THREE ESSAYS ON THE ECONOMICS OF COMPETITIVE ELECTRICITY MARKET

OLAYINKA WILLIAMS

THESIS SUBMITTED TO THE
FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DOCTORATE IN PHILOSOPHY DEGREE IN ECONOMICS

DEPARTMENT OF ECONOMICS
SOCIAL SCIENCE
UNIVERSITY OF OTTAWA

ABSTRACT

In this thesis we present three essays on the economics of competitive electricity.

In chapter one we use a stochastic dynamic model to examine the economic impact of storing wind power in a hydroelectric reservoir. Using a stochastic dynamic model allows us to explicitly model the uncertainty of wind power in the objective function of the hydroelectric generator. We show that the amount of water released by the hydroelectric generator in any period is a decreasing function of the number of periods that the hydroelectric generator has to deplete its reservoir. Furthermore if the hydroelectric generator has a high enough number of periods to deplete the water in its reservoir it will be optimal for it not to release any water at the beginning.

In chapter two we present a model of operating reserve in a competitive electricity market. Our paper departs from other papers in several respects. First, we concentrate on the possibility of generator failure and not demand uncertainty. Second, we allow demand to be highly inelastic rather than perfectly inelastic like other papers do. This allows price spikes to occur when there is a major generator failure. To moderate price spikes, options for operating reserves can be purchased and exercised when price spikes occur. Third, we model uncertainty on the supply side and not on the demand side. Fourth we adopt a two price approach where one price is used to reserve capacity and the other price is the strike price paid when the options are exercised unlike other models which use the spot price as the strike price. Finally we explicitly model the demand and supply in the market. Using the concept of rational expectation we develop and prove the existence and uniqueness of a rational expectations equilibrium and analyse its characteristics. Furthermore we show that the competitive electricity market will provide more operating reserve capacity than is socially optimal.

In the chapter three we formalize a model of reliability for an electric grid when consumers' preferences for electricity consumption are private information. In our model we design an optimal blackout strategy for the regulator. The model demonstrates that an ex post efficient social choice function is truthfully implementable in Bayesian Nash equilibrium. It also yields both the optimal level of generation capacity investment and the second-best blackout program.

Chapter Three: RELIABILITY IN AN ELECTRIC GRID: A MECHANISM DESIGN APPROACH

ACKNOWLEDGEMENTS

I would like to thank God for guiding me and giving me the strength to complete the program. He saw me through the highs and lows and set me upon solid ground. Anytime I needed help and encouragement he made them available in large quantities. Without God none of this would have been possible. Without him I am nothing.

I would like to thank my wife who stuck with me. She was a source of comfort and encouragement through late nights of comprehensive exams and research brain storming. She listened to my crazy ideas and when I complained about the problems I was having with my research, she offered encouraging advice. She never lost faith in me even though sometimes I lost faith in myself.

I would also like to thank my supervisor Dr. Nguyen Van Quyen for teaching me about economic research and for his advice on various aspects of life. He helped me understand the essence of the economic research and develop the ability to see and examine unanswered important research questions. I would also like to thank the members of my thesis committee for taking time out to read my work and for their helpful comments.

Finally I will like to thank my mother and my father who believed in me through the years. Thank you for the financial support and your prayers. You believed in me and never gave up on me through it all.

THANK YOU LORD!!!!

INTRODUCTION TO THE THESIS

This thesis contains three essays on the economics of electricity. The first essay deals with the subject of storing wind power in a hydroelectric reservoir. The second essay presents an analysis of operating reserves in a competitive electricity market. The third essay addresses the public good nature of operating reserves from the perspective of mechanism design.

Wind turbines produce clean electricity at little or no marginal costs, and this makes the green energy they produce a very attractive source of electricity generation. In many electricity markets programs such as Renewable Portfolio Standards and Feed in Tariffs have been implemented to encourage the integration of wind turbines and other renewable energy sources. However, the disadvantage with wind turbines is that they are only able to produce electricity when the wind is blowing. Unfortunately, wind flows are erratic and unpredictable. Therefore, the power produced by wind farms is volatile and unreliable.

To overcome the volatile and unreliable nature of wind power, independent system operators could increase the purchase of ancillary services. However the cost of ancillary services purchased from generators with fast ramp capabilities are usually high and could wipe out the cost savings attributed to wind power production. In addition, the ramp up and ramp down activities of these producers could cause more greenhouse gases emissions. Another way of overcoming the volatility and unreliable nature of wind power is to store the power produced by wind turbines as water in hydroelectric reservoirs. Due to their fast ramp capabilities, hydroelectric generators are able to ramp up to replace wind power when there is no wind and ramp down when the wind blows. This process enables wind power producers to store the power they produce as water which could be released by the hydroelectric generator at a constant rate over a period of time. The storage of wind power will however distort the optimal release program of a hydroelectric generator making electricity production by hydroelectric generators uncertain. Although the research on wind power integration is extensive, the impact of wind power storage on the optimal water release program of hydroelectric generators has yet to be examined. The main objective of the first essay is to examine the optimal water release program of a hydroelectric generator that stores wind power in its reservoirs and sells power in competitive electricity market.

To maintain the stability of any power system, an instantaneous and continuous balance between supply and demand of electricity is required. When a large generating plant fails, the spot price of electricity must rise sharply to align the remaining generation capacity with demand, assuming that prices are flexible. In order to protect themselves against unexpected price spikes, consumers often purchase operating reserves that help alleviating the cost of meeting their electric needs during difficult times. The second essay presents a model of the operating reserves and the real-time energy markets. In the market for operating reserves, call options for operating reserves are bought and sold, and trading in the market for operating reserves takes place before that of the real-time energy market. The model that is formalized in this chapter uses the concept of rational expectations as the solution concept.

In an electric grid, a major failure of transmission lines and generating plants will cause load shedding if the remaining generation capacity is inadequate to meet demand. A large installed generation capacity will be able to meet demand and withstand serious failure of transmission lines and generating plants. However, generation capacity investment is costly, and thus it is not optimal to invest in generation capacity to prevent all blackouts. The problem faced by the regulator is to find the optimal reliability of the system, i.e., the optimal generation capacity investment and the optimal blackout program to carry out when a major failure of generating plants occurs. Because a large installed generation capacity lowers the probability, the duration, and the magnitude of the blackouts, every consumer benefits from a high level of installed generation capacity investment, and thus reliability has the nature of a public good, and markets cannot be relied upon to obtain the optimal level of reliability. To find the optimal level of reliability, the regulator has to know the preferences for electricity consumption of each consumer. However, the consumers' preferences for electricity consumption are private information, and no consumer has the incentive to reveal her preferences for electricity consumption truthfully unless the contributions she is required to make in financing part of the funds needed to finance the investment are well structured. The third chapter presents a model of operating reserves from the perspective of mechanism design to capture the public good nature of reliability. It is demonstrated that an ex post efficient social choice function is truthfully implementable in Bayesian Nash equilibrium. The model yields both the optimal level of generation capacity investment and the second-best blackout program.

CHAPTER ONE: THE INTEGRATION OF WIND POWER GENERATION WITH HYDROELECTRICITY IN AN ELECTRIC GRID

1. INTRODUCTION

Wind turbines produce electricity by harnessing the kinetic energy from the wind. As wind blows, the blades attached to wind turbines spin and rotate magnets within a coil of wire, and this causes electrons to flow. The flow of electrons is what we call electricity. At the present time, the capital cost of building wind farms is high; however, once built the production cost of the electricity is low. The low production costs make it an attractive source of electricity generation.

The major disadvantage of wind energy is its intermittent nature. Wind blows at random times during the day, and the production of electricity from wind turbines is dependent on the wind speeds that prevail at any point in time. Therefore, when there are strong winds, wind turbines produce a lot of electricity. When wind speeds are low, the power produced is low. Due to the intermittent nature of wind, fast wind speeds do not always coincide with high demand periods. Therefore, wind power might be low when electricity demand is high and high when electricity demand is low.

To solve the problem associated with the intermittency and volatility of wind flows, fast ramp generating capacity or fast ramp storage capacities could be used to store wind power and therefore balance demand and supply. Today, hydroelectric generators with reservoirs provide the only viable option to store wind power. Hydroelectric generators with reservoirs are attractive because they have fast ramp capabilities, which allow them to keep up with the changes wind power over time. In addition, water reservoirs allow hydroelectric plant to move electricity across time. When wind power is high, the amount of water released could be reduced and when wind power is low more water can be released. This ability to store power gives wind turbine generators a means of delivering wind power to the electricity market at a constant rate.

In today's market, the potential storage ability provided by hydroelectric generators has developed into a new business opportunity for hydroelectric generators, and it is now offered as a service to wind turbine generators in several electricity markets. For example, the Bonneville Power Administration (BPA) offers storage services to wind farms that would like to balance their power production. The BPA will keep the power produced by wind farms in its reservoir for 7 days after which the power is sold in the market at a constant level over a period of time. This helps wind plants to remove any uncertainty in the supply of electricity.[1]

With the renewed efforts at wind integration, research into the implications of wind integration has picked up. Benitez et al (2008) analyze the integration of wind turbines into the Alberta's electricity market. To conduct their analyses, they use a constrained optimization approach, which minimizes the total cost of generation subject to the constraints on the various types of generating plants. They find that the integration of wind power in a fossil fuel market will increase total system costs. In addition to this result, they also find that with more hydroelectric capacity, system costs will go down with wind integration, and conclude that wind integration is better suited in power systems with a significant amount of hydro capacity.

Belanger and Gagnon (2002) analyze the integration of wind into a power system made up of mainly hydroelectric generators. In their research, they carry out a simulation using data on electricity demand in Quebec, wind speeds, and hydraulic flow rates. They find that when wind power is added into a hydro system, reserve hydro capacity will be needed to back up wind power when wind power is not. They point out that, if available, the extra reserve capacity could be sold as an ancillary service in competitive markets. This indicates that there are hidden costs of storage

[1] Appendix 1, Integration of Wind Generation with Power Systems in Canada: Overview of Technical and Economic Impacts

when wind power is integrated into a hydro system. They conclude that wind power cannot be the main power supply source in the system as it provides very little service in terms of extra capacity due to the fact that wind power available during peak periods is usually low.

Apart from the empirical analysis of wind integration, some of the literature has focused on the investment incentives involved in building wind farms. Ambec and Crampes (2012) study the impact of wind intermittency on profits and investment incentives. These authors assume that wind power has two states: when the wind is blowing, all wind turbines produce electricity at maximum capacity; however, when the wind is not blowing, no power is produced. In a static framework, they show that markets with regulated monopoly will result in strictly positive profit for wind generators. Their model also shows that when consumers cannot react to price changes, wind plants and fossil fuel plants are seen as complementary. This implies that wind capacity would have to be duplicated by building fossil fuel plants. This reinforces the previous result by Belanger and Gagnon (2002) that wind turbines have little in term of extra capacity. However, when consumers are allowed to react to electricity prices, wind power is seen as a substitute for power from fossil fuel plants.

Garcia and Alzate (2012) analyze the effectiveness of two regulatory mechanisms meant to encourage renewable energy: Feed in tariffs (FIT) and Renewable Portfolio standards (RPS). Their model is a two-stage game. In stage one, a decision on investment is made, and in stage two, firms compete in prices to dispatch their production capacity. The model shows that the lower the ratio of conventional fossil fuel generators to demand, the weaker the incentive needed to induce an optimal level of renewable generation. Also, feed in tariffs will be greater than the price of conventional electricity generated from fossil fuels if and only if the fixed costs of renewable resources are large, while RPS that induces an optimal investment in renewable technology will further induce under-investment in conventional fossil fuels.

Much of the research of wind power has focused on the impact of investment in wind turbine capacity and not much has been said on the impact of wind farms on the optimal water depletion policy under taken by a hydroelectric generator. When wind power is integrated with hydroelectricity, the intermittent nature of wind farms will transform the problem of the hydroelectric generator from a simple dynamic optimization problem to a stochastic dynamic optimization problem. Therefore the integration of wind power will change the optimal water release strategy of a hydroelectric generator and the price of electricity.

In this paper, the effect of wind intermittency on the water release policy of a hydroelectric generator is analysed. Using a stochastic dynamic model, the intermittent nature of wind power is explicitly modelled into the profit function of a hydroelectric generator. The model is then used to determine the optimal water release policy for a hydroelectric generator at various capacity levels. To solve the stochastic dynamic model, the backward induction method is used.

2. THE MODEL

Water resources are renewable and follow a cycle. In the model T is the time horizon, which represents the length, say one year, of the cycle. If one considers a day as a period in the cycle, then there are 365 periods under time horizon T. On each day of the cycle, water flows into the reservoir from precipitation and flows out through release or through evaporation. At the end of period T, the cycle repeats itself. In this manner, the total volume of water available for use during a cycle behaves like a non-renewable resource, but is renewable in the sense that the cycle occurs every year. Thus, it makes more sense to consider the time horizon as the length of a cycle. At the end of the cycle, the problem repeats itself.[2]

[2] When one lets $T \rightarrow \infty$, one obtains in the limit the solution of the infinite-time horizon problem. However, it is necessary to discount profits so that the present value of the stream of profits converges.

Let $D: p \rightarrow D[p]$ be the market demand curve for electricity in each period, where p is the price of electricity and $D[p]$ is the market demand at price p. On the supply side, there are two types of generators: a dominant firm[3] and a competitive fringe. Because of its market power, the dominant firm is able to set the market price of electricity. A generator in the fringe takes the price as given, and produces electricity at the level of output where marginal cost is equal to the market price of electricity.

The supply from the fringe is expected to rise with the market price: as the price of electricity rises, more inefficient generators will be able to enter the market. If we let $S: p \rightarrow S[p]$ denote the supply curve of the fringe, then the residual demand curve for the output of the dominant firm is given by $Q: p \rightarrow Q[p] = D[p] - S[p]$. For simplicity, we shall assume that the residual demand curve is linear, and its inverse is given by

(1) $p = a - b\, Q,$

where Q the volume of electricity is delivered by the dominant firm to the market in a period, and p is the price it obtains for this supply. Also, a and b are two positive parameters that characterize the residual demand curve.

The total revenue associated with the residual demand curve (1) is

(2) $\pi[Q] = Q\,(a - b\, Q),$

which rises with Q as Q rises from 0; reaches a maximum at $Q_{max} = \frac{a}{2\,b}$ and then falls as Q rises above $\frac{a}{2\,b}$.Thus, it is not rational for the dominant firm to deliver to the market a volume of electricity greater than Q_{max}. The maximum revenue that is possible is then given by

[3] It is generally accepted that hydroelectric generators have some form of market power in electricity market because of their ability to move electricity production over time.

(3) $\pi_{max} = \pi[Q_{max}] = \frac{a}{2\,b}\ (a - \frac{a}{2}) = \frac{a^2}{4\,b}$.

The generators in the competitive fringe produce electricity by burning fossil fuels. As for the dominant firm, its supply has two components: hydroelectric power and wind energy. The dominant firm produces hydroelectric power by releasing the water damped up in its reservoirs. It does not produce wind energy, but is obligated by law to purchase all the wind energy produced by privately-owned wind farms at a stipulated price, say γ.

Let $X[t]$ denote the volume of water remaining in the reservoir at the beginning of period t, and $Y[t]$ denote the output of wind energy - also in period $t, t = 1, \ldots, T$. The volume of water in the reservoir at the beginning of period $t = 1$, namely $X[1]$, is given. We assume that $Y[t], t = 1, \ldots, T$, are independent identically distributed random variables with common distribution function $F[Y]$, $0 < Y_{min} \leq Y \leq Y_{max}$, where Y_{min} and Y_{max} are, respectively, the lower and upper bounds of the wind energy generated in each period. The density of $F[Y]$, denoted by $f[Y]$, is assumed to be strictly positive and continuous in $[Y_{min}, Y_{max}]$. In what follows, we assume that wind power meets only a small part of the residual demand faced by the dominant firm. More precisely, we assume that $2\,Y_{max} < Q_{max}$. Also, we let μ and σ^2 denote, respectively, the mean and variance of Y.

Next, let

(4) $\bar{\omega} = \int_{Y_{min}}^{Y_{max}} \pi[Y] f[Y] dY$

denote the expected revenue generated by wind energy in a period when no hydroelectric power is produced.

Let $\gamma > 0$ denote the unit price that the dominant firm pays the wind farms for the power they generate in each period. The volume of power at which the marginal revenue is equal to γ is the value of Q, say Q_γ, that is given by

(5) $$Q_\gamma = \frac{a-\gamma}{2b}.$$

Because the dominant firm is obligated to purchase the entire output of the wind farms in each period, the marginal revenue - net of the marginal purchase cost - derived from wind power in any period is negative when the wind power output exceeds $Q_{\gamma.}$ We shall assume that $Q_{min} < Q_\gamma < Q_{max}$

The state of the system in each period t is the list $(X[t], Y[t])$, where $X[t]$ is the remaining volume of water in the reservoir at the beginning of that period, and $Y[t]$ is the realized volume of wind energy available also, in that period. A feed-back policy is a real-valued function $u: (X, Y, t) \to u[X, Y, t]$, which tells the dominant firm how much water to release in period t, given that (X, Y) is the state of the system in that period. For simplicity, we assume that the volume of hydroelectric power generated in a period is equal to the volume of water released in that period. Hence $u[X, Y, t]$ represents the volume of hydroelectricity produced under the policy u in period t, as a function of the state (X, Y) of the system in that period.

Suppose that the system begins in period 1 with $X[1]$ as the volume of water in the reservoir and $Y[1]$ as the realized volume of wind energy generated. Suppose also that a control policy $u: (X, Y, t) \to u[X, Y, t]$ is adopted. The profit, net of production cost and the payments to the wind energy producers, made by the dominant firm over the entire time horizon, under a particular realized time path$(Y[t])_{t=1}^{T}$ of wind energy output is given by

(6) $$\sum_{t=1}^{T} \pi\big[Y[t] + u[X[t], Y[t], t]\big] - \gamma Y[t],$$

where

(7) $X[t+1] = X[t] - u[X[t], Y[t]], \quad t = 1, \dots, T-1,$

$X[1]$ is given, and $Y[1]$ is also given.

Note that in (6) we have neglected the cost of producing hydroelectricity.

The expected profit made by the dominant firm over the entire time horizon, as a function of u and the initial condition $(X[1], Y[1])$, is then given by

(8) $\mathcal{E}\left[\sum_{t=1}^{T} \pi\left[Y[t] + u[X[t], Y[t], t]\right] - \gamma Y[t]\right].$

In (8), $\mathcal{E}$ is the expectation operator with respect to the joint distribution of $(Y[1], \dots, Y[t], \dots, Y[T])$.

The problem faced by the dominant firm can now be stated formally as follows:

(9) $v\left[X[1], Y[1]\right] = Max_u\, \mathcal{E}\left[\sum_{t=1}^{T} \pi\left[Y[t] + u[X[t], Y[t], t]\right] - \gamma Y[t]\right].$

In (9), $v[X[1], Y[1]]$ is the Bellman function for our problem, which gives the maximum expected profit made by the dominant firm, given that it begins period 1 with $X[1]$ as the volume of water in the reservoir and $Y[1]$ as the realized wind energy output in that period. We shall solve the maximization problem (8) by backward induction of the number of the remaining periods.

3. THE ONE-PERIOD PROBLEM

Suppose that the dominant firm has only one period to exploit a volume of water, say $X[1]$, in its reservoirs. Also, suppose that $Y[1]$ is the realized output of wind power that it is obligated to purchase.

If $X[1] + Y[1] < Q_{max}$, then the optimal strategy for the dominant firm to release all the water in the reservoir and send the hydroelectric cum wind power to the market. The profit it obtains by carrying out this action is

(10) $v_1[X[1], Y[1]] = \pi[X[1] + Y[1]] - \gamma Y[1], \quad X[1] < Q_{max} - Y[1]$.

Note that in (10) we have denoted the Bellman function for the one-period problem by v_1 , with the subscript indicating the number of remaining periods.

On the other hand, if $X[1] + Y[1] \geq Q_{max}$, then it is optimal to send the revenue-maximizing volume of power Q_{max} to the market by releasing the volume of water $x = Q_{max} - Y[1]$ for hydroelectric generation, and then delivering to the market the hydroelectricity produced x plus the wind power generated $Y[1]$. The profit made by the dominant firm is

(11) $v_1\big[X[1], Y[1]\big] = \pi_{max} - \gamma Y[1] , \qquad X[1] \geq \ Q_{max} - Y[1]$.

PROPOSITION 1: *For the one-period problem, the optimal value function (or the Bellman function) gives the profits made by the dominant firm - as a function of the volume of water in the reservoir* $X[1]$ *and the realized output of wind power* $Y[1]$ - *is given by*

(12) $v_1\big[X[1], Y[1]\big] = \begin{cases} \pi\big[X[1] + Y[1]\big] - \gamma Y[1] & \text{if } X[1] + Y[1] < Q_{max}, \\ \pi_{\max} - \gamma Y[1] & \text{if } X[1] + Y[1] \geq Q_{max}, \end{cases}$

The Bellman function $v_1\big[X[1], Y[1]\big]$ *has the following properties:*

(i) *It is concave in* $(X[1], Y[1])$;

(ii) *It is continuously differentiable;*

(iii) G*iven* $Y[1]$, *it is increasing in* $X[1]$. *However, given* $X[1]$, *the contributions of wind power to the profitability of the dominant firm might be positive or negative.*

PROOF: Because the dominant firm is obligated to purchase the entire output of wind farms in each period, it has no influence on the cost term $\gamma Y[1]$. If $X[1]+Y[1]<Q_{max}$, then clearly it is optimal to deliver all the power volume $X[1]+Y[1]$ to the market. However, when $X[1]+Y[1]\geq Q_{max}$, it is optimal to deliver to the market only the volume of power that maximizes total revenue, namely$Y[1]+x$, with $x=Q_{max}-Y[1]$ as the volume of water released. We have just established (12).

The concavity of $v_1\big[X[1],Y[1]\big]$ follows from the concavity of $\pi[Q]$. The fact that $v_1\big[X[1],Y[1]\big]$ is continuously differentiable is inherited from $\pi[Q]$. Also, $X[1]\rightarrow$ v1[X[1], Y[1]] is clearly increasing. As for the contribution of wind power to the profitability of the dominant firm, the result might be positive or negative depending on both γ and $Y[1]$. When the volume of water in the reservoir is low, so that at the margin the marginal revenue $\pi'\big[X[1]+Y[1]\big]>\gamma$, the revenue obtained from one more unit of wind power exceeds its purchase cost, and the profit made by the dominant firm improves with a slight rise in the output of wind power. However, when the volume of water in the reservoir and the output of wind power are not too low, so that $\pi'\big[X[1]+Y[1]\big]<\gamma$ a rise in the output of wind energy depresses the profits of the dominant firm: at the margin, the revenue the dominant firm earns on the last unit of wind power falls short of the price it has to pay for this power unit. ■

At the present time, renewable energy is not yet competitive with fossil fuels, especially with coal, because of its high capital cost. There are several major reasons why the governments of some countries choose to subsidize the production of renewable energy. First, the age of cheap oil is coming to an end. When fossil fuels run out, all the sources of renewable energy must be exploited to meet the energy needs of a country. Second, and this is perhaps the most frequently advanced argument for subsidizing wind energy, the burning of fossil fuels generates greenhouse gases which induces climate change. Wind energy, on the other hand, is non-polluting, and its marginal cost is negligible once the heavy capital investments have been made. It is hoped that the high capital costs will be alleviated by technological progress. In the US and Canada, regulators force public utilities to purchase all the wind power output at a substantial stipulated price, as is the case concerning Hydro Québec and wind power producers in this province. In our model, the price that the dominant firm pays for one unit of wind power reflects the policy adopted by regulators of electricity markets.

The shadow price of water in the reservoir is given by

$$(13) \quad D_1 v_1[X[1], Y[1]] = \begin{cases} \mathrm{a} - 2\,\mathrm{b}\,(\mathrm{X}[1] + \mathrm{Y}[1]) \text{ if } \mathrm{X}[1] + \mathrm{Y}[1] < \mathrm{Q}_{\max}, \\ 0 \text{ if } \mathrm{X}[1] + \mathrm{Y}[1] \geq \mathrm{Q}_{\max} \end{cases}.$$

Note that the schedule $X[1] \rightarrow D_1 v_1[X[1], Y[1]]$ is continuous, linearly decreasing in $X[1]$ for all $X[1] < Q_{max} - Y[1]$, and vanishes for $X[1] \geq Q_{max} - Y[1]$. Also, note that a rise in $Y[1]$ shifts this schedule downward. That is, a higher volume of water in the reservoir or a higher realized output of wind power lowers the contribution of a marginal unit volume of water.

Suppose that $X[1]$ is the known volume of water in the reservoir at the beginning of the period. Before the output of wind power is known, the expected value of the shadow price of water in the reservoir is given by

$$(14) \quad \phi_1[X[1]] = \int_{Y_{min}}^{Y_{max}} D_1 v_1[X[1], Y[1]] f[Y] dY.$$

LEMMA 1: *For the one-period problem, the expected shadow price of water before the output of wind power is known is a continuous and decreasing function of the volume of water in the reservoir. More explicitly, it is given by*

(15)

$$\phi_1[X[1]] = \begin{cases} a - 2\,b\,(X[1] + \mu) if\ X[1] < Q_{max} - Y_{max}\,, \\ (a - 2\,b\,X[1])F[Q_{max} - X[1]] - 2b\int_{Y_{min}}^{Q_{max}-Y_{max}} Y\,f[Y]dY\ if\ Q_{max} - Y_{max} \leq X[1] < Q_{max} - Y_{min}\,, \\ 0\ if\ X[1] \geq Q_{max} - Y_{min}. \end{cases}$$

That is, the expected shadow price of water, as a function of $X[1]$ *and before the output of wind power is known, is linearly decreasing for* $X[1]$ *in the interval* $0 \leq X[1] < Q_{max} - Y_{max}$, *strictly decreasing for* $X[1]$ *in the interval* $Q_{max} - Y_{max} \leq X[1] < Q_{max}$ *and vanishes for all* $X[1] \geq Q_{max}$.

PROOF: We have

$$\phi_1[X[1]] = \begin{cases} \int_{Y_{min}}^{Y_{max}} (a - 2b\,(X[1] + Y))f[Y]dY\ if\ X[1] < Q_{max} - Y_{max}, \\ \int_{Y_{min}}^{Q_{max}-X[1]} (a - 2b\,(X[1] + Y))f[Y]dY\ if\ Q_{max} - Y_{max} \leq X[1] < Q_{max} - Y_{min}, \\ 0\ if\ X[1] > Q_{max} - Y_{min}. \end{cases}$$

On the right side of the preceding expression, the first integral is given by

$$\int_{Y_{min}}^{Y_{max}} (a - 2b\,(X[1] + Y))f[Y]dY = a - 2b(X[1] + \mu),$$

which is linearly decreasing in $X[1]$, while the second integral, namely

$$\int_{Y_{min}}^{Q_{max}-X[1]} (a - 2b\,(X[1] + Y))f[Y]dY,\ \ Q_{max} - Y_{max} \leq X[1] < Q_{max} - Y_{min}.$$

is strictly and non-linearly decreasing in $X[1]$. Some computations yield the following more explicit form for the second integral:

$$\int_{Y_{min}}^{Q_{max}-X[1]} (a - 2b\,(X[1] + Y))f[Y]dY$$
$$= (a - 2bX[1])F[Q_{max} - X[1]] - 2b\int_{Y_{min}}^{Q_{max}-X[1]} Yf[Y]dY,\quad Q_{max} - Y_{max} \leq X[1] < Q_{max} - Y_{min}.$$

For $X[1] > Q_{max} - Y_{min}$, we have $X[1] + Y \geq Q_{max}$ and $D_1 v_1\big[X[1], Y[1]\big]$=0 for all Y. ■

For $X[1] = 0$, the expected shadow price of water for the one-period problem is given by

(16) $\phi_1[0] = a - 2\, b\mu$.

This will be used in solving the two-period problem.

4. THE TWO-PERIOD PROBLEM

Suppose that the dominant firm has two periods remaining to exploit the water in its reservoirs. Also, suppose that the volume of water in the reservoirs at the beginning of period 1 is $X[1]$ and that $Y[1]$ is the realized output of wind power in period 1.

If $x, 0 \leq x \leq X[1]$, is the volume of water released in the first period, then the profit - net of the cost of purchasing the wind power generated - that the dominant firm makes in the first period is $\pi\big[x + Y[1]\big] - \gamma Y[1]$. The volume of water that remains in the reservoirs at the beginning of the next period is $X[1] - x$. If Y is the realized volume of wind power generated in the next period, then the realized profit - net of the cost of purchasing the wind power - in the next period is $v_1[X[1] - x, Y]$. If we denote by $\psi_2\big[x, X[1], Y[1]\big]$expected profit made by the dominant firm over the two periods, then it is given by

(17) $\psi_2\big[x, X[1], Y[1]\big] = \pi\big[x + Y[1]\big] - \gamma Y[1] + \int_{Y_{min}}^{Y_{max}} v_1[X[1] - x, Y] f[Y] dY.$

The two-period problem can be stated formally as follows:

(18) $max_x \, \psi_2[x, X[1], Y[1]]$

subject to $0 \leq x \leq X[1]$.

To solve the maximization problem (18) when water is not abundant, we apply the Kuhn-Tucker theorem. First, write the Lagrangian

$$\mathcal{L} = \pi[x + Y[1]] - \gamma Y[1] + \int_{Y_{min}}^{Y_{max}} v_1[X[1] - x, Y] f[Y] dY - \lambda[1](-x) - \lambda[2](x - X[1]),$$

where $\lambda[1]$and $\lambda[2]$ are the multipliers associated, respectively, with the constraint $x \geq 0$ and the constraint $x \leq X[1]$.

The following first-order conditions characterize the optimal volume of water released in the current period:

(20) $$\frac{\partial \mathcal{L}}{\partial x} = \pi'[x + Y[1]] - \int_{Y_{min}}^{Y_{max}} D_1 v_1[X[1] - x, Y] f[Y] dY + \lambda[1] - \lambda[2]$$
$$= a - 2b(x + Y[1]) - \phi_1[X[1] - x] + \lambda[1] - \lambda[2] = 0;$$

(21) $\lambda[1] \geq 0, x \geq 0, \ \lambda[1]x = 0;$

(22) $\lambda[2] = 0, x \leq X[1], \lambda[2](x - X[1]) = 0;$

4.1.Water is Abundant

In attempting to find the solution of the expected profit maximization problem (19), we first note that if

(23) $X[1] \geq 2Q_{max} - Y[1] - Y_{min}$,

then it is feasible for the dominant firm to deliver the revenue-maximizing volume of power Q_{max} to the market in each of the two periods. Under this scenario, the volume of water that the dominant firm should release is $x = Q_{max} - Y[1]$, and then send the hydroelectricity thus produced cum the wind power output to the market in the current period. In the next period, if $Y[2]$ is the realized output of wind power, then the dominant firm should release a volume of water equal to $Q_{max} - Y[2]$, which is always feasible - when condition (23) holds - regardless of the value of $Y[2]$. The total profits over the two periods for the case in which there is an abundance of water is given by

(24) $v_2\big[X[1], Y[1]\big] = 2\,\pi_{max} - \gamma(Y[1] + \mu)$

The scenario characterized by condition (23) is depicted in Figure 1.

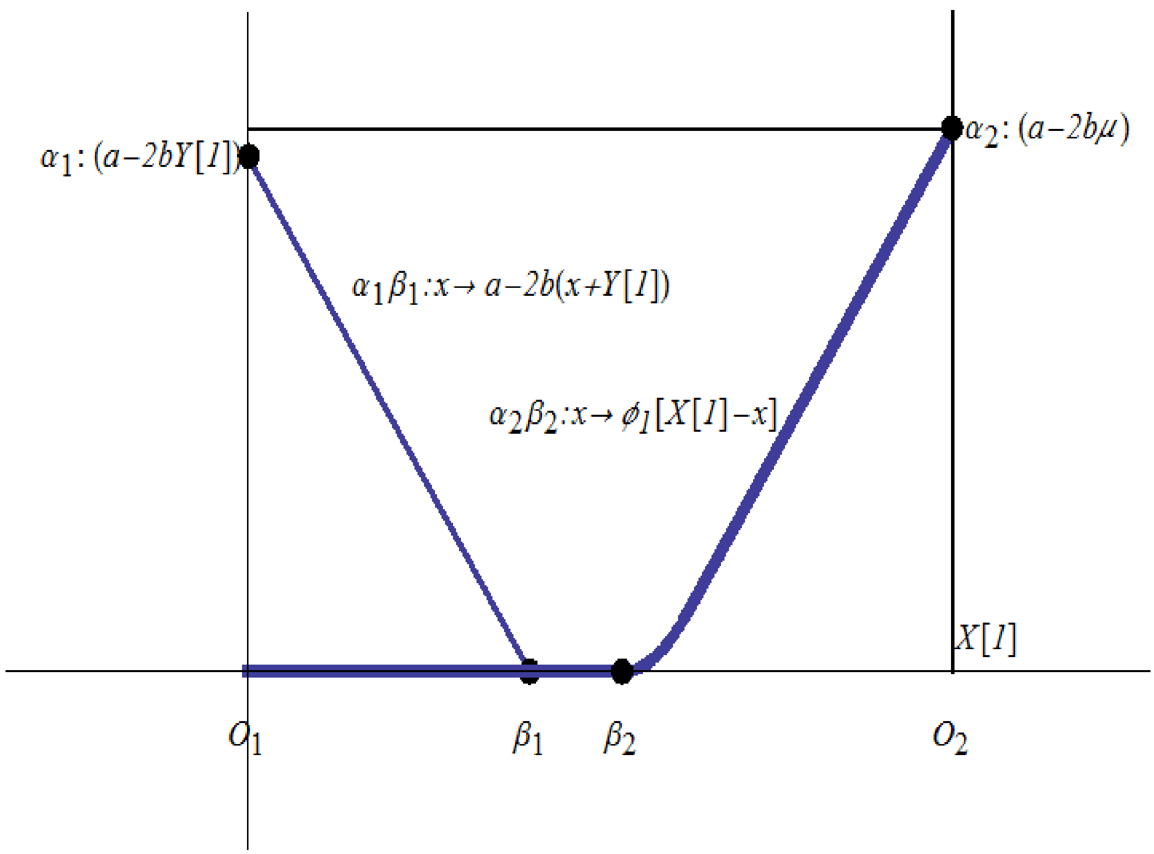

Figure 1. -- The Scenario "Water is abundant: $X[1] \geq 2Q_{max} - Y[1] - Y_{min}$

In Figure1 are drawn two curves: $\alpha_1\,\beta_1$ and $\alpha_2\,\beta_2$. The curve labelled $\alpha_1\,\beta_1$ represents the map $x \to\ a - 2\,b\,(x + Y[1]), 0 \leq\ x \leq\ X[1]$, which gives the contribution at the margin of one unit volume of water released in the current period. For this curve, the origin is O_1. The curve $\alpha_1\,\beta_1$ begins at point $\alpha_1$on the vertical axis, with $O_1\alpha_1 = \mathrm{a} - 2bY[1]$, and ends on the horizontal axis at point β_1, with $O_1\beta_1 = \mathrm{Q_{max}} - \mathrm{Y}[1]$.

The second curve, which is drawn thicker and is labelled $\alpha_2\,\beta_2 O_1$, is the map $x \to \phi_1[X[1] - x], 0 \leq x \leq X[1]$ which represents the expected shadow price of water at the margin in the next period. For $\alpha_2\,\beta_2 O_1$, the origin is O_2, with the length of $O_1 O_2$ being equal to $X[1]$, the volume of water in the reservoir at the beginning of the current period. The curve $\alpha_2\,\beta_2 O_1$, begins at α_2 on the vertical axis, descends to 0 at point β_2 on the horizontal axis, and then remains equal to 0 until it reaches point O_1, the origin of the other curve. For this curve, we have $O_2\alpha_2 = \mathrm{a} - 2b\mu$ and $O_2\beta_2 = \mathrm{Q_{max}} - \mathrm{Y_{min}}$.

Both curves, seen from their respective origins, are downward-sloping. In the figure, x is measured from left to right on the horizontal axes, while $X[1] - x$, the volume of water remaining at the beginning of the next period, is measured from right to left on the horizontal axis.

4.2.Water is neither Abundant nor Scarce

Having solved the expected profit maximization problem (18) for the case water is abundant let us now continue the analysis by considering the values of $X[1]$ below $2\,Q_{max} - Y[1] - Y_{min}$. For this exercise, let us lower $X[1]$ continuously, and note that this action leaves the curve $\alpha_1\,\beta_1$ undisturbed, but shifts the curve $\alpha_2\,\beta_2$ continuously to the left and parallel to itself. In Figure 1, the case of water abundance $\beta_2$is to the right of β_1. As the curve $\alpha_2\,\beta_2$ continuously shifts to the left, there comes a time when β_2 coincides with β_1, and then finds itself on the left of β_1. Figure 2 depicts the situation at a point in time when β_2 is in a left neighbourhood of β_1, and the two curves intersect at point E, which represents an interior solution for the expected profit maximization problem (18). Under this scenario, the dominant firm releases a positive volume of water in each of the two periods.

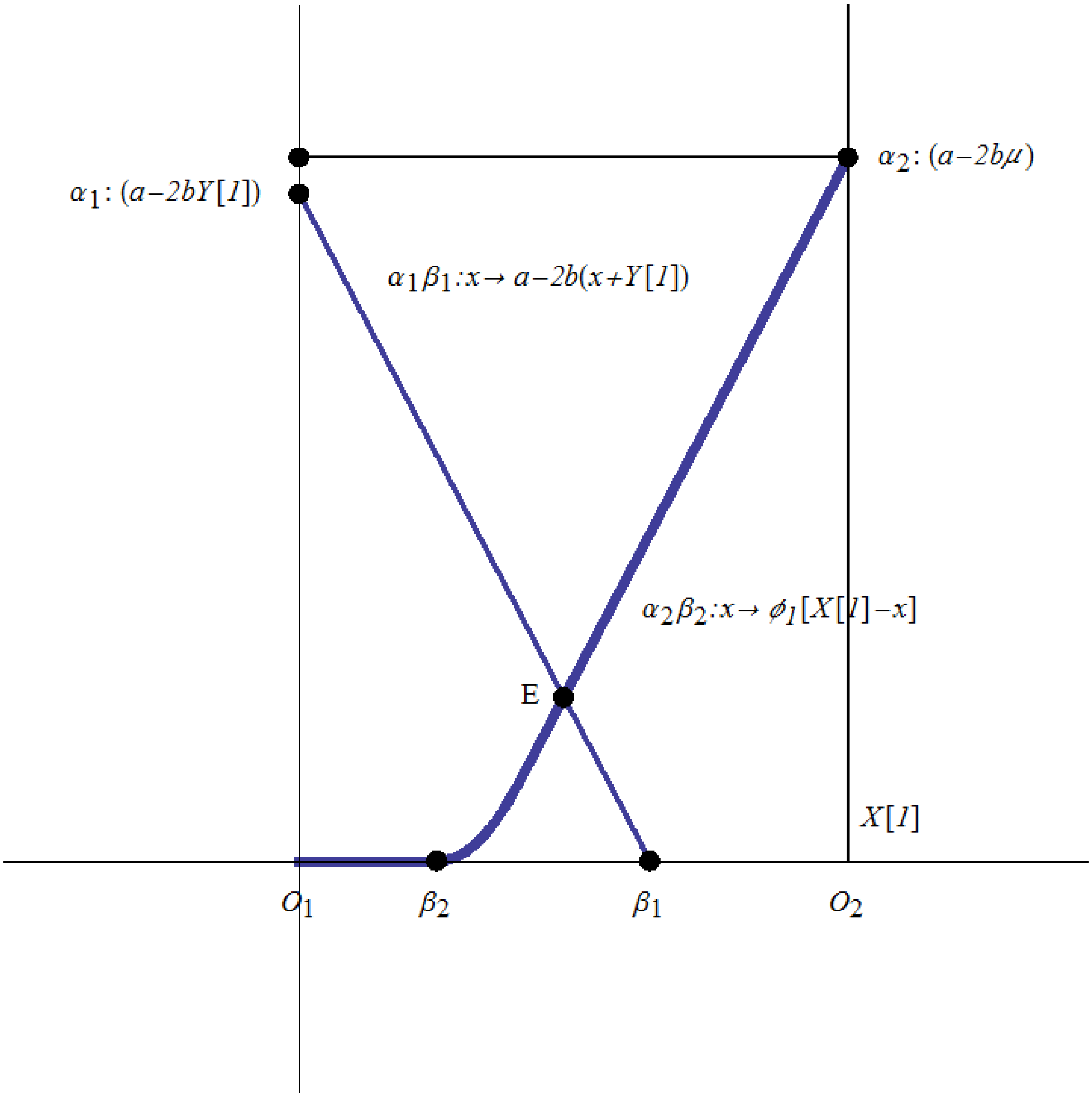

Figure 2.-- The Scenario "Water is neither abundant nor scarce: Interior solution"

For an interior solution, the two multipliers are both equal to 0, i.e., $\lambda[1] = \lambda[2] = 0$, and the first-order condition (20) is reduced to

(25) $a - 2\,b\,(x + Y[1]) - \phi_1[X[1] - x] = 0.$

4.3.Water is Scarce

As we continue to shrink $X[1]$, the optimal solution, as represented by point E, moves continuously up the curve $\alpha_1\,\beta_1$ as well as the curve $\alpha_2\,\beta_2$ until $X[1]$ descends to a critical value below which one curve lies below the other curve. If $a - 2b\mu \geq a - 2bY[1]$, or equivalently if $Y[1] \geq \mu$, then the critical value of $X[1]$, say $X[1] = \xi_2$, solves the following equation:

(26) $a - 2\,bY[1] - \phi_1[\xi_2] = 0$,

or

(27) $\xi_2\big[Y[1]\big] = (\phi_1)^{-1}\big[a - 2bY[1]\big]$,

where we have let $(\phi_1)^{-1}$denote the inverse for the strictly decreasing part of the curve $\phi_1\colon Z \rightarrow \phi_1[Z]$, $Z \geq 0$.

Furthermore, for all $X[1] \leq \xi\big[Y[1]\big]$, the curve $\alpha_1\,\beta_1$ lies below the curve $\alpha_2\,\beta_2$, and the optimal solution for the expected profit maximization (18) involves releasing no water in the current period. This scenario is depicted in Figure 3.

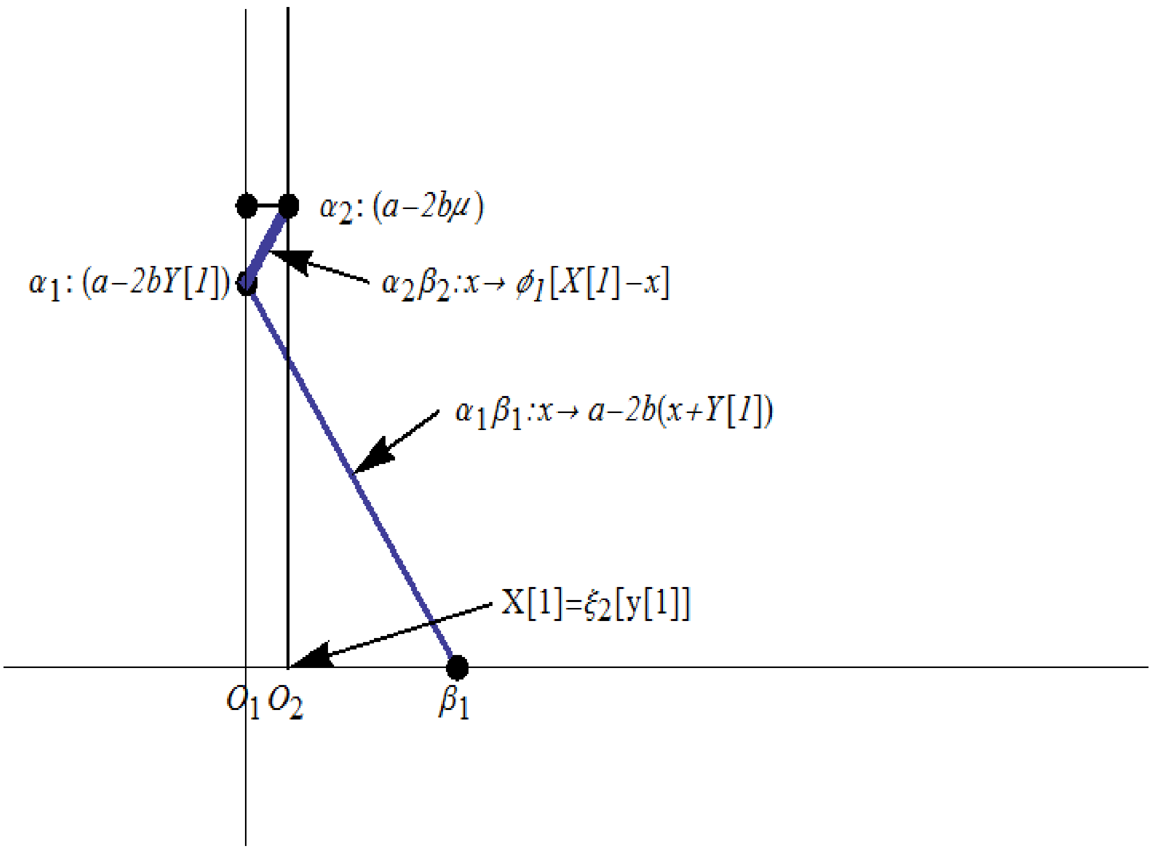

Figure 3. The Scenario "Water is scarce and the realized output of wind power is above its expected value: No water is released in the current period"

On the other hand, if $a - 2b\mu < a - 2bY[1]$, or equivalently if $Y[1] < \mu$, then the critical value of $X[1]$, say $X[1] = \eta_2$, solves the following equation:

(28) $a - 2\,b(\eta_2 + Y[1]) - \phi_1[0] = a - 2\,b(\eta_2 + Y[1]) - (a - 2\,b\mu) = 0$

i.e.

(29) $\eta_2\big[Y[1]\big] = \mu - Y[1].$

Furthermore, for all $X[1] \leq \eta_2\big[Y[1]\big]$, the curve $\alpha_1\beta_1$ lies above the curve $\alpha_2\beta_2$, and the optimal solution for the expected profit maximization (18) involves emptying the reservoir in the current period. This scenario is depicted in Figure 4.

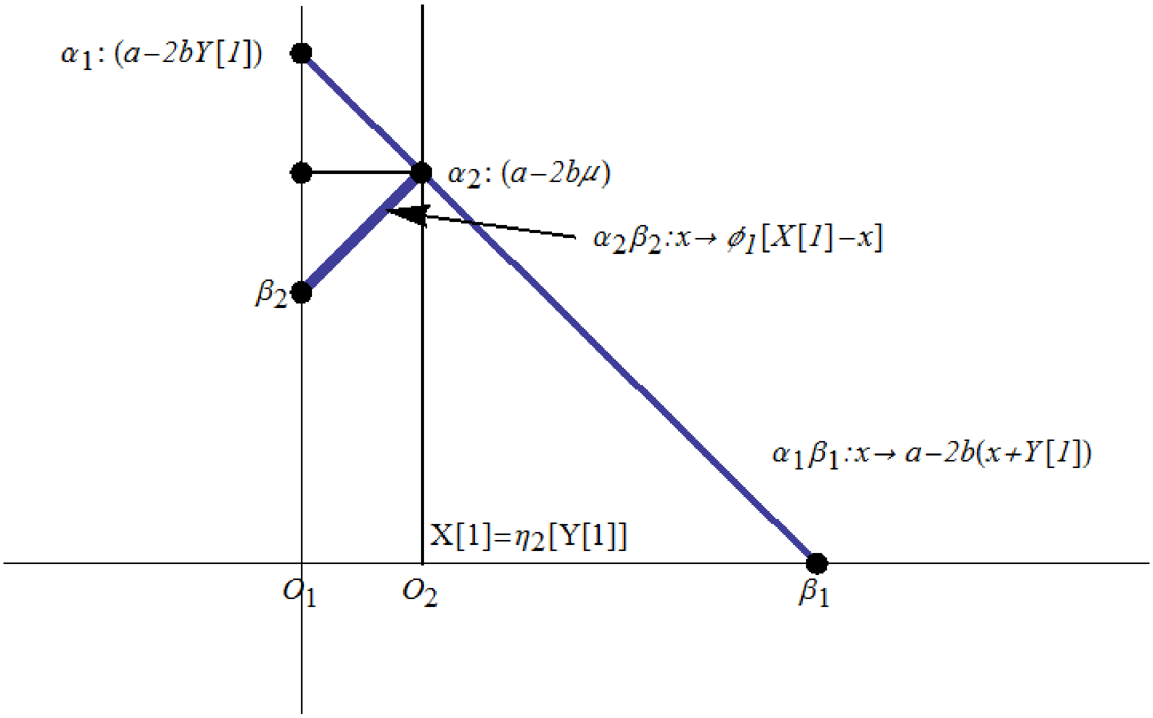

Figure 4. -- The Scenario "Water is scarce and the realized wind power output is below its expected value: Empty the reservoir in the current period"

4.4. The Properties of the Solution for the Two-Period Problem

PROPOSITION 2:

(i) *If $Y[1]$, the realized output of wind power in the current period, exceeds its expected value μ then the optimal solution for the two-period problem can be described as follows:*

   a. *For $X[1] < \xi_2\big[Y[1]\big]$, release no water in the current period.*

   b. *For $\xi_2\big[Y[1]\big] \leq X[1] < 2Q_{max} - Y[1] - Y_{min}$, release a positive volume of water in each of the two periods.*

   c. *For $X[1] \geq 2Q_{max} - Y[1] - Y_{min}$, deliver to the market in each period the revenue-maximizing volume of power Q_{max} - hydroelectric and wind - in each period.*

(ii) *If $Y[1]$ the realized output of wind power in the current period falls short of its expected value μ, then the optimal solution for the two-period problem can be described as follows:*

a. *For $X[1] < \eta_2\big[Y[1]\big]$ empty the reservoir in the current period.*

b. *For $\eta_2\big[Y[1]\big] \leq X[1] < 2Q_{max} - Y[1] - Y_{min}$, release a positive volume of water in each of the two periods.*

c. *For $X[1] \geq 2Q_{max} - Y[1] - Y_{min}$, deliver to the market in each period the revenue-maximizing volume of power Q_{max} - hydroelectric and wind - in each period.*

PROOF: Proposition 2 just summarizes the results obtained in Sub-sections 4.1-4.3.■

The economic contents of Proposition 2 are quite intuitive. When there is an abundance of water, the water resource constraint is not binding, and the dominant firm maximizes the expected profit over the two periods by delivering the revenue-maximizing volume of power Q_{max} to the market in each period. When water is not abundant, but not too scarce, the optimal solution dictates that a positive volume of water is released in each period. This is the case when the optimal solution is an interior solution. When water resources are really scarce, the optimal solution is a corner solution. If the realized current output of wind power is below its expected value, it is optimal to empty the reservoir in the current period. On the other hand, if the realized current output of wind power is above its expected value, it is optimal to reserve all the water for exploitation in the next period.

PROPOSITION 3: *For the two-period problem, the optimal volume of water released in the current period rises with the volume of water stored in the reservoir, but falls as the current realized output of wind power rises. In the latter case, the fall in the volume of water released is smaller than the rise in the realized output of wind power, with the ensuing consequence that the total volume of power - wind cum hydroelectric - that the dominant firm delivers to the market in the current period rises with the realized output of wind power.*

PROOF: A rise in $X[1]$ lengthens the distance O_1O_2 , the greater distance shifts the curve $x \to \phi_1[X[1]-x], 0 \leq x \leq X[1]$ to the right and lengthens it at the same time because there is now more water in the reservoir. As for the curve $x \to \mathrm{a} - 2\,\mathrm{b}(x + \mathrm{Y}[1]), 0 \leq x \leq X[1]$, the rise in $X[1]$ lengthens, but does not shift it. At the new equilibrium, the volume of water released will be either higher or at the same initial level. The latter case occurs if $X[1]$ has not risen sufficiently to preclude a corner solution.

As for the impact of the current realized output of wind power, a rise in $Y[1]$ shifts the curve $x \to \mathrm{a} - 2\,\mathrm{b}(x + \mathrm{Y}[1]), 0 \leq x \leq X[1]$, downward and parallel to itself, while leaving the other curve unchanged. At the new equilibrium, the volume of water released in the current period will either be higher or remain the same. Because $\pi'[x + Y[1]]$ is lower at the new equilibrium than at the initial equilibrium, $(x + Y[1])$ must be higher at the new equilibrium than at the initial equilibrium. ■

The fact that the generation of hydroelectric power rises with the volume of water in the reservoir, as asserted by Proposition 3, is not hard to understand. As for the statement that the higher is the realized current output of wind power, the lower is the generation of hydroelectric power in the current period, it captures the concept of integration of wind energy with hydroelectric production. As asserted by Proposition 3, the reduction in the volume of hydroelectric power is smaller than the rise in the current realized output of wind power, which implies that the price set by the dominant firm falls as the current realized output of wind power increases. The lower price set by the dominant firm induces a contraction in the output of the firms in the competitive fringe. Because the firms in the competitive firms burn fossil fuels to generate electricity, a lower output from the competitive fringe means a reduction in the emissions of greenhouse gases. Thus, a higher current realized output of wind power has a beneficial effect on climate change.

The intermittent nature of wind power and its impact on the supply decision of the dominant firm can be analyzed by considering a mean-preserving spread of the distribution of Y. To this end, let $\tilde{Y} = \theta Y + (1-\theta)\mu$, where $\theta \geq 1$. As defined, the random variable $\tilde{Y}$ has mean μ and variance $\theta^2\sigma^2$. An increase in θ above 1 represents a mean-preserving spread of the distribution of Y.

When there are two remaining periods to exploit the volume of water in the reservoir, and when wind speeds are more intermittent at secondary wind site, the dominant firm solves the following expected profit maximization problem:

(30) $$max_x \, \pi[x + Y[1]] - \gamma Y[1] + \int_{Y_{min}}^{Y_{max}} v_1[X[1] - x, \theta Y + (1-\theta)\mu] f[Y] dY$$

$subject\ to\ 0 \leq x \leq X[1]$.

The following first-order condition characterizes an interior solution of (30):

(31) $\pi'[x + Y[1]] - \int_{Y_{min}}^{Y_{max}} D_1 v_1[X[1] - x, \theta Y + (1-\theta)\mu] f[Y] dY = 0$.

The second-order condition obviously holds:

(32) $\pi''[x + Y[1]\,] + \int_{Y_{min}}^{Y_{max}} D_{11} v_1[X[1] - x, \theta Y + (1-\theta)\mu] f[Y] dY < 0$.

PROPOSITION 4*: Suppose that the dominant firm has two remaining periods to exploit the volume of water in the reservoir. A mean-preserving spread of the random output of wind power reduces the volume of water that is released in the current period.* [4]

PROOF: Suppose that $Y[1]$, the wind power output in the reference scenario, is given. Differentiating the first-order condition (31) with respect to θ, we obtain

(33) $\pi''[x + Y[1]] \frac{\partial x}{\partial \theta}$

[4] I want to thank Prof Jacques Robert for this Proposition

$$+\frac{\partial x}{\partial \theta}\int_{Y_{min}}^{Y_{max}} D_{11}v_1[X[1]-x,\theta Y+(1-\theta)\mu]f[Y]dY$$

$$-\int_{Y_{min}}^{Y_{max}} D_{12}v_1[X[1]-x,\theta Y+(1-\theta)\mu](Y-\mu)f[Y]dY=0.$$

Now if we differentiate (13) with respect to $Y[1]$, we obtain $D_{12}v_1\big[X[1],Y[1]\big]=-2b$ for all $Y\le Q_{max}-X[1]$ and $D_{12}v_1\big[X[1],Y[1]\big]=0$ for all $Y>Q_{max}-X[1]$. Using this last result to evaluate the last integral in (33), we obtain

(34) $\int_{Y_{min}}^{Y_{max}} D_{12}v_1[X[1]-x,\theta Y+(1-\theta)\mu](Y-\mu)f[Y]dY$

$=\int_{Y_{min}}^{Q_{max}-X[1]} -2b(Y-\mu)f[Y]dY$

$=2b\int_{Y_{min}}^{Q_{max}-X[1]}(\mu-Y)f[Y]dY>0$

Solving (33) for $\frac{\partial x}{\partial \theta}$, and using (34), we obtain

(35) $\frac{\partial x}{\partial \theta}=\frac{2b\int_{Y_{min}}^{Q_{max}-X[1]}(\mu-Y)f[Y]dY}{\pi''[x+Y[1]]+\int_{Y_{min}}^{Y_{max}} D_{11}v_1[X[1]-x,\theta Y+(1-\theta)\mu]f[Y]dY}<0$ ■

The following version of Proposition 1 holds for the two-period problem.

PROPOSITION 5: *For the two-period problem, the optimal value function (or the Bellman function)* $v_2[X[1],Y[1]]$ *gives the profits made by the dominant firm - as a function of the volume of water in the reservoir* $X[1]$ *and the realized output of wind power* $Y[1]$ *- has the following properties:*

(i) *It is concave in* $(X[1],Y[1])$.

(ii) *It is continuously differentiable.*

(iii) G*iven* $Y[1]$, *it is increasing in* $X[1]$. *However, given* $X[1]$, *the contributions of wind power to the profitability of the dominant firm might be positive or negative.*

PROOF: (i) The concavity of $v_2[X[1],Y[1]]$ follows from the concavity of $\pi[Q]$ and the concavity of v_1.

(ii)The fact that $v_2[X[1], Y[1]]$ is continuously differentiable follows from the fact that $\pi[Q]$ and v_1 are continuously differentiable.

(iii) As for the contribution of wind power to the profitability of the dominant firm, the result might be positive or negative depending on both γ and $Y[1]$. When the volume of water in the reservoir is low, so that at the margin the marginal revenue $\pi'\big[x + Y[1]\big] > \gamma$, the revenue obtained from one more unit of wind power exceeds its purchase cost, and the profit made by the dominant firm improves with a slight rise in the output of wind power. However, when the volume of water in the reservoir and the output of wind power are not too low, so that $\pi'\big[x + Y[1]\big] < \gamma$, a rise in the output of wind energy depresses the profits of the dominant firm: at the margin, the revenue the dominant firm earns on the last unit of wind power falls short of the price it has to pay for this power unit. ■

For any given $Y[1]$, the shadow price of water in the current period is given by $D_1 v_2\big[X[1], Y[1]\big]$. To evaluate $D_1 v_2\big[X[1], Y[1]\big]$, we need to consider two possibilities: (i) $Y[1] > \mu$ and (ii) $Y[1] \leq \mu$.

Under possibility (i), Figure 3 applies for small values of $X[1]$, namely $X[1] < \xi$. The analysis in the first half of Sub-section 4.3 indicates that when $X[1] < \xi$, no water is released in the current period. The expected shadow price of water is then given by

(36) $$D_1 v_2\big[X[1], Y[1]\big] = \phi_1\big[X[1]\big],\ 0 \leq X[1] < \xi$$

For $\xi \leq X[1] < 2Q_{max} - Y[1] - Y_{min}$, Figure 2 applies, and the analysis in Sub-section 4.2 indicates that the solution is interior. Applying the envelope theorem to the expected profit maximization (18), we obtain

(37) $$D_1 v_2\big[X[1], Y[1]\big] = \frac{\partial \mathcal{L}}{\partial X[1]} = \phi_1[X[1] - x] > \phi_1\big[X[1]\big], \xi \leq X[1] < 2Q_{max} - Y[1] - Y_{min}$$

For $X[1] \geq 2Q_{max} - Y[1] - Y_{min}$, Figure 1 applies, and the analysis in Sub-section 4.1 indicates that there is an abundance of water in the reservoir for the two-period problem and a fortiori an abundance of water in the reservoir for the one-period problem. In this case, we have

(38) $D_1 v_2\big[X[1], Y[1]\big] = \phi_1\big[X[1]\big] = 0 \quad \xi < X[1] \geq 2Q_{max} - Y[1] - Y_{min}$

Under possibility (ii), Figure 3 applies for small values of $X[1]$, namely $X[1] < \eta$. The analysis in the second half of Sub-section 4.3 indicates that when $X[1] < \eta$, the reservoir is emptied in the current period. The expected shadow price of water is then given by

(39) $D_1 v_2\big[X[1], Y[1]\big] = \pi'\big[X[1] + Y[1]\big] \geq \phi_1[0] > \phi_1\big[X[1]\big] \quad 0 \leq X[1] < \eta$

For $\eta \leq X[1] < 2Q_{max} - Y[1] - Y_{min}$, Figure 2 applies, and the analysis in Sub-section 4.2 indicates that the solution is interior. Applying the envelope theorem to the expected profit maximization (18), we obtain

(40) $D_1 v_2\big[X[1], Y[1]\big] = \frac{\partial \mathcal{L}}{\partial X[1]} = \phi_1[X[1] - x] > \phi_1\big[X[1]\big]$

$\eta \leq X[1] < 2Q_{max} - Y[1] - Y_{min}$

For $X[1] \geq 2Q_{max} - Y[1] - Y_{min}$, Figure 1 applies, and the analysis in Sub-section 4.1 indicates that there is an abundance of water in the reservoir for the two-period problem and a fortiori an abundance of water in the reservoir for the one-period problem. In this case, we have

(41) $D_1 v_2\big[X[1], Y[1]\big] = \phi_1\big[X[1]\big] = 0 \quad \eta < X[1] \geq 2Q_{max} - Y[1] - Y_{min}$

We summarize the preceding discussion in the following lemma.

LEMMA 2: *For any volume of water in the reservoir* $X[1]$ *and any value of the realized output of wind power* $Y[1]$ *in the current period, we have*

$$D_1 v_2\big[X[1], Y[1]\big] \geq \phi_1\big[X[1]\big]$$

with strict inequality holding when water is neither abundant nor scarce.

Suppose that $X[1]$ is the known volume of water in the reservoir at the beginning of the current period. Before the output of wind power is known, the expected value of the shadow price of water in the reservoir is given by

(42) $$\phi_2\big[X[1]\big] = \int_{Y_{min}}^{Y_{max}} D_1 v_2[X[1], Y] f[Y] dY$$

LEMMA 3: (i) *For the two-period problem, the expected shadow price of water before the output of wind power is known is a continuous and decreasing function of the volume of water in the reservoir.*

(ii) *Furthermore, if we let* Z, *considered as a variable, denote the volume of water in the reservoir in the current period, then* $\phi_2[Z] \geq \phi_1[Z]$ *with strict inequality holding when* Z *is neither too high nor too low.*

(iii) $a - 2bY_{min} > \phi_2[0] > \phi_1[0] = a - 2b\mu$

PROOF: Part (i) of Lemma 6 follows immediately from the fact that v_2 is concave and continuously differentiable, as asserted by Proposition 5. Part (ii) of Lemma 6 follows from the definition of $\phi_2[Z]$ and Lemma 5. As for part (iii) of Lemma 6, note that when $Y[1] \geq \mu$, we have

$$D_1 v_2\big[0, Y[1]\big] = a - 2b\mu = \phi_1[0] < a - 2bY_{min}$$

and when $Y[1] < \mu$, we have

$$a - 2bY_{min} > D_1 v_2\big[0, Y[1]\big] = a - 2bY[1] > a - 2b\mu$$

Taking the expectation of $D_1 v_2[0, Y[1]]$ with respect to $Y[1]$ before it is known, we obtain (iii) of Lemma 3.■

5. THE T-PERIOD PROBLEM: T>2

The arguments used to find the solution for the two-period problem can be used to find the solution for the three-period problem, which in turn can be used to obtain the solution for the four-period problem, and so forth. In general, for any integer $T > 2$, all the results established for the two-period problem hold for the T-period problem. In particular, a version of Proposition 5 holds for the T-period problem, i.e., the Bellman function for the T-period problem $v_T[\mathrm{X}[1], \mathrm{Y}[1]]$ is concave and continuously differentiable. Also, a version of Lemma 6 holds for the T-period problem, i.e.

(43) $\phi_T[Z] \geq \phi_{T-1}[Z],\ Z \geq 0$

and that

(44) $\phi_{T-1}[0] < \phi_T[0] < a - 2bY_{min},$

where

(45) $\phi_T[Z] = \int_{Y_{min}}^{Y_{max}} D_1 v_T[X[1], Y] f[Y] dY$

denote the expected shadow price of water for the T-period problem, given that the initial volume of water in the reservoir is Z and given that the output of wind power in the first period is not yet revealed.

As in the two-period problem, to find the solution for the T-period problem, there are three possible scenarios to consider: (i) water is abundant, (ii) water is not abundant, but not scarce, and (iii) water is scarce.

Water is abundant if $X[1] \geq TQ_{max} - Y[1] - (T-1)Y_{min}$. Under this scenario, it is feasible for the dominant firm to deliver the revenue-maximizing volume of power Q_{max} to the market in each period. When $X[1]$ is in a left neighbourhood of $TQ_{max} - Y[1] - (T-1)Y_{min}$, we are under scenario (ii), and it is optimal for the dominant firm to release part of the water in its reservoir: this is the case of an interior solution. Figure 5 depicts the scenario of an interior solution for the T-period problem.

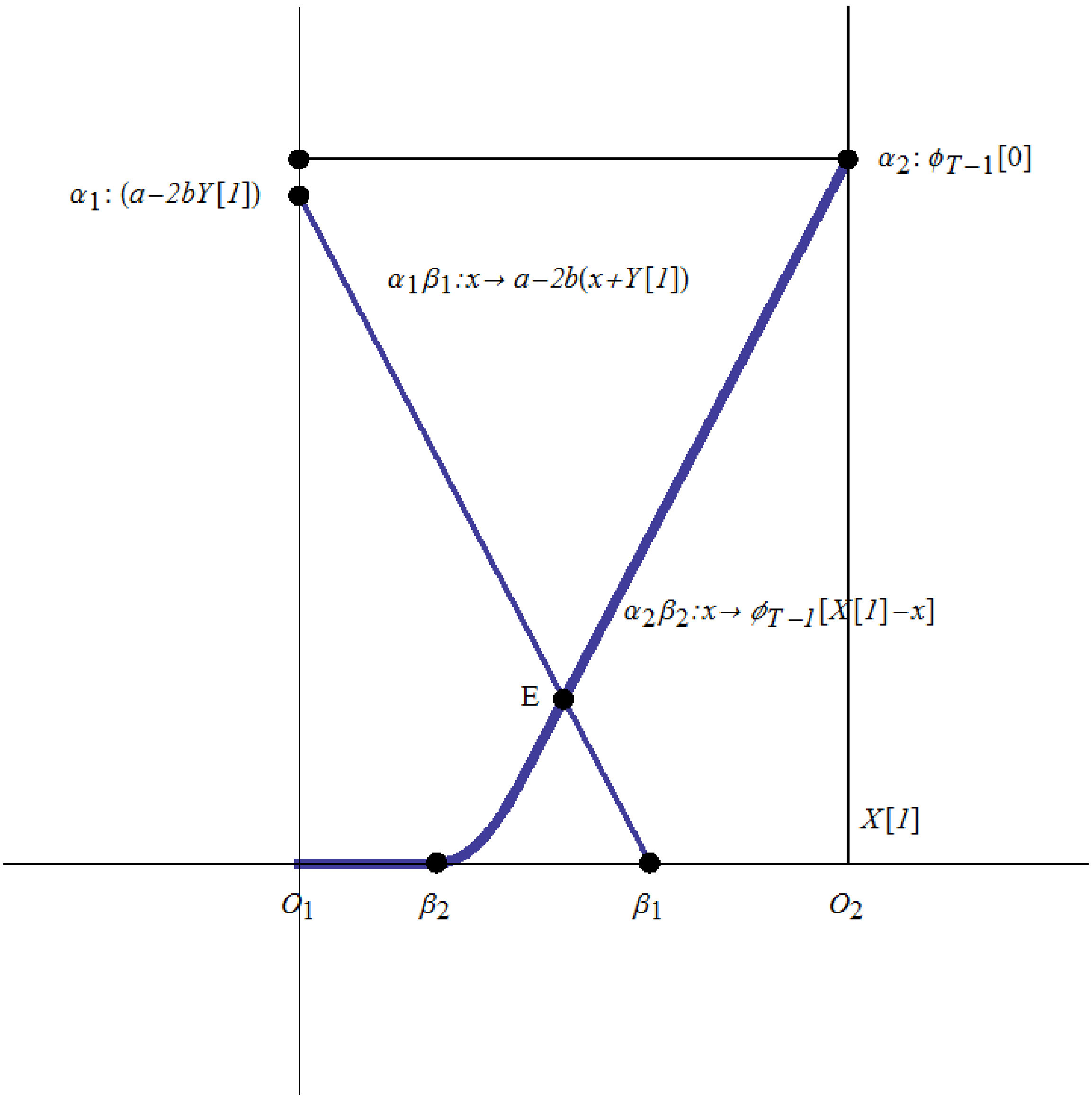

Figure 5.-- The Scenario "Water is neither abundant nor scarce: Interior solution"

If we shrink $X[1]$ further, there comes a point where the optimal solution is a corner solution. This is the scenario of scarcity, and it is optimal for the dominant firm either to empty the reservoir or not produce any hydroelectric power at all in the current period. The critical volume of water in the reservoir below which the optimal solution is a corner solution depends on $\phi_{T-1}[0]$ and $Y[1]$, can be determined as follows.

If $a - b\,Y[1] \leq \phi_{T-1}[0]$, let $\xi_T\big[Y[1]\big]$ be the value of ξ_T that solves the following equation:

(46) $a - b\,Y[1] = \phi_{T-1}[\xi_T\,]$

If $a - b\,Y[1] > \phi_{T-1}[0]$ let$\eta_T\big[Y[1]\big]$ be the value of η_T that solves the following equation:

(47) $a - b\,(\eta_T + Y[1]) = \phi_{T-1}[0]$

The following version of Proposition 2 characterizes the optimal solution for the T-period problem:

PROPOSITION 6: *Let T be a positive integer, and suppose that the dominant firm has T periods to exploit its water resources. Suppose further that $X[1]$ is the current volume of water in the reservoir and $Y[1]$ is the realized current output of wind power.*

(i) *If $a - b\,Y[1] \leq \phi_{T-1}[0]$, the optimal solution for the T-period problem can be described as follows:*

a. *For $X[1] < \xi_T\big[Y[1]\big]$, release no water in the current period.*
b. *For $\xi_T\big[Y[1]\big] \leq X[1] < TQ_{max} - Y[1] - (T-1)Y_{min}$, release a positive volume of water in the current period.*
c. *For $X[1] \geq TQ_{max} - Y[1] - (T-1)Y_{min}$, deliver to the market in each period the revenue-maximizing volume of power Q_{max} - hydroelectric and wind - in each period.*

(ii) *If $a - b\,Y[1] > \phi_{T-1}[0]$, the optimal solution for the* T-period *problem can be described as follows:*

a. *For $X[1] < \eta_T[Y[1]$ empty the reservoir in the current period.*

b. *For* $\eta_T\big[Y[1]\big] \le X[1] < TQ_{max} - Y[1] - (T-1)Y_{min}$ *release part of the volume of water in the reservoir in the current period.*

c. *For* $X[1] > TQ_{max} - Y[1] - (T-1)Y_{min}$, *deliver to the market in each period the revenue-maximizing volume of power* Q_{max} *- hydroelectric and wind - in each period.*

For the dominant firm, resource scarcity is a relative measure: whether resources are abundant or scarce depends on the volume of water damped up in the reservoir and the number of periods it has to exploit this volume of water. For a given volume of water in the reservoir, its depletion will be spread out thinner as the time horizon lengthens. The ensuing result is rising scarcity, as reflected by the upward shifting of the curve $\phi_T; Z \to \phi_T[Z], Z \ge 0$, as T rises. The following lemma gives the limiting value of the expected value of the shadow price of water at low volumes of water when the number of periods increases indefinitely.

LEMMA 4: *We have* $\lim_{T\to\infty} \phi_T[0] = a - 2bY_{min}$.

PROOF: According to (44), $\phi_T[0], T = 1,2,\ldots,$, is strictly monotone increasing and bounded above by $a - bY_{min}$.Hence it converges, and

(48) $\lim_{T\to\infty} \phi_T[0] \le a - 2bY_{min}$

We claim that (43) must hold with equality. Indeed, if

$$\lim_{T\to\infty} \phi_T[0] < a - 2bY_{min},$$

then there exists $\epsilon > 0$, such that

$$\lim_{T\to\infty} \phi_T[0] < a - 2b(Y_{min} + \epsilon)$$

Thus, when T is large enough, we will have $a - 2bY[1] > a - 2b(Y_{min} + \epsilon)$ for all $Y[1] \in [Y_{min}, Y_{min} + \epsilon]$, and this means $\phi_T[0] > \lim_{T\to\infty} \phi_T[0]$contradicting (43).■

PROPOSITION 7: *Let T be a positive integer, and suppose that the dominant firm has T periods to exploit its water resources. Suppose further that $X[1]$ is the current volume of water in the reservoir and $Y[1]$ is the realized current output of wind power. Then the volume of water that the dominant firm chooses to release in the current period is a decreasing function of T. In particular, when T is large enough, the dominant firm will choose not to exploit its water resources in the beginning.*

PROOF: The first statement of Proposition 7 follows immediately from the fact that when T rises, the curve $\phi_{T-1}: x \rightarrow \phi_{T-1}[X[1]-x], 0 \leq x \leq X[1]$ shifts upward, but the curve
$x \rightarrow \pi'\left[x+Y[1]\right], 0 \leq x \leq X[1]$ remains stationary.

As for the second statement of this proposition, note that $\phi_{T-1}[Z] > 0, \mathrm{Z} \leq (\mathrm{T}-1)\mathrm{Q}_{\max} - Y[1] - (\mathrm{T}-1)\mathrm{Y}_{\min}$ and this means that the part of this curve between O_1O_2 in Figure 5 will tend to a horizontal limiting position when $T \rightarrow \infty$. Furthermore, according to Lemma 4, this curve will be above the curve $x \rightarrow \pi'\left[x+Y[1]\right]$ for any $Y[1] > Y_{min}$, and this implies that it is optimal not to release any water in the current period. ■

6. CONCLUDING REMARKS

In our paper we analyse the impact of integrating wind power into a power system with a dominant hydro electric generator. Our paper departs from others in the literature by using a stochastic dynamic optimization model that explicitly models the intermittent nature of wind power into the hydroelectric generator's output decision. It develops a framework for optimal water release when a hydroelectric generator integrates wind power into its objective functions.

Our paper shows that when there is an abundance of water the hydroelectric generator will release the profit maximizing quantity of power and when water is neither scarce nor abundant positive amounts of water will be released. When water is scarce the release of water will be dependent the shadow and current period marginal profit of wind power price when there is no reservoir capacity. If shadow price of wind when there is no reservoir capacity is greater than the marginal profit of wind power, then it is optimal for the hydroelectric generator to release no water in the current period and it is optimal to release all the water if the opposite is true. Finally our model shows that the decision on when to release water and how much water should be released will depend on the capacity of the reservoir and the length of the time horizon. In essence, when the time horizon within the yearly cycle is long enough it will be optimal not to release any water at the beginning.

In many electricity markets, the use of hydroelectric reservoirs isn't an option. To solve the problems imposed by the randomness of wind power, independent system operators buy extra generating capacity which could be called upon when the reliability of their power systems is compromised. These services are bought and sold in the ancillary market which is a secondary market where reserve or standby capacity can be bought.

Reserves are provided by generators that can be dispatched when there is a sudden drop in supply. The sudden need for reserves requires that only fast ramp generators can provide ancillary services. Most low cost generators such as nuclear generators and hydroelectric generator are unable to provide ancillary services because they are used for base load supply or they have slow ramp speeds. For example, nuclear generators can take days to ramp up to maximum capacity. The implication is that generators that provide reserve capacity are usually high cost generators.

If an electricity market doesn't have an adequate number of fast ramp generators. The few generators that supply electricity in ancillary markets would have market power and prices in the ancillary market could be high. Apart from inadequate supply of fast ramp generators, high demand in the energy market will require the use of available fast ramp generator leaving an inadequate number of generators to supply standby services.

The growth of investments in wind turbine implies that there will be an increased need for reserve capacity in coming years. The additional reserve capacity will provide standby services in cases where wind power is unavailable. Therefore understanding the impact of uncertainty of supply and how it affect how much reserve capacity a generator is willing to supply will be important in determining cost effectiveness of wind power. Understanding this impact is the subject of our second chapter.

REFERENCES

1. Ambec, S. and Crampes C. (2012). "Electricity Production with Intermittent Sources of Energy," *Resource and Energy Economics 34(3), 319-336*
2. Bélanger, C. and Gagnon, L. (2002), “Adding wind energy to hydropower”, *Energy Policy,* 30, 1279-1284
3. Benitez, L.E. Benitez P.C. G. Cornelis Kooten C V, (2008), “The economics of wind power with energy storage”, *Energy Economics* 30 (4) 1973-1989
4. Bushnell J. (2003) “A mixed complementary model of Hydro-Thermal electricity competition in the western United States” *Operations Research* 51(1), 80-83
5. Dreyfus S.(1965), *Dynamic programming and the calculus of variations* (New York: Academic Press)
6. Garcia, A., and Alzate, J.M. (2012), Barrera J “Regulatory design and incentives for renewable energy” *Journal of Regulatory Economics* 2012 (41) 3, 315-336
7. Gil, H., Deslauriers, J.C., Dignard-Bailey, L., and Joos, G.(2006), *Integration of Wind Generation with Power Systems in Canada – Overview of Technical and Economic Impacts*, (CANMET Energy Technology Centre– Varennes, Natural Resources Canada)
8. REN21. 2011, *Renewables 2011 Global Status Report* (Paris: REN21 Secretariat)
9. Van Schaick P.L, Patton, D.B, Quarterly Report on the New York ISO Electricity Markets First Quarter 2010 *Potomac Economics Market Monitoring Unit* (May 2010).

CHAPTER TWO: THE OPERATING RESERVES IN A COMPETITIVE ELECTRICITY MARKET

1. INTRODUCTION

Electricity is generated by spinning a magnet placed at the center of a loop of wire. The energy needed to spin the magnet can be mechanical, high-pressured steam, falling water… As long as the magnet spins, an alternating current flows inside the loop of wire. The alternating current thus generated stops flowing when the magnet stops spinning. At the present time, no technology exists that can store large quantities of electricity economically. Thus, electricity must be produced and consumed at the same time, and trade in electricity always refers to a certain amount of electrical energy delivered over a specified period.

Electric flows in a network obey the laws of physics. As long as consumption is equal to output, the frequency and the interchanges remain constant, and the system is stable. However, the balance between the generation and the load is constantly perturbed by load fluctuations, by imprecise control of the outputs of generating plants, and by the sudden failures of some generating plants or of an interconnection. On a hot summer day, demand for electricity soars because of the need for air conditioning. The demand for electricity also strains the available generating capacity of the network during harsh winter nights when the electricity needed for home heating reaches its peak. In a large system, such as an electric grid, which consists of tens of thousands of components, the failure of a single component is not a rare event: transmission lines exposed to changing weather conditions might be down, and generating plants subjected to repeated changes in operating conditions might fail. All these events cause the generation in the network to fall. In an isolated electric grid, a surplus of generation over load induces a rise in the frequency, while a deficit depresses it. Because generating units are designed to operate within a relatively narrow range of frequencies, if the frequency falls below the lower end of the range, they will automatically disconnect themselves from the rest of the system for their own protection. Such a disconnection exacerbates the generation/load imbalance, causing a further drop in frequency and the ensuing disconnections of more generating units. The end result is a total system collapse, with catastrophic consequences.

To maintain the reliability of the system, there must be some stand-by generating capacity – called operating reserves – that is ready to maintain the balance of output and load at each instant. Operating reserves come in several qualities classified by how quickly the generator can respond. The highest quality of operating reserve is called "regulation" and is used continuously to handle rapid fluctuations in loads and small unintended changes in generation, not just for contingencies. It is bought separately and is often not included when operating reserves are discussed. Ten-minute spinning reserves can start responding almost instantly and deliver its full response within ten minutes. Typically the spinning reserve requirement of a system is roughly equal to the largest loss of power that could occur due to a single line or generator failure – a "single contingency." Spinning reserve (spin) is the most expensive type of reserve because a generator must be operating (spinning) to provide it. Spin is typically defined as the increase in output that a generator can provide in ten minutes. Steam units can typically ramp up (increase output) at a rate of 1% per minute, which allows them to provide spin equal to 10% of their capacity. Spin can also be provided by load that can reliably back down by a certain amount in ten minutes.

Until a few decades ago, a consumer of electricity had to buy it from a public utility that was a monopoly in the geographical area where the consumer was located. Electric utilities were vertically integrated: they generated the electricity, transmitted it from the generating plants over high-voltage transmission lines to load centers, and then distributed it to individual consumers. Electric utilities were regulated.

Under the traditional regulatory framework, an electric utility was allowed to earn a decent rate of return on its investment. In return for its monopoly status, an electric utility was responsible for ensuring an adequate and reliable supply of power within its own territory.

In liberalized electricity markets, the supply of energy and operating reserves are subject to market forces. It is expected that prices would provide the right signal for efficient investments in generation capacity as well as in operating reserves. The structure of a market for operating reserves is that of a monopsonist (the ISO) conducting an auction to obtain the operating reserves needed by the system from many competitive suppliers. To secure the co-operation of the suppliers, the ISO can either calculate the opportunity cost of holding capacity as reserves from the real-time energy price or ask generators to guess this value and include it in their bids. In some markets, such as the PJM (Pennsylvania-New Jersey-Maryland) market, the retailers – also known as LSE's (Load Serving Entities) – are required to own or purchase capacity resources greater than or equal to their expected peak load plus a reserve margin. The payments received by a generator for offering resources in the market for operating reserves consist of a price for committing her resources and a price for the energy when the resources are called to generate in the real-time energy market, with the latter price being the prevailing price in the real-time energy market.

In this paper, we present a model of operating reserves in a competitive electricity market. The main purpose of operating reserves is to moderate price spikes. In the model, there is a day-ahead market for operating reserves in which electricity producers sell call options that involve part of their generating capacity, with the remaining capacity being exploited to produce output for sale on the real-time energy market the following day. On the real-time market, there is uncertainty in the supply because some generating plants might fail at some instant during the following day. If such an event occurs, and if the failure is massive, there will be a price spike in the absence of reserves. Thus, consumers of electricity, being risk averse, will be willing to purchase call options on the market for operating reserves to lessen the negative impact of a price spike in the real-time market. For an electricity producer, she can offer part of her capacity under the form of call options and part of her capacity in the real-time market. The actions taken by consumers and producers of electricity depend on their expectations about the prices in the market for energy as well as the market for operating reserves, and our model takes into account the interdependence of these two markets. In the early years of liberalized electricity markets, energy and reserves were traded sequentially in separate markets, with the sequence determined by the speed of response of the service. According to this trading philosophy, the market for primary reserves was cleared first, to be followed by the market for secondary reserves, and then the market for energy. The rationale for this trading process was that the resources which were not successful in one market could be offered in the next market in the sequence where the performance standard is less demanding. The economic logic of this approach is clearly faulty and has been abandoned.[5] Some researchers, such as Just and Weber (2008) and Creti and Fabra (2003), have formalized the link between the market for operating reserves and the real-time energy market in an effort to deepen the understanding of the prices in these two inter-dependent markets. Our paper shares the same objectives as the works of these researchers in formalizing the interdependence between these two markets, but differs from their works in several aspects. In traditional studies on the market for operating reserves, the market demand curve for electrical energy is assumed to be vertical, i.e., market demand for electrical energy is perfectly inelastic. It is the complete price insensitivity of market demand – according to Crampton, Ockenfels, and Stoft (2013) – that justifies the need for the reserve market. Furthermore, traditional studies concentrate mostly on the demand side as the force that drives the reserve market although market supply is also inelastic as generating capacity becomes scarce. The model we formulate in this paper diverges from the approach taken by the conventional literature in several aspects. First, we concentrate on the possibility of generating plant failure in real time – not on demand uncertainty – as the main

[5] See Oren (2002).

factor that drives the reserve market. Second, unlike authors, such as Just and Weber (2008) and Creti and Fabra (2002), who model the demand for electricity in the real-time market as being perfectly inelastic,[6] we allow market demand to be highly, but not perfectly, inelastic so that price spikes might occur when there is a serious mismatch between supply and demand. Third, we model uncertainty on the supply side – not on the demand side, which is the approach taken by those researchers – as arising from generating plant failure in real time. Fourth, like those researchers, we also adopt a two-price approach – one price for committing resources on the reserve market and one price for the energy produced when the reserves are called to produce in the real-time market – but in our model the energy price paid to the reserves when they are called to produce in the real-time energy market is the strike price named by generators in the call options they sell, not the energy price that prevails in the real-time energy market that is assumed in the conventional literature. In our attempt to achieve our objectives, we have model explicitly the supply and demand in both markets as well as the possibility of generating plant failure, and the equilibrium concept that we use is that of a rational expectations equilibrium.

The paper is organized as follows. Because the model is long, it is developed step by step in several sections – from Section 2 to Section 4 and culminating in Section 6 with the definition of the rational expectations equilibrium. The existence and uniqueness of the rational expectations equilibrium is established in Section 6. Some properties of the rational expectations equilibrium are given in Section 7. Section 8 contains some concluding remarks.

2. PREFERENCES AND TECHNOLOGY

In the model, time is continuous and denoted by t. The model has two periods. In the first period, trading in call options for operating reserves is conducted. This is a day-ahead market, which – for simplicity – is assumed to occur at time $t = -1$. A call

[6] When demand is perfectly inelastic, a price cap must be imposed to maintain the internal consistency of the model.

option – when exercised – allows its owner to obtain one unit of electricity from the electricity producer who issued it. In the second period, trading in electrical energy takes place in real time, with the day ahead being the time interval $[0,1]$. Because the day ahead is not a long period, we ignore demand uncertainty. However, at any instant $t \in [0,1]$, there is always a chance that some generating plants or some inter-connections might fail, reducing the system's output at that instant. It is the possible occurrence of such an event that drives the behavior of producers and consumers in the model.

2.1. Preferences and Utility Maximization

At each instant, a consumer consumes two goods: electricity and the numéraire. Her preferences are assumed to be quasi-linear. Furthermore, consumers differ from each other according to their preferences for electricity consumption. More precisely, the preferences for electricity consumption – relative to the numéraire – of a consumer are represented by a parameter, say θ, and her utility function assumes the following quasi-linear form:

(1) $$u[x, m, \theta] = \theta\phi[x] + m\,, \qquad \left(0 < \underline{\theta} \leq \theta \leq \overline{\theta}\right).$$

In (1), x is her consumption of electricity; m is her consumption of the numéraire; and $\phi[x]$ is a function called the *base sub-utility function* for electricity consumption. As specified by (1), the sub-utility function for electricity consumption is a scaled – up or down – version of the base sub-utility function, and consumers' sub-utility functions of electricity consumption differ from each other by a scaling factor $\theta, \theta \in \left[\underline{\theta}, \overline{\theta}\right]$, where $\underline{\theta}$ and $\overline{\theta}$ are two parameters satisfying $0 \leq \underline{\theta} < \overline{\theta}$. The parameter θ represents the intensity of her preferences for electricity consumption. We shall let $h[\theta], \underline{\theta} \leq \theta \leq \overline{\theta},$ be the density function that characterizes the distribution of consumers of various types. That is, $h: \theta \to h[\theta], \underline{\theta} \leq \theta \leq \overline{\theta},$ is a non-negative continuous function with $\int_{\underline{\theta}}^{\overline{\theta}} h[\theta] d\theta = 1$.

In what follows, we shall assume that the base sub-utility function assumes the following form:

(2) $\phi[x] = \frac{x^{1-\frac{1}{\gamma}}}{1-\frac{1}{\gamma}}$,

where γ is a parameter satisfying $0 < \gamma < 1$. The functional form thus assumed by the base sub-utility function for electricity consumption leads to an iso-elastic individual demand curve for electricity which is quite in-elastic, with γ being the price elasticity of demand.[7]

In the absence of a market for operating reserves, a consumer meets her demand for electricity at each instant on the real-time energy market. Under such a scenario, the demand for electricity in real time by a consumer of type θ can be found by solving the following utility maximization problem:

(3) $max_x(\theta\phi[x] + y[\theta] - px)$.

In (2), $y[\theta]$ is her income at each instant, and p is the prevailing price of electrical energy.

The following first-order condition characterizes the solution of (3):

(4) $\theta\phi'[x] - p = 0$,

from which we obtain the following individual demand curve:[8]

(5) $x[p, \theta] = \left(\frac{p}{\theta}\right)^{-\gamma}$,

which can also be expressed under the following inverse form:

---

[7] A more general functional form for $\phi[x]$ makes the exposition much more burdensome without any gain in economic insight.

[8] Using the a data point collected on the electricity consumption of a consumer, and using the price that the consumer paid for 1Kwh from that data point, one can infer the value of θ.

(6) $p[x,\theta] = \theta x^{-\frac{1}{\gamma}}$.

The consumer's indirect utility function is given by

(7) $v[p,\theta] = \theta\phi\big[x[p,\theta]\big] + y[\theta] - px[p,\theta] = -\frac{\theta^{\gamma}}{1-\gamma}p^{1-\gamma} + y[\theta]$.

Let $X[p]$ be the market demand for electricity when p is the prevailing price of electrical energy. The market demand for electricity can be obtained by horizontally summing the demand for electricity of all the consumers; that is,

(8) $X[p] = \int_{\underline{\theta}}^{\overline{\theta}} x[p,\theta]h[\theta]d\theta = \int_{\underline{\theta}}^{\overline{\theta}} \left(\frac{p}{\theta}\right)^{-\gamma} h[\theta]d\theta = \Theta p^{-\gamma}$,

where we have let $\Theta = \int_{\underline{\theta}}^{\overline{\theta}} \theta^{\gamma}h[\theta]d\theta$. Using (8), we obtain the following expression for the inverse market demand for electricity:

(9) $p[X] = \Theta^{\frac{1}{\gamma}}X^{-\frac{1}{\gamma}}$.

2.2. Failure of Generating Plants

Let K denote the total generating capacity of the market, which we take to be a positive real number. For simplicity, we assume that each generating plant has a capacity of 1 and that an electricity producer owns exactly one generating plant. It follows from these simplifying assumptions that the number of electricity producers and the number of electric generating plants both have the cardinality of a continuum of size K. Also, we assume that all electricity producers have the same constant marginal cost – normalized to 0 – and that they are risk neutral.

At time $t = -1$, there is a day-ahead market for operating reserves. Certainly, one can also consider another forward market – the day-ahead market for electrical energy. We choose not to do so because we want to concentrate on the market for operating reserves. Another reason concerns the internal consistency of the model.

Indeed, if both markets are considered, then a consumer can avoid uncertainty completely by being active in both markets. However, in the event that some generating plants break down, the electricity producers who own these plants will not be able to honor the forward contracts for electrical energy they have signed because the system is closed. The way out for this inconsistency is to make the electric grid open so that electricity can be imported from the outside world.[9] Our model thus makes it possible for price spikes to occur, and a producer whose generating plant breaks down in the real time day-ahead market will not be able to earn any revenues after her generating plant has failed.

In the day-ahead market for operating reserves, call options are bought and sold, with each call option providing 1 unit of electric power when exercised. A call option for operating reserves can be represented by a pair (q, p^+), with q being the price of the option and p^+ the strike price. The market supply of and demand for operating reserves are denoted, respectively, by Y^+ and X^+.

Given that K is the generating capacity of the market, and given the market supply of operating reserves Y^+, the market supply of electrical energy on the day-ahead market is $Y = K - Y^+$. A subset of the Y generating plants – when they are running in real time during the time interval $0 \leq t \leq 1$ to generate electricity – might breakdown, and that such an event might occur at most once. We model such an event by assuming that it has the following density function:

(10) $$f_Y[Z, t] = \epsilon \frac{\lambda e^{-\lambda Z}}{1 - e^{-\lambda Y}},$$

where $\epsilon, 0 < \epsilon < 1$, and $\lambda > 0$ are two parameters. Here $f_Y[Z, t]$ is the density of the random event that Z generating plants, $0 \leq Z \leq Y$, break down at time $t, 0 < t \leq 1$, given that Y generating plants have been running on the real-time energy market.

According to the functional form specified by (10), the marginal density for the time a failure occurs, say τ, is uniformly distributed in the time interval $[0,1]$ and is given by

[9] This is the case considered by Creti and Fabra (2003).

(11) $f_Y[\tau] = \epsilon,$ $(0 < \tau \leq 1),$

while the conditional density for the number of generating plants breaking down, given that the breakdown has occurred, is given by

(12) $f_Y[Z|\tau] = \frac{\lambda e^{-\lambda Z}}{1-e^{-\lambda Y}},$ $(0 \leq Z \leq Y).$

It can be seen from the conditional density (12) that the higher the value of Z, the lower the probability that Z generating plants will fail when a breakdown occurs.

We shall assume that all the generating plants running during the time interval $0 \leq t \leq 1$ have the same chance of breaking down. Thus, if Z generating plants fail, then any one among the Y generating plants currently running will have a probability of $\frac{Z}{Y}$ for breaking down. Also, to keep the model from becoming unnecessarily complicated, we shall assume that a spinning generating plant – when instructed by the ISO (Independent System Operator) to inject electrical energy into the system – will not fail. One justification for this assumption is that given the short time interval of one day, the probability that part of the system breaks down and some spinning generating plants also break down is negligible.

3. EQUILIBRIUM IN THE REAL-TIME ENERGY MARKET

In this sub-subsection, we analyze the equilibrium in real time after some generating plants have failed, given the choices made by electric producers and consumers in the day-ahead market for operating reserves.

Consider a consumer of type θ. Let x^+ be the number of call options she bought in the day-ahead market for operating reserves. Now when a consumer exercises a call option she actually has to pay the effective price $q + p^+$ (the price of the option plus the strike price). Thus, the individual will not buy more than $x[q + p^+, \theta]$ call options because the marginal utility yielded by a call option is then less than the effective price of a call option. If we let X^+ denote the total number of call options bought in the day-ahead market for operating reserves, then

(13) $X^+ \leq \int_{\underline{\theta}}^{\overline{\theta}} x[q+p^+,\theta]h[\theta]d\theta = X[q+p^+] < X[p^+]$.

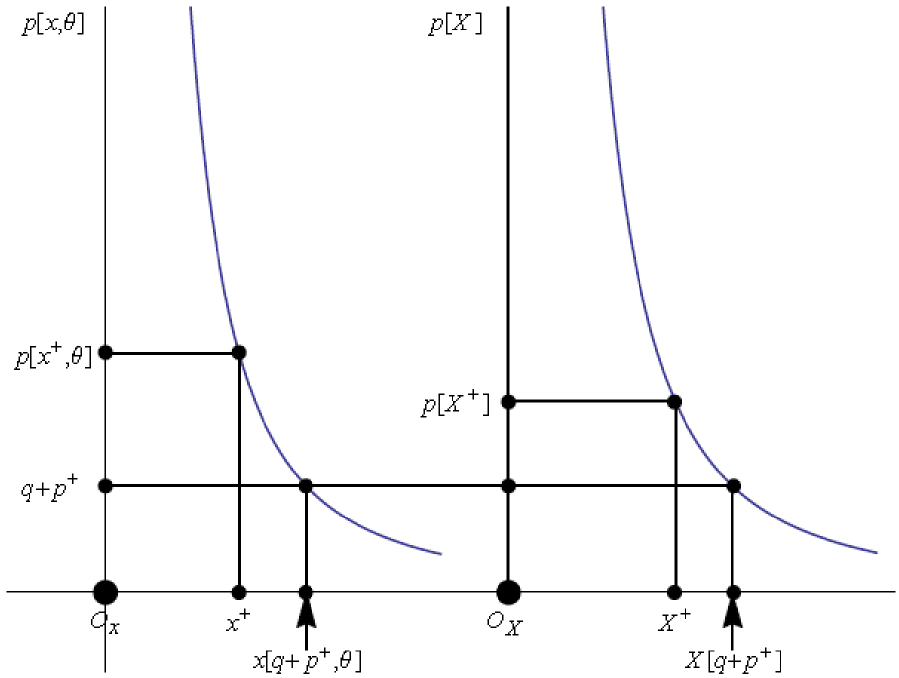

Figure 1. An upper bound on the number of call options bought by a consumer of type θ: $x^+ \leq x[q+p^+,\theta]$, and an upper bound on the market demand for call options: $X^+ \leq X[q+p^+]$

An upper bound on the number of call options bought by a consumer and an upper bound on the market demand for call options in the day-ahead market for operating reserves are depicted in Figure 1. In Figure 1, the individual demand curve for electricity and the market demand curve for electricity – in the absence of operating reserves – are drawn side by side, with O_x and O_X as the origins of the axes for these two curves, respectively.

If Y^+ is the total number of options bought and sold in the day-ahead market for operating reserves, then when the real-time market for energy opens at time $t = 0$, the market supply of electrical energy is given by

(14) $Y = K - Y^+$.

We presume that as long as no generating plant currently running to produce electricity has failed, no consumer will exercise her options. Under this presumption, the price of electricity in real time and before any generating plant breaks down is equal to $p[Y]$ and satisfies the following condition:

(15) $p[Y] < p^+$.

It follows from the presumption represented by (15) that $X[p^+] < Y$. We have also argued that $X^+ < X[q + p^+]$. Hence

(16) $Y^+ < X[q + p^+] < X[p^+] < Y$.

The chain of inequalities (16) is depicted in Figure 2.

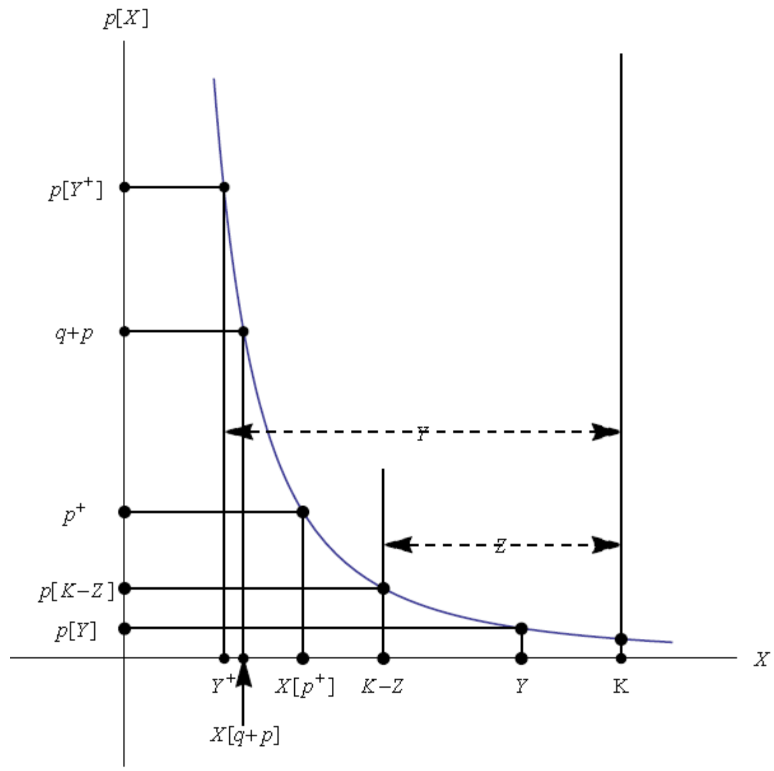

Figure 2. The critical number of failing generating plants: $K - X[p^{+}]$

If Z generating plants break down at time τ, then the number of generating plants that are still running is $Y - Z$. If no consumer chooses to exercise any of her options after the breakdown, then the owners of the Y^{+} spinning plants are no longer obliged to keep them spinning as operating reserves, and will then run them to produce electricity for profits in the real-time energy market. This follows from the assumption that the breakdown occurs at most once, and that no consumer chooses to exercise any of her options after the breakdown. Under this scenario, the total supply of energy in real time after the breakdown will be $Y - Z + Y^{+} = K - Z$, with $p[K - Z]$ as the ensuing price of electricity after the breakdown. If some of the options are exercised, then the spinning reserves that are not called will be supplied on the real-time market for profits. The total output produced by (i) the spinning reserves that are called, (ii) the generating plants that did not fail, and (iii) the spinning reserves that are not called is equal to $K - Z$. Under this scenario, the electricity price that prevails on the real-time market to equate demand (not including the options that are exercised) and supply after the breakdown will be $p[K - Z]$.

Note that the energy price that prevails in the real-time market after some generating plants have broken down might be higher or lower than the energy price that prevailed before the break-down, depending on the number of generating plants that have broken down. Indeed, if the number of generating plants that fail is small, then we have $K - Z > Y$, and $p[K - Z] < p[Y]$. On the other hand, if Z is close to Y, i.e., if a major break-down occurs, then the energy price after the breakdown will be close to $p[Y^+]$, which will be higher than $p[Y]$, the electricity price that prevailed before the breakdown.

The following proposition deals with the number of call options that will be exercised after some generating plants have broken down.

PROPOSITION 1: *Suppose that Z generating plants break down at time τ.*

(a) *If $Z \leq K - X[p^+]$, then no option will be exercised.*

(b) *If $Z > K - X[p^+]$, then all of the options bought and sold in the day-ahead market for operating reserves will be exercised, and the electricity price will take an upward jump at the time the breakdown occurs.*

PROOF: The price of electricity on the real-time energy market after the break-down is equal to $p[K - Z]$. Thus, if $Z \leq K - X[p^+]$, then $p[K - Z] \leq p^+$, and no option will be exercised. On the other hand, if $Z > K - X[p^+]$, then $p[K - Z] > p^+$, and it is not optimal for a consumer who holds some options not to exercise all of them. That is, under this scenario all the options will be exercised.

■

4. A CONSUMER'S DEMAND FOR CALL OPTIONS

Consider a consumer of type θ, and suppose that she anticipates a market supply of operating reserves equal to Y^+ and that she purchases x^+ call options in the day-ahead market for operating reserves. Let $\omega[x^+, q, p^+, Y^+, \theta]$ denote the net cumulative utility she obtains at the end of the day that she can expect at the time the call options are purchased. The value of $\omega[x^+, q, p^+, Y^+, \theta]$ can be computed as follows.

First, if no generating plant fails, then the realized cumulative utility she obtains at the end of the day is given by $v[p[Y], \theta]$, and this event has probability $(1 - \epsilon)$.

Second, if Z generating plants fail at time $\tau, 0 < \tau \leq 1$, there are two possible scenarios to consider: $Z \leq K - X[p^+]$ and $K - X[p^+] < Z \leq Y$. Under the scenario $Z \leq K - X[p^+]$, no consumer will exercise any of her options, and the realized cumulative utility that a consumer of type θ obtains at the end of the day is given by

(17) $\tau v[p[Y], \theta] + (1 - \tau) v[p[K - Z], \theta]$. $\qquad (Z \leq K - X[p^+])$

Under the scenario $K - X[p^+] < Z \leq Y$, all of the Y^+ options bought and sold on the day-ahead market for operating reserves will be exercised. A consumer's purchase of electricity on the real-time market after some generating plants have failed depends on the number of call options she bought in the day-ahead market for operating reserves and the number of generating plants that have failed. In computing her purchase of electricity on the real-time energy market after the failure, there are two possibilities to consider: (i) $x^+ \leq x[p[Y^+], \theta]$ and (ii) $x^+ > x[p[Y^+], \theta]$.

Under possibility (i), we have the $p[x^+, \theta] \geq p[Y^+]$, i.e., the price of electricity that prevails in the real-time energy market after the break-down if all the running generating plants fail is below the consumer's marginal utility of electricity consumption at the consumption level x^+, as depicted in Figure 3.

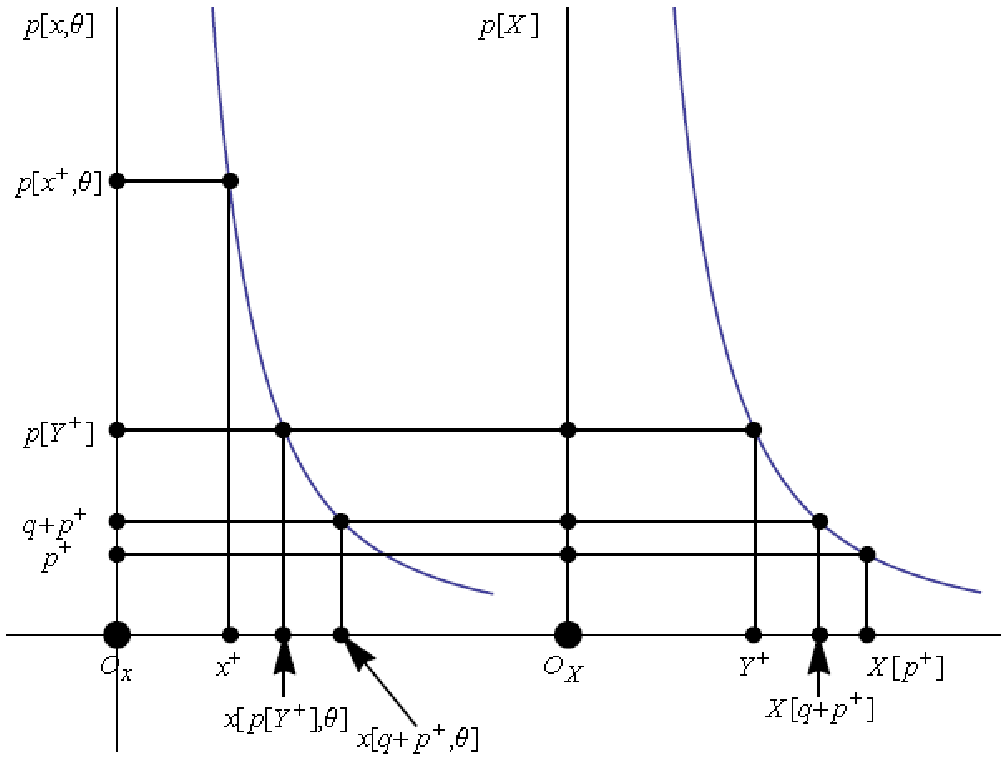

Figure 3. Possibility (i): $x^+ \leq x[p[Y^+],\theta]$

Under the scenario $K - X[p^+] < Z \leq Y$ and possibility (i), the cumulative utility she obtains at the end of the day is given by

(18) $\tau v[p[Y],\theta] + (1-\tau)(v[p[K-Z],\theta] + (p[K-Z] - p^+)x^+)$,

$(x^+ \leq x[p[Y^+],\theta], K - X[p^+] <$

$Z \leq Y)$.

Under possibility (ii), we have the $p[x^+,\theta] < p[Y^+]$, i.e., the price of electricity that prevails in the real-time energy market after the break-down if all the running generating plants fail is above the consumer's marginal utility of electricity consumption at the consumption level x^+, as depicted in Figure 4.

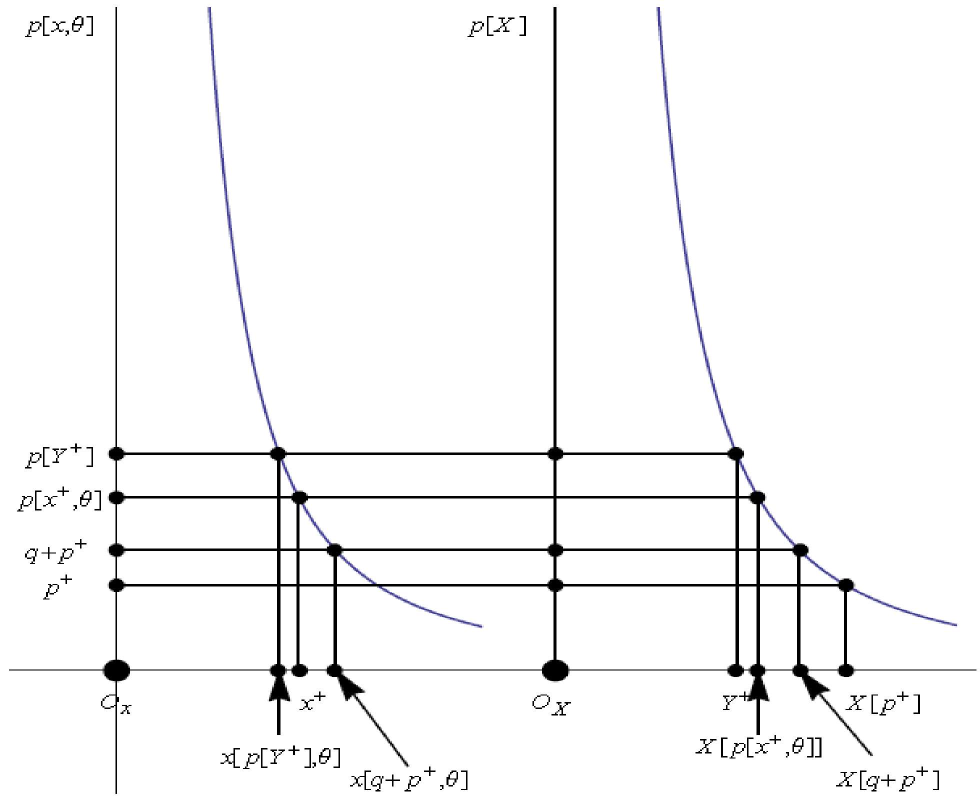

Figure 4. Possibility (ii): $x^+ > x[p[Y^+],\theta]$

Under the scenario $K - X[p^+] < Z \leq Y$ and possibility (ii), the cumulative utility she obtains at the end of the day is given by

(19) $\tau v[p[Y],\theta] + (1-\tau)(v[p[K-Z],\theta] + (p[K-Z] - p^+)x^+),$

$(x^+ > x[p[Y^+],\theta], K - X[p^+] < Z \leq K - X[p[x^+,\theta]]),$

and

(20) $\tau v[p[Y],\theta] + (1-\tau)(\theta\phi[x^+] - p^+x^+ + y[\theta]),$

$(x^+ > x[p[Y^+],\theta], K - X[p[x^+,\theta]] < Z \leq Y).$

Given (q, p^+) and Y^+, the cumulative utility – net of the cost of purchasing x^+ call options – that a consumer of type θ can expect – at the time the options are purchased – to obtain at the end of the day is as follows:

For $x^+ \leq x[p[Y^+], \theta]$, we have

(21) $$\omega[x^+, q, p^+, Y^+, \theta] = -qx^+ + (1-\epsilon)v[p[Y], \theta]$$

$$+\epsilon \int_0^1 \left(\tau v[p[Y], \theta] + (1 - \tau) \begin{pmatrix} \int_0^{K-X[p^+]} v[p[K-Z], \theta] \frac{\lambda e^{-\lambda Z}}{1-e^{-\lambda Y}} dZ + \\ \int_{K-X[p^+]}^{Y} \left(v[p[K-Z], \theta] + x^+(p[K-Z] - p^+)\right) \frac{\lambda e^{-\lambda Z}}{1-e^{-\lambda Y}} dZ \end{pmatrix} \right) d\tau$$

$$= -qx^+ + \left(1 - \frac{\epsilon}{2}\right) v[p[Y], \theta] + \frac{\epsilon}{2} \int_0^{K-X[p^+]} v[p[K-Z], \theta] \frac{\lambda e^{-\lambda Z}}{1-e^{-\lambda Y}} dZ$$

$$+ \frac{\epsilon}{2} \int_{K-X[p^+]}^{Y} \left(v[p[K-Z], \theta] + x^+(p[K-Z] - p^+)\right) \frac{\lambda e^{-\lambda Z}}{1-e^{-\lambda Y}} dZ.$$

For $x^+ > x[p[Y^+], \theta]$, we have

(22) $$\omega[x^+, q, p^+, Y^+, \theta] = -qx^+ + (1-\epsilon)v[p[Y], \theta]$$

$$+\epsilon\int_0^1\left(\tau v[p[Y],\theta]+(1-\tau)\left(\begin{array}{c}\int_0^{K-X[p^+]} v[p[K-Z],\theta]\frac{\lambda e^{-\lambda Z}}{1-e^{-\lambda Y}}dZ+\\ \int_{K-X[p^+]}^{K-X[\theta\phi'[x^+]]}\left(v[p[K-Z],\theta]+x^+(p[K-Z]-p^+)\right)\frac{\lambda e^{-\lambda Z}}{1-e^{-\lambda Y}}dZ\\ +\int_{K-X[\theta\phi'[x^+]]}^{Y}(\theta\phi[x^+]-p^+x^++y[\theta])\frac{\lambda e^{-\lambda Z}}{1-e^{-\lambda Y}}dZ\end{array}\right)\right)d\tau$$

$$=-qx^++\left(1-\frac{\epsilon}{2}\right)v[p[Y],\theta]+\frac{\epsilon}{2}\int_0^{K-X[p^+]} v[p[K-Z],\theta]\frac{\lambda e^{-\lambda Z}}{1-e^{-\lambda Y}}dZ$$

$$+\frac{\epsilon}{2}\left(\begin{array}{c}\int_{K-X[p^+]}^{K-X[\theta\phi'[x^+]]}\left(v[p[K-Z],\theta]+x^+(p[K-Z]-p^+)\right)\frac{\lambda e^{-\lambda Z}}{1-e^{-\lambda Y}}dZ\\ +\int_{K-X[\theta\phi'[x^+]]}^{Y}(\theta\phi[x^+]-p^+x^++y[\theta])\frac{\lambda e^{-\lambda Z}}{1-e^{-\lambda Y}}dZ\end{array}\right).$$

The demand for call options by a consumer with medium preferences for electricity consumption is obtained by solving the following maximization problem:

(23) $max_{x^+\geq 0}\omega[x^+,q,p^+,Y^+,\theta]$.

To solve the maximization problem stated by (23), we study how $\omega[x^+,q,p^+,Y^+,\theta]$ varies when x^+ rises from 0 to $x[p[Y^+],\theta]$, and then from $x[p[Y^+],\theta]$ to $x[q+p^+,\theta]$.[10]

For $0\leq x^+\leq x[p[Y^+],\theta]$, differentiating (21) with respect to x^+, we obtain

[10] Recall our discussion that the number of options purchased by a consumer of type θ never exceeds $x[q+p^+,\theta]$

(24) $\frac{\partial}{\partial x^+}\omega[x^+,q,p^+,Y^+,\theta] = -q + \frac{\epsilon}{2}\int_{K-X[p^+]}^{Y}(p[K-Z]-p^+)\frac{\lambda e^{-\lambda Z}}{1-e^{-\lambda Y}}dZ,$

$(0 \le x^+ \le x[p[Y^+],\theta])$.

For $x^+ > x[p[Y^+],\theta]$, differentiating (22) with respect to x^+, we obtain[11]

(25) $$\frac{\partial}{\partial x^+}\omega[x^+,q,p^+,Y^+,\theta] = -q + \frac{\epsilon}{2}\begin{pmatrix}\int_{K-X[p^+]}^{K-X\left[\theta\phi'[x^+]\right]}(p[K-Z]-p^+)\frac{\lambda e^{-\lambda Z}}{1-e^{-\lambda Y}}dZ \\ +(\theta\phi'[x^+]-p^+)\int_{K-X\left[\theta\phi'[x^+]\right]}^{Y}\frac{\lambda e^{-\lambda Z}}{1-e^{-\lambda Y}}dZ\end{pmatrix}.$$

LEMMA 1: *The function* $x^+ \to \omega[x^+,q,p^+,Y^+,\theta], 0 \le x^+ \le x[q+p^+,\theta]$, *is linear in the interval* $0 \le x^+ \le x[p[Y^+],\theta]$. *It is strictly concave in the interval* $x[p[Y^+],\theta] \le x^+ \le x[q+p^+,\theta]$.

PROOF: The first statement of Lemma 1 follows from the fact that the right-hand side of (24) does not depend on x^+. To prove the second statement of Lemma 1, differentiate (25) with respect to x^+ to obtain

(26) $$\frac{\partial}{\partial x^+}\frac{\partial}{\partial x^+}\omega[x^+,q,p^+,Y^+,\theta] = \frac{\epsilon}{2}\theta\phi''[x^+]\int_{K-X\left[\theta\phi'[x^+]\right]}^{Y}\frac{\lambda e^{-\lambda Z}}{1-e^{-\lambda Y}}dZ < 0.\ \blacksquare$$

The following lemma follows immediately from Lemma 1.

LEMMA 2: (i) *If*

(27) $$-q + \frac{\epsilon}{2}\int_{K-X[p^+]}^{Y}(p[K-Z]-p^+)\frac{\lambda e^{-\lambda Z}}{1-e^{-\lambda Y}}dZ < 0,$$

then no consumer will purchase any call options in day-ahead the market for operating reserves.

[11] The derivation of (25) is given in Appendix A.

(ii) *On the other hand, if inequality (27) is reversed, the number of call options that a consumer of type* θ *purchases in the day-ahead market for operating reserves is higher than* $x[p[Y^+],\theta]$.

(iii) *When (27) becomes an equality, the number of call options that a consumer of type* θ *purchases in the day-ahead market for operating reserves can assume any value between 0 and* $x[p[Y^+],\theta]$.

In what follows, let $x^+[q,p^+,Y^+,\theta]$ be the demand for options by a consumer of type $\theta, \theta \in [\underline{\theta},\overline{\theta}]$. To be more definite, when the maximization problem (23) has multiple solutions, i.e., when (27) becomes an equality, we shall assume that the consumer chooses the solution that involves the greatest number of options possible, i.e., we set $x^+[q,p^+,Y^+,\theta] = x[p[Y^+],\theta]$.

5. THE RATIONAL EXPECTATIONS EQUILIBRIUM

Let q be the price of a call option and p^+ be the strike price. Also, let Y^+ be the expectations of a consumer concerning the market supply of operating reserves. Given (q,p^+,Y^+), the market demand for options in the day-ahead market for operating reserves is given by

(28) $$X^+[q,p^+,Y^+] = \int_{\underline{\theta}}^{\overline{\theta}} x^+[q,p^+,Y^+,\theta]h[\theta]d\,\theta.$$

PROPOSITION 2: *If the consumers' expectations concerning the market supply of call options are correct, i.e., if* $X^+[q,p^+,Y^+] = Y^+$*, then (27) must be an equality; that is,*

(29) $$-q + \frac{\epsilon}{2}\int_{K-X[p^+]}^{K-Y^+}(p[K-Z]-p^+)\frac{\lambda e^{-\lambda Z}}{1-e^{-\lambda Y}}dZ = 0.$$

PROOF: Indeed, if the left-hand side of (29) is negative, then $X^+[q,p^+,Y^+] = 0$ according to Lemma 2. (i), which is certainly less than the anticipated market supply of operating reserves. On the other hand, if the left-hand side of (29) is positive, then according to Lemma 2.(ii), we have $x^+[q,p^+,Y^+,\theta] > x[p[Y^+],\theta]$ for all θ, with the ensuing consequence that $X^+[q,p^+,Y^+] > Y^+$, contradicting the hypothesis that expectations are correct. ■

Given (q,p^+,Y^+), the expected revenues earned by a generating plant that serves as operating reserves are given by

(30)

$$\pi^+[q,p^+,Y^+] =$$

$$q + \int_0^1 \epsilon(1-\tau)\left(\int_0^{K-X[p^+]} p[K-Z]\frac{\lambda e^{-\lambda Z}}{1-e^{-\lambda Y}}dZ + \int_{K-X[p^+]}^{Y} p^+\frac{\lambda e^{-\lambda Z}}{1-e^{-\lambda Y}}dZ\right)d\tau$$

$$= q + \frac{\epsilon}{2}\left(\int_0^{K-X[p^+]} p[K-Z]\frac{\lambda e^{-\lambda Z}}{1-e^{-\lambda Y}}dZ + \int_{K-X[p^+]}^{Y} p^+\frac{\lambda e^{-\lambda Z}}{1-e^{-\lambda Y}}dZ\right).$$

Given (q,p^+,Y^+), the expected revenues earned by a generating plant that sells its output in the real-time market are given by

(31)

$$\pi[q,p^+,Y^+] =$$

$$p[Y](1-\epsilon) + \int_0^1 \epsilon\left(\tau p[Y] + (1-\tau)\int_0^Y \left(1-\frac{Z}{Y}\right)p[K-Z]\frac{\lambda e^{-\lambda Z}}{1-e^{-\lambda Y}}dZ\right)d\tau$$

$$= p[Y]\left(1-\frac{\epsilon}{2}\right) + \frac{\epsilon}{2}\int_0^Y \left(1-\frac{Z}{Y}\right)p[K-Z]\frac{\lambda e^{-\lambda Z}}{1-e^{-\lambda Y}}dZ.$$

DEFINITION: A list (q,p^+,Y^+), with $0 \le Y^+ \le K$, is said to constitute a rational expectations equilibrium if the following two conditions are satisfied:

(32) $$-q + \frac{\epsilon}{2}\int_{K-X[p^+]}^{K-Y^+}(p[K-Z]-p^+)\frac{\lambda e^{-\lambda Z}}{1-e^{-\lambda Y}}dZ = 0,$$

(33) $$\pi^+[q,p^+,Y^+] = \pi[q,p^+,Y^+].$$

The condition (32), which is nothing than a rewritten version of (29), asserts that the consumers' expectations on the supply of operating reserves on the day-ahead market are correct, according to Proposition 2. As for (33), it asserts that in equilibrium generating resources earn the same expected revenues in both markets – the market for operating reserves and the real-time market for energy.

Note that the rational-expectations equilibrium involves three endogenous variables q, p^+, and Y^+. However, it is only defined by two equilibrium conditions, one for equalizing supply and demand in the day-ahead market for operating reserves, and one for making a generator indifferent between offering her resources in the reserve market and offering her resources in the real-time energy market. The system is thus under-identified. The problem lies in the fact that we have assigned two prices – the option price and the strike price – to a commodity (an option). One possible solution for the under-identified problem is to set the strike price equal to the price of energy on the real-time market.[12] However, doing so would remove the incentive for consumers to buy call options with the intention of exercising them when there is a price spike after a breakdown. Our solution for the under-identified problem is to allow the option price and the strike price to go up and down together while remaining in a fixed proportion relative to each other, say $q = \mu p^+$, where $0 < \mu < 1$ is a given parameter. Under this specification, (32) and (33) become, respectively,

$$(34) \quad -\mu p^+ + \frac{\epsilon}{2}\int_{K-X[p^+]}^{K-Y^+}(p[K-Z]-p^+)\frac{\lambda e^{-\lambda Z}}{1-e^{-\lambda Y}}\,dZ = 0,$$

and

$$(35) \quad \mu p^+ + \frac{\epsilon}{2}\left(\int_0^{K-X[p^+]} p[K-Z]\frac{\lambda e^{-\lambda Z}}{1-e^{-\lambda Y}}\,dZ + \int_{K-X[p^+]}^{Y} p^+\frac{\lambda e^{-\lambda Z}}{1-e^{-\lambda Y}}\,dZ\right)$$

$$-p[Y]\left(1-\frac{\epsilon}{2}\right) - \frac{\epsilon}{2}\int_0^Y\left(1-\frac{Z}{Y}\right)p[K-Z]\frac{\lambda e^{-\lambda Z}}{1-e^{-\lambda Y}}\,dZ = 0.$$

[12] This is the practice in the model of Just and Weber (2007) and the model of Creti and Fabra (2002).

In light of the identity $Y = K - Y^+$, the rational-expectations equilibrium is now defined by the two equations (34) and (35), which involves only the two unknowns p^+ and Y^+.

6. THE RATIONAL EXPECTATIONS EQUILIBRIUM: EXISTENCE AND UNIQUENESS

In this section we study the existence and uniqueness of a rational expectations equilibrium. The analysis involves the two conditions (34) and (35), which define the equilibrium.

6.1. *The Condition of Rational Expectations*

Now for any $Y^+, 0 < Y^+ < K$, and any $p^+ > 0$, let

(36) $$I[Y^+, p^+] = -\mu p^+ + \frac{\epsilon}{2}\int_{K-X[p^+]}^{K-Y^+}(p[K-Z] - p^+)\frac{\lambda e^{-\lambda Z}}{1-e^{-\lambda Y}}dZ.$$

The condition that expectations concerning the market supply of operating reserves are correct is equivalent to the condition that $I[Y^+, p^+] = 0$.

Differentiating (36) with respect to p^+, we obtain

(37) $$D_2 I[Y^+, p^+] = -\mu - \frac{\epsilon}{2}\int_{K-X[p^+]}^{K-Y^+}\frac{\lambda e^{-\lambda Z}}{1-e^{-\lambda Y}}dZ$$

$$+\frac{\epsilon}{2}X'[p^+](p[K-(K-X[p^+])] - p^+)\frac{\lambda e^{-\lambda(K-X[p^+])}}{1-e^{-\lambda Y}}$$

$$= -\mu - \frac{\epsilon}{2}\int_{K-X[p^+]}^{K-Y^+}\frac{\lambda e^{-\lambda Z}}{1-e^{-\lambda Y}}dZ.$$

Here D_i represents the differential operator with respect to the ith variable.

It follows from (37) that for any given $0 < Y^+ < K$, the map $I[Y^+,.]: p^+ \to I[Y^+, p^+], p^+ > 0$, is strictly decreasing with a slope strictly less than $-\mu < 0$, and this means $I[Y^+, p^+] \to -\infty$ when $p^+ \to \infty$. On the other hand, when $p^+ \to 0$, we have

(38) $I[Y^+, p^+] \to \frac{\epsilon}{2}\int_{-\infty}^{K-Y^+} p[K-Z]\,\frac{\lambda e^{-\lambda Z}}{1-e^{-\lambda Y}}\,dZ > 0$.

Therefore, for any $0 < Y^+ < K$, there exists a unique value of $p^+ > 0$, say $p^+ = \xi[Y^+]$, such that $I\big[Y^+, \xi[Y^+]\big] = 0$. The curve $\xi: Y^+ \to \xi[Y^+], 0 < Y^+ < K$, represents the strike price of an option so that when the consumers anticipate Y^+ as the market supply of operating reserves, the expectations are correct.

Applying the implicit function theorem to $I[Y^+, p^+]$, we obtain

(39) $$\xi'[Y^+] = -\frac{D_1 I\big[Y^+,\xi[Y^+]\big]}{D_2 I\big[Y^+,\xi[Y^+]\big]} = -\frac{-\frac{\epsilon}{2}\left(p[Y^+]-\xi[Y^+]\right)\frac{\lambda e^{-\lambda(K-Y^+)}}{1-e^{-\lambda Y}}}{-\mu-\frac{\epsilon}{2}\int_{K-X\left[\xi[Y^+]\right]}^{K-Y^+}\frac{\lambda e^{-\lambda Z}}{1-e^{-\lambda Y}}dZ}$$

$$= -\frac{\frac{\epsilon}{2}\left(p[Y^+]-\xi[Y^+]\right)\frac{\lambda e^{-\lambda(K-Y^+)}}{1-e^{-\lambda Y}}}{\mu+\frac{\epsilon}{2}\int_{K-X\left[\xi[Y^+]\right]}^{K-Y^+}\frac{\lambda e^{-\lambda Z}}{1-e^{-\lambda Y}}dZ} < 0.$$

We claim that $\ell im_{Y^+\to 0}\xi[Y^+] = \infty$. We prove the claim by reductio ad absurdum. Suppose then that $\ell im_{Y^+\to 0}\xi[Y^+] = m < \infty$. Letting $\zeta = K - Z$, and then using this change of variable to evaluate (36) at $(Y^+, \xi[Y^+])$, we obtain

(40) $$0 = I\big[Y^+,\xi[Y^+]\big] = -\mu\xi[Y^+] - \frac{\epsilon}{2}\int_{K-X\left[\xi[Y^+]\right]}^{Y^+}\left(p[\zeta]-\xi[Y^+]\right)\frac{\lambda e^{-\lambda(K-\zeta)}}{1-e^{-\lambda Y}}\,d\zeta$$

$$= -\mu\xi[Y^+] - \frac{\epsilon}{2}\xi[Y^+]\frac{\lambda e^{-\lambda K}}{1-e^{-\lambda Y}}\int_{Y^+}^{K-X\left[\xi[Y^+]\right]}\lambda e^{\lambda\zeta}d\zeta + \frac{\epsilon}{2}\frac{\lambda e^{-\lambda K}}{1-e^{-\lambda Y}}\int_{Y^+}^{K-X\left[\xi[Y^+]\right]}p[\zeta]\lambda e^{\lambda\zeta}d\zeta$$

Now let $Y^+ \to 0$. Using the premise of the reductio ad absurdum argument, we obtain

(41) $0 = \ell im_{Y^+\to 0} I\big[Y^+,\xi[Y^+]\big]$

$$= -\mu m - \frac{\epsilon}{2} m \frac{\lambda e^{-\lambda K}}{1-e^{-\lambda K}} \int_0^{K-X[m]} \lambda e^{\lambda\zeta} d\zeta + \frac{\epsilon}{2} \frac{\lambda e^{-\lambda K}}{1-e^{-\lambda K}} \int_0^{K-X[m]} p[\zeta] \lambda e^{\lambda\zeta} d\zeta$$

$$= -\mu m - \frac{\epsilon}{2} m \frac{\lambda e^{-\lambda K}}{1-e^{-\lambda K}} \lambda\left(e^{\lambda(K-X[m])} - 1\right) + \frac{\epsilon}{2} \frac{\lambda e^{-\lambda K}}{1-e^{-\lambda K}} \int_0^{K-X[m]} p[\zeta] \lambda e^{\lambda\zeta} d\zeta.$$

However, because $0 < \gamma < 1$, the integral on the last line of (41) diverges, i.e.,

$$\int_0^{K-X[m]} p[\zeta] \lambda e^{\lambda\zeta} d\zeta = \int_0^{K-X[m]} \Theta^{\frac{1}{\gamma}} \zeta^{-\frac{1}{\gamma}} \frac{\lambda e^{\lambda\zeta}}{1-e^{-\lambda Y}} d\zeta = \infty.$$

The premise of the reductio ad absurdum argument leads to a contradiction, and the claim is proved.

The curve $\xi: Y^+ \to \xi[Y^+], 0 < Y^+ < K,$ is depicted in Figure 5, which dictates the level of the strike price as a function of the anticipated market supply of operating reserves. Also, depicted in Figure 5 is the inverse market demand curve for electricity $p[X]$, drawn with O as the origin. Observe that in Figure 5, the distance between a point on the curve $\xi: Y^+ \to \xi[Y^+], 0 < Y^+ < K,$ and the vertical line located at a distance equal to K (the industry's generating capacity) represents the anticipated market supply of operating reserves Y^+.

Because $\xi[Y^+] \to \infty$ when $Y^+ \to 0$, the curve $\xi: Y^+ \to \xi[Y^+], 0 < Y^+ < K,$ is above the inverse market demand curve when Y^+ is small. Also, when $Y^+ \to K$, the curve $\xi: Y^+ \to \xi[Y^+], 0 < Y^+ < K,$ is below the inverse market demand curve. Hence these two curves cross each other exactly once, and this crossing is noted as A in Figure 5. The horizontal distance between point A and the vertical line located at K on the horizontal axis, which is noted as Y^+_{max}, represents the range in which the market supply of operating reserves is allowed to vary so that the price of electricity on the real-time market is below the strike price before any breakdown occurs.

6.2. *The Condition that the Revenues Earned by a Generating Plant is the Same in the Market for Operating Reserves and on the Real-Time Market for Energy*

Now for any $Y^+, 0 < Y^+ < K$, and any $p^+ > 0$, let

(42) $$J[Y^+, p^+] = \mu p^+ + \frac{\epsilon}{2}\left(\int_0^{K-X[p^+]} p[K-Z]\frac{\lambda e^{-\lambda Z}}{1-e^{-\lambda Y}}dZ + \int_{K-X[p^+]}^{K-Y^+} p^+\frac{\lambda e^{-\lambda Z}}{1-e^{-\lambda Y}}dZ\right)$$

$$-p[K-Y^+]\left(1-\frac{\epsilon}{2}\right) - \frac{\epsilon}{2}\int_0^{K-Y^+}\left(1-\frac{Z}{K-Y^+}\right)p[K-Z]\frac{\lambda e^{-\lambda Z}}{1-e^{-\lambda Y}}dZ.$$

The condition that producers are indifferent between offering their generating assets on the market for operating reserves and offering their generating assets on the real-time energy market is equivalent to the condition that $J[Y^+, p^+] = 0$.

Differentiating (42) with respect to p^+, we obtain

(43) $$D_2 J[Y^+, p^+] = \mu + \frac{\epsilon}{2}\begin{pmatrix} -X'[p^+]p[K-(K-X[p^+])]\frac{\lambda e^{-\lambda(K-X[p^+])}}{1-e^{-\lambda Y}} \\ +\int_{K-X[p^+]}^{Y}\frac{\lambda e^{-\lambda Z}}{1-e^{-\lambda Y}}dZ + X'[p^+]p^+\frac{\lambda e^{-\lambda(K-X[p^+])}}{1-e^{-\lambda Y}}\end{pmatrix}$$

$$= \mu + \frac{\epsilon}{2}\left(-X'[p^+]p^+\frac{\lambda e^{-\lambda(K-X[p^+])}}{1-e^{-\lambda Y}} + \int_{K-X[p^+]}^{Y}\frac{\lambda e^{-\lambda Z}}{1-e^{-\lambda Y}}dZ + X'[p^+]p^+\frac{\lambda e^{-\lambda(K-X[p^+])}}{1-e^{-\lambda Y}}\right)$$

$$= \mu + \frac{\epsilon}{2}\int_{K-X[p^+]}^{Y}\frac{\lambda e^{-\lambda Z}}{1-e^{-\lambda Y}}dZ.$$

It follows from (43) that for any given $Y^+, 0 < Y^+ < K$, as p^+ rises, $J[Y^+, p^+]$ also rises with p^+ at a slope exceeding μ, and this means $J[Y^+, p^+] \to \infty$ when $p^+ \to \infty$.

Next, note that for $p^+ = p[K]$, we have

(44) $$J[Y^+, p[K]] = \mu p[K] + \frac{\epsilon}{2}\left(\int_0^{K-X[p[K]]} p[K-Z]\frac{\lambda e^{-\lambda Z}}{1-e^{-\lambda Y}}dZ + \int_{K-X[p[K]]}^{K-Y^+} p^+\frac{\lambda e^{-\lambda Z}}{1-e^{-\lambda Y}}dZ\right)$$

$$-p[K-Y^+]\left(1-\frac{\epsilon}{2}\right) - \frac{\epsilon}{2}\int_0^{K-Y^+}\left(1-\frac{Z}{K-Y^+}\right)p[K-Z]\frac{\lambda e^{-\lambda Z}}{1-e^{-\lambda Y}}dZ$$

$$= \mu p[K] + \frac{\epsilon}{2} p[K] - p[K - Y^+]\left(1 - \frac{\epsilon}{2}\right) - \frac{\epsilon}{2}\int_0^{K-Y^+}\left(1 - \frac{Z}{K-Y^+}\right) p[K - Z]\frac{\lambda e^{-\lambda Z}}{1-e^{-\lambda Y}}\, dZ.$$

If $\mu \leq 1 - \epsilon$, and this we assume to hold throughout the paper, then the last line in (44) is negative. Therefore, for any $Y^+, 0 < Y^+ < K$, there exists a unique value of p^+, say $p^+ = \eta[Y^+]$, with $\eta[Y^+] > p[K]$, such that $J[Y^+, \eta[Y^+]] = 0$. Note that the curve $Y^+ \to \eta[Y^+], 0 < Y^+ < K$, which is depicted in Figure 5, is above the inverse market demand curve $p[X]$.

Differentiating (42) with respect to Y^+, we obtain

(45) $$D_1 J[Y^+, p^+] = -\frac{\epsilon}{2} p^+ \frac{\lambda e^{-\lambda(K-Y^+)}}{1-e^{-\lambda Y}} + p'[K - Y^+]\left(1 - \frac{\epsilon}{2}\right)$$

$$+\frac{\epsilon}{2}\left(1 - \frac{K-Y^+}{K-Y^+}\right) p[Y^+]\frac{\lambda e^{-\lambda(K-Y^+)}}{1-e^{-\lambda Y}} - \frac{\epsilon}{2}\int_0^{K-Y^+}\frac{Z}{(K-Y^+)^2} p[K - Z]\frac{\lambda e^{-\lambda Z}}{1-e^{-\lambda Y}}\, dZ$$

$$= -\frac{\epsilon}{2} p^+ \frac{\lambda e^{-\lambda(K-Y^+)}}{1-e^{-\lambda Y}} + p'[K - Y^+]\left(1 - \frac{\epsilon}{2}\right) - \frac{\epsilon}{2}\int_0^{K-Y^+}\frac{Z}{(K-Y^+)^2} p[K - Z]\frac{\lambda e^{-\lambda Z}}{1-e^{-\lambda Y}}\, dZ$$

$$< 0.$$

Applying the implicit function theorem to $J[Y^+, p^+]$, we obtain

(46) $$\eta'[Y^+] = -\frac{D_1 J[Y^+, \eta[Y^+]]}{D_2 J[Y^+, \eta[Y^+]]}$$

$$= -\frac{-\frac{\epsilon}{2} p^+ \frac{\lambda e^{-\lambda(K-Y^+)}}{1-e^{-\lambda Y}} + p\prime[K-Y^+]\left(1-\frac{\epsilon}{2}\right) - \frac{\epsilon}{2}\int_0^{K-Y^+}\frac{Z}{(K-Y^+)^2} p[K-Z]\frac{\lambda e^{-\lambda Z}}{1-e^{-\lambda Y}} dZ}{\mu + \frac{\epsilon}{2}\int_{K-X[\eta[Y^+]]}^{Y}\frac{\lambda e^{-\lambda Z}}{1-e^{-\lambda Y}} dZ} > 0.$$

The curve $Y^+ \to \eta[Y^+], 0 < Y^+ < K$, is thus upward-sloping, and a fortiori $\ell im_{Y^+ \to 0}\eta[Y^+] < \infty$.

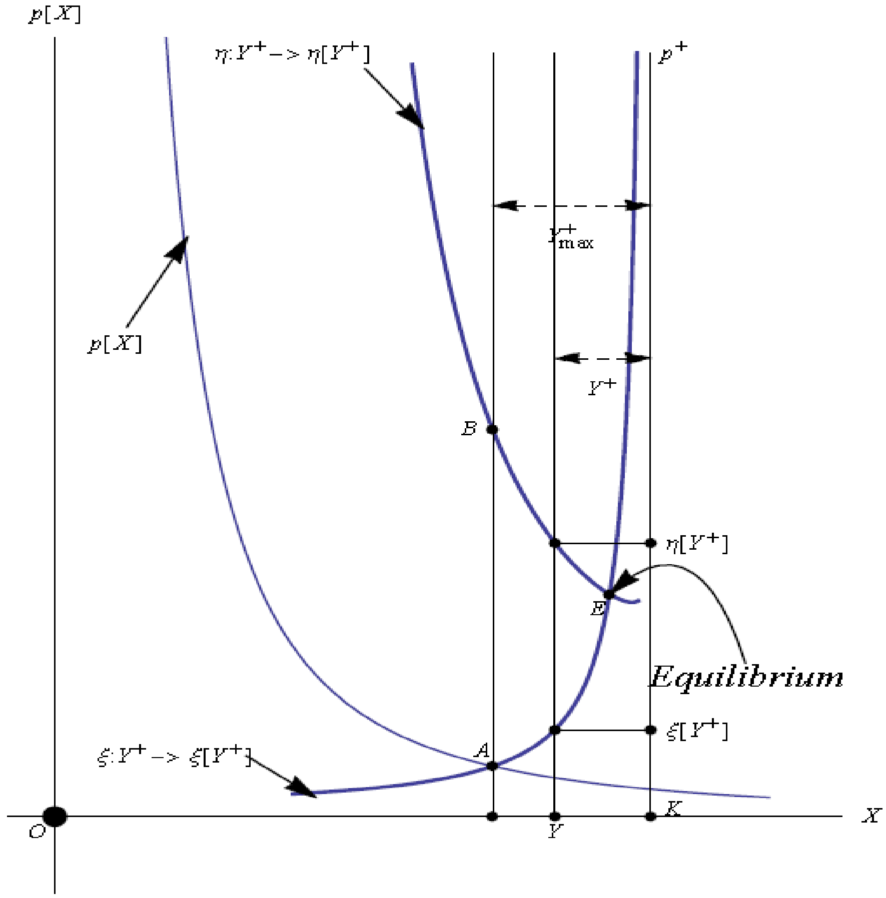

Figure 5 The rational expectations equilibrium

6.3. *Existence and Uniqueness of the Rational Expectations Equilibrium*

In Figure 5, the point $(Y^{+}_{max}, \eta[Y^{+}_{max}])$ is depicted as point B, which is above point A; that is the curve $Y^{+} \to \eta[Y^{+}], 0 < Y^{+} < K$, is above the curve $Y^{+} \to \xi[Y^{+}], 0 < Y^{+} < K$, when $Y^{+} = Y^{+}_{max}$. Also, when $Y^{+} \to 0$, the former curve is below the latter curve. Hence these two curves must cross each other at least once at some $Y^{+} \in (0, Y^{+}_{max})$. Furthermore, because the former curve is upward-sloping while the latter curve is downward-sloping, such a crossing is unique. We have just proved the following proposition:

PROPOSITION 3: *For each given value of $\mu, 0 < \mu < 1$, there exists a unique rational expectations equilibrium if $\mu \leq 1 - \epsilon$.*

7. PROPERTIES OF THE EQUILIBRIUM

According to Proposition 3, a rational expectations equilibrium exists, and is unique for each value of μ that satisfies the condition $\mu \leq 1 - \epsilon$. By varying the value of μ while maintaining this condition, we obtain a continuum of rational expectations equilibrium. However, all these rational expectations equilibria involve the same level of operating reserves, as asserted by the following proposition:

PROPOSITION 4: *The equilibrium strike price is a decreasing function of* μ. *However, the equilibrium level of operating reserves is independent of the ratio* μ *of the option price over the strike price. More specifically, the equilibrium level of operating reserves is the unique value of* Y^+ *that satisfies the following condition:*

(48) $$-p[K - Y^+]\left(1 - \frac{\epsilon}{2}\right) + \frac{\epsilon}{2}\int_0^{K-Y^+} \frac{Z}{K-Y^+} p[K - Z]\frac{\lambda e^{-\lambda Z}}{1-e^{-\lambda(K-Y^+)}}dZ = 0.$$

PROOF*:* Adding (34) and (35), we obtain

$$\frac{\epsilon}{2}\int_{K-X[p^+]}^{K-Y^+}(p[K-Z] - p^+)\frac{\lambda e^{-\lambda Z}}{1-e^{-\lambda Y}}dZ$$

$$+\frac{\epsilon}{2}\left(\int_0^{K-X[p^+]} p[K-Z]\frac{\lambda e^{-\lambda Z}}{1-e^{-\lambda Y}}dZ + \int_{K-X[p^+]}^{K-Y^+} p^+\frac{\lambda e^{-\lambda Z}}{1-e^{-\lambda Y}}dZ\right)$$

$$-p[K - Y^+]\left(1 - \frac{\epsilon}{2}\right) - \frac{\epsilon}{2}\int_0^{K-Y^+}\left(1 - \frac{Z}{K-Y^+}\right)p[K-Z]\frac{\lambda e^{-\lambda Z}}{1-e^{-\lambda Y}}dZ = 0,$$

which, after some simplifications, is reduced to

(48) $$-p[K - Y^+]\left(1 - \frac{\epsilon}{2}\right) + \frac{\epsilon}{2}\int_0^{K-Y^+} \frac{Z}{K-Y^+} p[K - Z]\frac{\lambda e^{-\lambda Z}}{1-e^{-\lambda(K-Y^+)}}dZ = 0.$$

Note that (48), which defines an equation in Y^+, is the condition stated in Proposition 4. It contains neither p^+ nor μ.

Now choose an arbitrary value for $\mu, 0 < \mu \leq 1 - \frac{\epsilon}{2}$, and then invoke Proposition 3 to obtain the existence of a unique rational expectations equilibrium associated with μ. The level of operating reserves associated with this equilibrium solves (48), i.e., (48) has at least one solution. We claim that such a solution is unique. To prove the claim, we use a reductio ad absurdum argument. Suppose then that (48) has more than one solution, say Y_1^+ and Y_2^+, with $Y_1^+ \neq Y_2^+$, and that (Y_1^+, p_1^+) is the unique rational expectations equilibrium associated with μ. According to (34), we have

(49) $$-\mu p_1^+ + \frac{\epsilon}{2}\int_{K-X[p_1^+]}^{K-Y_1^+}(p[K-Z]-p_1^+)\frac{\lambda e^{-\lambda Z}}{1-e^{-\lambda(K-Y_1^+)}}dZ = 0.$$

Now if we replace Y_1^+ by Y_2^+ in the expression on the left-hand side of (49), then we obtain

(50) $$-\mu p_1^+ + \frac{\epsilon}{2}\int_{K-X[p_1^+]}^{K-Y_2^+}(p[K-Z]-p_1^+)\frac{\lambda e^{-\lambda Z}}{1-e^{-\lambda(K-Y_2^+)}}dZ.$$

If (50) is equal to 0, then (Y_2^+, p_1^+) is also a rational expectations equilibrium associated with μ, which is not possible due to Proposition 3. On the other hand, if (50) is not equal to 0, then we can then choose a strike price, say $p_2^+, p_2^+ \neq p_1^+$, in (50) such that (50) becomes an equality; that is

(51) $$-\mu p_2^+ + \frac{\epsilon}{2}\int_{K-X[p^+]}^{K-Y_2^+}(p[K-Z]-p_2^+)\frac{\lambda e^{-\lambda Z}}{1-e^{-\lambda(K-Y_2^+)}}dZ = 0.$$

It follows from (51) that (Y_2^+, p_2^+) is also a rational expectations equilibrium associated with μ, which is not possible due to Proposition 3.

Having proved that the equilibrium level of operating reserves does not depend on μ, we now show that the strike price varies inversely with μ. To this end, pick some value for μ, and let (Y^+, p^+) be the unique rational expectations equilibrium associated with μ. According to (34), we have

(52) $$-\mu p^+ + \frac{\epsilon}{2}\int_{K-X[p^+]}^{K-Y^+}(p[K-Z]-p^+)\frac{\lambda e^{-\lambda Z}}{1-e^{-\lambda(K-Y^+)}}dZ = 0,$$

where Y^+ is the unique solution of (48).

Now if μ rises, then the left-hand side of (52) will become negative. To restore the equality, we must then reduce p^+. Hence p^+ falls when μ rises. ■

Proposition 4 makes it possible to compute the rational expectations equilibrium for each given μ in a recursive manner. First, solve (48) for Y^+. Once Y^+ has been found, use it in the condition (34) to solve for p^+.

According to Proposition 4, the equilibrium level of operating reserves depends only on preferences and technology, not on μ, the ratio option price/strike price. Furthermore, $x[p[Y^+],\theta]$, the demand for call options by a consumer of type θ, is also completely determined once Y^+ is given. On the other hand, the chosen value of μ serves only as a means to structure the rewards for operating reserves by varying the certain component – the option price – and the random component – the strike price and the probability of when an option is exercised. Thus, from the perspective of real resource allocation, the rational expectations equilibrium is unique.

Intuitively, we expect that the equilibrium level of operating reserves rises when the event of generating plant failure is more likely. Proposition 5 confirms this intuition.

PROPOSITION 5: *We have* $\frac{\partial Y^+}{\partial \epsilon} > 0$. *That is, a rise in the probability that a generating plant failure occurs induces a rise in the equilibrium level of operating reserves.*

PROOF*:* If we write $Y^+[\epsilon]$ to express the dependence of the equilibrium level of operating reserves on ϵ, then (48) assumes the following form:

$$(53) \quad -p[K - Y^+[\epsilon]]\left(1 - \frac{\epsilon}{2}\right) + \frac{\epsilon}{2}\int_0^{K-Y^+[\epsilon]} \frac{Z}{K-Y^+[\epsilon]}\, p[K - Z]\, \frac{\lambda e^{-\lambda Z}}{1-e^{-\lambda(K-Y^+[\epsilon])}}\, dZ = 0.$$

Differentiating (53) with respect to ϵ, and then solving for $\frac{\partial Y^+}{\partial \epsilon}$, we obtain

$$(54) \quad \frac{\partial Y^+}{\partial \epsilon} = -\frac{\left(\begin{array}{c} p'[K-Y^+]\left(1-\frac{\epsilon}{2}\right) - \frac{\epsilon}{2}p[Y^+]\frac{\lambda e^{-\lambda(K-Y^+)}}{1-e^{-\lambda Y}} - \frac{\epsilon}{2}\int_0^{K-Y^+}\frac{Z}{(K-Y^+)^2}p[K-Z]\frac{\lambda e^{-\lambda Z}}{1-e^{-\lambda(K-Y^+)}}dZ \\ -\frac{\epsilon}{2}\int_0^{K-Y^+}\frac{Z}{K-Y^+}p[K-Z]\frac{\lambda e^{-\lambda Z}}{1-e^{-\lambda(K-Y^+)}}\frac{\lambda e^{-\lambda(K-Y^+)}}{1-e^{-\lambda(K-Y^+)}}dZ \end{array}\right)}{\frac{1}{2}p[K-Y^+[\epsilon]]+\frac{1}{2}\int_0^{K-Y^+[\epsilon]}\frac{Z}{K-Y^+[\epsilon]}p[K-Z]\frac{\lambda e^{-\lambda Z}}{1-e^{-\lambda(K-Y^+[\epsilon])}}dZ} > 0.$$

■

One of the major questions concerning the market for operating reserves is whether it is efficient. This involves comparing the level of operating reserves under the competitive equilibrium with the socially optimal level.

To find the social optimum, we need to solve the central planner's problem. To this end, let Y^+ be the number of generating plants chosen by the central planner to serve as operating reserves. The number of generating plants that are run to produce electricity on the real-time energy market is then given by $Y = K - Y^+$. The market price that clears the volume of electricity generated by these generating plants is $p[Y] = \Theta^{\frac{1}{\gamma}} Y^{-\frac{1}{\gamma}}$, and the demand for electricity by a consumer of type θ who faces this price is $x[p[Y], \theta] = \left(\frac{\Theta^{\frac{1}{\gamma}} Y^{-\frac{1}{\gamma}}}{\theta}\right)^{-\gamma}$, which yields the following level of gross utility to the consumer in question:

$$(55) \quad \theta\phi\big[x[p[Y],\theta]\big] = \theta \frac{1}{1-\frac{1}{\gamma}} \left(\left(\frac{\Theta^{\frac{1}{\gamma}} Y^{-\frac{1}{\gamma}}}{\theta}\right)^{-\gamma}\right)^{1-\frac{1}{\gamma}} = \theta^{\gamma} \frac{1}{1-\frac{1}{\gamma}} \Theta^{\frac{1-\gamma}{\gamma}} Y^{1-\frac{1}{\gamma}}.$$

The social welfare obtained under the market allocation is

$$(56) \quad \Phi[Y] = \int_{\underline{\theta}}^{\overline{\theta}} \theta^{\gamma} \frac{1}{1-\frac{1}{\gamma}} \Theta^{\frac{1-\gamma}{\gamma}} Y^{1-\frac{1}{\gamma}} h[\theta] d\theta = \frac{1}{1-\frac{1}{\gamma}} \Theta^{\frac{1}{\gamma}} Y^{1-\frac{1}{\gamma}}.$$

Note that $\Phi'[Y] = p[Y]$. Because under the market allocation marginal utilities of electricity consumption are equalized across consumers, $\Phi[Y]$ also represents the social welfare obtained under the social optimum. Thus, if Y is the number of generating plants chosen by the central planner to run on the real-time energy market, then the expected social welfare obtained is

$$(57) \quad W[Y] = (1-\epsilon)\Phi[Y] + \epsilon \int_0^1 \left(\tau\Phi[K - Y^+] + (1-\tau)\int_0^Y \Phi[K - Z] \frac{\lambda e^{-\lambda Z}}{1-e^{-\lambda Y}} dZ\right) d\tau$$

$$= \left(1 - \frac{\epsilon}{2}\right)\Phi[Y] + \frac{\epsilon}{2}\int_0^Y \Phi[K - Z] \frac{\lambda e^{-\lambda Z}}{1-e^{-\lambda Y}} dZ.$$

The following first-order condition characterizes the socially optimal number of generating plants that solves the central planner's problem:

(58) $$W'[Y] = \left(1-\frac{\epsilon}{2}\right)p[Y] + \frac{\epsilon}{2}\left(\Phi[K-Y]\frac{\lambda e^{-\lambda Y}}{1-e^{-\lambda Y}} - \int_0^Y \Phi[K-Z]\frac{\lambda e^{-\lambda Z}}{1-e^{-\lambda Y}}\frac{\lambda e^{-\lambda Y}}{1-e^{-\lambda Y}}dZ\right)$$

$$= \left(1-\frac{\epsilon}{2}\right)p[Y] - \frac{\epsilon}{2}\left(\frac{\lambda e^{-\lambda Y}}{1-e^{-\lambda Y}}\int_0^Y (\Phi[K-Z]-\Phi[K-Y])\frac{\lambda e^{-\lambda Z}}{1-e^{-\lambda Y}}dZ\right) = 0,$$

which can be rearranged as

(59) $$\left(1-\frac{\epsilon}{2}\right)p[Y] = \frac{\epsilon}{2}\left(\frac{\lambda e^{-\lambda Y}}{1-e^{-\lambda Y}}\int_0^Y (\Phi[K-Z]-\Phi[K-Y])\frac{\lambda e^{-\lambda Z}}{1-e^{-\lambda Y}}dZ\right).$$

PROPOSITION 6: *If λ, the parameter of the exponential distribution that characterizes the number of failing generating plants conditional on the event that a breakdown occurs, is high enough, then the level of operating reserves is higher under the competitive equilibrium than under the social optimum. That is, the market over-provides operating reserves.*

PROOF: The proof is a little technical, and is relegated to Appendix B.

Proposition 6 implies that a social planner will allocate more generating capacity towards the production of the electricity than the competitive market would. When consumers are making decision on how much to consume, they are too small to affect the amount of operating reserve options that will be supplied in the market because of the competitive nature of the market. This implies that consumers are not able to optimally decide the optimal probability of generator failure. Using all the information available, the best consumers can do is to make a rational expectation on the amount of operating reserve capacity that would be supplied and make optimal consumption decisions based on such a guess. This inability to affect the supply of operating reserve options results in a too much capacity being committed to electricity production. The social planner is however able to maximise social welfare by choose the welfare maximizing probability of failure.

Proposition 6 can also be thought of from a public goods perspective.When a consumer is choosing how much operating reserve to consume, they will make rational expectations about the market supply of operating reserves. When expectations are correct the consumers will determine the amount of electricity that will be produced in real time. Hence the probability distribution of generator failure is determined by the rational expectation of operating reserves.

In a rational expectations model, consumers can't affect the amount of electricity that is produced in the real time market however collective action by all consumers will affect the distribution of generator failure. In this way the distribution of generator failure is a public good. From economics is clear that the market will under provide a public good. Hence the under provision of electricity will simply imply an over provision of operating reserves.[13]

8. CONCLUDING REMARKS

In this paper we have analysed a market for operating reserves where operating reserves are used to moderate price spikes that occur when generators fail. Our analysis uses the concept of rational expectations equilibria where each consumer demand for options was based on her expectations about the number of options that will be bought and sold in the market.

In our analysis, we show that a rational expectations equilibrium exists and it is unique under certain conditions. We also show that the optimal amount of operating reserves is unique and not dependent on μ. However, the strike price was a decreasing function of μ. Finally we show that under the rational expectations

[13] For more exposition on the public goods view refer to Stiglitz and Newberry (1982)

equilibrium, the competitive market will over-supply operating reserves relative to the socially optimal level of operating reserves.

Price spikes due to generator failure will cause individuals to lower their consumption of electricity. However in most electricity markets, price caps are used to moderate price spikes in the electricity markets. These price caps result in the need for involuntary blackouts to make sure that there is a balance between demand for and supply of electricity. Independent system operators have to choose which section of the power grid to blackout based on the preferences for electricity of consumers within that zone (i.e. how much the consumers are willing to pay to prevent a blackout). However the preferences for electricity are private information. In chapter 3 we design an optimal blackout strategy that will help determine when a zone will be blackout and the optimal level of reliability to offer.

APPENDIX A

Derivation of the First-Order Condition (22)

The problem faced by a consumer of type θ in the day-ahead market for operating reserves is

(A.1) $max_{x^+\geq 0}\omega[x^+,q,p^+,Y^+,\theta] =$

$$max_{x^+\geq 0}\left(\begin{array}{c} -qx^+ + \left(1-\frac{\epsilon}{2}\right)v[p[Y],\theta] + \frac{\epsilon}{2}\int_0^{K-X[p^+]} v[p[K-Z],\theta]\frac{\lambda e^{-\lambda Z}}{1-e^{-\lambda Y}}dZ \\ +\frac{\epsilon}{2}\left(\begin{array}{c}\int_{K-X[p^+]}^{K-X[\theta\phi'[x^+]]}\left(v[p[K-Z],\theta]+x^+(p[K-Z]-p^+)\right)\frac{\lambda e^{-\lambda Z}}{1-e^{-\lambda Y}}dZ \\ +\int_{K-X[\theta\phi'[x^+]]}^{Y}(\theta\phi[x^+]-p^+x^++y[\theta])\frac{\lambda e^{-\lambda Z}}{1-e^{-\lambda Y}}dZ\end{array}\right)\end{array}\right).$$

The following first-order condition characterizes the solution of (A.1):

(A.2) $\frac{\partial}{\partial x^+}\omega[x^+,q,p^+,Y^+,\theta]$

$$= -q + \frac{\epsilon}{2}\frac{\partial}{\partial x^+}\left(\begin{array}{c}\int_{K-X[p^+]}^{K-X[\theta\phi'[x^+]]}\left(v[p[K-Z],\theta]+x^+(p[K-Z]-p^+)\right)\frac{\lambda e^{-\lambda Z}}{1-e^{-\lambda Y}}dZ \\ +\int_{K-X[\theta\phi'[x^+]]}^{Y}(\theta\phi[x^+]-p^+x^++y[\theta])\frac{\lambda e^{-\lambda Z}}{1-e^{-\lambda Y}}dZ\end{array}\right)$$

$$= -q + \frac{\epsilon}{2}\left(\begin{array}{c}\int_{K-X[p^+]}^{K-X[\theta\phi'[x^+]]}(p[K-Z]-p^+)\frac{\lambda e^{-\lambda Z}}{1-e^{-\lambda Y}}dZ \\ -X'[\theta\phi'[x^+]]\theta\phi''[x^+]\left(\begin{array}{c}v[p[K-(K-X[\theta\phi'[x^+]])],\theta] \\ +x^+(p[K-(K-X[\theta\phi'[x^+]])]-p^+)\end{array}\right)\frac{\lambda e^{-\lambda(K-X[\theta\phi'[x^+]])}}{1-e^{-\lambda Y}} \\ +\int_{K-X[\theta\phi'[x^+]]}^{Y}(\theta\phi'[x^+]-p^+)\frac{\lambda e^{-\lambda Z}}{1-e^{-\lambda Y}}dZ \\ +X'[\theta\phi'[x^+]]\theta\phi''[x^+](\theta\phi[x^+]-p^+x^++y[\theta])\frac{\lambda e^{-\lambda(K-X[\theta\phi'[x^+]])}}{1-e^{-\lambda Y}}\end{array}\right)$$

$$=$$

$$-q+\frac{\epsilon}{2}\begin{pmatrix} \int_{K-X[p^+]}^{K-X\left[\theta\phi'[x^+]\right]}(p[K-Z]-p^+)\,\frac{\lambda e^{-\lambda Z}}{1-e^{-\lambda Y}}\,dZ \\ -X'\left[\theta\phi'[x^+]\right]\theta\phi''[x^+]\begin{pmatrix} v\left[p\left[X\left[\theta\phi'[x^+]\right]\right],\theta\right] \\ +x^+\left(p\left[X\left[\theta\phi'[x^+]\right]\right]-p^+\right)\end{pmatrix}\frac{\lambda e^{-\lambda\left(K-X\left[\theta\phi'[x^+]\right]\right)}}{1-e^{-\lambda Y}} \\ +\int_{K-X\left[\theta\phi'[x^+]\right]}^{Y}(\theta\phi'[x^+]-p^+)\,\frac{\lambda e^{-\lambda Z}}{1-e^{-\lambda Y}}\,dZ \\ +X'\left[\theta\phi'[x^+]\right]\theta\phi''[x^+](\theta\phi[x^+]-p^+x^++y[\theta])\,\frac{\lambda e^{-\lambda\left(K-X\left[\theta\phi'[x^+]\right]\right)}}{1-e^{-\lambda Y}}\end{pmatrix}$$

$$=-q+\frac{\epsilon}{2}\begin{pmatrix} \int_{K-X[p^+]}^{K-X\left[\theta\phi'[x^+]\right]}(p[K-Z]-p^+)\,\frac{\lambda e^{-\lambda Z}}{1-e^{-\lambda Y}}\,dZ \\ -X'\left[\theta\phi'[x^+]\right]\theta\phi''[x^+]\begin{pmatrix} v[\theta\phi'[x^+],\theta] \\ +x^+(\theta\phi'[x^+]-p^+)\end{pmatrix}\frac{\lambda e^{-\lambda\left(K-X\left[\theta\phi'[x^+]\right]\right)}}{1-e^{-\lambda Y}} \\ +\int_{K-X\left[\theta\phi'[x^+]\right]}^{Y}(\theta\phi'[x^+]-p^+)\,\frac{\lambda e^{-\lambda Z}}{1-e^{-\lambda Y}}\,dZ \\ +X'\left[\theta\phi'[x^+]\right]\theta\phi''[x^+](\theta\phi[x^+]-p^+x^++y[\theta])\,\frac{\lambda e^{-\lambda\left(K-X\left[\theta\phi'[x^+]\right]\right)}}{1-e^{-\lambda Y}}\end{pmatrix}$$

$$=$$

$$-q+\frac{\epsilon}{2}\begin{pmatrix} \int_{K-X[p^+]}^{K-X\left[\theta\phi'[x^+]\right]}(p[K-Z]-p^+)\,\frac{\lambda e^{-\lambda Z}}{1-e^{-\lambda Y}}\,dZ \\ -X'\left[\theta\phi'[x^+]\right]\theta\phi''[x^+]\begin{pmatrix} \theta\phi[x^+]-\theta\phi'[x^+]x^++y[\theta] \\ +x^+(\theta\phi'[x^+]-p^+)\end{pmatrix}\frac{\lambda e^{-\lambda\left(K-X\left[\theta\phi'[x^+]\right]\right)}}{1-e^{-\lambda Y}} \\ +\int_{K-X\left[\theta\phi'[x^+]\right]}^{Y}(\theta\phi'[x^+]-p^+)\,\frac{\lambda e^{-\lambda Z}}{1-e^{-\lambda Y}}\,dZ \\ +X'\left[\theta\phi'[x^+]\right]\theta\phi''[x^+](\theta\phi[x^+]-p^+x^++y[\theta])\,\frac{\lambda e^{-\lambda\left(K-X\left[\theta\phi'[x^+]\right]\right)}}{1-e^{-\lambda Y}}\end{pmatrix}$$

$$=-q+\frac{\epsilon}{2}\begin{pmatrix} \int_{K-X[p^+]}^{K-X\left[\theta\phi'[x^+]\right]}(p[K-Z]-p^+)\,\frac{\lambda e^{-\lambda Z}}{1-e^{-\lambda Y}}\,dZ \\ +(\theta\phi'[x^+]-p^+)\int_{K-X\left[\theta\phi'[x^+]\right]}^{Y}\frac{\lambda e^{-\lambda Z}}{1-e^{-\lambda Y}}\,dZ\end{pmatrix}\leq 0, \qquad (q+p^+<\theta\phi'[0]\leq p[Y^+]),$$

with equality holding if $x^+>0$. The first-order condition (A.2) is reproduced as (22) in Sub-section 4.2.

■

APPENDIX B

Proof of Proposition 6

If the social optimum can be decentralized as a competitive equilibrium, then Y^+ must satisfy simultaneously (59) and (48), which are reproduced, respectively, as follows:

(B.1) $\left(1-\frac{\epsilon}{2}\right)p[Y]=\frac{\epsilon}{2}\left(\frac{\lambda e^{-\lambda Y}}{1-e^{-\lambda Y}}\int_0^Y(\Phi[K-Z]-\Phi[K-Y])\frac{\lambda e^{-\lambda Z}}{1-e^{-\lambda Y}}dZ\right)$

and

(B.2) $p[Y]\left(1-\frac{\epsilon}{2}\right)=\frac{\epsilon}{2}\int_0^Y\frac{Z}{Y}p[K-Z]\frac{\lambda e^{-\lambda Z}}{1-e^{-\lambda Y}}dZ.$

It follows from (B.1) and (B.2) that if the social optimum can be decentralized, then the optimal level of operating reserves must satisfy the following equation:

(B.3) $\frac{\lambda e^{-\lambda Y}}{1-e^{-\lambda Y}}\int_0^Y(\Phi[K-Z]-\Phi[K-Y])\frac{\lambda e^{-\lambda Z}}{1-e^{-\lambda Y}}dZ=\int_0^Y\frac{Z}{Y}p[K-Z]\frac{\lambda e^{-\lambda Z}}{1-e^{-\lambda Y}}dZ.$

Now let n be a positive integer sufficiently large. We claim that when λ is sufficiently large, there is no value of $Y, \frac{1}{n}\leq Y\leq K-\frac{1}{n}$, that solves (B.3). To prove this claim, we prove specifically that

(B.4) $\frac{\lambda e^{-\lambda Y}}{1-e^{-\lambda Y}}\int_0^Y(\Phi[K-Z]-\Phi[K-Y])\frac{\lambda e^{-\lambda Z}}{1-e^{-\lambda Y}}dZ<\int_0^Y\frac{Z}{Y}p[K-Z]\frac{\lambda e^{-\lambda Z}}{1-e^{-\lambda Y}}dZ,\qquad\left(\frac{1}{n}\leq Y\leq K-\frac{1}{n}\right),$

or, equivalently, that

(B.5) $\int_0^Y(\Phi[K-Z]-\Phi[K-Y])\frac{\lambda e^{-\lambda Z}}{1-e^{-\lambda Y}}dZ<\int_0^Y\frac{Z}{Y}p[K-Z]e^{\lambda(Y-Z)}dZ.$

Note that the left-hand side of (B.5) is bounded above by $(\Phi[K] - \Phi[K - Y])$. As for the right-hand side of (B.5), it is bounded below by $\int_{Y/2}^{Y} \frac{1}{2} p[K] e^{\lambda(Y-Z)} dZ$, and this integral is large when λ is large. The claim is proved.

If $\overline{Y}$ is the number of generating plants that are run on the real-time market under the competitive equilibrium, then according to (B.2), we must have

(B.6) $p[\overline{Y}]\left(1 - \frac{\epsilon}{2}\right) = \frac{\epsilon}{2} \int_0^{\overline{Y}} \frac{Z}{\overline{Y}} p[K - Z] \frac{\lambda e^{-\lambda Z}}{1 - e^{-\lambda \overline{Y}}} dZ.$

Applying the claim, we can assert that

(B.7) $p[\overline{Y}]\left(1 - \frac{\epsilon}{2}\right) = \frac{\epsilon}{2} \int_0^{\overline{Y}} \frac{Z}{\overline{Y}} p[K - Z] \frac{\lambda e^{-\lambda Z}}{1 - e^{-\lambda \overline{Y}}} dZ > \frac{\epsilon}{2}\left(\frac{\lambda e^{-\lambda \overline{Y}}}{1 - e^{-\lambda \overline{Y}}} \int_0^{\overline{Y}} (\Phi[K - Z] - \Phi[K - Y]) \frac{\lambda e^{-\lambda Z}}{1 - e^{-\lambda \overline{Y}}} dZ\right).$

The strict inequality in (B.7) means that the curve

$$Y \to \left(1 - \frac{\epsilon}{2}\right) p[Y], \frac{1}{n} \leq Y \leq K - \frac{1}{n},$$

is above the curve

$$Y \to \frac{\epsilon}{2}\left(\frac{\lambda e^{-\lambda Y}}{1 - e^{-\lambda Y}} \int_0^{Y} (\Phi[K - Z] - \Phi[K - Y]) \frac{\lambda e^{-\lambda Z}}{1 - e^{-\lambda Y}} dZ\right), \frac{1}{n} \leq Y \leq K - \frac{1}{n},$$

at $Y = \overline{Y}$.

Next, note that $\int_0^{Y} (\Phi[K - Z] - \Phi[K - Y]) \frac{\lambda e^{-\lambda Z}}{1 - e^{-\lambda Y}} dZ$ tends to infinity when $Y \to K$. Indeed, applying the mean value theorem to Φ, we obtain

(B.8) $\Phi[K - Z] - \Phi[K - Y] = (Y - Z) p[K - Y + \theta(Y - Z)] > (Y - Z) p[K - Z]$

for some $0 < \theta < 1$. Hence

(B.9) $\int_0^{Y} (\Phi[K - Z] - \Phi[K - Y]) \frac{\lambda e^{-\lambda Z}}{1 - e^{-\lambda Y}} dZ > \int_0^{Y} (Y - Z) p[K - Z] \frac{\lambda e^{-\lambda Z}}{1 - e^{-\lambda Y}} dZ.$

Furthermore,

(B.9) $\int_0^Y (Y-Z)p[K-Z]\frac{\lambda e^{-\lambda Z}}{1-e^{-\lambda Y}}dZ = \int_0^Y (Y-Z)\Theta^{\frac{1}{\gamma}}(K-Z)^{-\frac{1}{\gamma}}\frac{\lambda e^{-\lambda Z}}{1-e^{-\lambda Y}}dZ,$

and when $Y \to K$, the integral on the right-hand side of (B.9) tends to

(B.10) $\int_0^K (K-Z)\Theta^{\frac{1}{\gamma}}(K-Z)^{-\frac{1}{\gamma}}\frac{\lambda e^{-\lambda Z}}{1-e^{-\lambda K}}dZ = \int_0^K \Theta^{\frac{1}{\gamma}}(K-Z)^{1-\frac{1}{\gamma}}\frac{\lambda e^{-\lambda Z}}{1-e^{-\lambda K}}dZ = \infty$

because $0 < \gamma < 1$.

We have just shown that the curve

$$Y \to \frac{\epsilon}{2}\left(\frac{\lambda e^{-\lambda Y}}{1-e^{-\lambda Y}}\int_0^Y (\Phi[K-Z]-\Phi[K-Y])\frac{\lambda e^{-\lambda Z}}{1-e^{-\lambda Y}}dZ\right), \frac{1}{n} \leq Y \leq K - \frac{1}{n},$$

is above the curve

$$Y \to \left(1-\frac{\epsilon}{2}\right)p[Y], \frac{1}{n} \leq Y \leq K - \frac{1}{n},$$

when Y is in a left neighborhood of K. Also, we have shown that the former curve is below the latter curve at $Y = \overline{Y}$. Hence the former curve must cross the latter curve at a point, say $Y = \hat{Y}$, with $\hat{Y} > \overline{Y}$, $\hat{Y}$ being the number of generating plants chosen to run in the real-time energy market under the social optimum.■

REFERENCES

CRAMPTON, PETER, AXEL OCKENFELS, AND STEVEN STOFT (2013): "Capacity Market Fundamentals," *Economics of Energy & Environmental Policy* 2, 27-46.

CRETI, ANNA AND NATALIA FABRA (2007): "Supply security and short-run capacity markets for electricity," *Energy Economics* 29, 259-276.

HOBBS, BENJAMIN, JAVIER INON AND STEVEN STOFT (2001): "Installed Capacity and Price Caps: Oil on Water or Fuel on the Fire?" *Electricity Journal* 14, 23-34.

JUST, SEBASTIAN AND CHRISTOPH WEBER (2008): "Pricing of Reserves: Valuing System Reserve Capacity against Spot Prices in Electricity Markets," *Energy Economics* 30, 3198-3221.

OREN, S. S. (2002): "Auction Design for Ancillary Reserve Products," *IEEE Power Engineering Society Summer Meeting,* Vol. 3, 1238-1239.

Newberry David and Stiglitz (1982): "The Choice of Techniques and The Optimality of Market Equilibrium with Rational Expectations" *Journal of Political Economy* 90, 2, 223-246.

JOSKOW, PAUL AND JEAN TIROLE (2007): "Reliability and Competitive Electricity Markets," *The Rand Journal of Economics* 38, 60-82.

Chapter Three: RELIABILITY IN AN ELECTRIC GRID: A MECHANISM DESIGN APPROACH

1. INTRODUCTION

In an ordinary market, demand often fluctuates randomly, and a firm – in order to meet unexpected demand – builds up inventory that can be drawn down when such an event arises. Because electricity cannot be stored economically, some generating plants must be spinning as operating reserves and be ready to inject power into the pool when an unexpected rise in demand occurs or when a large generating plant in the real-time energy market fails. If the price on the real-time energy market is flexible, it will rise when demand exceeds the available supply, and this gives the signal to consumers to cut back consumption until the equality between demand and supply is re-established. The high spot price in the real-time energy market also provides the incentive for generators to keep generating assets as operating reserves and run them to produce energy during peak periods or to build new generating capacity. In this manner, argue proponents of non-intervention, the market – if left alone – will function efficiently.

A non-interventionist policy, however, is not socially optimal because electricity suffers from a serious market failure. Because electricity is an essential commodity, its demand is extremely in-elastic. Furthermore, because most electricity consumers are on traditional meters which record their consumption over a period of time, and their contacts with the LSE's (Load Serving Entities or Electricity Retailers) allow them to consume any amount of electricity at a given price, they do not change their pattern of consumption when there is supply shortage. In this manner, there is no market mechanism to induce these consumers to cut back their consumption when demand exceeds supply, i.e., the market cannot clear. Thus, there is no other alternative for the ISO (Independent System Operator) than to blackout some geographical areas in the distribution network.[14] A high installed capacity will be able to serve a heavy load while leaving a considerable margin to serve as operating reserves and increase the reliability of the system. However, building new capacity is costly, and because a consumer cannot purchase reliability for her own protection and exclude other consumers from enjoying the reliability that she pays for, reliability has the characteristic of a public good, and we can expect that it will be under-provided. Crampton, Ockenfels, and Stoft (2013) assert that the market – if left alone – cannot optimize blackouts. The reason is simple: in order to find the optimal level of reliability, one has to know the preferences for electricity consumption of consumers. In electricity markets, the preferences for electricity consumption of a consumer are represented by her value of lost load (VOLL).[15] However, the VOLL of a consumer is private information, and – according to these authors – cannot be elicited from the consumers, and this is the reason why markets are unable to optimize blackouts.

The economic literature dealing with the public-good nature of operating reserves is sparse. Jaffe and Felder (1996) apply the traditional model of externality in pollution to determine the optimal level of operating reserves. Using the well-known prescription that to correct the under-provision of operating reserves a Pigovian tax – a subsidy in this case – should be applied, and the tax should be set at the level that induces an equality between marginal costs of capacity investment and marginal social benefits. However, as mentioned by these researchers, in their simple model of demand and supply of the Marshallian tradition, it is not clear how marginal social benefits should be measured. The same approach has also been adopted by Admunsen and Bergman (2007), who go further than Jaffe and Felder by introducing the probability of blackouts as a function of the level of operating reserves available. However, it is not explained how the probability of blackouts

[14] Technically, it is difficult to interrupt the electricity flowing into a single home. Thus when load shedding is required to maintain the integrity of the system, the ISO will choose to blackout an entire geographical area in the electric grid.

[15] The VOLL of a consumer is her willingness to pay to avoid the cutback of one unit of electricity consumption. Another related concept is the probability of lost load (LOLP), i.e., the probability that load must be shed to bring demand into alignment with supply. In an electric power system, VOLL and LOLP are the two practical measures often used to describe the security and reliability of the system.

arises. As in the model formulated by Jaffe and Felder, it is also not clear how marginal social benefits are measured in the model of Admunsen and Bergman. Joskow and Tirole (2007) formulate a model in which the behavior of consumers – those who can respond to real-time price changes and those on traditional meters who do not respond to real-time price changes – and the non-price rationing of retail consumers to equate demand and supply in real time are analyzed in detail. These researchers find the optimal generation capacity investment when demand is uncertain and when the system might collapse by solving the central planner's problem. The multipliers that come out of this programming problem serve as prices in a competitive market that decentralize the social optimum. However, as remarked by these researchers, the conditions required for the competitive market to duplicate the social optimum are quite strong. The specification of individual demand adopted by Joskow and Tirole makes it possible for the ISO to figure out the preferences for electricity consumption of each consumer, and thus their model does not deal with the issue of how to find the optimal level of reliability when the VOLL is not known.

In this paper, we present a model of reliability for an electric grid when consumers' preferences for electricity consumption are private information. In the model, time is continuous, and the time horizon is one cycle, say one day. Most of the economic models dealing with the problem of reliability in an electric grid assume that the uncertainty in the system arises from the random nature of demand, and ignore the uncertainty due to the failure of generating plants. Our model, in contrast, takes the load profile as a given deterministic function of time, and takes the failure of generating plants as the source of uncertainty in the system. The model we formalize thus provides an analysis of operating reserves from another angle. In the model, part of the generation capacity is online to serve the load at each instant. The remaining generation capacity serves as operating reserves. The number of generating units that serve the load at any instant – if there is enough generation capacity to serve the load – will provide exactly the volume of power that meets the load. The time at which a generating plant failure occurs follows a Poisson process with a parameter that depends on the number of generating units that are currently online to generate electricity in order to meet the load. When a generating plant failure occurs, the number of generating units that fail is also random. The higher the total generation capacity of the system and a fortiori the higher the operating reserves at any instant during the cycle, the better the system can handle a major generating breakdown. When a generating plant failure is massive, the generating capacity that remains will not be able to meet the load for the remaining part of the cycle. Under such a scenario, the ISO must shed the load in some parts of the electric grid, with the ensuing loss of welfare by the agents who reside in the areas of the electric grid whose loads are shed by the ISO. Certainly one can avoid load shedding by building a large generation capacity, which is twice the peak demand in our model. However, such a strategy is not optimal because capacity investment is costly and because of the low marginal social benefits associated with a high generation capacity investment. Thus, it is always optimal to accept some possible blackouts during the cycle. The model we formulate yields both the optimal level of generation capacity investment and the blackout strategy.

To model the public good nature of operating reserves and the asymmetric information concerning the consumers' preferences for electricity consumption, we adopt the mechanism design approach. In the model, consumers are asked to reveal their preferences for electricity consumption, and this information is then used to determine the generation capacity investment as well as the contribution required from each agent to finance the generation capacity investment. After the generation capacity has been built, the price that an agent pays for 1MW is equal to the marginal cost of generating electric power. We show (Proposition3) that an ex post efficient social choice function is truthfully implementable in Bayesian Nash equilibrium. The direct revelation mechanism used to implement such an ex post efficient social choice function satisfies the balance budget constraint (no external funds are needed to finance the generation capacity investment) and the ex post participation constraint (all agents are willing to participate in the mechanism). Our paper is not the first one to model the electricity market from the perspective of mechanism design. Hobbs et al. (2000) describe a model of auctions for capacity in electric power and gas using the Groves-Clarke mechanism. It is well known that truth telling is a dominant strategy for a bidder in the Groves-Clarke mechanism.

This result is powerful because there is no need for the auctioneer to know the prior distribution concerning the types of the agents. However, it is also widely known that the Groves-Clarke mechanism does not satisfy the balance budget constraint, and subsidies must be provided to support the mechanism. Finally, if a consumer does not obtain at least her reservation utility, she will not participate in the Groves-Clarke mechanism. The direct mechanism used in our model is known as the expected externality model satisfies the incentive compatibility condition and the balance budget condition. It also provides some sufficient conditions (Proposition 4) that ensure the participation constraint holds for each agent.

The paper is organized as follows. The model – because of its length and complexity – is presented in several stages. First, preferences and technologies are described in Section 2. The second-best blackout strategy is defined and found in Section 3. The formulation and solution of the blackout sub-model takes as given the generation capacity investment already made and the preferences of all the consumers. In Section 4, the social choice function, which dictates the generation capacity investment and the transfers needed to finance the investment, is presented. Section 5 proposes how to implement a social choice function by the direct revelation mechanism. In Section 6, we prove that an ex post efficient[16] social choice function is truthfully implementable in Bayesian Nash equilibrium. The proof of this result involves a minor extension of that in the expected externality mechanism.[17] The participation constraint is discussed in Section 7. Section 8 contains some concluding remarks.

2. PREFERENCES AND TECHNOLOGIES

Consider an electric grid, which – from the distribution perspective – consists of I non-overlapping geographical areas, where I is a positive integer. In what follows an

[16] An ex post efficient social choice function gives the generation capacity investment that maximizes social welfare, given the preferences of all the consumers.

[17] See MasColell et al. (1995), Chapter 23, p. 885-887.

area in the electric grid is indexed by $i, i = 1, \ldots, I,$ and is denoted by A_i. An area in the electric grid represents a distribution subsystem in the grid into which electric flow can be technically interrupted at any point in time. In the Netherlands, an area in the electric grid that the ISO can cut off from the remaining part of the electric grid is generally about 10 by 10km in size, which is the size of an average municipality.[18] For each $i = 1, \ldots, I,$ let N_i be a positive integer representing the number of consumers in A_i. A consumer of electricity is identified by a double subscript in, with i representing the area in which she resides and $n, n = 1, \ldots, N_i$, her co-ordinate in this area. In the model, time is continuous and denoted by t. The time horizon for the model is the unit interval $0 \leq t \leq 1$. We can interpret the unit time interval as the length of a cycle, say a day.

2.1. Preferences

At any instant, an agent consumes two goods: electricity and a numéraire. The instantaneous utility function of agent in is linear and assumes the following form:

(1) $u_{in}[x_{in}, y_{in}] = \theta_{in} \min[x_{in}, \bar{x}_{in}] + y_{in}$[19],

Where $\bar{x}_{in}$ is a parameter that represents maximum amount of electricity that is consumed while x_{in} and y_{in} denote, respectively, her consumption of electricity and her consumption of the numéraire. As for θ_{in}, it is a parameter that represents her preferences for electricity consumption, and is assumed to be private information. In the literature of electric power economics, θ_{in} is known as the value of loss load (VOLL) of this consumer. It represents the agent's willingness to pay to avoid the reduction in consumption of one unit of electricity. We suppose that $\theta_{in} \in [\underline{\theta}_{in}, \overline{\theta}_{in}]$, where $\underline{\theta}_{in}$ and $\overline{\theta}_{in}$ are two numbers satisfying $0 < \underline{\theta}_{in} < \overline{\theta}_{in}$.

A profile of agents' types is a list $\theta = (\theta_{in})_{i=1,\ldots,I,n=1,\ldots,N_i}$. The agents' profile of types is drawn from a commonly known prior probability distribution $\Phi[\theta]$, and only agent in observes the value of θ_{in}. In what follows, $\Phi_{in}[\theta_{in}]$ denotes the marginal

[18] See Noij et al. (2009).

[19] For simplicity we assume that $x_{in} < \bar{x}_{in}$. I would like to thank Professor Jacques Robert for pointing this functional form out.

probability distribution of the type of agent in. For our model, we assume that agents' types are statistically independent.

Let $x_{in}[t], 0 \leq t \leq 1$, be the demand for electricity at time t by agent in. Although the demand for electricity by a consumer fluctuates through time, we choose not to model demand uncertainty by assuming that $x_{in}[t]$ is non-random. We also assume that $x_{in}[t]$ is completely in-elastic. The assumption that demand is non-random allows us to concentrate on the failure of generating plants as the source of uncertainty. Incorporating demand uncertainty into the model makes the exposition more complicated without changing its message. The curve $t \to x_{in}[t], 0 \leq t \leq 1$, thus represents the load profile of consumer in. The utility that agent in obtains from consuming $x_{in}[t]$ units of electricity at time t is then given by $\theta_{in} x_{in}[t]$.

2.2. *Technologies*

Let K be the total generation capacity (in MW) chosen by the central planner. To eschew the problem of indivisibility, we take K to be a positive real number. Also, for simplicity, we assume that each generating plant has a capacity of 1 MW. The investment cost to build a generating plant is constant and normalized to be equal to 1. The marginal cost incurred by a generating plant at an instant when it is generating electricity is $c > 0$.

Let $X_i[t] = \sum_{n=1}^{N_i} x_{in}[t]$ denote the load in A_i at time t, and $X[t] = \sum_{i=1}^{I} X_i[t]$ denote the global load at time t. The number of generating units required to serve the global load at each instant is then given by $Y[t] = X[t], 0 \leq t \leq 1$. If the global load at an instant t does not exceed the generation capacity of the electric grid, then the operating reserves at that instant is $Y^+[t] = K - Y[t]$.

A complete modeling of generating plant failures must specify the properties of each generating plant and its probability of failure as well as the random time it takes to repair and bring it back online. At the present time, such a comprehensive approach is not available even among electric power engineers. At the aggregate and more manageable level for a simulation exercise, Anderson and Davison (2008) model the time to failure (TTF) and the time to repair (TTR) with the help of a two-parameter Weibull distribution, and the parameters are determined by the individual characteristics as well as the history of the generating unit. We eschew such a detailed modeling strategy[20] by assuming that the event of a subset of generating plants online breaking down occurs at most once during the time interval $0 \leq t \leq 1$, and that once a generating unit fails, it will not be repaired and be brought back online before time $t = 1$.

At any instant $t, 0 \leq t < 1$, when $Y[t]$ generating units are online and when no generating plant has failed, the probability that a subset of these generating units fail during an infinitesimal time interval $[t, t + \Delta t)$ is assumed to be given by

(2) $\epsilon[Y[t]]\Delta t + o[\Delta t]$,

where $\epsilon[Y]$ is a smooth, increasing function of $Y \geq 0$, and $\epsilon[0] = 0$. Also, $o[\Delta t]$ is an expression in Δt, with $\frac{o[\Delta t]}{\Delta t} \to 0$ when $\Delta t \to 0$.

The number of generating plants, say Z, that fail under such an event is also random, and is characterized by a distribution function $F_Y[Z], 0 \leq Z \leq Y$, with Y being the number of generating units being online at the time a breakdown occurs. We shall assume that $F_Y[Z]$ is smooth, non-decreasing, and satisfies the following conditions: $F_Y[0] = 0, F_Y[Y] = 1$. The density of $F_Y[Z]$ is denoted by $f_Y[Z]$. A convenient functional form for F_Y is the exponential function $F_Y[Z] = \frac{1-e^{-\lambda Z}}{1-e^{-\lambda Y}}$, $0 \leq Z \leq Y$, where $\lambda > 0$ is a parameter. The density of this distribution function is $f_Y[Z] = \frac{\lambda e^{-\lambda Z}}{1-e^{-\lambda Y}}$.

Let $\pi[t], 0 < t < 1$, denote the probability that no generating unit has failed by time t. The probability that no generating unit fail by time $t + \Delta t$ is then given by

[20] Adopting such a modeling strategy will turn the ISO's problem into a large stochastic control problem that can only be solved numerically.

$$\pi[t+\Delta t] = \pi[t]\big(1-\epsilon\big[X[t]\big]\Delta t + o[\Delta t]\big);$$

that is,

$$\pi[t+\Delta t] - \pi[t] = -\pi[t]\epsilon\big[X[t]\big]\Delta t + \pi[t]o[\Delta t].$$

Dividing the preceding expression by Δt, and then letting $\Delta t \to 0$, we obtain

$$\frac{d\pi}{dt} = -\pi[t]\epsilon\big[X[t]\big].$$

The solution of this differential equation is

(3) $$\pi[t] = \pi[0]e^{-\int_0^t \epsilon[X[t\prime]]dt\prime} = e^{-\int_0^t \epsilon[X[t\prime]]dt\prime},$$

where the second equality has been obtained by using the initial condition $\pi[0] = 1$. Using (3), we can then assert that a breakdown occurs by time $t, 0 \leq t \leq 1$, with probability

(4) $$p[t] = 1 - \pi[t] = 1 - e^{-\int_0^t \epsilon[X[t\prime]]dt\prime}.$$

Figure 1 depicts a possible evolution of the electric grid during the time interval $0 \leq t \leq 1$. Under the scenario depicted in Figure 1, the electric grid suffers a major breakdown at time τ, when Z generating units fail. Before the breakdown, the generating capacity of the electric grid is K. After the breakdown, the total generating capacity that remains and is available on the real-time energy market is $K - Z$, which stays at this level until time $t = 1$.

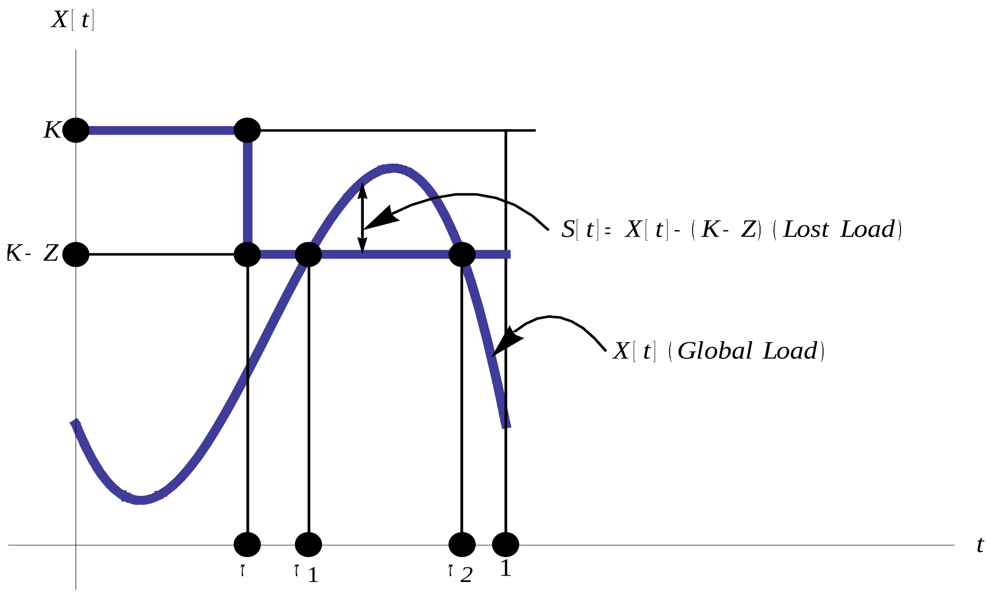

Figure 1. The lost load : $S[t] = X[t] - (K - Z), \tau_1 \leq t \leq \tau_2$

Define

$$\tau_1 = inf\{t \in [0,1] | X[t] \geq K - Z\},$$

$$\tau_2 = sup\{t \in [0,1] | X[t] \geq K - Z\}.$$

Note that τ_1 and τ_2 both depend on $(K - Z)$. We shall assume – for technical reasons – that $X[t] \geq K - Z$ for all $t \in [\tau_1, \tau_2]$. The time τ when the major breakdown occurs might be inside or outside the time interval $[\tau_1, \tau_2]$. If the breakdown occurs before time τ_1, as is the scenario depicted in Figure 1, then the constraint on the remaining generating capacity only begins at time τ_1, and the time interval $[\tau_1, \tau_2]$ represents the duration during which load must be shed to maintain the integrity of the system. If $\tau \in [\tau_1, \tau_2]$, then the time interval during which load must be shed is $(\tau, \tau_2]$. If $\tau > \tau_2$, the breakdown has no impact on the system under the time horizon considered for our model. In what follows, we let $S[t] = X[t] - (K - Z), \tau_1 \leq t \leq \tau_2$, be the load that must be shed at time t to maintain the integrity of the system.

Observe that for a given value of Z, the time interval $[\tau_1, \tau_2]$ shrinks when K increases. That is, the lost load, the duration of the lost load, and the probability of load shedding all decline when K increases. However, a larger generating capacity requires more investment. The optimal level of generating capacity is determined by balancing the benefits of power interruptions with the cost of building capacity.

3. THE BLACKOUT STRATEGY

Because it is technically much easier to blackout an entire region than cutting off the electricity flowing into a single household, a blackout strategy carried out by the ISO must be a second-best solution. That is, a blackout program – after a major breakdown has occurred – should specify which areas – not which consumer – in the electric grid must be cut off and at what time. This is a technically challenging problem in power engineering and applied mathematics.[21]

In our model, to solve the second-best problem of blackouts, it is necessary to presume that the agents' profile of types $\theta = (\theta_{in})_{i=1,\dots,I,n=1,\dots,N_i}$ is known. At each instant $t, 0 \leq t \leq 1$, the demand for electricity in area A_i is $X_i[t]$ and the total utilities of all the consumers in this area is $\sum_{n=1}^{N_i} \theta_{in} x_{in}[t]$ if the load $X_i[t]$ is served. The ordered pair $\left(X_i[t], \sum_{n=1}^{N_i} \theta_{in} x_{in}[t]\right)$ thus represents the load and the total utilities of all the consumers in area A_i at time t if the load is served. If the ISO chooses to shed this load at that instant, then instantaneous social welfare will fall by $\sum_{n=1}^{N_i} \theta_{in} x_{in}[t]$.

In our model, a blackout program involves the specification of I binary maps $\delta_i: [\tau, 1] \to \{0,1\}, i = 1, \dots, I$, with $\delta_i[t] = 0$ indicating that the load in area i is shed at time t, while $\delta_i[t] = 1$ indicating that the load in area i is served at time t. A blackout program $(\delta_i)_{i \in I}$ is feasible if (i) each binary map δ_i is piecewise right-continuous and (ii) if the total load served at any instant is less than or equal to the generation capacity that remains after the major breakdown, i.e., if

(5) $\sum_{i=1}^{I} \delta_i[t] X_i[t] \leq K - Z, \tau \leq t \leq 1$.

[21] See, for example, Trodden et al. (2013)

The social welfare yielded at time t by a feasible blackout program $(\delta_i)_{i \in I}$ is

(6) $\sum_{i=1}^{I} \delta_i[t] \sum_{n=1}^{N_i} (\theta_{in} - c) x_{in}[t], \tau \leq t \leq 1$.

The second-best blackout program after a major breakdown is obtained by solving the following constrained maximization problem

(7) $max_{(\delta_i)_{i \in I}} \int_{\tau}^{1} \left(\sum_{i=1}^{I} \delta_i[t] \sum_{n=1}^{N_i} (\theta_{in} - c) x_{in}[t] \right) dt$

subject to (5).

To solve the constrained maximization problem stated as (7), let $\mathfrak{P}[I]$ be the power set of $\{1, \dots, I\}$; that is, $\mathfrak{P}[I]$ is the set consisting of all the subsets of $\{1, \dots, I\}$. The set $\mathfrak{P}[I]$ has 2^I elements, and an element of $\mathfrak{P}[I]$, say $\mathcal{J}$, is a subset of $\{1, \dots, I\}$. By abuse of language, we shall refer to the collection of areas with indices in $\mathcal{J}$ as region $\mathcal{J}$.

For each $\mathcal{J} \in \mathfrak{P}[I]$ and each time $t, 0 \leq t \leq 1$, let

(8) $X_{\mathcal{J}}[t] = \sum_{i \in \mathcal{J}} X_i[t]$

and

(9) $W_{\mathcal{J}}[t] = \sum_{i \in \mathcal{J}, n=1,\dots,N_i} \theta_{in} x_{in}[t]$

denote, respectively, the load at time t in region $\mathcal{J}$ and the social welfare obtained – also at time t and in region $\mathcal{J}$.

At any instant after the breakdown, the ISO can only serve the load in region $\mathcal{J}$ if this load does not exceed the remaining generation capacity of the grid, i.e., if

(10) $X_{\mathcal{J}}[t] \leq K - Z$.

To capture the capacity constraint that the electric grid imposes upon a region, say region $\mathcal{J}$, first let

(11) $T[\mathcal{J}|Z, K] = \{t \in [0,1] | X_{\mathcal{J}}[t] > K - Z\}$

be the set of instants at which the constraint (10) is violated, given the generation capacity investment K and the number of generating units Z that fail when a major breakdown occurs. Note that the set $T[\mathcal{J}|Z,K]$ shrinks when K increases.

Next, let $t \to \chi_{T[\mathcal{J}|Z,K]}[t], 0 \leq t \leq 1$, be the characteristic function of $T[\mathcal{J}|Z,K]$, i.e.

(12) $$\chi_{T[\mathcal{J}|Z,K]}[t] = \begin{cases} 0 \; if t \in T[\mathcal{J}|Z,K] \\ 1, otherwise. \end{cases}$$

The time path of the total welfare in region $\mathcal{J}$ after a break down is then given by

(13) $$t \to \chi_{T[\mathcal{J}|Z,K]}[t] W_{\mathcal{J}}[t], \tau \leq t \leq 1, \qquad (\mathcal{J} \in \mathfrak{P}[I]).$$

PROPOSITION 1: *Suppose that for any two regions $\mathcal{J}$ and $\mathcal{J}'$ the two curves $t \to W_{\mathcal{J}}[t], \tau \leq t \leq 1$, and $t \to W_{\mathcal{J}'}[t], \tau \leq t \leq 1$, cross each other at most a finite number of times. For each $t, \tau \leq t \leq 1$, let*

(14) $$\mathcal{J}^*[t|\tau, Z, K, \theta] = argmax_{\mathcal{J} \in \mathfrak{P}[I]} \chi_{T[\mathcal{J}|Z,K]}[t] W_{\mathcal{J}}[t],$$

and

(15) $$\delta_i^*[t|\tau, Z, K, \theta] = \begin{cases} 1 \; if \; i \in \mathcal{J}^*[t|\tau, Z, K, \theta] \\ 0, otherwise. \end{cases}$$

Then the family of binary maps

(16) $$\delta_i^*[.|\tau, Z, K, \theta]: t \to \delta_i^*[t|\tau, Z, K, \theta], \tau \leq t \leq 1, \qquad (i = 1, \ldots, I)$$

constitute the second-best blackout strategy, given that (i) the breakdown occurs at time τ; (ii) Z generating units fail when the breakdown occurs; (iii) K is the generation capacity investment; and (iv) θ is the type profile of the agents.

PROOF: Using the assumption that for any two regions $\mathcal{J}$ and $\mathcal{J}'$ the curves $t \to W_{\mathcal{J}}[t], \tau \leq t \leq 1$, and $t \to W_{\mathcal{J}'}[t], \tau \leq t \leq 1$, cross each other at most a finite number of times, we can assert that the family of curves $\chi_{T[\mathcal{J}|Z,K]}[t]W_{\mathcal{J}}[t], \tau \leq t \leq 1, \mathcal{J} \in \mathfrak{P}[I]$, cross each other at most a finite number of times, and it follows from this last result that the map $t \to \mathcal{J}^*[t|\tau, Z, K, \theta], \tau \leq t \leq 1$, changes its value at most a finite number of time. Furthermore, by definition $\mathcal{J}^*[t|\tau, Z, K, \theta]$ is the region that maximizes social welfare at any instant after the major breakdown. Hence the family of binary maps defined by (16), solve the constrained maximization problem represented by (7). ■

Note that we have made explicit the dependence of the second-best blackout strategy on (i) the time (τ) the breakdown occurs, (ii) the size (Z) of the breakdown, (iii) the total generation capacity (K) of the market, and (iv) the agents' profile of types θ.

For an agent, say in, the utilities she obtains from electricity consumption over the entire period $0 \leq t \leq 1$ is given by

(17) $\int_0^\tau (\theta_{in} - c)x_{in}[t]dt + \int_\tau^1 \delta_i^*[t|\tau, Z, K, \theta](\theta_{in} - c)x_{in}[t]dt$.

Note that in (17) the first integral represents the cumulative utilities from electricity consumption she obtains at the time before the major breakdown occurs, and the second integral represents the cumulative utilities she obtains over the remaining time horizon after the breakdown and under the second-best blackout program. Also, note that (17) has been obtained conditioned on τ (the exact time the major breakdown occurs) and Z (the number of generating units that fail). It is also worth pointing out that the second integral depends on the agent's own type θ_{in} as well as on θ, the type profile of all the agents, through $\delta_i^*[t|\tau, Z, K, \theta]$, the second-best blackout strategy. It is the second-best blackout strategy that serves as a conduit for the externalities that the preferences for electricity of an agent have upon the welfare of all the other agents in the electric grid after a major breakdown.

Taking the expectation of (17) with respect to the random time of generating plant failure (τ), and the random number of generating plants that fail (Z), we obtain the following expression for the expected utility from electricity consumption that such an individual obtains over the entire period $0 \leq t \leq 1$:

(18)

$$\int_0^1 \left(\int_0^\tau (\theta_{in} - c)x_{in}[t]dt + \int_0^{X[\tau]} \int_\tau^1 \delta_i^*[t|\tau, Z, K, \theta](\theta_{in} - c)x_{in}[t]\, f_{X[\tau]}[Z]dtdZ\right) dp[\tau]$$

$$+\pi[1] \int_0^1 (\theta_{in} - c)x_{in}[t]dt.$$

Note that in (18) the first line represents the agent's expected payoff under the event that a major breakdown occurs, while the second line represents the payoff under the event that no breakdown occurs during the time interval $[0,1]$, weighted by $\pi[1]$, the probability of this event. Also, note that in the first line of (18) we have used the condition that as long as no major breakdown has occurred, the number of generating units online $Y[t]$ is exactly equal to the global load $X[t]$, i.e., $Y[t] = X[t]$, and this means that $f_{Y[t]}[Z] = f_{X[t]}[Z]$.

If we sum (18) across agents, we obtain the expected social welfare, given the generation capacity investment K and the type profile θ. However, this way of computing the expected social welfare conditioned on K and θ involves $\delta_i^*[t|\tau, Z, K, \theta], i = 1, \dots, I$. A more efficient way of computing the expected social welfare is to apply Proposition 1. To this end, let

(19) $$\Omega[K, \theta, \tau, Z] = \int_0^\tau W_{\{1,\dots,I\}}[t]dt + \int_\tau^1 \left(max_{\mathcal{J} \in \mathfrak{P}[I]} \chi_{T[\mathcal{J}|Z,K]}[t] W_{\mathcal{J}}[t]\right) dt$$

denote the social welfare obtained, given the generation capacity investment K, the type profile θ, the time the major breakdown occurs τ, and the number of generating units that fail when the breakdown occurs.

Taking the expectation of (19) with respect to τ and Z, we obtain the following expression for the expected social welfare, given the generation capacity investment K and the type profile θ:

(20) $\Omega[K,\theta] = \int_0^1 \int_0^{X[\tau]} \Omega[K,\theta,\tau,Z] f_{X[\tau]}[Z] dZ dp[\tau]$.

PROPOSITION 2: *During the load shedding phase, a rise in the preferences for electricity consumption of an agent, ceteris paribus, either improves or leaves unchanged the chance*[22] *that the load of a region in which the agent resides will be served and lowers or leaves unchanged the chances in a region in which the agent does not reside uniformly for all times after the breakdown.*

PROOF: Consider an agent, say agent in, and suppose that her preferences for electricity consumption rise from θ_{in} to $\theta_{in} + \Delta\theta_{in}$. The change in the agent's preferences for electricity consumption raises the welfare component of area A_i by $x_{in}[t]\Delta\theta_{in}$ at any time $t \geq \tau$. For any region $\mathcal{J}$ of which A_i is a part, its welfare also rises by $x_{in}[t]\Delta\theta_{in}$ at any time $t \geq \tau$. For each of the regions of which A_i is not a part, its welfare remains unchanged. Therefore, the chance that the load in A_i will be served during the blackouts will improve at the expense of all the regions of which A_i is not a part. Because the ISO cannot interrupt the electric flow into the residence of a single agent, but can either serve or shed the load of an entire area, the rise in the agent $in's$ preferences for electricity consumption benefit all the agents in A_i.■

An Example: To illustrate the economic contents of Propositions 1 and 2, consider an electric grid that consists of three areas, and in each area resides a single consumer. In this setting, we have $I = 3$, and $N_i = 1, i = 1, \ldots, 3$. The power set of $\{1,2,3\}$, namely $\mathfrak{P}[I]$, has 8 elements that we list as follows: $\emptyset(the\ empty\ set), \{1\}, \{2\}, \{3\}, \{1,2\}, \{1,3\}, \{2,3\}$, and $\{1,2,3\}$. Figure 2 depicts the load profiles of all the regions in the lectric grid, except for the empty region.

After the break down that has been depicted in Figure 1, the remaining generation capacity is insufficient to meet the global load of the entire grid during the time interval (τ_1, τ_2). Furthermore, the total load of region $\{2,3\}$ exceeds the remaining generation capacity during the time interval (τ_3, τ_4). However, it is possible for the

[22] At each instant after the breakdown, the chance that the load in a region is served is either 0 or 1.

ISO to serve exactly one of the regions other than $\{2,3\}$ and $\{1,2,3\}$. Observe that that the load profiles of any two regions in the grid – as depicted in Figure 2 – do not cross each other. This is certainly the case if the consumers' load profiles are identical up to a scaling factor.[23]

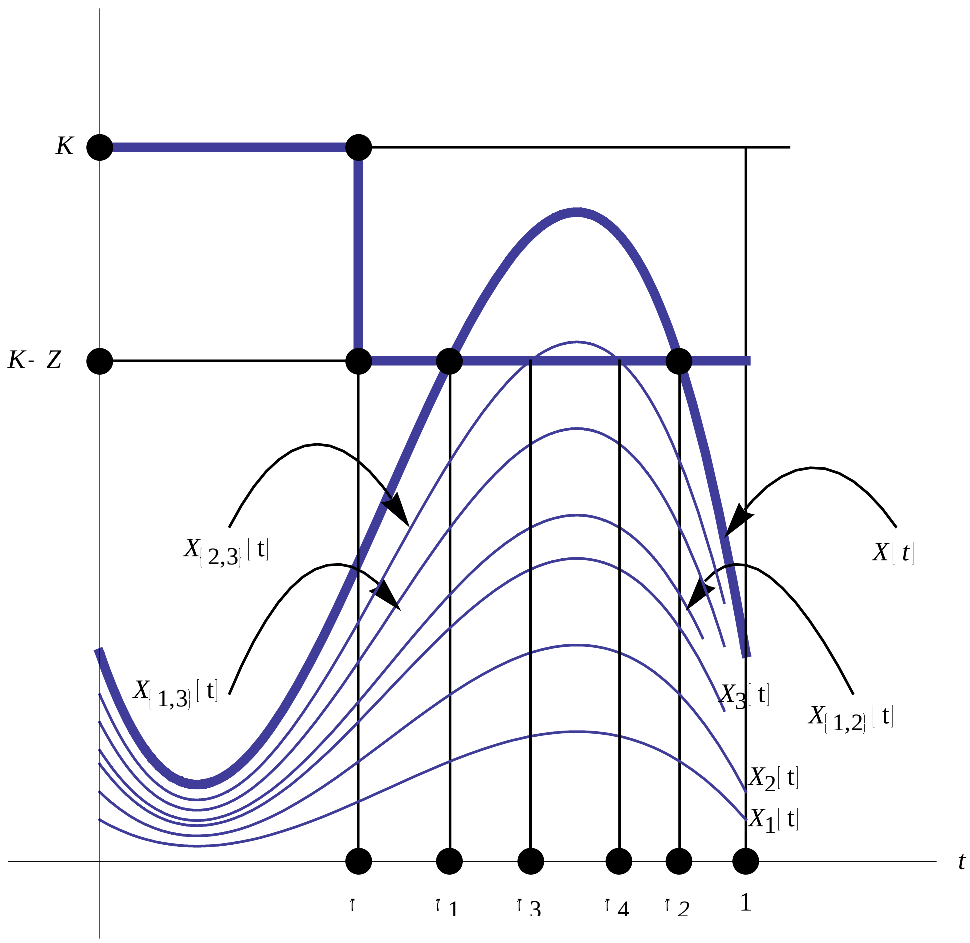

Figure 2. The load profiles of the various regions in the electric grid.

Associated with the load profile of each region is its welfare profile if the load in the region is not shed. Figure 3 depicts the welfare profiles of all the proper regions of the electric grid. In drawing these welfare profiles, we assume that the preferences

[23] This is the assumption made in the model of Joskow and Tirole.

for electricity consumption of the three consumers in the three areas that make up the electric grid are given, respectively, by $\theta_1 = 0.47, \theta_2 = 0.16, \theta_3 = 0.39$.

As can be seen from Figure 3, the welfare profile of region $\{1,3\}$ is above those of all the other regions. The second-best blackout strategy is obvious: after a major breakdown, the ISO should shed the load of area A_2 and serve only region $\{1,3\}$. The thick curve in Figure 3 represents the second-best blackout strategy. This result reflects the low preferences $(\theta_2 = 0.16)$ for electricity consumption of the consumer in A_2 relative to those of the other two consumers $(\theta_1 = 0.47,\ \theta_3 = 0.39)$. The social welfare obtained over the entire horizon condition on $K, \theta = (0.47, 0.16,\ 0.39), \tau$, and Z is

(21) $\Omega[K, (\theta_1 = 0.47, \theta_2 = 0.16, \theta_3 = 0.39), \tau, Z]$

$$= \int_0^{\tau_1} \sum_{i=1}^{3} \theta_i X_i[t]\, dt + \int_{\tau_1}^{\tau_2} (\theta_1 X_1[t] + \theta_3 X_3[t]) dt + \int_{\tau_2}^{1} \sum_{i=1}^{3} \theta_i X_i[t]\, dt.$$

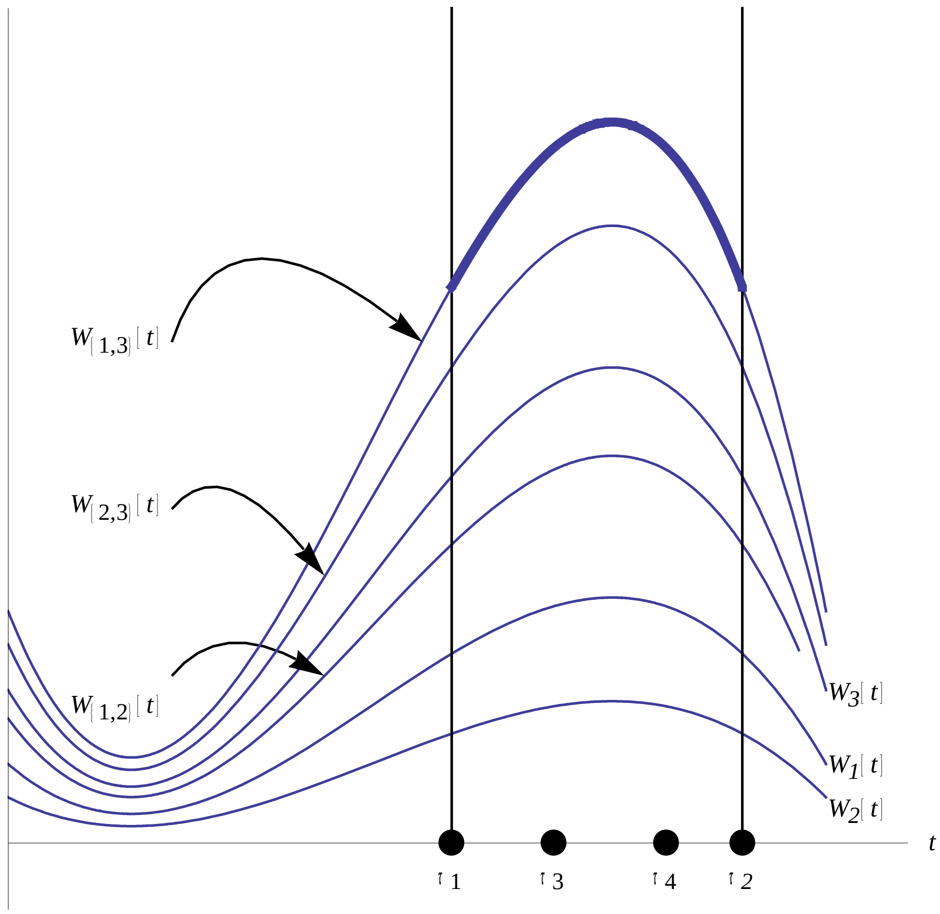

Figure 3. The welfare profiles of the various regions in the electric grid: ($\theta_1 = 0.47, \theta_2 = 0.16, \theta_3 = 0.39$)

Now let us raise the preferences for electricity consumption of the consumer in A_2, say to $\theta_2 = 0.45$, without changing those of the other two consumers. Also, let us maintain the load profile of each consumer so that Figure 2 still applies. Figure 4 now depicts the welfare profiles of the various regions. The higher preferences for electricity consumption of the consumer in A_2 now shift the welfare profile curve of region $\{2,3\}$ to the upper reaches of Figure 4.

The second-best blackout strategy sheds the load of area A_1 from time τ_1 till time τ_3. During the time interval (τ_3, τ_4), the total load of region $\{2,3\}$ exceeds the remaining generation capacity of the system, and so the ISO sheds the load of A_2. During the time interval (τ_4, τ_2), the ISO again chooses to meet the demand of area A_2 and sheds the load of area A_1.

During the shortage caused by the breakdown, the load in area A_3 is always served because the demand for electricity by and the preferences for electricity consumption of the consumer in A_3 are high. Although the preferences for electricity consumption of the consumers in area A_1 and area A_2 are the same, the demand for electricity is much lower in A_1 than in A_2, and this induces the ISO to favor area A_2 over A_1 at the beginning and near the end of the blackouts. In the middle phase of the shortage, the ISO chooses to shed the load in A_2 because she cannot serve both A_2 and A_3 and because the load in the latter area is much higher than that of the former although the preferences for electricity consumption in A_2 and A_3 are about the same.

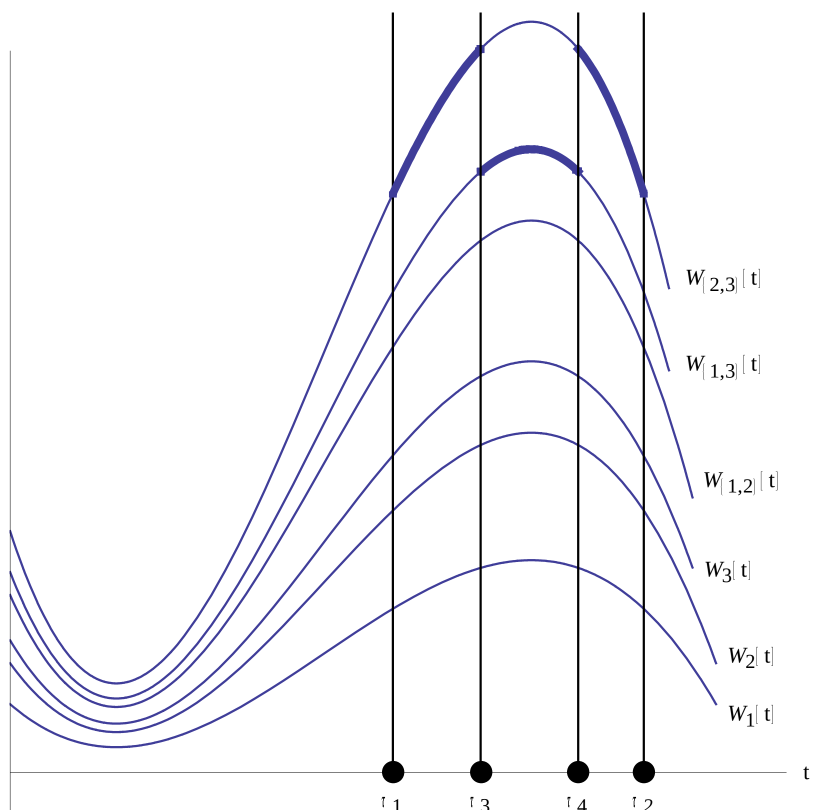

Figure 4.□ The evolution of social welfare (the curve constituted by the three thick arcs in the figure) during the blackouts: $(\theta_1 = 0.47, \theta_2 = 0.45, \theta_3 = 0.39)$

The broken curve constituted by the three thick arcs in Figure 4 represents the second-best blackout strategy. The social welfare obtained over the entire horizon condition on $K, \theta = (0.47, 0.16,\ 0.39), \tau$, and Z is

(22) $\Omega[K, (\theta_1 = 0.47, \theta_2 = 0.45, \theta_3 = 0.39), \tau, Z] =$

$$\int_0^{\tau_1} \sum_{i=1}^3 \theta_i X_i[t]\, dt + \int_{\tau_1}^{\tau_3} (\theta_2 X_2[t] + \theta_3 X_3[t]) dt$$

$$+ \int_{\tau_3}^{\tau_4} (\theta_1 X_{12}[t] + \theta_3 X_3[t]) dt + \int_{\tau_2}^{1} \sum_{i=1}^3 \theta_i X_i[t]\, dt.$$

On comparing Figure 3 and Figure 4, we can see that before the rise in the preferences for electricity consumption of the consumer in A_2, the ISO sheds the load in this area and serves the loads in A_1 and A_3 during the blackouts. After the rise, the load of A_3 is still served during the blackouts, but now A_2 displaces A_1 in the list of area being served in the two intervals $[\tau_1, \tau_3], [\tau_4, \tau_2]$. That is, the rise in the preferences for electricity consumption of the consumer in A_2 benefits A_2 at the expense of A_1 during the blackouts.

4. THE SOCIAL CHOICE FUNCTTION

A collective choice for the I agents is a list of non-negative real numbers, say $\left(K, (k_{in})_{i=1,\dots,I, n=1,\dots,N_i}\right)$, with the following properties:

(23) $K = \sum_{i=1,\dots,I, n=1,\dots,N_i} k_{in}$.

Here K is the total generation capacity investment for the I agents, and k_{in} is the part of the total generation capacity investment that agent in must contribute. Condition (23) is the balance budget constraint.

A social choice function is a map

(24) $\mathfrak{F}: \theta \to \mathfrak{F}[\theta] = \left(K[\theta], (k_{in}[\theta])_{i=1,\dots,I, n=1,\dots,N_i}\right)$

that associates a social alternative $\left(K[\theta],(k_{in}[\theta])_{i=1,\dots,I,n=1,\dots,N_i}\right)$ with each profile of types θ. The problem faced by the regulator is to find a social choice function, and then implement it.

If the regulator knows the type of each agent and if she implements a social choice function $\mathfrak{F}$, then the expected payoff for agent in is given by

$$(25)\quad \int_0^1 \begin{pmatrix} \int_0^\tau (\theta_{in}-c)x_{in}[t]dt \\ +\int_0^{X[\tau]}\int_\tau^1 \delta_i^*[t|\tau,Z,K[\theta],\theta](\theta_{in}-c)x_{in}[t]\, f_{X[\tau]}[Z]dtdZ \end{pmatrix} dp[\tau]$$

$$+\pi[1]\int_0^1(\theta_{in}-c)x_{in}[t]dt+\overline{y}_{in}-k_{in}[\theta],$$

Where $\overline{y}_{in}$ is agent $in's$ endowment of the numéraire at each instant $t, 0\le t\le 1$ and $k_{in}[\theta]$ contributions to the investment cost. In what follows, we normalize $\overline{y}_{in}=0$ for all agents.

Let us rewrite (25) as

$$(26)\quad \int_0^1 \begin{pmatrix} \int_0^\tau (\theta_{in}-c)x_{in}[t]dt \\ +\int_0^{X[\tau]}\int_\tau^1 \delta_i^*[t|\tau,Z,K[\theta],\theta](\theta_{in}-c)x_{in}[t]\, f_{X[\tau]}[Z]dtdZ \end{pmatrix} dp[\tau]$$

$$+\pi[1]\int_0^1(\theta_{in}-c)x_{in}[t]dt-\frac{K[\theta]}{I}+\left(\frac{K[\theta]}{I}-k_{in}[\theta]\right)$$

$$=v_{in}[K[\theta],\theta]+\kappa_{in}[\theta],$$

where we have let

$$(27)\quad v_{in}[K,\theta]=\int_0^1 \begin{pmatrix} \int_0^\tau (\theta_{in}-c)x_{in}[t]dt \\ +\int_0^{X[\tau]}\int_\tau^1 \delta_i^*[t|\tau,Z,K,\theta](\theta_{in}-c)x_{in}[t]\, f_{X[\tau]}[Z]dtdZ \end{pmatrix} dp[\tau]$$

$$+\pi[1]\int_0^1(\theta_{in}-c)x_{in}[t]dt-\frac{K}{I}$$

and

$$(28)\quad \kappa_{in}=\frac{K}{I}-k_{in}.$$

Using the balance budget constraint (23), we can assert that

(29) $\sum_{i=1,\ldots,I,n=1,\ldots,N_i} \kappa_{in}[\theta] = 0$.

As defined, $v_{in}[K,\theta]$ represents the expected payoff for agent in, given that (i) K is the generation capacity installed, (ii) θ is the agents' profile of types, and (iii) all agents share equally the cost of building K generating units. The expected payoff for agent in as represented by $v_{in}[K[\theta],\theta] + \kappa_{in}[\theta]$ can then be interpreted as follows. First, each agent is asked to contribute equally, i.e., contribute $\frac{K[\theta]}{I}$, to the collective generation capacity $K[\theta]$. The collective action provides sufficient funds to construct $K[\theta]$ generating units. In the second stage, agent in receives a transfer equal to $\kappa_{in}[\theta]$, which can be negative or positive. The balance budget constraint (29) ensures that no external funding is needed. In particular, if $\kappa_{in}[\theta] < 0$, agent in pays twice: first the egalitarian contribution $\frac{K[\theta]}{I}$, and then the contribution $-\kappa_{in}[\theta]$ on top of $\frac{K[\theta]}{I}$. If $\kappa_{in}[\theta] > 0$, agent in pays first the egalitarian contribution $\frac{K[\theta]}{I}$, and then receives a net transfer $\kappa_{in}[\theta]$ to offset part of the egalitarian contribution $\frac{K[\theta]}{I}$.

With $v_{in}[K[\theta],\theta]$ and $\kappa_{in}[\theta]$, thus defined, a social choice function now assumes the following form:

(30) $\mathfrak{F}: \theta \rightarrow \mathfrak{F}[\theta] = \left(K[\theta], (\kappa_{in}[\theta])_{i=1,\ldots,I,n=1,\ldots,N_i}\right)$,

where the transfers are required to satisfy the balance budget constraint (29).

A social choice function

$$\mathfrak{F}: \theta \rightarrow \mathfrak{F}[\theta] = \left(K[\theta], (\kappa_{in}[\theta])_{i=1,\ldots,I,n=1,\ldots,N_i}\right)$$

is said to be *ex post efficient* if for each profile of types θ, the following condition is satisfied:

(31) $\sum_{i=1,\ldots,I,n=1,\ldots,N_i} v_{in}[K[\theta],\theta] = max_{K\geq 0} \sum_{i=1,\ldots,I,n=1,\ldots,N_i} v_{in}[K,\theta]$.

PROPOSITION 3: *Suppose that the preferences for electricity consumption of the agents are much greater than the unit cost of generation capacity investment. Let*

$$\mathfrak{F}: \theta \to \mathfrak{F}[\theta] = \left(K[\theta], (\kappa_{in}[\theta])_{i=1,\dots,I,n=1,\dots,N_i}\right)$$

be an ex post efficient social choice function. Then $K[\theta]$*, the generation capacity investment dictated by* $\mathfrak{F}$*, satisfies the following conditions:*

(i) $max_{0\le t\le 1}X[t] < K[\theta] < 2max_{0\le t\le 1}X[t]$,

(ii) $\frac{\partial \Omega[K[\theta],\theta]}{\partial K} = 1$

PROOF: First, note that

$$\sum_{i=1,\dots,I,n=1,\dots,N_i} v_{in}[K,\theta] = \Omega[K,\theta] - K.$$

Next, note that if the VOLL's $\theta_{in}, i = 1, \dots, I, n = 1, \dots, N_i$, are much greater than the unit cost (which we normalize to be 1) – as such is the case of an essential commodity like electricity – it is socially optimal to invest in generation capacity to meet the peak global load, i.e., $K[\theta] > max_{0\le t\le 1}X[t]$. Furthermore, because a generating failure is assumed to occur at most once and because when a breakdown occur, say at time t, the number of generating units that fail cannot exceed the number of generating units that are online to serve the load $X[t]$, we must have $Z \le X[t]$, and this means that the number of generating units that fail cannot exceed $max_{0\le t\le 1}X[t]$. Thus, when $K \ge 2max_{0\le t\le 1}X[t]$, the generation capacity investment is large enough to ensure that load shedding will not be required after a major breakdown, and any investment beyond $2max_{0\le t\le 1}X[t]$ raises cost without raising expected gross social welfare. The optimal generation capacity investment then should satisfy the condition $max_{0\le t\le 1}X[t] < K[\theta] \le 2max_{0\le t\le 1}X[t]$, and is characterized by the first-order condition $\frac{\partial}{\partial K}\Omega[K[\theta],\theta] - 1 = 0$, which is (ii) of Proposition 3. Furthermore, as already argued, $\frac{\partial}{\partial K}\Omega[K,\theta] = 0$ for all $K \ge$

$2max_{0\leq t\leq 1}X[t]$, which when used with the first-order condition $\frac{\partial}{\partial K}\Omega[K[\theta],\theta]-1=0$, allows us to obtain (1) of Proposition 3.■

Condition (ii) of Proposition asserts that given the type profile θ of the agents, the generation capacity investment should be carried out until marginal social benefits are completely offset by marginal cost of investment. This rule is operational because we are able to measure precisely the social benefits with the help of the map $\Omega[.,\theta]: K\rightarrow \Omega[K,\theta]$. The ratio $\frac{K[\theta]-max_{0\leq t\leq 1}X[t]}{max_{0\leq t\leq 1}X[t]}$ is known as the operating reserve margin. Although the map $\Omega[.,\theta]: K\rightarrow \Omega[K,\theta]$ is not concave, we expect the operating reserve to be high if the preferences for electricity consumption of the agents are high.

5. THE DIRECT REVELATION MECHANISM

Let

(32) $\mathfrak{F}: \theta \rightarrow \mathfrak{F}[\theta] = \left(K[\theta], (\kappa_{in}[\theta])_{i=1,\ldots,I,n=1,\ldots,N_i}\right)$

be a social choice function that the regulator wishes to implement. Theoretically, there is a multitude of mechanisms that one can imagine of adopting to implement a social choice function, and thus there is no limit to our imagination in designing a mechanism that implements $\mathfrak{F}$. However, by invoking the *revelation principle*,[24] we can restrict the search for a mechanism that implements $\mathfrak{F}$ to the direct revelation mechanism. Under the direct revelation mechanism that implements the social choice function $\mathfrak{F}$, the regulator asks each agent to reveal her type, namely her preferences for the consumption of electricity, and then uses the information thus obtained to implement the social choice dictated by $\mathfrak{F}$.

[24] See Chapter 23 in MasColell et al. (1995).

Suppose that θ_{in} is the type of agent in and that she communicates $\hat{\theta}_{in}$ to the regulator as her preferences for electricity consumption. Let $\hat{\theta} = \left(\hat{\theta}_{in}\right)_{i=1,\dots,I,n=1,\dots,N_i}$ denote the type profile communicated by the agents to the regulator and K be a level of generation capacity that can meet peak demand. Define

$$\text{(33)} \quad \varphi_{jm}\left[K,\hat{\theta}\middle|\theta_{jm}\right] = \int_0^1 \begin{pmatrix} \int_0^{\tau}\left(\theta_{jm}-c\right)x_{jm}[t]dt \\ +\int_0^{X[\tau]}\int_{\tau}^1 \delta_j^*\left[t\middle|\tau,Z,K,\hat{\theta}\right]\left(\theta_{jm}-c\right)x_{jm}[t]f_{X[\tau]}[Z]dtdZ \end{pmatrix} dp[\tau]$$

$$+\pi[1]\int_0^1\left(\theta_{jm}-c\right)x_{jm}[t]dt - \frac{K}{I}, \qquad \left(j=1,\dots,I, m=1,\dots,N_j\right).$$

As defined, $\varphi_{jm}\left[K,\hat{\theta}\middle|\theta_{jm}\right]$ represents the expected payoff for agent jm, given that θ_{jm} is her type; K is the generation capacity investment; each agent is required to contribute equally, namely contribute $\frac{K}{I}$, to the generation capacity investment; and $\hat{\theta}$ is the profile of types announced by the agents. Note that the announced type profile $\hat{\theta}$ is used by the regulator to compute the second-best blackout program

$$\delta_i^*\left[t\middle|\tau,Z,K,\hat{\theta}\right], \tau \le t \le 1, \qquad (i=1,\dots,I),$$

already analyzed in Section 3.

Thus, when the social choice function $\mathfrak{F}$ is implemented, we obtain the following expression for the expected payoff of agent jm, given that θ_{jm} is her type and $\hat{\theta}$ is the profile of types that the agents communicate to the regulator:

$$\text{(34)} \quad \varphi_{jm}\left[K\left[\hat{\theta}\right],\hat{\theta}\middle|\theta_{jm}\right] + \kappa_{jm}\left[\hat{\theta}\right].$$

The direct revelation mechanism used to implement the social choice function $\mathfrak{F}$ together with the set of possible type profiles $\prod_{i=1}^{I}\prod_{n=1}^{N_i}\left[\underline{\theta}_{in},\overline{\theta}_{in}\right]$, the probability distribution Φ on $\prod_{i=1}^{I}\prod_{n=1}^{N_i}\left[\underline{\theta}_{in},\overline{\theta}_{in}\right]$, and the expected payoff functions (34) define a game with incomplete information that we call Γ. In the game Γ, a strategy for agent in is a map $s_{in}: \theta_{in} \rightarrow s_{in}[\theta_{in}]$, where $s_{in}[\theta_{in}]$ is the type that agent in communicates to the regulator when her type is θ_{in}. If $s_{in}[\theta_{in}] = \theta_{in}$ for all $\theta_{in} \in [\underline{\theta}_{in}, \overline{\theta}_{in}]$, then agent in tells the truth. Otherwise, the strategy s_{in} is not truthful. A profile of strategies for the agents is a list $s = (s_{in})_{i=1,\dots,I,n=1,\dots,N_i}$. If s is a profile of strategies adopted by the agents and $\theta = (\theta_{in})_{i=1,\dots,I,n=1,\dots,N_i}$ is a profile of types, then $s[\theta] = (s_{in}[\theta_{in}])_{i=1,\dots,I,n=1,\dots,N_i}$ is the profile of types revealed to the regulator. Unless all the agents choose to reveal their types truthfully, we will have $s[\theta] \neq \sigma$.

Because the preferences for electricity consumption of all the agents influence the total generation capacity investment and the transfers to each agent, there is no reason to expect that an agent will reveal her type truthfully unless the social choice function $\mathfrak{F}$ is well structured. If an agent, say in, exaggerates her preferences for electricity consumption, this action raises the welfare component associated with A_i, the area in which she resides, in the calculations of the second-best blackout program. The exaggeration leads probably first to a rise in the generation capacity investment, and the ensuing rise in the total cost to be borne by all the agents, agent in included. Furthermore, the exaggeration of her preferences for electricity consumption might induce the regulator to impose on agent in a higher share of the total generation capacity investment. On the other hand, the exaggeration benefits all the residents – agent in included – in A_i at the expense of the agents in all the other areas because it induces a rise in the priority of area A_i relative to that of any other area in the calculations of the ISO when she decides on which area to serve and which area to shed after a major failure. On the other hand, if the agent chooses to understate her preferences for electricity consumption, then she can expect to contribute less to the total generation capacity investment. However, this action also induces a higher chance of load shedding in her own area. Thus, only when the social choice function is well structured will an agent reveals her type truthfully.

DEFINITION: A strategy s_{in} is said to be a best response of agent in to $\left(s_{jm}\right)_{j=1,\ldots,I,m=1,\ldots,N_j,jm\neq in}$, a profile of strategies of all the agents other than in, in the game Γ if for all possible types $\theta_{in} \in \left[\underline{\theta}_{in}, \bar{\theta}_{in}\right]$, the following condition holds:

(35)

$$\int \varphi_{in}\left[\begin{matrix}K\left[s_{in}[\theta_{in}], \left(\sigma_{jm}\left[\tilde{\theta}_{jm}\right]\right)_{j=1,\ldots,I,m=1,\ldots,N_j,jm\neq in}\right], \\ s_{in}[\theta_{in}], \left(\sigma_{jm}\left[\tilde{\theta}_{jm}\right]\right)_{j=1,\ldots,I,m=1,\ldots,N_j,jm\neq in}\end{matrix}\middle|\theta_{in}\right]\prod_{j=1,\ldots,I,m=1,\ldots,N_j,jm\neq in} d\Phi_{jm}[\tilde{\theta}_{jm}]$$

$$\geq \int \varphi_{jm}\left[\begin{matrix}K\left[\hat{\theta}_{in}, \left(\sigma_{jm}\left[\tilde{\theta}_{jm}\right]\right)_{j=1,\ldots,I,m=1,\ldots,N_j,jm\neq in}\right], \\ \hat{\theta}_{in}, \left(\sigma_{jm}\left[\tilde{\theta}_{jm}\right]\right)_{j=1,\ldots,I,m=1,\ldots,N_j,jm\neq in}\end{matrix}\middle|\theta_{in}\right]\prod_{j=1,\ldots,I,m=1,\ldots,N_j,jm\neq in} d\Phi_{jm}[\tilde{\theta}_{jm}]$$

for all $\hat{\theta}_{in} \in \left[\underline{\theta}_{in}, \bar{\theta}_{in}\right]$.

Note that in computing (35) we have used the assumption that agents' types are statistically independent.

DEFINITION: A strategy profile $(s_{in}^*)_{i=1,\ldots,I,n=1,\ldots,N_i}$ constitutes a Bayesian Nash equilibrium for the game Γ if for any agent in, the strategy s_{in}^* is a best response to $\left(s_{jm}^*\right)_{j=1,\ldots,I,m=1,\ldots,N_j,jm\neq in}$.

DEFINITION: The social choice function $\mathfrak{F}$ is said to be implementable in Bayesian Nash equilibrium if the game Γ has a Bayesian Nash equilibrium, say $(s_{in}^*)_{i=1,\ldots,I,n=1,\ldots,N_i}$. Furthermore, if for each $i = 1, \ldots, I$, we have $s_{in}^*[\theta_{in}] = \theta_{in}$ for all $\theta_{in} \in \left[\underline{\theta}_{in}, \overline{\theta}_{in}\right]$, then $\mathfrak{F}$ is said to be truthfully implementable in Bayesian Nash equilibrium.

6. TRUTHFUL IMPLEMENTABILITY IN BAYESIAN NASH EQUILIBRIUM OF AN EX POST EFFICIENT SOCIAL CHOICE FUNCTION

For any profile of types $\theta = (\theta_{in})_{i=1,\dots,I,n=1,\dots,N_i}$, let

(36) $K^*[\theta] = argmax_K \sum_{i=1}^{I} \sum_{n=1}^{N_i} \varphi_{in}[K, \theta | \theta_{in}]$

$$= argmax_K (\Omega[K, \theta] - K)$$

and $\mathfrak{F}^*: \theta \to \mathfrak{F}^*[\theta]$ be the social choice function defined by

(37) $\mathfrak{F}^*[\theta] = \left(K^*[\theta], (\kappa_{in}^*[\theta])_{i=1,\dots,I,n=1,\dots,N_i}\right)$,

where

(38) $\kappa_{in}^*[\theta] =$

$$\int \left(\left(\sum_{j=1,\dots,I,m=1,\dots,N_j,jm\neq in} \varphi_{jm} \left[\begin{matrix} K^*\left[\theta_{in}, \left(\tilde{\theta}_{jm}\right)_{j=1,\dots,I,m=1,\dots,N_j,jm\neq in}\right], \\ \left(\theta_{in}, \left(\tilde{\theta}_{jm}\right)_{j=1,\dots,I,m=1,\dots,N_j,jm\neq in}\right) \end{matrix} \middle| \tilde{\theta}_{jm} \right] \right) \times \prod_{j=1,\dots,I,m=1,\dots,N_j,jm\neq in} d\Phi_{jm}\left[\tilde{\theta}_{jm}\right] \right)$$

$$+h_{in}^*\left[\left(\theta_{jm}\right)_{j=1,\dots,I,m=1,\dots,N_j,jm\neq in}\right].$$

Note that in (38) the integral represents the sum of the expected payoffs of all the agents other than agent in, conditioned on (i) all the agents other than agent in tell the truth and (ii) agent in communicate θ_{in} as her type to the ISO. Note also that in computing this integral we have used the assumption that agents' types are statistically independent.

As for $h_{in}^*\left[\left(\theta_{jm}\right)_{j=1,\dots,I,m=1,\dots,N_j,jm\neq in}\right]$, it is an arbitrary function of $\left(\theta_{jm}\right)_{j=1,\dots,I,m=1,\dots,N_j,jm\neq in}$, which is yet to be determined so that the balance budget constraint (29) is satisfied. Thus, when agent in varies the type she communicates to the ISO, i.e., when she varies θ_i, the change in her transfer above the egalitarian contribution to the generation capacity investment is equal exactly to the expected externality of the change in the expected benefits of all the agents other than agent in.

LEMMA 1: *In the game Γ, revealing her type truthfully is a best response for an agent when all the other agents reveal their types truthfully.*

PROOF: See Appendix A.

To simplify notations, for any $\theta_{in} \in [\underline{\theta}_{in}, \overline{\theta}_{in}]$, let

$$(39) \quad \xi_{in}[\theta_{in}] = \int \left(\begin{array}{c} \left(\sum_{j=1,\dots,I,m=1,\dots,N_j,jm\neq in} \varphi_{jm} \left[\begin{array}{c} K^*\left[\theta_{in}, \left(\tilde{\theta}_{jm}\right)_{j=1,\dots,I,m=1,\dots,N_j,jm\neq in}\right], \\ \left(\theta_{in}, \left(\tilde{\theta}_{jm}\right)_{j=1,\dots,I,m=1,\dots,N_j,jm\neq in}\right) \end{array} \middle| \tilde{\theta}_{jm} \right] \right) \\ \times \prod_{j=1,\dots,I,m=1,\dots,N_j,jm\neq in} d\Phi_{jm}\left[\tilde{\theta}_{jm}\right] \end{array} \right),$$

$(i = 1, \dots, I, n = 1, \dots, N_i)$,

and

$$(40) \quad h_{in}^*\left[\left(\theta_{jm}\right)_{j=1,\dots,I,m=1,\dots,N_j,jm\neq in}\right] = -\frac{1}{\left(\sum_{j=1}^{I} N_j\right)-1} \sum_{j=1,\dots,I,m=1,\dots,N_j,jm\neq in} \xi_{jm}\left[\theta_{jm}\right].$$

LEMMA 2: *The social choice function $\mathfrak{F}^*$ satisfies the balance budget constraint, i.e., for any type profile $\theta = (\theta_{in})_{i=1,\dots,I,n=1,\dots,N_i}$, we have $\sum_{i=1,\dots,I,n=1,\dots,N_i} \kappa_{in}^*[\theta] = 0$.*

PROOF: See Appendix B.

The following proposition follows immediately from Lemmas 1 and 2.

PROPOSITION 3: *The ex post efficient social choice function* $\mathfrak{F}^*$*, as defined by* (37)-(40), *is truthfully implementable in Bayesian Nash equilibrium. That is, the strategy profile under which each consumer reveals her preferences for electricity consumption truthfully constitutes a Bayesian Nash equilibrium for the game* Γ*, and the transfers under* $\mathfrak{F}^*$ *satisfy the balance budget constraint.*

7. THE PARTICIPATION CONSTRAINT

Proposition 3 asserts that the expected externality mechanism truthfully implements an ex post efficient social choice function in Bayesian Nash equilibrium. It says nothing about whether all the agents are willing to participate in this mechanism. It is well known in the context of bilateral trading with asymmetric information that there is no Bayesian incentive compatible social choice function that is ex post efficient and gives every buyer type and every seller type non-negative expected gains from participation.[25] Under the expected externality mechanism of Proposition 3, the expected payoff for agent, say agent in, is given by $v_{in}\left[K^*[\theta_i,\theta_j],\theta_i,\theta_j\right]+\kappa_{in}^*[\theta]$. Now if agent in refuses to participate in the mechanism, then her expected payoff is 0. Thus, agent in is only willing to participate in the mechanism if the following ex post participation constraint is satisfied:

(41) $v_{in}[K^*[\theta],\theta]+\kappa_{in}^*[\theta]\geq 0,$ $(i=1,\ldots,I, n=1,\ldots,N_i)$.

If we let

(42) $v_{in}^0[K[\theta],\theta]=$

$$\int_0^1\left(\begin{array}{c}\int_0^{\tau}(\theta_{in}-c)x_i[t]dt\\+\int_0^{X[\tau]}\int_{\tau}^1\delta_i^*[t|\tau,Z,K[\theta],\theta](\theta_{in}-c)x_{in}[t]\,f_{X[\tau]}[Z]dtdZ\end{array}\right)dp[\tau]$$

[25] See Proposition 23.E.1 (The Myerson-Satterthwaite Theorem) in MasCollel et al (1995), p. 895-896.

$$+\pi[1]\int_0^1 (\theta_{in} - c)x_{in}[t]dt$$

denote the expected payoff obtained by agent in, given that (i) she neither pays the egalitarian contribution in financing the generation capacity nor receives any transfer on top of the egalitarian contribution, then the participation constraint (41) becomes

(43) $v_{in}^0[K[\theta],\theta] - \frac{K}{I} + \kappa_{in}^*[\theta] \geq 0,$ $\qquad (i = 1, \dots, I, n = 1, \dots, N_i)$.

The following proposition gives sufficient conditions for the participation constraints (43) to hold.

PROPOSITION 4: *If $\overline{\theta}_{in} - \underline{\theta}_{in}$ is small, $i = 1, \dots, I, n = 1, \dots, N_i$, i.e., if the range over which the preferences for electricity consumption of each agent varies is narrow, and if there exists a type profile $\theta = (\theta_{in})_{i=1,\dots,I,n=1,\dots,N_i}$, such that the following conditions are satisfied:*

(44) $\frac{\sum_{j=1,\dots,I,m=1,\dots,N_j,jm\neq in} v_{jm}^0[K[\theta],\theta]}{(\sum_{j=1,\dots,I} N_j)-1} - \frac{1}{\sum_{j=1,\dots,I} N_j} K[\theta] > 0, \qquad (i = 1, \dots, I, n = 1, \dots, N_i),$

then the participation constraint holds with strict inequality for each agent under the expected externality mechanism that implements the social choice function $\mathfrak{F}^*$.

PROOF: See Appendix C.

The conditions represented by (44) are quite intuitive. They assert that for each consumer, say consumer in, the average expected gross payoff for the subset consisting of all the consumers other than consumer in exceeds the egalitarian contribution needed to finance the generation capacity investment. As an example, consider the case of two areas, say A_i and $A_{j,}$ with one consumer in each area, say consumer i and consumer j, the participation condition for consumer i is given by

$$(44i)\;\; v_j^0\left[K\left[\theta_i,\theta_j\right],\theta_i,\theta_j\right] - \frac{K[\theta_i,\theta_j]}{2} > 0,$$

while the participation constraint for consumer j is given by

$$(44j)\;\; v_i^0\left[K\left[\theta_i,\theta_j\right],\theta_i,\theta_j\right] - \frac{K[\theta_i,\theta_j]}{2} > 0.$$

If the two consumers are symmetric and if $\theta_i = \theta_j$, then both $(44i)$ and $(44j)$ are satisfied because

$$(44i)\;\; v_i^0\left[K\left[\theta_i,\theta_j\right],\theta_i,\theta_j\right] + v_j^0\left[K\left[\theta_i,\theta_j\right],\theta_i,\theta_j\right] - K\left[\theta_i,\theta_j\right] > 0.$$

8. CONCLUDING REMARKS

At the present time, no technology exists that can economically store large quantities of electricity. Thus, electricity must be produced and consumed at the same time. However, the balance between the generation and the load is constantly perturbed by load fluctuations, by imprecise control of the outputs of generating plants, and by the sudden failures of some generating plants or of an interconnection. To maintain the reliability of the system, there must be some stand-by generating capacity – called operating reserves – that is ready to maintain the balance between output and load at each instant.

Because electricity is an essential commodity, its demand is extremely in-elastic, and thus will not be met when a major failure of generating plants occurs. Under such a scenario, the ISO has no choice but to interrupt the electricity supply to some geographical areas of the network. A high installed capacity will be able to serve a heavy load while leaving a considerable margin to serve as operating reserves and increase the reliability of the system. However, building new capacity is costly, and therefore it is optimal to accept some blackouts. Furthermore, because large operating reserves make blackouts less likely and thus benefit all consumers, they have the nature of a public good. For each given level of generation capacity installed, to find the optimal blackout strategy, the ISO has to know the VOLL of every consumer so that welfare can be computed and maximized. Once the optimal blackout program for any given level of installed generation capacity has been found, the ISO can work backward and find the optimal level of generation capacity investment. The ISO can try to obtain the information needed for the welfare calculations through (i) survey interviews, (ii) the production approach, (iii) market behavior (expenditures on backups and interruptible contracts), and (iv) case studies. Each of these methods has its own advantages and disadvantages. In survey interviews, consumers are asked how much damage they suffer as a result of supply interruptions. Because of the public good nature of operating reserves there is no incentive for the consumers to reveal their VOLL's truthfully. The production function approach estimates the consequences of supply interruptions for firms' loss production and households' lost time by quantitative techniques. A main disadvantage of the production approach involves the difficulty of accounting for the restart time of businesses and the stress in households. As for the approach of market behavior, the frequencies of backup expenditures and interruptible contracts are so low that the revealed preference for these types of economic behavior to estimate VOLL's cannot be relied upon. Case studies do yield actual damages caused by supply interruptions. However, the results obtained are specific to the cases being studied, and thus cannot be easily generalized.

In this paper, we have formalized a model of reliability for an electric grid using the mechanism design approach. With the transfers structured according to the expected externality mechanism, it is possible to elicit the preferences for electricity consumption of all the consumers. More specifically, we demonstrate that it is possible to truthfully implement an ex post efficient social choice function in Bayesian Nash equilibrium. The model provides both the optimal level of generation capacity investment and the optimal blackouts. Under the traditional regulatory system, the electric utilities were responsible for providing adequate capacity to meet demand and ensure a given level of reliability for the system. The model we formalize in this paper applies to search an environment. After liberalization, the provision of capacity to meet demand and ensure the desired level of reliability is left to market forces. In a large market, such as the PJM (Pennsylvania-New Jersey-Maryland) market, the LSE's are required to contract with electric producers for enough capacity to meet the monthly peak demand plus a reserve margin. The task of eliciting the VOLL's from consumers now falls on the LSE's, and these retailers can use the information thus obtained through the expected externality mechanism to find out the optimal level of capacity to contract for and the consumers whose supply must be interrupted in case of blackouts. This is the next problem in our research agenda.

APPENDIX A

The Proof of Lemma 1

Consider an agent, say agent in. Suppose that all the agents other than agent in reveal their types truthfully. If agent in reveals her type truthfully, then her expected utility under the social choice function Γ^* is given by

(A.1)

$$\int \left(\begin{array}{c} \varphi_{in} \left[\begin{array}{c} K^* \left[\theta_{in}, (\theta_{jm})_{j=1,\dots,I,m=1,\dots,N_j,jm\neq in} \right], \\ \theta_{in}, (\theta_{jm})_{j=1,\dots,I,m=1,\dots,N_j,jm\neq in} \end{array} \middle| \theta_{in} \right] \\ +\kappa^*_{in} \left[\theta_{in}, (\theta_{jm})_{j=1,\dots,I,m=1,\dots,N_j,jm\neq in} \right] \end{array} \right) \prod_{j=1,\dots,I,m=1,\dots,N_j,jm\neq in} d\Phi_{jm}[\theta_{jm}]$$

$$= \int \varphi_{in} \left[\begin{array}{c} K^* \left[\theta_{in}, (\theta_{jm})_{j=1,\dots,I,m=1,\dots,N_j,jm\neq in} \right], \\ \theta_{in}, (\theta_{jm})_{j=1,\dots,I,m=1,\dots,N_j,jm\neq in} \end{array} \middle| \theta_{in} \right] \prod_{j=1,\dots,I,m=1,\dots,N_j,jm\neq in} d\Phi_{jm}[\theta_{jm}]$$

$$+ \int \kappa^*_{in} \left[\theta_{in}, (\theta_{jm})_{j=1,\dots,I,m=1,\dots,N_j,jm\neq in} \right] \prod_{j=1,\dots,I,m=1,\dots,N_j,jm\neq in} d\Phi_{jm}[\theta_{jm}].$$

Now using (38), we can write

(A.2) $$\int \kappa^*_{in} \left[\theta_{in}, (\theta_{jm})_{j=1,\dots,I,m=1,\dots,N_j,jm\neq in} \right] \prod_{j=1,\dots,I,m=1,\dots,N_j,jm\neq in} d\Phi_{jm}[\theta_{jm}] =$$

$$\iint \left(\sum_{j=1,\dots,I,m=1,\dots,N_j,jm\neq in} \varphi_{jm} \left[\begin{array}{c} K^* \left[\theta_{in}, (\tilde{\theta}_{j'm'})_{j'=1,\dots,I,m'=1,\dots,N_{j'},j'm'\neq in} \right], \\ \left(\theta_{in}, (\tilde{\theta}_{j'm'})_{j'=1,\dots,I,m'=1,\dots,N_{j'},j'm'\neq in} \right) \end{array} \middle| \tilde{\theta}_{jm} \right] \right)$$
$$\times \prod_{j=1,\dots,I,m=1,\dots,N_j,jm\neq in} d\Phi_{jm}[\tilde{\theta}_{jm}] \times \prod_{j=1,\dots,I,m=1,\dots,N_j,jm\neq in} d\Phi_{jm}[\theta_{jm}]$$

$$+ \int h_{in} \left[\theta_{in}, (\theta_{jm})_{j=1,\dots,I,m=1,\dots,N_j,jm\neq in} \right] \prod_{j=1,\dots,I,m=1,\dots,N_j,jm\neq in} d\Phi_{jm}[\theta_{jm}]$$

$$= \int \left(\sum_{j=1,\dots,I,m=1,\dots,N_j,jm\neq in} \varphi_{jm} \left[\begin{matrix} K^* \left[\theta_{in}, \left(\tilde{\theta}_{j\prime m\prime}\right)_{j\prime=1,\dots,I,m\prime=1,\dots,N_{j\prime},j\prime m\prime\neq in} \right], \\ \left(\theta_{in}, \left(\tilde{\theta}_{j\prime m\prime}\right)_{j\prime=1,\dots,I,m\prime=1,\dots,N_{j\prime},j\prime m\prime\neq in} \right) \end{matrix} \middle| \tilde{\theta}_{jm} \right] \right) \times$$

$$\times \prod_{j=1,\dots,I,m=1,\dots,N_j,jm\neq in} d\Phi_{jm}\left[\tilde{\theta}_{jm}\right]$$

$$+ \int h_{in} \left[\theta_{in}, \left(\theta_{jm}\right)_{j=1,\dots,I,m=1,\dots,N_j,jm\neq in} \right] \prod_{j=1,\dots,I,m=1,\dots,N_j,jm\neq in} d\Phi_{jm}\left[\theta_{jm}\right]$$

$$= \int \left(\sum_{j=1,\dots,I,m=1,\dots,N_j,jm\neq in} \varphi_{jm} \left[\begin{matrix} K^* \left[\theta_{in}, \left(\theta_{j\prime m\prime}\right)_{j\prime=1,\dots,I,m\prime=1,\dots,N_{j\prime},j\prime m\prime\neq in} \right], \\ \left(\theta_{in}, \left(\theta_{j\prime m\prime}\right)_{j\prime=1,\dots,I,m\prime=1,\dots,N_{j\prime},j\prime m\prime\neq in} \right) \end{matrix} \middle| \theta_{jm} \right] \right) \times$$

$$\times \prod_{j=1,\dots,I,m=1,\dots,N_j,jm\neq in} d\Phi_{jm}\left[\theta_{jm}\right]$$

$$+ \int h_{in} \left[\theta_{in}, \left(\theta_{jm}\right)_{j=1,\dots,I,m=1,\dots,N_j,jm\neq in} \right] \prod_{j=1,\dots,I,m=1,\dots,N_j,jm\neq in} d\Phi_{jm}\left[\theta_{jm}\right]$$

Using (A.2) in (A.1), we obtain

(A.3)

$$\int \left(\begin{matrix} \varphi_{in} \left[\begin{matrix} K^* \left[\theta_{in}, \left(\theta_{jm}\right)_{j=1,\dots,I,m=1,\dots,N_j,jm\neq in} \right], \\ \theta_{in}, \left(\theta_{jm}\right)_{j=1,\dots,I,m=1,\dots,N_j,jm\neq in} \end{matrix} \middle| \theta_{in} \right] \\ +\kappa_{in}^* \left[\theta_{in}, \left(\theta_{jm}\right)_{j=1,\dots,I,m=1,\dots,N_j,jm\neq in} \right] \end{matrix} \right) \prod_{j=1,\dots,I,m=1,\dots,N_j,jm\neq in} d\Phi_{jm}\left[\theta_{jm}\right]$$

$$= \int \sum_{j=1,\dots,I,m=1,\dots,N_{j\prime}} \varphi_{jm} \left[\begin{matrix} K^* \left[\theta_{in}, \left(\theta_{j\prime m\prime}\right)_{j\prime=1,\dots,I,m\prime=1,\dots,N_{j\prime},j\prime m\prime\neq in} \right], \\ \theta_{in}, \left(\theta_{j\prime m\prime}\right)_{j\prime=1,\dots,I,m\prime=1,\dots,N_{j\prime},j\prime m\prime\neq in} \end{matrix} \middle| \theta_{jm} \right] \times$$

$$\prod_{j=1,\dots,I,m=1,\dots,N_j,jm\neq in} d\Phi_{jm}\left[\theta_{jm}\right]$$

$$+ \int h_{in} \left[\theta_{in}, \left(\theta_{j\prime m\prime}\right)_{j\prime=1,\dots,I,m\prime=1,\dots,N_{j\prime},j\prime m\prime\neq in} \right] \prod_{j=1,\dots,I,m=1,\dots,N_j,jm\neq in} d\Phi_{jm}\left[\theta_{jm}\right].$$

On the other hand, if agent in announces $\hat{\theta}_{in}$ as her type while all the other agents reveal their types truthfully, then agent in can expect the following level of payoff:

$$\text{(A.4)} \quad \int \left(\sum_{j=1,\ldots,I,m=1,\ldots,N_j} \varphi_{jm} \left[\begin{matrix} K^*\left[\hat{\theta}_{in}, \left(\theta_{j'm'}\right)_{j'=1,\ldots,I,m'=1,\ldots,N_{j'},j'm'\neq in}\right], \\ \hat{\theta}_{in}, \left(\theta_{j'm'}\right)_{j'=1,\ldots,I,m'=1,\ldots,N_{j'},j'm'\neq in} \end{matrix} \middle| \theta_{jm} \right] \right) \times$$

$$\prod_{j=1,\ldots,I,m=1,\ldots,N_j,jm\neq in} d\Phi_{jm}\left[\theta_{jm}\right].$$

$$+ \int h_{in}\left[\theta_{in}, \left(\theta_{j'm'}\right)_{j'=1,\ldots,I,m'=1,\ldots,N_{j'},j'm'\neq in}\right] \prod_{j=1,\ldots,I,m=1,\ldots,N_j,jm\neq in} d\Phi_{jm}\left[\theta_{jm}\right].$$

Subtracting (A.4) from (A.3), we obtain

$$\text{(A.5)} \quad \int \left(\begin{matrix} \sum_{j=1,\ldots,I,m=1,\ldots,N_j} \varphi_{jm} \left[\begin{matrix} K^*\left[\theta_{in}, \left(\theta_{j'm'}\right)_{j'=1,\ldots,I,m'=1,\ldots,N_{j'},j'm'\neq in}\right], \\ \theta_{in}, \left(\theta_{j'm'}\right)_{j'=1,\ldots,I,m'=1,\ldots,N_{j'},j'm'\neq in} \end{matrix} \middle| \theta_{jm} \right] \\ - \sum_{j=1,\ldots,I,m=1,\ldots,N_j} \varphi_{jm} \left[\begin{matrix} K^*\left[\hat{\theta}_{in}, \left(\theta_{j'm'}\right)_{j'=1,\ldots,I,m'=1,\ldots,N_{j'},j'm'\neq in}\right], \\ \hat{\theta}_{in}, \left(\theta_{j'm'}\right)_{j'=1,\ldots,I,m'=1,\ldots,N_{j'},j'm'\neq in} \end{matrix} \middle| \theta_{jm} \right] \end{matrix} \right) \times.$$

$$\prod_{j=1,\ldots,I,m=1,\ldots,N_j,jm\neq in} d\Phi_{jm}\left[\theta_{jm}\right]$$

To show that it is best for an agent to reveal her type truthfully when all the other agents reveal their types truthfully, we now show that (A.5) is non-negative. To this end, using the definition of $\varphi_{jm}\left[K, \hat{\theta} \middle| \theta_{jm}\right]$, given by (33), we can write

$$\text{(A.6)} \quad \varphi_{jm} \left[\begin{matrix} K^*\left[\hat{\theta}_{in}, \left(\theta_{j'm'}\right)_{j'=1,\ldots,I,m'=1,\ldots,N_{j'},j'm'\neq in}\right], \\ \hat{\theta}_{in}, \left(\theta_{j'm'}\right)_{j'=1,\ldots,I,m'=1,\ldots,N_{j'},j'm'\neq in} \end{matrix} \middle| \theta_{jm} \right]$$

$$= \pi[1] \int_0^1 \left(\theta_{jm} - c\right) x_{jm}[t] dt - \frac{K^*\left[\hat{\theta}_{in}, \left(\theta_{j'm'}\right)_{j'=1,\ldots,I,m'=1,\ldots,N_{j'},j'm'\neq in}\right]}{I}$$

$$+\int_0^1 \left(\begin{array}{c} \int_0^{\tau} (\theta_{j\prime m\prime} - c) x_{j\prime m\prime}[t] dt \\ + \int_0^{X[\tau]} \int_{\tau}^1 \delta_j^* \left[t \left| \begin{array}{c} \tau, Z, K^*\left[\hat{\theta}_{in}, (\theta_{j'm'})_{j'=1,\dots,I,m'=1,\dots,N_{j'},j'm'\neq in}\right], \\ \hat{\theta}_{in}, (\theta_{j'm'})_{j'=1,\dots,I,m'=1,\dots,N_{j'},j'm'\neq in} \end{array} \right. \right] (\theta_{jm} - c) x_{jm}[t] f_{X[\tau]}[Z] dt dZ \end{array} \right) dp[\tau].$$

Summing (A.6) over $j = 1, \dots, I, m = 1, \dots, N_j$, we obtain

(A.7) $$\sum_{j=1,\dots,I,m=1,\dots,N_j} \varphi_{jm} \left[\begin{array}{c} K^*\left[\hat{\theta}_{in}, (\theta_{j\prime m\prime})_{j\prime=1,\dots,I,m\prime=1,\dots,N_{j}\prime,j\prime m\prime\neq in}\right], \\ \left(\hat{\theta}_{in}, (\theta_{j\prime m\prime})_{j\prime=1,\dots,I,m\prime=1,\dots,N_{j}\prime,j\prime m\prime\neq in}\right) \end{array} \middle| \theta_{jm} \right]$$

$$= \sum_{j=1,\dots,I,m=1,\dots,N_j} \left(\pi[1] \int_0^1 (\theta_{jm} - c) x_{jm}[t] dt - \frac{K^*\left[\hat{\theta}_{in}, (\theta_{j\prime m\prime})_{j\prime=1,\dots,I,m\prime=1,\dots,N_{j\prime},j\prime m\prime\neq in}\right]}{I} \right)$$

$$+ \sum_{j=1,\dots,I,m=1,\dots,N_j} \int_0^1 \int_0^{\tau} (\theta_{jm} - c) x_{jm}[t] dt\, dp[\tau]$$

$$+ \sum_{j=1,\dots,I,m=1,\dots,N_j} \int_0^1 \int_0^{X[\tau]} \int_{\tau}^1 \delta_j^* \left[t \left| \begin{array}{c} \tau, Z, K^*\left[\hat{\theta}_{in}, (\theta_{j\prime m\prime})_{j\prime=1,\dots,I,m\prime=1,\dots,N_{j\prime},j\prime m\prime\neq in}\right], \\ \hat{\theta}_{in}, (\theta_{j\prime m\prime})_{j\prime=1,\dots,I,m\prime=1,\dots,N_{j\prime},j\prime m\prime\neq in} \end{array} \right. \right] (\theta_{jm} - c) x_{jm}[t] f_{X[\tau]}[Z] dt dZ\, dp[\tau].$$

Now because

(A.8) $\delta_j^* \left[t \middle| \tau, Z, K^*[\hat{\theta}_{in},], \left(\hat{\theta}_{in}, (\theta_{j\prime m\prime})_{j\prime=1,\dots,I,m\prime=1,\dots,N_{j}\prime,j\prime m\prime\neq in}\right) \right],$ $(\tau \leq t \leq 1, j = 1, \dots, I)$,

is the second-best blackout program executed by the ISO when $K^*\left[\hat{\theta}_{in}, (\theta_{j\prime m\prime})_{j\prime=1,\dots,I,m\prime=1,\dots,N_{j}\prime,j\prime m\prime\neq in}\right]$ is the generation capacity investment and when the ISO receives $\left(\hat{\theta}_{in}, (\theta_{j\prime m\prime})_{j\prime=1,\dots,I,m\prime=1,\dots,N_{j}\prime,j\prime m\prime\neq in}\right)$ as the agents' type profile,[26] it is not necessarily the second-best blackout program that the ISO would carry out when $K^*\left[\hat{\theta}_{in},(\theta_{j\prime m\prime})_{j\prime=1,\dots,I,m\prime=1,\dots,N_{j\prime},j\prime m\prime\neq in}\right]$ is the generation capacity investment if the ISO knew that $\left(\theta_{in},(\theta_{j\prime m\prime})_{j\prime=1,\dots,I,m\prime=1,\dots,N_{j\prime},j\prime m\prime\neq in}\right)$ is the true type profile, and this means

(A.9) $$\sum_{j=1,\dots,I,m=1,\dots,N_j}\left(\int_\tau^1 \delta_j^*\left[t\left|\begin{matrix}\tau,Z,K^*\left[\hat{\theta}_{in},(\theta_{j\prime m\prime})_{j\prime=1,\dots,I,m\prime=1,\dots,N_{j\prime},j\prime m\prime\neq in}\right],\\ \hat{\theta}_{in},(\theta_{j\prime m\prime})_{j\prime=1,\dots,I,m\prime=1,\dots,N_{j\prime},j\prime m\prime\neq in}\end{matrix}\right.\right](\theta_{jm}-c)x_{jm}[t]dt\right)$$

$$\leq\sum_{j=1,\dots,I,m=1,\dots,N_j}\left(\int_\tau^1 \delta_j^*\left[t\left|\begin{matrix}\tau,Z,K^*\left[\hat{\theta}_{in},(\theta_{j\prime m\prime})_{j\prime=1,\dots,I,m\prime=1,\dots,N_{j\prime},j\prime m\prime\neq in}\right],\\ \theta_{in},(\theta_{j\prime m\prime})_{j\prime=1,\dots,I,m\prime=1,\dots,N_{j\prime},j\prime m\prime\neq in}\end{matrix}\right.\right](\theta_{jm}-c)x_{jm}[t]dt\right).$$

Furthermore, because $K^*\left[\hat{\theta}_{in},(\theta_{j\prime m\prime})_{j\prime=1,\dots,I,m\prime=1,\dots,N_{j\prime},j\prime m\prime\neq in}\right]$ is not necessarily the ex post efficient generation capacity investment when $\left(\theta_{in},(\theta_{j\prime m\prime})_{j\prime=1,\dots,I,m\prime=1,\dots,N_{j\prime},j\prime m\prime\neq in}\right)$ is the true type profile, we must also have

(A.10) $$\sum_{j=1,\dots,I,m=1,\dots,N_j}\left(\int_\tau^1 \delta_j^*\left[t\left|\begin{matrix}\tau,Z,K^*\left[\hat{\theta}_{in},(\theta_{j\prime m\prime})_{j\prime=1,\dots,I,m\prime=1,\dots,N_{j\prime},j\prime m\prime\neq in}\right],\\ \theta_{in},(\theta_{jm})_{j=1,\dots,I,m=1,\dots,N_j,jm\neq in}\end{matrix}\right.\right](\theta_{j\prime m\prime}-c)x_{j\prime m\prime}[t]dt\right)$$

$$\leq\sum_{j\prime=1,\dots,I,m\prime=1,\dots,N_{j\prime}}\left(\int_\tau^1 \delta_j^*\left[t\left|\begin{matrix}\tau,Z,K^*\left[\theta_{in},(\theta_{j\prime m\prime})_{j\prime=1,\dots,I,m\prime=1,\dots,N_{j\prime},j\prime m\prime\neq in}\right],\\ \theta_{in},(\theta_{j\prime m\prime})_{j\prime=1,\dots,I,m\prime=1,\dots,N_{j\prime},j\prime m\prime\neq in}\end{matrix}\right.\right](\theta_{j\prime m\prime}-c)x_{j\prime m\prime}[t]dt\right).$$

Hence

---

[26] The blackout program represented by (A.8) solves the constrained maximization problem (7) for $K=K^*\left[\hat{\theta}_{in},(\theta_{jm})_{j=1,\dots,I,m=1,\dots,N_j,jm\neq in}\right]$ and the type profile $\left(\hat{\theta}_{in},(\theta_{jm})_{j=1,\dots,I,m=1,\dots,N_j,jm\neq in}\right)$ that the ISO believes to be the true type profile.

$$\text{(A.11)} \sum_{j=1,\dots,I,m=1,\dots,N_j} \varphi_{jm}\left[\begin{matrix}K^*\left[\hat{\theta}_{in},\left(\theta_{j\prime m\prime}\right)_{j\prime=1,\dots,I,m\prime=1,\dots,N_{j\prime},j\prime m\prime\neq in}\right], \\ \hat{\theta}_{in},\left(\theta_{j\prime m\prime}\right)_{j\prime=1,\dots,I,m\prime=1,\dots,N_{j\prime},j\prime m\prime\neq in}\end{matrix}\middle|\theta_{jm}\right]$$

$$= \sum_{j=1,\dots,I,m=1,\dots,N_j}\left(\pi[1]\int_0^1(\theta_{jm}-c)x_{jm}[t]dt - \frac{K^*\left[\hat{\theta}_{in},\left(\theta_{j\prime m\prime}\right)_{j\prime=1,\dots,I,m\prime=1,\dots,N_{j\prime},j\prime m\prime\neq in}\right]}{I}\right)$$

$$+\sum_{j=1,\dots,I,m=1,\dots,N_j}\int_0^1\int_0^\tau(\theta_{jm}-c)x_{jm}[t]dt\,dp[\tau] +$$

$$\sum_{j=1,\dots,I,m=1,\dots,N_j}\int_0^1\int_0^{X[\tau]}\int_\tau^1 \delta_j^*\left[t\middle|\begin{matrix}\tau,Z,K^*\left[\hat{\theta}_{in},\left(\theta_{j\prime m\prime}\right)_{j\prime=1,\dots,I,m\prime=1,\dots,N_{j\prime},j\prime m\prime\neq in}\right], \\ \hat{\theta}_{in},\left(\theta_{j\prime m\prime}\right)_{j\prime=1,\dots,I,m\prime=1,\dots,N_{j\prime},j\prime m\prime\neq in}\end{matrix}\right](\theta_{jm}-c)x_{jm}[t]f_{X[\tau]}[Z]dtdZ\,dp[\tau]$$

$$\leq \sum_{j=1,\dots,I,m=1,\dots,N_j}\left(\pi[1]\int_0^1(\theta_{jm}-c)x_{jm}[t]dt - \frac{K^*\left[\hat{\theta}_{in},\left(\theta_{j\prime m\prime}\right)_{j\prime=1,\dots,I,m\prime=1,\dots,N_{j\prime},j\prime m\prime\neq in}\right]}{I}\right)$$

$$+\sum_{j=1,\dots,I,m=1,\dots,N_j}\int_0^1\int_0^\tau(\theta_{jm}-c)x_{jm}[t]dt\,dp[\tau] +$$

$$\sum_{j=1,\dots,I,m=1,\dots,N_j}\int_0^1\int_0^{X[\tau]}\int_\tau^1 \delta_j^*\left[t\middle|\begin{matrix}\tau,Z,K^*\left[\theta_{in},\left(\theta_{j'm'}\right)_{j'=1,\dots,I,m'=1,\dots,N_{j'},j'm'\neq in}\right], \\ \theta_{in},\left(\theta_{j'm'}\right)_{j'=1,\dots,I,m'=1,\dots,N_{j'},j'm'\neq in}\end{matrix}\right](\theta_{jm} - c)x_{jm}[t]f_{X[\tau]}[Z]dtdZ\,dp[\tau],$$

$$= \sum_{j=1,\dots,I,m=1,\dots,N_j} \varphi_{jm}\left[\begin{matrix}K^*\left[\theta_{in},\left(\theta_{j\prime m\prime}\right)_{j\prime=1,\dots,I,m\prime=1,\dots,N_{j\prime},j\prime m\prime\neq in}\right], \\ \theta_{in},\left(\theta_{j\prime m\prime}\right)_{j\prime=1,\dots,I,m\prime=1,\dots,N_{j\prime},j\prime m\prime\neq in}\end{matrix}\middle|\theta_{jm}\right],$$

with the inequality following from (A.9). We have just shown that (A.5) is non-negative. ■

APPENDIX B

The Proof of Lemma 2

PROOF: Summing $\kappa_{in}^{*}[\theta]$ over $i = 1, \dots, I, n = 1, \dots, N_i$, we obtain

$$\sum_{i=1}^{I} \sum_{n=1}^{N_i} \kappa_{in}^{*}[\theta] = \sum_{i=1}^{I} \sum_{n=1}^{N_i} \left(\xi_{in}[\theta_{in}] + h_{in}^{*} \left[\left(\theta_{jm} \right)_{j=1,\dots,I,m=1,\dots,N_j,jm \neq in} \right] \right)$$

$$= \sum_{i=1}^{I} \sum_{n=1}^{N_i} \xi_{in}[\theta_{in}] + \sum_{i=1}^{I} \sum_{n=1}^{N_i} - \frac{1}{(\sum_{i=1}^{I} N_i) - 1} \sum_{j=1,\dots,I,m=1,\dots,N_j,jm \neq in} \xi_{jm}\left[\theta_{jm}\right]$$

$$= \sum_{i=1}^{I} \sum_{n=1}^{N_i} \xi_{in}[\theta_{in}] - \frac{1}{(\sum_{i=1}^{I} N_i) - 1} \sum_{i=1}^{I} \sum_{n=1}^{N_i} \left((\sum_{i=1}^{I} N_i) - 1 \right) \xi_{in}[\theta_{in}]$$

$$= \sum_{i=1}^{I} \sum_{n=1}^{N_i} \xi_{in}[\theta_{in}] - \sum_{i=1}^{I} \sum_{n=1}^{N_i} \xi_{in}[\theta_{in}] = 0.$$

■

APPENDIX C

Proof of Proposition 4

We prove Proposition 4 for the case an electric grid is made up of three areas with one consumer in each area. Generalization to the case the electric grid is made up of an arbitrary number of areas, with an arbitrary number of consumers in each area is straightforward.

We label the three agents i, j, and k. Let $\theta = (\theta_i, \theta_j, \theta_k)$ be a profile of types for the three agents. For the case of three areas, with one consumer in each area, (39) assumes the following form for A_i, A_j, and A_k, respectively,

(C.1) $\xi_i[\theta_i]$

$$= \int \left(v_j^0\left[K\left[\theta_i, \tilde{\theta}_j, \tilde{\theta}_k\right], \theta_i, \tilde{\theta}_j, \tilde{\theta}_k\right] - \frac{K}{3} + v_k^0\left[K\left[\theta_i, \tilde{\theta}_j, \tilde{\theta}_k\right], \theta_i, \tilde{\theta}_j, \tilde{\theta}_k\right] - \frac{K}{3} \right) d\Phi_j\left[\tilde{\theta}_j\right] d\Phi_k\left[\tilde{\theta}_k\right],$$

(C.2) $\xi_j\left[\theta_j\right]$

$$= \int \left(v_i^0\left[K\left[\tilde{\theta}_i, \theta_j, \tilde{\theta}_k\right], \tilde{\theta}_i, \theta_j, \tilde{\theta}_k\right] - \frac{K}{3} + v_k^0\left[K\left[\tilde{\theta}_i, \theta_j, \tilde{\theta}_k\right], \tilde{\theta}_i, \theta_j, \tilde{\theta}_k\right] - \frac{K}{3} \right) d\Phi_i\left[\tilde{\theta}_i\right] d\Phi_k\left[\tilde{\theta}_k\right],$$

and

(C.3) $\xi_k[\theta_k]$

$$= \int \left(v_i^0\left[K\left[\tilde{\theta}_i, \tilde{\theta}_j, \theta_k\right], \tilde{\theta}_i, \tilde{\theta}_j, \theta_k\right] - \frac{K}{3} + v_j^0\left[K\left[\tilde{\theta}_i, \tilde{\theta}_j, \theta_k\right], \tilde{\theta}_i, \tilde{\theta}_j, \theta_k\right] - \frac{K}{3} \right) d\Phi_i\left[\tilde{\theta}_i\right] d\Phi_j\left[\tilde{\theta}_j\right].$$

For consumer i, the participation constraint (43) now assumes the following form:

(C.4) $v_i^0\left[K\left[\theta_i, \theta_j, \theta_k\right], \theta_i, \theta_j, \theta_k\right] - \frac{K}{3} + \xi_i[\theta_i] - \frac{1}{2}\left(\xi_j\left[\theta_j\right] + +\xi_k[\theta_i]\right)$

$$= v_i^0\left[K\left[\theta_i, \theta_j, \theta_k\right], \theta_i, \theta_j, \theta_k\right] - \frac{K\left[\theta_i, \theta_j, \theta_k\right]}{3}$$

$$+\int\begin{pmatrix} v_j^0\left[K\left[\theta_i,\tilde{\theta}_j,\tilde{\theta}_k\right],\theta_i,\tilde{\theta}_j,\tilde{\theta}_k\right]-\frac{K\left[\theta_i,\tilde{\theta}_j,\tilde{\theta}_k\right]}{3} \\ +v_k^0\left[K\left[\theta_i,\tilde{\theta}_j,\tilde{\theta}_k\right],\theta_i,\tilde{\theta}_j,\tilde{\theta}_k\right]-\frac{K\left[\theta_i,\tilde{\theta}_j,\tilde{\theta}_k\right]}{3}\end{pmatrix}d\Phi_j\left[\tilde{\theta}_j\right]d\Phi_k\left[\tilde{\theta}_k\right]$$

$$-\quad\frac{1}{2}\begin{pmatrix}\int\begin{pmatrix} v_i^0\left[K\left[\tilde{\theta}_i,\theta_j,\tilde{\theta}_k\right],\tilde{\theta}_i,\theta_j,\tilde{\theta}_k\right]-\frac{K\left[\tilde{\theta}_i,\theta_j,\tilde{\theta}_k\right]}{3} \\ +v_k^0\left[K\left[\tilde{\theta}_i,\theta_j,\tilde{\theta}_k\right],\tilde{\theta}_i,\theta_j,\tilde{\theta}_k\right]-\frac{K\left[\tilde{\theta}_i,\theta_j,\tilde{\theta}_k\right]}{3}\end{pmatrix}d\Phi_i\left[\tilde{\theta}_i\right]d\Phi_k\left[\tilde{\theta}_k\right] \\ +\int\begin{pmatrix} v_i^0\left[K\left[\tilde{\theta}_i,\tilde{\theta}_j,\theta_k\right],\tilde{\theta}_i,\tilde{\theta}_j,\theta_k\right]-\frac{K\left[\tilde{\theta}_i,\tilde{\theta}_j,\theta_k\right]}{3} \\ +v_j^0\left[K\left[\tilde{\theta}_i,\tilde{\theta}_j,\theta_k\right],\tilde{\theta}_i,\tilde{\theta}_j,\theta_k\right]-\frac{K\left[\tilde{\theta}_i,\tilde{\theta}_j,\theta_k\right]K}{3}\end{pmatrix}d\Phi_i\left[\tilde{\theta}_i\right]d\Phi_j\left[\tilde{\theta}_j\right]\end{pmatrix}\geq 0.$$

Now suppose there exists a type profile $\theta=\left(\theta_i,\theta_j,\theta_k\right)$ such that

(C.5) $$v_i^0\left[K\left[\theta_i,\theta_j,\theta_k\right],\theta_i,\theta_j,\theta_k\right]-\frac{K}{3}+\begin{pmatrix} v_j^0\left[K\left[\theta_i,\theta_j,\theta_k\right],\theta_i,\theta_j,\theta_k\right]-\frac{K\left[\theta_i,\theta_j,\theta_k\right]}{3} \\ +v_k^0\left[K\left[\theta_i,\theta_j,\theta_k\right],\theta_i,\theta_j,\theta_k\right]-\frac{K\left[\theta_i,\theta_j,\theta_k\right]}{3}\end{pmatrix}$$

$$-\quad\frac{1}{2}\begin{pmatrix}\begin{pmatrix} v_i^0\left[K\left[\theta_i,\theta_j,\theta_k\right],\theta_i,\theta_j,\theta_k\right]-\frac{K\left[\theta_i,\theta_j,\theta_k\right]}{3} \\ +v_k^0\left[K\left[\theta_i,\theta_j,\theta_k\right],\theta_i,\theta_j,\theta_k\right]-\frac{K\left[\theta_i,\theta_j,\theta_k\right]}{3}\end{pmatrix} \\ +\begin{pmatrix} v_i^0\left[K\left[\theta_i,\theta_j,\theta_k\right],\theta_i,\theta_j,\theta_k\right]-\frac{K\left[\theta_i,\theta_j,\theta_k\right]}{3} \\ +v_j^0\left[K\left[\theta_i,\theta_j,\theta_k\right],\theta_i,\theta_j,\theta_k\right]-\frac{K\left[\theta_i,\theta_j,\theta_k\right]}{3}\end{pmatrix}\end{pmatrix}$$

$$=\frac{1}{2}\left(v_j^0\left[K\left[\theta_i,\theta_j,\theta_k\right],\theta_i,\theta_j,\theta_k\right]+v_k^0\left[K\left[\theta_i,\theta_j,\theta_k\right],\theta_i,\theta_j,\theta_k\right]\right)-\frac{K\left[\theta_i,\theta_j,\theta_k\right]}{3}>0.$$

Then if the range in which the preferences for electricity consumption of each consumer is narrow, then using (C.5) and a continuity argument, we can show that the participation constraint (C.4) hold with strict inequality for consumer i.

In the same manner, if the range in which the preferences for electricity consumption of each consumer is narrow, and if

(C.6) $\frac{1}{2}\left(v_i^0\left[K\left[\theta_i,\theta_j,\theta_k\right],\theta_i,\theta_j,\theta_k\right]+v_k^0\left[K\left[\theta_i,\theta_j,\theta_k\right],\theta_i,\theta_j,\theta_k\right]\right)-\frac{K\left[\theta_i,\theta_j,\theta_k\right]}{3}>0,$

(C.7) $\frac{1}{2}\left(v_i^0\left[K\left[\theta_i,\theta_j,\theta_k\right],\theta_i,\theta_j,\theta_k\right]+v_j^0\left[K\left[\theta_i,\theta_j,\theta_k\right],\theta_i,\theta_j,\theta_k\right]\right)-\frac{K\left[\theta_i,\theta_j,\theta_k\right]}{3}>0,$

hold, then we can show

(C.8) $v_j^0\left[K\left[\theta_i,\theta_j,\theta_k\right],\theta_i,\theta_j,\theta_k\right]-\frac{K}{3}+\xi_j\left[\theta_j\right]-\frac{1}{2}(\xi_i[\theta_i]++\xi_k[\theta_k])>0$

and

(C.9) $v_k^0\left[K\left[\theta_i,\theta_j,\theta_k\right],\theta_i,\theta_j,\theta_k\right]-\frac{K}{3}+\xi_k[\theta_k]-\frac{1}{2}\left(\xi_i[\theta_i]++\xi_j\left[\theta_j\right]\right)>0,$

i.e., the participation constraint also holds with strict inequality for consumer j and consumer k.

■

<u>REFERENCES</u>

Admunsen, E. S. AND L. Bergman (2007): "Provision of operating reserve capacity: Principles and practices on the Nordic Electricity Market," Competition and Regulation in Network Industries, 2(1): 73-98.

Anderson, Lindsay C. AND Matt Davison (2008): "A Hybrid System-Econometric Model for Electricity Spot Prices Considering Spike Sensitivity to Forced Outage Distributions," *IEEE Transactions on Power Systems*, 23(3), 927-937.

Crampton Peter, Axel Ockenfels and Steven Stoft (2013): "Capacity Market Fundamentals," *Journal Economics of Energy and Environmental Policy,* 2, 27-46.

Creti, Anna and Natalia Fabra (2007): "Supply security and short-run capacity markets for electricity," *Energy Economics* 29, 259-276.

Hobbs, B. F., M. H. Rothkopf, L. C. Hyde, AND R. P. O'Neill (2000): "Evaluation of a Truthful Revelation Auction in the Context of Energy Markets with Non-concave Benefits," *Journal of Regulatory Economics*, 18(1), 5-32.

Jaffe, Adam B. AND Frank A. Felder (1996): "Should Electricity Markets Have A Capacity Requirement? If So, How Should It Be Priced?" *The Electricity Journal*, 9(1), 52-60.

Joskow, Paul and Jean Tirole (2007): "Reliability and Competitive Electricity Markets," *The Rand Journal of Economics* 38 (1), 60-82.

Mas-Colell, A, Whinston, M.D. and Green, J (1995): "Microeconomic Theory," Oxford University Press.

Nooij, M., R. Lieshout, AND C. Koopmans (2009): “Optimal Blackouts: Empirical Results on Reducing the Social Cost of Electricity Outages through Efficient Regional Rationing,” *Energy Economics* 31, 342–347.

Trodden, P. A., W. A. Bukhsh, A. Grothey, AND K. I. M. Mckinnon(2013): “MILP Formulation for Islanding of Power Networks,” *International Journal of Electrical Power and Energy Systems,* February 2013, Vol.45(1), 501-508.