved between the long run and the short run perspective. This new insight arises from the fact that we explicitly model strategic investment decisions of firms, whereas previous arguments relied on the hypothesis of competitive investment behavior.<sup>10</sup>

Since the problem of missing money has been discussed extensively in the context of investment decisions in liberalized electricity markets, we fit our model to a specific electricity market (the German market) in order to quantitatively illustrate our results. We find that restriction of scarcity prices below 260 €/MWh increase total industry investment for any market structure based on up to 20 firms.<sup>6</sup> Especially for concentrated markets (up to 4 firms) surprisingly low caps on scarcity prices ranging at around 110 €/MWh lead to a dramatic increase of industry capacities and overall welfare (for the case of duopoly total investment increases by 30%). Our quantitative results thus suggest that for appropriately chosen caps on scarcity prices welfare can be increased substantially both from a short run and a long run perspective.

In the literature most of the contributions on price caps under fluctuating demand or uncertainty analyze static frameworks, which do not consider investment decisions. Among those most focus on the analysis of cost uncertainty. The case of symmetric information between firms and regulator regarding cost of production has been analyzed for example by Pint (1992), or Biglaiser and Riordan (2000). The literature on incentive regulation analyzes the case of asymmetric information in a static framework, for a survey of the seminal contributions see Laffont and Tirole (1993), or Laffont and Tirole (2001). All those approaches do not take into account firms' investment incentives prior to competing on the regulated spot markets. As Cowan (2002) and more recently Joskow and Tirole (2007), or Hogan (2005) point out, those incentives are crucial, however, when evaluating the desirability of caps on scarcity prices, since they limit price spikes in those markets and thus distort investment decisions.

Another strand of literature focuses on optimal timing of investment decisions in stochas-

---

<sup>10</sup>Since strategic behavior of firms at the spot markets is the justification for caps on scarcity prices, a thorough analysis of the strategic investment decisions seems to be a natural contribution to this discussion.

tic continuous time models. The seminal contributions of Dixit (1991) and Dixit and Pindyk (1994), for the case of perfect competition, has been extended by Dobbs (2004), and in a recent contribution Roques and Savva (2006) analyze the case of strategically interacting firms. In this literature, demand typically follows a Brownian process, capacities are assumed to be always fully utilized<sup>11</sup>. Roques and Savva (2006) show that strategic firms delay their investment decisions, relative to the case of perfect competition (that is, the price trigger for investment is higher under strategic interaction, than for a perfectly competitive industry). They find that introduction of a price cap can reduce this problem and show that the optimal price cap in their framework is given by the investment trigger-price of a competitive industry. Development of industry demand in those models of optimal timing is best interpreted as changes in the level of average market demand where changed price levels trigger instantaneous investment adjustment. Those contributions thus provide very rich insights on the dynamical timing of strategic firms' investment decisions. They abstract, however, from scarcity pricing and price spikes due to fluctuating demand at the spot markets. Since the discussion of the missing money problem precisely analyzes the impact of reduced scarcity rents, such framework of optimal timing of investment decisions does not serve our purpose to contribute to this discussion for the case of strategic investment decisions.

The article is organized as follows: In section 2 we introduce the model. In Section 3 we derive the market equilibrium of the investment game and its existence and uniqueness. Section 4 characterizes desirable and section 5 optimal caps on scarcity prices. In section 6 we quantify our results for the example of the German electricity market. Section 7 concludes.

## 2 The Model

Firms can invest in different technologies which allow production at many spot markets. Spot markets are subject to fluctuating demand, additionally firms might be uncertain regarding the precise pattern of such fluctuation. Once investment decisions are made, firms make their production decisions for each single spot market. At the time of production firms know demand at each spot market. Price spikes at the spot market are limited by a cap  $\bar{p}$ . We denote by  $q(\theta) = (q_1, \dots, q_n)$  the vector of outputs of the  $n$  firms and by  $Q(\theta) = \sum_{i=1}^n q_i$  total quantity produced at each spot market.

---

<sup>11</sup>This allows to disregard interaction at the spot market.

Inverse Demand is given by the function  $P(Q, \theta)$ , which depends on  $Q \in \mathbb{R}^+$ , and the variable  $\theta \in \mathbb{R}$  which represents both fluctuation and uncertainty of demand. The variable  $\theta \in \mathbb{R}$  is distributed according to  $F(\theta)$  with bounded support  $[\underline{\theta}, \bar{\theta}]$ . Notice that the distribution function  $F(\theta)$  couples frequencies originating from fluctuating demand and probability mass originating from uncertainty.<sup>12</sup>

We denote by  $\bar{Q}(\theta)$  the quantity where prices start to be zero,<sup>13</sup> the following assumption has to be satisfied for quantities  $0 \leq q_i \leq Q < \bar{Q}(\theta)$ .

ASSUMPTION 1 [*Demand*]

(i) *Inverse demand  $P(Q, \theta)$  is twice continuously differentiable<sup>14</sup> in  $Q$  with  $P_q(Q, \theta) < 0$  and  $P_q(Q, \theta) + P_{qq}(Q, \theta)q_i < 0$ .*

(ii)  *$P(Q, \theta)$  is differentiable in  $\theta$ , with  $P_\theta(Q, \theta) > 0$  and  $P_\theta(Q, \theta) + P_{q\theta}(Q, \theta)q_i > 0$ .*

We are interested in a situation, where firms can invest into several different technologies. Those come at different levels of investment cost and allow for producing at different levels of marginal cost. In the case of electricity markets this corresponds to the different electricity generation technologies such as nuclear, lignite-, coal-, and gas-fired power plants. For the formal model, we will analyze only the case of two different technologies ("peak-load" and "base-load"), in order to keep the notational burden limited. Throughout the article we will comment, however, how our results could be extend to the case of more technologies.

ASSUMPTION 2 [*Technologies*] *Firms can invest in the two different technologies  $(k, c)$  and  $(k_1, c_1)$ . Constant marginal cost of investment are given by  $k$  and  $k_1$  respectively and the corresponding marginal cost of production are given by  $c$ , and  $c_1$  respectively. Wlg we assume<sup>15</sup>  $c > c_1$  and  $k < k_1$ .*

Since demand fluctuates over time, it is optimal for firms to invest into a mix of technologies. Technology  $(k, c)$  is relatively cheap to build, but produces at high marginal cost. Those units will serve only higher realizations of demand ("peak-load"). Technology  $(k_1, c_1)$

---

<sup>12</sup>That is, the associated density  $f(\theta)$  will not integrate to 1, but to the total number of spot markets analyzed. Even though this detail will not be relevant when solving our general framework, it should be explicitly emphasized at this point.

<sup>13</sup>In case prices remain positive we can set  $\bar{Q}(\theta) = \infty$ . In order to ensure a bounded solution we then have to assume  $\lim_{Q \rightarrow \infty} P(Q, \theta) < c$  for each  $\theta \in (-\infty, \infty]$ .

<sup>14</sup>We denote the derivative of a function  $g(x, y)$  with respect to the argument  $x$ , by  $g_x(x, y)$ , the second derivative with respect to that argument by  $g_{xx}(x, y)$ , and the cross derivative by  $g_{xy}(x, y)$ .

<sup>15</sup>Whenever one technology has both lower production and lower investment cost we obtain a corner solution with one technology strictly dominating.

allows to produce more cheaply, however, it has to run a larger fraction of time, in order to justify its high investment cost ("base-load"). We denote investment of some firm  $i$  in technology  $(k_1, c_1)$  by  $x_i^1$  and total investment<sup>16</sup> of firm  $i$  in both technologies by  $x_i$  (that is, investment in technology  $(k, c)$  is then given by  $(x_i - x_i^1)$ ). For given investment  $(x_i, x_i^1)$  we can now pin down cost of firm  $i$  for producing some quantity  $q_i$  at a spot market:

$$C(q_i, x_i, x_i^1) = \begin{cases} c_1 q_i & \text{for } 0 < q_i \leq x_i^1 \\ c q_i & \text{for } x_i^1 < q_i \leq x_i \\ \infty & \text{for } x_i < q_i \end{cases} \quad (1)$$

That is firms' investment determines their cost function, allowing them to produce at the spot markets. Firms make simultaneous production decisions for each spot market. Since demand is known to firms at the time of making their production decisions, produced quantities are contingent on the demand scenarios  $\theta \in [\underline{\theta}, \bar{\theta}]$ . We can now state firm  $i$ 's expected profit when choosing investment given by  $(x_i, x_i^1)$  and production schedules given by  $q(\theta)$  for all  $\theta$ .

$$\pi_i(x_i, x_i^1, q) = \int_{\underline{\theta}}^{\bar{\theta}} [\min\{\bar{p}, P(Q(\theta), \theta)\} q_i(\theta) - C(q_i(\theta), x_i, x_i^1)] dF(\theta) - k_1 x_i^1 - k(x_i - x_i^1).$$

We are interested in pure strategy, perfect Bayesian equilibria of the game. We denote the equilibrium investment profile by  $(x^*, x^{1*})$  and the corresponding industry investments by  $(X^*, X^{1*})$ .

We conclude this section by identifying those cases, where indeed investment in technology  $(k, c)$  could take place.<sup>17</sup> First, in absence of a cap on scarcity prices, investment in technology  $(k, c)$  can only be profitable whenever  $k < \mathbf{E}_\theta[P(0, \theta) - c]$ . If this condition is not satisfied, firms will not invest in technology  $(k, c)$ , independently of caps on scarcity prices at the spot markets. Second, for caps very close to marginal cost of production  $c$ , firms will not find it profitable to invest in technology  $(k, c)$ , since they cannot recover cost of investment. The lowest cap which yields positive investment  $\bar{p}^0$  is defined by:  $\bar{p}^0 = \{\bar{p} : \mathbf{E}_\theta[\min\{\bar{p}^0, P(0, \theta)\} - c] = k\}$ . We thus restrict attention to caps  $\bar{p} \geq \bar{p}^0 \geq c$  throughout the article.

<sup>16</sup>The cumulative representation of investment levels crucially simplifies the derivation of the optimal investment strategies, allowing to entirely disentangle the decision regarding total investment  $x_i$  and investment in the "base-load" technology  $x_i^1$  (i.e. cross derivatives of firms' profits wrt  $x_i, x_i^1$  will cancel, making the verification of second order conditions tractable).

<sup>17</sup>We will identify those case where investment in both technologies takes place after having analyzed the market game in section 3.



### 3 Market Equilibrium

In this section we characterize equilibrium investment of strategic firms for a given cap on scarcity prices  $\bar{p}$ . We first analyze the case of strategic investment in a single technology (theorem 1) and then the case of two different technologies (theorem 2). As a benchmark, we finally determine investment decisions under perfect competition (theorem 3).

**THEOREM 1 (MARKET EQUILIBRIUM, SINGLE TECHNOLOGY)** *The market game where firms can invest in the technology  $(k, c)$  has a unique equilibrium which is symmetric. Industry investment  $X^*(\bar{p})$  is characterized by*

$$X^*(\bar{p}) = \left\{ X : \int_{\tilde{\theta}^X(x)}^{\tilde{\theta}^{\bar{p}}(x, \bar{p})} \left[ P(X, \theta) + P_q(X, \theta) \frac{X}{n} - c \right] dF(\theta) + \int_{\tilde{\theta}^{\bar{p}}(x, \bar{p})}^{\bar{\theta}} (\bar{p} - c) dF(\theta) = k \right\}. \quad (2)$$

$\tilde{\theta}^X(x) \in [\underline{\theta}, \bar{\theta}]$  is the demand scenario from which on firms are capacity constrained and  $\tilde{\theta}^{\bar{p}}(x, \bar{p}) \in [\underline{\theta}, \bar{\theta}]$  is the demand scenario from which on the cap starts to be binding.

**PROOF** See appendix 7. □

In order to provide some intuition for the equilibrium–characterization in theorem 1, we first determine the range of relevant caps, which are binding at least at some spot markets. We denote by  $\bar{\rho}^\infty$  that cap, which in equilibrium is reached only for the highest demand realization  $\bar{\theta}$ , it is defined by:

$$\bar{\rho}^\infty = \{ \bar{p} : \bar{p} = P(X^*(\bar{p}), \bar{\theta}) \}.$$

All higher caps  $\bar{p} \geq \bar{\rho}^\infty$  are never binding. In other words, the demand realization where the cap starts to be binding  $\tilde{\theta}^{\bar{p}}(x, \bar{\rho}^\infty)$  is at or above the range of relevant demand realizations.<sup>18</sup> Formally this corresponds to  $\tilde{\theta}^{\bar{p}}(x, \bar{\rho}^\infty) > \bar{\theta}$ , leading to the elimination of the second term on the LHS of (2). For the absence of caps on scarcity prices, i.e never–binding caps, the first order condition has a straightforward intuition: The LHS is expected marginal revenue of investment  $X$ , whereas the RHS is just marginal cost  $k$  of investment in technology  $(k, c)$ . Note that on the LHS expectation is taken only over those spot markets, where the additional unit installed would actually be used (i. e. the lower limit of integration is  $\tilde{\theta}^X(x)$  and not  $\underline{\theta}$ ). The reason is, that additional investment only contributes to marginal revenue for those spot markets, where it is actually used.

We thus observe that only caps in the range  $\bar{p} \in [\bar{\rho}^0, \bar{\rho}^\infty]$  yield truly interior solutions with positive investment and binding cap at some spot markets. Characterization of the

<sup>18</sup>Remember  $\theta$  has been defined on  $\mathbb{R}$ , the distribution of uncertainty has the support  $[\underline{\theta}, \bar{\theta}]$ , however.

corresponding equilibrium in expression (2) again equates expected marginal revenue of investment with its marginal cost. As in the case without binding cap, when calculating expected marginal revenue, the firm considers only those spot markets, where the additional investment in technology  $(k, c)$  is actually being used (i.e. integration starts at  $\tilde{\theta}^X$ ). Marginal revenue in scenario  $\theta$  is given by the well known expression  $(P - P_q x - c)$ , until the cap is binding at  $\tilde{\theta}^{\bar{p}}$ . From there on, marginal revenue equals to  $(\bar{p} - c)$ .

After having analyzed the case of investment in a single technology, we turn towards extending those results to the case of investment in two different technologies. We will focus on those cases where the technologies  $(k_1, c_1)$  and  $(k, c)$  have sufficiently different cost of production. This allows to ensure existence of a well behaved symmetric equilibrium. Firms' investment decisions for a continuum of infinitely close technologies has been treated in Zoetl (2008). We are able to determine a unique equilibrium, which is given as follows:

**THEOREM 2** [*Market Equilibrium, Two Technologies*] *The market game, where firms can invest in the two different technologies  $(k, c)$  and  $(k_1, c_1)$  has a unique equilibrium, which is symmetric. Industry investment  $(X^*(\bar{p}), X^{1*})$  is characterized by*

$$X^*(\bar{p}) = \left\{ X : \int_{\tilde{\theta}^X(x)}^{\tilde{\theta}^{\bar{p}}(x, \bar{p})} \left[ P(X, \theta) + P_q(X, \theta) \frac{X}{n} - c \right] dF(\theta) + \int_{\tilde{\theta}^{\bar{p}}(x, \bar{p})}^{\bar{\theta}} (\bar{p} - c) dF(\theta) = k \right\} \quad (3)$$

$$X^{1*} = \left\{ X : \int_{\tilde{\theta}^{X^1}(x)}^{\tilde{\theta}^X(x)} \left[ P(X, \theta) + P_q(X, \theta) \frac{X}{n} - c_1 \right] dF(\theta) + \int_{\tilde{\theta}^X(x)}^{\bar{\theta}} (c - c_1) dF(\theta) = k_1 - k \right\}. \quad (4)$$

$\tilde{\theta}^{X^1}(x, \bar{p})$  is the demand scenario, from which on investment in the base-load technology  $(k_1, c_1)$  is binding, and  $\tilde{\theta}^X(x, \bar{p})$  is the demand scenario from which on firms produce at the marginal cost of the peak-load technology  $c$ .

**PROOF** See appendix 7. □

When comparing the case of investment in a single technology (theorem 1) and investment in two technologies (theorem 2), we observe that total investment  $X^*(\bar{p})$  remains entirely unaffected, if firms can choose to additionally invest in the base-load technology  $(k_1, c_1)$ . Clearly, this does not imply that investment in technology  $(k, c)$  remains unchanged with respect to the single technology case. In contrast, having the option to choose also from technology  $(k_1, c_1)$ , firms in the market equilibrium find it cheaper to invest the amount  $X^{1*}$  in technology  $(k_1, c_1)$  and only the remaining units  $X^* - X^{1*}$  will be chosen from technology  $(k, c)$ . That is, in the market solution, as given by theorem 2, firms determine their

total investment  $X^*$  based on the "least efficient"<sup>19</sup> technology available as characterized by expression (3).

The investment decision  $X^{1*}$  in the more efficient technology solves the trade off of replacing some units of technology  $(k, c)$  by units of technology  $(k_1, c_1)$ , as formalized in expression (4): The cost of replacing the marginal unit of technology  $(k, c)$  by a unit of technology  $(k_1, c_1)$  is given by  $(k_1 - k)$ , the RHS of (4). This has to equal the extra profits generated by that unit. Those come through two channels, which are given by the LHS of expression (4). First marginal spot market profits are changed for all demand realizations where capacity of technology  $(k_1, c_1)$  is binding, i.e. for  $\theta \in [\tilde{\theta}^{X^1}, \tilde{\theta}^X]$ . Second it allows to produce one extra unit of output at (lower) marginal cost  $c_1$  instead of  $c$  for all demand realizations  $\theta \in [\tilde{\theta}^X, \bar{\theta}]$ .

We conclude the discussion of theorem 2 by identifying those cases where indeed an interior solution with  $0 < X^{1*} < X^*$  obtains. In section 2 we already established conditions such that  $X^* > 0$ . We now have to verify that investment  $X^{1*}$  in technology  $(k_1, c_1)$  remains within the admissible bounds of  $X^{1*} > 0$  and  $X^{1*} < X^*$ . Violation of either of those two conditions brings us back to the single technology case of theorem 1. First, whenever technology  $(k_1, c_1)$  is very unattractive,<sup>20</sup> then only technology  $(k, c)$  is active. Second, if technology  $(k_1, c_1)$  is always more attractive<sup>21</sup> than technology  $(k, c)$ , then only technology  $(k_1, c_1)$  is active in the market equilibrium.

In the subsequent sections 4 and 5 we will analyze the impact of a cap on scarcity prices on the market equilibrium and determine desirable and optimal levels of such cap. Evaluation of the market equilibrium already at this point reveals a striking property: Investment in the base-load technology  $(k_1, c_1)$  is independent of the cap  $\bar{p}$ , which only has an impact on investment in the peak-load technology  $(k, c)$ . This result follows immediately from the equilibrium conditions specified in theorem 2. Caps matter only at the spot markets with high demand, when prices are high due to tight supply at the capacity bound  $X^*$  (i.e. for  $\theta \in [\tilde{\theta}^{\bar{p}}, \bar{\theta}]$ ). Investment in the base-load technology  $(k_1, c_1)$  is driven exclusively by the trade-off of replacing technology  $(k, c)$  by technology  $(k_1, c_1)$ . This tradeoff is independent of the level of the cap on scarcity prices, however. This observation will have significant importance in the subsequent sections, and is thus briefly summarized in corollary 1.

**COROLLARY 1** *In any market equilibrium, where firms invest in both technologies (i.e.*

---

<sup>19</sup>More precisely that is the technology which is cheapest to invest, but most expensive to produce with, given by  $(k, c)$ .

<sup>20</sup>That is expression (4) yields  $X^{1*} \leq 0$ .

<sup>21</sup>That is expressions (3) and (4) yield  $X^{1*} \geq X^*$ .

$0 < X^{1*} < X^*$ ), a change of the cap  $\bar{p}$  leading to higher total investment  $X^*(\bar{p})$ , leaves investment  $X^{1*}$  in technology  $(k_1, c_1)$  unchanged.<sup>22</sup>

We briefly discuss, how those results would extend to the case of more than two technologies. Suppose firms can invest also in a third, even more efficient technology  $(k_2, c_2)$  with  $c_2 < c_1$  and  $k_2 > k_1$ . Adopting the cumulative definition of investment levels, we denote total investment by  $X^*$ , total investment in technologies  $(k_2, c_2)$  and  $(k_1, c_1)$  by  $X^{1*}$  and investment in technology  $(k_2, c_2)$  by  $X^{2*}$ . The equilibrium conditions for investment levels  $X^*$  and  $X^{1*}$  remain unchanged and are given by expressions (3) and (4). Congruent to expression (3), investment  $X^{2*}$  solves the trade off of replacing units of technology  $(k_1, c_1)$  by units of technology  $(k_2, c_2)$ . Analogous to corollary 1, again only total investment  $X^*$ , but not investment in the base-load technologies (now  $(k_1, c_1)$  and  $(k_2, c_2)$ ) are affected by a change of the cap on scarcity prices.

It should be explicitly emphasized that a unique symmetric equilibrium obtains only for investment in a limited number of sufficiently different technologies, but not for the limit case of finitely many technologies. Then a symmetric equilibrium generally fails to exist. As Zöttl(2008) shows for a continuum of technologies, existence and uniqueness of the symmetric equilibrium can be reestablished for smooth investment functions, however.<sup>23</sup> That is, we obtain existence and uniqueness of a symmetric equilibrium both for the case of a limited number and a continuum of technologies, but not for the limit case of finitely many technologies, where asymmetric equilibria of the investment game arise.<sup>24</sup>

Throughout this article we focus on those cases, where a unique symmetric equilibrium of the investment game obtains. We do not formally analyze the limit case of sufficiently similar technologies, leading to asymmetric market outcomes. This would increase the

---

<sup>22</sup>In principle the same holds true for caps that lead to lower total investment. A reduction of total investment could lead to a corner solution where firms stop to invest in the peaking technology  $(k, c)$ , however. Investment in the single technology  $(k_1, c_1)$  would then depend on the cap  $\bar{p}$ , however. Such corner solution cannot result from an increase of total investment.

<sup>23</sup>For a finite number of technologies firms investment decisions necessarily have to be step functions. In a continuous framework firms are able to choose smooth investment schedules.

<sup>24</sup>Interestingly this parallels the findings on supply function equilibria. In a seminal article Klemperer and Meyer (1989) show existence and uniqueness of differentiable supply functions for a continuous distribution of demand uncertainty. Bolle (1992) and Green and Newberry (1992) apply those findings in order to model firms' behavior at electricity spot markets. In a subsequent contribution Fehr and Harbord (1993) analyze the market game where firms bid discrete supply functions and find that symmetric equilibria in discrete supply functions fail to exist for the limit case of finitely small bid-steps. Nevertheless, depending on the precise purpose, both discrete and continuous supply function frameworks enjoy unchanged popularity when modeling firms' behavior at electricity spot markets.

notational and computational burden of the present article considerably. Furthermore, as shown in appendix <sup>25</sup>, potentially arising asymmetric equilibria are asymmetric only with respect to investment in technology  $(k_1, c_1)$ , but still exhibit symmetric total investment level  $X^*$ . Following the insights provided by corollary 1, a cap on scarcity prices only has an impact on total investment  $X^*$  characterized by expression (3), however. Desirable and optimal caps, as analyzed in the remainder of this article, can thus be determined independently of the nature of equilibrium investment in technology  $(k_1, c_1)$ . That is, also for the limit case of similar technologies, when asymmetric equilibria arise, the findings of theorems 4, 5, and 6 would continue to hold true.

As a benchmark for the strategic market game, we now briefly state the solution for the case of perfect competition. The critical spot market scenarios are determined analogously to those of theorem 2, now for the case of perfectly competitive behavior at all spot markets, however.<sup>26</sup>

**THEOREM 3 (EQUILIBRIUM UNDER PERFECT COMPETITION)** *In a perfectly competitive market industry investment  $(\widehat{X}^1, \widehat{X}(\bar{p}))$  is uniquely characterized by*

$$\widehat{X}(\bar{p}) = \left\{ X : \int_{\widehat{\theta}^X(x)}^{\widehat{\theta}^{\bar{p}}(x, \bar{p})} [P(X, \theta) - c] dF(\theta) + \int_{\widehat{\theta}^{\bar{p}}(x, \bar{p})}^{\bar{\theta}} (\bar{p} - c) dF(\theta) = k \right\} \quad (5)$$

$$\widehat{X}^1 = \left\{ X : \int_{\widehat{\theta}^{X^1}(x)}^{\widehat{\theta}^X(x)} [P(X, \theta) - c_1] dF(\theta) + \int_{\widehat{\theta}^X(x)}^{\bar{\theta}} (c - c_1) dF(\theta) = k_1 - k \right\}. \quad (6)$$

*The welfare maximum is implemented without cap on scarcity prices. Any binding cap leads to a reduction of investment and total welfare.*

**PROOF** See appendix 7 □

It is worthwhile to notice that the lowest non binding cap under perfect competition,  $\widehat{\rho}^\infty = P(\widehat{X}(\widehat{\rho}^\infty), \bar{\theta})$ , will play an important role in the following section. It is given by the market price obtained under perfect competition for the highest possible demand realization. Under perfect competition overall welfare is maximized, if scarcity prices remain uncapped. Distortion of this market outcome by a binding cap will always reduce equilibrium investment choices and welfare. We can thus conclude that for a perfectly competitive industry a limitation of scarcity prices is never desirable under demand fluctuation. This

<sup>25</sup>See appendix 7.

<sup>26</sup>From  $\widehat{\theta}^{X^1}(x, \bar{p})$  on investment in the base-load technology  $(k_1, c_1)$  is binding, from  $\widehat{\theta}^X(x, \bar{p})$  on firms produce at marginal cost  $c$ , from  $\widehat{\theta}^X(x)$  on total investment is binding, and from  $\widehat{\theta}^{\bar{p}}(x, \bar{p})$  on also the cap is binding.

replicates the arguments of the missing money problem as put forth e.g. by Hogan(2005) or Joskow(2007). As we show in the subsequent section, however, this is not true for strategically behaved firms.

## 4 Desirable Caps on Scarcity Prices

As argued by Hogan(2005) or Joskow(2007), limiting scarcity prices in a framework of fluctuating spot market demand leads to problems of missing money, implying a reduction of firms' investment incentives and welfare. This argument up to date has been made exclusively under the hypothesis of a perfectly competitive market. The extension to the case of strategic investment seems to be an important contribution to this discussion: Caps on scarcity prices are designed in order to limit the exercise of market power of strategically behaved firms at the spot markets, indeed only then the use of such caps is justifiable. It thus seems most natural to also consider strategic behavior of firms when making investment decisions. We want to shed light on this issue and provide insight on the design of such caps for strategic investment decisions prior to fluctuating demand.

As already established in section 2, also under strategic behavior caps close to marginal cost  $c$  of production (i.e.  $\bar{p} \in [c, \bar{p}^0]$ ) will completely eliminate equilibrium investment in technology  $(k, c)$  and lead to a reduction (or elimination) of investment in technology  $(k_1, c_1)$ . Such caps on scarcity prices are thus never desirable to introduce, the results obtained under perfect competition and under strategic behavior coincide.

It remains to explore, however, if higher caps  $\bar{p} \in [\bar{p}^0, \bar{p}^\infty]$  could be beneficial for the case of strategically behaved firms. As we find, under imperfect competition, indeed there always exist binding caps on scarcity prices which lead to an increase of total investment and overall welfare. In line with corollary 1, only investment in technology  $(k, c)$  is affected, whereas investment  $X^{1*}$  in technology  $(k_1, c_1)$  remains unchanged.

### THEOREM 4 (INVESTMENT AND WELFARE ENHANCING CAPS ON SCARCITY PRICES)

- (i) Denote by  $\overline{MR}$  the highest marginal equilibrium revenue without cap on scarcity prices<sup>27</sup>. A cap  $\bar{p} \in [\overline{MR}, \bar{p}^\infty]$  increases equilibrium investment in technology  $(k, c)$  and leaves investment in technology  $(k_1, c_1)$  unchanged. It furthermore leads to an increase of consumer surplus and total welfare, and lowers average prices as compared to the market game without cap.

---

<sup>27</sup>I.e.  $\overline{MR} = \max_{\theta} \{P(X^*(\bar{p}^\infty), \theta) + P_q(X^*(\bar{p}^\infty), \theta) \frac{X^*(\bar{p}^\infty)}{n}\} - c$

(ii) Total investment  $X^*$ , investment  $X^{1*}$  in technology  $(k_1, c_1)$ , and total welfare under imperfect competition, remain strictly below the first best solution for any cap  $\bar{p} \geq 0$ .

PROOF See appendix 7. □

Let us first discuss why caps  $\bar{p} \in [\overline{MR}, \bar{p}^\infty]$  lead to increased total investment in the market equilibrium. The equilibrium characterization (3) of theorem 2 reveals that any cap at or above the highest possible marginal revenue increases marginal spot market profits of investment in those spot markets where it is binding, as compared to a situation without cap. In the new market equilibrium with binding cap, however, expected marginal spot market profits must again be equal to marginal cost of investment  $k$  of the peaking technology  $(k, c)$ . This can only be achieved by decreasing marginal spot market profits in scenarios with non-binding cap, i.e. by increasing investment in technology  $(k, c)$ .<sup>28</sup> Notice that the findings of theorem 4 do not allow to characterize the most desired cap from the interval  $[\overline{MR}, \bar{p}^\infty)$ , but just establishes that introduction of any of those caps is desirable.

Furthermore, all those caps lead to an increase of investment in technology  $(k, c)$ , but have no impact on investment  $X^{1*}$  in the base-load technology  $(k_1, c_1)$ . The intuition for this result has already been provided in the context of corollary 1. We are finally able to show that total investment and investment in the base-load technology  $(k_1, c_1)$  remain always strictly below the investment levels of the first best solution.

In the subsequent theorem we provide a characterization of the desirable caps established in theorem 4 in terms of the highest market price reached under perfect competition. Under relatively mild assumptions on the demand function<sup>29</sup> the highest possible market price reached under perfect competition  $\widehat{p}^\infty$  could lend itself as a welfare enhancing cap on scarcity prices:

**THEOREM 5 (CAPPING SCARCITY PRICES AT THE HIGHEST COMPETITIVE PRICE)**

*Suppose demand can be decomposed such that  $P(X, \theta) = a(\theta) + b(\theta)\tilde{P}(X)$ . A cap  $\bar{p} \in [\widehat{p}^\infty, \bar{p}^\infty]$  strictly increases investment in technology  $(k, c)$ , but leaves investment in*

---

<sup>28</sup>This generalizes the intuition obtained for a single deterministic spot market. In the absence of a price cap, marginal profits of the single spot market have to equal marginal cost of investment (i.e.  $P(X^*) - P_q(X^*)/n - c = k$ ). All price caps at or above marginal revenue will increase capacity choice. The optimal price cap in the deterministic framework is suited to implement the first best outcome and is given by marginal revenue, i.e.  $\bar{p} = P(X^*) - P_q(X^*)/n = c + k$ , i.e. the optimal price cap equals to total marginal cost. With fluctuating demand at the spot market this logic can be applied only to the highest demand realizations, typically the first best outcome cannot be reached, however.

<sup>29</sup>These do not seem necessary, but they allow for a very intuitive proof.



technology  $(k_1, c_1)$  unchanged. It leads to an increase of consumer surplus and total welfare, and lowers average prices as compared to the market game without cap.

PROOF See appendix 7. □

Note the implications of this results from a policy maker's perspective. For example in electricity markets, regulators typically have quite detailed information on demand and cost. Thus, calculating  $\widehat{p}^\infty$  by modeling the *competitive benchmark* seems relatively easy, while basically nothing is known for the case of market power. Theorem 5 connects the cases of perfect competition and oligopoly, giving a clear cut policy result.

Let us finally point out that theorems 4 and 5 do not contradict the result of theorem 3, where we established that under perfect competition a distortion of scarcity prices is never desirable. As we approach the case of perfect competition ( $n \rightarrow \infty$ ), the range of desirable caps on scarcity prices vanishes.

## 5 Optimal Cap on Scarcity Prices

As we have shown in section 4, for strategically behaved firms we can always identify a range of binding caps on scarcity prices that strictly increase total equilibrium investment and thus are desirable from a welfare point of view. Interestingly such caps leave investment in the base-load technology  $(k_1, c_1)$  unchanged. We have not yet determined the optimal limit on scarcity prices, which is characterized in the following theorem. In order to do so we need to make regularity assumptions on the distribution of  $\theta$ .<sup>30</sup>

**THEOREM 6 (OPTIMAL CAP ON SCARCITY PRICES)** *Suppose demand can be decomposed such that  $P(X, \theta) = \theta + \tilde{P}(X)$  and the hazard rate  $h(\theta) := \frac{f(\theta)}{1-F(\theta)}$  is increasing.<sup>31</sup> Then there exists a unique optimal cap  $\bar{p}^*$  which is uniquely characterized by:*

$$-P_q(X^*(\bar{p}^*), \tilde{\theta}^{\bar{p}}) \frac{X^*(\bar{p}^*)}{n} = \frac{1 - F(\tilde{\theta}^{\bar{p}})}{f(\tilde{\theta}^{\bar{p}})}. \quad (7)$$

$\bar{p}^*$  is increasing in the number of firms  $n$ .

PROOF See appendix 7. □

---

<sup>30</sup>Recall that in theorem 4 and 5, we could not make statements about monotonicity without further assumptions on the distribution.

<sup>31</sup>In order to prove the theorem it is sufficient to assume  $P_{q\theta} = 0$ , such that demand can be decomposed such that  $P(Q, \theta) = v(\theta) + \tilde{P}(Q)$ . Then, the hazard rate of the transformed random variable  $v(\theta)$  would have to be increasing. In the theorem we present a slightly less general statement for easier exposition.



Let us point out the trade off that has to be solved for the optimal cap on scarcity prices. Recall from expression (3) of theorem 2 that firms choose total investment in order to equate expected marginal revenue of investment with marginal cost. Lowering a given cap affects expected marginal revenue and thus, the firms' investment decisions. We observe two effects: On the one hand, lowering the cap by an increment decreases marginal revenue in all scenarios where the cap has been binding. The expected marginal loss equals  $[1 - F(\tilde{\theta}^{\bar{p}})]$ . On the other hand, the cap becomes binding also in lower demand scenarios and thereby marginally increases expected revenue by  $-P_q(X^*, \tilde{\theta}^{\bar{p}})x_i f(\tilde{\theta}^{\bar{p}})$ .<sup>32</sup> Condition (7), as stated in theorem 6, balances those two effects. Note that the first effect lowers investment incentives, while the second effect encourages investment. At high caps, the first effect is necessarily small (the cap is binding only at few spot markets), such that the second effect always dominates. It depends on the exact distribution of  $\theta$ , whether the optimal cap lies in  $[\overline{MR}, \widetilde{\rho}^\infty)$ , or below, however.

## 6 Quantitative Illustration: Restricting Scarcity Prices in the German Electricity Market

Since the problem of missing money has been intensively discussed in the context of investment decisions in liberalized electricity markets, we now provide some empirical illustration of our results for the case of a specific electricity market. For reasons of data availability we analyze the German electricity market, the quantitative results obtained are unlikely to change substantially for other electricity markets, however. Our aim is to fit the theoretical model as closely as possible to the data of the German Electricity market for the year 2006. As established in Corollary 1, desirable caps do only have an impact on total investment, but not on investment in base-load technologies. We can thus restrict attention to quantify total market investment, according to expression (3) this is determined by the cost of the peak-load technology, given by open cycle gas turbines.

In order to use our theoretical model for the analysis, we chose to make the following specifications. We assume linear fluctuating demand  $P(Q) = \theta - bQ$ . We furthermore determine average production- and investment-cost. We then determine firms' investment choice for different degrees of market power and different price-caps by making use of expressions (3) and (5). In order to assess the robustness of our results, we do not perform

---

<sup>32</sup>Note that marginal revenue in scenario  $\theta$  without binding cap is  $P_q(X^*, \theta)x_i + P(X^*, \theta)$  and thus, a binding cap  $\bar{p} = P(X^*, \tilde{\theta}^{\bar{p}})$  increases marginal revenue in scenario  $\tilde{\theta}^{\bar{p}}$  by  $-P_q(X^*, \tilde{\theta}^{\bar{p}})x_i > 0$ .

the analysis for single parameter values, but rather for plausible ranges of parameter distributions. This concerns the following parameters of the model: The demand elasticity (determined by the slope of the demand function,  $b$ ), marginal cost of production,  $c$ , and marginal investment cost,  $k$  of the peak-load technology. From the possible ranges of those parameters, our algorithm selects one random combination in each iteration. The resulting distributions of investment and welfare differences give an impression of the sensitivity of our results to changes in the parameters. In the following we provide some details on the relevant ranges of our cost and demand parameters.

**Market demand:** To construct fluctuating market demand, we start with hourly market prices (from the European Energy Exchange (EEX)<sup>33</sup>) and hourly quantities consumed (from the Union for the Co-ordination of Transmission of Electricity (UCTE)<sup>34</sup>) for the year 2006. We chose the value of  $b$  in line with other studies on energy markets. Most studies that estimate demand for electricity<sup>35</sup> find short run elasticities between 0.1 and 0.5 and long run elasticities between 0.3 and 0.7. The relevant range of prices is around  $P = 100$  €/MWh and corresponding consumption is approximately  $Q = 50$  GW. In our simulations we thus use a uniform distribution of  $b$  on the interval  $[0.004, 0.007]$ , which corresponds to elasticities between 0.5 and 0.29. In order to illustrate our results concerning welfare maximizing caps on scarcity prices as analyzed in section 5 we fit a shifted Weibull distribution over the observed frequencies.

**Production cost:** The major components of variable production cost are gas prices<sup>36</sup> and prices for  $CO_2$  emission allowances.<sup>37</sup> The average TTF gas price in 2006 was 20 €/MWh and  $CO_2$  permissions traded on average for 9.30 €/MWh.<sup>38</sup> The efficiency of gas turbines currently ranges at around 37, 5%.<sup>39</sup> The resulting daily production cost for the year 2006 was on average 66.30 €/MWh. In order to assess robustness of our results, however, we consider uniformly distributed errors ( $\pm 10\%$ ) and thus consider average marginal cost to be in the interval  $[59.67, 72.93]$  €/MWh.

---

<sup>33</sup>See [www.EEX.com](http://www.EEX.com)

<sup>34</sup>See [www.UCTE.org](http://www.UCTE.org)

<sup>35</sup>See, for example, Lijssen (2006) for an overview of recent contributions on that issue.

<sup>36</sup>Daily values from the Dutch Hub TTF, corrected for transportation cost.

<sup>37</sup>Daily data taken from the EEX. The emission-coefficient for natural gas is set by the German ministry of environment at 56t  $CO_2$ /TJ which corresponds to 0.2016t  $CO_2$ /MWh.

<sup>38</sup>Recall that we do not use the averages but the daily values in our simulation.

<sup>39</sup>See Gas Turbine World Handbook (2006).

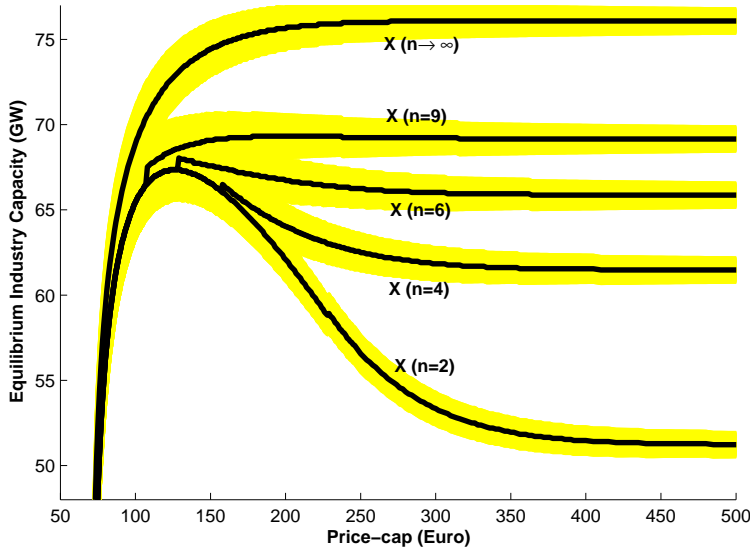


Figure 1: Total Industry investment over caps on scarcity prices for different degrees of concentration. Means and 95%-confidence intervals.

**Investment Cost:** Since we analyze investment incentives based solely on one year, we break down investment cost to annuities.<sup>40</sup> In order to take construction time of gas turbine plants into account, we consider investment cost on the basis of data from the year 2000. We assume perfect foresight, i.e. all cost components have been predicted accurately by the firms at the time of their investment decision. We base investment cost on the following two studies: First, a study on the German energy market commissioned by the German Parliament, with scenarios for investment decisions summarized in Weber and Swider (2004). Second, Energiereport III, a study conducted by the Institute of Energy Economics EWI (2000) for the German Ministry of Economics. Based on those studies we determine a plausible range of relevant annuities given by [28558, 39692] €/MWa.

**Results:** We determine industry capacity choice for different number of firms, and for different caps on scarcity prices. Figure 1 shows mean values (solid black lines) and 95% confidence intervals for industry capacity choice as a function of those caps (for different number of firms). We observe that especially for concentrated markets surprisingly low caps lead to a dramatic increase of total industry investment (in the case of duopoly total

<sup>40</sup>The results will thus only yield a benchmark for current profitability of investment. Provided, however, that yearly demand is increasing over time (and that strategic timing of investment is not an issue) our procedure should yield accurate predictions, even though once installed capacities cannot be removed the subsequent year.

investment is increased by up to 30%). Observe furthermore that, in line with the findings of theorem 3 (and the existing literature, e.g. see Joskow (2007)), for the case of perfect competition any binding cap leads to reduced industry capacity. We observe that caps below 260 €/MWh increase industry capacity for any market structure based on up to 20 firms.

We then analyze the impact of limiting scarcity prices on overall welfare. The left graph of

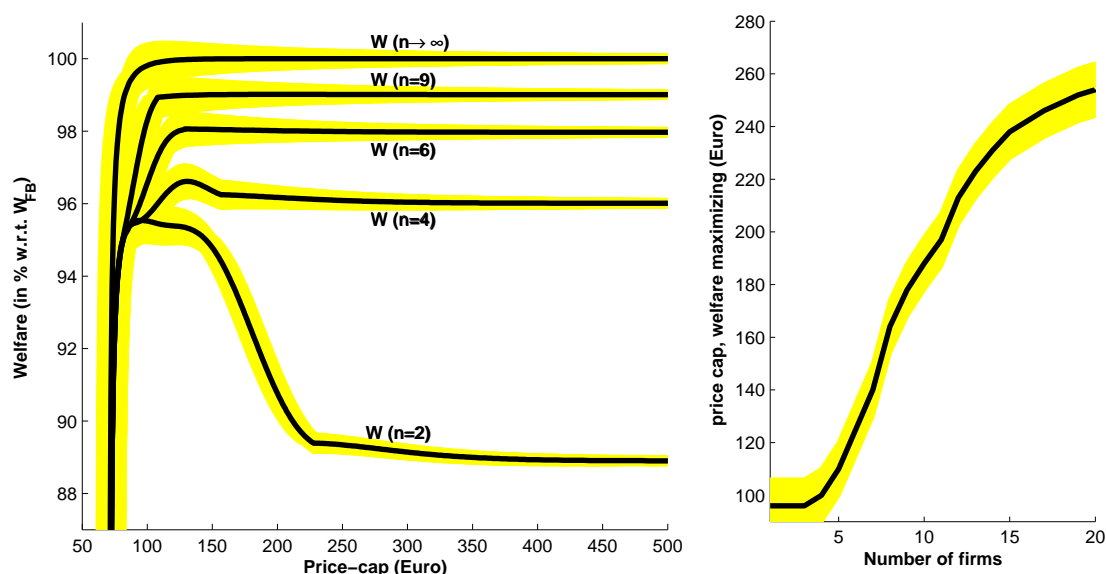


Figure 2: Left: Welfare over caps on scarcity prices for different degrees of concentration. Right: Optimal caps over market concentration. Means and 95%-confidence intervals.

figure 2 shows mean values (solid black lines) and 95% confidence intervals for Welfare<sup>41</sup> as a function of caps on scarcity prices (for different number of firms). The right graph of figure 2 gives welfare maximizing caps as a function of the number of firms in the industry. We observe that caps below 260 €/MWh increase industry capacity for any market structure based on up to 20 firms. For concentrated markets we find surprisingly low optimal caps at around 110 €/MWh for the case of up to 4 firms in the industry. The empirical results show that the impact of choosing appropriate caps on scarcity prices can be substantial, the range of caps having those beneficial properties is rather large.

<sup>41</sup>Notice that all values are normalized with respect to welfare generated in the first best scenario, which is given by roughly 100 bn €.

## 7 Conclusion

In this article we have analyzed the impact of restricting scarcity prices in markets with fluctuating demand and its implications for investment decisions of strategic firms. Especially during hours of high demand and tight supply many of those markets are suspected to be subject to the exercise of market power (with a prominent example being electricity markets). Caps on scarcity prices have been proposed and are currently used in order to mitigate this exercise of market power. In the short run, when firms' investment decisions can be assumed to remain fixed, there is undisputed consensus regarding the impact of such caps. They reduce firms' incentives to withhold production and are thus considered to be a powerful tool to discipline market power. As an increasing number of experts in the field points out there seems to be a trade-off to be solved when also taking into account firms' investment incentives in the long run. In a competitive market scarcity prices, which result during the hours of high demand and tight supply, convey optimal investment signals to firms. Restricting those price spikes leads to the so called problem of missing money, and to suboptimal investment activity of firms (see Hogan (2005), or Joskow (2007)).

It has been the purpose of the present article, to contribute to this discussion, by precisely modeling investment behavior of strategic firms. The rationale for doing so is straight forward: If firms are suspected to behave strategically at the spot markets (being the sole justification for the currently active caps in those markets), then it seems natural to consider firms to also behave strategically when making their investment decisions. We furthermore established that a meaningful contribution to the discussion of missing money necessarily has to take into account fluctuating demand at several spot markets.<sup>42</sup>

In the present article we thus did analyze investment decisions of strategic firms in two different technologies (i.e. base-load and peak-load) for the case of fluctuating demand<sup>43</sup> at several spot markets. In section 3 we first characterized the market equilibrium for investment in both technologies for a given cap on scarcity prices and derived its existence and uniqueness. We then showed in section 4 that appropriately chosen caps always lead to an increase of investment in the peak-load technology  $(k, c)$ , leaving investment in the base-load technology  $(k_1, c_1)$  unaffected. We characterized them and also derived the optimal cap. That is, if firms are suspected to have market power in a market with fluctuating demand and a cap on scarcity prices is put in place to mitigate the exercise of market

---

<sup>42</sup>Also compare footnote 9.

<sup>43</sup>Firms could furthermore be uncertain regarding the precise pattern of fluctuations at the time of making their investment.

power at the spot market, then an appropriately chosen cap does not lead to a reduction of firms' investment incentives. Unlike suggested in the current discussion of missing money and suppressed scarcity rents, this implies that for strategic investment choice there is no trade-off to be solved between the long run and the short run perspective.

Since the problem of missing money has been discussed extensively in the context of investment decisions in recently liberalized electricity markets, we then quantified desirable and optimal caps on scarcity prices for the case of a specific electricity market (the German market). We found a surprisingly large interval of caps leading to a substantial increase of total investment for almost any market structure. Those results confirmed the relevance of our framework for the currently rather intensive discussion of proper design of caps on scarcity prices in those markets.<sup>44</sup>

## References

- Biglaiser, G. and M. Riordan (2000). Dynamics of price regulation. *Rand Journal of Economics* 31, 744–767.
- Borenstein, S., J. Bushnell and F. Wolak (2002). Measuring Market Inefficiencies in Californias Restructured Wholesale Electricity Market. *American Economic Review*, 92, 1376–1405.
- Cowan, S. (2002). Price cap regulation. *Swedish Economic Policy Review* 9, 167–188.
- Crew, M., C. Fernando, and P. Kleindorfer (1995). The Theory of Peak-Load Pricing: A Survey. *Journal of Regulatory Economics* 8, 215 - 248.
- Dixit, A. (1991). Irreversible Investment with Price Ceilings. *Journal of Political Economy* 99, 541–557.
- Dobbs, I. (2004). Intertemporal price cap regulation under uncertainty. *Economic Journal* 114, 421–440.

---

<sup>44</sup>In order to encourage investment in new generation facilities, the cap on scarcity prices in the Californian electricity market has very recently been adjusted from originally 250\$/MWh to 2500\$/MWh, see FERC(2000b) and FERC(2009). In our quantitative example we found that caps around 260/*MWh* lead to an increase of total investment and overall welfare for any number of firms with  $n \leq 20$ . Clearly in the present version our quantitative results are mainly illustrative and a more detailed analysis based on our framework would be needed in order to allow for policy recommendations.

- EWI (2000). Energy Report III, Study commissioned by the German Ministry of Economics and Technology, in Schäfer Pöschel Press, Stuttgart.
- FERC (2000). Federal Energy Regulatory Commissions Order, 92 FERC 61,073.
- FERC (2000b). Federal Energy Regulatory Commissions Order, 92 FERC 61,112.
- FERC (2009). Federal Energy Regulatory Commissions Order, 127 FERC 61,156.
- Gas Turbine World Handbook (2006). Vol. 25, Pequot Publishing, Inc.
- von der Fehr, N.-H. and D. Harbord (1993). Spot Market Competition in the UK Electricity Industry. *Economic Journal* 103 (418), 531 – 46.
- Green, R.J. and D. Newbery (1992). Competition in the British Electricity Spot Market. *Journal of Political Economy* 100 (5), 929 – 53.
- Hogan, W. (2005). On an Energy Only Electricity Market Design for Resource Adequacy. Harvard working paper series.
- Joskow, P. , and J. Tirole (2007). Reliability and competitive electricity markets. *Rand Journal of Economics*, 38, 60–84.
- Joskow, P. (2007). Competitive Electricity Markets and Investment in New Generating Capacity. The New Energy Paradigm (D. Helm, Editor). *Oxford Univ. Press*.
- Kreps, D. and J. Scheinkman (1983). Quantity Precommitment and Bertrand Competition Yield Cournot Outcomes. *Bell Journal of Economics*, 14, 326–337.
- Laffont, J.-J. and J. Tirole (1993). A Theory of Incentives in Procurement and Regulation. MIT Press.
- Lijzen, M. (2006). The Real-Time Price Elasticity of Electricity. *Energy Economics*, 29, 249–258.
- Pint, E. (1992). Price cap versus rate-of-return regulation in a stochastic cost model. *Rand Journal of Economics* 23, 564–578.
- Roques, F. and N. Savva (2006). Price Cap Regulation and Investment Incentives under Demand Uncertainty. University of Cambridge working paper CWPE 0636.
- Walsh, C. (2001). Worst may be Ahead for West Coast Power Market, *Dow Jones Newswires*, May 23.

Weber, C. and D. Swider (2004). Power Plant Investments Under Fuel and Carbon Price Uncertainty, mimeo.

Wolfram, C. (1999). Measuring Duopoly Power in the British Electricity Spot Market. *American Economic Review*, 89, 805–26.

Zöttl, G. (2008). A Framework of Peak–Load Pricing with Strategic Firms. *CORE working paper series*, 2008/41.

## Appendix

### Proof of Theorem 1

**I) Spot Market Equilibria.** We first determine the equilibrium which obtains at a given spot market  $\theta$  and given investment choices  $x$ . Note that we have to consider also asymmetric investments, w.l.g. we order firms according to their investment levels, i. e.  $x_1 \leq x_2 \leq \dots \leq x_n$ . Depending on their investment–choices and on demand at the spot market  $\theta$ , firms produce either the unconstrained Cournot–output, or are capacity constrained by their investment decision. Consider a spot market  $\theta$ , where firms  $i = 1, \dots, m$  are capacity constrained, producing their invested capacities  $x_i$ . Firms  $i = m + 1, \dots, n$  produce their unconstrained Cournot output  $y$ . We first define the range of spot markets  $\theta \in [\tilde{\theta}_m^X, \tilde{\theta}_{m+1}^X]$ , where this is indeed the case. Critical demand realizations  $\tilde{\theta}_m^X$  solve:

$$P\left(\sum_{i=1}^m x_i + (n-m)x_m, \tilde{\theta}_m^X\right) + P_q\left(\sum_{i=1}^m x_i + (n-m)x_m, \tilde{\theta}_m^X\right) x_m = c.$$

Output  $y$  of unconstrained firms at the spot markets  $\theta \in [\tilde{\theta}_m^X, \tilde{\theta}_{m+1}^X]$  solves:

$$P\left(\sum_{i=1}^m x_i + (n-m)y, \theta\right) + P_q\left(\sum_{i=1}^m x_i + (n-m)y, \theta\right) \tilde{q}_i^{Cm} = c.$$

We denote the spot market output of each player by  $\tilde{q}_i(x, \theta) = \min\{x_i, y\}$ , industry output is then given by  $\tilde{Q}(x, \theta) = \sum_{i=1}^n \tilde{q}_i$ . Profit of firm  $i$  in a spot market  $\theta \in [\tilde{\theta}_m^X, \tilde{\theta}_{m+1}^X]$  is given as follows:

$$\tilde{\pi}_i^m = \left(P(\tilde{Q}(x, \theta), \theta) - c\right) \tilde{q}_i(x, \theta).$$

If demand at a spot market is high and capacity is binding, prices will rise, leading to scarcity prices. We consider caps designed to limit those scarcity prices. We denote by  $\tilde{\theta}^{\bar{p}}(x) = \{\theta : \bar{p} = P(X, \theta)\}$  the spot market realization, where the cap starts to be binding. Profits of firm  $i$  in spot market  $\theta \in [\tilde{\theta}^{\bar{p}}, \bar{\theta}]$  are given by:

$$\tilde{\pi}_i^{\bar{p}} = (\bar{p} - c) \tilde{q}_i(x, \theta).$$



The results obtained for the spot market equilibria enable us to derive firm  $i$ 's profit when investing  $x_i$  in technology  $(k, c)$ , given that the other firms invest  $x_{-i}$  and output at the spot markets is given by  $\tilde{q}(x, \theta)$  for  $\theta \in [\tilde{\theta}_m^X, \tilde{\theta}_{m+1}^X]$ . A firm's profit from given levels of investments,  $x$ , is the integral over equilibrium profits at each  $\theta$  given  $x$  on the entire domain of spot market realizations  $[\underline{\theta}, \bar{\theta}]$ . In order to simplify the exposition we define  $\tilde{\theta}_0^X := \underline{\theta}$  and  $\tilde{\theta}_{n+1}^X := \bar{\theta}$ . Then, the profit of firm  $i$  can be written as follows

$$\pi_i(x) = \sum_{m=0}^{m=n} \int_{\tilde{\theta}_m^X}^{\tilde{\theta}_{m+1}^X} \tilde{\pi}_i^m(x, \theta) dF(\theta) + \int_{\tilde{\theta}^{\bar{p}}}^{\bar{\theta}} \tilde{\pi}_i^{\bar{p}}(x, \theta) dF(\theta) - kx_i. \quad (8)$$

Note that at each critical value  $\tilde{\theta}_m^X$ ,  $m = 1, \dots, n$  it holds that  $\tilde{\pi}_i^{m-1}(x, \tilde{\theta}_m^X) = \tilde{\pi}_i^m(x, \tilde{\theta}_m^X)$ . Thus,  $\pi_i(x)$  is continuously differentiable. Differentiating  $\pi_i(x)$  yields<sup>45</sup>

$$\frac{d\pi_i(x, \tilde{q})}{dx_i} = \sum_{m=i}^n \int_{\tilde{\theta}_m^X(x)}^{\tilde{\theta}_{m+1}^X(x)} \frac{d\tilde{\pi}_i^m(x, \theta)}{dx_i} dF(\theta) + \int_{\tilde{\theta}^{\bar{p}}(x)}^{\bar{\theta}} \frac{d\tilde{\pi}_i^{\bar{p}}(x, \theta)}{dx_i} dF(\theta) - k \quad (9)$$

**II) Existence Symmetric Equilibrium.** In the following we show that a symmetric equilibrium of the investment game exists, and that equilibrium choices  $x_i^* = \frac{1}{n}X^*$ ,  $i = 1, \dots, n$ , are defined by equation (2). For this purpose it is sufficient to show quasiconcavity of firm  $i$ 's profit given the other firms invest  $x_{-i}^*$ ,  $\pi_i(x_i, x_{-i}^*)$ , which we do in the following. Note that  $\pi_i(x_i, x_{-i}^*)$  is defined piecewisely. For  $x_i \leq x_i^*$ , we have to examine the profit of firm 1 (by convention the lowest capacity firm) given that  $x_2 = x_3 = \dots = x_n$ . Since this implies that  $\tilde{\theta}_2^X = \dots = \tilde{\theta}_n^X$  and thus it follows from (8) that

$$\begin{aligned} \pi_1(x_1, x_{-1}^*) &= \int_{\underline{\theta}}^{\tilde{\theta}_1^X(x)} \tilde{\pi}_1^0(x, \theta) dF(\theta) + \int_{\tilde{\theta}_1^X(x)}^{\tilde{\theta}_n^X(x)} \tilde{\pi}_1^1(x, \theta) dF(\theta) \\ &\quad + \int_{\tilde{\theta}_n^X(x)}^{\tilde{\theta}^{\bar{p}}} \tilde{\pi}_1^n(x, \theta) dF(\theta) + \int_{\tilde{\theta}^{\bar{p}}}^{\bar{\theta}} \tilde{\pi}_1^{\bar{p}}(x, \theta) dF(\theta) - kx_1 \end{aligned} \quad (10)$$

For  $x_i \geq x_i^*$ , the profit of firm  $i$  is the profit of the highest capacity firm (firm  $n$  according to our convention), given all other firm have invested the same, i. e.  $x_1 = \dots = x_{n-1}$ .

$$\begin{aligned} \pi_n(x_n, x_{-n}^*) &= \int_{\underline{\theta}}^{\tilde{\theta}_{n-1}^X(x)} \tilde{\pi}_n^0(x, \theta) dF(\theta) + \int_{\tilde{\theta}_{n-1}^X(x)}^{\tilde{\theta}_n^X(x)} \tilde{\pi}_n^{n-1}(x, \theta) dF(\theta) \\ &\quad + \int_{\tilde{\theta}_n^X(x)}^{\tilde{\theta}^{\bar{p}}} \tilde{\pi}_n^n(x, \theta) dF(\theta) + \int_{\tilde{\theta}^{\bar{p}}}^{\bar{\theta}} \tilde{\pi}_n^{\bar{p}}(x, \theta) dF(\theta) - kx_n \end{aligned} \quad (11)$$

---

<sup>45</sup>Note that continuity of  $\tilde{\pi}_i$  implies that due to Leibnitz' rule the derivatives of the integration limits cancel out. Moreover,  $\tilde{\pi}_i^m$  only changes in  $x_i$  if firm  $i$  is constrained, i. e.  $i \leq m$ . Thus, the sum does not include the cases where firm  $i$  is unconstrained, i. e.  $m < i$ .

(i) For  $x_i \geq x_i^*$ , the second derivative of the profit function  $\pi_n$  is given by:

$$\frac{d^2\pi_n}{(dx_n)^2} = -\frac{d\tilde{\theta}_n^X(x)}{dx_n} \underbrace{\left[ \frac{d\tilde{\pi}_n^n(x, \tilde{\theta}_n^X)}{dx_n} \right]}_{=0 \text{ (} x_n \text{ is opt. at } \tilde{\theta}_n^X)} f(\tilde{\theta}_n^X) + \int_{\tilde{\theta}_n^X(x)}^{\tilde{\theta}^P} \underbrace{\frac{d^2\tilde{\pi}_n^n(x, \theta)}{(dx_n)^2}}_{<0 \text{ by A1 part (iv)}} f(\theta) d\theta < 0. \quad (12)$$

Note that the first term cancels out and the second term is negative by concavity of the spot market profit function (implied by assumption 1). We find that for  $x_i \geq x_i^*$ ,  $\pi_i(x_i, x_{-i}^*)$  is concave, which implies that upwards deviations are not profitable.

(ii) For  $x_i \leq x_i^*$ , the second order condition is more difficult to analyze since the profit function  $\pi_1(x_1, x_{-1}^*)$  is not concave. We can, however, show quasiconcavity of  $\pi_1(x_1, x_{-1}^*)$ . For this purpose we need property 1 in order to complete the proof of existence (part I). We can show quasiconcavity of  $\pi_1(x_1, x_{-1}^*)$  by showing that

$$\frac{d\pi_1(x_1^0, x_{-1}^*)}{dx_1} > \frac{d\pi_1(x_1^*, x_{-1}^*)}{dx_1} = 0 \quad \text{for all } x_1^0 < x_1^*.$$

This holds true, since [compare also equation (9)]

$$\begin{aligned} \frac{d\pi_1(x_1^0, x_{-1}^*)}{dx_1} &= \underbrace{\int_{\tilde{\theta}_1^X(x_1^0, x_{-1}^*)}^{\tilde{\theta}_1^X(x_1^0, x_{-1}^*)} \frac{d\tilde{\pi}_1^n(x_1^0, x_{-1}^*, \theta)}{dx_1} dF(\theta)}_{\geq 0 \text{ by property 1, part (i)}} + \int_{\tilde{\theta}_n^X(x_1^0, x_{-1}^*)}^{\tilde{\theta}^P} \frac{d\tilde{\pi}_1^n(x_1^0, x_{-1}^*, \theta)}{dx_1} dF(\theta) \\ &\geq \int_{\tilde{\theta}_n^X(x_1^0, x_{-1}^*)}^{\tilde{\theta}^P} \frac{d\tilde{\pi}_1^n(x_1^0, x_{-1}^*, \theta)}{dx_1} dF(\theta) \\ &= \underbrace{\int_{\tilde{\theta}_n^X(x_1^0, x_{-1}^*)}^{\tilde{\theta}_n^X(x_{-1}^*, x_{-1}^*)} \frac{d\tilde{\pi}_1^n(x_1^0, x_{-1}^*, \theta)}{dx_1} dF(\theta)}_{\geq 0 \text{ by property 1, part (ii)}} \\ &\quad + \underbrace{\int_{\tilde{\theta}_n^X(x_1^*, x_{-1}^*)}^{\tilde{\theta}^P} \left[ \frac{d\tilde{\pi}_1^n(x_1^0, x_{-1}^*, \theta)}{dx_1} - \frac{d\tilde{\pi}_1^n(x_1^*, x_{-1}^*, \theta)}{dx_1} \right] dF(\theta)}_{>0 \text{ by property 1, part (ii)}} \\ &\quad + \underbrace{\int_{\tilde{\theta}_n^X(x_1^*, x_{-1}^*)}^{\tilde{\theta}^P} \frac{d\tilde{\pi}_1^n(x_1^*, x_{-1}^*, \theta)}{dx_1} dF(\theta)}_{= \frac{d\pi_1(x_1^*, x_{-1}^*)}{dx_1} = 0 \text{ [recall that } \tilde{\theta}_1^X(x_1^*) = \tilde{\theta}_n^X(x_1^*)]} \geq 0. \end{aligned}$$

To summarize, in part II) (i) and (ii) we have shown that  $\pi_i(x_i, x_i^*)$  is quasiconcave. We conclude that the first order condition given in theorem 1 indeed characterizes equilibrium investment in technology  $(k, c)$ .

PROPERTY 1 [PROPERTIES OF MARGINAL PROFITS AT THE SPOT MARKETS] *Suppose all firms but firm 1 have invested symmetric capacities summarized in the vector  $x_{-1}^0$ . Firm 1 has invested  $x_1$ , less than each of the other firms. We obtain:*

$$(i) \frac{d\tilde{\pi}_1^1(x_1^0, x_{-1}^0, \theta)}{dx_1} \geq 0 \text{ for } \tilde{\theta}_1^X \leq \theta \leq \tilde{\theta}_n^X.$$

$$(ii) \frac{d\tilde{\pi}_1^n(x_1', x_{-1}^0, \theta)}{dx_1} \geq \frac{d\tilde{\pi}_1^n(x_1'', x_{-1}^0, \theta)}{dx_1} \geq 0 \text{ for } x_1' < x_1'', \tilde{\theta}_n^X \leq \theta \leq \bar{\theta}.$$

PROOF (i) The first part holds due to the fact in case firm 1 is constrained, i. e. ( $\theta \geq \tilde{\theta}_1^X$ ), firm 1 would like to produce more than  $x_1$  for all demand realizations  $\theta \geq \tilde{\theta}_1^X$ , which, however, is not possible due to the capacity constraint.

(ii) The first inequality follows from concavity of the profit functions in the spot markets, which is implied by assumption 1. Thus, the first order condition at each spot-market is decreasing in  $x_1$  until  $\tilde{q}_1$ , which immediately yields the first inequality of part (ii). The second inequality is due to the fact that in case all firms are constrained, i. e. ( $\theta \in [\tilde{\theta}_n^X, \infty]$ ), firm 1 would like to produce more for all demand realizations  $\theta$  (which is not possible because it is constrained).  $\square$

**III) Uniqueness of the Equilibrium.** We now show that (i)  $x^*$  is the unique symmetric equilibrium and (ii) that there are no asymmetric equilibria.

**(i)  $x^*$  is the unique symmetric equilibrium.** If capacities are equal, i. e.  $x_1^0 = x_2^0 = \dots = x_n^0$ , we have

$$\frac{d\pi_i(x^0)}{dx_i} = \int_{\tilde{\theta}_n^X(x^0)}^{\tilde{\theta}^{\bar{p}}} [P(n x_i^0, \theta) + P_q(n x_i^0, \theta) x_i^0 - c] f(\theta) d\theta + \int_{\tilde{\theta}^{\bar{p}}}^{\bar{\theta}} [\bar{p} - c] f(\theta) d\theta - k.$$

Differentiation yields<sup>46</sup>

$$\frac{d^2\pi_i(x^0)}{(dx_i)^2} = \int_{\tilde{\theta}_n^X(x^0)}^{\tilde{\theta}^{\bar{p}}} [(n+1)P_q(n x_i^0, \theta) + n P_{qq}(n x_i^0, \theta) x_i^0 - C_{qq}(x_i^0, \theta)] dF(\theta) - K_{xx}(x_i^0) < 0,$$

which is negative due to assumption 1. Thus, since  $\frac{d\pi_i(x^*)}{dx_i} = 0$  and moreover  $\pi_i(x)$  is concave along the symmetry line, no other symmetric equilibrium can exist.

**(ii) There cannot exist an asymmetric equilibrium.** Any candidate for an asymmetric equilibrium  $\hat{x}$  can be ordered such that  $\hat{x}_1 \leq \hat{x}_2 \leq \dots \leq \hat{x}_n$ , where at least one inequality has to hold strictly. This implies  $\hat{x}_1 < \hat{x}_n$ . The profit of firm  $n$  can be obtained by setting  $i = n$  in equation (8), and the first derivative is given by

$$\frac{d\pi_n}{dx_n} = \int_{\tilde{\theta}_n^X(x)}^{\tilde{\theta}^{\bar{p}}} \frac{d\tilde{\pi}_n^n(x, \theta)}{dx_n} f(\theta) d\theta + \int_{\tilde{\theta}^{\bar{p}}}^{\bar{\theta}} (\bar{p} - c) f(\theta) d\theta - k.$$

It is easy to show that firm  $n$ 's profit function is concave by examination of the second derivative [see equation (12)]. Thus, any asymmetric equilibrium  $\hat{x}$ , if it exists, must satisfy

<sup>46</sup>Differentiation works as in (12).

$\frac{d\pi_n(\hat{x})}{dx_n} = 0$ . We now show that whenever it holds that  $\frac{d\pi_n(\hat{x})}{dx_n} = 0$ , firm 1's profit is increasing in  $x_1$  at  $\hat{x}$  (which implies that no asymmetric equilibria exist).

From equation (9) it follows that the first derivative of firm 1's profit function is given by

$$\frac{d\pi_1}{dx_1} = \int_{\tilde{\theta}_1^X(x)}^{\tilde{\theta}_2^X(x)} \frac{d\tilde{\pi}_1^1(x, \theta)}{dx_1} f(\theta) d\theta + \dots + \int_{\tilde{\theta}_n^X(x)}^{\tilde{\theta}^{\bar{p}}} \frac{d\tilde{\pi}_1^n(x, \theta)}{dx_1} f(\theta) d\theta + \int_{\tilde{\theta}^{\bar{p}}}^{\bar{\theta}} (\bar{p} - c) dF(\theta) - k.$$

Note that all the integrals in  $\frac{d\pi_1}{dx_1}$  are positive since firm 1 is constrained at all demand realizations and therefore would want to increase its production. Thus, we have

$$\frac{d\pi_1}{dx_1} > \int_{\tilde{\theta}_n^X(x)}^{\tilde{\theta}^{\bar{p}}} \frac{d\tilde{\pi}_1^n(x, \theta)}{dx_1} dF(\theta) + \int_{\tilde{\theta}^{\bar{p}}}^{\bar{\theta}} (\bar{p} - c) dF(\theta) - k,$$

where the RHS are simply the last three terms of  $\frac{d\pi_1}{dx_1}$ . Now we can conclude that

$$\frac{d\pi_1(\hat{x})}{dx_1} > \int_{\tilde{\theta}_n^X(\hat{x})}^{\tilde{\theta}^{\bar{p}}} \frac{d\tilde{\pi}_1^n(\hat{x}, \theta)}{dx_1} dF(\theta) + \int_{\tilde{\theta}^{\bar{p}}}^{\bar{\theta}} (\bar{p} - c) dF(\theta) - k > \int_{\tilde{\theta}_n^X(\hat{x})}^{\tilde{\theta}^{\bar{p}}} \frac{d\tilde{\pi}_n^n(\hat{x}, \theta)}{dx_n} dF(\theta) + \int_{\tilde{\theta}^{\bar{p}}}^{\bar{\theta}} (\bar{p} - c) dF(\theta) - k = 0.$$

The last equality is due to the fact that this part is equivalent to the first order condition of firm  $n$ , which is satisfied at  $\hat{x}$  by construction. To summarize, we have shown that  $\frac{d\pi_1}{dx_1} > 0$ , which implies that there exist no asymmetric equilibria, since at any equilibrium candidate, firm 1 has an incentive to increase its capacity.

## Proof of Theorem 2

**I) Spot Market Equilibria.** For given investment choice  $(x^1, x)$ , the spot market equilibria are determined analogously to appendix 7, when analyzing the case of investment in a single technology. Notice, however, that for the case of investment in two technologies, marginal cost of production are not constant, but given by a step function as characterized by expression 1. Additionally to the critical spot market realizations introduced in appendix 7, we have to define those demand realizations, where investment in technology  $(k_1, c_1)$  starts and stops to be binding for each firm. We denote by  $\tilde{\theta}_l^{X1}$  the demand realization, when firm  $l = 1, \dots, n$  starts to produce the constrained quantity  $x_l^1$  at marginal cost  $c_1$ . We denote by  $\tilde{\theta}_l^X$  the demand realization, when firm  $l$  stops to produce at marginal cost  $c_1$  of the base-load technology, and decides to produce at marginal cost  $c$  of the peak-load technology. Formally, those critical spot market realizations are defined as follows:

$$\begin{aligned} \tilde{\theta}_l^{X1} : \quad & P \left( \sum_{i=1}^l x_i^1 + (n-l)x_l^1, \tilde{\theta}_l^{X1} \right) + P_q \left( \sum_{i=1}^l x_i^1 + (n-l)x_l^1, \tilde{\theta}_l^{X1} \right) x_l = c_1 \\ \tilde{\theta}_l^X : \quad & P \left( \sum_{i=1}^l x_i^1 + (n-l)x_l^1, \tilde{\theta}_l^X \right) + P_q \left( \sum_{i=1}^l x_i^1 + (n-l)x_l^1, \tilde{\theta}_l^X \right) x_l = c. \end{aligned}$$

That is, for  $\theta \in [\underline{\theta}, \tilde{\theta}_i^{X1}]$  firm  $i$  produces its unconstrained Cournot output at marginal cost of the base-load technology  $c_1$ . For  $\theta \in [\tilde{\theta}_i^{X1}, \tilde{\theta}_i^X]$  firm  $i$  produces  $x_i^1$ , the amount invested

in the base-load technology, still at marginal cost  $c_1$ . For  $\theta \in [\tilde{\theta}_i^X, \bar{\theta}_i^X]$  firm  $i$  produces again the unconstrained Cournot output, at marginal cost of the peak-load technology  $c$ , however. And finally for  $[\bar{\theta}_i^X, \bar{\theta}]$ , firm  $i$  produces at its overall capacity bound. Analogous to expression (9), we can now determine first order conditions for firms' investment decisions. For each player  $i$ , the derivative wrt  $x_i^1$  (Investment in technology  $(k_1, c_1)$ ) is given by: <sup>47</sup>

$$\frac{d\pi_i(x^1, x)}{dx_i^1} = \sum_{l=i}^n \int_{\tilde{\theta}_i^{X^1}}^{\tilde{\theta}_{l+1}^{X^1}} \frac{d\tilde{\pi}_i^l(x^1, \theta)}{dx_i^1} dF(\theta) + \sum_{l=1}^i \int_{\tilde{\theta}_i^X}^{\tilde{\theta}_{l+1}^X} \frac{d\tilde{\pi}_i^l(x^1, \theta)}{dx_i} dF(\theta) + \int_{\tilde{\theta}_i^X}^{\bar{\theta}} (c - c_1) dF(\theta) - (k_1 - k) \quad (13)$$

For each player  $i$ , the derivative wrt  $x_i$  (Investment in technology  $(k, c)$ ) is given by:<sup>48</sup>

$$\frac{d\pi_i(x^1, x)}{dx_i} = \sum_{m=i}^n \int_{\tilde{\theta}_m^X}^{\tilde{\theta}_{m+1}^X} \frac{d\tilde{\pi}_i^m(x, \theta)}{dx_i} dF + \int_{\tilde{\theta}_i^{\bar{p}}}^{\bar{\theta}} \frac{d\tilde{\pi}_i^{\bar{p}}(x, \theta)}{dx_i} dF(\theta) - k \quad (14)$$

We observe that for the case of investment in two technologies, the first order condition of investment in technology  $(k, c)$  remains unchanged. That is, expressions (14) and (9) are identical, independently of firms' possibility to also invest in the second technology  $(k_1, c_1)$ . We can thus directly apply the results derived in appendix 7, and conclude that any market equilibrium where firms can invest in two technologies gives rise to unique and symmetric total investment  $x^*$ , as characterized in expression (3). In order to derive the entire market equilibrium for the case of two technologies, we now have to analyze firms' equilibrium investment in technology  $(k_1, c_1)$ .

**II) Existence Symmetric Equilibrium** The symmetric equilibrium candidate  $(x^*, x^{1*})$  is characterized by the first order conditions (3 ) and (4) respectively. As already established, firms have no incentive to deviate from total investment choice  $X^*/n$ . We have to verify, however, if some firm  $i$  has an incentive to deviate from the candidate investment  $x_i^{1*} = X^{1*}/n$  in technology  $(k_1, c_1)$ .

**(i) We first verify downwards deviations from the candidate, i.e.  $x_i^1 \leq X^{1*}/n$ :** The derivative of firm  $i$ 's profits for  $x_i^1 \leq X^{1*}/n$  is given as follows:

$$\begin{aligned} \frac{d\pi_i(x_i^1, x_{-i}^{1*})}{dx_i^1} &= \int_{\tilde{\theta}_i^{X^1}}^{\tilde{\theta}_{-i}^{X^1}} [P(x_i, BR_{-i}(x_i^1), \theta) + (1 + BR'_{-i}(x_i^1)) P_q(x_i^1, BR_{-i}(x_i^1), \theta) x_i^1 - c_1] dF(\theta) + \\ &\int_{\tilde{\theta}_{-i}^{X^1}}^{\tilde{\theta}_i^X} [P(x_i^1 + X_{-i}^{1*}, \theta) + P_q(x_i^1 + X_{-i}^{1*}, \theta) x_i^1 - c_1] dF(\theta) + \int_{\tilde{\theta}_i^X}^{\bar{\theta}} (c - c_1) dF(\theta) - (k_1 - k) \end{aligned} \quad (15)$$

<sup>47</sup>For a condensed representation of the first derivative, define  $\tilde{\theta}_{n+1}^{X^1} := \tilde{\theta}_1^X$  and  $\tilde{\theta}_{i+1}^X := \tilde{\theta}_i^X$ .

<sup>48</sup>As in expression (9), for condensed representation, define  $\tilde{\theta}_{n+1}^X := \tilde{\theta}^{\bar{p}}$ .

The critical demand realizations have been defined above. They characterize those spot markets, where firm  $i$  (or all other firms  $-i$ ) start to produce at the capacity bound of technology  $(k_1, c_1)$  (i.e.  $\tilde{\theta}_i^{X^1}$  and  $\tilde{\theta}_{-i}^{X^1}$ ). At the demand realization  $\tilde{\theta}_i^X$  firm  $i$  starts to produce with technology  $(k, c)$ . Since we analyze a downwards deviation, firm  $i$  will be constrained for lower demand realizations than its rivals, which play the unconstrained Cournot equilibrium for  $\theta \in [\tilde{\theta}_i^{X^1}, \tilde{\theta}_{-i}^{X^1}]$  (the first integral of expression (15)). The unconstrained Cournot spot market output of all firms other than  $i$  in those scenarios is denoted by  $BR_{-i}(x_i^1)$ .

Comparing the first order condition of firm  $i$  obtained for the case of downwards deviation with the one obtained in the symmetric candidate given by (4) we can establish:

$$0 = \frac{d\pi_i(X^{1*})}{dx_i^1} < \frac{d\pi_i(x_i^1, x_{-i}^{1*})}{dx_i^1}.$$

This result is driven by the following observations:

- The integrand of the first integral of expression (15), is strictly positive, since by definition, marginal profits of firm  $i$  are strictly positive whenever firm  $i$  is capacity constrained. Since furthermore  $\tilde{\theta}_{-i}^{X^1} < \tilde{\theta}_i^{X^1} < \tilde{\theta}^{X^1}$ , the integrand of the corresponding demand realizations  $\theta$  equals to zero.
- The integrand (marginal spot market profits) of the second integral of expression (15) is decreasing in output, since profits are concave in output (see assumption 1). A downwards deviation  $x_i \leq X^{1*}/n$  implies thus a point wise increase for each corresponding demand realization  $\theta$ , as compared to expression (4).

We can conclude that downwards deviation from the equilibrium candidate is never profitable for firm  $i$ .

**(ii) Analyze upwards deviations from the candidate equilibrium, i.e.  $x_i^1 \geq X^{1*}/n$ :**  
The derivative of firm  $i$ 's profits for  $x_i^1 \geq X^{1*}/n$  is given as follows:

$$\begin{aligned} \frac{d\pi_i(x_i^1, x_{-i}^{1*})}{dx_i^1} &= \int_{\tilde{\theta}_i^{X^1}}^{\tilde{\theta}_{-i}^X} [P(x_i^1 + X_{-i}^{1*}, \theta) + P_q(x_i^1 + X_{-i}^{1*}, \theta) x_i^1 - c_1] dF(\theta) - (k_1 - k) \\ &\int_{\tilde{\theta}_{-i}^X}^{\tilde{\theta}_i^X} [P(x_i, BR_{-i}(x_i^1)\theta) + (1 + BR'_{-i}(x_i^1)) P_q(x_i^1, BR_{-i}(x_i^1)\theta) x_i^1 - c_1] dF(\theta) + \int_{\tilde{\theta}_i^X}^{\bar{\theta}} (c - c_1) dF(\theta) \end{aligned} \quad (16)$$

The critical demand realizations have been defined above. For the spot markets  $\theta \in [\tilde{\theta}_i^{X^1}, \tilde{\theta}_{-i}^X]$ , all firms produce at the capacity bound  $(x_i^1, x_{-i}^1)$  of investment in technology  $(k_1, c_1)$ . Since we analyze an upwards deviation of firm  $i$ , for higher spot market realizations,  $\theta \in [\tilde{\theta}_{-i}^X, \tilde{\theta}_i^X]$ , firm  $i$  can still produce precisely  $x_i^1$  units with technology  $(k_1, c_1)$ , but all other

firms already produce with the less efficient technology  $(k, c)$ . Those firms will produce their unconstrained Cournot output which is denoted by  $BR_{-i}(x_i^1)$ . In order to verify second order conditions we determine the second derivative  $\frac{d^2\pi_i(x_i^1, x_{-i}^{1*})}{(dx_i^1)^2}$  of firm  $i$ 's profits for  $x_i^1 > X^*/n$ . The resulting second order condition reads as follows:

$$\int_{\tilde{\theta}_i^{X^1}}^{\tilde{\theta}_{-i}^X} [2P_q(\cdot) + P_{qq}(\cdot)x_i^1] dF(\theta) + \int_{\tilde{\theta}_{-i}^X}^{\tilde{\theta}_i^X} [2P_q(\cdot)(1 + BR'_{-i}(\cdot)) + P_{qq}(\cdot)x_i^1(1 + BR'_{-i}(\cdot))^2 + P_q(\cdot)x_i^1 BR''_{-i}(\cdot)] dF(\theta) - \frac{d\tilde{\theta}_{-i}^X}{dx_i^1} f(\tilde{\theta}_{-i}^X) P_q(\cdot)x_i^1 BR'_{-i}(\cdot) + \frac{d\tilde{\theta}_i^X}{dx_i^1} f(\tilde{\theta}_i^X) P_q(\cdot)x_i^1 BR'_{-i}(\cdot) < 0. \quad (17)$$

Due to assumption 1 (and since  $-1 < BR'_{-i}(x_i^1) < 0$  and  $BR''_{-i}(x_i^1) > 0$ ) both integrals of expression (17) are strictly negative. The critical demand realizations  $\tilde{\theta}_i^{X^1}$  and  $\tilde{\theta}_i^X$  depend on marginal cost of production  $c_1$  and  $c$  respectively. As marginal cost of production  $c_1$  and  $c$  of both technologies get infinitely close,  $\tilde{\theta}_i^{X^1}$  and  $\tilde{\theta}_i^X$  collapse and both integrals vanish, however. Then concavity of firm  $i$ 's profits (for upwards deviation) cannot be ensured for every distribution of spot market demand<sup>49</sup>, and the symmetric equilibrium fails to exist, since firms have an incentive to deviate upwards increasing their investment in technology  $(k_1, c_1)$ . If on the contrary, however,  $c_1$  and  $c$  are sufficiently different<sup>50</sup>, the negative sign of both integrals dominates expression (17). Then concavity of profits can be ensured and upwards deviation is thus not profitable for firm  $i$ .

**III) Uniqueness of the Equilibrium** We now show that (i)  $x^{1*}$  is uniquely characterized by expression (4), and (ii) that there are no asymmetric equilibria.

(i) Proof that expression (4) indeed uniquely characterizes unique equilibrium investment  $x^{1*}$  in technology  $(k_1, c_1)$ . We differentiate (4) wrt to symmetric investment  $x^{1S}$  to verify concavity of profits for all symmetric solutions. This yields

$$\frac{d^2\pi_i(x^{1S}, x^*)}{(dx_i^{1S})^2} = \int_{\tilde{\theta}^{X^1}}^{\tilde{\theta}^X} [(n+1)P_q(nx_i^{1S}, \theta) + nP_{qq}(nx_i^{1S}, \theta)x_i^{1S}] dF(\theta) < 0,$$

which is negative due to assumption 1. We can conclude that  $x^{1*}$  is unique.

(ii) There cannot exist an asymmetric equilibrium. Any candidate for an asymmetric equilibrium  $\check{x}^1$  can be ordered such that  $\check{x}_1^1 \leq \check{x}_2^1 \leq \dots \leq \check{x}_n^1$ , where at least one inequality holds strictly. The first derivative  $\frac{d\pi_1(\check{x}^1, x^*)}{dx_1^1}$  of firm 1's profit is given by:

<sup>49</sup>The sign of the sum of the third and fourth term of expression (17) changes with on the precise distribution of demand at the spot markets.

<sup>50</sup>To give a simple example: for the case of a uniform distribution  $F(\theta)$  and linear demand, if  $c - c_1 \geq \frac{1}{13}$ , then condition (17) is satisfied for all  $x_i^1 \leq X^*/n$ .

$$\begin{aligned}
& \int_{\tilde{\theta}_1^{X^1}}^{\tilde{\theta}_2^{X^1}} [P(\check{x}_1^1, BR_{-1}(\check{x}_1^1)\theta) + (1 + BR'_{-1}(\check{x}_1^1)) P_q(\cdot) \check{x}_1^1 - c_1] dF(\theta) + \int_{\tilde{\theta}_2^{X^1}}^{\tilde{\theta}_3^{X^1}} \dots dF(\theta) \\
& + \dots + \int_{\tilde{\theta}_{(n-1)}^{X^1}}^{\tilde{\theta}_n^{X^1}} \dots dF(\theta) + \int_{\tilde{\theta}_1^X}^{\tilde{\theta}_1^X} [P(\check{X}^1, \theta) + P_q(\cdot) \check{x}_1^1 - c_1] dF(\theta) + \int_{\tilde{\theta}_1^X}^{\bar{\theta}} (c - c_1) dF(\theta) - (k_1 - k)
\end{aligned} \tag{18}$$

The first derivative  $\frac{d\pi_n(\check{x}_1^1, x^*)}{dx_n^1}$  of firm  $n$ 's profit is given by:

$$\begin{aligned}
& \int_{\tilde{\theta}_n^{X^1}}^{\tilde{\theta}_1^X} [P(\check{X}^1, \theta) + P_q(\cdot) \check{x}_n^1 - c_1] dF(\theta) + \int_{\tilde{\theta}_1^X}^{\tilde{\theta}_2^X} [P(\check{x}_n^1, BR_{-n}(\check{x}_n^1), \theta) + (1 + BR'_{-n}(\check{x}_n^1)) P_q(\cdot) \check{x}_n^1 - c_1] dF(\theta) \\
& + \int_{\tilde{\theta}_2^X}^{\tilde{\theta}_3^X} \dots dF(\theta) + \dots + \int_{\tilde{\theta}_{(n-1)}^X}^{\tilde{\theta}_n^X} \dots dF(\theta) + \int_{\tilde{\theta}_1^X}^{\bar{\theta}} (c - c_1) dF(\theta) - (k_1 - k)
\end{aligned} \tag{19}$$

In order to proof nonexistence of an asymmetric equilibrium, we show that  $\check{x}_n^1$  cannot be optimal for firm  $n$ , given  $\check{x}_1^1$  is the optimal choice of firm 1. Consider investment  $\check{x}_n^1$  of firm  $n$  in technology  $(k_1, c_1)$  and analyze the impact of firm  $n$  lowering its investment level to  $x_n^1 < \check{x}_n^1$ . As we show, this always leads to an increase of marginal profits of firm  $n$ .

For concave profits as established in expression (17), deviations  $\check{x}_{n-1}^1 \leq x_n^1 < \check{x}_n^1$  always lead to an increase of marginal profits of firm  $n$ . A further decrease to  $\check{x}_{n-2}^1 \leq x_n^1 < \check{x}_{n-1}^1$  will have a positive impact on marginal profits through two channels: First, a reduction of investment levels above  $\check{x}_{n-2}^1$  leads to an increase of marginal profits of firm  $i$  due to concavity established in expression (17). Second a reduction of investment levels below  $\check{x}_{n-1}^1$  leads to an increase of marginal profits of firm  $i$  (compare part II) of this proof).

We proceed analogously in order to establish an increase of firm  $n$ 's marginal profits as  $x_n^1$  is further reduced to  $\check{x}_{n-2}^1 \leq x_n^1 < \check{x}_{n-1}^1$ . Such reduction of the investment  $x_n^1$  of firm  $n$  in technology  $(k_1, c_1)$  is then carried through till  $x_n^1 = \check{x}_1^1$ . Since for the asymmetric candidate equilibrium the inequality  $\check{x}_1^1 < \check{x}_n^1$  is strict, we can conclude that marginal profits of firm  $n$  are strictly higher when choosing  $x_n^1 = \check{x}_1^1$  and not  $x_n^1 = \check{x}_n^1$ . Since an asymmetric equilibrium has to satisfy  $\frac{d\pi_1(\check{x}_1^1, x^*)}{dx_1^1} = 0$ , marginal profits when choosing  $x_n^1 = \check{x}_n^1$  will be strictly lower, becoming necessarily negative. To sum up, whenever it is optimal for firm 1 to choose  $\check{x}_1^1$ , it cannot be optimal for firm  $n$  to indeed choose  $\check{x}_n^1 > \check{x}_1^1$ , since it could strictly increase its profits by reducing its investment by choosing  $x_n^1 < \check{x}_n^1$ . We can conclude that an asymmetric equilibrium cannot exist.



## Proof of Theorem 3

The proof of existence and uniqueness for the case of perfect competition follows the same lines as the proof in case of oligopoly. Moreover under the hypothesis of price taking behavior, firms will not change their output in response to changed investment decision of its rivals. More specifically concavity of profits can always be ensured, the second order condition for upwards deviation, as given by expression (17), is always satisfied, there always exists a unique and symmetric competitive market equilibrium given by (5) and (6).

In order to prove the second part of the theorem, we derive welfare optimal total investment  $X^{FB}$  and show that it coincides with the solution under perfect competition in the absence of a cap. Welfare given  $x$  and  $\bar{p}$  is

$$W(x, \bar{p}) = \int_{\underline{\theta}}^{\hat{\theta}^X(x)} \int_0^Q (P(y, \theta) - c) dy dF(\theta) + \int_{\hat{\theta}^X(x)}^{\bar{\theta}} \int_0^X (P(y, \theta) - c) dy dF(\theta) - nK \left( \frac{X}{n} \right)$$

Thus, welfare is maximized at

$$X^{FB} = \left\{ X : \int_{\hat{\theta}^X(x)}^{\bar{\theta}} (P(X, \theta) - c) dF(\theta) = k \right\}. \quad (20)$$

Note that this coincides with the solution under perfect competition for  $\hat{\theta}^{\bar{p}}(x) = \bar{\theta}$ , i. e. for a non-binding cap. Pointwise comparison of expressions (5) and (20) directly reveals that for any fixed  $X$  the LHS of (5) is smaller than the LHS of (20) for all binding caps. This yields lower industry capacity choice under perfect competition for any binding cap than in the first best solution.

## Proof of Theorem 4

**1) Investment:** First we show  $X^*(\bar{p}) > X^{*\infty}$  for all  $\bar{p} \in [\overline{MR}, \bar{p}^\infty]$ . Denote equilibrium production in absence of a cap by  $X^{*\infty} := X^*(\bar{p}^\infty)$ , which solves the following first order condition:

$$\int_{\theta^X(x^{*\infty})}^{\bar{\theta}} \left[ P(X^{*\infty}, \theta) + P_q(X^{*\infty}, \theta) \frac{X^{*\infty}}{n} - c \right] dF(\theta) - k = 0 \quad (21)$$

Remember  $\overline{MR}$  is defined to be given by  $\overline{MR} = [P(X^{*\infty}, \theta) + P_q(X^{*\infty}, \theta) \frac{X^{*\infty}}{n} - c]_{\theta=\bar{\theta}}$ , which is the highest value the integrand of expression (21) can take. In order to prove

that a cap  $\bar{p} \in [\overline{MR}, \bar{\rho}^\infty]$  increases equilibrium capacity choice, we compare pointwise the first order condition without cap (given above by (21)) with the first order condition with binding cap as given by expression (2):

$$\int_{\theta^{X^*}(x^*)}^{\theta^{\bar{p}}(x^*)} \left[ P(X^*, \theta) + P_q(X^*, \theta) \frac{X^*}{n} - c \right] dF(\theta) + \int_{\tilde{\theta}^{\bar{p}}(x^*)}^{\bar{\theta}} (\bar{p} - c) dF(\theta) - k = 0 \quad (22)$$

By construction of  $\overline{MR}$ , for any fixed value  $X^0$ , the LHS of (22) is strictly bigger than the LHS of (21). Since profits of firms along the symmetry line are concave and cost of investment weakly convex, both expressions (22) and (21) are strictly decreasing in  $X$ . This implies however that condition (22) can only be satisfied for strictly bigger investment levels than (21), which proves  $X^*(\bar{p}) > X^{*\infty}$  for all  $\bar{p} \in [\overline{MR}, \bar{\rho}^\infty]$ .

**2) Welfare:** In order to show that total welfare always increases if a cap  $\bar{p} \in [\overline{MR}, \bar{\rho}^\infty]$  is chosen (as compared to the case without a cap) we just need to show that from a welfare perspective investment chosen by strategic firms for a given cap are always below the welfare maximizing level and thus any increase in total capacity increases welfare.

Welfare maximizing capacity choice solves expression (20) and equilibrium investment of strategic firms for any binding cap on scarcity prices solves (22). As above we observe that for any fixed investment level  $X^0$ , the LHS of (20) is strictly bigger than the LHS of (22), since the LHS of both expressions are decreasing in capacity choice  $X$ , first best capacity choice is strictly bigger than capacity choice of strategic firms for any cap.

**3) Consumer Surplus:** Since production is always increasing for any cap  $\bar{p} \in [\overline{MR}, \bar{\rho}^\infty]$  (as compared to the case without a cap), consumer surplus also increases.

**4) Average Prices:** We show that average prices decrease in the relevant range. For a given  $X^*(\bar{p})$ , average prices are uniquely given by

$$\mathbf{E} \left[ P \right] = \int_{\underline{\theta}}^{\theta^X(x^*(\bar{p}))} 0 dF(\theta) + \int_{\tilde{\theta}^X(x^*(\bar{p}))}^{\theta^{\bar{p}}(x^*(\bar{p}))} P(X^*(\bar{p}), \theta) dF(\theta) + \int_{\theta^{\bar{p}}(x^*(\bar{p}))}^{\bar{\theta}} \bar{p} dF(\theta). \quad (23)$$

If we lower caps from  $\bar{\rho}^\infty$  to some  $\bar{p}^0 \in [\overline{MR}, \bar{\rho}^\infty]$ , average prices are affected as follows. For the limits of integration in (23) we obtain  $\theta^X(x^*(\bar{p}))$  and  $\theta^{\bar{p}}(x^*(\bar{p}))$  increase. For the integrands we obtain  $P(X^{*\infty}, \theta) > P(X^*(\bar{p}), \theta)$  and  $\bar{\rho}^\infty > \bar{p}^0$ . Thus, average prices decrease.

## Proof of Theorem 5

We proof the theorem by showing that  $\widehat{p}^\infty = \overline{MR}$  under the assumptions of the theorem. This allows then to apply theorem 4. In other words we show that in absence of a cap, the perfectly competitive equilibrium price just equals marginal equilibrium revenue:

$$P(X_{PC}^*, \theta) = P(X^{*\infty}, \theta) + P_q(X^{*\infty}, \theta) \frac{X^{*\infty}}{n} \quad \forall \theta \in [\tilde{\theta}^X(x^{*\infty}) = \theta^X(x^{*\infty}), \bar{\theta}],$$

Thus, setting the highest price observed in the perfectly competitive benchmark is equivalent to setting the lowest price from the interval of welfare increasing caps identified in theorem 4.

In order to do so we show that for undistorted scarcity prices the equilibria of the market game under perfect competition and the Cournot market game can be characterized by the same condition. To this end, define  $T(X^{FB}) := \tilde{P}(X^{FB})$  and  $T(X^{*\infty}) := \tilde{P}(X^{*\infty}) + \tilde{P}_q(X^{*\infty}) \frac{X^{*\infty}}{n}$  (recall  $X^{FB}$  denotes first best equilibrium quantity, while  $X^{*\infty}$  denotes the market outcome for uncapped scarcity prices). Observe that the unique equilibrium production in both cases is characterized by the following equation,

$$\int_{\tilde{\theta}^T}^{\bar{\theta}} (a(\theta) + b(\theta)T - c) dF(\theta) = k, \quad (24)$$

where  $\tilde{\theta}^T = \{\theta : a(\theta) + b(\theta)T = c\}$ .

Thus both, the solution under perfect competition and under imperfect competition involve an identical value  $T^*$  solving (24). For all  $\theta \in [\theta^{T^*}, \bar{\theta}]$  we obtain

$$a(\theta) + b(\theta)T^* = P(X_{PC}^*, \theta) = P(X^{*\infty}, \theta) + P_q(X^{*\infty}, \theta) \frac{X^{*\infty}}{n}.$$

## Proof of Theorem 6

Welfare generated in the market equilibrium is given as follows:

$$\begin{aligned} & \int_{\underline{\theta}}^{\tilde{\theta}^X(x)} \left[ \int_0^Q (P(Y, \theta) - c)Y dY \right] dF(\theta) + \int_{\tilde{\theta}^X(x)}^{\tilde{\theta}^{\bar{p}}(x)} \left[ \int_0^{X^*} (P(Y, \theta) - c)Y dY \right] dF(\theta) + \\ & \int_{\tilde{\theta}^{\bar{p}}(x)}^{\bar{\theta}} \left[ \int_0^{X^*} (\bar{p} - c)Y dY \right] dF(\theta) - kX^* \end{aligned} \quad (25)$$

In order to derive the optimal pricecap  $\bar{p}^*$  we form the first derivative of (25) with respect to  $\bar{p}$ , which yields after rearranging:

$$\frac{dX^*(\bar{p})}{d\bar{p}} \left( \int_{\tilde{\theta}^X(x)}^{\tilde{\theta}^{\bar{p}}(x)} \left[ \int_0^{X^*} (P(X, \theta) - c)X dX \right] dF(\theta) + \int_{\tilde{\theta}^{\bar{p}}(x)}^{\bar{\theta}} \left[ \int_0^{X^*} (\bar{p} - c)X dX \right] dF(\theta) - c \right) = 0$$

As established in theorem 3, total investment  $X^*$  always remains strictly below the first best level, marginal Welfare (i.e. the term in brackets) is thus strictly positive for any cap  $\bar{p}$ . Accordingly, the welfare maximizing cap solves  $\frac{dX^*}{d\bar{p}} = 0$ . In the following we derive the optimal cap  $\bar{p}^*$  as a solution to  $\frac{dX^*}{d\bar{p}} = 0$  and show its uniqueness.

We first derive properties and slopes of  $\tilde{\theta}^{\bar{p}}(x, \bar{p})$ . Notice that at the demand realization where the cap is met in equilibrium we obtain

$$P(X, \tilde{\theta}^{\bar{p}}(x, \bar{p})) \equiv \bar{p} \quad (26)$$

We determine the partial derivative of (26) with respect to  $\bar{p}$  given some fixed  $X$ ,<sup>51</sup>

$$\frac{\partial \tilde{\theta}^{\bar{p}}(x, \bar{p})}{\partial \bar{p}} = \frac{1}{P_\theta(X, \tilde{\theta}^{\bar{p}})} = 1.$$

And then determine the total derivative of (26), evaluated at  $X^*(\bar{p})$ ,

$$P_q(\cdot) \frac{dX^*(\bar{p})}{d\bar{p}} + P_\theta(\cdot) \frac{d\tilde{\theta}^{\bar{p}}(x^*(\bar{p}), \bar{p})}{d\bar{p}} = 1 \quad \Leftrightarrow \quad \frac{d\tilde{\theta}^{\bar{p}}(x^*(\bar{p}), \bar{p})}{d\bar{p}} = 1 - P_q(X^*, \theta) \frac{dX^*(\bar{p})}{d\bar{p}}.$$

**Characterization of  $\bar{p}^*$**  We can apply the implicit function theorem to the equilibrium condition (4) in order to obtain  $\frac{dX^*(\bar{p})}{d\bar{p}}$ :

$$\frac{dX^*(\bar{p})}{d\bar{p}} = -\frac{\partial^2 \pi_i(x, \bar{p})}{\partial x_i \partial \bar{p}} \Big/ \frac{\partial^2 \pi_i(x, \bar{p})}{\partial x_i \partial X} = \left[ \frac{P_q(X^*, \tilde{\theta}^{\bar{p}})}{P_\theta(X^*, \tilde{\theta}^{\bar{p}})} \frac{X^*(\bar{p})}{n} f(\tilde{\theta}^{\bar{p}}) + 1 - F(\tilde{\theta}^{\bar{p}}) \right] \Big/ -\frac{\partial^2 \pi_i(x, \bar{p})}{\partial x_i \partial X}. \quad (27)$$

Rewriting equation (27) using the hazard rate yields

$$\frac{dX^*(\bar{p})}{d\bar{p}} = \left[ P_q(\cdot) \frac{X^*(\bar{p})}{n} + \frac{1}{h(\tilde{\theta}^{\bar{p}})} \right] \Big/ \left[ -\frac{\partial \pi_i(x, q^*)}{\partial x_i \partial X} \frac{1}{f(\tilde{\theta}^{\bar{p}})} \right]. \quad (28)$$

Define  $\xi(\bar{p}) = \left[ -\frac{\partial \pi_i(x, q^*)}{\partial x_i \partial X} \frac{1}{f(\tilde{\theta}^{\bar{p}})} \right]$ . Note that  $\xi(\bar{p}) > 0$  since profits (along the symmetry line) are concave. Thus, as established above at the welfare maximum satisfies  $\frac{dX^*(\bar{p})}{d\bar{p}} = 0$ , which gives the characterization of the optimal cap as stated in the theorem.

In order to show existence and uniqueness of these maximizers, however, we also need to show that second order conditions are satisfied, namely  $\frac{d^2 X^*(\bar{p})}{d\bar{p}^2} \Big|_{\frac{dX^*(\bar{p})}{d\bar{p}}=0} < 0$ . Differentiation of (28) with respect to  $\bar{p}$  yields<sup>52</sup>

$$\frac{d^2 X^*(\bar{p})}{d\bar{p}^2} = \frac{1}{\xi(\bar{p})^2} \left[ \left( \frac{dX^*(\bar{p})}{d\bar{p}} \left( P_{qq} \frac{X^*(\bar{p})}{n} + P_q \right) - \frac{h_\theta(\tilde{\theta}^{\bar{p}})}{h^2(\tilde{\theta}^{\bar{p}})} \frac{d\tilde{\theta}^{\bar{p}}}{d\bar{p}} \right) \xi(\bar{p}) - \frac{dX^*(\bar{p})}{d\bar{p}} \xi(\bar{p}) \xi_{\bar{p}}(\bar{p}) \right]$$

<sup>51</sup>Note that  $P_\theta = 1$ , given the assumptions on demand established for theorem 6.

<sup>52</sup>Notice that  $P_{q\theta} = P_{\theta\theta} = 0$ , given the assumptions on demand established for theorem 6. Furthermore  $\frac{d\tilde{\theta}^{\bar{p}}}{d\bar{p}} = 1$  at  $\frac{dX^*(\bar{p})}{d\bar{p}} = 0$ ,  $\frac{dX^*(\bar{p})}{d\bar{p}} \xi_i(\bar{p})$  is the numerator of expression (28).

This can be evaluated at the optimal  $\bar{p}^*$ , which satisfies  $\frac{dX^*(\bar{p})}{d\bar{p}} = 0$ , we obtain

$$\left. \frac{d^2 X^*(\bar{p})}{d\bar{p}^2} \right|_{\frac{dX^*(\bar{p})}{d\bar{p}}=0} = -\frac{1}{\xi(\bar{p})} \frac{h_\theta(\tilde{\theta}^{\bar{p}})}{h^2(\tilde{\theta}^{\bar{p}})} < 0.$$

Thus, under our assumptions all (local) extrema are necessarily maxima. Since we have local minima at  $p = \bar{p}^0$  and  $p = \bar{p}^\infty$ , we conclude that the above condition characterizes the unique welfare maximizing cap.

Finally show that the optimal cap  $\bar{p}^*$  is weakly increasing in the number of firms. In order to do so we start with the unique characterization of the optimal cap as derived above and given by expression (28). Applying the envelope theorem yields:

$$\frac{d\bar{p}^*(n)}{dn} = -\frac{\partial}{\partial n} \left[ P_q(X^*, \tilde{\theta}^{\bar{p}}) \frac{X^*}{n} + \frac{1}{h(\tilde{\theta}^{\bar{p}})} \right] \bigg/ \frac{\partial}{\partial \bar{p}^*} \left[ P_q(X^*, \tilde{\theta}^{\bar{p}}) \frac{X^*(p)}{n} + \frac{1}{h(\tilde{\theta}^{\bar{p}})} \right].$$

Since at  $\bar{p} = \bar{p}^*$  we obtain  $\frac{dX^*(\bar{p})}{d\bar{p}} = 0$ , this simplifies to give:

$$\frac{d\bar{p}^*(n)}{dn} = -P_q X^* \frac{-1}{n^2} \bigg/ \frac{-h_\theta(\tilde{\theta}^{\bar{p}})}{h(\tilde{\theta}^{\bar{p}})^2} \frac{d\tilde{\theta}^{\bar{p}}}{d\bar{p}^*} > 0$$

Notice that the above expression is positive, since (i) the hazard rate is increasing (i.e.  $h_\theta > 0$ ) and (ii) for higher caps the cap starts to be binding for higher demand realizations (i.e.  $\frac{d\tilde{\theta}^{\bar{p}}}{d\bar{p}^*} > 0$ ).