DISCUSSION PAPER SERIES

No. 2998

CAPACITY CHOICES IN LIBERALIZED ELECTRICITY MARKETS

Fidel Castro-Rodriguez, Pedro L Marín
and Georges Siotis

INDUSTRIAL ORGANIZATION

Centre for Economic Policy Research

www.cepr.org

Available online at: **www.cepr.org/pubs/dps/DP2998.asp**

ISSN 0265-8003

CAPACITY CHOICES IN LIBERALIZED ELECTRICITY MARKETS

Fidel Castro-Rodriguez, Universidade de Vigo
Pedro L Marín, Universidad Carlos III de Madrid and CEPR
Georges Siotis, Universidad Carlos III de Madrid

Discussion Paper No. 2998
October 2001

Centre for Economic Policy Research
90–98 Goswell Rd, London EC1V 7RR, UK
Tel: (44 20) 7878 2900, Fax: (44 20) 7878 2999
Email: cepr@cepr.org, Website: www.cepr.org

CEPR Discussion Paper No. 2998

October 2001

ABSTRACT

Capacity Choices in Liberalized Electricity Markets*

We develop a theoretical model of long-run investment decisions on capacity in the context of a liberalized electricity market. The sector's idiosyncrasies such as the uncertainty surrounding future supply and demand, as well as technological constraints, are explicitly modelled. The model is sufficiently flexible to describe the situation in different systems. We derive the level of capacity that maximizes social welfare, and compare it to a decentralized outcome. We show that in the absence of any regulation, private investment decisions on capacity unambiguously lead to a socially sub-optimal outcome, and we illustrate these results using simulations.

JEL Classification: L13, L43 and L94
Keywords: capacity, electricity, liberalization, long-run investment and regulation

Fidel Castro-Rodriguez
Departamento de Fundamentos del Analisis Economico
Universidade de Vigo
Campus Lagoas-Marcosende
36200-Vigo (Pontevedra)
SPAIN
Tel: (34 98) 681 2526
Fax: (34 98) 681 2401
Email: fcastro@uvigo.es

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=155766

Pedro L Marín
Departamento de Economía
Universidad Carlos III de Madrid
C./ Madrid, 126
28903 Getafe (Madrid)
SPAIN
Tel: (34 91) 624 9646
Fax: (34 91) 624 9875
Email: marin@eco.uc3m.es

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=131657

Georges Siotis
Departamento de Economia
Universidad Carlos III de Madrid
Calle Madrid 126
28903 Getafe
Madrid
SPAIN
Tel: (34 91) 624 9312
Fax: (34 91) 624 9875
Email: siotis@eco.uc3m.es

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=119325

* This research was partially financed by Union Electrica Fenosa (UEF), whose financial support is gratefully acknowledged, and whose staff greatly helped us navigate through the technological and regulatory intrinsicacies of the electricity sector. Maria Angeles de Frutos and Jose-Luis Moraga provided very insightful suggestions and comments. All three authors also gratefully acknowledge financial support from the Ministerio de Ciencia y Technologia (SEC1999-1236-C02-01). This Paper is produced as part of a CEPR Research Network on 'The Evolution of Market Structure in Network Industries', funded by the European Commission under the Training and Mobility of Researchers Programme (Contract No. ERBFMRX-CT98-0203)

Submitted 15 September 2001

Capacity Choices in Liberalised Electricity Markets*

Fidel Castro*, Pedro L. Marín**, and Georges Siotis***†

This version: July 2001

Abstract

We develop a theoretical model of long-run investment decisions in capacity in the context of a liberalised electricity market. The sector's idiosyncrasies such as the uncertainty surrounding future demand and supply as well as technological constraints are explicitly modelled. The model is sufficiently flexible to describe the situation in different systems. We derive the level of capacity that maximizes social welfare, and compare it to a decentralised outcome. We show that in the absence of any regulation, private investment decisions in capacity unambiguously lead to a socially sub-optimal outcome, and we illustrate these results using simulations.

JEL classification codes: L13; L43; L94.

Keywords: Electricity; Capacity; Long-run investment; Regulation; Liberalisation

1. Introduction

Following England and Wales' lead, many countries have embarked on a process of liberalisation of their respective electricity sectors. This trend has been observed in developed economies (e.g. Scandinavian Noordpool, Spain), as well as

*This research was partially financed by Unión Eléctrica Fenosa (UEF), whose financial support is gratefully acknowledged, and whose staff greatly helped us navigate through the technological and regulatory intrinsicacies of the electricity sector. Maria Angeles de Frutos and José-Luis Moraga provided very insightful suggestions and comments. All three authors also gratefully acknowledge financial support from the Ministerio de Ciencia y Technología (SEC1999-1236-C02-01).

†* Universidad de Vigo, ** Universidad Carlos III de Madrid and CEPR, *** Universidad Carlos III de Madrid. Corresponding author: siotis@eco.uc3m.es

in emerging market economies (e.g. Chile, Argentina, or Peru). In some cases, liberalisation has been first implemented regionally (California, PJM -Pennsylvania, New-Jersey, and Maryland-, Victoria), though there are instances of uniform deregulation at the country level (Spain, Noordpool, and most emerging market economies mentioned earlier).

Liberalisation in all these markets has led to the establishment of spot wholesale markets for electricity. The common characteristic of all electricity pools is that generators make bids to supply a given amount of electricity at a certain price. A market operator orders these bids from highest to lowest constructing the market offer curve, and the intersection of this curve with a demand curve yields a price at which all trades occur. In all countries, including England and Wales (E&W thereafter), deregulation has been a recent phenomenon, which means that none of the aforementioned electricity sectors can be considered to have reached their long-run steady state. In addition, regulatory adjustments have been frequent, and it is as yet unclear what the long-run "rules of the game" will be in each of these liberalised electricity markets. For instance, the E&W centralised and compulsory pool has been replaced by the New Electricity Trading Arrangements (NETA) in early 2001. Spain's regulatory framework has been adjusted almost every year since deregulation started.

With the exception of von der Fehr and Harbord (1997), the literature, both theoretical and empirical, has focused on the final stages of the liberalised electricity game, that is how firms compete to supply energy. For instance, Green and Newbery (1992), von der Fehr and Harbord (1993), Borenstein and Bushnell (1999), Borenstein, Bushnell, and Stoft (2000), and García-Díaz and Marín (2000) provide a theoretical analysis of the English, Californian and Spanish markets. The empirical literature has mainly focused on identifying the degree of individual or collective market power exercised by firms in the pool. Wolfram (1999) examines the English case, while Joskow and Kahn (2000), and Borenstein, Bushnell, and Wolak (2000) provide evidence pertaining to California. All these papers focus on short-run outcomes, and take existing capacity as exogenously given.

Both theory and evidence suggest that decentralised markets yield outcomes that approximate a socially optimal situation. In the case of the electricity industry, the sector's idiosyncracies may not warrant such an a priori judgement. Indeed, the power crisis that has been unfolding in California since the Summer of 2000 suggests that deregulation may have unexpected adverse effects for social welfare. In that context, the issue of whether a decentralised market approximates the social optimum is a relevant one. This is the central issue dealt with in this paper.

von der Fehr and Harbord (1997) present a model of long-run investment choices. They demonstrate that, under competitive conditions, a no-intervention private outcome will yield an insufficient level of installed capacity. Consequently, it is necessary to allow firms to charge a mark-up which in turn generates a sufficient amount of revenue to cover fixed costs. Under this scenario, the decentralised outcome may yield under as well as over investment in capacity, depending on the mark-up that is being charged. Logically, there always exists a level of mark-up such that the private and social outcomes coincide. This implies that the social optimal level of capacity is attained at the cost of market power. As a general statement, von der Fehr and Harbord (1997) only obtain over-investment under fairly restrictive conditions, such as very high mark-ups or a convex distribution of demand (the latter implying that the bulk of consumption is either very high or very low). Under reasonable assumptions, such as a concave distribution of demand, over-investment never occurs.

Our paper shares the same motivation as that of von der Fehr and Harbord, but our model differs from theirs in a number of respect. First, our model distinguishes between consumers that can undertake demand side bidding and those that cannot. Second, we explicitly allow for brown-outs and black-outs (in von der Fehr and Harbord, there always exists a price that clears the market). Third, we run simulations to illustrate our results. This allows us to take into account the sector's idiosyncracies, such as the degree of demand and supply side variability as well as the existence of distinct technologies. Simulation results suggests that the model adequately describes existing situations. For instance, our results confirm the widely held view that the amount of excess capacity in Spain is rather narrow, an assessment shared by industry observers (see, for instance, *The Financial Times* 29/01/2001, Lex Column which states that "Only Spain seems likely in the medium term to have inadequate generation capacity").

We represent firms' decisions as a two stage game. During the first stage, agents decide on the level of capacity to install, a decision that we consider to be a long-run one. During stage two, firms compete to supply energy. We focus our analysis on the long-run decision, namely the decision to install new capacity. The main finding is that deregulated markets under-invest in capacity. This result is robust to changes in the number of firms in the industry, the cost of installing new capacity, or the nature of competition in the final stage of the game. We illustrate our findings with simulations of the model using publicly available data. In that exercise, we are careful to explicitly take into account the idiosyncracies of electricity systems such as demand and well as supply side uncertainty.

The paper is organised as follows: section 2 motivates the exercise, while section 3 presents the model. In section 4, we describe how we build our simulations and present the central results. Section 5 discusses some policy implications and concludes.

2. Motivation

The fact that a decentralised market may sometimes yield an outcome distinct from the social optimum is related to the sector's idiosyncracies. Before presenting the model, we briefly recall its main characteristics.

First of all, electricity energy can not be stored, except at a prohibitive cost. In addition, in an electricity system, demand must almost exactly equal supply in real time to ensure that the system's integrity is maintained.[1] These two characteristics imply that fluctuations in supply can not be smoothed via changes in inventories, and that installed capacity may only be dispatched a few hours per year. Second, prices are an imperfect signal compared to other markets. On the consumption side, demand is inelastic save for very large consumers. Even for those, the elastic part of their demand is quite inelastic, and in any case, this represents a tiny proportion of their total consumption.[2] This is the so-called interruptible load, for which large customers pay lower overall rates for their energy in exchange of bearing the risk of being cut-off when the system is under strain. For small residential customers, the price often consists in a two-part, fixed price, tariff. To date, residential customers in deregulated electricity markets have not yet been exposed to real time energy pricing.[3] Consequently, they are not

[1] The system operator enjoys some limited room for manoeuvre, as the system's voltage may be allowed to vary slightly.

[2] The Californian situation in January 2001 illustrates this point. Some large consumers entered into contracts that stipulated the number of hours during which they could be interrupted during an entire calendar year, in exchange for lower energy prices. By the end of January, this maximum number of hours had already been reached, following the wave of brown-outs and black-outs. This is direct evidence that the proportion of total demand that is price sensitive is quite limited (due to the intrinsic inelasticity of demand, and the existence of some regulated prices).

[3] In some markets that are still regulated (e.g. France or Greece), the publicly owned utilities offer to charge distinct residential prices at different moments of the day (e.g. cheap night rate). However, the menus are often restricted to two or three prices that remain fixed during a number of hours, and thus do not allow real time pricing when load reaches its peak (an event of relatively short duration that is difficult to predict with a degree of precision sufficient

sensitive to market prices and it is not possible to distinguish residential customers according to their valuation of the electricity they purchase.[4]

On the supply side, when demand approaches the level of installed capacity, the supply schedule becomes steeper, and ends-up vertical when installed capacity is reached. This simply reflects that supply is totally inelastic in the short-run when the level of installed capacity is reached. According to industry sources, it takes a minimum of ten months to install new capacity in the best of cases. On average, it takes a bit more than two years to expand supply. Also, all capacity is not always available to generate power because of failures or maintenance work.

In addition, both future supply and demand are subject to uncertainty. On the demand side, events such as weather conditions or economic growth can generate important fluctuations. On the supply side, unexpected outages and/or repair and maintenance work, the price and availability of inputs, as well as hydrologic conditions (in systems that partially rely on dams), all contribute to generate a considerable amount of medium to long term uncertainty. Finally, the electricity industry differs from large chunks of the economy, as it is a basic input for many activities. Consequently, electricity outages have the potential to generate large social costs.[5] In that context, the sector's oligopolistic structure gains special importance.

3. The Model

The basic architecture of the model is presented in figure 1. The demand is made-up of two parts. The flat segment represents the demand that is non-modulable in a given hour.[6] v is the average value that consumers give to one unit of electricity

to maintain the system's integrity). Consequently, these systems do not allow for intra-hour consumption smoothing.

[4] Real time energy pricing may in the future be common for residential customers as the metering technology is by and large available at a non-prohibitive cost. Assuming that residential customers react in a similar manner to large customers, this would result in a higher elasticity of demand only for a fraction of total consumption. Thus, even with real time pricing, total demand would remain highly inelastic.

[5] A recent example is that of aluminium smelters that had set-up operations in the North-West of the US to benefit from cheap electricity rates. The Californian crisis has resulted in much higher energy prices, that eventually spilled over to the North-West. Some smelters have seen their profits driven to zero, and about a third of smelting capacity has been shut down and staff laid off (*The Financial Times* (22/12/2000).

[6] What we call non-modulable load are consumers that cannot adjust their within the hour consumption, either because they do not receive price signals, or because technological con-

net of transportation, distribution, and administrative costs. Thus, v indicates the maximum amount that consumers are willing to pay for energy. In principle, this fraction of demand should be downward sloping. However, it is not possible to differentiate consumers according to their valuation of electricity at each moment in time. As a result, they are cut-off randomly when the system is short of capacity. Consequently, we make use of this average valuation of electricity for modelling purposes.

The second segment of the demand curve is downward sloping. It represents the demand for energy that stems from the bids that electricity suppliers make on behalf of interruptible load, as well the energy used by pump storage.[7] For that segment where we say that demand is modulable, demand is given by:

$$q(p) = a - bp$$

where p is the spot price and b is the slope of the demand curve.[8] Thus, the inverse demand function is defined as:

$$P(q) = \frac{a - q}{b}$$

Figure 1 depicts three realisations of demand, which reflect the fluctuations in sales that a plant faces during its lifetime. This variability stems from real time fluctuations, as well as mid to long term uncertainty that stems from the business cycle. The last source of uncertainty are hydrological conditions. We assume that hydro capacity is despatched when demand is at its highest, that is

ditions impedes them to do so. Examples of the latter involve some refrigeration operations, or metallurgical production. Regarding the former, residential customers that pay a regulated tariff per KWh do not receive within the hour price signals.

[7] Electricity suppliers are agents that buy electricity in the pool, pay a fee for its transmission, and sell it to a final consumer. "Pump storage" refers to a common practice (at least in Spain), whereby dams use electricity to pump-up water into the reservoir. Thus, the amount of "pump storage" is at its maximum when spot prices are at their minimum (the reservoir can be refilled cheaply), and zero when prices are highest (the cost of pumping water upstream is larger than the expected revenue). This practice, which is the closest to inter-temporal smoothing of supply, represents a small proportion of total capacity.

[8] Note that we implicitly assume that $\frac{a}{b} = v$, while the existence of modulable demand could result in bids higher than v. We do not contemplate this possibility for the sake of simplicity. Moreover, modulable demand only represents a small fraction of total demand, and allowing for bids above v would not change the essence of the results.

there is peak-load shaving.[9] Consequently, the demand curve depicted in figure 1 pertains to the residual demand for conventional thermic generators (total demand minus the part covered by dams and renewable energy sources that operate under a special regime). Last, we assume that v and the slope remain constant, so that shifts in demand are parallel within the interval $[\underline{a}, \overline{a}]$. More precisely, the uncertainty parameter a is distributed in $[\underline{a}, \overline{a}]$, with a distribution function $F(a)$ whose density is given by $f(a)$, and where $\underline{a}$ is associated with the lowest possible level of demand, and $\overline{a}$ with the maximum level of demand.

Insert Figure 1 about here

On the cost side, we assume that the existing technological mix yields an upward marginal cost schedule up to the level of installed capacity, which we denote k. Once the level of available capacity is reached, the marginal cost becomes infinite, which reflects the fact that supply can not be expanded in the short-run. Concretely, we specify the marginal cost function as:

$$MC(q) = \begin{cases} c_0 + c_1 q \text{ for } q \leq RAk \\ +\infty \text{ for } q > RAk \end{cases}$$

where c_0 and c_1 are the intercept and slope of the marginal cost schedule, and q is the industry's aggregate output. R and A stem from technical restrictions (defined below). The intercept and the slope will be given by the values of marginal costs for the cheapest and most expensive technologies available. Last, we denote f as the unitary cost associated with one unit of capacity, that we assume to be constant.

We model firms' decisions as a two stage game. At stage one, which we take to represent long-run decisions, firms simultaneously decide on how much capacity to install. In the second stage (short-run decision) firms compete by making supply bids in a spot market. With respect to this second decision, we assume that firms supply energy competitively, so that price is equal to the industry's marginal cost when capacity is not binding.[10] Once the capacity constraint is reached, prices increase along the demand curve up to v. In this wholesale market, supply

[9] It seems reasonable to assume that hydro capacity is exogenously given. The assumption we adopt is simply that hydro power will be sold when it is most valuable. This is what would occur when prices increases monotonically with demand (this encompasses a monopoly situation, centralised planning, and perfect competition). See Garcia-Diaz and Marín (2000) for a discussion of this issue.

[10] We discuss the importance of this assumption for our results in section 3.2.1.

and demand are matched in real time by the market operator. When demand at price v exceeds installed capacity, some consumers are rationed randomly, so that there is a zero probability of total collapse of the system. Thus, our model features rolling browns-out (when interruptible load is cut-off), and rolling black outs (when an entire geographical area is left without energy for a finite period of time). For a low level of realised demand such as D_1, installed capacity is sufficient to cover the entire demand for equilibrium price p_1. When demand at marginal cost exceeds installed capacity, such as D_2, the price is given by the demand schedule. Finally, if realised demand is D_3, then demand at price v exceeds capacity and some consumers (both interruptible and non-modulable) will be blacked-out. Consequently, we have that:

$$p = \begin{cases} c_0 + c_1 q & \text{if} \quad k > q_c(a) \\ P(k) & \text{if} \quad q(a,v) < k \leq q_c(a) \\ v & \text{if} \quad k \leq q(a,v) \end{cases}$$

where $q_c(a)$ represents the level of demand at which the demand curve intersects the marginal cost schedule, and $q(a,v)$ is the intersect between the flat and downward segments of the demand curve.

Clearly, the choice of capacity becomes the central variable in this model. Figure 1 also allows a decomposition of total welfare. Suppose that demand is at D_1, then area B_1 represents firms' surplus, while B_2 gives that of consumers. Because of maintenance work and unexpected outages, only a fraction of total installed capacity k is available at any point in time. We denote this availability ratio A, so that the capacity that can be effectively despatched is Ak. Also, the system's operator maintains a reserve margin, that we denote r, so that the final amount of capacity dispatched in the system is $(1-r)Ak$. To simplify notation, we set $R = (1-r)$. Finally, we assume that technological conditions as well as the possible realisations of demand are common knowledge among agents.

3.1. Characterisation of the Social Optimum

We consider a situation in which a benevolent regulator maximizes the surplus of producers and final consumers. We give equal weight to both groups in the social welfare function. This will allow us to establish the relevant benchmark that can then be compared to a decentralised outcome.

Efficient use of installed capacity implies marginal cost pricing as long as the capacity constraint is not binding. By contrast, when the price that would result in efficient rationing is above v, then a proportion of non-modulable consumers and the whole of interruptible load will have to be blacked-out.

In order to compute the social welfare function, we define the following values for a:

Definition:

i) Let $a_c \in [\underline{a}, \overline{a}]$ be the value of a associated with a demand function such that installed capacity just covers demand at a price equal to marginal cost. If we define $q_c(a) = \frac{a - bc_0}{1 + bc_1}$ as the level of demand which corresponds to the intersection of the demand curve with the marginal cost schedule, then $q_c(a_c) = \frac{a_c - bc_0}{1 + bc_1} = RAk$. Consequently, we have that:

$$a_c = bc_0 + (1 + bc_1)RAk$$

ii) Let $a_m \in [\underline{a}, \overline{a}]$ be such that $q(a_m, v) = (a_m - bv) = RAk$. That is, a_m is associated with a realisation of demand for which installed capacity just covers demand at the maximum price (given by v). Consequently, we have that:

$$a_m = RAk + bv$$

In order to define the social welfare function, we have to identify three different cases.

Case 1: $a \in [\underline{a}, a_c]$, i.e. installed capacity is sufficient to cover demand at price p_c, where p_c is equal to marginal cost ($p_c = \frac{c_0 + ac_1}{1 + bc_1}$). Thus, social welfare amounts to:

$$W_1 = \int_{p_c}^{v} (a - bp)dp + \left[p_c q_c - \int_0^{q_c} (c_0 + c_1 q)dq \right] - fk$$

where the first term represents the surplus obtained by consumers (modulable and non-modulable), and the second term the surplus obtained by generators.[11]

Case 2: $a \in [a_c, a_m]$, that is, capacity is sufficient to cover the whole of fixed rate customers, but only a fraction of interruptible load. In that case, welfare is given by:

[11] We are implicitly assuming that the final price, net of transport and distribution costs, is equal to the pool's price p_c. If prices to non-modulable consumers are regulated by a tariff, then the difference between the latter and the pool's price would represent the per unit profit obtained by energy suppliers. The difference between the regulated tariff and v would represent net consumer surplus per unit of energy consumed.

$$W_2 = \int_{\frac{a-RAk}{b}}^{v} (a - bp)dp + \left[\left(\frac{a - RAk}{b} \right) Ak - \int_0^{Ak} (c_0 + c_1 q)dq \right] - fk$$

where the first term represents the surplus consumers, and the second term are generators profits.

Case 3: $a \in [a_m, \overline{a}]$, i.e. it is not possible to provide energy to all customers, even at v. Thus, welfare becomes:

$$W_3 = \left[vAk - \int_0^{Ak} (c_0 + c_1 q)dq \right] - fk$$

In this case, welfare only consists of generator's profits.

To complete our description of expected welfare, we have to specify the relationship between possible realisations of demand and the level of installed capacity. To this end, we have to define five distinct regions, as depicted in Figure 2. In terms of notation, upper key letters are associated with our three cases, while lower key letters refer to expected values in the five distinct regions (e.g., W_1 is welfare in case 1, while w_1 is expected welfare in region one). Lower key letters that are not indexed refer to the global function (e.g., w is the global welfare function).

Region 1: $k \in \left[0, \frac{(\underline{a}-bv)}{RA}\right]$, that is, installed capacity is insufficient to cover the minimum level of demand, even if price is set equal to v. Under these circumstances, case 3 will occur with probability one, and cases 1 and 2 will never arise. Consequently, welfare in this region is given by:

$$w_1 = \int_{\underline{a}}^{\overline{a}} \left[(vAk - \int_0^{Ak} (c_0 + c_1 q)dq - fk \right] f(a)da$$

Insert Figure 2 about here

Region 2: $k \in \left[\frac{(\underline{a}-bv)}{RA}, \frac{(\underline{a}-bc_0)}{(1+bc_1)RA}\right]$ which implies case 1 will never occur, while cases 2 and 3 materialise with a positive probability. Thus, the welfare function is given by:[12]

[12] In what follows, note that a_m and a_c are both functions of k. For notational simplicity, we write a_m and a_c instead of $a_m(k)$ and $a_c(k)$.

$$
\begin{aligned}
w_2 &= \int_{\underline{a}}^{a_m}\left[\int_{\frac{a-RAk}{b}}^{v}(a-bp)dp+\left[(\frac{a-RAk}{b})Ak-\int_0^{Ak}(c_0+c_1q)dq\right]-fk\right]f(a)da \\
&\quad+\int_{a_m}^{\overline{a}}\left[vAk-\int_0^{Ak}(c_0+c_1q)dq-fk\right]f(a)da
\end{aligned}
$$

Region 3: $k\in\left[\frac{(\underline{a}-bc_0)}{(1+bc_1)RA},\frac{(\overline{a}-bv)}{RA}\right]$. Under this scenario, all three cases occur with a positive probability. Welfare becomes:

$$
\begin{aligned}
w_3 &= \int_{\underline{a}}^{a_c}\left[\int_{p_c}^{v}(a-bp)dp+\left[p_cq_c-\int_0^{q_c}(c_0+c_1q)dq\right]-fk\right]f(a)da \\
&\quad+\int_{a_c}^{a_m}\left[\int_{\frac{a-RAk}{b}}^{v}(a-bp)dp+\left[(\frac{a-RAk}{b})Ak-\int_0^{Ak}(c_0+c_1q)dq\right]-fk\right]f(a)da \\
&\quad+\int_{a_m}^{\overline{a}}\left[\left[vAk-\int_0^{Ak}(c_0+c_1q)dq\right]-fk\right]f(a)da
\end{aligned}
$$

Region 4: $k\in\left[\frac{(\overline{a}-bv)}{RA},\frac{(\overline{a}-bc_0)}{(1+bc_1)RA}\right]$. In this case, the level of installed capacity stands in between the maximum level of demand at v, and the maximum demand at a price equal to marginal cost. Consequently, case 3 never occurs, while cases 1 and 2 have a positive probability associated to them. Welfare thus becomes:

$$
\begin{aligned}
w_4 &= \int_{\underline{a}}^{a_c}\left[\int_{p_c}^{v}(a-bp)dp+\left[p_cq_c-\int_0^{q_c}(c_0+c_1q)dq\right]-fk\right]f(a)da \\
&\quad+\int_{a_c}^{\overline{a}}\left[\int_{\frac{a-RAk}{b}}^{v}(a-bp)dp+\left[(\frac{a-RAk}{b})Ak-\int_0^{Ak}(c_0+c_1q)dq\right]-fk\right]f(a)da
\end{aligned}
$$

Region 5: $k\in\left[\frac{(\overline{a}-bc_0)}{(1+bc_1)RA},+\infty\right]$. This last situation corresponds to one where installed capacity is greater than maximum demand at a price equal to marginal cost. In that region, welfare is given by:

$$
w_5=\int_{\underline{a}}^{\overline{a}}\left[\int_{p_c}^{v}(a-bp)dp+\left[p_cq_c-\int_0^{q_c}(c_0+c_1q)dq\right]-fk\right]f(a)da
$$

To summarise, global expected welfare is given by:

$$
w = \begin{cases} w_1 \text{ if } k \in \left[0, \frac{(a-bv)}{RA}\right] \\ w_2 \text{ if } k \in \left[\frac{(\underline{a}-bv)}{RA}, \frac{(\underline{a}-bc_0)}{(1+bc_1)RA}\right] \\ w_3 \text{ if } k \in \left[\frac{(\underline{a}-bc_0)}{(1+bc_1)RA}, \frac{(\overline{a}-bv)}{RA}\right] \\ w_4 \text{ if } k \in \left[\frac{(\overline{a}-bv)}{RA}, \frac{(\overline{a}-bc_0)}{(1+bc_1)RA}\right] \\ w_5 \text{ if } k \in \left[\frac{(\overline{a}-bc_0)}{(1+bc_1)RA}, +\infty\right] \end{cases}
$$

w can not be globally differentiated since it is made-up of five distinct functions. In section 3.2.1, we derive some properties of w that allow us to draw welfare results. We also simulate the model by giving values for all parameters of the model, save for the endogenous variable k, to obtain the level of capacity that maximizes social welfare. We thus obtain the values of k that maximize w for each of the five segments. Consequently, the optimal level of capacity is the one that gives the highest value to w.[13]

Note that w yields a lower bound of the level of socially optimal capacity. For instance, we have assumed that the benevolent social maximiser is risk-neutral, a condition unlikely to be met in practice. Also, there may be non-linearities in the social cost associated with repeated black-outs. Organised crime could potentially exploit a situation of capacity shortages. Moreover, electricity shortages could produce dynamic costs, particularly in countries or regions that heavily rely on Foreign Direct Investment (a condition that is clearly met in the Spanish case). We have not modelled these (very real) effects, as we would have to rely on adhoc judgements. Nonetheless, it is certainly the case that the true level of socially optimal capacity lies above the one we obtain algebraically.

Before turning to the construction of the simulation exercise, we first present the results pertaining to a decentralised outcome.

3.2. Capacity investments in decentralised markets

In liberalised markets, firms make their profit maximizing long-run investment decisions taking into account regulations (e.g. environmental) as well as the structure of the market and behaviour of its participants. As mentioned above, we

[13] Within each region, there may be an interior solution, or a corner one if the function is strictly increasing or decreasing in that segment. Consequently, we also compute the value of W_i at the extremes of each interval, and compare the k thus obtained with that the k derived from first order conditions.

assume that firms behave competitively during the second stage of our game, that is when they make bids to supply energy to the pool. A number of papers have analysed strategic bidding in the pool, taking capacity as fixed (see Introduction). Our interest is in long-run investment in capacity, and we therefore model competition during the second stage in the simplest manner. More importantly, our qualitative results are barely sensitive to the nature of competition in the second stage of the game.

We do not consider situations in which a generator may be unwilling to increase prices along the downward sloped segment of the demand curve, since we assume that market is totally "deregulated" in the sense that there are neither price caps, nor vertical integration with suppliers that face regulated final rates. This may be viewed as an approximation of the current Californian situation, and to the Spanish market in 2007 when all regulated prices are supposed to disappear.[14]

3.2.1. Private investment decisions in the absence of regulatory intervention

We index firms in the industry with $i \in [1, n]$ where n is a finite number, and we assume that firms are symmetric. We proceed to define the profits obtained by firm i that owns capacity k_i, while the rest of the industry's aggregate capacity is k_{-i}. The analysis follows the steps of the previous section, that is we distinguish three cases and five regions.

Case 1: $a \in [\underline{a}, a_c]$ Under this case, individual profits are given by:

$$\Pi_i^1 = p_c \frac{q_c}{n} - \int_0^{\frac{q_c}{n}} (c_0 + nc_1 q) dq - fk_i$$

This corresponds to the profits obtained from generation.[15]

Case 2: $a \in [a_c, a_m]$ Profits are given by:

$$\Pi_i^2 = \left(\frac{a - RAk}{b} \right) Ak_i - \int_0^{Ak_i} (c_0 + nc_1 q) dq - fk_i$$

Case 3: $a \in [a_m, \overline{a}]$, then profits are:

[14] We refer to an "approximation", since Investor Owned Utilities (IOUs) still possess some generation capacity. However, given the level of demand, this capacity is infra-marginal most of the time, i.e. IOUs are unable to set the pool's marginal price.

[15] Note that $c_0 + nc_1 q_i$ represent firm marginal costs, while $c_0 + c_1 q$ is the industry's marginal cost.

$$\Pi_i^3 = vAk_i - \int_0^{Ak_i} (c_0 + nc_1 q)dq - fk_i$$

As with the social welfare analysis, we have to define five regions:

Region 1: $k \in \left[0, \frac{(\underline{a}-bv)}{RA}\right]$Case 3 occurs with probability one, so that expected profits for a generator are given by:

$$\pi_{i,1} = \int_{\underline{a}}^{\overline{a}} \left[vAk_i - \int_0^{Ak_i} (c_0 + nc_1 q)dq - fk_i \right] f(a)da$$

Region 2: $k \in \left[\frac{(\underline{a}-bv)}{RA}, \frac{(\underline{a}-bc_0)}{(1+bc_1)RA}\right]$

$$\begin{aligned} \pi_{i,2} &= \int_{\underline{a}}^{a_m} \left[\left(\frac{a - RAk}{b} \right) Ak_i - \int_0^{Ak_i} (c_0 + nc_1 q)dq - fk_i \right] f(a)da \\ &\quad + \int_{a_m}^{\overline{a}} \left[vAk_i - \int_0^{Ak_i} (c_0 + nc_1 q)dq - fk_i \right] f(a)da \end{aligned}$$

Region 3: $k \in \left[\frac{(\underline{a}-bc_0)}{(1+bc_1)RA}, \frac{(\overline{a}-bv)}{RA}\right]$

$$\begin{aligned} \pi_{i,3} &= \int_{\underline{a}}^{a_c} \left[p_c \frac{q_c}{n} - \int_0^{\frac{q_c}{n}} (c_0 + nc_1 q)dq - fk_i \right] f(a)da \\ &\quad + \int_{a_c}^{a_m} \left[\left(\frac{a - RAk}{b} \right) Ak_i - \int_0^{Ak_i} (c_0 + nc_1 q)dq - fk_i \right] f(a)da \\ &\quad + \int_{a_m}^{\overline{a}} \left[vAk_i - \int_0^{Ak_i} (c_0 + nc_1 q)dq - fk_i \right] f(a)da \end{aligned}$$

Region 4: $k \in \left[\frac{(\overline{a}-bv)}{RA}, \frac{(\overline{a}-bc_0)}{(1+bc_1)RA}\right]$

$$\begin{aligned} \pi_{i,4} &= \int_{\underline{a}}^{a_c} \left[p_c \frac{q_c}{n} - \int_0^{\frac{q_c}{n}} (c_0 + nc_1 q)dq - fk_i \right] f(a)da \\ &\quad + \int_{a_c}^{\overline{a}} \left[\left(\frac{a - RAk}{b} \right) Ak_i - \int_0^{Ak_i} (c_0 + nc_1 q)dq - fk_i \right] f(a)da \end{aligned}$$

Region 5: $k \in \left[\frac{(\overline{a}-bc_0)}{(1+bc_1)RA}, +\infty\right]$

$$\pi_{i,5} = \int_{\underline{a}}^{\overline{a}} \left[p_c \frac{q_c}{n} - \int_0^{\frac{q_c}{n}} (c_0 + nc_1 q) dq - f k_i \right] f(a) da$$

Consequently, the expected profit is made-up of the five functions defined above.

Under this scenario, it is intuitively obvious that there will be under-investment in capacity from a social point of view. This reason is simple: given that the number of firms is finite and the second stage of the game is characterised by perfect competition, the only route available to obtain extraordinary profits is to reduce the level of installed capacity. With a low level of capacity, generators are able to charge high prices during many hours of the year, even if they behave competitively during the second stage of the game.

The key point is that generators will not install capacity even if spot prices amply allow them to recover their fixed and variable costs. The reason is that in this oligopolistic context, every extra unit of capacity installed depresses prices during many hours of the year. This allows us to state Proposition 1:

Proposition 1: In the absence of regulatory intervention and with a finite number of firms (generators), a decentralised outcome unambiguously yields a socially sub-optimal level of installed capacity.

Proof: See appendix.

The proof of proposition 1 yields an additional result: the socially optimal level of capacity does not cover peak demand.

Corollary 1: The socially optimal level of capacity lies below the capacity that would cover expected peak demand.

We briefly return to the issue of competitive conditions in the spot market. The simulations under alternative competitive scenarios during the second stage indicate that the degree of spot market competition is of no qualitative importance under any reasonable parameter constellation. The intuition for this result runs as follows. Under competitive conditions $p = MC$, which implies that firms can only increase total profits by reducing quantities supplied. By contrast, colluding firms would set monopoly prices ($p_{monopoly} = v$) all year long. Consequently, collusion or other type of anti-competitive behaviour would reduce firms incentives to undercut capacity. However, this would be antimonic to the objectives pursued by deregulation, since more investment in capacity would be obtained at the cost of high prices and monopoly power.

4. Simulations: data and results

In order to further illustrate our results and to quantify the difference between the social and private outcomes in real-world economies, we have carried out simulations. As will be seen below, the data necessary to simulate the model is readily available. Consequently, our model could be easily tailored to the specific situation of other electricity systems.

4.1. Data and methodology

To measure demand variability, we obtained data on the Spanish hourly distribution of demand for 1999. This information is made available by the market operator (OMEL), and can be downloaded from the web. Hourly demand was updated with projections of future demand up to 2010 provided by the system operator, Red Eléctrica de España (REE).[16] Given the time required to build new capacity and the fact that long term demand projections are less precise, we focus on the year 2005 in our simulations.

As mentioned before, we assume that renewable energy and hydro generation capacity are exogenously determined, and that no new dams will be built in the foreseeable future. We net out hydro capacity in the following manner. First, we use historical data on rainfall since the last dam started producing energy to obtain the average, minimum, and maximum hydro capacity in a year. Since electricity from hydro sources can be "stored", hydro power will be sold when demand (and prices) are highest. We apply the usual peak-load shaving technique, that is hydro capacity is distributed to serve the hours of peak demand till the entire hydro capacity is used-up. In doing so, we took into account the relevant technical restrictions such as the fact that there are maximum and minimum of MWh that can be despatched during an hour. Combining with the predictions of demand provided by REE, we obtained the probabilistic distribution of hourly net thermic demand for the year 2005. Adjusting a lognormal distribution to the data yielded an average thermic demand of 22.7 GWh with a standard deviation of 2.302 GWh.

Our demand schedule is made-up of two segments, as depicted in Figure 1. For our central simulations, we have imposed a value of v of 3005 euros per MWh. We approximated this value via two different routes. The first is to use the values that

[16] The projections provided by REE oscillate within some interval; we have used the central projections.

exist in countries that specifically incorporate the "value of lost load" (VOLL) in their regulatory systems. The second is to compute average annual industrial and service GDP per MWh. The results are included in Tables 1 and 2. As can be seen from these Tables, our cap is very close to the VOLL that has been used in England and Wales. Moreover, all the different approximations yield very similar values. We further assumed that in the presence of excess demand, firms would be able to increase prices up to v. Alternative scenarios where firms' prices are capped below v yield results of the same order of magnitude for reasonable values of the cap (e.g. for any cap above 200 euros per MWh).

The slope of the downward segment accounts for the fact that some customers are willing to be interrupted if prices sky rocket. We assumed that pump storage would be at its maximum when prices are zero, and would disappear when prices reach their maximum. We then augmented the demand for hydro pumping with that derived from large interruptible consumers.[17] We assume that in an hour of average demand, customers would be able to reduce their consumption within the hour by 5% if prices were to reach their maximum. Finally, we assume that the slope of our demand schedule remains constant throughout the year.

On the cost side, we have used the technology mix existing in Spain. As is well known, electricity generation in Spain and California have much in common in terms of technological mix. Currently installed generating capacity gives rise to the upward slope of the marginal cost schedule. Accordingly, we approximate the industry's cost schedule with an increasing linear marginal cost function. The intercept of the marginal cost curve with the vertical axis is 7.99 euros per MWh, which corresponds to the marginal cost of a nuclear plant.[18] Marginal cost then increases linearly up to the level of socially optimal capacity. At this level, marginal cost is equal to 29.15 euros per MWh, which corresponds to the marginal cost of gas-fuel plants. Above installed capacity, marginal costs are assumed to tend towards infinity, reflecting the fact that capacity is fixed in the short-run. Further, we assume that all new investment will consist of combined gas cycle, which is deemed to be the most efficient today.[19] We further assume that the

[17] Under current Spanish regulation, a "large consumer" is defined as one with an aggregate annual demand greater or equal to 1GWh.

[18] Recall that we have netted-out hydro capacity. In Spain, hydro energy is the cheapest to produce, so that the cost curve depicted in Figure 1 has its first segment (that corresponds to hydro costs) cut-off.

[19] In the Spanish case, the technology that is currently dominant is generation powered by a combined gas cycle. All new capacity (recently installed and projected) use this technology. An example of technology "inherited" from the regulated period are nuclear plants, that are

plant is despatched 7500 hours per year on average during 20 years, which implies $f = 6.53$ euros per MWh produced.

We also take into account technical restrictions pertaining to the system's operation as well as temporary outages resulting from maintenance work. The existing technology mix in Spain yields an average availability ratio $A = 92.5\%$.[20] The reserve margin r maintained by the system operator is given by the following formula: $r = 3\sqrt{k}$, where k is total installed capacity measured in MWh.

Our last parametrisation pertains to the number of firms active in the Spanish market. Currently, there are four generators that compete to supply energy in the pool. However, they are of different sizes, with the largest accounting for about 40% of generation, and the smallest a bit more than 5%. For our simulations we used the benchmark of $n = 4$. All our findings carry over to a situation with an alternative (finite) number of firms.

Finally, it should be noted that our model yields a lower bound of what true social welfare probably consists of. Concretely, we have assumed no risk aversion on the part of the regulator, a condition that is most unlikely to be met in the real world.[21] Also, we have not factored in all possible social costs derived from black-outs. We believe that these additional costs are very real; however, we are unable to quantify them in a consistent manner.[22] In what follows, we report the results for the parameter values reported above. We carried out a sensitivity analysis, and our results are qualitatively robust to sensible changes in parameter values.

characterised by low marginal costs.

[20] This parameter results from aggregating technologies with different availability ratios. All the results are robust to changes in the parameter.

[21] It is indeed plausible to think that the agents that are the counterparts of our social planner are somewhat risk averse. Political parties may lost elections as a result of a power crisis. Senior civil servants may see their career prospects severely impaired. Industry regulators may be publicly blamed for energy shortages and sky rocketing prices.

[22] A group of distinguished economists that includes a Nobel laureate, former public officials, and consultants have produced a "Manifesto on the California Electricity Crisis" (21/1/2001). In the Preamble they state that: "Rolling blackouts impose tremendous social and economic costs on California society and threaten to wreck its economy. The situation is very serious and endangers the livelihoods of many citizens in and out of state".

4.2. Simulation results

Our main results are reported in Table 3. The social optimum yields a level of 30731 MW of conventional thermic capacity for the year 2005.[23] As can be seen, the socially optimal level of capacity does not cover peak demand, as 10 hours experience shortages. Given that our social optimum is very conservative (see comments above), we obtained the level of capacity that would meet peak demand, taking into the system's technical restrictions.[24] Actually, the difference between these two levels of capacity is quite small. The socially optimal level of capacity would only result in 10 hours of expected excess demand and in a tiny amount of unsatisfied demand at marginal cost pricing.

The most striking results pertain to the deregulated scenario. Our simulations clearly indicate that a decentralised outcome would be sub-optimal. Generators would only install 74% of the conventional thermic capacity that maximises social welfare (and 70% of the capacity necessary to meet peak demand). This would imply rolling black-outs almost every single hour of the year, and would result in a large proportion of demand being unsatisfied. As such, this result is not surprising. In our oligopolistic setting, a profit maximising strategy consists in maintaining capacity shortages for the purpose of increasing prices along the (inelastic) demand schedule.

Insert Table 3 about here

Socially sub-optimal outcomes are common in oligopolistic industries. What is special about the electricity industry is the large gap that exists between a deregulated outcome and a situation where social welfare is maximised.

5. Conclusions

We model the electricity industry as an oligopoly, and identify the potential welfare effects of deregulation. Previous models in the literature have focused their

[23] This corresponds to a total of 59500 MW of installed capacity if we include hydro capacity and renewable energy sources. The latter operate under very specific rules, and are not exposed to market signals (i.e., renewable capacity is set exogenously). In addition, availability (which is typically very low) is beyond the firms' control, as it depends on climatological factors such as sunshine and winds.

[24] Note that the social optimum that our model generates is very close to the simple rule that involves covering peak demand. Our results are almost identical to the estimations of necessary capacity for 2005 made by Red Eléctrica de España (the system operator). This suggests that our model does a good job of reproducing the real-world situation found in Spain.

attention on short run competition in decentralised electricity pools, always assuming that there is excess capacity. By contrast, we center our attention on capacity choices in the context of a two-stage game, in which firms first chose capacity and then compete in the product market. We explicitly take into account the industry's idiosyncracies such as the level of uncertainty surrounding supply and demand, the near-impossibility to store electricity, as well as the technical restrictions that characterise electricity generation.

Our main finding is that a deregulated market will result in under-investment in generation capacity. We prove this result analytically, and the use of simulations indicate that potential welfare costs are very large. Inadequate generation capacity is the main culprit behind California's woes, and it could possibly become an issue in other deregulated electricity systems, such as the Spanish one.

Since short-run competition is easier to monitor for the regulator, we believe that capacity choices may become the main instrument through which generators will attempt to exercise market power. This suggests that there is a need for some regulatory mechanism that provides the right incentives to install a socially desirable level of capacity. Overall, our findings indicate that blind faith in market mechanisms will not yield a satisfactory outcome from a welfare point of view.

References

[1] Borenstein, S., Bushnell, Stoft, S., (2000), "The Competitive Effects of Transmission Capacity in a Deregulated Electricity Market", *Rand Journal of Economics*, Vol. 31 (2), pp. 294-325.

[2] Borenstein, S., Bushnell, J., Wolak, F., (2000), "Diagnosing Market Power in California's Restructured Wholesale Electricity Market", *NBER* Working Paper No 7868.

[3] Borenstein, S., Bushnell, J., (1999), "An Empirical Analysis of the Potential for Market Power in California's Electricity Market", *Journal of Industrial Economics*, Vol. 47, pp.

[4] García-Diaz, A., Marín, P.L., 2000, "Strategic Bidding in Electricity Pools with Short-Lived Bids: An Application to the Spanish Market", *CEPR* Discussion Paper, 2567.

[5] Green, R., Newbery D., (1992), "Competition in the British Electricity Spot Market," *Journal of Political Economy*, Vol. 100, pp. 929-953.

[6] Joskow, P., Kahn, E., (2001), "A Quantitative Analysis of Pricing Behavior in California's Wholesale Electricity Market During Summer 2000", *NBER* Working Paper No 8157.

[7] Puller, S., (2000), "Pricing and Firm Conduct in California's Deregulated Electricity Market", *POWER* Working Paper PWP-080, University of California Energy Institute.

[8] Solow, R., *et al.*, (2001), "Manifesto on the California Electricity Crisis", available at http://haas.berkely.edu/news/california_electricity_crisis.html

[9] von der Fehr, N-H., Harbord, D.C., (1993), "Spot Market Competition in the UK Electricity Industry", *The Economic Journal*, Vol. 103, pp. 531-546.

[10] von der Fehr, N-H., Harbord, D.C., (1997), "Capacity investment and competition in decentralized electricity market", *Department of Economics Working Paper*, University of Oslo.

[11] Wolak, F.A., Patrick, R., (1997), "The Impact of Market Rules and Market Structure on the Price Determination Process in the England and Wales Electricity Market", *POWER* Working Paper PWP-047, University of California Energy Institute.

[12] Wolfram, C.D., (1998),"Strategic Bidding in a Multi-Unit Auction: An Empirical Analysis of Bids to Supply Electricity in England and Wales", *Rand Journal of Economics*, Vol. 29(4), pp. 703-725.

[13] Wolfram, C.D., (1999), "Measuring Duopoly Power in the British Electricity Spot Market", *American Economic Review*, Vol. 89(4), pp. 805-826.

Web Sites:

[14] Comisión Nacional de la Energía: www.cne.es

[15] Financial Times: www.ft.com

[16] Los Angeles Time: www.latimes.com

[17] OMEL (Spanish market operator): www.omel.es

[18] Red Eléctrica de España (REE, Grid Operator): www.ree.es

[19] Yorkshire Electricity: www.yeg.co.uk/business/electricity

Appendix

The intuition of the proof is straightforward. Both the social planner and a private firm optimise by equalising marginal revenue to marginal cost. The latter is the same in both expressions, while the marginal revenue of the social planner is always larger than that of private firms. This is because it includes both producer and consumer surplus, while firm's individual revenue function only includes their individual producer surplus.

Proof: Throughout the proof, we assume that the number of firms (n) is finite. We proceed by steps:

i) The marginal social welfare function in region 1 is given by:

$$\frac{\partial w_1}{\partial k} = [(v - c_0)A - c_1 A^2 k - f] \int_{\underline{a}}^{\bar{a}} f(a)da \tag{A.1.1.}$$

and the marginal individual profit function is given by:

$$\frac{\partial \pi_i}{\partial k} = [(v - c_0)A - c_1 A^2 k_i - f] \int_{\underline{a}}^{\bar{a}} f(a)da \tag{A.1.2.}$$

Consequently, both functions are increasing in k in region 1 iff $(v - c_0)A - c_1 A^2 k_i - f > 0$. This implies that if the unit cost of capacity (f) is not prohibitive, the social optimum and the decentralised outcome will never be located in region 1.

ii) The marginal social welfare function and the marginal individual profit function in region 5 are given by:

$$\frac{\partial w_1}{\partial k} = \frac{\partial \pi_i}{\partial k_i} = \int_{\underline{a}}^{\bar{a}} [-f]\, f(a)da < 0 \tag{A.2.1.}$$

Consequently, these functions are always decreasing in k. Therefore, the social optimum and the decentralised outcome will never be located in the region 5.

Thus, both the social optimum and the decentralised outcome can only be located in regions 2, 3 or 4. We proceed to show that the decentralised outcome always yields a lower level of installed capacity compared to the social optimum in these three regions.

iii) Suppose that the decentralised outcome is located in region 2. Let the level of capacity chosen by the social planner be given by: $k = nk_i$ (where k_i is the per firm level of capacity chosen by the planner). The social welfare function can thus be written as:

$$\begin{aligned} w_2 &= \int_{\underline{a}}^{a_m} \left[\int_{\frac{a-RAnk_i}{b}}^{v} (a-bp)dp \right] f(a)da \qquad \text{(A.3.1.)} \\ &+ \int_{\underline{a}}^{a_m} \left[(\frac{a-RAnk_i}{b})Ank_i - \int_0^{Ank_i} (c_0 + c_1 q)dq - fnk_i \right] f(a)da \\ &+ \int_{a_m}^{\overline{a}} \left[\left[vAnk_i - \int_0^{Ank_i} (c_0 + c_1 q)dq \right] - fnk_i \right] f(a)da \end{aligned}$$

The expression given above can be re-written as:

$$\begin{aligned} w_2 &= \int_{\underline{a}}^{a_m} \left[\int_{\frac{a-RAnk_i}{b}}^{v} (a-bp)dp \right] f(a)da \qquad \text{(A.3.2.)} \\ &+n \int_{\underline{a}}^{a_m} \left[(\frac{a-RAnk_i}{b})Ak_i - \int_0^{Ak_i} (c_0 + nc_1 q)dq - fk_i \right] f(a)da \\ &+n \int_{a_m}^{\overline{a}} \left[\left[vAk_i - \int_0^{Ak_i} (c_0 + nc_1 q)dq \right] - fk_i \right] f(a)da \end{aligned}$$

Consequently:

$$w_2 = \int_{\underline{a}}^{a_m} \left[\int_{\frac{a-RAnk_i}{b}}^{v} (a-bp)dp \right] f(a)da + n\pi_{i,2} \qquad \text{(A.3.3.)}$$

Let $k^s = \sum_{i=1}^{n} k_i^s$ and $k^d = \sum_{i=1}^{n} k_i^d$ be the capacity levels associated with the social optimum and the decentralized outcome respectively.

If we can show that $\left.\frac{dw_2}{dk_i}\right|_{k_i=k_i^d} > 0$, we have that $k_i^s > k_i^d$, and by consequence, $k^s > k^d$.

Using A.3.3., we can re-write $\left.\frac{dw_2}{dk_i}\right|_{k_i=k_i^d}$ as:

$$\begin{aligned} \left.\frac{dw_2}{dk_i}\right|_{k_i=k_i^d} &= \frac{d}{dk_i} \left[\int_{\underline{a}}^{a_m} \left[\int_{\frac{a-RAnk_i}{b}}^{v} (a-bp)dp \right] f(a)da \right]_{k_i=k_i^d} \qquad \text{(A.3.4.)} \\ &+n \left[\frac{d\pi_{i,2}}{dk_i} \right]_{k_i=k_i^d} \end{aligned}$$

The second term is equal to zero given that k_i^d solves the first-order condition of the decentralised problem. We thus have to demonstrate that the first term is positive.

Let $z(a,k_i)$ denote the integral $\int_{\frac{a-RAnk_i}{b}}^{v}(a-bp)dp$. Note that $z(a,k_i)>0$ and $\frac{dz(a,k_i)}{dk_i}>0$. Recall that $a_m = bv + RAnk_i$. We thus have that:

$$\frac{d}{dk_i}\left[\int_{\underline{a}}^{a_m} z(a,k_i)f(a)da\right]\Bigg|_{k_i=k_i^d} = z(a_m,k_i)\ RAn\ f(a_m)|_{k_i=k_i^d} + \int_{\underline{a}}^{a_m}\frac{dz(a,k_i)}{dk_i}f(a)da\Bigg|_{k_i=k_i^d} \quad \text{(A.3.5.)}$$

which is positive, and this implies that $k_i^s > k_i^d$ in region 2.

iv) Suppose that the decentralised outcome is located in region 3. As before, we use the relation $k = nk_i$, to re-write the social welfare in region 3 as:

$$\begin{aligned} w_3 = & \int_{\underline{a}}^{a_c}\left[\int_{p_c}^{v}(a-bp)dp\right]f(a)da \qquad \text{(A.4.1.)}\\ & + \int_{a_c}^{a_m}\left[\int_{\frac{a-RAnk_i}{b}}^{v}(a-bp)dp\right]f(a)da \\ & + \int_{\underline{a}}^{a_c}\left[\left[p_cq_c - \int_0^{q_c}(c_0+c_1q)dq\right]-fnk_i\right]f(a)da \\ & + \int_{a_c}^{a_m}\left[(\frac{a-RAnk_i}{b})Ank_i - \int_0^{Ank_i}(c_0+c_1q)dq - fnk_i\right]f(a)da \\ & + \int_{a_m}^{\overline{a}}\left[\left[vAnk_i - \int_0^{Ank_i}(c_0+c_1q)dq\right]-fnk_i\right]f(a)da \end{aligned}$$

Which is equivalent to:

$$\begin{aligned} w_3 = & \int_{\underline{a}}^{a_c}\left[\int_{p_c}^{v}(a-bp)dp\right]f(a)da \qquad \text{(A.4.2.)}\\ & + \int_{a_c}^{a_m}\left[\int_{\frac{a-RAnk_i}{b}}^{v}(a-bp)dp\right]f(a)da \\ & + n\int_{\underline{a}}^{a_c}\left[\left[\frac{p_cq_c}{n} - \int_0^{\frac{q_c}{n}}(c_0+nc_1q)dq\right]-fk_i\right]f(a)da \end{aligned}$$

$$+n\int_{a_c}^{a_m}\left[(\frac{a-RAnk_i}{b})Ak_i-\int_0^{Ak_i}(c_0+nc_1q)dq-fk_i\right]f(a)da$$

$$+n\int_{a_m}^{\overline{a}}\left[\left[vAk_i-\int_0^{Ak_i}(c_0+nc_1q)dq\right]-fk_i\right]f(a)da$$

Thus:

$$\begin{aligned} w_3 &= \int_{\underline{a}}^{a_c}\left[\int_{p_c}^{v}(a-bp)dp\right]f(a)da \\ &+\int_{a_c}^{a_m}\left[\int_{\frac{a-RAnk_i}{b}}^{v}(a-bp)dp\right]f(a)da+n\pi_{i,3} \end{aligned} \quad \text{(A.4.3.)}$$

If we can show that $\left.\frac{dw_3}{dk_i}\right|_{k_i=k_i^d}>0$, we have that $k_i^s>k_i^d$, and by consequence, $k^s>k^d$ in region 3.

$\left.\frac{dw_3}{dk_i}\right|_{k_i=k^d}$ can be written as:

$$\begin{aligned} \left.\frac{dw_3}{dk_i}\right|_{k_i=k_i^d} &= \frac{d}{dk_i}\left[\begin{array}{c}\int_{\underline{a}}^{a_c}\left[\int_{p_c}^{v}(a-bp)dp\right]f(a)da \\ +\int_{a_c}^{a_m}\left[\int_{\frac{a-RAnk_i}{b}}^{v}(a-bp)dp\right]f(a)da\end{array}\right]_{k_i=k_i^d} \\ &+n\left[\frac{d\pi_{i,3}}{dk_i}\right]_{k_i=k_i^d} \end{aligned} \quad \text{(A.4.4.)}$$

The second term is equal to zero given that k_i^d solves the first-order condition of the decentralised problem. We thus have to demonstrate that the first term is positive.

Let the integral $\int_{p_c}^{v}(a-bp)dp$ be denoted by $z_1(a)$, and $\int_{\frac{a-RAnk_i}{b}}^{v}(a-bp)dp$ by $z_2(a,k_i)$. Note that $z_1(a)>0$, $z_2(a,k_i)>0$ and $\frac{dz_2(a,k_i)}{dk_i}>0$. Also recall that $a_c=bc_0+(1+bc_1)RAnk_i$ and $a_m=bv+RAnk_i$.

We thus have that:

$$\begin{aligned} &\left.\frac{d}{dk_i}\left[\int_{\underline{a}}^{a_c}z_1(a)f(a)da\right]\right|_{k_i=k_i^d}+\left.\frac{d}{dk_i}\left[\int_{a_c}^{a_m}z_2(a,k_i)f(a)da\right]\right|_{k_i=k_i^d} \quad \text{(A.4.5.)} \\ =\ & z_1(a_c)\ (1+bc_1)RAn\ f(a_c)|_{k_i=k_i^d}+\left.\int_{a_c}^{a_m}\frac{dz_2(a,k_i)}{dk_i}f(a)da\right|_{k_i=k_i^d} \\ &+z_2(a_m,k_i)\ RAn\ f(a_m)|_{k_i=k_i^d}-z_2(a_c,k_i)\ (1+bc_1)RAn\ f(a_c)|_{k_i=k_i^d} \end{aligned}$$

$$= \int_{a_c}^{a_m} \frac{dz_2(a,k_i)}{dk_i} f(a)da\Big|_{k_i=k_i^d} + z_2(a_m,k_i)\ RAn\ f(a_m)|_{k_i=k_i^d}$$
$$+ [z_1(a_c) - z_2(a_c,k_i)]\ (1+bc_1)RAn\ f(a_c)|_{k_i=k_i^d}$$

The first two terms are positive. With respect to the third term, given the definition of de a_c, we have that $p_c(a_c) = \frac{a_c - RAnk_i}{b}$, and consequently, $z_1(a_c) = z_2(a_c,k_i)$. Thus, $\frac{dw_3}{dk_i}\Big|_{k_i=k_i^d} > 0$ which implies that $k^s > k^d$ in region 3.

v) Suppose that private outcome is located in region 4. As before, we use the relation $k = nk_i$, to write the social welfare in region 4 as:

$$\begin{aligned} w_4 &= \int_{\underline{a}}^{a_c} \left[\int_{p_c}^{v} (a-bp)dp\right] f(a)da \\ &+ \int_{a_c}^{\overline{a}} \left[\int_{\frac{a-RAnk_i}{b}}^{v} (a-bp)dp\right] f(a)da \\ &+ \int_{\underline{a}}^{a_c} \left[\left[p_c q_c - \int_0^{q_c} (c_0 + c_1 q)dq\right] - fnk_i\right] f(a)da \\ &+ \int_{a_c}^{\overline{a}} \left[(\frac{a-RAnk_i}{b})Ank_i - \int_0^{Ank_i} (c_0 + c_1 q)dq - fnk_i\right] f(a)da. \end{aligned} \tag{A.5.1.}$$

Which can be re-written as:

$$\begin{aligned} w_4 &= \int_{\underline{a}}^{a_c} \left[\int_{p_c}^{v} (a-bp)dp\right] f(a)da \\ &+ \int_{a_c}^{\overline{a}} \left[\int_{\frac{a-RAnk_i}{b}}^{v} (a-bp)dp\right] f(a)da \\ &+n \int_{\underline{a}}^{a_c} \left[\left[\frac{p_c q_c}{n} - \int_0^{\frac{q_c}{n}} (c_0 + nc_1 q)dq\right] - fk_i\right] f(a)da \\ &+n \int_{a_c}^{\overline{a}} \left[(\frac{a-RAnk_i}{b})Ak_i - \int_0^{Ak_i} (c_0 + nc_1 q)dq - fk_i\right] f(a)da. \end{aligned} \tag{A.5.2.}$$

We thus have that:

$$\begin{aligned} w_4 &= \int_{\underline{a}}^{a_c} \left[\int_{p_c}^{v} (a-bp)dp\right] f(a)da \\ &+ \int_{a_c}^{\overline{a}} \left[\int_{\frac{a-RAnk_i}{b}}^{v} (a-bp)dp\right] f(a)da + n\pi_{i,4} \end{aligned} \tag{A.5.3.}$$

We need to show that $\left.\frac{dw_4}{dk_i}\right|_{k_i=k_i^d} > 0$ to conclude that $k^s > k^d$ in region 4.

We calculate $\left.\frac{dw_4}{dk_i}\right|_{k_i=k^d}$ as:

$$\left.\frac{dw_4}{dk_i}\right|_{k_i=k_i^d} = \frac{d}{dk_i}\left[\begin{array}{c} \int_{\underline{a}}^{a_c}\left[\int_{p_c}^{v}(a-bp)dp\right]f(a)da \\ +\int_{a_c}^{\overline{a}}\left[\int_{\frac{a-RAnk_i}{b}}^{v}(a-bp)dp\right]f(a)da \end{array}\right]_{k_i=k_i^d} + n\left[\frac{d\pi_{i,3}}{dk_i}\right]_{k_i=k_i^d} \qquad \text{(A.5.4.)}$$

The second term is equal to zero given that k_i^d solves the first-order condition of the decentralised problem. We thus have to demonstrate that the first term is positive.

Denote the integral $\int_{p_c}^{v}(a-bp)dp$ as $z_1(a)$ and $\int_{\frac{a-RAnk_i}{b}}^{v}(a-bp)dp$ as $z_2(a,k_i)$. Note that $z_1(a) > 0$, $z_2(a,k_i) > 0$ and $\frac{dz_2(a,k_i)}{dk_i} > 0$. Recall that $a_c = bc_0 + (1 + bc_1)RAnk_i$.

We thus have that:

$$\begin{aligned} & \left.\frac{d}{dk_i}\left[\int_{\underline{a}}^{a_c} z_1(a)f(a)da\right]\right|_{k_i=k_i^d} + \left.\frac{d}{dk_i}\left[\int_{a_c}^{\overline{a}} z_2(a,k_i)f(a)da\right]\right|_{k_i=k_i^d} \qquad \text{(A.5.5.)} \\ = \; & z_1(a_c)\;(1+bc_1)RAn\;f(a_c)|_{k_i=k_i^d} + \left.\int_{a_c}^{\overline{a}}\frac{dz_2(a,k_i)}{dk_i}f(a)da\right|_{k_i=k_i^d} \\ & - z_2(a_c,k_i)\;(1+bc_1)RAn\;f(a_c)|_{k_i=k_i^d} \\ = \; & \left.\int_{a_c}^{\overline{a}}\frac{dz_2(a,k_i)}{dk_i}f(a)da\right|_{k_i=k_i^d} \\ & + [z_1(a_c) - z_2(a_c,k_i)]\;(1+bc_1)RAn\;f(a_c)|_{k_i=k_i^d} \end{aligned}$$

By the definition of a_c we have that $p_c(a_c) = \frac{a_c - RAnk_i}{b}$, which implies $z_1(a_c) = z_2(a_c,k_i)$. Thus, $\left.\frac{dw_4}{dk_i}\right|_{k_i=k_i^d} > 0$, which implies that $k^s > k^d$ in region 4.

Q.E.D.

Table 1. Value of lost load in Euro/MWh

Country	1995	1996	1997	1998	1999
Victoria (Australia)	2788.42	2979.88	3273.83	2825.72	3028.38
England and Wales	2704.16	2810.39	3488.45	3716.42	3924.49

Exchange rate: IMF, June 2000.
Inflation ratesIMF, January 1997 and June 2000.
Note: Average exchange rate for the year

Table 2. Spanish GDP per MWh, 1999

Quarter	Agric. and fishing	Energy	Industrial	Construc.	Services	Other	Total
1^{st}	118.22	112.93	529.91	198.33	1688.84	277.43	2926.15
2^{nd}	159.57	117.80	595.00	236.74	1903.04	288.67	3300.82
3^{rd}	124.89	112.21	510.38	228.50	1809.89	269.37	3055.24
4^{th}	151.21	107.82	545.54	241.07	1756.58	263.78	3065.94

GDP from Boletín Trimestral de Coyuntura, INE, March 2000, nº 75.

Table 3. Optimal conventional thermic capacity.

	Conservative social optimum	Capacity to supply peak demand	Private optimum
Capacity (MW)	30730	32280	22620
No. Of hours with loss of load (per year)	10	≅0	7336
Unsatisfied demand (%)	0.0023%	≅0.00%	9.62%

Note: 1) These figures have been obtained applying a deterministic availability ratio, i.e. the average availability ratio applies to peak demand hours. A common industry practice is to use a stochastic availability ratio. If we were to apply a stochastic ratio, the number of hours with loss of load would be larger.

2) The figures pertain to conventional thermic capacity. Adding hydro capacity and energy produced under the regime that applies to renewable energy sources would yield the following values: 59500 (conservative social optimum), 61050 (capacity that meets peak demand), and 51390 (private outcome).

Figure 1

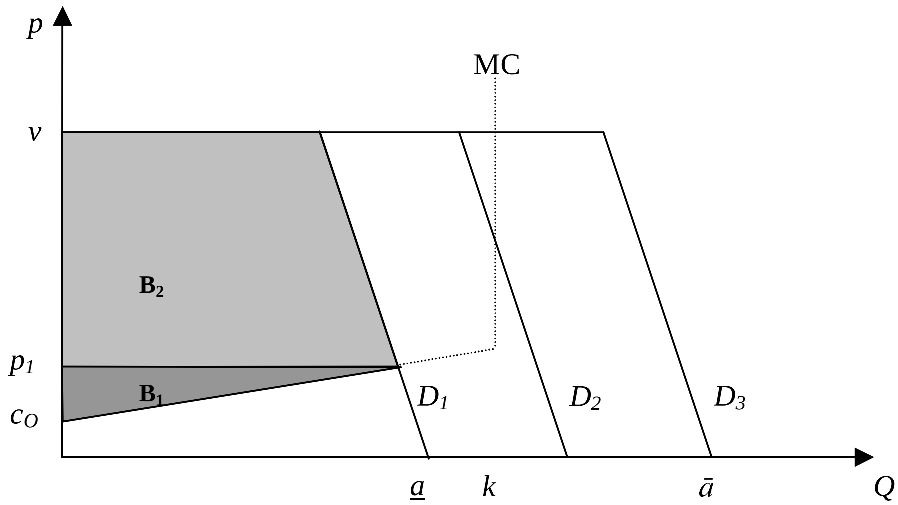

Figure 2