Queen's University at Kingston

An essay submitted to the Department of Economics in partial fulfilment of the requirements for the degree of Master of Arts

Maintenance of Electrical Generation in a Market Context

Author:
Daniel Jensen

Supervisor:
Dr. Roger Ware

I wish to thank everyone who has given me support over the past year, in particular Robin Shaban and my supervisor, Roger Ware. I would also like to thank Matt Ayres and everyone else at the Alberta Market Surveillance Administrator for inspiring this paper. Finally, I would like to thank my sister Glynnis.

Contents

1 Introduction

One of the fundamental concepts in the electricity industry is reliability. Virtually every electrical system in the world is administered with the goal of keeping the power flowing above and beyond all other priorities. In the context of deregulated electricity markets the question of reliability takes on additional dimensions. Instead of a regulated government entity controlling all aspects of generation, maintenance decisions are made by private firms governed by a profit motive. Maintenance becomes a strategic variable, potentially a very important one. Even if sufficient capacity remains to power the grid, the failure of a key plant can increase the price of electricity significantly. Power system reliability prevents inconvenience and economic loss resulting from electrical shortages and encourages efficient outcomes in deregulated markets.

The following will develop a theoretical framework to study the incentives firms face when choosing to how much to invest in plant maintenance. In a two stage game firms will first simultaneously choose maintenance levels, determining how likely it is that their plants malfunction. We will then introduce two variants on second stage competition in order to explore the relationship between market outcomes and maintenance investment. First, we will consider the effects of forward electricity sales on maintenance levels. Second, we will develop a variant on the traditional demand curve that better reflects the demand curve that strategic firms face. [1] We find that in symmetric duopoly an increase in identical contracts imposed on both firms increases system reliability (for 'reasonable' levels of contracts), as does additional demand beyond what the competitive fringe may supply.

There are three areas of literature relevant to this paper within the sub field of

[1] Short run electrical demand is typically considered totally inelastic. One explanation as to why strategic firms face a downward sloping demand curve is a competitive fringe consisting of imports and price taking firms. Underlying demand is inelastic, but the *residual* demand curve becomes price responsive. We will consider the effect of this fringe being unable or unwilling to supply total demand, that is, there being a totally elastic region on the demand curve where additional output does not lower price.

electricity: General modelling, plant maintenance choices, and the theory of futures contracts. Deregulated electricity markets have received a good deal of attention from economists. Notable references include Borenstein, Bushnell, and Wolak (2002); Hogan (1997). Ventosa et al. (2005) and Garcia, Momoh, and Milli (2008) provide overviews of general modelling trends.

System operators and economists are traditionally very concerned about reliability from the long-run perspective of whether sufficient generation will be built to meet demand. How much generation capacity is deemed necessary includes a 'reserve margin', excess capacity beyond projected demand to accommodate plant failures. The Brattle Group (2012) is recent example of how concerns about maintaining a sufficient capacity buffer can inform market design, while Joskow (2006) is a more academic treatment of investment incentives. Very little however has been written on reliability and plant maintenance from an economic (rather than engineering) perspective. Deltas and Hadjicostis (2009) propose a model of plant reliability to compare monopoly and duopoly outcomes to the social optimum. Joskow and Tirole (2004) focuses on investment and rationing when demand outstrips supply, and includes an extension based on endogenous reliability. In many ways, this model of maintenance resembles traditional models of two stage competition (for instance, see Fudenberg and Tirole (1984)), but has one critical difference. In this model, first stage choices (maintenance) affect which type of competition will occur in the second stage, but do not change the nature of that competition. The choice of maintenance levels will dictate how likely a firm is to be a monopolist, but will not alter the choices the monopolist will make.

The theory of two stage competition is much more applicable to futures contracts and competition. In a seminal paper Allaz and Villa (1993) show that in a symmetric Cournot model signing futures contracts confer a strategic advantage but reduce overall profits. Subsequently due to their prevalence of contracting in real markets many influential papers have included financial obligations in models of electricity market dynamics, such as Green (1999) and Newberry (1998). Empirical studies of the effect of hedge contracts have been carried out including Wolak (2000). Other features of market structure can have similar effects on incentives; Bushnell et al. (2008) is an empirical study of generating firms vertically integrating into retail markets.

After concluding the introduction with a brief explanation of electricity markets, the next section will develop a theoretical model of plant maintenance. As plant maintenance choices are fundamentally linked to second stage market outcomes, we then explore the theory of maintenance in relation to variations in competition. A section will examine the effects of futures contracts on maintenance levels, then the next will study a growth in market demand beyond what a competitive fringe may supply respectively. We first find that low levels of selling production forward increases the incentive to maintain plants. We then show that additional market demand increases equilibrium maintenance spending and overall reliability.

1.1 An Overview of Electricity Markets

Electricity markets are sufficiently unique to warrant a brief primer. [2] Most of the peculiarities come from the physical nature of electricity itself, and the necessity of power being delivered through a network that connects many producers and consumers anonymously. There are two physical characteristics of electricity that truly differentiate it from other commodities:

- *Electricity is Non-Storable:* With a few (very limited) exceptions electricity cannot be stored on a significant scale. Supply and demand in electricity markets must match each other at all times.
- *Electricity is Homogeneous:* At the point of use, no two watts of electricity are

[2] The following is based on Alberta MSA (2010b), Hogan (1998), Kirschen et al. (2000)

distinct. Unlike, say, crude oil which has many grades and qualities depending on source, electricity is a single indistinguishable product.

Electrical systems with deregulated markets and regulated systems alike are controlled by a centralized organization (usually styled the Independent System Operator "ISO" in market systems). The ISO has the task of dispatching generation to produce electricity equal to current demand. Because the ISO coordinates and monitors production it typically administers an electricity market.

Market procedures differ between jurisdictions, [3] but most follow the same general contours. A day is divided up into periods, known as settlement intervals. For every settlement interval generators submit offers to the ISO. These are usually a number of price/quantity blocks that effectively form a supply curve for each generating unit. For instance, 350 MW coal generator could offer in 100 MW at $0, 150 MW for $10, and the final 100 MW for $100. Horizontally aggregating all of these offers provides the market supply curve, known as the Merit Order. Most markets are subject to price caps and price floors [4] designed to prevent speculative, excessively high offers from clearing the market in extreme circumstances.

The job of the ISO is to ensure there is exactly as much supply (generation) as demand (load) at any point in time, which it accomplishes by creating a market for electricity and then setting a price that clears the market. Unlike generation, the ISO does not have control of how much electricity is consumed – there is no reasonable way to compel ten thousand homes to stop using their ovens. Electricity demand is then considered completely inelastic for the purposes of clearing the market. The instantaneous price of electricity is however much the market requires to completely

[3] Due to the Author's experience, the following will consider the market in Alberta as the baseline. Alberta has a relatively simple market as, in the jargon, it has a 'Single Price, Energy Only' design. Single Price means that a single price prevails everywhere instead of different prices at different locations within the market. Energy Only means that firms only get paid for the energy they produce; in some other markets generators get payments for having capacity available.

[4] Some markets allow negative prices, or paying load to consume electricity. Naturally, negative prices are rare but are occasionally posted in certain markets. For instance, see Energy Information Administration (2012) at `http://www.eia.gov/todayinenergy/detail.cfm?id=6730`

fill demand. A single price prevails for the settlement interval, which is determined by taking the average of the instantaneous price.

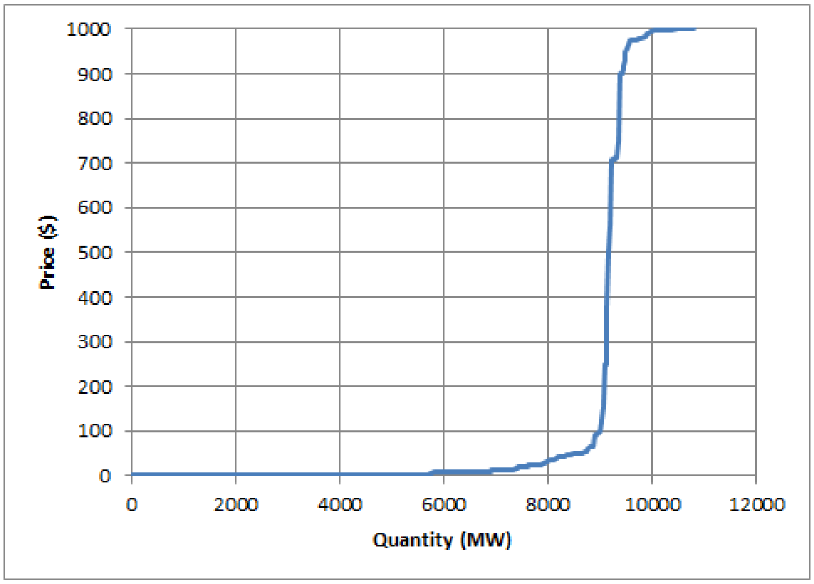

Figure 1: Sample Demand Curve

The supply curves that arise from offers are fairly unusual. Figure 1 provides a 'typical' offer curve for Alberta, from 12 AM to 1 AM on January 1 2013. [5] A significant fraction of provincial generation is actually offered for $0. A common argument is this is not because they desire to produce for free but because for many plants turning off and on again is slow and costly. There is a minimum level most plants can produce, below which the machinery cannot operate safely. Offering the minimum stable generation for free ensures that the plant is never shut off if it can be avoided. At higher demand levels, the typical supply curve becomes highly inelastic, climbing very quickly to the price cap. This is typically argued to be partly for strategic reasons, and partly due to increasing marginal cost. Price is usually set on the region of the supply curve greater than the $0 offers but lower than when it becomes highly inelastic. Figure 2 shows the duration curve for the 2012 Alberta pool price, that is, the amount of the time the price was less or equal to a given value.

[5] Data from the Alberta Electric System Operator website `http://ets.aeso.ca/` (AESO 2013)

Price is lower than \$10 in only about 5% of the hours (and virtually never zero), and less than \$100 90% of the time.

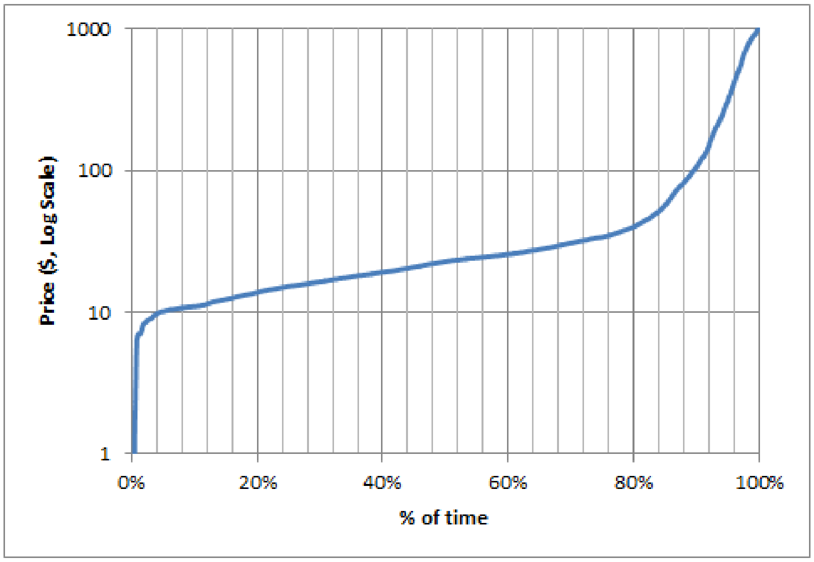

Figure 2: Alberta 2012 Pool Price Duration Curve

When a plant malfunctions and is forced off-line, that supply is removed from the merit order, shifting it to the left. If market demand is close to the 'cliff', and that plant has bid its supply in for a low price, removing a few hundred megawatts from the market has the potential to increase the price by an order of magnitude. For instance on the merit order depicted in figure 1, there is roughly 450 MW between a price of \$90 and \$900. Coal power plants in Alberta range between 144 MW and 466 MW in capacity [6] so such a swing is hardly impossible.

2 The Choice of Maintenance Investment

To examine the incentives for a firm to invest in plant maintenance, consider the following game:

1. In stage 1, firms simultaneously choose some maintenance level, m_i, which will give their plant some probability of availability $F_i(m_i)$, which is independent

[6] AESO (2013) at `http://ets.aeso.ca/`

of the availability realizations of the other firms. If the plant breaks down the firm cannot produce any electricity at all.

2. Uncertainty is resolved, and firms engage in Cournot competition

For simplicity, the following will consider the risk neutral duopoly case where each firm controls a single plant. Firms are assumed to be risk neutral. We will assume $F(m_i)$ is twice differentiable and continuous, $F_i(m_i) : [0, \infty) \to [0, 1]$, $F_i' = f_i(m_i) > 0$, $F_i''(m_i) < 0$, and $F_i'(0) = \infty$ to ensure an interior solution in maintenance levels.

There are a few things to note about $F_i(m_i)$. First, it has an alternate interpretation: instead of the probability of failure in a single period, it can be interpreted as the fraction of the time the plant is expected to be unavailable. Instead of a 30% chance of failure, a plant could be off-line 30% of the time. Second, while it must be bounded by 0 and 1, these need not be the greatest lower bound nor the lowest upper bound respectively. Even an infinite amount of maintenance expenditure may not reduce the probability of failure to zero, nor must a plant with \$0 invested in maintenance fail. Third, this is a very stylized presentation of failure, with a focus on mathematical tractability. For instance, in the real world generation levels can have a significant impact on failure: repeatedly starting up and shutting down a coal power plant, for instance, places great stress on the machinery. This would however require adding a dynamic element to the model. Moreover, maintenance need not reduce overall down time, as often maintenance must be performed when the plant is off-line. This would suggest that maintenance expenditure reduces *unplanned* downtime, but increases *planned* downtime, which could provide another strategic variable for the firm.

There are four possible states of the world following the realization of plant availability: either both firms have working electrical plants, both fail, or one of the two is left with a functioning plant but the other has malfunctioned.

Let:

$\pi_i^M \equiv$ The profits if the firm is the only one active in the market. (Monopoly)

$\pi_i^D \equiv$ The profits if both firms remain active. (Duopoly)

$\pi_i^E \equiv$ The profits if the firm experiences a malfunction but its rival does not. (Exclusion)

$\pi_i^B \equiv$ The profits if both plants break down. (Total Breakdown)

2.1 The Problem of the Firm

When choosing their maintenance level, firms maximize expected profits:

$$\max_{m_i} E(\pi_i) = F_i(m_i)F_j(m_j)\pi_i^D + F_i(m_i)(1 - F_j(m_j))\pi_i^M$$
$$+ (1 - F_i(m_i))F_j(m_j)\pi_i^E + (1 - F_i(m_i))(1 - F_j(m_j))\pi_i^B - m_i \quad (1)$$

Which provides the first order condition

$$\frac{\partial E(\pi_i)}{\partial m_i} = f_i(m_i)F_j(m_j)\pi_i^D + f_i(m_i)(1 - F_j(m_j))\pi_i^M$$
$$- f_i(m_i)F_j(m_j)\pi_i^E - f_i(m_i)(1 - F_j(m_j))\pi_i^B - 1 = 0 \quad (2)$$

Which can be better interpreted as

$$1 = f_i(m_i)[E(\pi_i|\text{Active}) - E(\pi_i|\text{Breakdown})] \quad (3)$$

To understand the first order condition it is helpful to dissect precisely what a dollar of maintenance expenditure will buy the firm. The overall benefit of a dollar of maintenance expenditure is the difference between the expected profit if their plant is active and if it breaks down, multiplied by the increase in the probability the plant will remain active that the dollar will affect. The larger the difference between the two profit levels, the more incentive there is to invest in plant maintenance.

How equilibrium incentives to invest in maintenance change in response to second stage market dynamics depends on three things. First, $f_i(m_i)$, the effect a dollar of maintenance will have on probabilities. Second, the change in Cournot profit levels. Third, how the other firm's maintenance choice will change.

It is informative to consider what is necessary for maintenance levels to increase in the symmetric equilibrium. Because we have assumed that $F(m)$ is strictly concave, we know the increase in m causes $f(m)$ to decrease: to maintain equality, the difference in the two expected profit levels must increase. However, as we are considering the symmetric equilibrium, if firm i increases its maintenance spending so must firm j. When firm j (the rival) increases its spending, it increases the weight put on π_i^D and π_i^E over pi_i^M and π_i^B respectively. Equation (4) shows what the difference in expectations evaluates to:

$$E(\pi_i|\text{Active}) - E(\pi_i|\text{Breakdown}) = (1 - F(m_j))[\pi^M - \pi^B] + F(m_j)[\pi^E - \pi^D] \quad (4)$$

As $F(m)$ increases, were we to hold the profit levels constant, the second term becomes more heavily weighted at the expense of the first term. We then need for $\pi^M - \pi^B$ or $\pi^E - \pi^D$ to increase (in the case of the latter, become less negative), while the other does not decrease sufficiently to dominate the effect.

2.2 The First Order Condition and Comparative Statics

We can rearrange (2) to get the equilibrium condition for firm j's probability of failure in terms of firm i's profits and maintenance expenditure:

$$F_j(m_j) = \frac{(\pi_i^M - \pi_i^B) - \frac{1}{f_i(m_i)}}{(\pi_i^M - \pi_i^B) - (\pi_i^D - \pi_i^E)} \quad (5)$$

We need not invoke symmetry to know that firm i will choose m_i such that (5) holds, however in the symmetric case this equation will prove particularly valuable. The same intuition for the f.o.c. applies to (5), but the latter highlights important implications of the boundedness of $F_j(m_j)$ for the choice of m_i.

We can easily show that both the numerator and the denominator of (5) are positive. Because $F_j(m_j)$ is greater than zero, the choice of m_i must be such that both the numerator and the denominator have the same sign. As usual, we know $\pi_i^M > \pi_i^D$, so $\pi_i^M - \pi_i^D > 0$. If we assume that $\pi_i^E - \pi_i^B \geq 0$ [7] the denominator (and therefore the numerator) must be positive. This means that for $F(m_j)$ to be appropriately bounded between zero and one, we require that $\pi_i^M - \pi_i^B \geq 1/f_i(m_i) \geq \pi_i^D - \pi_i^E$.

If we assume symmetry and then totally differentiate with respect to m and some model parameter that affects profit levels α we can show that

$$\begin{aligned}\text{sign}\left(\frac{\partial m}{\partial \alpha}\right) = \text{sign}\Bigg(&\frac{\partial}{\partial \alpha}\left[\pi^M - \pi^B\right]\left(\pi^M - \pi^B + \pi^E - \pi^D\right) \\ &-\frac{\partial}{\partial \alpha}\left[\pi^M - \pi^B + \pi^E - \pi^D\right]\left(\pi^M - \pi^B - \frac{1}{f(m)}\right)\Bigg) \qquad (7)\end{aligned}$$

If α has a positive effect on m all firms will invest more in maintenance. This will reduce the chance that both plants will fail simultaneously, and increase the time spent in a duopoly, the most efficient market structure. Confirming the intuition above, it can be shown that (7) is equivalent to increasing the difference between expected profit if the firm's generation remains active and if it fails.

[7] In the following discussion, the reason why π_i^E and π^B will be non-zero is that the generator has sold production forward through futures contracts. Every period they must pay the market price times the contract quantity. We assume that if the market breaks down (both generators are forced offline) the price will rise to the vertical intercept of the demand curve, which may be the market price cap. This vertical intercept is greater or equal to the price which a monopolist will charge. When a firm breaks down but its rival remains, the rival acts as a monopolist. By this logic, letting $\phi_i > 0$ be the contract quantity and $\bar{P}$ be the vertical intercept

$$\pi_i^E - \pi_i^B = -\phi_i(P_j^M - \bar{P}) = \phi_i(\bar{P} - P_j^M) \geq 0 \qquad (6)$$

Prior to invoking symmetry, this framework is sufficiently general that it can be applied to any formulation of second stage duopoly competition. In the following analysis we will assume firms face a linear downward sloping demand curve. [8] The inverse demand curve is $P(x_1, x_2) = a - x_1 - x_2$, where the slope is normalized to one. Capacity constraints are assumed to be non-binding, that is, both firms have invested in capacity at least equal to the unconstrained monopoly output. Further, both firms are assumed to have identical, constant marginal cost c and that $a > c$.

3 Futures Contracts

One lens through which we can examine how incentives for maintenance change is the introduction of futures contracts. In real electricity markets, firms usually sell their production forward in some way. Options include futures contracts, long term purchase agreements with system operators, or participation in the retail electricity market. The following will examine the effect of a change in the level of identical contracts exogenously imposed on both firms (e.g. government procurement programs [9]). One could conceive of a three stage game where firms choose both their level of contract cover and maintenance expenditure before competing in the market. However, this comes at a significant cost in terms of tractability. Assuming firms have the same level of contract cover is not as limiting as it may appear. This model is the same as if the contract choice stage has already occurred and the symmetric firms have chosen a symmetric level of contract cover.

[8] In the short run, electricity demand curves are generally taken to be inelastic. This downward sloping demand curve is usually taken to be *residual* supply curve left for strategic firms, after the supply of price taking firms and imports have been netted out.

[9] This is how investment in Ontario is procured. The Ontario Power Authority has 22 282 MW of contracted capacity (Ontario Power Authority 2013), compared to a record peak demand in 2006 of 27 005 MW (IESO 2012).

3.1 Introduction to Futures Markets

Like most commodities, electricity is traded on both spot and forward markets. [10] A forward contract is simply an agreement to exchange an amount of electricity at an agreed price (known as the strike price) at a certain time. A firm may offer to sell 50 MW of electricity at every moment in the coming month for $20 a MW.

Most other commodities are underpinned by physical delivery. At least in theory, an individual who buys a futures contract for 100 tonnes of wheat and carries it to maturity will find themselves in possession of a silo's worth of grain. The nature of the electrical grid makes the same thing impossible when transacting in electricity. Generators inject power into the grid and load draws power from it, but there is no way for generator A to provide 50 MWh of power to factory B in the same way grain can be delivered. Instead, electricity futures are purely financial exchanges that can simulate the effects of physical delivery, known as Contracts for Difference (CFD).

A CFD obliges the buyer of the contract to pay the seller the strike price for each unit, and the seller pays the buyer the market price. To illustrate, imagine a 100 MW contract with a strike price of $70, while the spot price is $50. The seller then pays the buyer $5000 ($50 per unit) that settlement interval, and the buyer pays the seller $7000 ($70 per unit). To see how this mimics the physical delivery of 100 MW of electricity, imagine the seller is a generator that physically sells 100 MW on the spot market, and the buyer is a load that consumes 100 MW. Without the contract, when the generator sells the power into the spot market it receives $7000 (the spot price for 100 MW). Likewise, the load must pay $7000 for the 100 MW. With the contract, the payout to the generator is $7000 (the market revenue) minus $7000 (the payment to the load) plus $5000 (the payment from the contract) – the generator gets $50 per megawatt delivered to the grid. Likewise, the load must pay $7000 for the power it uses in the spot market, but gets $7000 from the generator in exchange for $5000.

[10] For more information, see Alberta MSA (2010a), Hull (2010).

The load also pays \$50 per megawatt, exactly as if physical delivery had occurred.

Different agents participate in futures markets with different motives. Most finance textbooks [11] will provide at least three goals: hedgers, speculators, and arbitrageurs. Hedgers are firms that are exposed to the spot market because they produce or consume the commodity and seek to manage their price risk. The example just given of how the futures market can simulate physical delivery would function as a hedge. Speculators exploit the difference between the current futures price and the price they expect to prevail to make a profit. Arbitrageurs attempt to make a margin buying in low markets and selling in higher priced venues. In electricity, locational arbitrage is possible through imports and exports (and futures may aid in this strategy). However, inter temporal arbitrage based on exploiting differences between a single spot and future market is impossible due to the difficulty of storing electricity.

There is a literature on the effects of forwards contracting on oligopoly. When a firm sells production forward, it will prefer a lower price than it will otherwise, as the payments to the buyer increase with the market price. In Cournot models, this could be considered a reduction in the marginal cost of production, [12] encouraging greater output. For instance, see Allaz and Villa (1993).

3.2 The Duopoly Problem

The problem facing the firm in the duopoly outcome is

$$\max_{x_i} \pi^D = (P(x_1, x_2) - c)x_i - \phi_i P(x_1, x_2) \tag{8}$$

Where ϕ_i is the level of futures contracts the firm has sold, which is assumed to be greater than zero. While generators can buy forward as well, the case of forward

[11] See Hull (2010) for one standard text.

[12] One way to think of it is that for every additional unit of output the market price is lowered, reducing the opportunity cost of having sold forward. Other interpretations, including increases in marginal revenue, are possible but equivalent.

sales is generally more applicable. For simplicity, the revenue from contracts, $S_i\phi_i$, is omitted. [13]

Following standard solution techniques, we get

$$x_i^D = \frac{a - c + 2\phi_i - \phi_j}{3} \tag{9}$$

$$P^D = \frac{a + 2c - \phi_i - \phi_j}{3} \tag{10}$$

Which, when both firms have identical contract volumes, corresponds with the profit level

$$\pi_i^D = \left(\frac{a - c - 2\phi}{3}\right)\left(\frac{a - c + \phi}{3}\right) - \phi\left(\frac{a + 2c - 2\phi}{3}\right) \tag{11}$$

$$= \frac{4\phi^2 - (4a + 5c)\phi + (a - c)^2}{9} \tag{12}$$

3.3 The Monopoly Problem

Similar to the duopoly problem, firms seek to maximize their objective function:

$$\max_{x_i} \pi^M = (a - x_i - c)x_i - \phi_i(a - x_i) \tag{13}$$

Which is solved by the output (and price)

$$x_i^M = \frac{a - c + \phi_i}{2} \tag{14}$$

$$P_i^M = \frac{a + c - \phi_i}{2} \tag{15}$$

[13] Recall, every period for each contract unit (ϕ_i), the firm pays the market price ($P(x_1, x_2)$) and receives the strike price (call it S_i). Revenue from contracts, $S_i\phi_i$ is the same regardless of their production choice and market structure, and therefore does not affect output or maintenance choices.

Which give the maximum value function

$$\pi_i^M = \left(\frac{a-c-\phi}{2}\right)\left(\frac{a-c+\phi}{2}\right) - \phi\left(\frac{a+c-\phi}{2}\right) \quad (16)$$

$$= \frac{\phi^2 - 2(a+c)\phi + (a-c)^2}{4} \quad (17)$$

Further, we note that if a firm is 'excluded' from the market due to breakdown (i.e. the other firm remains as a monopolist) then the same monopoly price will prevail and the profits of the excluded firm will be $\pi_i^E = -\phi_i P_j^M$. Naturally, when both plants fail the price is $P^B = a$ and $\pi_i^B = -\phi_i a$ (as $x_1 = x_2 = 0$). As the price when both firms experience breakdowns is higher than the monopoly price, $\pi_i^E > \pi_i^B \implies \pi_i^E - \pi_i^B > 0$ confirming the denominator of (5) must be positive.

3.4 The effect of Contracts on Maintenance Levels

As it turns out, the effect of contracts on the symmetric maintenance equilibrium is generally ambiguous. We will first restrict our analysis to two points: Evaluating the derivative when contracts are zero and when they are high enough to induce marginal cost pricing in the duopoly outcome. We will then show that for all points in between, the effect of contracts on maintenance is positive.

We can substitute the equilibrium profit levels into (7) to find that

$$\text{sign}\left(\frac{\partial m}{\partial \phi}\right) = \text{sign}\left(\frac{11\phi + 8(a-c)}{18}\frac{1}{f(m)} - \frac{(a-c)(a-c+\phi)^2}{24}\right) \quad (18)$$

3.4.1 No Contracts ($\phi = 0$)

Evaluating (18) at $\phi = 0$, we get the expression

$$\text{sign}\left(\frac{\partial m}{\partial \phi}\right) = \text{sign}\left(\frac{8(a-c)}{18}\frac{1}{f(m)} - \frac{(a-c)^3}{24}\right) \quad (19)$$

The sign will then be given by

$$\text{sign}\left(\frac{\partial m}{\partial \phi}\right) = \text{sign}\left(\frac{1}{f(m)} - \frac{3}{32}(a-c)^2\right) \tag{20}$$

Which at first sight seems to be ambiguous. However, if we evaluate (5) at $\phi = 0$ as well (i.e. where $\pi_i^B = \pi_i^E = 0$), for $F(m)$ to be less than one it is easy to see it must be the case that $1/f(m) > \pi_i^D = \frac{1}{9}(a-c)^2$. As $\pi_i^D > \frac{3}{32}(a-c)^2$, it must be the case that the first unit of symmetric contract cover increases the equilibrium level of maintenance.

3.4.2 Contracts Induce Marginal Cost Pricing in Duopoly

Suppose instead the firms were endowed with sufficient contracts that the market price would equal marginal cost in duopoly. Using (10), we can see that this will happen when $\phi = \frac{1}{2}(a-c)$. Evaluating (18) at that level shows that additional contracts will increase maintenance levels if

$$\frac{(1}{f(m)} > \frac{(a-c)^2}{8} \tag{21}$$

Using the same technique as before, we can see that $\pi^D - \pi^E = \frac{1}{8}(a-c)^2$ as well. Therefore, contracts have a positive effect on maintenance levels when ϕ is 'high' as well.

3.4.3 Intermediate Values

It can be shown that contracts will have a positive effect on maintenance levels for all values of ϕ between 0 and $\frac{1}{2}(a-c)$. We can rearrange (18) to see that $\frac{\partial m}{\partial \phi}$ will be positive if

$$\frac{1}{f(m)} > \frac{3(a-c)(a-c+\phi)^2}{4(11\phi + 8(a-c))} \tag{22}$$

We can use the same technique as above (that is, exploiting the boundedness of $F(m)$) so long as $\pi^D - \pi^E \geq$ RHS of (22). This will be the case so long as

$$-22\phi^2 + (a-c)\phi + 5(a-c)^2 \geq 0 \tag{23}$$

Which will hold if $\frac{1}{2}(a-c) \geq \phi \geq -\frac{5}{11}(a-c)$. In other words, between zero contracts and contracts high enough to induce marginal cost pricing in duopoly, increasing the level of contracts strictly increases the equilibrium level of expenditure. It should be noted that ϕ *may* increase maintenance levels outside of this interval. However we cannot use the same technique to sign the derivative, so without additional assumptions no claims can be made for other values. [14]

3.5 Interpretation and Conclusions about Contracts and Maintenance

Firms will want to invest more in maintenance if additional contracts increase the return to reliability. Specifically, recalling (3), we require the difference between the expected profits levels conditional on plant status to grow. This is complicated by the fact that in the symmetric equilibrium the increase in maintenance spending applies to both firms, which tends to make maintenance less attractive. According to (4), this is tantamount to increasing the difference between $\pi^M - \pi^B$ and $\pi^E - \pi^D$ sufficiently to offset the increased probability of the rival plant remaining active.

In general, the effect of contracts on the profit levels are ambiguous. For low levels of ϕ [15] additional contracts reduce π^M and π^D, as they encourage output beyond the traditional monopoly and duopoly optimum. They also tend to make π^B and π^E more

[14] One may question whether we can derive any results using the other boundedness condition of (5), namely, $1/f(m) < \pi^M - \pi^B$. A similar analysis shows that contracts will decrease maintenance levels when $-(a-c) < \phi < -\frac{5}{11}(a-c)$. However, we are focusing on positive levels of ϕ, where generators have sold rather than bought electricity forward.

[15] For high levels of ϕ, all of these effects reverse other than π^B. We ignore these cases as they imply price is very low, either below marginal cost or negative.

strongly negative. The cumulative effect is to make $\pi^M - \pi^B$ larger (for all levels of ϕ), but decreases $\pi^E - \pi^D$ (when ϕ is low). Coupled with the confounding effect of the other firm increasing maintenance spending, we are forced to investigate for what ranges of ϕ additional contracts cause an increase or decrease in m, by investigating the effect on $E(\pi_i|\text{Active}) - E(\pi_i|\text{Breakdown})$ analytically.

We have shown that it must be the case that between the levels of no contracts and contracts high enough to induce marginal cost pricing in the duopoly market structure, the effect of contracts on maintenance is positive. This suggests that maintenance spending will be maximized at some level of contracts *beyond* that where marginal cost duopoly pricing prevails. Putting aside the question of what the efficient level of maintenance is, this suggests the policy of encouraging quantity based forward sales as a method of increasing system reliability and effective competition.

4 Cournot Competition with Perfectly Elastic Demand Regions

A second variation on market structure that can be used to explore choices of maintenance levels is to allow for a region of totally elastic demand. One of the peculiarities of electricity markets is in the short run demand is almost totally inelastic. Downward sloping demand curves have been argued to be the result of a competitive fringe, including price taking generators and imports – see, for instance, Borenstein et al. (2002) where downward sloping demand is a result of price sensitive imports or Bushnell et al. (2008). These smaller producers do not act strategically and take price as given. The downward sloping demand curve the strategic firms face is then the residual demand curve after the competitive fringe's supply has been removed from the underlying inelastic demand. It should be noted that this practise is not universal in the literature, but competitive fringes are often mentioned in relation to the residual

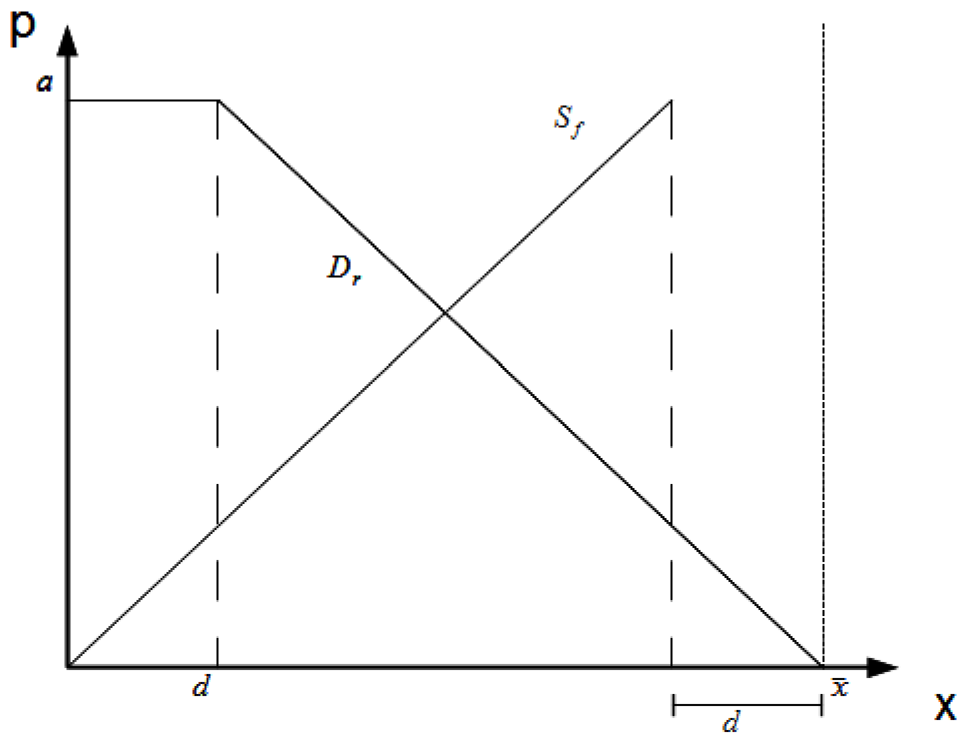

Figure 3: Residual Demand Curve

demand curve, as in Green (1999).

However, a second peculiarity of electricity markets are binding supply constraints. The following assumes that the competitive fringe does not have the capacity to supply the entire market at any price. As a result, there will be a perfectly elastic region on the demand curve where expanding output does not come at the cost of decreasing price.

Figure 3 illustrates the situation. Present market demand is given by $\bar{x}$, which is totally inelastic. However, once supply from the competitive fringe S_f is subtracted, we get the residual demand curve D_r. A strategic firm could supply any quantity up to d and not decrease the market price, as even at the market price cap the competitive fringe can only supply $\bar{x} - d$.

An alternate explanation for this type of supply curve would be to remove the fringe and allow market demand to be price sensitive, but subject to a price cap. As long as some consumers value the consumption of electricity more than the price cap the same phenomenon can arise. Under this interpretation the results of this section would apply to a growth in 'high value' demand.

4.1 Monopoly and Duopoly Market Outcomes

Consider the problem of a monopolist in this market. We assume that the market price cap, a is greater than their marginal cost, c. They will then want to generate at least d MW of electricity, as each successive MW strictly increases profit. Beyond d, the firm faces the problem

$$\max_x \pi^M = (a - x - c)(x + d) \tag{24}$$

Where x is output *in addition to d*. It is easy to show that profits will be maximized by

$$x^M = \tfrac{1}{2}(a - c - d) \tag{25}$$

If $d > a - c$, the firm then wishes to produce no output beyond d as the gain from the marginal output is less than the loss incurred on all infra marginal units. We can then represent the monopolist's supply schedule in terms of d, a, and c:

$$x^M = \begin{cases} 0 & \text{if } d > a - c \\ \frac{1}{2}(a - c - d) & \text{if } d \leq a - c \end{cases} \tag{26}$$

Similarly, we can solve for the price and profits that the monopolist will receive:

$$P^M = \begin{cases} a & \text{if } d > a - c \\ \frac{1}{2}(a + c + d) & \text{if } d \leq a - c \end{cases} \tag{27}$$

$$\pi^M = \begin{cases} (a - c)d & \text{if } d > a - c \\ \frac{1}{4}(a - c + d)^2 & \text{if } d \leq a - c \end{cases} \tag{28}$$

Market equilibrium in the duopoly case is more complex due to the necessity of allocating demand on the perfectly elastic region. We will assume that if market output is greater than d the firms split the region equally. We will focus on the case of symmetric equilibria but depending on the assumptions made others can exist. [16] To find the symmetric equilibrium, we will assume that each firm is 'endowed' with $\frac{1}{2}d$ and then ask to what extent they have an incentive to expand output. Beyond the inelastic region firms maximize profits as usual taking their rival's output as given:

$$\max_{x_i} \pi_i^D = (a - x_i - x_j - c)(x_i + \tfrac{1}{2}d) \tag{29}$$

Which has the interior solution

$$x^D = \tfrac{1}{3}(a - c - \tfrac{1}{2}d) \tag{30}$$

[16] To illustrate the intuition, assume firm B has chosen production equal to $0.75d$. Firm A will want to produce at least $0.25d$. Like in the symmetric case, the desire to expand market output into the downward sloping region of the demand curve will depend on the marginal loss on the $0.25d$ units from a lowered price compared to the gain from the marginal output. If firm A does not find it optimal to increase output then firm B which holds a greater share of d will not either. The greater d is compared to the margin between cost and the price cap the greater the range of asymmetric equilibria that can be sustained.

We can then characterize the outputs, price, and profits of each firm in terms of the model parameters:

$$x^D = \begin{cases} 0 & \text{if } d > 2(a-c) \\ \frac{1}{3}(a+c+\frac{1}{2}d) & \text{if } d \leq 2(a-c) \end{cases} \tag{31}$$

$$P^D = \begin{cases} a & \text{if } d > 2(a-c) \\ \frac{1}{3}(a+2c+d) & \text{if } d \leq 2(a-c) \end{cases} \tag{32}$$

$$\pi^D = \begin{cases} \frac{1}{2}(a-c)d & \text{if } d > 2(a-c) \\ \frac{1}{9}(a-c+d)^2 & \text{if } d \leq 2(a-c) \end{cases} \tag{33}$$

4.2 The Effect of d on Maintenance

We can now evaluate (7) to obtain the effect of additional inelastic demand on maintenance expenditure. There are three cases to consider:

1. High d ($d > 2(a-c)$): Exactly d is produced under monopoly and duopoly.
2. Intermediate d ($2(a-c) \geq d > a-c$)): A monopolist will only produce d, but duopoly output is greater than d.
3. Low d ($d \leq a-c$): Both monopoly and duopoly market structures result in market output greater than d.

4.2.1 Case 1: High d

Here we can evaluate (7) to get

$$\text{sign}\left(\frac{\partial m}{\partial d}\right) = \text{sign}\Big((a-c)((a-c)d - \tfrac{1}{2}(a-c)d) \\ -[(a-c) - \tfrac{1}{2}(a-c)]((a-c)d - \frac{1}{f(m)}\Big) \tag{34}$$

Which immediately simplifies to

$$\text{sign}\left(\frac{\partial m}{\partial d}\right) = \text{sign}\left(\tfrac{1}{2}(a-c)\frac{1}{f(m)}\right) > 0 \tag{35}$$

4.2.2 Case 2: Medium d

Evaluating (7), we obtain the expression

$$\begin{aligned}\text{sign}\left(\frac{\partial m}{\partial d}\right) = \text{sign}\Bigg((a-c)\left[(a-c)d - \frac{(a-c+d)^2}{9}\right] \\ - \left[(a-c) - \frac{2(a-c+d)}{9}\right]\left((a-c)d - \frac{1}{f(m)}\right)\Bigg)\end{aligned} \tag{36}$$

This is strictly positive. To see this, we evaluate (36) at $d = a - c$ to obtain

$$\text{sign}\left(\frac{\partial m}{\partial d}\right) = \text{sign}\left(\tfrac{5}{9}(a-c)\frac{1}{f(m)}\right) \tag{37}$$

Which is positive. We then differentiate (36) with respect to d to get

$$\frac{2}{9}\left((a-c)d - \frac{1}{f(m)}\right) = \frac{2}{9}\left(\pi^M - \frac{1}{f(m)}\right) > 0 \tag{38}$$

By the boundedness of (5). The sign of the effect of d on m in equilibrium is positive at the lowest possible bound, and the equation that defines the sign increases with d, meaning that an increase in inelastic demand increases equilibrium maintenance levels for all of case 2.

4.2.3 Case 3: Low d

Now, evaluating (7) provides

$$\text{sign}\left(\frac{\partial m}{\partial d}\right) = \text{sign}\Bigg(\tfrac{1}{2}(a-c+d)\left[\frac{(a-c+d)^2}{4} - \frac{(a-c+d)^2}{9}\right] \\ - \left[\frac{a-c+d}{2} - \frac{2(a-c+d)}{9}\right]\left(\frac{(a-c+d)^2}{4} - \frac{1}{f(m)}\right)\Bigg) \quad (39)$$

which simplifies to

$$\text{sign}\left(\frac{\partial m}{\partial d}\right) = \text{sign}\left(\tfrac{5}{18}(a-c+d)\frac{1}{f(m)}\right) > 0 \quad (40)$$

We note that in all three cases, an increase in the quantity of electricity that the competitive fringe is unable to supply at any price increases the equilibrium level of maintenance.

4.3 Interpretation and Conclusion

As we know, the incentive to increase maintenance spending depends on the difference between the expected earning when a firm has a functioning generating unit and when it breaks down. In this variation on second stage competition the expected firm profits when its plant breaks down is zero. For it to be optimum to increase maintenance spending, expected profit when the plant remains active must increase. Expected Profit is given by

$$F(m)\pi^D + (1 - F(m))\pi^M \quad (41)$$

More market demand causes profits to increase in all cases. Moreover, because the monopolist captures the entirety of the increase relative to the duopolist the effect is strong enough to overwhelm the simultaneous increase in spending by the rival firm.

One interpretation of an increase in d is an increase in short run market demand.

In the short run, the capacity of the competitive fringe is fixed and any additional demand must be supplied by the strategic firms.

The traditional view is that an increase in market demand reduces overall system reliability. More power must be supplied by the same number of plants; the loss of the same generation brings the system closer to supply shortfall than it did before.

That view however takes the probability of any plant failing as fixed. This analysis suggests that an increase in market demand also increases the incentive to maintain existing infrastructure. In this model the probability of supply shortfall is equivalent to the probability that both plants fail, as a single plant can satisfy market demand (due to non-binding capacity constraints). As a result, we see an overall reduction in the chance of having insufficient generation to meet demand. In the real world capacity constraints often apply, so multiple functioning plants are necessary to serve demand. However, these results suggest that the increased incentive to invest in maintenance may greatly reduce, or even offset completely, the negative reliability effects of increased demand.

5 Conclusion

In the sections above, we have derived a model of plant maintenance, and applied it to study the effects of forward electricity sales and increases in short run electricity demand. We found, in relation to forward sales, that additional commitments tended to increase the equilibrium level of maintenance for relatively low levels of contracts. We also found that an increase in short run demand tended to increase maintenance levels as well.

These results have direct implications for the design of electricity markets. It is well understood that producers selling production forward encourages more vigorous competition in spot markets. These results suggest that the benefits may be more extensive: in addition to lowering prices, contracts encourage additional reliability and more efficient market structures. The results from an increase in short run demand imply that negative reliability effects from increased demand are, perhaps significantly, offset by increased maintenance spending. In light of these results, worries about long term adequacy of supply in various markets may be overblown.

This model is intended to be relatively abstract; as such, there are many possible extensions. Possible variations in market structure include an arbitrary number of firms – including consideration of entry and exit decisions and the potential for maintenance to discourage entry; different forms of competition such as supply function equilibria; [17] and single firms controlling multiple plants, each with the chance of failure. There are other options for modelling failure, including partial failure, the chance of breakdown being correlated across plants, or maintenance downtime as a strategic variable. Very little has been written on the economics of maintenance for electrical generation facilities in a market context, so the scope for new insights is substantial.

Given the importance of electricity and the significant role that system reliability plays in the industry it is both surprising and disappointing that plant maintenance choices have received so little attention. Plant failures can increase price manyfold and bring the system closer to supply shortfall and attendant blackouts. By understanding how firms choose to maintain their plants in accordance with their profit motive we can glean insights into alternative policies to encourage reliable, efficient electricity markets.

[17] In a SFE, firms compete in supply functions, rather than price or quantity. As these are a fair approximation of the bidding procedure in most markets, they have gained some traction in market modelling. However, they are more complicated to work with and tend not to produce unique results. For a more complete discussion of the use of SFE in modelling electricity markets see Ventosa et al. (2005).

References

Alberta Electric System Operator (2013) "Current and Historical Market Reports," URL: http://ets.aeso.ca/.

Alberta Market Surveillance Administrator (2010a) "An Introduction to Albertas Financial Electricity Market," April, URL: http://albertamsa.ca/files/Financial_Electricity_Market.pdf.

——— (2010b) "The Alberta Wholesale Electricity Market," September, URL: http://albertamsa.ca/files/Financial_Electricity_Market.pdf.

Allaz, Blaise and Jean-Luc Vila (1993) "Cournot Competition, Forward Markets and Efficiency," *Journal of Economic Theory*, Vol. 59, pp. 1–16.

Borenstein, Severin, James B. Bushnell, and Frank A. Wolak (2002) "Measuring Market Inefficiencies in California's Restructured Wholesale Electricity Market," *The American Economic Review*, Vol. 92, pp. 1376–1405.

Bushnell, James B., Erin T. Mansur, and Celeste Saravia (2008) "Vertical Arrangements, Market Structure, and Competition: An Analysis of Restructured US Electricity Markets," *The American Economic Review*, Vol. 98, pp. 237–266.

Cardell, Judith B., Carrie Cullen Hitt, and William W. Hogan (1997) "Market power and strategic interaction in electricity networks," *Resource and Energy Economics*, Vol. 19, pp. 109–137.

Deltas, George and Christoforos Hadjicostis (2009) "Plant Reliability in Monopolies and Duopolies: A Comparison of Market Outcomes with Socially Optimal Levels," in James Momoh and Lamine Miliessor eds. *Economic Market Design and Planning for Electric Power Systems*: John Wiley & Sons, Inc. pp. 113–129.

Energy Information Administration (2012) "Negative wholesale electricity prices occur in RTOs," June, URL: http://www.eia.gov/todayinenergy/detail.cfm?id=6730.

Fudenberg, Drew and Jean Tirole (1984) "The Fat-Cat Effect, the Puppy-Dog Ploy, and the Lean and Hungry Look," *The American Economic Review*, Vol. 74, pp. 361–366.

Garcia, Alfredo, Lamine Mili, and James Momoh (2009) "Modeling Electricity Markets: A Brief Introduction," in James Momoh and Lamine Miliessor eds. *Economic Market Design and Planning for Electric Power Systems*: John Wiley & Sons, Inc. pp. 21–44.

Green, Richard (1999) "The Electricity Contract Market in England and Wales," *The Journal of Industrial Economics*, Vol. 47, pp. 107–124.

Hogan, William (1998) "Competitive electricity market design: A wholesale primer.," December, URL: http://www.hks.harvard.edu/fs/whogan/empr1298.pdf.

Hogan, William W. (1997) "A Market Power Model with Strategic Interaction in Electricity Networks," *The Energy Journal*, Vol. Volume18, pp. 107–141.

Hull, John (2010) *Options, Futures and Other Derivatives*: Pearson Education, Limited.

Joskow, Paul L. (2006) "Competitive Electricity Markets and Investment in New Generating Capacity," SSRN Scholarly Paper ID 902005, Social Science Research Network, Rochester, NY.

Kirschen, D.S., G. Strbac, P. Cumperayot, and D. de Paiva Mendes (2000) "Factoring the elasticity of demand in electricity prices," *IEEE Transactions on Power Systems*, Vol. 15, pp. 612–617.

Newbery, David M. (1998) "Competition, Contracts, and Entry in the Electricity Spot Market," *The RAND Journal of Economics*, Vol. 29, pp. 726–749.

Ontario Independent System Operator (2012) "2012 Annual Report," URL: http://www.ieso.ca/imoweb/pubs/corp/IESO_2012AnnualReport.pdf.

Ontario Power Authority (2013) "A Progress Report on Contracted Electricity Supply 2013 Q1," URL: http://www.powerauthority.on.ca/sites/default/files/news/Q1-2013-Progress-Report.pdf.

The Brattle Group (2012) "ERCOT Investment Incentives and Resource Adequacy," URL: http://www.ercot.com/content/news/presentations/2013/Brattle%20ERCOT%20Resource%20Adequacy%20Review%20-%202012-06-01.pdf.

Ventosa, Mariano, Alvaro Baillo, Andres Ramos, and Michel Rivier (2005) "Electricity market modeling trends," *Energy Policy*, Vol. 33, pp. 897–913.