The visible hand: ensuring optimal investment in electric power generation

Thomas-Olivier Léautier

Toulouse School of Economics (IAE, IDEI, University of Toulouse)

21 allée de Brienne

F31 000 Toulouse

thomas.leautier@iae-toulouse.fr. and +33 6 33 34 60 68

August 19, 2012

Abstract

This article formally analyzes the causes of underinvestment in electric power generation, and the various corrective market designs that have been proposed and implemented. It yields four main analytical findings. First, using a simple numerical example, (a linear demand function, calibrated on the French power load duration curve), strategic supply reduction is shown to be a more important cause of underinvestment than the imposition of a price cap. Second, physical capacity certificates markets implemented in the United States restore optimal investment, but increase producers' profits beyond the imperfect competition level. Third, financial reliability options, proposed in many markets, fail to restore investment incentives. If a "no short sale" condition is added, they are equivalent to physical capacity certificates. Finally, if competition is perfect, energy only markets yield a negligible underinvestment compared to the optimum. Taken together, these findings suggest that, to ensure generation adequacy, policy makers should put more effort on enforcing competitive behavior in the energy markets, and less on designing additional markets.

Keywords: imperfect competition, market design, investment incentives

JEL Classification: L13, L94

1 Introduction

An essential objective of the restructuring of the electric power industry in the 1990s was to "push to the market" decisions and risks associated with investment in power generation, i.e., to have market forces, not bureaucrats, determine how much investment is required, and to have investors, not rate-payers, bear the risk of excess capacity, construction cost overruns and delays.

However, since the early 2000s, generation adequacy has become an issue of concern for policy makers, power System Operators (*SOs*), and economists. It would appear that, contrary to the initial belief, the "market" does not necessarily provide for the adequate level of generation capacity. Britain, that pioneered the restructuring of the electricity industry in 1990, constitutes the most recent and striking example: Ofgem, the energy regulator warns of possible power shortages around 2015 (Ofgem (2010)).

Operating and regulatory practices aimed at preventing the exercise of market power are often considered to be the primary cause of this "market failure". As shown in Marcel Boiteux (1949)'s seminal analysis, high prices in some states of the world are required to finance the optimal capacity. However, in most jurisdictions *SOs* impose *de jure* or *de facto* price caps, that deprive producers of these high prices. This revenue loss, called "missing money", is considered an important driver of underinvestment in generation (Joskow (2007)).

Therefore, *SOs* and policy makers worldwide have designed and implemented a variety of market designs to correct this apparent "market failure" (Finon and Pignon (2008)). For example, most US power markets have adopted highly structured and prescriptive physical certificates markets, and many European countries are considering, designing or implementing capacity mechanisms[1].

These mechanisms are extremely complex, hence expensive to set up and run. Furthermore, they constitute a partial reversion towards central planning, which restructuring precisely attempted to eliminate: using a centralized system reliability model, the *SO* sets a generation capacity target, and organizes its procurement. Risk of overcapacity is borne by consumers, while risk of cost overrun is borne by investors. A rigorous economic analysis of the causes of underinvestment in generation and the various market designs implemented by *SOs* to restore investment incentives is therefore required.

[1] France formally instituted a capacity obligation mechanism in March 2012, to be effective in 2015. Britain, Germany, and Belgium are designing mechanisms to ensure adequate capacity.

This is the objective of this article. I am not aware of any previous systematic analytical comparison of these designs.

This work draws on a rich literature, that can be structured along two themes. A first group of articles examines generation investment in restructured power markets. While these works differ in important aspects, most model two stage games: in stage 1, producers decide on installed capacity; in stage 2 they produce and sell in the spot markets, subject to the installed capacity constraint. For example, Borenstein and Holland (2005) and Joskow and Tirole (2007), building on Boiteux (1949), have developed the "benchmark" model of optimal investment and production when (i) demand is uncertain at the time the investment decision is made, and (ii) a fraction of the demand does not react to price. The former article considers the perfect competition case, while the latter introduces some elements of imperfect competition. Murphy and Smeers (2005) have developed models of closed- and open-loop Cournot competition at the investment and spot market stages, and characterized the equilibria of these games. Boom (2009) has examined the impact of vertical integration on equilibrium investment, while Fabra et al. (2011) have examined the impact of the structure of the auction in the spot market on the equilibrium investment. This article builds on the two-stage Cournot game formalized in Zöttl (2011). A more recent literature (e.g., Garcia and Shen (2010)) examine multiperiod investment decisions.

A second group of works describes and analyzes the possible "corrective" market designs. Stoft (2002) discusses average Value of Lost Load pricing, Hogan (2005) proposes an energy cum operating reserves markets, and Cramton and Stoft (2006 and 2008) and Cramton and Ockenfels (2011) propose a financial reliability options mechanism[2]. Joskow and Tirole (2007) show that a capacity market and a price cap do not restore the first best with more than two states of the world. Chao and Wilson (2005) examine the impact of options on spot market equilibrium, investment, and welfare. Zöttl (2011) determines the welfare maximizing price cap in the spot market. However, none of these works presents a rigorous comparison of these mechanisms in a general and common setting.

This article bridges these two strands of literature, that analyzes the proposals described in the second group of articles using a rigorous economic model developed in the first group: an extension of the two-stage Cournot model developed by Zöttl (2011) to include both "price reactive" customers and

[2]Since these mechanisms are described extensively in the article, they are not developped further here.

"constant price customers", the latter being unable to react to spot energy prices and being rationed in some instances (Borenstein and Holland (2005), Joskow and Tirole (2007), Stoft (2002), and Hogan (2005)). Its contribution is to propose clear policy recommendations, building on the economic analysis of underinvestment and corrective market designs. While this work's primary focus is the electric power industry, the analysis presented here can serve as a basis to examine (under)investment issues in other industries where participants must select capacity in the presence of significant demand variability and uncertainty and limited storage possibilities, for example telecommunications and transport networks.

The analysis yields four main analytical findings, that can inform policy makers.

First, Proposition 1 characterizes two sufficient conditions for underinvestment: (i) strategic supply reduction from imperfectly competitive producers, *or* (ii) a binding price cap, the missing money problem identified earlier. This work's contribution is not the result *per se*, but the analysis of the relative impact of the two causes of underinvestment. If competition is perfect, using an illustrative numerical example (a linear demand function, calibrated on the French power load duration curve), a 3 000 €$/MWh$ price cap, which is the current level in European markets, is never binding. Any underinvestment would thus be entirely caused strategic supply reduction. A much lower price cap at 1 000 €$/MWh$ yields installed capacity 0.9% lower than optimal if 5% of the load is price responsive. If rationing is anticipated and proportional, strategic supply reduction is more important than a 1 000 €$/MWh$ price cap in reducing capacity as soon as more than 3.9% of the load is price responsive. For example, strategic supply reduction is 1.5 times more important than a 1 000 €$/MWh$ price cap in reducing capacity if 5% of the load is price responsive. These results are robust to using a much lower price elasticity.

This observation is original to this work, that casts a new light on the cause of underinvestment, emphasizing strategic supply reduction as opposed to the specificities of the power industry.

Second, Proposition 2 examines the equilibrium of markets where energy and forward physical installed capacity certificates are separately exchanged. This is the case for example in the Northeast of the United States: 3 to 5 years ahead, the SO procures from producers physical capacity certificates (usually 15 to 20% higher than anticipated peak load to protect against supply and demand fluctuations). The cost of these purchases is then passed on to customers. Proposition 2 shows that the SO must impose a "no short sale" requirement, i.e., require producers to sell less certificates than have installed capacity (or to build as much capacity as they have sold certificates). If she does, a physical capacity certificates market restores investment incentives: the resulting capacity installed is optimal. Social welfare is thus maximized. However, producers profits are higher than the imperfect competition outcome without the capacity market.

The last result is original to this work. Coupled with the previous analysis of the causes of underinvestment, it casts a new light on physical certificates markets: the common wisdom is that these markets are required to compensate the missing money. If, as suggested by the illustrative numerical example, strategic supply reduction is indeed the main driver of underinvestment, physical certificates markets then sur-remunerate this strategic behavior, hence should not be implemented. As in most industries, increasing effective competition is the main policy tool to achieve generation adequacy.

Third, Proposition 3 analyzes the equilibrium of another form of forward markets, where producers are required to sell financial call options to customers, covering all the demand up to a certain level at a given strike price. Option sellers pay customers the difference between the actual spot energy price and the strike price (Cramton and Stoft (2006 and 2008), Cramton and Ockenfels (2011)). The sale of options is expected to solve the underinvestment problem, as producers are financially exposed to high power prices.

However, Proposition 3 proves this intuition holds only partially: options sale reduces but does not eliminate underinvestment. Installed capacity is higher with options sale than without, but still lower than socially optimal. To ensure optimal investment, the SO must again impose a "no short sale" requirement. If she does, Proposition 4 shows that "dual markets" relying on financial reliability options or on physical capacity certificates are equivalent if the "technical" parameters are identical (e.g., if the option strike price equals the wholesale price cap). Reliability options thus also sur-remunerate strategic underinvestment. While Propositions 3 and 4 are consistent with Chao and Wilson (2005) and Allaz and Villa (1993)'s theoretical analysis of the interaction between forward and spot markets, they are new to the literature.

The last analytical finding is a clarification of the economics of two "energy only" market designs, where prices in the energy market are expected to be high enough to restore investment incentives.

These market designs have limited impact on strategic supply reduction: if generation is imperfectly competitive, underinvestment will remain. Hence, the research question is whether they restore the missing money under perfect competition.

Consider first average Value of Lost Load ($VoLL$) pricing, presented for example by Steven Stoft (2002) and implemented in Texas: the SO sets a price cap at the average $VoLL$, an estimate of the average value for users of not being curtailed, typically around 10 000 to 20 000 $\$/MWh$. Proposition 1 applies: a price cap, even very high, reduces investment incentives. However, if generation is competitive and the cap high enough, the numerical illustration suggests the distortion is negligible.

Finally, consider the "energy cum operating reserves market" proposed by Hogan (2005). SOs procure operating reserves to protect against an unplanned generation outage. Hogan (2005) proposes the SO balances supply against demand for energy *and* operating reserves, subject to the average $VoLL$ being used as a price cap. Producers receive additional revenues since: (i) the resulting power price is higher than when the SO balances supply against demand for energy alone, and (ii) capacity providing operating reserves – but no energy – is remunerated. This additional revenue is expected to resolve the missing money problem, hence restores investment incentives. However, Proposition 5 shows this intuition is invalid: since these additional revenues are already accounted for in the determination of the installed capacity, the situation is isomorphic to average $VoLL$ pricing, hence Proposition 1 applies. If generation is competitive, a high price cap induces a negligible underinvestment.

The analysis summarized above suggests that, despite its specific physical features, the power industry is not that different from other industries: strategic supply reduction is the main cause of underinvestment. Assuming this observation is confirmed using other specifications of power demand, the analysis yields clear policy recommendations.

If policy makers and the SO are confident a market is sufficiently competitive, as may be the case in Texas, there is no need to set up a forward capacity market (physical or financial), which are complex and costly to administer. Average $VoLL$ pricing or an energy cum operating reserves market are simple to set up and cause limited distortion compared to the optimum. Furthermore, an energy cum operating reserves market remunerates flexibility, an important issue which is not covered in this work.

On the other-hand, policy makers may determine that generation is insufficiently competitive in their jurisdiction. This may be the case in European markets, where in most markets less than 10 generation companies actually compete. This may also be the case where congestion on the transmission grid separates the market in smaller submarkets, and producers may be able to exert local market power. Then, policy makers should first seek to increase competition among generators before setting up a forward capacity market.

The article is structured as follows. Part I examines the causes of underinvestment. Section 2 presents the model structure and derives the optimal capacity. Section 3 estimates the impact of the causes of underinvestment. Part II analyzes "dual markets" designs. Section 4 examines markets for physical installed capacity certificates. Section 5 analyzes financial reliability options. Part III analyzes "energy only" market designs. Section 6 analyzes average $VoLL$ pricing. Section 7 analyzes the "energy cum operating reserves market". Finally, Section 8 suggests future research directions. Technical proofs are included in the Appendix.

Part I

Underinvestment

2 Model structure

2.1 Uncertainty

Uncertainty is an essential feature of power markets. In this work, demand uncertainty is explicitly modeled, while production uncertainty is taken into account implicitly through operating reserves (presented in Section 7). This representation is suitable for markets that rely mostly on controllable generation technologies, such as thermal and nuclear (see for example Chao and Wilson (1987)). Extension to markets where intermittent sources constitute an important portion of the generation portfolio are left for further work.

The number of possible states of the world is infinite, and these are indexed by $t \in [0, +\infty)$. The functions $f(t)$ and $F(t)$ are respectively the ex ante probability and cumulative density functions of state t. Since all market participants have the same information about future demand projections and construction plans, $f(t)$ and $F(t)$ are common to all stakeholders.

2.2 Supply and demand

Supply This article considers a single generation technology, with marginal cost $c > 0$ and investment cost r. A single technology is sufficient to analyze total installed capacity, that depends solely on the characteristics of the marginal technology (see for example Boiteux (1949) for the perfect competition case and Zöttl (2011) for the imperfect competition case).

Underlying demand

Assumption 1 *Customers all have the same load profile: denoting p the electric power price, all customers have the same underlying demand $D(p;t)$ in state t, up to a scaling factor.*

Assumption 1 greatly simplifies the derivations, while preserving the main economics insights. Inverse demand is $P(q;t)$ defined by $D(P(q;t);t) = q$, and gross consumers surplus is $S(p;t) = \int_0^{D(p;t)} P(q;t)\, dq$.

Assumption 2 Properties of the inverse demand function[3]*. For all $t \geq 0$, $Q \geq q \geq 0$: (i) $P_t(Q;t) > 0$, (ii) $P_q(Q;t) < 0$, (iii) $P_q(Q;t) + qP_{qq}(Q;t) < 0$, (iv) $P_t(Q;t) + qP_{qt}(Q;t) > 0$, and (v) $\lim_{Q\to+\infty} P(Q;t) < c$.*

Assumption 2 (*i*) simply orders the states of the world by increasing demand, (*ii*) requires the demand to be downward sloping, (*iii*) guarantees that marginal revenue is decreasing, (*iv*) requires the marginal revenue to be increasing with the state of the world, and (*v*) requires the marginal value of power to be lower than its marginal cost beyond some level. Assumption 2 holds for example if the inverse demand is linear with constant slope: $P(q;t) = a(t) - bq$, where $a(t) > 0$, $a'(t) > 0$, and $b > 0$.

Constant price customers and curtailment Currently, only a fraction $\alpha > 0$ of customers face and react to real time wholesale price ("price reactive" customers), while the remaining fraction $(1 - \alpha)$ of customers face constant price p^R in all states of the world ("constant price" customers).

Since a fraction of customers does not react to real time price, there may be instances when the *SO* has no alternative but to curtail demand, i.e., to interrupt supply.

[3] For any function $f(x,y)$, $f_x(x,y) = \frac{\partial f}{\partial x}(x,y)$, $f_{xx}(x,y) = \frac{\partial^2 f}{\partial x^2}(x,y)$, and $f_{xy}(x,y) = \frac{\partial^2 f}{\partial x \partial y}(x,y)$.

Assumption 3 *The SO has the technical ability to curtail "constant price" consumers while not curtailing "price reactive" customers.*

Assumption 3 holds only partially today: most *SOs* can only organize curtailment by geographical zones, and cannot differentiate by type of customer. However, most price reactive customers are large enough that they are connected directly to individual transformers, or to the high voltage grid, hence they need not be curtailed when the *SO* curtail constant price customers. Assumption 3 will hold fully in a few years, when "smart meters" are rolled out, as is mandated in most European countries and many US states. *SOs* will then be able to differentiate among adjacent customers, on the basis of the information provided by power suppliers.

As discussed for example in Joskow and Tirole (2007), there exists multiple rationing technologies. Curtailment is represented by a *serving ratio* $\gamma \in [0,1]$: $\gamma = 0$ represents no serving (i.e., all energy to all consumers is curtailed), while $\gamma = 1$ represents full serving (i.e., no customer is curtailed). $\mathcal{D}(p,\gamma;t)$ is the demand for price p and serving ratio γ in state t, that verifies $\mathcal{D}_t(p,\gamma;t) > 0$. All rationing technologies satisfy: (i) $\mathcal{D}(p,0;t) = 0$, (ii) $\frac{\partial \mathcal{D}}{\partial \gamma} > 0$ for $\gamma \in [0,1]$, and (iii) $D(p;t) \equiv \mathcal{D}(p,1;t)$ and $S(p;t) \equiv \mathcal{S}(p,1;t)$ where $\mathcal{S}(p,\gamma;t) = \int_0^{\mathcal{D}(p,\gamma;t)} \mathcal{P}(q,\gamma;t)\,dq$ is the gross consumer surplus, and $\mathcal{P}(q,\gamma;t)$ is the inverse demand for a given serving ratio γ: $\mathcal{D}(\mathcal{P}(q,\gamma;t),\gamma;t) = q$. One verifies that: $\frac{\partial \mathcal{S}(p,\gamma;t)}{\partial p} = p\frac{\partial \mathcal{D}}{\partial p}$.

When consumers are curtailed, the marginal Value of Lost Load ($VoLL$) represents the value they place on an extra unit of electricity (Joskow and Tirole (2007), Stoft (2002)), formally defined as

$$v(p,\gamma;t) = \frac{\frac{\partial \mathcal{S}}{\partial \gamma}}{\frac{\partial \mathcal{D}}{\partial \gamma}}(p,\gamma;t).$$

2.3 Socially optimal production and investment

The optimal production and investment is derived for example by for example Borenstein and Holland (2005) and Joskow and Tirole (2007), and summarized below for the reader's convenience.

Define $\rho(Q;t)$ the residual inverse demand curve with possible curtailment of constant price customers

$$\rho(Q;t) = P\left(\frac{Q-(1-\alpha)\mathcal{D}\left(p^R,\gamma^*;t\right)}{\alpha};t\right) \tag{1}$$

where γ^* is the optimal serving ratio in state t for production Q.

Price reactive customers face the wholesale spot price $\rho(Q;t)$, hence are never curtailed at the optimum. Off-peak, demand is low, and production $Q(t)$ is determined by $\rho(Q(t);t) = c$. On-peak, demand is set by installed capacity K, and the wholesale price is $\rho(K;t)$.

As long as $\rho(K;t) \leq v\left(p^R, 1; t\right)$, constant price customers are not curtailed in state t. If $\rho(K;t) > v\left(p^R, 1; t\right)$, then, $\gamma^* < 1$ is set to equalize constant price customers' $VoLL$ and the wholesale price:

$$v\left(p^R, \gamma^*; t\right) = \rho(K;t)$$

Define $\widehat{t}(K)$ the first state of the world when curtailment may occur[4]. Assumption 4 below insures that, if curtailment occurs in state $\widehat{t}(K)$, it also occurs in all states $t \geq \widehat{t}(K)$. If curtailment never occurs, $\widehat{t}(K) \to +\infty$.

Define

$$\Psi(K,c) = \int_{\bar{t}(K,c)}^{+\infty} (\rho(K;t) - c) f(t)\, dt$$

where $\bar{t}(K,c)$ is the first on-peak state of the world, where price equals marginal cost for production K

$$\rho(K; \bar{t}(K,c)) = c.$$

$\Psi(K,c)$ is the marginal social value capacity, decreasing in both arguments. We assume (c,r) are such that there exists a solution to:

$$\Psi(K^*, c) = r. \tag{2}$$

The optimal capacity K^* is unique since $\Psi(K,c)$ is decreasing in its first argument. Off-peak, as long as capacity is not constrained, price equals marginal cost, hence marginal capacity generates no economic profit. On-peak, when capacity is constrained, price exceeds marginal cost. The optimal capacity is set such that the marginal social value capacity is exactly equal to the marginal capacity cost r.

Additional assumptions on the inverse demand and rationing technology are required to ensure

[4] $\widehat{t}$ is a function of all the parameters. The notation $\widehat{t}(K)$ is used since the dependency on installed capacity K is the most important in this analysis.

that $\rho(K;t)$ satisfies Assumption 2, for which we use the following notation: if no rationing occurs, $\rho_q = \frac{1}{\alpha} P_q \left(\frac{Q-(1-\alpha)D\left(p^R;t\right)}{\alpha}; t \right)$; when rationing occurs, $\rho_q = \frac{\partial v}{\partial K} = \frac{\partial v}{\partial \gamma} \frac{\partial \gamma^*}{\partial K}$.

Assumption 4 Properties of the inverse demand and rationing technology. *For all* $t \geq 0$, $Q \geq q \geq 0$, $\alpha \in (0,1]\,0$, $p^R > 0$, *and* $\gamma \in (0,1]$*: (i) the marginal revenue decreases as production increases,* $\rho_q(Q;t) + q\rho_{qq}(Q;t) < 0$, *(ii) the marginal revenue increases as the state of the world increases,* $\rho_t(Q;t) + q\rho_{qt}(Q;t) > 0$, *(iii) the VoLL decreases as the serving ratio increases,* $\frac{\partial v}{\partial \gamma} \leq 0$ *and increases as the state of the world increases* $\frac{\partial v}{\partial t} > 0$, *and (iv)* $\left(\alpha\left(\frac{\partial D}{\partial p}\frac{\partial v}{\partial t} + \frac{\partial D}{\partial t}\right) + (1-\alpha)\frac{\partial \mathcal{D}}{\partial t}\right) > 0$.

If no rationing occurs, Assumption 2 is sufficient to guarantee that $\rho_t(Q;t) > 0$, $\rho_q(Q;t) < 0$, and $\lim_{Q\to+\infty} \rho(Q;t) < c$. However, if $P_{qq} > 0$, $P_q(Q;t) + QP_{qq}(Q;t) < 0$ does not guarantee that $\rho_q(Q;t) + Q\rho_{qq}(Q;t) < 0$. Additional conditions (i) and (ii) from Assumption 4 are required.

Suppose now rationing occurs for $t \geq \widehat{t}(K)$. As shown in Appendix A, $\frac{\partial v}{\partial \gamma} \leq 0$ guarantees that $\frac{\partial \gamma^*}{\partial K} > 0$ and $\frac{\partial v}{\partial K} \leq 0$. Conditions (iii) and (iv) guarantee that the optimal serving ratio decreases as the state of the world increases: $\frac{\partial \gamma^*}{\partial t} < 0$. If curtailment occurs in state $\widehat{t}$, it also occurs in all states $t \geq \widehat{t}$. Furthermore, price increases as the state of the world increases: $\frac{dv}{dt} > 0$. Finally, conditions (i) and (ii) ensure that the second derivatives of $v\left(p^R, \gamma^*; t\right)$ have the desired properties.

Assumption 4 holds for example if inverse demand is linear with constant slope: $P(q;t) = a(t) - bq$ and rationing anticipated and proportional: $\mathcal{S}(p,\gamma;t) = \gamma S(p;t)$ and $\mathcal{D}(p,\gamma;t) = \gamma D(p;t)$. If no rationing occurs,

$$\rho(Q;t) = \frac{a(t) - bQ - (1-\alpha)p^R}{\alpha} \tag{3}$$

is linear, hence satisfies conditions (i) and (ii).

Since rationing is anticipated and proportional,

$$v\left(p^R, \gamma; t\right) = \frac{S\left(p^R;t\right)}{D\left(p^R;t\right)} = a(t) - b\frac{D\left(p^R;t\right)}{2} = \frac{a(t) + p^R}{2}. \tag{4}$$

$\frac{\partial v}{\partial \gamma} = 0$, and $\frac{\partial v}{\partial t} = \frac{a^{'}(t)}{2} > 0$, and $v\left(p^R, \gamma^*; t\right)$ satisfies conditions (iii) to (v).

3 Imperfect competition, price cap, and underinvestment

3.1 Capacity constrained Cournot competition under uncertainty

Consider now N producers, that play a two-stage game: in stage 1, producer n installs capacity k^n; in stage 2 he produces $q^n(t) \leq k^n$ in the spot market in state t. Producers are assumed to compete à la Cournot in the spot markets, facing inverse demand $\rho(Q,t)$ defined by equation (1). Stage 2 can be interpreted as a repetition of multiple states of the world over a given period (for example one year), drawn from the distribution $F(.)$.

The game is solved by backwards induction: producers first compute profits from a Nash equilibrium given installed capacities $\left(k^1, ..., k^N\right)$ in the energy spot market for each state of the world t; then they make their investment choice in stage 1 based on the expectation of these spot market profits.

$Q(t) = \sum_{n=1}^{N} q^n(t)$ and $K = \sum_{n=1}^{N} k^n$ are respectively aggregate production in state t and aggregate installed capacity. $\Pi^n(k^n, \mathbf{k}^{-n})$ is producer's n profit for the two-stage game.

As will be discussed, the results hold for other forms of imperfect competition in the spot market, as long they yield an equilibrium price higher than the marginal cost c, and a profit function $\Pi^n(k^n, \mathbf{k}^{-n})$ with the required concavity. Cournot competition is used as it provides simple analytical expressions that can be illustrated numerically.

To limit the exercise of market power, the SO may impose a cap $\bar{p}^W$ on the wholesale power price[5]. If she does, $\bar{p}^W$ verifies:

$$c + r < \bar{p}^W \leq \rho(0,0) \tag{5}$$

A price cap lower than the full marginal cost of the first unit of energy would block any investment. The upper bound on the price cap is sufficient to ensure existence of an equilibrium. The price cap may be a formal cap, or the result of operational practices that depress prices (see Joskow (2007)). $\bar{t}\left(K, \bar{p}^W\right)$ is the first state of the world where the cap may be binding for production K:

$$\rho\left(K; \bar{t}\left(K, \bar{p}^W\right)\right) = \bar{p}^W.$$

[5]In practice, most SOs in the United States impose a cap on *bids* into the wholesale markets, not a cap on wholesale price. A wholesale price cap simplifies the analysis, while preserving the main economic insights.

If the price cap is never binding, $\bar{t}\left(K, \bar{p}^W\right) \to +\infty$.

K^C is the equilibrium capacity referred to as the "Cournot" capacity in the following. $t^N(K)$ is the first on-peak state of the world under imperfect competition, i.e., where the marginal revenue for production K equals marginal cost:

$$\rho\left(K; t^N(K)\right) + \frac{K}{N}\rho_q\left(K; t^N(K)\right) = c.$$

The aggregate capacity constraint may be binding before or after the price cap constraint in the relevant range, i.e., $t^N(K) < \bar{t}\left(K, \bar{p}^W\right)$ or $t^N(K) \geq \bar{t}\left(K, \bar{p}^W\right)$. This issue is original to this analysis. Zöttl (2011) does not include constant price customers, i.e., sets $\alpha = 1$. Thus, with the values of the parameters he estimates, only the case $t^N(K) < \bar{t}\left(K, \bar{p}^W\right)$ occurs.

In this analysis, $t^N(K) > \bar{t}\left(K, \bar{p}^W\right)$ is a distinct possibility, in particular if $\rho(Q, t)$ is very inelastic, i.e., if α is very low. We therefore derive in Appendix B the equilibrium investment if $t^N(K) > \bar{t}\left(K, \bar{p}^W\right)$. Both cases yield the same economic insights, with slightly different equations. Only the case $t^N(K) \leq \bar{t}\left(K, \bar{p}^W\right)$ is presented here, since it is the only possibility under perfect competition ($\lim_{N\to+\infty} t^N(K) = \bar{t}(K, c) < \bar{t}\left(K, \bar{p}^W\right)$), and is consistent with average $VoLL$ pricing and operating reserves pricing presented in Sections 6 and 7.

For states $t \leq \bar{t}\left(K, \bar{p}^W\right)$, producers face inverse demand $\rho(K; t)$, while they face "horizontal" inverse demand $\bar{p}^W$ for $t \geq \bar{t}\left(K, \bar{p}^W\right)$. Imposition of the price cap $\bar{p}^W$ leads to curtailment: since both price reactive and constant price customers face a constant price, demand is perfectly inelastic. Curtailment does not impact producers' profit, since they already produce their maximum capacity and receive $\bar{p}^W$. To simplify the notation, I assume that price reactive customers are never curtailed. The social value of energy thus remains $\rho(K; t)$.

Since $\rho(K; t)$ verifies Assumption 2, Zöttl (2011)'s Proposition 1 applies: the profit function $\Pi^n\left(\frac{K}{N}, ..., \frac{K}{N}\right)$ is concave, and there exists a unique symmetric equilibrium of the investment game $\left(\frac{K^C}{N}, ..., \frac{K^C}{N}\right)$ defined by:

$$\Omega\left(K^C, \bar{p}^W\right) = r \tag{6}$$

where

$$\Omega\left(K,\bar{p}^W\right) = \int_{t^N(K)}^{\bar{t}\left(K,\bar{p}^W\right)} \left(\rho\left(K;t\right)+\frac{K}{N}\rho_q\left(K;t\right)-c\right) f\left(t\right)dt + \int_{\bar{t}\left(K,\bar{p}^W\right)}^{+\infty}\left(\bar{p}^W-c\right) f\left(t\right)dt \tag{7}$$

$$= \Psi\left(K,c\right)+\frac{K}{N}\int_{t^N(K)}^{\bar{t}\left(K,\bar{p}^W\right)}\rho_q\left(K;t\right) f\left(t\right)dt - \int_{\bar{t}(K,c)}^{t^N(K)}\left(\rho\left(K;t\right)-c\right) f\left(t\right)dt - \Psi\left(K,\bar{p}^W\right)$$

is the marginal value of capacity for a producer at the symmetric equilibrium. $\Omega(.)$ is decreasing in its first argument since $\Pi^n\left(\frac{K}{N},...,\frac{K}{N}\right)$ is concave.

Equation (7) illustrates the marginal value of capacity for producers. As in equation (2), only states of the world where the aggregate capacity is binding (i.e., $t \geq t^N$) appear in equation (7): marginal generation capacity remunerates its investment cost only when it is constrained. When capacity is constrained but the price cap is not yet binding (i.e., $t^N(K) \leq t < \bar{t}\left(K,\bar{p}^W\right)$), adding a marginal MW of generation capacity increases the capacity which receives margin per unit $\left(\rho\left(K^C;t\right)-c\right)$ and decreases the margin on all inframarginal units $\frac{K^C}{N}\rho_q\left(K^C;t\right) < 0$. The net effect is an increase in profit, since $\rho\left(K^C;t^N\left(K^C\right)\right)+\frac{K^C}{N}\rho_q\left(K^C;t^N\left(K^C\right)\right)=c$ and $\rho_t\left(Q;t\right)+\frac{Q}{N}\rho_{qt}\left(Q;t\right)>0$ by Assumption 4. When the price cap is binding (i.e., $t \geq \bar{t}\left(K,\bar{p}^W\right)$), adding a marginal MW of generation generates additional margin $\left(\bar{p}^W-c\right)$.

3.2 Underinvestment and missing money

Observing that

$$\Omega\left(K,\bar{p}^W\right)-\Psi\left(K,c\right)=\frac{K}{N}\int_{t^N(K)}^{\bar{t}\left(K,\bar{p}^W\right)}\rho_q\left(K;t\right) f\left(t\right)dt - \int_{\bar{t}(K,c)}^{t^N(K)}\left(\rho\left(K;t\right)-c\right) f\left(t\right)dt - \Psi\left(K,\bar{p}^W\right)$$

illustrates the two distortions that reduce investment. First, imperfect competition reduces the marginal value of capacity by two terms: the reduction in profit on the inframarginal units as in all Cournot competition models $\left(\frac{K}{N}\int_{t^N(K)}^{\bar{t}\left(K,\bar{p}^W\right)}\rho_q\left(K;t\right) f\left(t\right)dt\right)$, but also the lost margin $\left(\rho\left(K;t\right)-c\right)$ in the states of the world $t\in\left[\bar{t}\left(K,c\right),t^N\left(K\right)\right]$. Both effects are negative. Second, the price cap reduces the marginal value by $\Psi\left(K,\bar{p}^W\right)$: the SO values energy at $\rho\left(K;t\right)$, while producers receive only $\bar{p}^W < \rho\left(K;t\right)$. This is the "missing money" discussed for example by Joskow (2007), and Cramton and Stoft (2006).

The previous discussion can be summarized in the following Proposition:

Proposition 1 *Define* K^C *and* $K^C\left(\bar{p}^W\right)$ *respectively the Cournot capacity absent any price cap and for price cap* $\bar{p}^W$*, and* $K^*\left(\bar{p}^W\right)$ *the installed capacity under perfect competition for price cap* $\bar{p}^W$*. Imperfect competition always reduce installed capacity, and, under perfect competition, a binding price cap reduces installed capacity:*

$$K^C < K^* \text{ and } K^C\left(\bar{p}^W\right) < K^*\left(\bar{p}^W\right) \leq K^*$$

Proof. *Suppose no price cap is imposed,*

$$\Omega\left(K^*\right) < \Psi\left(K^*, c\right) = r = \Omega\left(K^C\right) \Leftrightarrow K^* > K^C$$

since $\Omega\left(.\right)$ *is decreasing. Suppose now competition is perfect, and a cap is imposed,*

$$\Omega\left(K^*, \bar{p}^W\right) < \Psi\left(K^*, c\right) = r = \Omega\left(K^*\left(\bar{p}^W\right), \bar{p}^W\right) \Leftrightarrow K^* > K^*\left(\bar{p}^W\right).$$

Finally, since $\Psi\left(K^*\left(\bar{p}^W\right), c\right) - \Psi\left(K^*\left(\bar{p}^W\right), \bar{p}^W\right) = r$*,*

$$\begin{aligned}
\Omega\left(K^*\left(\bar{p}^W\right), \bar{p}^W\right) &= r - \int_{\bar{t}\left(K^*\left(\bar{p}^W\right), c\right)}^{t^N\left(K^*\left(\bar{p}^W\right)\right)} \left(\rho\left(K^*\left(\bar{p}^W\right); t\right) - c\right) f\left(t\right) dt \\
&\quad + \frac{K^\infty\left(\bar{p}^W\right)}{N} \int_{t^N\left(K^*\left(\bar{p}^W\right)\right)}^{\bar{t}\left(K^*\left(\bar{p}^W\right), \bar{p}^W\right)} \rho_q\left(K^*\left(\bar{p}^W\right); t\right) f\left(t\right) dt \\
&< r = \Omega\left(K^C\left(\bar{p}^W\right), \bar{p}^W\right)
\end{aligned}$$

$\Leftrightarrow$

$$K^*\left(\bar{p}^W\right) > K^C\left(\bar{p}^W\right).$$

■

The price cap does not have a monotonic impact on the Cournot capacity: K^C and $K^C\left(\bar{p}^W\right)$ cannot be compared in general. Zöttl (2011) proves that, if the price cap is sufficiently high, $\frac{\partial K^C}{\partial \bar{p}^W} < 0$. Conversely, Earle et al. (2007)) prove that if the price cap is sufficiently low, $\frac{\partial K^C}{\partial \bar{p}^W} > 0$. In this article, I take the price cap as given, and do not presume it is optimally chosen. Proposition 1 shows that, for any binding price cap, investment is lower than socially optimal.

Proposition 1 holds for any specification of the imperfect competition game in the spot market that yields (i) a spot price higher than marginal cost, hence $\bar{t}(K,c) \leq t^N(K)$, and (ii) a unique symmetric equilibrium of the investment game, with a concave profit function $\Pi^n(k,...,k)$.

3.3 Numerical illustration of underinvestment

Using a numerical example developed in Léautier (2012), we now compare the capacity reductions caused by (i) a binding price cap $(1 - \frac{K^*(\bar{p}^W)}{K^*})$, and (ii) imperfect competition absent any price cap $(1 - \frac{K^C}{K^*})$.

The model is specified as follows: (i) inverse demand is linear: $P(q;t) = a(t) - bq$ where $a(t) = a_0 - a_1 e^{-\lambda_2 t}$, (ii) states of the world are distributed according to $f(t) = \lambda_1 e^{-\lambda_1 t}$, and (iii) rationing is anticipated and proportional. This specification provides an adequate representation of actual demand (shape of load duration curve, and elasticity), while leading to simple expressions, as shown below.

In order to isolate the impact of curtailment (if it occurs), define

$$\Phi(K,c) = \int_{\bar{t}(K,c)}^{+\infty} \left[P\left(\frac{K - (1-\alpha) D\left(p^R,t\right)}{\alpha}; t \right) - c \right] f(t)\, dt$$

the marginal social value of capacity if curtailment does not occur, and

$$I(K) = \int_{\widehat{t}(K)}^{+\infty} \left[P\left(\frac{K - (1-\alpha) D\left(p^R,t\right)}{\alpha}; t \right) - \rho(K;t) \right] f(t)\, dt.$$

the impact of curtailment. Equation (2) is then

$$\Psi(K^*,c) = \Phi(K^*,c) - I(K^*) = r.$$

Substituting in the residual inverse demand (3), then integrating by parts yields:

$$\begin{aligned} \Phi(K,c) &= \int_{\bar{t}(K,c)}^{+\infty} \frac{a'(t)}{\alpha} (1 - F(t))\, dt = \frac{a_1 \lambda_2}{\alpha} \int_{\bar{t}(K,c)}^{+\infty} e^{-(\lambda_1+\lambda_2)t} dt \\ &= \frac{a_1}{\alpha(1+\lambda)} \left(\left(\frac{a_0 - bK - \left(\alpha c + (1-\alpha) p^R\right)}{a_1} \right)^{1+\lambda} \right), \end{aligned}$$

where $\lambda = \frac{\lambda_1}{\lambda_2}$. Similarly, substituting in the residual inverse demand (3) and $VoLL$ (4), then integrating by parts yields:

$$\begin{aligned} I(K) &= \int_{\widehat{t}(K)}^{+\infty}\left(\frac{a^{'}(t)}{\alpha}-\frac{a^{'}(t)}{2}\right)(1-F(t))\,dt = \frac{a_1(2-\alpha)\lambda_2}{\alpha}\int_{\widehat{t}(K)}^{+\infty}e^{-(\lambda_1+\lambda_2)t}dt \\ &= \begin{cases} \frac{(2-\alpha)a_1}{2\alpha(1+\lambda)}\left(\frac{2b}{a_1(2-\alpha)}\left(\bar{K}-K\right)\right)^{1+\lambda} & if\ K<\bar{K} \\ 0 & otherwise \end{cases} \end{aligned}$$

where $b\bar{K} = \frac{2-\alpha}{2}\left(a_0 - p^R\right)$. Curtailment occurs if and only if installed capacity is lower than the threshold $\bar{K}$. Then, its marginal impact is proportional to $\left(\bar{K}-K\right)^{1+\lambda}$.

a_0, a_1, λ, and bQ^{∞}, where $Q^{\infty} = \frac{a_0-p_0}{b}$ is the maximum demand for price p_0, are the parameters to be estimated. λ is estimated by Maximum Likelihood using the load duration curve for France in 2010. The same load duration curve provides an expression of a_0 and a_1 as a function of bQ^{∞}. The average demand elasticity η is then used to estimate bQ^{∞}. As a base case, I choose $\eta = -0.05$ at price $p_0 = 100$ €$/MWh$, which is lower than most other studies, and the upper bound of Lijesen (2007) estimates of real time price elasticity. As a robustness check, I use $\eta = -0.01$, which is the lower bound of Lijesen (2007) estimates.

Following this procedure, Léautier (2012) estimates

$$\begin{matrix} for\ \eta=-0.05 & & for\ \eta=-0.01 \\ \begin{cases} bQ^{\infty} = 3\ 745 \text{ €}/MWh \\ a_0 = 3\ 845 \text{ €}/MWh \\ a_1 = 2\ 472 \text{ €}/MWh \\ \lambda = 1.78 \end{cases} & and & \begin{cases} bQ^{\infty} = 18\ 727 \text{ €}/MWh \\ a_0 = 18\ 827 \text{ €}/MWh \\ a_1 = 12\ 360 \text{ €}/MWh \\ \lambda = 1.78 \end{cases} \end{matrix}.$$

Investment and operating costs are those of a Combined Cycle Gas Combustion Turbine, and provided by the International Energy Agency (median case, IEA (2010)): $c = 49$ €$/MWh$ and $r = 8$ €$/MWh$. The regulated price is $p^R = 100$ €$/MWh$, consistent with most European power markets.

3.3.1 Perfect competition benchmark

Suppose first the market is perfectly competitive ($N \to +\infty$, hence $t^N(K) \to \bar{t}(K,c)$). $\Psi(K^*, c) = r$ is solved numerically for the optimal capacity K^*. For $\eta = -0.05$, curtailment of constant price customers no longer occurs after $\alpha = 10.0\%$ of load is price responsive. K^* is decreasing with α, as illustrated on the table below:

α (%)	5	10	20
K^*/Q^∞	.958	.948	.935

For $\eta = -0.01$, curtailment of constant price customers no longer occurs after $\alpha = 4.6\%$ of load is price responsive. $K^*(\alpha)$ is presented on the table below:

α (%)	5	10	20
K^*/Q^∞	.976	.970	.961

If demand is more elastic (higher $|\eta|$), optimal capacity is lower, thus curtailment occurs until a higher fraction of demand is price responsive.

3.3.2 Capacity reduction due to missing money

We continue to assume that competition is perfect, i.e., $N \to \infty$. Equation (6) then becomes:

$$\Psi\left(K^*\left(\bar{p}^W\right), c\right) - \Psi\left(K^*\left(\bar{p}^W\right), \bar{p}^W\right) = r.$$

$K^*\left(\bar{p}^W\right)$ is then determined numerically. For $\eta = -0.05$, a price cap $\bar{p}^W = 3000$ €$/MWh$, the level in effect in European wholesale markets, is never binding at the optimal capacity. Therefore, the table below presents the relative capacity reduction $\left(1 - \frac{K^*(\bar{p}^W)}{K^*}\right)$ for a much lower cap, maybe resulting from operational practices, set at $\bar{p}^W = 1\ 000$ €$/MWh$, and for different values of α:

α (%) $\bar{p}^W$ (€$/MWh$)	5	10	20
1000	0.9	0.5	0.1

For $\eta = -0.01$, a price cap $\bar{p}^W = 3000$ €$/MWh$ is binding. The table below presents the relative capacity reduction $\left(1 - \frac{K^*(\bar{p}^W)}{K^*}\right)$ for $\bar{p}^W = 1000$ €$/MWh$ and $\bar{p}^W = 3000$ €$/MWh$ and for different values of α:

α (%) / $\bar{p}^W$ (€/MWh)	5	10	20
1000	2.3	1.8	1.2
3000	0.4	0.1	0.0

The tables above illustrate that the missing money has a small impact on installed capacity. Even with demand elasticity at a lower end of estimates $\eta = -0.01$, a low price cap at $\bar{p}^W = 1000$ €$/MWh$ reduces investment by only 2.3% if $\alpha = 5\%$. We would have to assume $\bar{p}^W = 500$ €$/MWh$ and $\alpha = 1\%$ for missing money to reduce capacity by 5.4%.

3.3.3 Capacity reduction due to imperfect competition

Suppose now no price cap is imposed. Given the specification of the demand function and the rationing technology, $\widehat{t}(K) < t^N(K)$ for the relevant values of the parameters. We prove in Appendix B that

$$\Omega(K) = \int_{\widehat{t}(K)}^{+\infty} (v(t) - c) f(t)\, dt.$$

Intuition can be obtained by letting $\bar{p}^W \to +\infty$ and observing $\rho_q = v_q = 0$ for $t \geq \widehat{t}(K)$ in equation (7). We then prove in Appendix B that

$$\Omega(K) = \left(\frac{2b}{(2-\alpha)a_1}\left(\bar{K} - K\right)\right)^{\lambda}\left(p^R - c + \frac{b}{(2-\alpha)}\left(\frac{\bar{K} + \lambda K}{1+\lambda}\right)\right).$$

We are now able to compare the relative capacity reduction attributable to imperfect competition $\left(1 - \frac{K^C}{K^*}\right)$ to the relative capacity reduction attributable to a price cap $\left(1 - \frac{K^*(\bar{p}^W)}{K^*}\right)$.

Consider first the base case elasticity $\eta = -0.05$. As previously observed, a price cap set at $\bar{p}^W = 3\,000$ €$/MWh$ is never binding, hence all underinvestment is caused by strategic supply reduction. Consider now $\bar{p}^W = 1\,000$ €$/MWh$. The table below present the ratio $\left(1 - \frac{K^C}{K^*}\right) / \left(1 - \frac{K^*(\bar{p}^W)}{K^*}\right)$ for different values of α:

α (%) / $\bar{p}^W$ (€/MWh)	3.9	5	10	20
1000	1	1.5	6.4	107.4

As soon as $\alpha \geq 3.9\%$, strategic supply reduction is larger than reduction attributable to missing money. For $\alpha = 5\%$, strategic underinvestment is 1.5 times more important than missing money.

Consider now the lower elasticity $\eta = -0.01$:

α (%) $\bar{p}^W$ (€/MWh)	2.7	5	6.9	10	20
1000	0.3	0.6	1	1.8	6.2
3000	1	3.1	6.6	19.0	68.2

A price cap $\bar{p}^W = 3\ 000$ €/MWh is binding for low values of α. However, as soon as $\alpha \geq 2.7\%$, strategic supply reduction is larger than reduction attributable to missing money. For $\alpha = 5\%$, strategic underinvestment is 3.1 times more important than missing money. A much lower price cap $\bar{p}^W = 1\ 000$ €/MWh is the main driver of underinvestment up to $\alpha = 6.9\%$.

These results are of course dependant on the joint specification of demand and uncertainty, and on the parameters estimated, thus need to be validated using a richer specification. This is an important avenue for further research. Still, I believe these results will prove robust. The impact of imperfect competition is not overstated. While Cournot competition is an extremely severe form of imperfect competition, experience in England and Wales before 2000 and California in 2000/2001 suggest competition in power markets can be highly imperfect. Second, the impact of the price cap is not understated. The joint specification of demand and uncertainty yields maximum oligopoly prices higher than the price caps. For $\alpha = 5\%$, $\rho_{\max} = \lim_{t \to +\infty} \rho\left(K^C, t\right)$ is 4 000 €/MWh for $\eta = -0.05$ and 14 000 €/MWh for $\eta = -0.01$.

Thus, this analysis casts a different light on reliability in the power industry.

Part II

"Dual markets" designs

This part analyzes two "dual markets" designs: Installed Capacity Markets, where physical capacity certificates are exchanged, and markets for financial reliability options, that rely on financial obligations. To simplify the notation and analysis, operating reserves are ignored: as will be proven in Section 7, including them would not modify the economic insights. All units (old and new) receive the same compensation in these markets.

The timing is common to all "dual markets" designs:

1. The SO designs the rules of the energy and capacity (or options) markets. All parameters are set
2. Producers sell physical capacity certificates or financial reliability options, according to the rules set up previously
3. Producers build new capacity if needed
4. The spot markets are played. In each state, producers compete à la Cournot facing $\rho(Q;t)$, given their installed capacity and their capacity obligation (physical or financial)

Steps 2 and 3 can be inverted or simultaneous: generators first build the plants, then sell physical capacity certificates or financial reliability options, or build and sell simultaneously. As proven below, the analysis is identical for all three timings.

4 Physical capacity certificates

The SO imposes price cap $\bar{p}^W$ on the energy markets and procures at least K^* physical capacity certificates from producers. ϕ^n and $\Phi = \sum_{m=1}^{N} \phi^m$ are respectively the certificates sold by producer n and the aggregate volume of certificates sold. In practice, SOs offer a "smoothed" (inverse) demand curve:

$$H(\Phi) = \begin{cases} r & if \quad \Phi \leq K^* \\ h(\Phi) & if \quad K^* < \Phi < K^* + \Delta\bar{K} \\ 0 & if \quad \Phi \geq K^* + \Delta\bar{K} \end{cases}$$

where (i) r, the capital cost of capacity, is the maximum price the SO is offering for capacity, (ii) $\Delta\bar{K} > 0$ is an arbitrary capacity increment, and (iii) $h(.)$ is such that $H(.)$ is C^2, except maybe at

K^* and $K^* + \Delta\bar{K}$, $h^{'}(\Phi) < 0$, $2h'(\Phi) + \phi h''(\Phi) < 0$ for all ϕ, and

$$\left|h^{'}(K^*)\right| \geq \frac{Nr}{K^*}. \tag{8}$$

As will be discussed below, condition (8) simplifies the exposition, but is not essential. It is met in practice. For example, Cramton and Ockenfels (2011) suggest a linear form for $h(.)$ with $\frac{\Delta\bar{K}}{K^*} = 4\%$. Condition (8) is then equivalent to $N\frac{\Delta\bar{K}}{K^*} \leq 1$, and holds as long as less than 25 producers compete.

Proposition 2 *If the SO imposes and monitors that the installed capacity exceeds the capacity certificates sold by each generator: $k^n \geq \phi^n$, then (i) producers issue as many credits as they install capacity, and (ii) $K^*$is the unique symmetric equilibrium investment level. Compared to the no installed capacity market situation, producer's profit and overall welfare are increased.*

Proof. *The full proof is presented in Appendix C. Existence of a physical capacity certificates market alone does not alter investment incentives. The SO must impose $k^n \geq \phi^n$, otherwise K^C remains the installed capacity.*

If she does, producers sell exactly as many certificates as they have installed capacity (or install exactly as much capacity as they have sold certificates) since incremental capacity is unprofitable unless it collects capacity markets revenues. Then, since $k^n = \phi^n$ at the equilibrium, producer n program is:

$$\max_{k^n} \Pi^n_{CM}(k^n; \mathbf{k}_{-n},) = \Pi^n(k^n; \mathbf{k}_{-n}) + k^n H(K)$$

Given the shape of the inverse demand function $H(.)$, $k^n = \frac{K^*}{N}$ for all n is the unique symmetric equilibrium, and producers' profit is:

$$\Pi^n_{CM}\left(\frac{K^*}{N}, ..., \frac{K^*}{N}\right) = \Pi^n\left(\frac{K^*}{N}, ..., \frac{K^*}{N}\right) + \frac{K^*}{N}r.$$

Then, since $\Pi^n(k, ..., k)$ is concave and $K^C \leq K^*$:

$$\Pi^n\left(\frac{K^C}{N}, ..., \frac{K^C}{N}\right) \leq \Pi^n\left(\frac{K^*}{N}, ..., \frac{K^*}{N}\right) + \left(\frac{K^C - K^*}{N}\right)\frac{\partial \Pi^n}{\partial k}\left(\frac{K^*}{N}, ..., \frac{K^*}{N}\right)$$

$\Leftrightarrow$

$$\begin{aligned}\Pi_{CM}^{n}\left(\frac{K^{*}}{N},...,\frac{K^{*}}{N}\right) &\geq \Pi^{n}\left(\frac{K^{C}}{N},...,\frac{K^{C}}{N}\right)-\frac{K^{C}}{N}\frac{\partial\Pi^{n}}{\partial k}\left(\frac{K^{*}}{N},...,\frac{K^{*}}{N}\right)+\frac{K^{*}}{N}\left(\frac{\partial\Pi^{n}}{\partial k}\left(\frac{K^{*}}{N},...,\frac{K^{*}}{N}\right)+r\right)\\ &> \Pi^{n}\left(\frac{K^{C}}{N},...,\frac{K^{C}}{N}\right)\end{aligned}$$

since $\frac{\partial\Pi^{n}}{\partial k}\left(\frac{K^{*}}{N},...,\frac{K^{*}}{N}\right)<0$ *and* $\frac{\partial\Pi^{n}}{\partial k}\left(\frac{K^{*}}{N},...,\frac{K^{*}}{N}\right)+r=\Omega\left(K^{*}\right)>0$.

Producers' profits increase compare to the no installed capacity market situation. Finally, since overall welfare $W(K)$ *increases up to to* $K=K^{*}$, $W(K^{*})\geq W\left(K^{C}\right)$. ■

Capacity markets do not automatically restore investment incentives. The *SO* must ensure that producers cannot sell short, i.e., sell more certificates than their installed capacity. This observation is by now well accepted by economists, policy makers, and *SOs*. It has been articulated for example by Wolak (2006). As a result, *SOs* monitor that existing generation assets providing certificates are still operational, and that planned capacity having received certificates has indeed be installed. *SOs* then impose a penalty on producers that, when requested, do not offer in the spot market energy up to the certificates they have sold forward. This monitoring process is extremely expensive.

Matters are slightly different in this model, since producers exercise market power by reducing capacity *ex ante*, and not by withholding output on-peak. Spot-market penalty is replaced in this model by an obligation not to sell certificates short, i.e., to control more physical capacity (existing or planned) than certificates sold.

Physical capacity markets increase overall welfare, and also increase transfers from customers to producers. To my best knowledge, this result is original to this work. It is also very general. Denote K^{E} (not necessarily equal to K^{*}) the equilibrium capacity including the certificates markets. As long as $\Pi^{n}(k,...,K)$ is concave, and $K^{E}>K^{C}$, the marginal value of capacity at the K^{E} is negative: $\frac{\partial\Pi^{n}}{\partial k}\left(\frac{K^{E}}{N},...,\frac{K^{E}}{N}\right)<0$. The equilibrium price in the capacity market (r in this case) must compensate for this negative marginal value, otherwise K^{E} would not be an equilibrium: $\frac{\partial\Pi^{n}}{\partial k}\left(\frac{K^{E}}{N},...,\frac{K^{E}}{N}\right)+r\geq 0$. This is sufficient for the proof.

Coupled with the underinvestment analysis of Part I, Proposition 2 casts a new light on physical certificates markets: the common wisdom is that these markets are required to compensate the missing money. If, as suggested by the illustrative numerical example presented in Part I, strategic supply reduction is indeed the main driver of underinvestment, physical certificates markets then sur-remunerate this strategic behavior. This suggests that competition authorities should play a greater role than SOs in alleviating underinvestment.

If condition 8 is not met, the aggregate capacity at the unique symmetric equilibrium is $K^C_{CM} \in \left(K^*, K^* + \Delta\bar{K}\right]$. Welfare increases if and only if $\Delta\bar{K}$ is small enough that $W\left(K^* + \Delta\bar{K}\right) \geq W\left(K^C\right)$.

5 Financial reliability options

Financial contracts constitute another approach to alleviate underinvestment. This Section examines financial reliability options, proposed by Cramton and Stoft (2006 and 2008), and more recently Cramton and Ockenfels (2011). Options and not forward contracts are the financial instruments analyzed here, since Chao and Wilson (2005), that examine a slightly different option design, argue that options are in general preferable. These options constitute an insurance against spot energy prices higher than a pre-agreed strike price $\bar{p}^S$, sold by producers to customers. If the spot price $p(t)$ is lower than $\bar{p}^S$, producer n does not make any payment. If $p(t) > \bar{p}^S$, producer n pays $\left(p(t) - \bar{p}^S\right)$ times a fraction of the realized demand equal to his fraction of the total options sale.

The SO does not impose a cap on wholesale prices, and runs an auction for financial reliability options. θ^n and $\Theta = \sum_{m=1}^{N} \theta^m$ are respectively the options sold by producer n and the aggregate volume of options sold. The notation is identical to the capacity market case, except that the subscript $_{RO}$ is added when appropriate. A very simple auction setup is assumed, similar to the one suggested by Cramton and Stoft (2008): the SO determines the volume she desires to purchase, assumed to be K^*, sets the capital cost of capacity r as the reserve price for the auction, and proposes a downward sloping inverse demand curve for options:

$$H_{RO}(\Theta) = \begin{cases} r & if \quad \Theta \leq K^* \\ h_{RO}(\Theta) & if \quad K^* < \Theta < K^* + \Delta\bar{K}_{RO} \\ 0 & if \quad \Theta \geq K^* + \Delta\bar{K}_{RO} \end{cases}$$

where (i) $\Delta\bar{K}_{RO} > 0$ is an arbitrary capacity increment, and (ii) $h_{RO}(.)$ is such that $H_{RO}(.)$ is C^2, except maybe at K^* and $K^* + \Delta\bar{K}$, $h'_{RO}(\phi) < 0$, $2h'_{RO}(\phi) + \phi h''_{RO}(\Phi) < 0$ for all ϕ, and $h_{RO}(.)$

verifies condition (8). To limit the potential exercise of market power, Cramton and Ockenfels (2011) propose the SO impose $\theta^n \geq k^n$: all capacity must be committed forward through option sales.

We assume $\bar{p}^S$ satisfies

$$\Psi\left(K^C\left(\bar{p}^S\right), \bar{p}^S\right) \leq r. \tag{9}$$

Condition (9) simplifies the exposition, as it guarantees that $\Theta = K^*$ is the unique equilibrium of the options market. As shown by Zöttl (2011), if $\bar{p}^S$ is high enough, $\frac{\partial K^C}{\partial \bar{p}^S} \geq 0$, thus $K^C\left(\bar{p}^S\right)$ converges, and $\lim_{\bar{p}^S \to +\infty} \Psi\left(K^C\left(\bar{p}^S\right), \bar{p}^S\right) = 0$. Condition (9) is met for $\bar{p}^S$ high enough.

When the spot price exceeds the strike price, price-reactive consumers then pay $\bar{p}^S$ as the effective price, i.e., they know when making their consumption decision they receive rebate $\max\left(\rho\left(Q, t\right) - \bar{p}^S, 0\right)$ per unit of energy purchased. Then, actual demand does not depend on the spot price, which leads to rationing.

As previously, $\bar{t}\left(K, \bar{p}^S\right)$ is the first state of the world where the spot price exceeds the strike price, and is defined by $\rho\left(K, \bar{t}\left(K, \bar{p}^S\right)\right) = \bar{p}^S$. The capacity constraint is assumed to be binding before the spot price reaches the option price: $t^N\left(K\right) \leq \bar{t}\left(K, \bar{p}^S\right)$.

Chao and Wilson (2005) examine a slightly different market structure: they consider physical options paired (or not) with a complementary price insurance, and compute the linear supply function equilibrium for options forward sales and power spot sales. Their findings are aligned with those presented below.

5.1 Expected profits with financial reliability options

The producers profit function is characterized below:

Lemma 1 *The expected profit of producer n is:*

$$\Pi^n_{RO}\left(k^n, \theta^n; \mathbf{k}_{-n}, \boldsymbol{\theta}_{-n}\right) = \theta^n H_{RO}\left(\Theta\right) + \Pi^n\left(k^n; \mathbf{k}_{-n}\right) + \left(k^n - \frac{\theta^n}{\Theta} K\right) \Psi\left(K, \bar{p}^S\right) \tag{10}$$

with the convention that $\bar{p}^S$ acts as the price cap in Π^n.

Proof. *Producer n receives the revenues from options sale $\theta^n H_{RO}\left(\Theta\right)$, plus profits from the energy market. A possible decomposition is as follows: first, the producer receives profit $\Pi^n\left(k^n; \mathbf{k}_{-n}\right)$ pre-*

viously computed, assuming a price cap at $\bar{p}^S$*. Second, since there is no price cap, he receives the difference between the spot price* $\rho(K,t)$ *and the cap* $\bar{p}^S$ *for every unit produced when the price exceeds* $\bar{p}^S$*. Since we have assumed* $t^N(K) \leq \bar{t}(K,\bar{p}^S)$*, he produces his entire capacity* k^n*, hence he receives* $k^n \int_{\bar{t}(K,\bar{p}^S)}^{+\infty} \left(\rho(K,t) - \bar{p}^S\right) f(t)\,dt = k^n \Psi\left(K,\bar{p}^S\right)$*.*

Finally, when the spot price exceeds the strike price $\bar{p}^S$*, each generator must pay* $\left(\rho(K,t) - \bar{p}^S\right)$ *times his fraction* $\frac{\theta^n}{\Theta}$ *of the total demand. Since we have assumed* $t^N(K) \leq \bar{t}(K,\bar{p}^S)$*, total demand is equal to total capacity* K *and the payment is proportional to* $\frac{\theta^n}{\Theta}K$*. Total expected payment from generator* n *is thus:* $\frac{\theta^n}{\Theta}K \int_{\bar{t}(K,\bar{p}^S)}^{+\infty} \left(\rho(K,t) - \bar{p}^S\right) f(t)\,dt = \frac{\theta^n}{\Theta}\Psi\left(K,\bar{p}^S\right)$*. Summing these terms yields equation (10).* ■

The profit realized in states higher than $\bar{t}\left(K,\bar{p}^S\right)$ is $\pi^n_{RO}(K;t) = k^n \left(\left(1 - \frac{\theta^n}{\Theta}\frac{K}{k^n}\right)\rho(K;t) + \frac{\theta^n}{\Theta}\frac{K}{k^n}\bar{p}^S - c\right)$. Producers face a weighted average of the spot price and the option price, hence are less sensitive to an increase in spot price. Consistent with Allaz and Villa (1993) and Chao and Wilson (2005), a producer holding forward contracts faces lower incentives to exert market power in the spot market.

5.2 Equilibrium capacity with financial reliability options

Proposition 3 *Reliability options reduce but do not solve the underinvestment problem.* K^C_{RO}*, the unique symmetric equilibrium of the options and investment game, verifies:*

$$K^C\left(\bar{p}^S\right) \leq K^C_{RO} < K^*$$

with equality occurring when $N = 1$*.*

Proof. *Appendix D proves that, if producers invest first then sell options, there exists a unique symmetric equilibrium. If producers sell options first, then invest, we assume existence of a symmetric equilibrium, then prove its unicity. In all cases, equilibrium capacity satisfies:*

$$\frac{\partial \Pi^n_{RO}}{\partial k^n}\left(\frac{K^C_{RO}}{N}, ..., \frac{K^C_{RO}}{N}\right) = \Omega\left(K^C_{RO},\bar{p}^S\right) - r + \frac{N-1}{N}\Psi\left(K^C_{RO},\bar{p}^S\right) = 0 \tag{11}$$

Then,

$$\Omega\left(K^C_{RO},\bar{p}^S\right) = r - \frac{N-1}{N}\Psi\left(K^C_{RO},\bar{p}^S\right) \leq r.$$

Hence $K_{RO}^{C} \geq K^{C}\left(\bar{p}^{S}\right)$. *Then,*

$$\begin{aligned}
\frac{\partial \Pi_{RO}^{n}}{\partial k^{n}}\left(\frac{K^{*}}{N},...,\frac{K^{*}}{N}\right) &= -\int_{\bar{t}(K,c)}^{t^{N}(K)}\left(\rho\left(K^{*};t\right)-c\right)f\left(t\right)dt+\frac{K^{*}}{N}\int_{t^{N}(K)}^{\bar{t}\left(K,\bar{p}^{S}\right)}\rho_{q}\left(K^{*};t\right)f\left(t\right)dt \\
&\quad -\frac{1}{N}\Psi\left(K^{*},\bar{p}^{S}\right) \\
&< 0.
\end{aligned}$$

Then $K^{*} > K_{RO}^{C}$ *since we prove in Appendix D that* $\Pi_{RO}^{n}\left(\frac{K}{N},...,\frac{K}{N}\right)$ *is concave.* ■

For $N > 1$, reliability options are more effective than physical certificates: the resulting installed capacity is higher than the Cournot capacity, while physical certificates alone (i.e., without the added no short sale obligation) have no impact on installed capacity. However, reliability options are not sufficient to restore optimal investment incentives.

This result may appear surprising, since reliability options impose a penalty of $\left(\rho\left(K;t\right)-\bar{p}^{S}\right)$ on each unit a producer is "short" energy. However, a closer examination of the mechanism reveals that, at the symmetric equilibrium, this penalty represents only $\frac{N-1}{N}\left(\rho\left(K;t\right)-\bar{p}^{S}\right)$, which is not sufficient to fully compensate for the "missing money" $\left(\rho\left(K;t\right)-\bar{p}^{S}\right)$.

Proposition 3 mirrors Allaz and Villa (1993) analysis of the interaction between spot and forward markets: assuming Cournot competition in both, they show that introducing forward markets reduces but does not eliminate market power, and has not impact on a monopoly ($N = 1$).

5.3 Equivalence between the two "dual markets" designs when "no short sale" condition is added

Proposition 4 *If* (i) *the SO imposes and monitors that the installed capacity exceeds the options sold by each generator:* $\theta^{n} \leq k^{n}$, (ii) *the wholesale price cap in the capacity market is set equal to the strike price of the reliability option* $\left(\bar{p}^{S}=\bar{p}^{W}\right)$ *and satisfies condition* (9)*, and* (iii) *the demand functions for reliability options and for capacity credits are identical and satisfy condition* (8)*, then* (i) *producers sell as many options as they install capacity, and* (ii) *both market designs yield the same symmetric equilibrium.*

Proof. *The proof of the first point is identical to that of Proposition 2. Then, substituting* $\theta^{n} = k^{n}$

into equation (10) yields:

$$\Pi_{RO}^{n}\left(k^{n};\mathbf{k}_{-n}\right)=\Pi^{n}\left(k^{n};\mathbf{k}_{-n}\right)+k^{n}H_{RO}\left(K\right)$$

Hence, $\Pi_{RO}^{n}=\Pi_{CM}^{n}$ *if* $\bar{p}^{S}=\bar{p}^{W}$ *and* $H_{RO}\left(.\right)=H\left(.\right)$*. Hence the equilibria are identical.* ■

As mentioned earlier, since producers sell exactly as many options as their installed capacity (or install as much capacity as they sold options), the profit net of the payment on the option is equivalent to a cap on prices. Therefore, if the "technical parameters" are identical, both approaches are equivalent. Policy-makers considering implementing them should therefore select one or the other approach based on other factors not included in this analysis.

Part III

"Energy-only" market designs

6 Average $VoLL$ pricing

Hogan (2005) describes average $VoLL$ pricing as follows: the SO approximates the true demand curve by valuing *all* lost load from constant price customers at average $VoLL$ $\overline{v}$, determined exogenously. Thus, when load is shed, producers receive average $VoLL$ $\overline{v}$. Estimates of average $VoLL$ range between between 10 000 and 20 000 €$/MWh$. Does this mechanism restore investment incentives?

Average $VoLL$ pricing has limited impact on strategic supply reduction: if generation is imperfectly competitive, underinvestment will remain, although it may be slightly reduced as shown by Zöttl (2011).

Consider now the perfectly competitive case. The analysis presented in Section 3 proves that, since average $VoLL$ is a cap on wholesale prices, it reduces investment incentives. However, the cap is high enough that the distortion is negligible.

Underinvestment arises since, when curtailment must occur, producers receive $\overline{v}$ which is by construction lower than the true marginal value of energy $\rho\left(K,t\right)$. What if the SO sets $\overline{v}$ exactly at the average $VoLL$, conditional on curtailment occurring? This approach is suggested Stoft (2002, pages 136-138 and 157-159):

> "the regulator acts as a surrogate for load by setting the price when power is being shed. The appropriate price equals the average value that shed load places on power".

This question is of course theoretical, as computing $VoLL$, let alone average $VoLL$ conditional on curtailment, is impossible in practice. Even if the SO could compute the correct average $VoLL$, underinvestment would still occur, as proven below:

When curtailment occurs, producers receive the average $VoLL$, which eliminates the wedge between private and public values. However, the profit function is discontinuous at $\hat{t}(K)$, the first state when curtailment occurs: just before $\hat{t}(K)$, producers receive $\rho\left(K, \hat{t}(K)\right)$. Immediately after $\hat{t}(K)$, producers receive $\overline{v}$. Since $\overline{v}$ is the average $VoLL$ over states of the world higher than $\hat{t}(K)$, $\overline{v} > \rho\left(K, \hat{t}(K)\right)$. This induces under-investment: a marginal increase in capacity reduces the occurrence of curtailment, hence increases $\hat{t}(K)$, thus producers receive negative margin $\left(\rho\left(K, \hat{t}(K)\right) - \overline{v}\right)$ on these states of the world. This point is formally proven in Appendix E. The impact however, is likely to be small.

7 Energy cum operating reserves market

SOs must secure operating reserves to protect the system against catastrophic failure. Hogan (2005) suggests that remuneration of these operating reserves can solve the missing money problem.

The representation of operating reserves is that of Borenstein and Holland (2005). For simplicity, only one type of reserves is considered, the non-spinning one (i.e., plants that are not running, but can start up and produce energy within a short pre-agreed time frame). Since the plant is not running, the marginal cost of providing reserves is normalized to zero. In reality, *SOs* run multiple markets for operating reserves, for example, spinning, 10-minutes, 30-minutes. The economic insights are not modified, as long as the no-arbitrage condition presented below holds.

Hogan (2005) proposes that the SO runs a single market for energy and operating reserves. Generating units called to produce receive the wholesale price $w(t)$, generating units that provide operating reserves receive the wholesale price $w(t)$ less the marginal cost of generation c, assumed to be perfectly known by the SO. Generators are therefore indifferent between producing energy or providing reserves, an essential condition (Borenstein and Holland (2005)). When an unscheduled generation outage occurs, operating reserves produce energy and receive the full price $w(t)$.

Operating reserves requirements are expressed as a percentage of demand, denoted $h(t)$, and taken as given here[6]. Defining the optimal $h(t)$ is beyond the scope of this work. Joskow and Tirole (2007) show the optimal reserve ratio increases with the state of the world; hence $h(t)$ is assumed to be nondecreasing.

The retail price $p(t)$ must be higher than wholesale price $w(t)$ to cover generators' revenues from the operating reserves market. A natural choice is to directly include the cost of reserves in the retail price faced by "price reactive" customers[7]:

$$p(t) = w(t) + h(t)(w(t) - c)$$

$\Leftrightarrow$

$$p(t) - c = (1 + h(t))(w(t) - c) \tag{12}$$

Throughout this section, the retail and wholesale prices are assumed to be related by equation (12). The notation and model structure are identical to the previous Sections, except that the subscript $_{OR}$ is added when appropriate.

Only the fraction $\frac{1}{1+h(t)}$ of installed capacity is used to meet demand in state t, hence $\frac{K}{1+h(t)}$ and not K is the output appearing in the function $\rho(.;t)$. Thus, the marginal social value of capacity in state t is

$$w(K,t) - c = \frac{p(t) - c}{1 + h(t)} = \frac{\rho\left(\frac{K}{1+h(t)};t\right) - c}{1 + h(t)}.$$

The marginal social value of capacity is

$$\Psi_{OR}(K) = \int_{\bar{t}_{OR}(K,c)}^{+\infty} \frac{\rho\left(\frac{K}{1+h(t)};t\right) - c}{1 + h(t)} f(t)\, dt$$

[6] In practice, various metrics for operating reserves are used, including absolute values expressed in MW. Expressing reserves as a percentage of peak demand simplifies the analysis while preserving the main economic intuition.

[7] Borenstein and Holland (2005) show it to be the perfect competition outcome.

where $\bar{t}_{OR}(K,c)$ is uniquely defined[8] by $\rho\left(\frac{K}{1+h(t)};\bar{t}_{OR}(K,c)\right)=c$.

Consider now the producers' problem. By construction, producers are indifferent between producing energy or providing reserves. In state t, they offer $s^n(t)$ into the energy cum operating reserves market. $S(t)=\sum_{n=1}^{N}s^n(t)$ is the total offer. Energy available to meet demand is $Q(t)=\frac{S(t)}{1+h(t)}$. The SO then (i) verifies that $s^n(t)\leq k^n$, and (ii) allocates each $s^n(t)$ between energy $q^n(t)$ and reserves $b^n(t)$. Producer n profit is then:

$$\begin{aligned}\pi^n(t) &= (q^n(t)+b^n(t))(w(t)-c)\\ &= \frac{s^n(t)}{1+h(t)}\left(\rho\left(\frac{S(t)}{1+h(t)}\right)-c\right)\end{aligned}$$

since (i) energy and operating reserves receive same net revenue by construction, and (ii) wholesale $(w(t))$ and retail $\left(\rho\left(\frac{S(t)}{1+h(t)}\right)\right)$ prices are linked by equation (12). The problem is then isomorphic to the previous Sections, except that $\frac{s^n(t)}{1+h(t)}$ replaces production $q^n(t)$.

$t_{OR}^N(K)$, the first on-peak state of the world under imperfect competition, is uniquely defined[9] by $\rho\left(\frac{K}{1+h(t)};t\right)+\frac{1}{N}\frac{K}{1+h(t)}\rho\left(\frac{K}{1+h(t)};t\right)=c$.

The SO imposes a wholesale price cap $\overline{v}$ equal to the best estimate of $VoLL$. $\bar{t}_{OR}(K,\overline{v})$, the first state of the world where the cap may be binding, is uniquely defined by $\rho\left(\frac{K}{1+h(t)};\bar{t}_{OR}(K,\overline{v})\right)=\overline{v}$. As in the previous Sections, $\overline{v}$ is assumed to be binding after the capacity constraint under imperfect competition: $t_{OR}^N(K)\leq\bar{t}_{OR}(K,\overline{v})$. The inverse demand function for producers is then: $\rho\left(\frac{K}{1+h(t)};t\right)$ as long as price cap is not reached, and a horizontal inverse demand at $\overline{v}$ afterwards.

Anticipating on the next Lemma, the marginal value of capacity for a producer at the symmetric equilibrium is

$$\begin{aligned}\Omega_{OR}(K) &= \Psi_{OR}(K,c)-\Psi_{OR}(K,\overline{v})-\int_{\bar{t}_{OR}(K,c)}^{t_{OR}^N(K)}\frac{\rho\left(\frac{K}{1+h(t)};t\right)-c}{1+h(t)}f(t)\,dt\\ &\quad+\frac{1}{N}\int_{t_{OR}^N(K)}^{\bar{t}_{OR}(K,\overline{v})}\frac{K}{1+h(t)}\rho_q\left(\frac{K}{1+h(t)};t\right)f(t)\,dt\end{aligned}$$

[8]Since $h(t)$ is nondecreasing, $m_1(K;t)=\rho\left(\frac{K}{1+h(t)};t\right)$ is increasing in t: $\frac{\partial m_1}{\partial t}=-\rho_q\frac{Kh'(t)}{(1+h(t))^2}+\rho_t>0$.

[9]Similarly, $m_2(t)=\rho\left(\frac{K}{1+h(t)};t\right)+\frac{1}{N}\frac{K}{1+h(t)}\rho\left(\frac{K}{1+h(t)};t\right)$ is increasing in t since $m_2'(t)=-\left(\frac{N+1}{N}\rho_q+\frac{1}{N}\frac{K}{1+h(t)}\rho_{qq}\right)\frac{Kh'(t)}{(1+h(t))^2}+\rho_t>0$.

Lemma 2 *The socially optimal capacity K^*_{OR} for a energy-cum-operating reserves market is defined by:*

$$\Psi_{OR}\left(K^*_{OR}\right) = r \tag{13}$$

There exists a unique symmetric equilibrium for which each generator invests $k^C_{OR} = \frac{K^C_{OR}}{N}$ defined by:

$$\Omega_{OR}\left(K^C_{OR}\right) = r \tag{14}$$

Proof. *The proof is presented in Appendix F.* ■

Proposition 5 *If the SO runs an energy cum operating reserves market and imposes a price cap $\overline{v}$, underinvestment occurs unless* (i) *generation is perfectly competitive,* and (ii) *the price cap is never expected to be binding.*

Proof. *The result follows immediately from equations* (14) *and* (13). ■

Including an operating reserve market leads to the same investment incentives as average $VoLL$ pricing. This result is surprising: one would have expected the operating reserves market to alleviate the missing money problem, since (i) all producing units receive a higher price, and (ii) units providing capacity but not energy are remunerated.

However, Lemma 2 shows these two effects are already included in the determination of the socially and privately optimal capacities K^*_{OR} and K^C_{OR}. Then, units providing reserve capacity receive the same profit $(w(t) - c)$ as units producing electricity, to avoid arbitrage between markets. No additional profit is generated. The operating reserves market remunerates reserves, which are needed, not capacity investment.

8 Conclusion

This article formally analyzes the causes of underinvestment, and the various corrective market designs that have been proposed and implemented. It yields four main analytical findings. First, using a simple numerical example, (a linear demand function, calibrated on the French power load duration curve), strategic supply reduction is shown to be a more important cause of underinvestment than the imposition of a price cap. For example, a 3 000 €$/MWh$ price cap, which is the current level in

European markets, is shown to never be binding at the optimal capacity. Any underinvestment would thus be entirely caused by strategic supply reduction. If a low 1 000 €$/MWh$ price cap is imposed, strategic supply reduction is the primary cause of underinvestment as soon as more than 3.9% of the load is price responsive. Second, physical capacity certificates markets implemented in the United States restore optimal investment, but increase producers' profits beyond the imperfect competition level. Third, financial reliability options, proposed in many markets, fail to restore investment incentives. If a "no short sale" condition is added, they are equivalent to physical capacity certificates. Finally, if competition is perfect, energy only markets yield a negligible underinvestment compared to the optimum. Taken together, these findings suggest that, to ensure generation adequacy, policy makers should put more effort on enforcing competitive behavior in the energy markets, and less on designing additional markets.

These results provide a sound basis for policy makers decision making. Different avenues for further work would increase their applicability. First, confirm for other specifications of demand the importance of imperfect competition in leading to a binding price cap.

Second, expand the economic models to other types of technologies: (*i*) intermittent and uncontrollable production technologies such as photovoltaic and on- and off-shore wind mills, which will provide an increasingly important share of power supply; (*ii*) reservoir hydro production, which has almost zero marginal cost, but limited overall production capacity, and (*iii*) voluntary curtailment, i.e., consumers reducing their consumption upon the SO's request.

Finally, expand the model to multiple investment periods. Observation suggests the power industry, like many capital-intensive industries, displays cycle of over- and under-investment ("boom bust" cycles). Understanding how various market designs perform in a dynamic setting is therefore extremely important.

9 Acknowledgments

I am grateful to Claude Crampes, Bill Hogan, Patrick Mouttapa, Marie-Anne Plagnet, David Salant, Jean Tirole, and participants to the 2010 IDEI conference on the economics of energy markets for insightful discussions and comments on previous drafts, and for a grant from Electricité de France that supported this research. All errors and omissions are my responsibility.

References

[1] B. Allaz and J.-L. Vila. Cournot competition, futures markets and efficiency. *Journal of Economic Theory*, 59:1–16, 1993.

[2] M. Boiteux. La tarification des demandes en pointe: Application de la théorie de la vente au cout marginal. *Revue Générale d'Electricité*, 58:321–40, 1949.

[3] A. Boom. Vertically integrated firm's investments in electricity generating capacities. *International Journal of Industrial Economics*, 27:544–551, October 2009.

[4] S. Borenstein and S. Holland. On the efficiency of competitive electricity markets with time-invariant retail prices. *RAND Journal of Economics*, 36:469–493, 2005.

[5] H. Chao and R. Wilson. Priority service: Pricing, investment, and market organization. *American Economic Review*, 77:899–916, 1987.

[6] H. Chao and R. Wilson. Resource adequacy and market power mitigation via option contracts. Technical Report 1010712, Electric Power Research Institute, 2005. http://stoft.com/metaPage/lib/Chao-Wilson-2003-04-resource-adequacy-options.pdf.

[7] P. Cramton and A. Ockenfels. Economics and design of capacity markets for the power sector. Mimeo, University of Maryland, May 2011. http://www.cramton.umd.edu/papers2010-2014/cramton-ockenfels-economics-and-design-of-capacity-markets.pdf.

[8] P. Cramton and S. Stoft. The convergence of market designs for adequate generation capacity. White Paper for the California Electricity Oversight Board, April 2006.

[9] P. Cramton and S. Stoft. Forward reliability markets: Less risk, less market power, more efficiency. *Utilities Policy*, 16(3):194–201, September 2008.

[10] D. Finon and V. Pignon. Capacity mechanisms in imperfect electricity markets. *Utilities Policy*, 16(3), September 2008.

[11] W. Hogan. On an "energy only" electricity market design for resource adequacy. Mimeo, Center for Business and Government, John F. Kennedy School of Government, Harvard University, September 2005.

[12] IEA. *Projected Cost of Generating Electricity.* OECD/IEA, 2010.

[13] P. Joskow and J. Tirole. Reliability and competitive electricity markets. *RAND Journal of Economics*, 38(1):60–84, Spring 2007.

[14] P. L. Joskow. Competitive electricity markets and investment in new generating capacity. in: Helm, D. (Ed.), The New Energy Paradigm, Oxford University Press, 2007.

[15] T.-O. Léautier. Is mandating smart meters smart? Mimeo, Toulouse School of Economics, June 2012.

[16] M. Lijesen. The real-time price elasticity of electricity. *Energy Economics*, 29:249–258, 2007.

[17] F. Murphy and Y. Smeers. Generation capacity expansion in imperfectly competitive restructured electricity markets. *Operations Research*, 53(4):646–661, July-August 2005.

[18] Ofgem. Project discovery – updated energy markets scenarios. htttp://www.ofgem.ogv.uk/, February 2010. Ref 16a/10.

[19] S. Stoft. *Power System Economics.* Wiley-Interscience, 2002.

[20] F. Wolak. What's wrong with capacity markets? http://stoft.com/metaPage/lib/WolaK-2004-06-contract-adequacy.pdf, June 2004. Mimeo, Department of Economics, Stanford University.

[21] G. Zoettl. On optimal scarcity prices. *International Journal of Industrial Organization*, 29(5):589–605, September 2011.

A Properties of the $VoLL$

When capacity is constrained and if curtailment occurs, total differentiation of the energy balance

$$K = \alpha D\left(v\left(p^R, \gamma^*; t\right), t\right) + (1-\alpha)\, \mathcal{D}\left(p^R, \gamma; t\right)$$

with respect to K yields

$$\frac{\partial \gamma^*}{\partial K} = \frac{1}{\alpha \frac{\partial D}{\partial p}\frac{\partial v}{\partial \gamma} + (1-\alpha)\frac{\partial \mathcal{D}}{\partial \gamma}}.$$

$\frac{\partial v}{\partial \gamma} \leq 0$ guarantees that $\frac{\partial \gamma^*}{\partial K} > 0$ and $\frac{\partial v}{\partial K} = \frac{\partial v}{\partial \gamma}\frac{\partial \gamma^*}{\partial K} \leq 0$.

Total differentiation of the energy balance with respect to t yields

$$\frac{\partial \gamma^*}{\partial t} = -\frac{\alpha\left(\frac{\partial D}{\partial p}\frac{\partial v}{\partial t} + \frac{\partial D}{\partial t}\right) + (1-\alpha)\frac{\partial \mathcal{D}}{\partial t}}{\alpha \frac{\partial D}{\partial p}\frac{\partial v}{\partial \gamma} + (1-\alpha)\frac{\partial \mathcal{D}}{\partial \gamma}}.$$

Conditions (iii) and (iv) guarantee that the optimal serving ratio decreases as the state of the world increases: $\frac{\partial \gamma^*}{\partial t} < 0$. If curtailment occurs in state $\widehat{t}$, it also occurs in all states $t \geq \widehat{t}$. Furthermore, price increases as the state of the world increases:

$$\frac{dv}{dt} = \frac{\partial v}{\partial \gamma}\frac{\partial \gamma^*}{\partial t} + \frac{\partial v}{\partial t} > 0.$$

B Equilibrium capacity when price constraint binds before the capacity constraint

B.1 General demand function

Technical supplement G derives the equilibrium capacity when the capacity constraint binds before the price cap constraint, i.e., $t^N(K) \leq \bar{t}\left(K, \bar{p}^W\right)$ in the relevant range for K. In all the derivations, we also ignore the possibility that curtailment of constant price customers occurs before the capacity constraint, i.e., $\widehat{t}(K) < t^N(K)$. If $v\left(p^R, \gamma^*, t\right)$ does not depend on capacity K, for example if rationing is anticipated and proportional, $\Omega(K)$ may be modified.

We examine both possibilities in this Appendix. Denote $\mu(t)$ the maximum possible price. If we consider a constant price cap, $\mu(t) = \bar{p}^W$. If we consider anticipated and proportional rationing,

$\mu(t) = v(t)$.

Consider N producers, each with installed capacity k^n, ordered such that $k^1 \leq ... \leq k^N$. Lemma 3 presented in Technical supplement G characterizes the equilibrium of the energy markets as follows: there exists N critical states of the world $0 \leq t^1 \leq t^2 \leq ... \leq t^N$ such that, for $t \in [t^j, t^{j+1}]$, all producers $i \leq j$ produce their entire capacity k^i, while all producers $n > j$ produce $\phi^{j+1}(k^1, ..., k^j; t)$ defined on $[t^j, t^{j+1}]$ as the solution of a "modified Cournot condition":

$$\rho\left(\sum_{i=1}^{j} k^i + (N-j)\,\phi^{j+1}; t\right) - c + \phi^{j+1}\rho_q\left(\sum_{i=1}^{j} k^i + (N-j)\,\phi^{j+1}; t\right) = 0.$$

$\widehat{Q}(t) = \sum_{i=1}^{j} k^i + (N-j)\,\phi^{j+1}(t)$ is increasing in t. For $t \geq t^N$, all producers produce their entire capacity k^n, hence $\widehat{Q}(t) = \sum_{n=1}^{N} k^n$.

Since we are ultimately looking for a symmetric equilibrium, suppose the price cap is binding before t^1, i.e., there exists $\widehat{t} \in [0, t^1]$ such that $\rho\left(\widehat{Q}\left(\widehat{t}\right); \widehat{t}\right) = \mu\left(\widehat{t}\right)$.

Assumption 5 *1. If there exists $\widehat{t} \in [0, t^1]$ such that $\rho\left(\widehat{Q}\left(\widehat{t}\right); \widehat{t}\right) = \mu\left(\widehat{t}\right)$, then $\rho\left(\widehat{Q}(t); t\right) > \mu(t)$ for all $t > \widehat{t}$.*

2. $\rho(0, t) > v(t)$ for all $t \geq 0$.

The first part of Assumption 5 greatly simplifies the derivation of the equilibrium, but may not be essential. It holds if $\rho_t(Q; t) - q\rho_{qt}(Q; t) > 0$ for $Q \geq q \geq 0$. The second part of Assumption 5 guarantees the existence of an equilibrium.

Define $\tilde{q}^1(t)$ by $\rho\left(N\tilde{q}^1(t); t\right) = \mu(t)$. $q^n = \tilde{q}^1(t)$ for all $n \geq 1$ is the unique equilibrium on $\left(\widehat{t}, \tilde{t}^1\right]$. To prove the result, first observe that $\tilde{q}^1(t) > \frac{\widehat{Q}(t)}{N}$ since

$$\rho\left(N\tilde{q}^1(t); t\right) = \mu(t) < \rho\left(\widehat{Q}(t); t\right).$$

This is the classical result: in each state of the world, a price cap reduces firms' ability to reduce output. Suppose all firms $n > 1$ produce $\tilde{q}^1(t)$, while firm 1 considers deviating. A negative deviation is unprofitable since it reduces output but cannot increase price, which is capped at $v(t)$. Then, since

$\tilde{q}^1(t) > \frac{\widehat{Q}(t)}{N}$ and the marginal revenue is decreasing,

$$\begin{aligned}\frac{\partial \pi^1}{\partial q^1}\left(\tilde{q}^1(t), ..., \tilde{q}^1(t); t\right) &= \rho\left(N\tilde{q}^1(t); t\right) + \tilde{q}^1(t)\,\rho_q\left(N\tilde{q}^1(t); t\right) - c \\ &< \rho\left(\widehat{Q}(t); t\right) + \frac{\widehat{Q}(t)}{N}\rho_q\left(\widehat{Q}(t); t\right) - c = \frac{\partial \pi^n}{\partial q^n}\left(\frac{\widehat{Q}(t)}{N}, ..., \frac{\widehat{Q}(t)}{N}; t\right) = 0.\end{aligned}$$

A positive deviation is not profitable. $q^n = \tilde{q}^1(t)$ for all $n \geq 1$ is a symmetric equilibrium. Since the profit function is concave, this equilibrium is unique.

When $t = \tilde{t}^1$ characterized by $\tilde{q}^1(t) = k^1 \Longleftrightarrow \rho\left(Nk^1; t^1\right) = \mu\left(t^1\right)$, producer 1 is constrained.

Similarly, for $t \in \left[\tilde{t}^j, \tilde{t}^{j+1}\right]$ for $j = 1, ..., N-1$, the unique symmetric equilibrium for the $(N-j)$ remaining producers is $q^n = (N-j)\,\tilde{q}^{j+1}(t)$ where $\rho\left(\sum_{i=1}^{j} k^i + (N-j)\,\tilde{q}^{j+1}(t); t\right) = \mu(t)$.

Consider now producer n expected profit, given capacities $k^1 \leq ... \leq k^N$, and the structure of the equilibrium described above:

$$\begin{aligned}\Pi^n\left(k^n; \mathbf{k}_{-n}\right) &= \int_0^{\widehat{t}} \phi^1\left(\rho\left(\hat{Q}\right) - c\right) f(t)\,dt + \int_{\widehat{t}}^{\tilde{t}^1} \tilde{q}^1(t)\left(\mu(t) - c\right) f(t)\,dt \\ &\quad + \sum_{i=1}^{n-1} \int_{\tilde{t}^i}^{\tilde{t}^{i+1}} \tilde{q}^{i+1}(t)\left(\mu(t) - c\right) f(t)\,dt + k^n\left(\int_{\tilde{t}^n}^{+\infty}\left(\mu(t) - c\right) f(t)\,dt - r\right).\end{aligned}$$

Thus, for all $n \leq N$,

$$\frac{\partial \Pi^n}{\partial k^n} = \int_{\tilde{t}^n}^{+\infty}\left(\mu(t) - c\right) f(t)\,dt - r$$

since output is continuous, and

$$\frac{\partial^2 \Pi^n}{\left(\partial k^n\right)^2} = -\left(\mu\left(\tilde{t}^n\right) - c\right) f\left(\tilde{t}^n\right)\frac{\partial \tilde{t}^n}{\partial k^n} < 0.$$

If there exists K such that

$$\frac{\partial \Pi^n}{\partial k^n}\left(\frac{K}{N}, ..., \frac{K}{N}\right) = \int_{\tilde{t}(K)}^{+\infty}\left(\mu(t) - c\right) f(t)\,dt - r = 0$$

where $\widetilde{t}(K)$ is such that $\rho\left(K; \widetilde{t}(K)\right) = \mu\left(\widetilde{t}(K)\right)$, then $\frac{K}{N}$ for all $n = 1, ..., N$ is the unique symmetric equilibrium. For any $t \geq 0$, $\lim_{K\to+\infty} \rho(K; t) < c$, thus $\lim_{K\to+\infty} \widetilde{t}(K) = +\infty$, hence $\lim_{K\to+\infty} \frac{\partial \Pi^n}{\partial k^n}\left(\frac{K}{N}, ..., \frac{K}{N}\right) = -r < 0$.

To prove that $\frac{\partial \Pi^n}{\partial k^n}(0,...,0) > 0$, we need to consider separately a price cap and rationing. Consider first a price cap $\bar{p}^W$. Since $\rho(0,t) \geq \bar{p}^W$ for all t≥ 0 by condition (5), $\bar{t}\left(0,\bar{p}^W\right) = 0$ and

$$\frac{\partial \Pi^n}{\partial k^n}(0,...,0) = \int_0^{+\infty} \left(\bar{p}^W - c\right) f(t)\, dt - r = \bar{p}^W - (c+r) > 0$$

by condition (5).

Consider now rationing. The second part of Assumption 5 implies that $\widehat{t}(0) = 0$, thus

$$\frac{\partial \Pi^n}{\partial k^n}(0,...,0) = \int_0^{+\infty} (v(t) - c) f(t)\, dt - r \geq \int_0^{+\infty} (\rho(K^*,t) - c) f(t)\, dt - r > \Psi(K^*;t) - r = 0.$$

Thus, in both cases, $\frac{\partial \Pi^n}{\partial k^n}(0,...,0) > 0$. There exists a unique symmetric equilibrium.

If the price constraint is a price cap $\mu(t) = \bar{p}^W$, $\widetilde{t}(K) = \bar{t}\left(K,\bar{p}^W\right)$, and

$$\Omega\left(K,\bar{p}^W\right) = \int_{\bar{t}(K,\bar{p}^W)}^{+\infty} \left(\bar{p}^W - c\right) f(t)\, dt = \Psi(K,c) - \Psi\left(K,\bar{p}^W\right).$$

If the price constraint is curtailment of constant price customers, $\mu(t) = v(t)$, $\widetilde{t}(K) = \widehat{t}(K)$, and

$$\Omega(K) = \int_{\widehat{t}(K)}^{+\infty} (v(t) - c) f(t)\, dt = \Psi(K,c) - \int_{\bar{t}(K,c)}^{\widehat{t}(K)} (\rho(t) - c) f(t)\, dt.$$

In both cases, the Cournot capacity K^C does not appear to depend on the number of competing firms N. The dependence is in fact embedded in the constraint $\widetilde{t}(K) < t^N(K)$. This point will be illustrated on the specific case of linear demand examined below.

B.2 Linear inverse demand function

Suppose inverse demand is $\rho(Q,t) = \frac{a(t) - bQ - (1-\alpha)p^R}{\alpha}$. The first part of Assumption 5 is met, since $\rho_{qt} = 0$. $t^N(K)$ is defined by

$$a\left(t^N(K)\right) = \frac{N+1}{N} bK + (1-\alpha)\, p^R + \alpha c,$$

while $\bar{t}\left(K, \bar{p}^W\right)$ is defined by

$$a\left(\bar{t}\left(K, \bar{p}^W\right)\right) = bK + (1-\alpha)\, p^R + \alpha \bar{p}^W,$$

and $\widehat{t}(K)$ by

$$a\left(\widehat{t}(K)\right) = \frac{2}{2-\alpha} bK + p^R.$$

Thus,

$$t^N(K) \leq \bar{t}\left(K, \bar{p}^W\right) \Leftrightarrow \frac{bK}{N} \leq \alpha\left(\bar{p}^W - c\right)$$

and

$$t^N(K) \leq \widehat{t}(K) \Leftrightarrow \frac{bK}{N}\left(\frac{1}{N} - \frac{\alpha}{2-\alpha}\right) \leq \alpha\left(p^R - c\right).$$

With the numerical values estimated, and for $N \leq 10$, which is the case of interest, $t^N(K) > \bar{t}\left(K, \bar{p}^W\right)$ and $t^N(K) > \widehat{t}(K)$ for all relevant values of K for $\alpha \leq 0.2$.

Suppose now no price cap is imposed. Since $\widehat{t}(K) < t^N(K)$,

$$\Omega(K) = \int_{\widehat{t}(K)}^{+\infty} (v(t) - c)\, f(t)\, dt.$$

Integration by parts yields

$$\Omega(K) = \left(1 - F\left(\widehat{t}(K)\right)\right)\left(v\left(\widehat{t}(K)\right) - c\right) + \int_{\widehat{t}(K)}^{+\infty} \frac{\partial v}{\partial t}(t)\,(1 - F(t))\, dt.$$

$$\left(1 - F\left(\widehat{t}(K)\right)\right) = \left(\frac{2b}{a_1(2-\alpha)}\left(\bar{K} - K\right)\right)^{\lambda},$$

hence

$$\left(1 - F\left(\widehat{t}(K)\right)\right)\left(v\left(\widehat{t}(K)\right) - c\right) = \left(\frac{2b}{a_1(2-\alpha)}\left(\bar{K} - K\right)\right)^{\lambda}\left(p^R - c + \frac{bK}{2-\alpha}\right).$$

$$\int_{\widehat{t}(K)}^{+\infty} \frac{\partial v}{\partial t}(t)\,(1 - F(t))\, dt = \frac{\alpha}{2-\alpha} I(K) = \frac{a_1}{2(1+\lambda)}\left(\frac{2b}{a_1(2-\alpha)}\left(\bar{K} - K\right)\right)^{1+\lambda},$$

thus

$$\Omega(K) = \left(\frac{2b}{(2-\alpha)\,a_1}\left(\bar{K}-K\right)\right)^{\lambda}\left(p^R - c + \frac{b}{(2-\alpha)}\left(\frac{\bar{K}+\lambda K}{1+\lambda}\right)\right).$$

C Physical capacity certificates

C.1 No short sale condition

Suppose first the SO imposes no condition on certificates sales. Producer n's expected profit, including revenues from the capacity market is: $\Pi^n_{CM}\left(k^n, \phi^n; \mathbf{k}_{-n}, \boldsymbol{\phi}_{-n}\right) = \Pi^n\left(k^n; \mathbf{k}_{-n}\right) + \phi^n H(\Phi)$. Since ϕ^n does not enter $\Pi^n\left(k^n; \mathbf{k}_{-n}\right)$,

$$\frac{\partial \Pi^n_{CM}}{\partial k^n}\left(\frac{K}{N}, ..., \frac{K}{N}\right) = \frac{\partial \Pi^n}{\partial k^n}\left(\frac{K}{N}, ..., \frac{K}{N}\right):$$

the certificate market has no impact on equilibrium investment.

Suppose now the SO imposes $k^n \geq \phi^n$. Consider the case where producers first sell credits, then install capacity. When selecting capacity, each producer maximizes $\bar{\Pi}^n_{CM}\left(k_n; \mathbf{k}_{-n}\right)$ subject to $k^n \geq \phi^n$. The first-order condition is then:

$$\frac{\partial \mathcal{L}^n}{\partial k^n} = \frac{\partial \bar{\Pi}^n}{\partial k^n} + \mu^n_1$$

where μ^n_1 is the shadow cost of the constraint $k^n \geq \phi^n$. Suppose first $\hat{\phi}^n < \hat{k}^n \ \forall n$, then $\mu^n_1 = 0 \ \forall n$ and $\hat{k}^n = \frac{K^C}{N}$ at the symmetric equilibrium. When selecting the amount of credits sold, the producers then maximize $\phi^n H(\Phi)$. Given the shape of $H(.)$, the symmetric equilibrium is $\hat{\phi}^n \geq \frac{K^*}{N}$. But then, $K^C > \Phi \geq K^*$, which is a contradiction, hence $\hat{\phi}^n = \hat{k}^n$.

Consider now the case where producers first build capacity, then sell credits. The proof proceeds along the same lines. At the second stage, producers maximize $\phi^n H(\Phi)$ subject to $\phi^n \leq k^n$. If $\hat{\phi}^n < \hat{k}^n \ \forall n$, the symmetric equilibrium is $\hat{\phi}^n \geq \frac{K^*}{N}$, hence $\hat{\Phi} \geq K^*$. Then, when selecting capacity, producers maximize $\bar{\Pi}^n\left(k^n; \mathbf{k}_{-n}\right)$, hence the equilibrium installed capacity is $\hat{K} = K^C$. But then, $K^C > \hat{\Phi} \geq K^*$, which is a contradiction, hence $\hat{\phi}^n = \hat{k}^n$.

C.2 Equilibrium investment with physical capacity certificates market

Since $k^n = \phi^n$ at the equilibrium, producer n program is:

$$\max_{k^n} \Pi^n_{CM}\left(k^n; \mathbf{k}_{-n},\right) = \Pi^n\left(k^n; \mathbf{k}_{-n}\right) + k^n H\left(K\right)$$

We prove that $\left(\frac{K^*}{N}, ..., \frac{K^*}{N}\right)$ is the unique symmetric equilibrium. Consider first a negative deviation, i.e., $k^1 < \frac{K^*}{N}$ while $k^n = \frac{K^*}{N}$ for all $n > 1$. Since $K = k^1 + \frac{N-1}{N}K^* < K^*$,

$$\frac{\partial \Pi^1_{CM}}{\partial k^1}\left(k^1, \frac{K^*}{N}, ..., \frac{K^*}{N}\right) = \frac{\partial \Pi^1}{\partial k^1}\left(k^1, \frac{K^*}{N}, ..., \frac{K^*}{N}\right) + r.$$

Analysis of the two-stage Cournot game (Zöttl (2001), presented in the technical supplement G yields:

$$\begin{aligned} \frac{\partial \Pi^1}{\partial k^1}\left(k^1, \frac{K^*}{N}, ..., \frac{K^*}{N}\right) \quad = \quad & \int_{t^1}^{t^N(K)} \left(\rho\left(\hat{Q}\left(k^1;t\right)\right) + k^1 \rho_q\left(\hat{Q}\left(k^1;t\right)\right)\frac{\partial \hat{Q}}{\partial k^1} - c\right) f\left(t\right) dt \qquad (15) \\ & + \int_{t^N(K)}^{\bar{t}\left(K, \bar{p}^W\right)} \left(\rho\left(K\right) + k^1 \rho_q\left(K\right) - c\right) f\left(t\right) dt \\ & + \int_{\bar{t}\left(K, \bar{p}^W\right)}^{+\infty} \left(\bar{p}^W - c\right) f\left(t\right) dt - r \end{aligned}$$

where t^1 is the first state of the world where producer 1 is constrained, $\hat{Q}\left(k^1;t\right) = k^1 + (N-1)\,\phi^N\left(k^1;t\right)$ is the aggregate production, and $\phi^N\left(k^1;t\right)$ is the equilibrium production from the remaining $(N-1)$ identical producers, that solves:

$$\rho\left(k^1 + (N-1)\,\phi^N\left(k^1;t\right)\right) + \phi^N\left(k^1;t\right)\rho_q\left(k^1 + (N-1)\,\phi^N\left(k^1;t\right)\right) = c.$$

$\phi^N\left(k^1;t\right) \geq k^1$ for $t \in \left[t^1, t^N\left(K\right)\right]$: lower-capacity producer 1 is constrained, while the $(N-1)$ higher capacity producers are not. Since quantities are strategic substitutes, $\frac{\partial \phi^N}{\partial k^1} < 0$ and:

$$0 < \frac{\partial \hat{Q}}{\partial k^1} = 1 + (N-1)\frac{\partial \phi^N}{\partial k^1} < 1.$$

$\rho\left(\hat{Q}\right) + k^1 \rho_q\left(\hat{Q}\right) - c = \left(k^1 - \phi^N\right)\rho_q\left(\hat{Q}\right)\frac{\partial \hat{Q}}{\partial k^1} \geq 0$ for $t \in \left[t^1, t^N\left(K\right)\right]$. $\rho\left(K; t^N\left(K\right)\right) + k^1 \rho_q\left(K; t^N\left(K\right)\right) =$

c, and $\rho_t(K) + k^1\rho_{qt}(K) \geq 0$, hence $\rho(K) + k^1\rho_q(K) - c \geq 0$ for $t \geq t^N(K)$. Therefore:

$$\frac{\partial \Pi^1}{\partial k^1}\left(k^1, \frac{K^*}{N}, ..., \frac{K^*}{N}\right) + r > 0$$

for $k^1 < \frac{K^*}{N}$, hence no negative deviation is profitable.

Consider now a positive deviation, i.e., $k^N > \frac{K^*}{N}$ while $k^n = \frac{K^*}{N}$ for all $n < N$. Since $K = k^N + \frac{N-1}{N}K^* > K^*$:

$$\frac{\partial \Pi^N_{CM}}{\partial k^N}\left(\frac{K^*}{N}, ..., \frac{K^*}{N}, k^N\right) = \frac{\partial \Pi^N}{\partial k^N}\left(\frac{K^*}{N}, ..., \frac{K^*}{N}, k^N\right) + k^N H'(K) + H(K),$$

and

$$\frac{\partial^2 \Pi^N_{CM}}{(\partial k^N)^2}\left(\frac{K^*}{N}, ..., \frac{K^*}{N}, k^N\right) = \frac{\partial^2 \Pi^N}{(\partial k^N)^2}\left(\frac{K^*}{N}, ..., \frac{K^*}{N}, k^N\right) + k^N H''(K) + 2H'(K).$$

Analysis presented in the technical supplement G shows that, for $k^N > \frac{K}{N}$,

$$\begin{aligned} \frac{\partial^2 \Pi^N}{\partial (k^N)^2}\left(\frac{K^C}{N}, ..., \frac{K^C}{N}, k^N\right) &= \int_{t^N}^{\bar{t}(K,\bar{p}^W)} \left[2\rho_q\left(\hat{K}; t\right) + k^N \rho_{qq}\left(\hat{K}; t\right)\right] f(t)\, dt \\ &\quad + k^N \rho_q\left(\hat{K}; \bar{t}\left(K, \bar{p}^W\right)\right) f\left(\bar{t}\left(K, \bar{p}^W\right)\right) \frac{\partial \bar{t}\left(K, \bar{p}^W\right)}{\partial k^N} \\ &< 0. \end{aligned} \quad (16)$$

Thus,

$$\frac{\partial \Pi^N_{CM}}{\partial k^N}\left(\frac{K^*}{N}, ..., \frac{K^*}{N}, k^N\right) < \frac{\partial \Pi^N}{\partial k^N}\left(\frac{K^*}{N}, ..., \frac{K^*}{N}\right) + \frac{K^*}{N}H'(K^*) + r < 0$$

since condition (8) implies $\frac{K^*}{N}H'(K^*) + r < 0$.

Hence, $\left(\frac{K^*}{N}, ..., \frac{K^*}{N}\right)$ is a symmetric equilibrium. Finally, no other symmetric equilibrium exists since $\Pi^n\left(\frac{K}{N}, ..., \frac{K}{N}\right) + \frac{K}{N}H(K)$ is concave.

D Financial reliability options

D.1 Investment decision before options market

The equilibrium is solved by backwards induction. In the second stage, producers solve the equilibrium of the option market, taking $(k^n, \mathbf{k}^{-n})$ as given.

We assume that including the option market does not decrease investment, i.e., $K \geq K^C\left(\bar{p}^S\right)$. As suggested by Cramton and Ockenfels, the SO imposes the restriction that all capacity is sold forward: $\theta^n \geq k^n$. This restriction is made operational by conditioning profits from the option market to $\theta^n \geq k^n$. Since these profits are positive, $\theta^n \geq k^n$ is a dominant strategy, hence holds.

Producers' expected profit is:

$$\Pi_{RO}^n\left(k^n, \theta^n; \mathbf{k}_{-n}, \boldsymbol{\theta}_{-n}\right) = \Pi^n\left(k^n; \mathbf{k}_{-n},\right) + \theta^n H_{RO}\left(\Theta\right) + \left(k^n - \frac{K\theta^n}{\Theta}\right)\Psi\left(K, \bar{p}^S\right)$$

D.1.1 Equilibrium in the options market

$$\frac{\partial \Pi_{RO}^n}{\partial \theta^n}\left(\theta^n; \boldsymbol{\theta}_{-n}\right) = H_{RO}\left(\Theta\right) + \theta^n H_{RO}^{'}\left(\Theta\right) - \frac{\Theta - \theta^n}{\Theta^2} K\Psi\left(K, \bar{p}^S\right).$$

We prove that $\theta^n = \frac{K^*}{N} \geq k^n$ for all n is a symmetric equilibrium. Consider first a negative deviation, i.e., $\theta^1 < \frac{K^*}{N}$ while $\theta^n = \frac{K^*}{N}$ for all $n > 1$. Since $\theta^1 \geq k^1$, $\Theta = \theta^1 + \frac{N-1}{N}K^* < K^*$,

$$\frac{\partial \Pi_{RO}^1}{\partial \theta^1}\left(\theta^1, \frac{K^*}{N}, ..., \frac{K^*}{N}\right) = r - \frac{\frac{N-1}{N}K^*}{\Theta^2} K\Psi\left(K, \bar{p}^S\right).$$

Since $\theta^n \geq k^n$ for all n, $\Theta \geq K$,

$$\frac{\frac{N-1}{N}K^*}{\Theta^2} K\Psi\left(K, \bar{p}^S\right) \leq \frac{\frac{N-1}{N}K^*}{\theta^1 + \frac{N-1}{N}K^*}\Psi\left(K, \bar{p}^S\right) < \Psi\left(K, \bar{p}^S\right),$$

hence

$$\frac{\partial \Pi_{RO}^1}{\partial \theta^1}\left(\theta^1, \frac{K^*}{N}, ..., \frac{K^*}{N}\right) > r - \Psi\left(K, \bar{p}^S\right) \geq r - \Psi\left(K^C\left(\bar{p}^S\right), \bar{p}^S\right) \geq 0$$

by condition (9). No negative deviation is profitable.

Consider now a positive deviation, i.e., $\theta^N > \frac{K^*}{N} \geq k^N$ while $\theta^n = \frac{K^*}{N}$ for all $n < N$.

$$\frac{\partial \Pi_{RO}^N}{\partial \theta^N}\left(\frac{K^*}{N},...,\frac{K^*}{N},\theta^N\right) = H_{RO}(\Theta) + \theta^N H_{RO}'(\Theta) - \frac{\Theta - \theta^N}{\Theta^2} K\Psi\left(K, \bar{p}^S\right).$$

$\Theta = k^N + \frac{N-1}{N}K^* > K^*$, therefore $H_{RO}(\Theta) + \theta^N H_{RO}'(\Theta) < H_{RO}(K^*) + \frac{K^*}{N} H_{RO}'(K^*) < 0$ by condition (8), hence $\frac{\partial \Pi_{RO}^N}{\partial \theta^N}\left(\frac{K^*}{N},...,\frac{K^*}{N},\theta^N\right) < 0$ for all $\theta^N > \frac{K^*}{N}$. No positive deviation is profitable.

$\theta^n = \frac{K^*}{N}$ for all n is therefore an equilibrium.

We now prove $\theta^n = \frac{K^*}{N} \geq k^n$ for all n is the unique symmetric equilibrium. Consider first $\theta^n = \frac{\Theta}{N} < \frac{K^*}{N}$ for all n:

$$\frac{\partial \Pi_{RO}^n}{\partial \theta^n}\left(\frac{\Theta}{N},...,\frac{\Theta}{N}\right) = r - \frac{N-1}{N}\frac{K}{\Theta}\Psi\left(K,\bar{p}^S\right) > r - \Psi\left(K,\bar{p}^S\right) > 0.$$

No symmetric equilibrium exists with $k^n \leq \frac{\Theta}{N} < \frac{K^*}{N}$ for all n.

Finally, consider the case $\theta^n = \frac{\Theta}{N} > \frac{K^*}{N}$ for all n:

$$\frac{\partial \Pi_{RO}^n}{\partial \theta^n}\left(\frac{\Theta}{N},...,\frac{\Theta}{N}\right) = H_{RO}(\Theta) + \frac{\Theta}{N} H_{RO}'(\Theta) - \frac{N-1}{N}\frac{K}{\Theta}K\Psi\left(K,\bar{p}^S\right) < 0.$$

There exists no symmetric equilibrium with $\frac{\Theta}{N} > \frac{K^*}{N}$.

D.1.2 Equilibrium investment

In the first stage, producers decide on capacity, taking into account the equilibrium of the options market. Denote $V^n(k^n;\mathbf{k}_{-n})$ producer n profit function:

$$V^n(k^n;\mathbf{k}_{-n}) = \Pi_{RO}^n\left(k^n, \frac{K^*}{N}; \mathbf{k}_{-n}, \frac{\mathbf{K}^*}{\mathbf{N}}\right) = \Pi^n(k^n;\mathbf{k}_{-n},) + \frac{K^*}{N}r + \left(k^n - \frac{K}{N}\right)\Psi\left(K,\bar{p}^S\right)$$

$$\frac{\partial V^n}{\partial k^n} = \frac{\partial \Pi^n}{\partial k^n} + \frac{N-1}{N}\Psi\left(K,\bar{p}^S\right) + \left(k^n - \frac{K}{N}\right)\frac{\partial \Psi}{\partial K}.$$

A necessary condition for a symmetric equilibrium $k^n = \frac{K}{N}$ is:

$$\frac{\partial V^n}{\partial k^n}\left(\frac{K}{N},...,\frac{K}{N}\right) = \frac{\partial \Pi^n}{\partial k^n}\left(\frac{K}{N},...,\frac{K}{N}\right) + \frac{N-1}{N}\Psi\left(K,\bar{p}^S\right)$$

$\frac{\partial V^n}{\partial k^n}\left(\frac{K}{N},...,\frac{K}{N}\right)$ is decreasing since $\frac{\partial \Pi^n}{\partial k^n}\left(\frac{K}{N},...,\frac{K}{N}\right)$ is decreasing and $\frac{\partial \Psi}{\partial K} < 0$. $\frac{\partial V^n}{\partial k^n}(0,...,0) = \frac{\partial \Pi^n}{\partial k^n}(0,...,0) + \frac{N-1}{N}\Psi\left(0,\bar{p}^S\right) > 0$ since (i) $\frac{\partial \Pi^n}{\partial k^n}(0,...,0) > 0$ and (ii) $\Psi\left(0,\bar{p}^S\right) > 0$ by construction. $\lim_{K\to+\infty}\frac{\partial V^n}{\partial k^n}(K) = -r < 0$. Hence, there exists a unique $K_{RO}^C > 0$ such that $\frac{\partial \Pi_{RO}^n}{\partial k^n}\left(\frac{K_{RO}^C}{N},...,\frac{K_{RO}^C}{N}\right) = 0$. This is equation (11).

We now prove that $k^n = \frac{K_{RO}^C}{N}$ for all n is an equilibrium. Consider first a negative deviation: $k^1 < \frac{K_{RO}^C}{N}$ while $k^n = \frac{K_{RO}^C}{N}$ for all $n > 1$. Total installed capacity is $K = k^1 + \frac{N-1}{N}K_{RO}^C < K_{RO}^C$. Substituting expression (15) for $\frac{\partial \Pi^n}{\partial k^n}\left(k^1,\frac{K_{RO}^C}{N},...,\frac{K_{RO}^C}{N}\right)$,

$$\begin{aligned}
\frac{\partial V^1}{\partial k^1}\left(k^1,\frac{K_{RO}^C}{N},...,\frac{K_{RO}^C}{N}\right) &= \int_{t^1}^{t^N(K)}\left(\rho\left(\hat{Q}\left(k^1;t\right)\right) + k^1\rho_q\left(\hat{Q}\left(k^1;t\right)\right)\frac{\partial \hat{Q}}{\partial k^1} - c\right)f(t)\,dt \\
&+ \int_{t^N(K)}^{\bar{t}\left(K,\bar{p}^S\right)}\left(\rho(K) + k^1\rho_q(K) - c\right)f(t)\,dt \\
&+ \int_{\bar{t}\left(K,\bar{p}^S\right)}^{+\infty}\left(\left(\bar{p}^S - c\right) + \frac{N-1}{N}\left(\rho(K,t) - \bar{p}^S\right) + \left(k^1 - \frac{K}{N}\right)\rho_q(K,t)\right)f(t)\,dt - r.
\end{aligned}$$

Substituting in equation (11), observing that $t^N(K) < t^N\left(K_{RO}^C\right)$ and $\bar{t}\left(K,\bar{p}^S\right) < \bar{t}\left(K_{RO}^C,\bar{p}^S\right)$ since $K < K_{RO}^C$, and rearranging yields

$$\begin{aligned}
\frac{\partial V^1}{\partial k^1}\left(k^1,\frac{K_{RO}^C}{N},...,\frac{K_{RO}^C}{N}\right) &= \int_{t^1}^{t^N(K)}\left(\rho\left(\hat{Q}\right) + k_q^1\rho\left(\hat{Q}\right)\frac{\partial \hat{Q}}{\partial k^1} - c\right)f(t)\,dt \\
&+ \int_{t^N(K)}^{t^N\left(K_{RO}^C\right)}\left(\rho(K) + k_q^1\rho(K) - c\right)f(t)\,dt \\
&+ \int_{t^N\left(K_{RO}^C\right)}^{\bar{t}\left(K,\bar{p}^S\right)}\left(\rho(K) + k_q^1\rho(K) - \left(\rho\left(K_{RO}^C\right) + \frac{K_{RO}^C}{N}\rho_q\left(K_{RO}^C\right)\right)\right)f(t)\,dt \\
&+ \int_{\bar{t}\left(K,\bar{p}^S\right)}^{\bar{t}\left(K_{RO}^C,\bar{p}^S\right)}\begin{pmatrix} \bar{p}^S - \rho\left(K_{RO}^C;t\right) - \frac{K_{RO}^C}{N}\rho_q\left(K_{RO}^C\right) \\ +\frac{N-1}{N}\left(\rho(K;t) - \bar{p}^S + \rho_q(K;t)\left(k^1 - \frac{K_{RO}^C}{N}\right)\right)\end{pmatrix}f(t)\,dt \\
&+ \frac{N-1}{N}\int_{\bar{t}\left(K_{RO}^C,\bar{p}^S\right)}^{+\infty}\left(\rho(K;t) - \rho\left(K_{RO}^C;t\right) + \rho_q(K;t)\left(k^1 - \frac{K_{RO}^C}{N}\right)\right)f(t)\,dt.
\end{aligned}$$

Each term is positive:

1. $\rho\left(\hat{Q}\right)+k_q^1\rho\left(\hat{Q}\right)\frac{\partial\hat{Q}}{\partial k^1}-c=\left(k^1-\phi^N\right)\rho_q\left(\hat{Q}\right)\frac{\partial\hat{Q}}{\partial k^1}\geq 0$ for $t\in\left[t^1,t^N\left(K\right)\right]$

2. $\rho\left(K;t^N\left(K\right)\right)+k_q^1\rho\left(K;t^N\left(K\right)\right)=c$, and $\rho_t\left(K\right)+k^1\rho_{qt}\left(K\right)\geq 0$, hence $\rho\left(K\right)+k_q^1\rho\left(K\right)-c\geq 0$ for $t\in\left[t^N\left(K\right),t^N\left(K_{RO}^C\right)\right]$

3. $\rho_q\left(Q\right)+q\rho_{qq}\left(Q\right)<0$, hence $\rho\left(K\right)+k^1\rho_q\left(K\right)\geq\rho\left(K\right)+\frac{K_{RO}^C}{N}\rho_q\left(K\right)\geq\rho\left(K_{RO}^C\right)+\frac{K_{RO}^C}{N}\rho_q\left(K_{RO}^C\right)$ for $t\in\left[t^N\left(K_{RO}^C\right),\bar{t}\left(K,\bar{p}^S\right)\right]$

4. $\rho\left(K_{RO}^C;t\right)\leq\bar{p}^S$ for $t\leq\bar{t}\left(K_{RO}^C,\bar{p}^S\right)$ and $\rho\left(K;t\right)\geq\bar{p}^S$ for $t\geq\bar{t}\left(K,\bar{p}^S\right)$, hence

$$\left(\bar{p}^S-\rho\left(K_{RO}^C;t\right)-\frac{K_{RO}^C}{N}\rho_q\left(K_{RO}^C\right)+\frac{N-1}{N}\left(\rho\left(K;t\right)\geq\bar{p}^S\right)\right)\geq 0$$

for $t\in\left[t^{\bar{p}^S}\left(K\right),t^{\bar{p}^S}\left(K_{RO}^C\right)\right]$

5. $K\leq K_{RO}^C$, yields $\rho\left(K;t\right)\geq\rho\left(K_{RO}^C;t\right)$ for all t

Thus, $\frac{\partial\Pi_{RO}^1}{\partial k^1}\left(k^1,\frac{K_{RO}^C}{N},...,\frac{K_{RO}^C}{N}\right)>0$: a negative deviation is not profitable.

Consider now a positive deviation, $k^N>\frac{K_{RO}^C}{N}$ while $k^n=\frac{K_{RO}^C}{N}$ for all $n<N$. $K=k^N+\frac{N-1}{N}K_{RO}^C>K_{RO}^C$.

$$\begin{aligned}\frac{\partial^2V^N}{\left(\partial k^N\right)^2}\left(\frac{K_{RO}^C}{N},...,\frac{K_{RO}^C}{N},k^N\right) &= \frac{\partial^2\Pi^N}{\left(\partial k^N\right)^2}++2\frac{N-1}{N}\frac{\partial\Psi}{\partial K}+\left(k^N-\frac{K}{N}\right)\frac{\partial^2\Psi}{\left(\partial K\right)^2}\\ &= \frac{\partial^2\Pi^N}{\left(\partial k^N\right)^2}+\frac{N-1}{N}\int_{\bar{t}\left(K,\bar{p}^S\right)}^{+\infty}\left[2\rho_q\left(K;t\right)+\left(k^N-\frac{K_{RO}^C}{N}\right)\rho_{qq}\left(K;t\right)\right]f\left(t\right)dt\\ &\quad -\left(k^N-\frac{K}{N}\right)\rho_q\left(K;\bar{t}\left(K,\bar{p}^S\right)\right)f\left(\bar{t}\left(K,\bar{p}^S\right)\right)\frac{\partial\bar{t}\left(K,\bar{p}^S\right)}{\partial K}.\end{aligned}$$

Substituting in $\frac{\partial^2\Pi^N}{\left(\partial k^N\right)^2}$ from equation (16),

$$\begin{aligned}\frac{\partial^2V^N}{\left(\partial k^N\right)^2}\left(\frac{K_{RO}^C}{N},...,\frac{K_{RO}^C}{N},k^N\right) &= \int_{t^N}^{\bar{t}\left(K,\bar{p}^S\right)}\left[2\rho_q\left(\hat{K};t\right)+k^N\rho_{qq}\left(\hat{K};t\right)\right]f\left(t\right)dt\\ &\quad +\frac{N-1}{N}\int_{\bar{t}\left(K,\bar{p}^S\right)}^{+\infty}\left[2\rho_q\left(K;t\right)+\left(k^N-\frac{K_{RO}^C}{N}\right)\rho_{qq}\left(K;t\right)\right]f\left(t\right)dt\\ &\quad +\frac{K}{N}\rho_q\left(K;\bar{t}\left(K,\bar{p}^S\right)\right)f\left(\bar{t}\left(K,\bar{p}^S\right)\right)\frac{\partial\bar{t}\left(K,\bar{p}^S\right)}{\partial K}\\ &< 0.\end{aligned}$$

A positive deviation is not profitable.

Therefore $\left(\frac{K_{RO}^C}{N},...,\frac{K_{RO}^C}{N}\right)$ constitute an equilibrium. Furthermore,

$$\begin{aligned}\frac{\partial^2 V^n}{\partial (k^n)^2}\left(\frac{K}{N},...,\frac{K}{N}\right) &= \int_{t^N}^{t^{\bar{p}^S}}\left[2\rho_q(K;t)+\frac{K}{N}\rho_{qq}(K;t)\right]f(t)\,dt+2\frac{N-1}{N}\int_{t^{\bar{p}^S}}^{+\infty}\rho_q(K;t)f(t)\,dt \\ &\quad +\frac{K}{N}\rho_q\left(K;\bar{t}\left(K,\bar{p}^S\right)\right)f\left(\bar{t}\left(K,\bar{p}^S\right)\right)\frac{\partial\bar{t}\left(K,\bar{p}^S\right)}{\partial K} \\ &< 0\end{aligned}$$

hence $\left(\frac{K_{RO}^C}{N},...,\frac{K_{RO}^C}{N}\right)$ is the unique symmetric equilibrium.

Finally, we prove in the main text that $K^C\left(\bar{p}^S\right)\leq K_{RO}^C<K^*$.

D.2 Options market before investment decision

D.2.1 Equilibrium investment

In the second stage, producers decide on capacity, taking θ^n as given. Producer's n profit is:

$$\Pi_{RO}^n\left(k^n,\theta^n;\mathbf{k}_{-n},\boldsymbol{\theta}_{-n}\right)=\Pi^n\left(k^n;\mathbf{k}_{-n},\right)+\theta^n H_{RO}(\Theta)+\left(k^n-\frac{K\theta^n}{\Theta}\right)\Psi\left(K,\bar{p}^S\right)$$

Then:

$$\frac{\partial\Pi_{RO}^n}{\partial k^n}=\frac{\partial\Pi^n}{\partial k^n}+\left(1-\frac{\theta^n}{\Theta}\right)\Psi\left(K,\bar{p}^S\right)+\left(k^n-\frac{K}{N}\right)\frac{\partial\Psi}{\partial K}$$

Suppose $\theta^n=\frac{\Theta}{N}$ for all n. $k^n=\frac{K}{N}$ for all n yields the following necessary condition:

$$\frac{\partial\Pi^n}{\partial k^n}\left(\frac{K}{N},...,\frac{K}{N}\right)+\frac{N-1}{N}\Psi\left(K,\bar{p}^S\right)=0$$

which is equation (11). $k^n=\frac{K_{RO}^C}{N}$ for all n is an equilibrium candidate.

Using the same argument as in the previous section, $k^n=\frac{K_{RO}^C}{N}$ for all n is the unique symmetric equilibrium of the investment game, assuming that $\theta^n=\frac{\Theta}{N}$ for all n.

D.2.2 Equilibrium in the options market

In the first-stage, producers solve the equilibrium of the option market, replacing θ^n by $\theta^n(k^n, \mathbf{k}^{-n})$. Suppose a symmetric equilibrium $\theta^n = \frac{\Theta}{N}$ exists. Then, $k^n = \frac{K^C_{RO}}{N}$ for all n and producer n expected profit is

$$V^n(\theta^n; \boldsymbol{\theta}_{-n}) = \Pi^n\left(\frac{K^C_{RO}}{N}, ..., \frac{K^C_{RO}}{N}\right) + \theta^n H_{RO}(\Theta) + \left(\frac{1}{N} - \frac{\theta^n}{\Theta}\right)\Psi\left(K^C_{RO}, \bar{p}^S\right)$$

Hence:

$$\frac{\partial V^n}{\partial \theta^n} = H_{RO}(\Theta) + \theta^n H^{'}_{RO}(\Theta) - \frac{\Theta - \theta^n}{\Theta^2}\Psi\left(K^C_{RO}, \bar{p}^S\right)$$

since K^C_{RO} is independent of θ^n by equation (). The previous analysis then applies: $\Theta = K^*$ is the only symmetric equilibrium.

Proving existence of a symmetric equilibrium in the general case is beyond the scope of this work. I provide below a heuristic argument, in the case $N = 2$. Consider first $\theta^2 < \theta^1 = \frac{K^*}{2}$. Suppose there exists a unique interior equilibrium of the investment game, $\left(\bar{k}^1(\theta_1, \theta_2), \bar{k}^2(\theta_1, \theta_2)\right)$, defined by:

$$\begin{cases} \frac{\partial \Pi^1_{RO}}{\partial k^1} = \frac{\partial \Pi^1}{\partial k^1} + \frac{\theta^2}{(\theta^1+\theta^2)}\Psi + \frac{\bar{k}^1\theta^2 - \bar{k}^2\theta^1}{\Theta}\frac{\partial \Psi}{\partial K} = 0 \\ \frac{\partial \Pi^2_{RO}}{\partial k^2} = \frac{\partial \Pi^2}{\partial k^2} + \frac{\theta^1}{(\theta^2+\theta^1)}\Psi + \frac{\bar{k}^2\theta^1 - \bar{k}^1\theta^2}{\Theta}\frac{\partial \Psi}{\partial K} = 0 \end{cases}.$$

Then,

$$\frac{\partial V^1}{\partial \theta^1} = \frac{\partial \Pi^1_{RO}}{\partial k^1}\frac{\partial \bar{k}^1}{\partial \theta_1} + \frac{\partial \Pi^1_{RO}}{\partial k^2}\frac{\partial \bar{k}^2}{\partial \theta_1} + \frac{\partial \Pi^1_{RO}}{\partial \theta_1}.$$

$\frac{\partial \Pi^1_{RO}}{\partial k^1}\left(\bar{k}^1, \bar{k}^2\right) = 0$ since $\left(\bar{k}^1, \bar{k}^2\right)$ is the equilibrium, and

$$\frac{\partial \Pi^1_{RO}}{\partial k^2} = \left(\frac{\partial \Pi^1}{\partial k^2} - \frac{\theta^1}{(\theta^1+\theta^2)}\Psi\left(\bar{K}, \bar{p}^S\right) + \frac{\bar{k}^1\theta^2 - \bar{k}^2\theta^1}{(\theta^1+\theta^2)}\frac{\partial \Psi}{\partial K}\right)$$

$\frac{\partial \Pi^1}{\partial k^2} < 0$ since capacities are strategic complements, $\Psi\left(\bar{K}, \bar{p}^S\right) > 0$, thus, if $\frac{\bar{k}^1\theta^2 - \bar{k}^2\theta^1}{(\theta^1+\theta^2)}\frac{\partial \Psi}{\partial K}$ is small enough, $\frac{\partial \Pi^1_{RO}}{\partial k^2} < 0$: capacities remain strategic complements including the options market. I provide in the technical supplement H sufficient conditions for $\frac{\partial \bar{k}^2}{\partial \theta_1} < 0$. A heuristic argument is that increasing θ_1 increases firm 1 exposure to spot prices, hence increases $\bar{k}^1$, hence reduces $\bar{k}^2$ since capacities are strategic complements. Thus, we expect $\frac{\partial \Pi^1_{RO}}{\partial k^2}\frac{\partial \bar{k}^2}{\partial \theta_1} > 0$. Since $\frac{\partial \Pi^1_{RO}}{\partial \theta_1} > 0$, $\frac{\partial V^1}{\partial \theta^1} > 0$, a negative deviation is not profitable.

Consider now a positive deviation, $\theta^2 > \theta^1 = \frac{K^*}{2}$. Then,

$$\frac{\partial V^2}{\partial \theta^2} = \frac{\partial \Pi^2_{RO}}{\partial k^1}\frac{\partial \bar{k}^1}{\partial \theta^2} + H_{RO}\left(\theta^1+\theta^2\right) + \theta^2 H^{'}_{RO}\left(\theta^1+\theta^2\right) - \frac{\theta^1}{\left(\theta^1+\theta^2\right)^2} K\Psi\left(K,\bar{p}^S\right)$$

$\frac{\partial \Pi^2_{RO}}{\partial k^1}\frac{\partial \bar{k}^1}{\partial \theta^2} > 0$, which goes in the "wrong" direction. However, since $\left(-\frac{\theta^1}{\left(\theta^1+\theta^2\right)^2} K\Psi\left(K,\bar{p}^S\right)\right) > 0$ and $H_{RO} + \theta^2 H^{'}_{RO} > 0$, we should still have $\frac{\partial V^2}{\partial \theta^2} < 0$.

D.3 Simultaneous option market and investment decision

If the decisions are simultaneous, the previous first-order conditions determine a symmetric equilibrium candidate. The previous arguments apply, and it constitutes the unique symmetric equilibrium.

E Average $VoLL$ pricing

Suppose a fraction of constant price customers must be curtailed for capacity K. $\hat{t}(K) < +\infty$, and the average $VoLL$ when constant price customers are curtailed is:

$$\overline{v}(K) = \frac{\int_{\hat{t}(K)}^{+\infty} \rho(K,t) f(t)\,dt}{\int_{\hat{t}(K)}^{+\infty} f(t)\,dt}.$$

The rule proposed by Stoft (2002) leads to the following inverse demand function for producers: $\rho(K;t)$ for $t \leq \widehat{t}$ and a horizontal inverse demand at $\overline{v}$ for $t > \widehat{t}$. This creates a discontinuity in producers' profit at $t = \hat{t}(K)$ since

$$\rho\left(K;\hat{t}(K)\right) < \overline{v}(K) = \lim_{t \to \hat{t}(K)^+} \overline{v}(K).$$

Thus,

$$\begin{aligned}\Omega(K) &= \Psi(K,c) + \left(\rho\left(K;\hat{t}\right) - \overline{v}(K)\right) f\left(\hat{t}\right)\frac{\partial \hat{t}}{\partial K} - \int_{\hat{t}(K)}^{+\infty} \left(\rho(K,t) - \overline{v}(K)\right) f(t)\,dt \\ &= \Psi(K,c) + \left(\rho\left(K;\hat{t}\right) - \overline{v}(K)\right) f\left(\hat{t}\right)\frac{\partial \hat{t}}{\partial K}\end{aligned}$$

since $\int_{\hat{t}(K)}^{+\infty} (\rho(K,t) - \overline{v}(K)) f(t)\, dt = 0$ by definition of $\overline{v}(K)$. Therefore:

$$\Omega(K^*) = r - \left(\rho\left(K^*; \hat{t}\right) - \overline{v}(K^*)\right) f\left(\hat{t}\right) \frac{\partial \hat{t}}{\partial K} < r$$

since $\frac{\partial \hat{t}}{\partial K} > 0$.

F Energy cum operating reserves market

Define the total surplus

$$\mathring{\mathcal{S}}(p,\gamma;t) = \alpha S(p(t);t) + (1-\alpha)\mathcal{S}(p,\gamma;t)$$

and total demand

$$\mathring{\mathcal{D}}(p,\gamma;t) = \alpha S(p(t);t) + (1-\alpha)\mathcal{D}(p,\gamma;t).$$

The social planner's program is:

$$\begin{array}{cc} \max\limits_{\{p(t),\gamma(t)\},K} & \mathbb{E}\left\{\mathring{\mathcal{S}}(p(t),\gamma(t);t) - c\mathring{\mathcal{D}}(p(t),\gamma(t);t)\right\} - rK \\ st: & (1+h(t))\,\mathring{\mathcal{D}}(p(t),\gamma(t);t) \leq K \qquad (\lambda(t)) \end{array}$$

The associated Lagrangian is:

$$\mathcal{L} = \mathbb{E}\left\{\mathring{\mathcal{S}}(p(t),\gamma(t);t) - c\mathring{\mathcal{D}}(p(t),\gamma(t);t) + \lambda(t)\left[K - (1+h(t))\,\mathring{\mathcal{D}}(p(t),\gamma(t);t)\right]\right\} - rK$$

and:

$$\begin{cases} \frac{\partial \mathcal{L}}{\partial p(t)} = \{p(t) - [c + (1+h(t))\lambda(t)]\} \frac{\partial \mathring{\mathcal{D}}}{\partial p(t)} \\ \frac{\partial \mathcal{L}}{\partial \gamma(t)} = \left\{v_t\left[\mathring{\mathcal{D}}(p(t),\gamma(t)),\gamma(t)\right] - [c + (1+h(t))\lambda(t)]\right\} \frac{\partial \mathring{\mathcal{D}}}{\partial \gamma(t)} \\ \frac{\partial \mathcal{L}}{\partial K} = \mathbb{E}[\lambda(t)] - r \end{cases}$$

First, off-peak $\lambda(t) = 0$ and $\gamma(t) = 1$. Then $p(t) = c = w(t)$. This holds as long as $\rho\left(\frac{Q}{1+h(t)}, t\right) = c$ $for\ Q \leq K \Leftrightarrow t \leq \bar{t}_{OR}(K,c)$.

Second, on-peak, if constant price customers are not curtailed, $(1+h(t))\,\mathring{\mathcal{D}}(p(t);1,t) = K$ hence $\lambda(t) > 0$ and $\gamma(t) = 1$. Then $p(t) = c + \lambda(t)(1+h(t)) = \rho\left(\frac{K}{1+h(t)};t\right)$ and $\lambda(t) = w(t) - c = \frac{p(t)-c}{1+h(t)} > 0$.

Finally, constant price customers may have to be curtailed, $(1+h(t))\mathring{\mathcal{D}}(p(t),\gamma^*(t);t)=K$ for $\gamma^*(t)<1$ such that $(1+h(t))\mathring{\mathcal{D}}(\bar{v},\gamma^*(t);t)=K$. Then $(1+h(t))\lambda(t)=\rho\left(\frac{K}{1+h(t)};t\right)-c$ as before.

The optimal capacity K^*_{OR} is then defined by $\mathbb{E}[\lambda(t)]=r$ which yields equation (13).

As discussed in Section 7, the producers' problem is isomorphic to the previous Sections, except that $\frac{s^n(t)}{1+h(t)}$ replaces production $q^n(t)$. The proof of equation (14) then follows the steps of Lemmata 3 to 6 presented in the technical report G.

Technical Report

Not part of the paper

Available upon request

G Derivation of the Cournot capacity

For the reader's convenience, the derivation of the Cournot equilibrium capacity is presented here. The proof follows Zöttl (2011). The main difference is the introduction of rationing.

G.1 Capacity constrained Cournot equilibrium without a price cap

Producers are ordered by increasing capacity: $k^1 \leq k^2 ... \leq k^N$.

The unconstrained Cournot aggregate output $Q^C(t)$ in state t is defined by $\rho(Q;t) - c + \frac{Q}{N}\rho_q(Q;t) = 0$. Under Assumption 4, the implicit function theorem yields:

$$\frac{dQ^C}{dt} = -\left(\rho_t + \frac{Q}{N}\rho_{qt}\right)\left(\rho_q + \frac{\rho_q + Q\rho_{qq}}{N}\right)^{-1} > 0$$

Lemma 3 *Define $t^0 = 0$. For a given vector $\mathbf{k} \in \mathbb{R}^N$ of generation capacities, there exists N critical states of the world $0 \leq t^1 \leq t^2 \leq ... \leq t^N$ such that $\hat{q}^n(\mathbf{k};t)$, the equilibrium output for producer n, is characterized by*

$$\hat{q}^n(\mathbf{k};t) = \begin{cases} \phi^{j+1}(k^1, ..., k^j; t) & if \quad t \in [t^j, t^{j+1}] \ \ for\ j < n \\ k^n & if \quad t \geq t^n \end{cases}$$

where $\phi^{j+1}(k^1, ..., k^j; t)$ defined on $[t^j, t^{j+1}]$ is the solution of a "modified Cournot condition":

$$\rho\left(\sum_{i=1}^{j} k^i + (N-j)\phi^{j+1}; t\right) - c + \phi^{j+1}\rho_q\left(\sum_{i=1}^{j} k^i + (N-j)\phi^{j+1}; t\right) = 0$$

$\forall n \leq N$, $\hat{q}^n(\mathbf{k};t)$ is continuous in all its arguments and increasing in t.

Producers expected profit is:

$$\Pi^n(k^n; \mathbf{k}_{-n}) = \sum_{j=0}^{n-1} \int_{t^j}^{t^{j+1}} \phi^{j+1}\left(\rho\left(\hat{Q}\right) - c\right) f(t)\,dt + k^n \sum_{j=n}^{N}\left[\int_{t^j}^{t^{j+1}} \left(\rho\left(\hat{Q}\right) - c\right) f(t)\,dt - r\right]$$

where $\hat{Q}(\mathbf{k};t) = \sum_{n=1}^{N} \hat{q}^n(\mathbf{k};t)$ is the aggregate output.

Proof. *Construction of the equilibrium proceeds by induction on $n \leq N$. As seen previously, $Q^C(t)$ is increasing in t. Denote t^1 the first state such that $Q^C(t^1) = k^1$. Suppose $t^1 \to +\infty$, then $\forall t \geq 0$,*

$Q^C(t) < k^1 \leq k^2 \leq ... \leq k^N$. Then:

$$\Pi^1\left(k^1;\mathbf{k}_{-1}\right) = \mathbb{E}\left[Q^C(t)\left(\rho\left(Q^C(t);t\right)-c\right)\right] - rk^1$$

and $\frac{\partial\Pi^1}{\partial k^1} = -r < 0$, hence $\Pi^1\left(k^1\right) < 0$ since $\Pi^1(0) = 0$. This contradicts producer 1 individual rationality, hence t^1 exists by contradiction. Denote $\hat{Q}^1(t) = Q^C(t)$ and $\phi^1(t) = \frac{Q^C(t)}{N}$.

Suppose now we have characterized the equilibrium up until t^n, defined by $\phi^n\left(k^1,...,k^{n-1};t^n\right) = k^n$. We now search for an equilibrium for $t \geq t^n$. Suppose all generators $i = 1,..,n$ produce up to their capacity k^i. The profit of any generator $j > n$ is:

$$\pi^j\left(q^j;\mathbf{q}_{-j};t\right) = q^j\left(\rho\left(\sum_{i=1}^{n}k^i + q^j + \sum_{\substack{i=n+1\\ i\neq j}}^{N} q^i;t\right) - c\right)$$

As long as producer $(n+1)$ is not constrained, we have:

$$\frac{\partial\pi^j}{\partial q^j} = \rho\left(\sum_{i=1}^{n}k^i + q^j + \sum_{\substack{i=n+1\\ i\neq j}}^{N} q^i;t\right) - c + q^j\rho_q\left(\sum_{i=1}^{n}k^i + q^j + \sum_{\substack{i=n+1\\ i\neq j}}^{N} q^i;t\right)$$

Since the first-order conditions are symmetric, a symmetric interior equilibrium $\phi^{n+1}\left(k^1,...,k^n;t\right)$ is characterized by:

$$\rho\left(\sum_{i=1}^{n}k^i + (N-n)\phi^{n+1};t\right) - c + \phi^{n+1}\rho_q\left(\sum_{i=1}^{n}k^i + (N-n)\phi^{n+1};t\right) = 0$$

Since $\pi^j(.;t)$ is strictly concave, ϕ^{n+1} is a maximum, hence it constitutes a best response to the others' strategies. ϕ^{n+1} is increasing in t, since by the implicit function theorem:

$$\frac{\partial\phi^{n+1}}{\partial t} = -\frac{\rho_t + \phi^{n+1}\rho_{qt}}{(N-n+1)\left(\rho_q + \frac{N-n}{N-n+1}\phi^{n+1}\rho_{qq}\right)} > 0$$

Furthermore, $\phi^{n+1}(t^n) = \phi^n(t^n) = k^n$. To see that, we observe that $\phi^{n+1}(t^n)$ verifies:

$$\frac{\partial\pi^j}{\partial q^j}\left(t^n\right) = \rho\left(\sum_{i=1}^{n}k^i + (N-n)\phi^{n+1};t^n\right) - c + \phi^{n+1}\rho_q\left(\sum_{i=1}^{n}k^i + (N-n)\phi^{n+1};t^n\right) = 0$$

while $\phi^n(t)$ *solves:*

$$\rho\left(\sum_{i=1}^{n-1}k^i+(N-n+1)\,\phi^n;t\right)-c+\phi^n\rho_q\left(\sum_{i=1}^{n-1}k^i+(N-n+1)\,\phi^n;t\right)=0$$

hence, since $\hat{q}^n(t^n)=k^n$ *by construction:*

$$\rho\left(\sum_{i=1}^{n}k^i+(N-n)\,k^n;t^n\right)-c+k_n\rho_q\left(\sum_{i=1}^{n}k^i+(N-n)\,k^n;t^n\right)=0$$

Since $\pi^j(.;t^n)$ *is strictly concave, there exists a unique value such that* $\frac{\partial\pi^j}{\partial q^j}(t^n)=0$ *hence* $\phi^{n+1}(t^n)=k^n=\phi^n(t^n)$.

Producer $(n+1)$ *produces* $\phi^{n+1}(k^1,...,k^n;t)$ *up to* t^{n+1} *defined by* $\phi^{n+1}(k^1,...,k^n;t)=k^{n+1}$. *As before, we can show by contradiction that* t^{n+1} *exists. Furthermore, since* $k^{n+1}\geq k^n$, *then* $t^{n+1}\geq t^n$.

We now show that $q^l=k^l$ *for* $l\leq n$ *is a best response to* $q^j=\phi^{n+1}$ *for* $j>n$. *Suppose* $q^i=k^i\ \forall i\leq n, i\neq l$ *and* $q^j=\phi^{n+1}$ *for* $j>n$. *Then:*

$$\frac{\partial\pi^l}{\partial q^l}=\rho\left(\sum_{\substack{i=1\\ i\neq j}}^{n}k^i+q^l+(N-n)\,\phi^{n+1};t\right)-c+q^l\rho_q\left(\sum_{\substack{i=1\\ i\neq j}}^{n}k^i+q^l+(N-n)\,\phi^{n+1};t\right)$$

hence

$$\begin{aligned}\left.\frac{\partial\pi^l}{\partial q^l}\right|_{q^l=k^l} &= \rho\left(\sum_{i=1}^{n}k^i+(N-n)\,\phi^{n+1};t\right)-c+k^l\rho_q\left(\sum_{i=1}^{n}k^i+(N-n)\,\phi^{n+1};t\right)\\ &= -\left(\phi^{n+1}-k^l\right)\rho_q\left(\sum_{i=1}^{n}k^i+(N-n)\,\phi^{n+1};t\right)>0\end{aligned}$$

since $\phi^{n+1}(t)>\phi^{n+1}(t^n)=k^n\geq k^j$ *for* $t>t^n$.

We have therefore completed step $(n+1)$. *By induction, the structure of the equilibria holds up until* $n=N$, *as long as we adopt the convention:* $t^{N+1}\to+\infty$, $\phi^{N+1}(t)=0$ *and* $\hat{Q}(t^N)=K=\hat{Q}(t)$ $\forall t\geq t^N$.

From the previous discussion, $\hat{q}^n(\mathbf{k};t)$ *and* $\hat{Q}(\mathbf{k};t)$ *are continuous and increasing in* t.

The expression of profits follow directly from the characterization of equilibria above. ■

G.2 Capacity constrained Cournot equilibrium with price cap

Lemma 4 *Suppose the capacity constraint binds before the price cap constraint, i.e.,* $t^N \leq \bar{t}\left(K, \bar{p}^W\right)$ *at the equilibrium. Producer* n*'s equilibrium profit for the constrained Cournot game in state* t *is:*

$$\begin{aligned}\Pi^n\left(k^n;\mathbf{k}_{-n}\right) &= \sum_{j=0}^{n-1}\int_{t^j}^{t^{j+1}} \phi^{j+1}\left(\rho\left(\hat{Q}\right)-c\right)f(t)\,dt \\ &\quad + k^n\left\{\sum_{j=n}^{N}\left[\int_{t^j}^{t^{j+1}}\left(\rho\left(\hat{Q}\right)-c\right)f(t)\,dt\right]+\int_{\bar{t}(K,\bar{p}^W)}^{+\infty}\left(\bar{p}^W-c\right)f(t)\,dt-r\right\}\end{aligned} \tag{17}$$

where t^j *for* $j = 1, ..., N$ *is the first state of the world such that all producers up to* j *are capacity constrained,* $t^0 = 0$ *and the convention* $t^{N+1} = \bar{t}\left(K, \bar{p}^W\right)$*,* ϕ^{j+1} *is producer* n*'s output in state* $t \in \left[t^j, t^{j+1}\right]$ *for* $j < n$*, and* $\hat{Q}$ *is the equilibrium aggregate output.*

Proof. *In the off-peak states of the world where at least one generator is unconstrained, i.e., with our previous notation* $t < t^N(K) \leq \bar{t}\left(K, \bar{p}^W\right)$*, imposition of the price cap has no impact on the equilibrium in these states, and Lemma 3 applies.*

Consider now the peak states of the world $t \geq t^N(K)$*, and* $\hat{Q}(t) = K = \sum_{m=1}^{N} k^m$*. As long as* $\rho(K;t) \leq \bar{p}^W$*, imposition of the price cap has no impact on the equilibrium in these states, and Lemma 3 applies.*

States of the world may exist where $\rho(K;t) > \bar{p}^W$*. Then, for* $t \geq \bar{t}\left(K, \bar{p}^W\right)$*,* $p(t) = \bar{p}^W$*. Facing a constant price, generators individually maximize production to maximize profit, hence* $q^n(t) = k^n$ *for all* n *is an equilibrium. This then yields equation (17). However, the SO must ration demand.* ■

G.3 Equilibrium investment

Lemma 5 *For any* $(k^n, \mathbf{k}_{-n})$*:*

$$\frac{\partial \Pi^n}{\partial k^n}\left(k^n,\mathbf{k}_{-n}\right)=\sum_{j=n}^{N}\left[\int_{t^j}^{t^{j+1}}\left(\rho\left(\hat{Q}\right)-c+k^n\rho_q\left(\hat{Q}\right)\frac{\partial \hat{Q}}{\partial k^n}\right)f(t)\,dt\right]+\int_{\bar{t}(K,\bar{p}^W)}^{+\infty}\left(\bar{p}^W-c\right)f(t)\,dt-r. \tag{18}$$

For any $k^N \geq \frac{K}{N}$

$$\frac{\partial^2 \Pi^N}{\left(\partial k^N\right)^2}\left(\frac{K}{N},...,\frac{K}{N},k^N\right) = \int_{t^N}^{\bar{t}\left(K,\bar{p}^W\right)}\left[2\rho_q\left(\hat{K};t\right)+k^N\rho_{qq}\left(\hat{K};t\right)\right]f(t)\,dt + k^N\rho_q\left(\hat{K};\bar{t}\left(\hat{K},\bar{p}^W\right)\right)f\left(\bar{t}\left(\hat{K},\bar{p}^W\right)\right)\frac{\partial\bar{t}\left(\hat{K},\bar{p}^W\right)}{\partial k^N} < 0 \tag{19}$$

where $\hat{K} = k^n + \frac{N-1}{N}K$*. Furthermore,* $\Pi^n\left(\frac{K}{N},..,\frac{K}{N}\right)$ *is globally concave.*

Proof. *The first-order derivative of profit function is:*

$$\frac{\partial \Pi^n}{\partial k^n} = \sum_{j=n}^{N}\left[\int_{t^j}^{t^{j+1}}\left(\rho\left(\hat{Q}\right)-c+k^n\rho_q\left(\hat{Q}\right)\frac{\partial\hat{Q}}{\partial k^n}\right)f(t)\,dt\right]+\int_{\bar{t}(K,\bar{p}^W)}^{+\infty}\left(\bar{p}^W-c\right)f(t)\,dt - r+\Delta_1$$

where

$$\begin{aligned}\Delta_1 &= k_n\sum_{j=n}^{N}\left[\left(\rho\left(\hat{Q}\left(t^{j+1}\right);t^{j+1}\right)-c\right)f\left(t^{j+1}\right)\frac{\partial t^{j+1}}{\partial k^n}-\left(\rho\left(\hat{Q}\left(t^j\right);t^j\right)-c\right)f\left(t^j\right)\frac{\partial t^j}{\partial k^n}\right]\\&\quad+\hat{q}^n\left(t^n\right)\left(\rho\left(\hat{Q}\left(t^n\right);t^n\right)-c\right)f\left(t^n\right)\frac{\partial t^n}{\partial k^n}-k^n\left(\bar{p}^W-c\right)f\left(\bar{t}\left(K,\bar{p}^W\right)\right)\frac{\partial\bar{t}\left(K,\bar{p}^W\right)}{\partial k^n}\\&= 0\end{aligned}$$

since $\hat{q}^n\left(t^n\right)=k^n$ *and* $\rho\left(\hat{Q}\left(\bar{t}\left(K,\bar{p}^W\right)\right);\bar{t}\left(K,\bar{p}^W\right)\right)=\bar{p}^W$*. This proves equation (18).*

Suppose $k^N \geq \frac{K^C}{N}$ *while* $k^n = \frac{K^C}{N}$ *for all* $n < N$*. Equation (18) yields:*

$$\frac{\partial \Pi^N}{\partial k^N}\left(\frac{K^C}{N},...,\frac{K^C}{N},k^N\right)=\int_{t^N(K)}^{\bar{t}\left(K,\bar{p}^W\right)(K)}\left(\rho(K;t)+k^N\rho_q(K;t)-c\right)f(t)\,dt+\int_{\bar{t}(K,\bar{p}^W)(K)}^{+\infty}\left(\bar{p}^W-c\right)f(t)\,dt-r.$$

Thus:

$$\frac{\partial^2 \Pi^N}{\partial\left(k^N\right)^2}\left(\frac{K^C}{N},...,\frac{K^C}{N},k^N\right) = \int_{t^N}^{\bar{t}\left(K,\bar{p}^W\right)}\left[2\rho_q\left(\hat{K};t\right)+k^N\rho_{qq}\left(\hat{K};t\right)\right]f(t)\,dt + k^N\rho_q\left(\hat{K};\bar{t}\left(K,\bar{p}^W\right)\right)f\left(\bar{t}\left(K,\bar{p}^W\right)\right)\frac{\partial\bar{t}\left(K,\bar{p}^W\right)}{\partial k^N} < 0.$$

This proves equation (19). Then, selecting $k^N = \frac{K}{N}$ proves the global concavity of $\Pi^n\left(\frac{K}{N}, .., \frac{K}{N}\right)$. ■

Lemma 6 *K^C solution of*

$$\int_{t^N(K^C)}^{\bar{t}\left(K,\bar{p}^W\right)(K^C)} \left(\rho\left(K^C;t\right) + \frac{K^C}{N}\rho_q\left(K^C;t\right) - c\right) f(t)\,dt + \int_{\bar{t}(K^C,\bar{p}^W)(K^C)}^{+\infty} \left(\bar{p}^W - c\right) f(t)\,dt = r$$

is the only symmetric equilibrium investment.

Proof. *If $k^n = \frac{K}{N}$, for all n, all producers are constrained simultaneously: $t^n = t^N$ for all n. The first order derivative (18) then becomes:*

$$\frac{\partial \Pi^n}{\partial k^n}\left(\frac{K}{N},...,\frac{K}{N}\right) = \int_{t^N(K)}^{\bar{t}\left(K,\bar{p}^W\right)(K)} \left(\rho(K;t) + \frac{K}{N}\rho_q(K;t) - c\right) f(t)\,dt + \int_{\bar{t}(K,\bar{p}^W)(K)}^{+\infty} \left(\bar{p}^W - c\right) f(t)\,dt - r.$$

$\frac{\partial \Pi^n}{\partial k^n}(0,...,0) = \int_0^{\bar{t}\left(K,\bar{p}^W\right)} (\rho(0;t) - c) f(t)\,dt + \int_{\bar{t}(K,\bar{p}^W)}^{+\infty} \left(\bar{p}^W - c\right) f(t)\,dt - r > \bar{p}^W - (c+r) > 0$ *since $\rho(0;t) > \bar{p}^W > (c+r)$ by equation (5).* $\lim_{K\to+\infty} \frac{\partial \Pi^n}{\partial k^n}\left(\frac{K}{N},...,\frac{K}{N}\right) = -r < 0$. *Hence $K^C > 0$ such that $\frac{\partial \Pi^n}{\partial k^n}\left(\frac{K^C}{N},...,\frac{K^C}{N}\right) = 0$ exists.*

We now prove that $\left(\frac{K^C}{N},...,\frac{K^C}{N}\right)$ is an equilibrium. Consider first a negative deviation: $k^1 \leq \frac{K^C}{N}$ while $k^n = \frac{K^C}{N}$ for all $n > 1$. Total installed capacity is $K = k^1 + \frac{N-1}{N}K^C \leq K^C$.

$$\begin{aligned}
\frac{\partial \Pi^1}{\partial k^1}\left(k^1, \frac{K^C}{N},...,\frac{K^C}{N}\right) &= \int_{t^1}^{t^N(K)} \left(\rho\left(\hat{Q}\right) + k^1\rho_q\left(\hat{Q}\right)\frac{\partial \hat{Q}}{\partial k^1} - c\right) f(t)\,dt \\
&\quad + \int_{t^N(K)}^{\bar{t}\left(K,\bar{p}^W\right)(K)} \left(\rho(K) + k^1\rho_q(K) - c\right) f(t)\,dt \\
&\quad + \int_{\bar{t}(K,\bar{p}^W)(K)}^{+\infty} \left(\bar{p}^W - c\right) f(t)\,dt - r \\
&= \int_{t^1}^{t^N(K)} \left(\rho\left(\hat{Q}\right) + k^1\rho_q\left(\hat{Q}\right)\frac{\partial \hat{Q}}{\partial k^1} - c\right) f(t)\,dt \\
&\quad + \int_{t^N(K)}^{\bar{t}\left(K,\bar{p}^W\right)} \left(\rho(K) + k^1\rho_q(K) - c\right) f(t)\,dt \\
&\quad - \int_{t^N(K^C)}^{\bar{t}\left(K^C,\bar{p}^W\right)} \left(\rho\left(K^C\right) + \frac{K^C}{N}\rho_q\left(K^C\right) - c\right) f(t)\,dt \\
&\quad + \int_{\bar{t}(K,\bar{p}^W)}^{+\infty} \left(\bar{p}^W - c\right) f(t)\,dt - \int_{\bar{t}(K^C,\bar{p}^W)}^{+\infty} \left(\bar{p}^W - c\right) f(t)\,dt
\end{aligned}$$

$t^N(K) \leq t^N(K^C)$ *and* $\bar{t}(K,\bar{p}^W)(K) \leq t^{\bar{p}^W}(K^C)$ *since* $K < K^C$*. Then:*

$$\begin{aligned}\frac{\partial \Pi^1}{\partial k^1}\left(k^1, \frac{K^C}{N}, ..., \frac{K^C}{N}\right) &= \int_{t^1}^{t^N(K)} \left(\rho\left(\hat{Q}\right) + k^1 \rho_q\left(\hat{Q}\right)\frac{\partial \hat{Q}}{\partial k^1} - c\right) f(t)\, dt \\ &+ \int_{t^N(K)}^{t^N(K^C)} \left(\rho(K) + k^1\rho_q(K) - c\right) f(t)\, dt \\ &+ \int_{t^N(K)}^{\bar{t}(K,\bar{p}^W)} \left(\rho(K) + k^1\rho_q(K) - \left(\rho\left(K^C\right) + \frac{K^C}{N}\rho_q\left(K^C\right)\right)\right) f(t)\, dt \\ &+ \int_{\bar{t}(K,\bar{p}^W)}^{\bar{t}(K^C,\bar{p}^W)} \left(\bar{p}^W - \rho\left(K^C\right) - \frac{K^C}{N}\rho_q\left(K^C\right)\right) f(t)\, dt.\end{aligned}$$

$\rho\left(\hat{Q}\right)+k^1\rho_q\left(\hat{Q}\right)-c = \left(k^1 - \phi^N\right)\rho_q\left(\hat{Q}\right)\frac{\partial \hat{Q}}{\partial k^1} \geq 0$ *for* $t \in \left[t^1, t^N(K)\right]$*.* $\rho\left(K; t^N(K)\right)+k^1\rho_q\left(K; t^N(K)\right) = c$*, and* $\rho_t(K) + k^1\rho_{qt}(K) \geq 0$*, hence* $\rho(K) + k^1\rho_q(K) - c \geq 0$ *for* $t \in \left[t^N(K), t^N\left(K^C\right)\right]$*.* $\rho_q(Q) + q\rho_{qq}(Q) < 0$*, hence* $\rho(K)+k^1\rho_q(K) \geq \rho(K)+\frac{K^C}{N}\rho_q(K) \geq \rho\left(K^C\right)+\frac{K^C}{N}\rho_q\left(K^C\right)$ $\left(k^1 - \frac{K^C}{N}\right)\rho_q\left(\hat{Q}\right)\frac{\partial \hat{Q}}{\partial k^1} \geq 0$ *for* $t \in \left[t^N(K), \bar{t}\left(K,\bar{p}^W\right)(K)\right]$*. Finally,* $\rho\left(K^C\right) \leq \bar{p}^W$ *for* $t \leq t^{\bar{p}^W(K^C)}$*, hence* $\bar{p}^W - \rho\left(K^C\right) - \frac{K^C}{N}\rho_q\left(K^C\right) \geq 0$ *for* $t \in \left[\bar{t}\left(K,\bar{p}^W\right)(K), t^{\bar{p}^W(K^C)}\right]$*. Thus,* $\frac{\partial \Pi^1}{\partial k^1}\left(k^1, \frac{K^C}{N}, ..., \frac{K^C}{N}\right) \geq 0$*: a negative deviation is not profitable.*

From equation (19), $\frac{\partial \Pi^N}{\partial k^N}\left(\frac{K^C}{N}, ..., \frac{K^C}{N}, k^N\right)$ *is decreasing for* $k^N \geq \frac{K^C}{N}$*, hence* $\frac{\partial \Pi^N}{\partial k^N}\left(\frac{K^C}{N}, ..., \frac{K^C}{N}, k^N\right) \leq 0$*: a positive deviation is not profitable.*

Therefore, $\left(\frac{K^C}{N}, ..., \frac{K^C}{N}\right)$ *is a symmetric equilibrium. Furthermore,* K^C *is the only symmetric equilibrium since* $\Pi^n\left(\frac{K}{N}, ..., \frac{K}{N}\right)$ *is globally concave.* ■

H Asymmetric equilibrium: sign of $\frac{\partial \bar{k}^2}{\partial \theta^1}$

Suppose there exists a unique interior equilibrium of the investment game, $\left(\bar{k}^1(\theta_1, \theta_2), \bar{k}^2(\theta_1, \theta_2)\right)$, defined by:

$$\begin{cases} \frac{\partial \Pi^1_{RO}}{\partial k^1} = \frac{\partial \Pi^1}{\partial k^1} + \frac{\theta^2}{(\theta^1+\theta^2)}\Psi + \frac{\bar{k}^1\theta^2 - \bar{k}^2\theta^1}{\Theta}\frac{\partial \Psi}{\partial K} = 0 \\ \frac{\partial \Pi^2_{RO}}{\partial k^2} = \frac{\partial \Pi^2}{\partial k^2} + \frac{\theta^1}{(\theta^2+\theta^1)}\Psi + \frac{\bar{k}^2\theta^1 - \bar{k}^1\theta^2}{\Theta}\frac{\partial \Psi}{\partial K} = 0 \end{cases}.$$

Define

$$\psi\left(k^1, k^2, \theta^1, \theta^2\right) = \frac{\partial \Pi^1}{\partial k^1}\left(k^1, k^2\right) + \frac{\theta^2}{(\theta^1+\theta^2)}\Psi\left(k^1+k^2\right) + \frac{\bar{k}^1\theta^2 - \bar{k}^2\theta^1}{\Theta}\frac{\partial \Psi}{\partial K}\left(k^1+k^2\right).$$

The first order conditions defining the equilibrium can recast as:

$$\psi\left(\bar{k}^1, \bar{k}^2, \theta^1, \theta^2\right) = \psi\left(\bar{k}^2, \bar{k}^1, \theta^2, \theta^1\right) = 0.$$

Full differentiation of the system with respect to θ^1 yields:

$$\begin{cases} \left(\psi_1 \frac{\partial \bar{k}^1}{\partial \theta^1} + \psi_2 \frac{\partial \bar{k}^2}{\partial \theta^1} + \psi_3\right)\left(\bar{k}^1, \bar{k}^2, \theta^1, \theta^2\right) = 0 \\ \left(\psi_1 \frac{\partial \bar{k}^2}{\partial \theta^1} + \psi_2 \frac{\partial \bar{k}^1}{\partial \theta^1} + \psi_4\right)\left(\bar{k}^2, \bar{k}^1, \theta^2, \theta^1\right) = 0 \end{cases},$$

where ψ_k corresponds to the derivative with respect to the k^{th} argument. Thus, assuming $\Delta = \psi_1\left(\bar{k}^1, \bar{k}^2, \theta^1, \theta^2\right)\psi_1\left(\bar{k}^2, \bar{k}^1, \theta^2, \theta^1\right) - \psi_2\left(\bar{k}^1, \bar{k}^2, \theta^1, \theta^2\right)\psi_2\left(\bar{k}^2, \bar{k}^1, \theta^2, \theta^1\right) \neq 0$,

$$\frac{\partial \bar{k}^2}{\partial \theta^1} = \frac{\psi_3\left(\bar{k}^1, \bar{k}^2, \theta^1, \theta^2\right)\psi_2\left(\bar{k}^2, \bar{k}^1, \theta^2, \theta^1\right) - \psi_4\left(\bar{k}^2, \bar{k}^1, \theta^2, \theta^1\right)\psi_1\left(\bar{k}^1, \bar{k}^2, \theta^1, \theta^2\right)}{\Delta}.$$

$$\psi_3\left(\bar{k}^1, \bar{k}^2, \theta^1, \theta^2\right) = -\frac{\theta^2}{\left(\theta^1 + \theta^2\right)^2}\left(\Psi + K\frac{\partial \Psi}{\partial K}\right) = -\psi_4\left(\bar{k}^2, \bar{k}^1, \theta^2, \theta^1\right),$$

thus

$$\frac{\partial \bar{k}^2}{\partial \theta^1} = -\frac{\theta^2}{\left(\theta^1 + \theta^2\right)^2}\left(\Psi + K\frac{\partial \Psi}{\partial K}\right)\frac{\psi_1\left(\bar{k}^1, \bar{k}^2, \theta^1, \theta^2\right) + \psi_2\left(\bar{k}^2, \bar{k}^1, \theta^2, \theta^1\right)}{\Delta}.$$

We examine the sign of each term in turn. Suppose demand follows the specification presented in Section 3. Then,

$$\left(\Psi + K\frac{\partial \Psi}{\partial K}\right)\left(K, \bar{p}^S\right) = \frac{1}{\alpha(1+\lambda)}\left(\frac{a_0 - bK - \left(\alpha\bar{p}^S + (1-\alpha)p^R\right)}{a_1}\right)^{\lambda}\left(a_0 - \left(\alpha\bar{p}^S + (1-\alpha)p^R\right) - (2+\lambda)bK\right)$$

which is negative for most values of K and $\bar{p}^S$. We assume this property holds for other specifications.

Then,

$$\psi_1\left(\bar{k}^1, \bar{k}^2, \theta^1, \theta^2\right) = \frac{\partial^2 \Pi^1}{\left(\partial k^1\right)^2} + \frac{2\theta^2}{\left(\theta^1 + \theta^2\right)}\frac{\partial \Psi}{\partial K} + \frac{\bar{k}^1\theta^2 - \bar{k}^2\theta^1}{\left(\theta^1 + \theta^2\right)}\frac{\partial^2 \Psi}{(\partial K)^2}$$

and

$$\psi_2\left(\bar{k}^1, \bar{k}^2, \theta^1, \theta^2\right) = \frac{\partial^2 \Pi^1}{\partial k^1 \partial k^2} + \frac{\theta^2 - \theta^1}{\left(\theta^1 + \theta^2\right)}\frac{\partial \Psi}{\partial K} + \frac{\bar{k}^1\theta^2 - \bar{k}^2\theta^1}{\left(\theta^1 + \theta^2\right)}\frac{\partial^2 \Psi}{(\partial K)^2}.$$

Assuming $\frac{\partial^2 \Pi^1}{\left(\partial k^1\right)^2}$ and $\frac{\bar{k}^1\theta^2 - \bar{k}^2\theta^1}{\left(\theta^1+\theta^2\right)}\frac{\partial^2 \Psi}{(\partial K)^2}$ are small, $\psi_1\left(\bar{k}^1, \bar{k}^2, \theta^1, \theta^2\right) < 0$, which guarantees that $\bar{k}^1$ is indeed the best response to $\bar{k}^2$.

$$\psi_1\left(\bar{k}^1,\bar{k}^2,\theta^1,\theta^2\right)+\psi_2\left(\bar{k}^2,\bar{k}^1,\theta^2,\theta^1\right)=\frac{\partial^2\Pi^1}{\left(\partial k^1\right)^2}+\frac{\partial^2\Pi^1}{\partial k^1\partial k^2}+\frac{\partial\Psi}{\partial K}.$$

$\frac{\partial\Psi}{\partial K}<0$, and $\frac{\partial^2\Pi^1}{\partial k^1\partial k^2}<0$ since capacities are strategic complements. $\frac{\partial^2\Pi^1}{\left(\partial k^1\right)^2}<0$ for $k^1>k^2$. Suppose $\frac{\partial^2\Pi^1}{\left(\partial k^1\right)^2}$ is small enough, then

$$\psi_1\left(\bar{k}^1,\bar{k}^2,\theta^1,\theta^2\right)+\psi_2\left(\bar{k}^2,\bar{k}^1,\theta^2,\theta^1\right)<0.$$

Finally,

$$\psi_1\left(\bar{k}^1,\bar{k}^2,\theta^1,\theta^2\right)-\psi_2\left(\bar{k}^1,\bar{k}^2,\theta^1,\theta^2\right)=\frac{\partial^2\Pi^1}{\left(\partial k^1\right)^2}-\frac{\partial^2\Pi^1}{\partial k^1\partial k^2}+\frac{\partial\Psi}{\partial K}.$$

Assuming that $\frac{\partial\Psi}{\partial K}$ dominates, $\psi_1\left(\bar{k}^1,\bar{k}^2,\theta^1,\theta^2\right)-\psi_2\left(\bar{k}^1,\bar{k}^2,\theta^1,\theta^2\right)<0$.

Consider first the case $\theta^1<\theta^2$. Then, assuming $\frac{\bar{k}^1\theta^2-\bar{k}^2\theta^1}{\left(\theta^1+\theta^2\right)}\frac{\partial^2\Psi}{(\partial K)^2}$ is small, $\psi_2\left(\bar{k}^1,\bar{k}^2,\theta^1,\theta^2\right)<0$. Then:

$$\psi_1\left(\bar{k}^1,\bar{k}^2,\theta^1,\theta^2\right)<\psi_2\left(\bar{k}^1,\bar{k}^2,\theta^1,\theta^2\right)\Rightarrow\psi_1\left(\bar{k}^1,\bar{k}^2,\theta^1,\theta^2\right)\psi_1\left(\bar{k}^2,\bar{k}^1,\theta^2,\theta^1\right)>\psi_2\left(\bar{k}^1,\bar{k}^2,\theta^1,\theta^2\right)\psi_1\left(\bar{k}^2,\bar{k}^1,\theta^2,\theta^1\right)$$

$$\psi_1\left(\bar{k}^2,\bar{k}^1,\theta^2,\theta^1\right)<\psi_2\left(\bar{k}^2,\bar{k}^1,\theta^2,\theta^1\right)\Rightarrow\psi_2\left(\bar{k}^1,\bar{k}^2,\theta^1,\theta^2\right)\psi_1\left(\bar{k}^2,\bar{k}^1,\theta^2,\theta^1\right)>\psi_2\left(\bar{k}^2,\bar{k}^1,\theta^2,\theta^1\right)\psi_2\left(\bar{k}^1,\bar{k}^2,\theta^1,\theta^2\right)$$

thus

$$\psi_1\left(\bar{k}^1,\bar{k}^2,\theta^1,\theta^2\right)\psi_1\left(\bar{k}^2,\bar{k}^1,\theta^2,\theta^1\right)>\psi_2\left(\bar{k}^2,\bar{k}^1,\theta^2,\theta^1\right)\psi_2\left(\bar{k}^1,\bar{k}^2,\theta^1,\theta^2\right)\Leftrightarrow\Delta>0.$$

Thus, $\frac{\partial\bar{k}^2}{\partial\theta^1}=-\frac{\theta^2}{\left(\theta^1+\theta^2\right)^2}\underbrace{\left(\Psi+K\frac{\partial\Psi}{\partial K}\right)}_{-}\frac{\overbrace{\psi_1\left(\bar{k}^1,\bar{k}^2,\theta^1,\theta^2\right)+\psi_2\left(\bar{k}^2,\bar{k}^1,\theta^2,\theta^1\right)}^{-}}{\underset{+}{\Delta}}<0$. If $\theta^1>\theta^2$, $\psi_2\left(\bar{k}^2,\bar{k}^1,\theta^2,\theta^1\right)<0$. The same argument also yields $\frac{\partial\bar{k}^2}{\partial\theta^1}<0$.