Energy demand change under uncertainty

Matti Liski and Pauli Murto*

September 10, 2009

Abstract

We consider a model of capital replacement under uncertainty of energy costs. Adjustment delays in the replacement of energy-intensive capital follow from two fundamentals in the problem: uncertainty and Ricardian rents in the existing capital structure. They imply a simple dichotomy where short-run output contracts, but the long-run output recovers and increases above the initial output, despite the increasing energy costs. To provide a quantitative assessment of the consumer price increase needed for the replacement, adjustment delays, and policies expediting the change, the model is calibrated using electricity market data. We find that a market-driven large scale entry of green energy requires unprecedented energy cost and consumer price increases mainly due to rents of the existing capital. Subsidies to green energy can greatly benefit the consumer side.

JEL Classification: D00; D40; L00 Q30.

Keywords: energy cost, investment, real options, electricity

* <liski@hse.fi> Helsinki School of Economics, Department of Economics, HECER, and MIT-CEEPR. <murto@hse.fi > Helsinki School of Economics, Department of Economics, and HECER. Funding from the Academy of Finland, Nordic Energy Research Program, and Yrjö Jahnsson Foundation is gratefully acknowledged.

1 Introduction

The replacement of energy-intensive capital plays a central role in efforts to reduce energy consumption. Energy use response to prices is dynamic, and existing approaches to understanding it are based on capital adjustment costs coming either exogenously (Pindyck and Rothemberg, 1983) or from a "putty-clay" structure (Atkeson and Kehoe, 1999).[1] However, it appears to have been overlooked that uncertainty of energy prices has unique implications for the price-induced capital replacement, and yet uncertainty seems an almost defining feature of energy prices. We argue that the uncertainty of conventional energy costs is essential for understanding adjustment delays in the energy use patterns, and the role of policies expediting the change. To provide a quantitative assessment of these effects and policy experiments, we consider an equilibrium model of capital replacement under uncertainty of conventional energy costs and irreversibilities.

The primary energy demand is derived from capital goods used, e.g., in electricity generation, transportation, housing, and manufacturing. As primary energy inputs such as fossil fuels become sufficiently expensive — either due to finiteness of their availability or policies making their use more costly — the value of the capital structure employing these inputs decreases. The old capital is no longer well suited to the new economic conditions. But clearly different capital goods suffer differently when energy costs increase, as the capital goods differ in their efficiency in using the energy input.[2] This "energy quality" heterogeneity between the capital goods leads to the presence of Ricardian rents. Such rents are perhaps most visible in electricity production where the marginal producing unit is usually very sharply identified, leaving a price-cost margin to the remaining producers.[3]

Energy input prices are likely to exhibit a long-run upward trend but they are also extremely volatile, as recent developments in the oil market vividly illustrate. The energy cost volatility creates uncertainty not only about the prospects of the old energy-intensive

[1]Empirical research has found that the energy use is much more responsive to prices in the short run than in the long run (Berndt and Wood 1975, Griffin and Gregory 1976; see also Thompson and Taylor 1995). The putty-clay model can better explain this difference, but there are also other explanations. See Linn (2008) for a discussion, and for a plant-level empirical analysis. For empirical work on innovation induced by energy prices (rather than capital replacement), see Popp (2001, 2002).

[2]This can be due to various combinations of *ex ante* sunk costs and *ex post* variable costs such that the firms were indifferent between the combinations when they entered the industry in the past. See, e.g., Roques et. al. (2006) for analysis of the choice between nuclear power and gas technologies.

[3]It is a standard practice in electricity market studies to evaluate these rents to isolate them from rents arising from market power, for example. See Wolfram (2000) and Borenstein et al. (2002).

capital goods but also about the profitability of the new energy-saving capital goods — the social value of the replacement depends crucially on the expected value of the capital replaced. The green energy investments obviously face multiple uncertainties, but the future cost of conventional energy seems the most fundamental uncertainty for the economists to include in the analysis of the energy demand change.

Our analysis builds on the above two elements, old capital rents and uncertainty, and produces a pattern that, in spirit, is similar with that implied by the putty-clay model: short-run output contracts as a response to energy cost increases, but in the long-run the output recovers while simultaneously the energy cost keeps on increasing. In contrast with any previous result known to us, the long-run output expands even beyond that prevailing before the capital replacement started, despite the fact that the only exogenous change is the increasing energy cost.

The output contraction-expansion pattern follows from the existence of Ricardian rents and uncertainty. Rents reflect heterogeneity in the social value of the energy-intensive structure and, therefore, it will be socially optimal to replace the energy-intensive capital gradually as the energy cost increases. However, because conventional energy costs are uncertain, the social value of each replacement is also uncertain, generating equilibrium real options for green technology entrants and thus a separation of marginal costs and prices (see, e.g., Dixit and Pindyck 1994). Our benchmark description involves no distortions, so the equilibrium achieves the social first best.

The output contraction-expansion pattern can be understood in terms of the two rents, the Ricardian rents of the existing capital and the price-cost mark-up of the energy-saving capital. The short-run output contracts in order to create a mark-up for early green entrants. The mark-up must exist to compensate for the downside risk that the conventional energy costs decline in the near future — the replacement is socially wasteful *ex post* if, e.g., the oil price sufficiently declines, or the externality costs of fossil fuels diminish. The greater are the Ricardian rents, the more valuable is the existing structure from the social point of view, and the larger is the output contraction needed for the capital replacement to take place. In contrast, the long-run output expands because, at sufficiently high energy cost levels, there will be less uncertainty about the value of the energy-intensive capital. The decline in uncertainty reduces mark-ups and boosts investment, leading to the recovery of output. Because the needed mark-up is largest for the early entrants while being smallest for the late entrants, the overall output expands beyond that prevailing initially.

Our model is a simple supply and demand framework that is conceptually an application of Leahy (1993), but the substantial implications cannot be seen from the Leahy's model. We believe the framework is well suited for gauging the consumer price increase needed for the capital replacement to take place in particular industries. To illustrate this, we calibrate the model using data from the Nordic electricity market.[4] We estimate the key elements of the model from data, and simulate the replacement of the fossil-fuel intensive generation under various scenarios for the costs of the new technologies. We find that the uncertainty of conventional energy costs are alone a significant source of investor caution: even under the most optimistic scenarios, the price-cost margin exceeds investors' costs by multiple factors during the transition. As a result of this inertia, the transition in the electricity sector is likely to be very costly to consumers. Green energy subsidies can be extremely beneficial to consumers, even when they are distorting the overall welfare: the cost of the intervention falls to a large extent on the old capital rents. Our quantitative assessment suggests that subsidies can considerably expedite the transition, and increase the consumer welfare, even without externalities justifying the need to expedite the phase out of energy-intensive capital.

In Section 2, we describe the basic model, and in Section 3, we develop the simple analytics of the model and use figures to explain the basic mechanism. In Section 4, we describe the equilibrium more generally and connect the equilibrium to Leahy's results. In Section 5, we calibrate the model using data from the Nordic electricity market, and perform the policy experiments.

2 The model

We describe the change of the energy-input demand in a simple final-good demand and supply framework. Denote the inverse demand of the final good by $p = D(q)$, where p is the price and q is the final good demand, and assume that the function is monotone and non-increasing. An example of the good produced is electricity, a case that we consider in our application, but we do not want to limit ourselves to this interpretation. Nothing in the structure of the model prevents thinking of any final good market whose supply side uses energy-inputs, or alternatively, energy-saving technologies.[5]

There are two basic sources of supply, namely the old energy using technology and

[4]The approach is general and can be applied to other electricity markets as well.

[5]We can thus think of a market where primary energy (crude oil, natural gas, coal) is used to produce secondary energy (electricity, gas, refined petrolium), or the output can be the final consumption good.

the new input-free technology.[6] We can think of a continuum of old and new technology firms, each producing one unit of output. An old firm uses one unit of energy input to produce one unit of output, and old firms differ in their efficiency in using the input. Let x denote the price of a unit of energy, and let q^f denote the total final-good supply coming from the energy-using firms. We assume that the marginal cost of the last producing unit, denoted by $MC(q^f, x)$, is strictly increasing and differentiable in q^f for a given energy price x. We also assume that the marginal cost is strictly increasing in x for all $q^f > 0$.

The new technology firms produce the same output but use no energy input. We denote the number of these firms by k, i.e., k is the existing energy-saving capital stock. We will introduce the entry problem of a new firm shortly, and take for a moment the existing k as given. Since we are interested in describing a situation where the new supply from k replaces the old supply structure, we can set its variable cost to zero; the exact level of this cost does not matter as long as the new capital is the least-cost option, once in place.[7] The combined total inverse supply can then be written as

$$S(q,k,x) = \begin{cases} 0 \text{ if } q \leqslant k \\ MC(q-k,x) \text{ otherwise.} \end{cases} \tag{1}$$

We will introduce a capital structure that supports the old supply curve in a later section. For now, we just assume a static supply curve captured by $MC(q^f, x)$.[8] The usage of old structure, i.e., production $q^f = q - k$, is what clears the final-good market,

$$p = D(q) = S(q,k,x) > 0, \tag{2}$$

for a given k and x. [9]

Let us introduce time and uncertainty into the analysis. We assume that time is continuous and that the energy cost x_t is the only source of uncertainty over time. We

[6]The new technology may still use energy but this energy is not coming from the fossil-fuel inputs. It may also be the case the new technology saves primary energy in absolute terms. Both interpretations are consistent with the model, and we use interchangeably the wordings "energy-input saving" and "energy-saving" technology

[7]Under this assumption, any positive flow cost can eliminated and incorporated into the initial investment cost of a new entrant.

[8]In Section 4.2 , we will show that the results are not substantially altered, if the the old capital has a choice between idleness and exit at any time period. In this section, without exit, the old capital replacement has the meaning that the increase in k forces part of the old structure to idleness.

[9]In equilibrium, where the amount of new capital k is determined endogenously, the price will remain positive, i.e., the lower bound for prices is $D(k) > 0$.

assume that the energy cost follows a general diffusion process of the form

$$dx = \alpha(x)dt + \sigma(x)dw, \tag{3}$$

where w is a Wiener process.[10] This formulation admits the commonly used specifications used in the analysis of irreversible investments under uncertainty. In particular, if $\alpha(x) = \alpha x$ and $\sigma(x) = \sigma x$, then the process is a Geometric Brownian Motion (GBM). We do not have to be specific about the process, as long as it is defined for arbitrarily high levels of x to induce the transition to the new technology. If there is not much uncertainty in the process (σ is close to zero), we assume a positive trend ensuring the high prices in the end (α is strictly positive). The assumptions on x_t introduce persistence into the fuel prices. If, for example, there is no clear trend in prices but there is significant volatility we capture the idea that "to predict the price of oil one quarter, one year, or one decade ahead, it is not at all naive to offer as a forecast whatever the price currently happens to be" (Hamilton, 2008).

Since the old fuel-using supply clears the output market for a given k, the stochastic fuel price implies that the output price is a stochastic process too, see equation (2). Because each entering new capital unit supplies one unit of output, the output price process is the revenue process for entrants. The new capital units thus make irreversible entry decisions under uncertainty. We assume that there is a continuum of potential entrants who can each invest in one capital unit by paying an irreversible upfront investment cost $I > 0$. The investors are risk neutral and face a constant market interest rate $r > 0$. Once in place, the new capital unit lives forever.

3 The simple analytics

We provide first the simple analytics of the model under the following assumptions. First, we assume that the variable x_t follows GBM with drift $\alpha > 0$ and standard deviation $\sigma \geqslant 0$,

$$dx_t = \alpha x_t dt + \sigma x_t dz_t. \tag{4}$$

Second, the final-good demand is linear

$$p_t = A - Bq_t, \tag{5}$$

[10]We assume that functions $\alpha(x)$ and $\sigma(x)$ satisfy standard requirements for the solution to exist. See, e.g., Leahy 1993.

and, third, the marginal cost for the old supply is additive in x_t and linear in $q_t^f > 0$,

$$MC(q_t^f, x) = x_t + Cq_t^f, \tag{6}$$

where the positive constants A, B and C satisfy the assumptions outlined above.

3.1 Transition without uncertainty

Let us first describe the build-up of the new energy-saving capital by eliminating uncertainty and assuming that the energy cost is on a deterministic upward trend, shifting gradually the old supply curve upwards. That is, set $\alpha > 0$ and $\sigma = 0$ in (4). We can now explain the distinct role of Ricardian rents as the source of gradualism in the transition.

In Figure 1, x enters as the intercept of the old supply curve — x can be thought of as the direct purchase cost of the fuel. The producer surplus (shaded area) illustrates the presence of Ricardian rents. The entry cost of one new capital unit is rI, expressed as a flow cost. When the energy cost is sufficiently low so that the output price satisfies $p < rI$, the new technology units cannot enter, and the old structure satisfies the full demand. But since the energy price is on an upward trend, the output price must meet rI at some point. The first new capital then unit enters the market, as its present-value revenue p/r covers the investment cost I. This is the situation depicted in Figure 1.

As the energy cost keeps on shifting the supply curve up, there is a tendency for the output price to increase. But because of free entry, the consumer price cannot exceed $p = rI$, the entry cost of alternative supply. In Figure 2 we depict a situation where the fuel price has reached x_t, and there are k_t units of new capital in place. Recall that the new supply is the least-cost option, once in place, so the inverse supply is zero up to k_t, and then increasing for $q \geqslant k$, as depicted. The input-saving new technology has reduced the rents of the old supply structure when compared to the initial situation. Because these rents are sandwiched between the constant final-good price $p = rI$ and the increasing energy cost, they will vanish altogether at the moment the fuel cost meets the price $p = rI$. From this point onwards, the new technology serves the market alone.

3.2 Output Contraction

We assume now that there is volatility in the energy cost process, i.e., $\sigma > 0$ in (4). We look for the conditions under which the new technology starts to enter the market.

Recall that there is persistence in the fuel price process: both input and output prices are expected to remain high longer, the higher is x_t. The current level of x_t measures

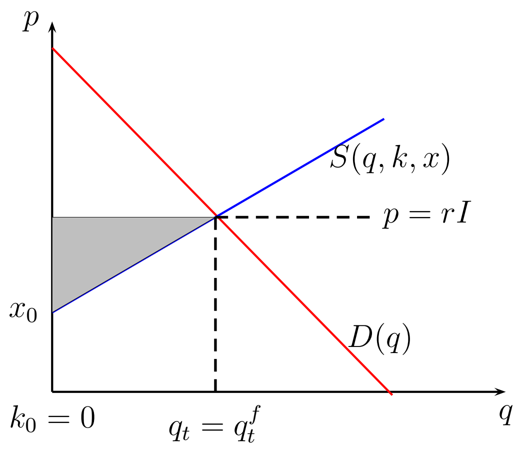


Figure 1: Entry of the first energy-saving capital unit under certainty.

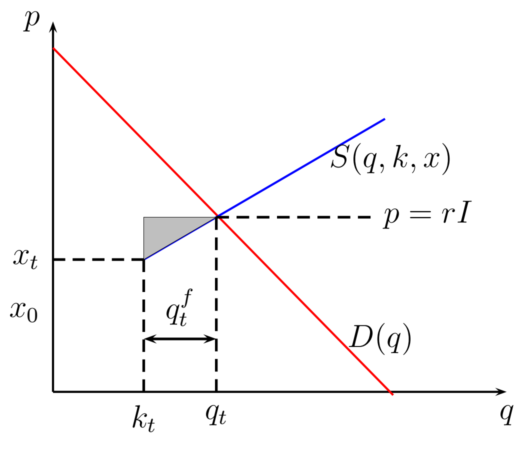


Figure 2: Entry of additional energy saving capital at k_t under certainty.

the profitability of entry directly, and it makes sense to enter only when x_t reaches new record levels. Let $\hat{x}_t$ to denote the highest energy price level seen by time point t, so that new entry will take place only if the fuel price process beats this record — previous entry considerations were made under fuel prices (weakly) lower than $\hat{x}_t$, so, all else remaining unchanged, additional entry requires a higher energy price. As a normalization, we denote the time where the first new firm enters by $t = 0$.

In Figure 3, the first new firm enters when the energy input price level reaches $\hat{x}_0$. The output price at the moment of entry is $P_H(\hat{x}_0)$, denoted this way to emphasize that it is the highest output price observed so far. We have drawn the Figure such that the new firm earns a mark-up above its entry costs, $P_H(\hat{x}_0) > rI$. This must hold because the new firm faces the downside risk that the fuel price and thus the final-good price starts to decrease after the entry; the lowest price conceivable is $P_L(\hat{x}_0)$.[11] For the mark-up to arise, the output must contract.

From the real options theory (Dixit and Pindyck, 1994), we know that the first entering unit requires a mark-up above its deterministic entry cost, $P_H(\hat{x}_0) > rI$, when facing the above described uncertainty.[12] Consider now how this mark-up develops in

[11] We denote the lowest price this way to emphasize that it depends on the state of the record $\hat{x}_0$.

[12] For a moment, we put aside the issues of how this firm conceives the future market development when more firms enter to the market. We come back to this issue in the formal characterization of the equilibrium. We can think that the first entrant mistakenly believes that it will be the only new firm ever entering the market. In fact, Leahy (1993) shows that the entry considerations can be correctly described this way. Based on this, it is clear from standard real options arguments that the first entrant requires the above mark-up, when there is uncertainty.

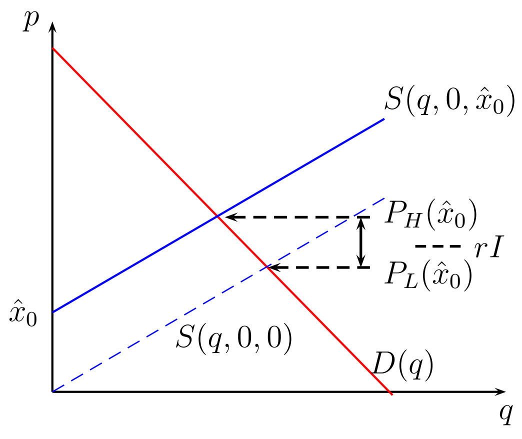


Figure 3: Entry of the first energy-saving capital unit under uncertainty.

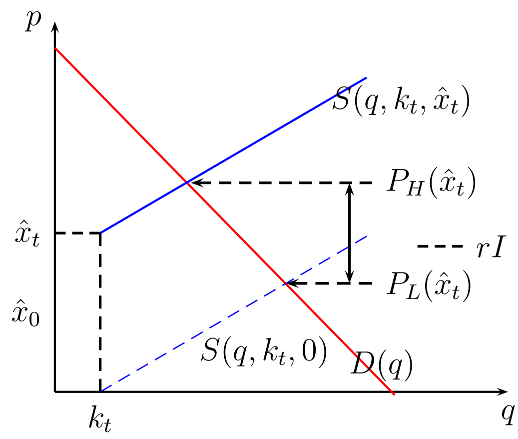


Figure 4: Entry of additional energy saving capital at k_t under uncertainty.

equilibrium when the energy demand continues to change. We argue that the mark-up and thus consumer price must increase as more energy-saving capital enters the market.

Figure 4 depicts time point $t > 0$ where there are k_t new units in place. We have drawn Figure 4 in the way that the output price is higher than the initial price at which the first new capital unit entered: $P_H(\hat{x}_t) > P_H(\hat{x}_0)$. This must hold, because the new entrant faces a higher risk of lower output prices when there is some new capital already in place — the overall supply capacity has increased while the process for x and the old technology supply curve remain the same.[13] As a result, the entry price must be higher and the output lower, to compensate for the increased downside risk. We see therefore that the energy price increase induces more entry but also higher consumer price levels, even though the substitute cost remains unaltered.

The consumer price increase follows from the combination of two elements in the model. First, the Ricardian rents ensure that the old production structure is replaced only gradually. Second, since the old structure remains in the market, it can benefit from the potential downside development in the fuel market, making the output market potentially extremely competitive.[14] This downside risk for the new entrant implies that

[13]That is, the output price process becomes less favorable to the entrant at each given x. This argument will be made formal in the next section.

[14]One may ask if this property arises only because the old structure remains existing by assumption. We show in Section 4.2 that this is not case by assuming that the old structure decides also when to exit the market. In fact, we first developed the model for this case. Since that framework is considerably more complicated and the substantial results are the same, we can without loss of generality build up on the insights provided by this simpler model.

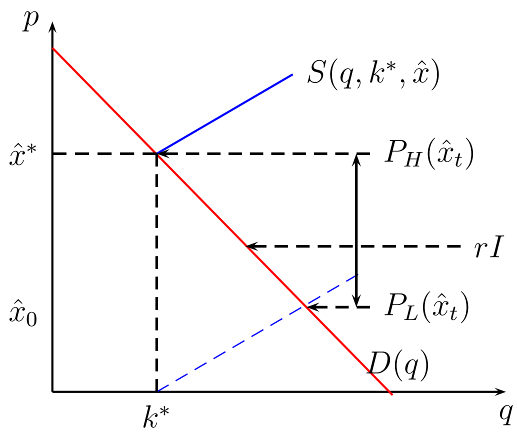


Figure 5: The consumer price peak

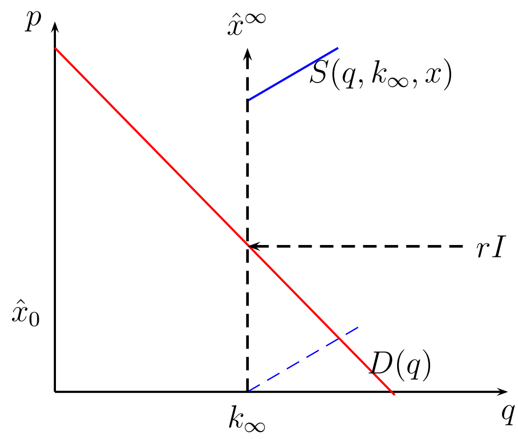


Figure 6: The long-run equilibrium

the input-saving substitute can only enter when the not only the energy but also the consumer price reaches record high levels. We see that the uncertainty protects the rents of the old supply structure, as more extreme energy prices are needed to trigger entry than without uncertainty.

3.3 Output expansion

We have just demonstrated that the higher input price induces more entry but also higher output prices and lower output. We show now that the consumer price reaches a peak during the transition, after which the output recovers even though the energy prices increase. The final consumer prices will be lower and the output higher than the initial prices at which the transition started.

Consider now an input price so high that the entire old structure is just idle, i.e., the most efficient old unit is indifferent between idleness and production. We denote this input price by $\hat{x}^*$, and the corresponding new capital that serves the entire demand at that point by k^*. See Figure 5 for this situation. The market environment cannot become more risky for a new entrant than the situation described here, i.e., the mark-up above costs reaches its peak, $P_H(\hat{x}^*) - rI > P_H(\hat{x}_t) - rI$ for all $\hat{x}_t$.

Note that for the capital to increase above k^*, the energy price must reach values higher than $\hat{x}^*$. But then the old structure is not only idle at input prices $\hat{x}_t > \hat{x}^*$ but is also expected to remain idle in the near future; the input price must decline by the discrete amount $\hat{x}_t - \hat{x}^* > 0$ before the old structure can consider producing again. In this sense, the output price is expected to remain isolated from the input price uncertainty in the near future. For this precise reason, the new technology's prospects improve, and therefore it requires lower equilibrium entry prices. That is, the entry price declines in $\hat{x}_t$ after peaking at $P_H(\hat{x}^*)$.

As sufficiently high $\hat{x}_t$ values are reached, the old structure produces with probability zero in the relevant future, implying that the entry becomes practically free of risk, and the consumer price approaches the deterministic entry cost, with which we started the analysis. The output has now fully recovered, and is larger than the initial output both at the entry point and at all other states.

Let us pull together this description more formally in the following Proposition.

Proposition 1 *There exists $\sigma^* > 0$ such that for $0 < \sigma < \sigma^*$, the equilibrium output contracts at investment points $0 < \hat{x} \leqslant \hat{x}^*$, and expands for $\hat{x} > \hat{x}^*$. Furthermore:*

- *peak price $P_H(\hat{x}^*)$ increases in C and σ*
- *replacement becomes instantaneous as $C \to 0$*
- *investor mark-up disappears in the long run, $P_H(\hat{x}_t) \to rI$ as $\hat{x} \to \infty$.*

Proof. See Appendix. ■

The proof in the Appendix (supplementary material) is based on a connection to Leahy (1993) that we explain in detail in the next Section. The upper bound on uncertainty measured by σ ensures that entry starts before the market shuts down, i.e., without this assumption the market would contract extremely before any new entry.[15]

The Proposition also shows how the resistance to entry depends on the Ricardian rents, measured by the slope of the supply curve, C. The greater is the slope C, the lower is the responsiveness of entry to energy prices, or the more protected are the old units from entry. On the other hand, when C is close to zero, the Ricardian rents are absent, and that there is a large one-time replacement of old units as soon as the output price reaches a certain treshold level. Similarly, the energy cost uncertainty, captured by σ provides protection to the old supply structure.

For illustration, see Figures 7-9 depicting equilibrium paths based on the specification (4)-(6). In Figure 7, we show a sample path for the energy price (solid line) and the historical maximum (dotted line). Figure 8 depicts the path for the energy saving capital. Figure 9 depicts the output price path, declining to the deterministic investment cost towards the end of the path. Note that the overall volatility of consumer prices increases temporarily during the transition because (i) for geometric Brownian motion larger x

[15] In this case, the peak output price would be the first entry price, and $\hat{x}^* = 0$.

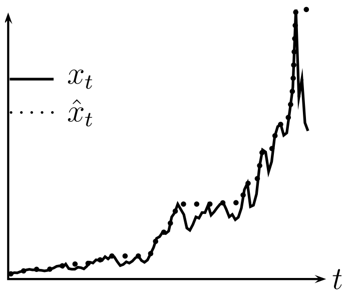


Figure 7: Sample path for the energy price

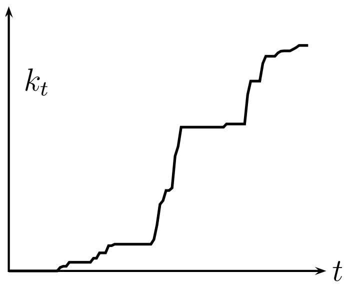


Figure 8: The energy saving capital stock

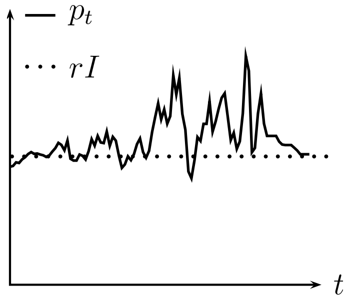


Figure 9: The consumer price path

means higher absolute volatility for x, and (ii) the domain for conceivable consumer prices increases for reasons explained above.

4 General analysis

4.1 Equilibrium

In this section we define formally the industry equilibrium, and show how it can be computed. We follow closely Leahy (1993). There is a continuum of identical potential entrant firms, and each may at any time enter by installing an infinitesimal capacity addition dk at cost Idk. Let k_t denote the aggregate capacity level at time t, and let $\{k_t\}$ denote the capacity path, i.e., the stochastic process governing its evolution in time. New entries increase k_t, and since there is no exit, $\{k_t\}$ must be an increasing process. To find an equilibrium, we must specify an entry strategy profile for the potential entrants and a corresponding capacity path such that (i) given the capacity path, the entry profile is optimal for each individual firm, and (ii) the entry profile induces the capacity path.

The profit flow to a holder of a capacity unit is given by the ouput price. In Section 2 output price was defined by equation (2), but the results hold generally for an output price function $p = P(k_t, x_t)$, where we assume that $P(k, x)$ is continuous in k and x, increasing in x, and decreasing in k. In addition, to ensure that profit for a unit of capacity is always finite, we assume that for any fixed value of k,

$$E \int_0^\infty P(k, x_\tau) e^{-r\tau} d\tau < \infty.$$

The information upon which the entering firms base their behavior at period t consists of the historical development of x_t and k_t up to time t. However, since $\{x_t\}$ is a Markov process, the state of the economy at any point in time is fully summarized by the current values (x_t, k_t). It is therefore natural to restrict to Markovian strategies. Moreover, in the current context we may restrict further to strategies that can be expressed in cut-off form.

Definition 1 *A Markovian cut-off strategy is a mapping*

$$x^* : [0, \infty) \to \mathbb{R} \cup \infty \cup -\infty,$$

where $x^*(k)$ gives the lowest level for the shock variable at which the firm is willing to enter, given capacity k.

We use $x^*(k) = \infty$ to indicate that the firm does not enter at any level of x_t, and $x^*(k) = -\infty$ to indicate that the firm enters immediately for any value of x_t.

Remark 1 *Leahy (1993) expresses strategies as a cut-off level for the ouput price. If $P(k, x)$ is strictly increasing in x (which is assumed by Leahy), this is equivalent to our formulation: instead of $x^*(k)$, one could just as well use a strategy $p^*(k) = P(k, x^*(k))$, which defines the cut-off price at capacity k that triggers new entry. We express strategies in terms of x, because we do not require $P(k, x)$ to be strictly increasing. Despite this less demanding requirement for $P(k, x)$, the main results of Leahy (1993) hold in our context with some notational modifications.*

We will see that there is a symmetric equilibrium, where all the firms adopt such a cut-off strategy. To formalize this, we must derive the capacity path that such a profile induces. Let us assume an arbitrary symmetric cut-off profile x^*. There is a large number of potential entrants, each with a strategy commanding them to enter as soon as x_t hits $x^*(k)$. Of course, it would make no sense to assume that all the firms actually enter at the same time. Instead, as soon as the entry threshold is hit, capacity k_t will immediately increase up to the point where entry stops. This happens every time x_t hits the relevant cut-off level $x^*(k_t)$, and consequently, we end up with a capacity path along which entry takes place only at such time moments where x_t hits new record-values. Note that we are not interested in the identities of actual entrants, but the aggregate capacity path.

Let us denote by $\hat{x}_t$ the historical maximum value of x_t up to time t:

$$\hat{x}_t \equiv \sup_{\tau \leq t} \{x_\tau\}. \tag{7}$$

We can now formalize the above discussion by defining the aggregate capacity path as a function of $\hat{x}_t$.

Definition 2 *The capacity path induced by a symmetric cut-off strategy x^* is the stochastic process $\{k_t\} \equiv \{\mathbf{k}^*(\hat{x}_t; x^*)\}$, where*

$$\mathbf{k}^*(\hat{x}_t; x^*) = \inf\{k \geq 0 \,|x^*(k) > \hat{x}_t \;\}. \tag{8}$$

Note that $\{\mathbf{k}^*(\hat{x}_t; x^*)\}$ is an increasing stochastic process, and its value at time t is fully specified by the development of x_t up to time t.[16]

Equation (8) together with (3) defines the law of motion of k for a given symmetric entry strategy x^*. To find an equilibrium, we must also check the optimality of the entry strategy agains a given capacity path. The entry problem of an individual firm can be written as follows. Let $\mathbf{k} : [x_0, \infty) \to \mathbb{R}_+$ denote an arbitrary increasing function that defines the aggregate capacity as a function of the historical maximum value of x_t. A potential entrant is effectively holding an option to install one capacity unit at cost I, so the entrant solves the following stopping problem:

$$F(x_t, \hat{x}_t; \mathbf{k}) = \sup_{\tau^* \geq t} E\left[\int_{\tau^*}^{\infty} P(\mathbf{k}(\hat{x}_\tau), x_\tau) e^{-r(\tau - t)} d\tau - I e^{-r(\tau^* - t)}\right], \tag{9}$$

where $F(\cdot)$ is the value of the option to enter.

The potential entrants are all alike and solve the same entry problem, but in equilibrium with unrestricted entry each entrant must remain indifferent between entering and staying out. Following Leahy (1993), we now define formally the competitive equilibrium as a rational expectations Nash equilibrium in entry strategies. Consider a symmetric candidate profile x^* and the induced capacity process $\{k_t\} \equiv \{\mathbf{k}(\hat{x}_t; x^*)\}$. We need two conditions. First, free entry eliminates any profits to the potential entrants. That is, for all x_t and $\hat{x}_t$, we have

$$F(x_t, \hat{x}_t; \mathbf{k}) = 0. \tag{10}$$

Second, whenever x_t hits $x^*(\mathbf{k}(\hat{x}_t; x^*))$, entrants must find it (weakly) optimal to enter, otherwise they would rather stay idle:

$$E\left[\int_{t}^{\infty} P(\mathbf{k}(\hat{x}_\tau), x_\tau) e^{-r(\tau - t)} d\tau\right] - I = 0. \tag{11}$$

Definition 3 *The industry equilibrium is a trigger strategy profile x^* and corresponding capacity path $\{k_t\} = \{\mathbf{k}^*(\hat{x}_t; x^*)\}$ such that*

- *(10) holds for all $\hat{x}_t$ and $x_t \leq \hat{x}_t$ when $\mathbf{k} = \mathbf{k}^*(\hat{x}_t; x^*)$*

[16] That is, $\{k_t\}$ is a stochastic process adapted to the filtration $\{\mathcal{F}_t\}$.

- *(11) holds whenever* $x_t = x^*\left(\mathbf{k}^*\left(\hat{x}_t; x^*\right)\right)$

- $\mathbf{k}^*\left(\hat{x}_t; x^*\right)$ *is given by (8)*

The key to finding such an equilibrium is the observation that a marginal firm which understands the stochastic process $\{x_t\}$ but disregards the other firms' entry decisions will choose the same entry time as a firm that optimizes against the equilibrium capacity path $\mathbf{k}^*$. This myopia result, due to Leahy (1993), can be formalized as follows. An entering firm that thinks the current capacity $\mathbf{k}^*\left(\hat{x}_t; x^*\right)) = k_t$ remains unchanged in the future solves the exit time from

$$F^m\left(x_t; k\right) = \sup_{\tau^m \geq t} E\left[\int_{\tau^m}^{\infty} P(k, x_\tau) e^{-r(\tau - t)} d\tau - I e^{-r(\tau^m - t)}\right]. \tag{12}$$

Note first that the solution to (12) can be expressed as a cut-off rule.

Lemma 1 *The optimal solution to (12) can be expressed as a cut-off rule* $x^m\left(k\right)$*, so that the optimal stopping time* τ^m *is the first moment when* x *hits* $x^m\left(k\right)$ *from below:*

$$\tau^m = \inf\left\{\tau^m \geq t\, | x_{\tau^m} \geq x^m\left(k\right)\right\}.$$

Proof. The problem (12) is a standard excercise problem of a perpertual call option, where the value of the underlying asset at time t is given by

$$V^m\left(x_t, k\right) = \int_t^{\infty} P(k, x_\tau) e^{-r(\tau - t)} d\tau,$$

and the cost of excercise is constant I. By assumption, $P\left(k, x\right)$ is increasing in x, so under our assumptions on $\{x_t\}$, V^m is also increasing in x. It is then clear that if excercising at x' is optimal, it must be optimal to excercise also at any $x'' > x'$. Conversely, if it is not optimal to excercise at x', it is neither optimal to excercise at $x'' < x'$. Thus, the solution is a cut-off rule. ■

The following proposition, based on Leahy (1993), states that the model has an equilibrium that can be computed by solving the myopic problem (12) for all fixed values of k.

Proposition 2 *Under the assumptions stated, the model has an industry equilibrium, where the entry threshold* $x^*\left(k\right)$ *is given by the solution to (12). The corresponding capacity path is given by (8). The entry threshold* $x^*\left(k\right)$ *is increasing in* k*.*

Proof. By Lemma 1, the solution to (12) for any $k \geq 0$ is a cutoff policy, which we can denote $x^*(k)$. Since $P(k,x)$ is increasing in x and decreasing in k, it is a standard comparative static property of this type of a problem that $x^*(k)$ is increasing in k. The proof that the solution to (12) constitutes a competitive equilibrium can be constracted following the steps given in Leahy (1993). The only difference is that our assumptions on $P(k,x)$ are slightly less demanding than similar assumptions in Leahy (1993), but those differences are not crucial for this result. ■

4.2 Extension

One may argue that our description depends on the specific assumption that the old production structure is static and cannot respond by exit decisions to the changing market situation. This is not at all the case. In our working paper Liski and Murto (2006), we analyze a considerably more complicated model, where the substance-related results are essentially the same. In that framework, we assume that there is a continuum of infinitesimal firms, and each active firm has one unit of capital of either type. If we let k_t^f and k_t^b denote the respective total fuel-dependent and fuel-free capacities at time t, then k_t^f and k_t^b denote also the numbers of firms at t. By k_0^f and k_0^b we refer to exogenously given initial capacity levels. Each factor-dependent firm that is still in the industry at some given t must choose one of the following options: produce, remain idle, or exit. To make the choice between idleness and exit interesting, we assume that staying in the industry implies an unavoidable cost per period. Let $c > 0$ denote this fixed flow cost. A producing unit in period t thus incurs cost $c + x_t$, where x_t is the factor price. An idle unit pays just c. An exiting unit pays a one-time cost $I_f > 0$ and, of course, avoids any future costs. Note that, in equilibrium, firms (discrete) choices between production and idleness determine the overall utilization of the old capacity. Let q_t^f denote the total output from the factor dependent capacity. Then, q_t^f is also the number of producing firms which satisfies $q_t^f = k^f$ if all remaining firms produce, and $0 \leq q_t^f < k^f$ if utilization is adjusted.

Assume

$$I_f < \frac{c}{r} < I_f + I_b.$$

The first inequality implies that exit saves on unavoidable costs for an old capacity unit. The second inequality implies that replacing an old unit by a new unit is costly. Without the former restriction, old plants would never exit. Without the latter, the factor-dependent capacity would be scrapped and new capacity built immediately.

For this structure, the capital replacement path is a pair $\mathbf{k}(\hat{x}) = \left(\mathbf{k}^f(\hat{x}), \mathbf{k}^b(\hat{x})\right)$ with the following properties.

Proposition 3 *The model with two capital stocks has an equilibrium with the following properties:*

- $\mathbf{k}^f$ *is everywhere continuous, strictly decreasing on some interval* $(a_f, b) \subset \mathbb{R}^+$*, and constant on* $\mathbb{R}^+ \setminus (a_f, b)$.
- $\mathbf{k}^b$ *is everywhere continuous, strictly increasing on some interval* $(a_b, b) \subset \mathbb{R}^+$*, and constant on* $\mathbb{R}^+ \setminus (a_b, b)$.
- *total capacity* $\mathbf{k}^b + \mathbf{k}^f$ *increases on* (a, b) *where* $a = min\{a_b, a_f\}$.

Proof. See Liski and Murto 2006. ■

Before turning back to our current framework, let us note two basic implications of the result. The exit of the old technology may start before or after the entry of the new one (i.e., $a_f \neq a_b$), but both transitions end at the same factor market condition, $\hat{x} = b$. The result implies that as long as the transition is going on for both technologies, there is both exit and entry every time $\hat{x}$ reaches a new record value. Also in this framework the transition is gradual because of Ricardian rents. Finally, the consumer price must increase and the output contract during the transition because there is technology overlap: the new technology does not replace the old one-to-one, but is built to co-exist with the old structure during the transition (the third result in the Proposition). This feature is the key for the result that increasing consumer prices are needed for the transition to take place. In our current simpler model, the technology overlap is extreme as the the new input-saving technology is built to coexist with the old structure forever.

5 Application to the Nordic electricity market

We now provide a quantitative assessment of the mark-ups needed for green electricity entry using electricity market data. Electricity generation uses primary energy (e.g., fossil fuels) to produce secondary energy (electricity), with long-lived capacity and relatively clear green electricity options such as those relying on wind and renewable energy sources. The data for the assessment comes from the Nordic electricity market, but the procedure is not specific to the Nordic market.[17] We believe that the main lessons apply to electricity

[17] The Nordic case has the advantage that the data needed is publicly available, and that the relevant supply curve can be estimated without using engineering approach on plant-level cost characteristics.

markets in general because the key properties of electricity generation are common across markets. The capital replacement is of particular importance in this sector: for example in the US, the electricity sector uses 42 percent of primary energy, 34 percent of fossil fuels, and produces about 40 percent of CO2 emissions (Joskow, 2008).

5.1 Institutions

The Nordic wholesale power market developed to its current form through a series of steps when the four continental Nordic countries (Finland, Denmark, Norway, Sweden) underwent electricity market liberalization at different times in the 1990's. Today, it is perhaps the most tightly integrated cross-border wholesale electricity market in Europe, serving majority of the ca. 400 Twh annual demand in the Nordic region.

Wholesale electricity trade is organized through a common pool, Nord Pool, which is a power exchange owned by the national transmission system operators.[18] Market participants submit quantity-price schedules to the day-ahead hourly market (Elspot market). The demand and supply bids are aggregated, and the hourly clearing price is called the system price. The Nordic market uses a zonal pricing system, in which the market is divided into separate price areas. If the delivery commitments at the system price lead to transmission congestion, separate price areas are established. For our study, the price areas are not important since we aggregate prices to the weekly level, and from these we construct annual revenues for new entrants to this market. At this level of aggregation, there is no loss of generality from working with the system price.[19]

When estimating the supply, we focus on period 2000-05 because the institutional and economic environment was relatively stable; that is, the market was not yet affected by the European emissions trading scheme and further integration to the continental Europe.[20]

[18]For more information about the pool, see www.nordpool.com. For a succint description of the Nordic market, see Amundsen and Bergman (2006).

[19]The direction of congestion in the transmission links varies within the year, and also between the years depending on the division of labor between hydro-intensive and thermal-intensive regions in the market. The deviations from the system price tend to cancel out over time. See Juselius and Stenbacka (2008) for a study focusing on the degree of integration of the Nordic price areas at the hourly level.

[20]To recap the market development, we may call the years 2000-01 as years of abundant availability of hydroelectricity which led to low prices during these years. The year 2002 in turn was exceptional: the Fall rainfall and thus inflow was scant and the stocks were drawn down to approach historical minimums by the turn of the year. The price spike resulted, and it took almost two years for the stocks to recover. See Kauppi and Liski (2008) detailed explanation and analysis of the price spike.

Roughly one half of annual Nordic generation is produced by hydro plants. In 2000-05, 61 per cent of hydroelectricity was generated in Norway and 33 per cent in Sweden. Sweden is the largest producer of thermoelectricity with a share of 46 per cent of annual Nordic mean production, followed by Finland and Denmark, with shares of 35 and 19 per cent, respectively. Hydro availability is the one single market fundamental that causes significant swings in demand for other production technologies. These swings are exploited in our estimation of the non-hydro supply curve for this market.

In the Nordic area, the non-hydro production capacity consists of nuclear, thermal (coal-, gas-, biofuel-, waste- and oil-fired plants), and wind power. An important part of thermal capacity is combined heat and power (CHP) plants which primarily serve local demand for heating but also generate power for industrial processes and very cost-efficient electricity as a side product. An implication of CHP capacity is that the non-hydro market supply experiences temperature-related seasonal shifts, which we also seek to capture in our estimation procedure detailed later. Table 1 provides a breakdown of capacity forms over the period 2000-2005. At the market level, there is thus a rich portfolio of capacities with large number of plants in each category determining a relatively smooth aggregate supply function.

	TWh			
	Denmark	Finland	Norway	Sweden
Total generation	37.3	73.4	125.2	146.5
Hydro	.0	12.7	124.1	67.8
Other renewable	5.8	2.0	.3	1.9
Thermal	31.5	58.8	.8	76.7
-Nuclear	.0	21.8	.0	66.6
-CHP, industry	29.4	26.3	.1	5.8
-CHP, other	2.1	10.7	.4	4.3
-gas turbines, etc.	.0	.0	.3	.0

Table 1: Average annual production breakdown by technology 2000-05

5.2 Empirical implementation

We estimate the monthtly supply function of the thermal sector from data on the weekly system price and total non-hydro output in 2000-05.[21] We regress the thermal supply (non-hydro supply) on the price of electricity, the prices of fossil fuels and the time of year. A majority of the marginal cost of thermal plants consists of the price of the fuel. As explained, the thermal generation costs vary within the year for reasons related to heating demand and maintenance, both of which follow a seasonal pattern (nuclear plants and other large thermal power plants follow a seasonal maintenance schedule). To capture these effects, we include month dummies d_t in the regression equation,

$$q_t^f = \beta_0 + \beta_1 \ln p_t^{elec} + \delta x_t + \gamma d_t + \varepsilon_t, \tag{13}$$

where q_t^f is the thermal supply, and x_t is the vector of fuel prices. The thermal generation is composed of all other production than hydro, including wind power and the net import of electricity to the Nordic region.

The output price depends on thermal generation, and is thus endogenous. There are two natural candidates for instruments, the hydro production and the level of reservoirs, both of which influence the price level but not the cost of thermoelectricity. Given the slightly better fit in the first stage, we use the model with reservoir levels as instruments. We omit the precise estimation results and refer to Table 2 in Kauppi and Liski (2008), who estimate the same equation for a different purpose. We note that fossil fuel prices are strongly multicollinear, and all other fuel than oil prices can be dropped from the final estimation.[22]

We can now construct the price function $P(k_t, x_t)$ that determines the annual revenue for a new technology unit that enters this market, given the existing capacity k_t and oil the price x_t. Using this price function we will generate equilibrium capacity path as well as the annual price $P(\mathbf{k}(\hat{x}), \hat{x})$ as a function of the oil price history. To this end, we compute the historical annual monthly supply profile of thermal power over the years from data, and use this profile together with the estimated thermal supply to generate

[21] We use weekly demand data for the Nordic market in 2000-05 as published by the Organization for Nordic Transmission System Operators (Nordel). The system price data is published by Nord Pool, while electricity production by technology is reported by Nordel. We used the European Brent spot price for the price of fuel oil as reported by Reuters.

[22] Figure 3 in Kauppi and Liski (2008) illustrates the fit with observed prices, when actual thermal supply and oil prices are inserted to the estimated equation (13) to produce an estimate for the weekly electricity price.

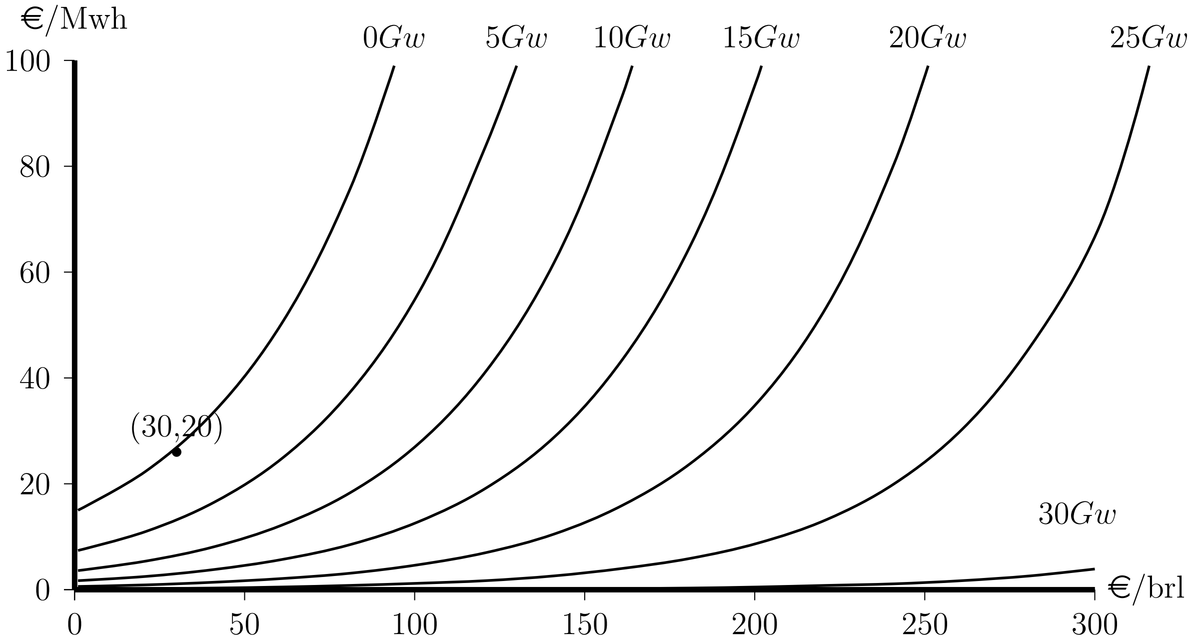


Figure 10: Average annual electricity price (€/Mwh) as given by estimated $P(x, k)$. Oil price €/barrel on the horizontal axis. Installed new capacity in Gigawatts indicated for each curve.

an annual revenue for new plant that runs at full capacity through the year. An increase in the total new capacity k is assumed to decrease the residual demand for thermal power one-to-one — the estimated supply curve (13) is shifted horizontally to the right at a given electricity price level.

More formally, let q_i^h and q_i^d denote the realized monthly hydro production and final demands, respectively. The difference $D_i = q_i^d - q_i^h$ is the realized residual demand that thermal production must meet in the absence of new capital. When there is k units of new capital, thermal production is $D_i - k$, and the implied market price is given by the estimated supply curve (13). We assume that D_i follows a normal distribution with the first and second moments estimated from data.[23] The annual revenue can be expressed as follows:

[23]We construct observations for D_i using data for demand and hydro production over the years 2000-2007, as published by Nordel. We use a slightly longer period 2000-07 for this estimation than that used for the thermal supply estimation, where we use the six years 2000-05. Including years 2006-07 in supply estimation is problematic because of regime changes such as introduction of emission permit markets. However, these changes do not significantly influence demand realizations and hydro availability.

$$P(k,x) = \sum_{i=1}^{12} \int \Pi(i,x,D_i - k)dF_i, \tag{14}$$

where the monthly price $p_i = \Pi(.)$ is given by the inverse of (13), and F_i is the cumulative distribution function for D_i in month i. Note that $P(k,x)$ is in fact the expected annual revenue but because the uncertainty is idiosyncratic we can apply $P(k,x)$ in the subsequent analysis exactly as before — $P(k,x)$ satisfies the assumptions of Section 4.1.

Figure 10 depicts the basic properties of the estimated revenue function $P(k,x)$. We express the revenue as the average annual price to make it comparable with historical prices observed in the market.[24] Each graph depicts the relationship between electricity price (€/Mwh) and oil price (€/barrel) for a given k. The upper-most graph corresponds to the historical capacity, i.e., $k = 0$. During the period 2000-05 the average price pair was close to 26 €/Mwh and 30€/barrel, which is quite precisely what the $k = 0$ -graph indicates. In the Figure, we add new capacity in 5000 Mw chunks ending at 30000 Mw which is close to the existing amount of thermal capacity in this market (excluding nuclear).

5.3 Results

For the counterfactual simulations, we take the fuel price as given by a Geometric Brownian Motion, matched with a long crude oil price series 1970-2007. Using the data in Nordhaus (2007) or Hamilton (2008), we conclude that there is no reasonable way of estimating a trend in prices, so we set $\alpha = 0$. Both data sets imply an extremely high annual volatility, $\sigma = .3$. We experiment with different volatility levels and take $\sigma = .2$ as our benchmark case; $\sigma = .3$ implies extreme investor caution as explained in more detail below. We have no single estimate for the investment cost flow rI, but the subsidy levels applied in practice imply that new green capacity can enter when they receive a fixed-price in the range 25 €/Mwh to 80 €/Mwh.[25]

See first Figure 11 which shows the equilibrium price-cost mark-up for new entrants as a function of the fuel price level (the shaded area). In this Figure, we assume $rI = 25$ and $\sigma = .2$, but individual mark-ups are also show for the higher uncertainty case $\sigma = .3$ (see dots). Note that the assumed entry cost is at the low end of the empirical support,

[24] That is, the Figure plots $P(k,x)$ divided by the number of hours in a year.

[25] The cost obviously varies across technologies but also for the same technology depending, e.g., on the site properties. For a review of costs for wind power, see the IEA (2008) report and Benitez et al. (2008). For a cost comparison across technologies, see Heptonstall (2007).

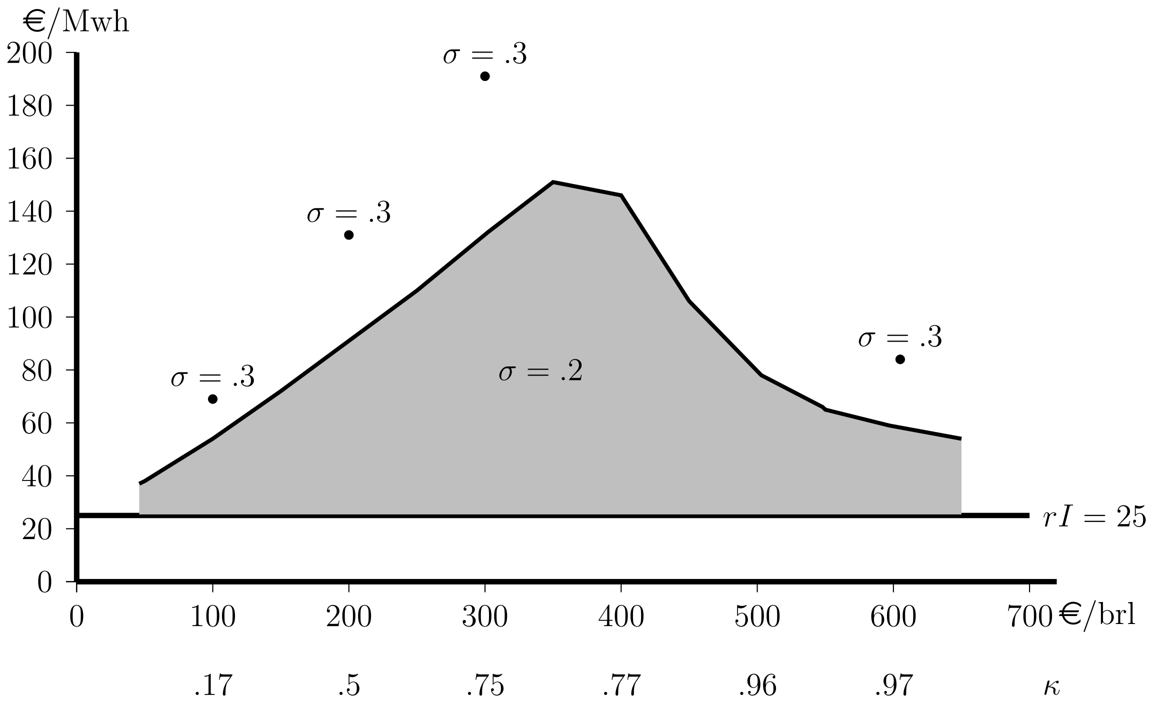


Figure 11: Equilibrium price-cost margin when $rI = 25$ and $\sigma = .2$. 'Dots' indicate individual mark-ups for ($\sigma = .3$). κ=fraction of capacity replaced (1=30 000 Mw).

and the uncertainty is lower than suggested by the historical fuel prices. Yet, the peak electricity price 160 €during the transition implies a 500 percent mark-up! However, a large fraction of the existing capital is replaced at much lower prices. We indicate this fraction by variable κ which gives the fraction of capital replaced as a function of the fuel price (the second horizontal axis below the Figure). Note that 50 percent of replacement requires that fuel price reaches 200 €/barrel but the electricity price is still in the domain of historical observations. But matching the historical uncertainty for fuels ($\sigma = .3$) leads to much higher resistance in replacement: the first 10 percent replacement requires a 100 €fuel price, and unprecedented output price increases for electricity consumers.[26]

Figure 12 is otherwise the same but the investment cost $rI = 50$ is in the middle range of the empirical support.

There are two points worth emphasizing when interpreting the results. First, as usual in electricity markets, the final consumer demand is relatively inelastic and, therefore, it is not clear how the output contraction-expansion pattern arises. In this market, the pattern should be understood as the contraction of the demand by the industrial sources

[26]The peak electricity price is 300 €/Mwh under $\sigma = .3$ (not depicted due to the scale).

which are price sensitive.[27] A fraction of the industrial consumers are on both sides of the wholesale market, and their net position depends on the price level. Sufficient price increases mean contraction of industrial demand as more of this demand is met by own production facilities; lower prices lead to expansion in the market demand from these sources.[28]

Second, for large uncertainty such as $\sigma = .3$, the implied output price goes off the support of historical prices used in the estimation of $P(k_t, x_t)$. This part of the results are not well supported by the calibration. However, this does not undermine the reliability of the results for the part where prices remain in a reasonable domain, say below 100 €/Mwh: the optimality of the entry at time t depends only on the price history up to t, and not on the properties of the revenue process defined for all prices higher than that at t.[29]. For this reason, we can modify the price function $P(k_t, x_t)$ relevant for future investments, without influencing the implications for prices ranges considered reasonable.

5.4 Policy experiments

Primary energy inputs, mostly fossil fuels, are often imposing external costs to the society, when their use releases unabated pollutants leading to a variety damages. If the social cost is fully internalized through a first-best penalty on the use of the inputs, the model description remains valid, with the modification that the social cost is added to the private supply curve. The gradualism and price dynamics are efficient features of the transition, even when externality prices are included.

However, the problem is that the social cost is in most cases not exactly known and its presence has emerged as a surprise to policy makers and citizens, and therefore there is a need to expedite the demand change, e.g., due to accumulated pollutant stocks such as greenhouse gases. There are multiple policy instruments currently in use, or under planning, in countries interested in inducing a faster than market-led demand change. [30]

Perhaps the most important policy instrument applied in the electricity sector is a price subsidy called feed-in tariff. There are different versions of the feed-in tariff in use,

[27]The consumers' price insensitivity is partly due to the order of moves in the clearing of the wholesale market; current consumption technologies do not allow for significant responses to prices (excluding large industrial consumers). In this sense, the trading institution is imperfect and implies welfare loss.

[28]This description applies well to the paper and pulp industry. We have not undertaken a separate industrial demand estimation but this supply is included in the aggregate supply used in our estimation. See Johnsen et al. (1999) for a discussion of the industrial demand in Norway.

[29]This follows from the Leahy's myopia result, as explained in Section 4

[30]For a discussion on existing subsidies in the EU, see European Commission (2005).

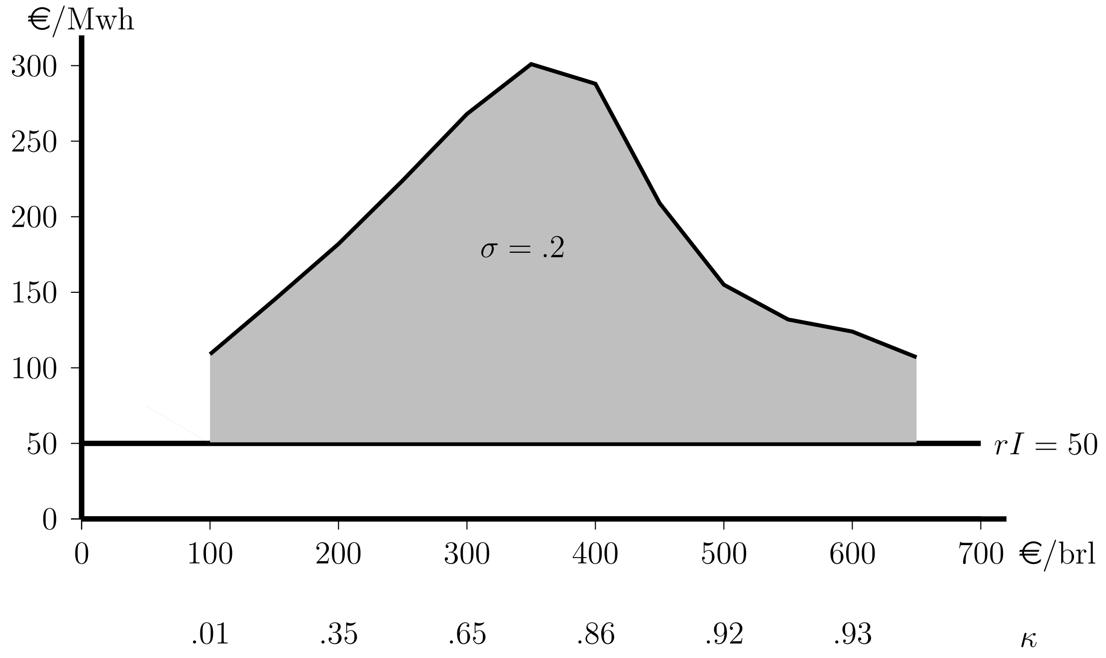


Figure 12: Equilibrium price-cost margin when $rI = 50$ and $\sigma = .2$. κ=fraction of capacity replaced (1=30 000 Mw).

but the common idea is to provide a price insurance to the new technology producer, i.e., a fixed-price or variable-price subsidy providing a pre-determined minimum revenue over time.[31]

We consider the following case: the tariff is a price floor ensuring that the new technology producer's sales price does not drop below a certain pre-determined level. Let τ denote the tariff level and assume

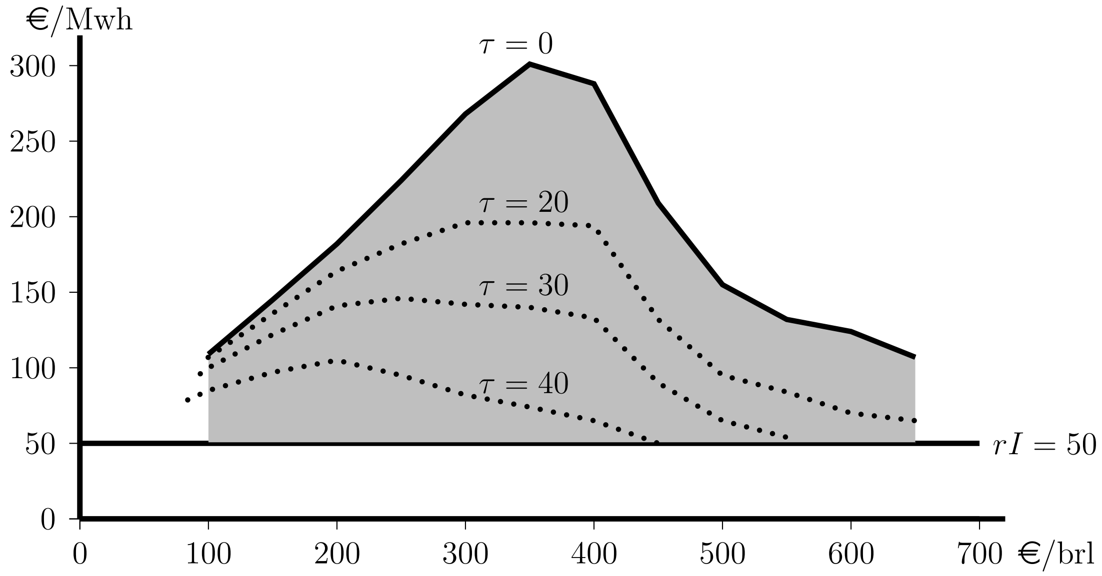
$$\tau < rI. \tag{15}$$

Whenever the final-good price falls below τ, the producer is compensated for the difference $p - \tau$. We assume that the tariff cost is collected from the consumers in a non-distorting manner. The effect of the tariff on the equilibrium can be understood by studying a price ceiling of the form

$$P_L(\hat{x}_t) \geqslant \tau, \forall \hat{x}_t. \tag{16}$$

The tariff pre-determines the lowest sales price for an entering new capital unit and, therefore, it influences the riskiness of the environment to which the new technology

[31]The subsidy is collected from consumers as part of the electricity bill, explaining in part the popularity the instrument; the costs do not appear in the government budget (in contrast to direct subsidies).



.01	.35	.65	.86	.92	.93	$\kappa_{\tau=0}$
.01	.39	.71	.89	.94	.96	$\kappa_{\tau=20}$
.03	.43	.75	.92	.96	.98	$\kappa_{\tau=30}$
.07	.49	.81	.95	.99	.99	$\kappa_{\tau=40}$

Figure 13: Equilibrium price-cost margin at tariff levels $\tau = 0, 20, 30, 40$ €/Mwh, when $rI = 50$ and $\sigma = .2$. κ_τ=fraction of capacity replaced when tariff is τ (1=30 000 Mw).

enters. For τ sufficiently close to rI, the new technology faces practically no risk and, as a result, entry takes place whenever the market price reaches rI. On the other hand, when τ is sufficiently low, it can be below the lowest prices conceivable during some part of the transition, and then it does not essentially change our description of the demand change without the tariff. However, a tariff that is between these extreme levels—full elimination of uncertainty, or no change in uncertainty faced by new firms)—it has interesting implications for the transition from the consumers' point of view.

In Figure 13, we depict the equilibrium entry mark-up for different tariff levels when the investment cost and uncertainty are as in Figure 12 ($rI = 50, \sigma = .2$). The mark-up and capital replacement levels without tariff ($\tau = 0$) at a given price are is thus as depicted in Figure 12. In general, the tariff speeds up the replacement rate and lowers the entry output price at each level of the energy cost. The tariff of 20 €/Mwh has a relatively moderate effect on replacement speed, although the effect on peak output prices is significant. The highest tariff considered covers 4/5 of the investment cost. The price profile falls into the historical empirical support but, despite the large subsidy rate, the capital replacement still requires extremely high fuel prices.

Consumers can find the tariff beneficial due to risk aversion that we have not explicitly modeled. The tariff can be seen as an insurance against extreme electricity prices. Yet another un-modeled reason for tariffs is an exogenous benefits from the decline in the energy input use (import dependence or pollution externalities). Finally, there is a potential reason for consumers to prefer tariffs that is included in our model. As such, the tariff distorts the efficiency of the equilibrium but the loss falls mainly on the Ricardian rents of the existing production structure. Since there rents are delaying the entry of lower-cost production, consumers may benefit in expected terms from imposing the tariff. We explore this next...

6 Concluding remarks

Our results is at the crossroad of several branches of the previous literature on energy costs. We conclude by discussing the potential extensions to the directions suggested by the literature.

There is a large literature on the exhaustible-resource nature of the energy commodity supply and the so called backstop technologies (for early papers, see Nordhaus 1973, Dasgupta and Heal 1974, Heal 1976; and for a later application, see, e.g., Chakravorty el al. 1997). This research casts the adoption problem in an exhaustible-resource framework without uncertainty. The models from the 70s typically feature a switch to the backstop as soon as the resource is physically or economically depleted. While such models are helpful in gauging the limits to resource prices using the backstop cost data (see the seminal work by Nordhaus), the predictions for the backstop technology entry are not entirely plausible if one accepts uncertainty and adjustment delays as characteristic features of the energy demand change.[32] However, while being less explicit about the capital replacement on the demand side, the exhaustible-resource approach is needed for understanding the long-run supply of the energy resource commodities. The inclusion of the resource supply would be a step towards a general equilibrium description of the energy demand change.[33]

Yet another step towards general equilibrium relates to the macroeconomics effects of the energy demand change. Macroeconomists have found it puzzling that the oil prices

[32] A more realistic approach to the energy demand change is described in Chakravorty et al. (1997) where the demand for exhaustible factors is heterogenous and backstop technologies such as solar energy have a declining trend in adoption costs. We provide a complementary and simpler approach to gradual energy demand technology transition, capturing similar features, but arising from Ricardian rents and persistent uncertainty in the energy input supply.

[33] Pindyck (1978) characterizes the traditional Hotelling model under uncertainty.

have an aggregate effect despite the low cost share of oil in GDP (See, e.g., Barsky and Kilian (2004) and Hamilton (2008)). One potential explanation is that factor price changes are propagated through movements in other factor prices they induced. We believe our explanation for the consumer price increase is different from previously identified propagation channels but, as such, it cannot be used to explain the historical macroeconomic experiences. It would be valuable to have a quantitative assessment of the effects identified in this paper in a macroeconomic context.[34]

7 Proof of Proposition 1 (Supplementary material)

This proof builds on the myopia result explained in Section 4. We derive the stopping rule for a myopic investor when the aggregate capacity k is taken as given, and from this we derive the equilibrium path $k = \mathbf{k}(\hat{x})$ and its properties. Define

$$\begin{aligned}\beta_1 &= \frac{1}{2} - \frac{(r-\delta)}{\sigma^2} + \sqrt{\left[\frac{(r-\delta)}{\sigma^2} - \frac{1}{2}\right]^2 + \frac{2r}{\sigma^2}} > 1, \\ \beta_2 &= \frac{1}{2} - \frac{(r-\delta)}{\sigma^2} - \sqrt{\left[\frac{(r-\delta)}{\sigma^2} - \frac{1}{2}\right]^2 + \frac{2r}{\sigma^2}} < 0.\end{aligned}$$

Lemma 2 *Given the specification (4)-(6), the optimal cut-off rule for a myopic investor as defined in Lemma 1 is*

$$x^m(k) = \begin{cases} \frac{\delta\beta_1(B+C)}{rB(\beta_1-1)}\left(rI - \frac{AC}{B+C} + \frac{BC}{B+C}k\right) \; \text{for } x \leq A - Bk \\ \left(-\frac{\beta_1(B+C)\left(\frac{A-Bk}{r}-I\right)}{(A-Bk)^{1-\beta_2}B\left(\frac{\beta_1}{r}+\frac{(1-\beta_1)}{\delta}\right)}\right)^{\frac{1}{\beta_2}} \; \text{for } x > A - Bk. \end{cases} \tag{17}$$

Proof. Given k, the revenue process for an existing new plant is defined by

$$\begin{aligned} P(x;k) &= \begin{cases} \frac{C(A-Bk)}{B+C} + \frac{B}{B+C}k, & \text{for } x \leq A - Bk \\ A - Bk, & \text{for } x > A - Bk \end{cases} \\ &= \begin{cases} Q(k) + Rk, & \text{for } x \leq A - Bk \\ A - Bk, & \text{for } x > A - Bk \end{cases} \end{aligned}$$

where we use the definitions

$$Q(k) = \frac{C(A-Bk)}{B+C}, R = \frac{B}{B+C}.$$

[34] See Wei (2003) for a general equilibrium assessment of frictions in capital replacement under a putty-clay approach.

The value of an existing plant, denoted by $V(x;k)$, satisfies the following dynamic programming equation (DPE)

$$\frac{1}{2}\sigma^2 X^2 V'' + (r-\delta)\, xV' - rV + P = 0,$$

where r is the discount rate, and $\delta = r - \alpha$. The general solution of the DPE is

$$\begin{aligned} V(x;k) &= \begin{cases} V_0\,(x;k)\,, \text{ for } x \leq A - Bk \\ V_+\,(x;k)\,, \text{ for } x > A - Bk \end{cases} \\ &= \begin{cases} B_1^0 x^{\beta_1} + B_2^0 x^{\beta_2} + \frac{Q(k)}{r} + \frac{Rx}{\delta}, \text{ for } x \leq A - Bk \\ \quad B_1^+ x^{\beta_1} + B_2^+ x^{\beta_2} + \frac{A-Bk}{r}, \text{ for } x > A - Bk. \end{cases} \end{aligned}$$

where

The two boundary conditions $\lim_{x\to 0^+} V\,(x;k) = \frac{Q(k)}{r}$ and $\lim_{x\to\infty} V\,(x;k) = \frac{A-Bk}{r}$ imply that $B_2^0 = 0$ and $B_1^+ = 0$. The two remaining parameters can be solved from value matching and smooth pasting conditions (omitted).

Denote the value of the option to install such a plant by $F\,(x;k)$. This must satisfy the following DPE

$$\frac{1}{2}\sigma^2 X^2 F'' + (r-\delta)\, XF' - rF = 0,$$

which has the general solution

$$F\,(x;k) = C_1 x^{\beta_1} + C_2 x^{\beta_2}.$$

The boundary condition $\lim_{x\to 0^+} F\,(x;k) = 0$ implies that $C_2 = 0$. The problem is to find C_1 and the myopic investment treshold x^m. There are two possible cases that must be considered separately: (1) $x^m \leq A - Bk$, and (2) $x^m > A - Bk$.

The boundary conditions in case $x^m \leq A - Bk$ are

$$\begin{aligned} C_1 x^{\beta_1} &= B_1^0 x^{\beta_1} + B_2^0 x^{\beta_2} + \frac{Q}{r} + \frac{Rx}{\delta} - I \\ \beta_1 C_1 x^{\beta_1 - 1} &= \beta_1 B_1^0 x^{\beta_1 - 1} + \beta_2 B_2^0 x^{\beta_2 - 1} + \frac{R}{\delta}. \end{aligned}$$

Because of the myopia-result, the ceiling $A - Bk$ is irrevelant in this case, and one can solve variable $C_1 - B_1^0$ instead of C_1. To see this, write these equations as

$$\begin{aligned} \left(C_1 - B_1^0\right) x^{\beta_1} &= B_2^0 x^{\beta_2} + \frac{Q(k)}{r} + \frac{Rx}{\delta} - I \\ \beta_1 \left(C_1 - B_1^0\right) x^{\beta_1 - 1} &= \beta_2 B_2^0 x^{\beta_2 - 1} + \frac{R}{\delta}. \end{aligned}$$

But since $B_2^0 = 0$ by the boundary condition above, we obtain the following linear equation defining x^m,

$$x^m = \frac{-\delta\beta_1\left(\frac{Q(k)}{r} - I_r\right)}{R(\beta_1 - 1)} = \frac{\delta\beta_1(B+C)}{rB(\beta_1-1)}(rI - \frac{AC}{B+C} + \frac{BC}{B+C}k).$$

The boundary conditions in case $x^m > A - Bk$ are

$$\begin{aligned} C_1 x^{\beta_1} &= B_2^+ x^{\beta_2} + \frac{A-Bk}{r} - I \\ \beta_1 C_1 x^{\beta_1 - 1} &= \beta_2 B_2^+ x^{\beta_2 - 1}. \end{aligned}$$

This implies that the investment trigger is given by the non-linear equation

$$x^m = \left(-\frac{\beta_1 (B+C)\left(\frac{A-Bk}{r} - I\right)}{(A-Bk)^{1-\beta_2} B\left(\frac{\beta_1}{r} + \frac{(1-\beta_1)}{\delta}\right)} \right)^{\frac{1}{\beta_2}}.$$

■

For the properties of the equilibrium it is enough to focus on the case $x^m \leq A - Bk$. Let us now use the notation $\hat{x}$ for the equilibrium investment trigger which is defined by the myopic trigger $x^m(k)$. We can see from (17) that for $x^m \leq A - Bk$, the myopic investment trigger $x^m(k)$ defines the equilibrium capacity as a linear function of the current record $\hat{x}$

$$\mathbf{k}(\hat{x}) = \frac{r(\beta_1 - 1)}{\beta_1 \delta C}\hat{x} + \frac{AC - rI(B+C)}{BC}.$$

Let us now explain the role of volatility for the equilibrium description to apply. Recall that $\hat{x}^*$ is the equilibrium investment trigger at which $\hat{x}^* = x^m = P = A - Bk^*$. We can thus use the derived $x^m(k)$ to solve for k^* from

$$\frac{\delta\beta_1(B+C)}{rB(\beta_1-1)}(rI - \frac{AC}{B+C} + \frac{BC}{B+C}k) = A - Bk, \tag{18}$$

which gives

$$\begin{aligned} k^* &= \frac{\beta_1(\delta AC + rAB - \delta rI(B+C)) - rAB}{B(\beta_1(rB + \delta C) - rB)}, \\ \hat{x}^* &= \frac{r\delta\beta_1(B+C)}{\beta_1\delta C + rB(\beta_1 - 1)} \end{aligned}$$

where the latter equation is obtained by evaluating $x^m(k)$ at k^*. Consider now $k = 0$ and the condition (18). The ratio $\beta_1/(1-\beta_1)$ increases in σ monotonically so that the left-hand side of (18) exceeds the right-hand side even at $k = 0$. This would imply that the market must shut down before new entry can take place. There is exists a unique σ^*

such that equation (18) holds as equality when $k = 0$. For all $\sigma < \sigma^*$ we can therefore find strictly positive value for k^* and thus for $\hat{x}^*$.

The investment trigger in terms of output price is

$$P_H(\hat{x}) = \frac{C(A - Bk)}{B + C} + \frac{B}{B + C}\hat{x} = rI + \frac{\beta_1 B(\delta - r) + rB}{\beta_1 \delta (B + C)}\hat{x} \text{ for } x \leq \hat{x}^*.$$

We see that the price is increasing in $\hat{x}$, implying contraction of output for $x \leq \hat{x}^*$. The price trigger is

$$P_H(\hat{x}) = A - B\mathbf{k}(\hat{x}) \text{ for } x > \hat{x}^*,$$

which is decreasing in $\hat{x}$. The output thus expands for $x > \hat{x}^*$.

The peak price follows by direct substitution

$$P_H(\hat{x}^*) = \frac{\beta_1 \delta r I(B + C)}{\beta_1(rB + \delta C) - rB},$$

which is increasing in C and σ. When $C \to 0$, the myopic investment trigger approaches

$$x^m \to \frac{\delta \beta_1 B}{rB(\beta_1 - 1)} rI,$$

which is independent of k. Thus, once this trigger is reached, there is a one-time jump to the long-run equilibrium. This completes the proof of the Proposition.

References

[1] Amundsen, E. and Bergman, L. (2006), "Why Has the Nordic Electricity Market Worked So Well?", *Utilities Policy*, 14, 148-157.

[2] A. Atkeson; P. J. Kehoe, , 1999. Models of Energy Use: Putty-Putty versus Putty-Clay, The American Economic Review, Vol. 89, No. 4, pp. 1028-1043.

[3] R. B. Barsky, and L. Kilian, 2004. Oil and the Macroeconomy Since the 1970s, Journal of Economic Perspectives, Vol. 18, No. 4, pp. 115-134.

[4] Benitez, L.E., P.C. Benitez, P.C., and G.C. van Kooten, 2008. The Economics of Wind Power with Energy Storage. Energy Economics, 30, 1973-1989.

[5] E.R. Berndt, and D.O. Wood, 1975. Technology, Prices, and the Derived Demand for Energy, Review of Economics and Statistics, Vol. 57, No. 3, pp. 259-68.

[6] Borenstein, S., Bushnell, J. and Wolak, F. (2002), "Measuring Market Inefficiencies in California's Restructured Wholesale Electricity Market", *American Economic Review,* 92, 1376-1405.

[7] Chakravorty, U., Roumasset, J., Tse, K., 1997. Endogenous substitution among energy resources and global warming. Journal of Political Economy 195, 1201-1234.

[8] Dasgupta, P., Heal, G., 1974. The optimal depletion of exhaustible resources. Review of Economic Studies (Symposium), 3-28.

[9] Dixit, A.,1989. Hysteresis, Import Penetration, and Exchange Rate Pass-Through, The Quarterly Journal of Economics, Vol. 104, No. 2., 205-228.

[10] International Energy Agency (IEA), 2007. IEA Wind Energy Annual Report.

[11] Dixit, A.K, and R.S. Pindyck, 1994. Investment under Uncertainty, Princeton University Press, Princeton, New Jersey.

[12] Griffin, J. M. and P. R Gregory, 1976. An Intercountry Translog Model of Energy Substitution Responses American Economic Review. vol. 66, issue. 5, 845-57.

[13] Hamilton, J. D. 2005. Oil and the Macroeconomy, prepared for the New Palgrave Dictionary of Economics.

[14] Heal, G., 1976. The relationship between price and extraction cost for a resource with a backstop technology. Bell Journal of Economics 7, pp. 371-378.

[15] Heptonstall, P. 2007. A review of electricity unit cost estimates. Working Paper, December 2006, Updated May 2007. Report for UK Energy Research Centre, Reference.

[16] Johnsen,T.A., S. K. Verma, and C. Wolfram, 1999. Zonal Pricing and Demand-Side Responsiveness in the Norwegian Electricity Market,with Tor Arnt Johnsen and Shashi Kant Verma, University of California Energy Institute. Power Working Paper 063.

[17] Joskow, P., 2008. Challenges for Creating a Comprehensive National Electricity Policy, MIT-CEEPR WP-2008-019.

[18] Juselius M., and R. Stenbacka, 2008. The Relevant Market for Production and Wholesale of Electricity in the Nordic Countries: An Econometric Study. HECER Discussion Paper No 222 / May 2008.

[19] Kauppi, O., and M. Liski, 2008. An empirical model of imperfect dynamic competition and application to hydroelectricity storage. MIT-CEEPR working paper WP-2008-11.

[20] Leahy, J.V. 1993. Investment in Competitive Equilibrium: The Optimality of Myopic Behaviour, The Quarterly Journal of Economics, Vol. 108, No. 4, 1105-1133.

[21] Linn, J., 2008. Energy Prices and the Adoption of Energy Saving Technology. The Economic Journal, vol. 118(533), 1986-2012.

[22] Liski, M., and P. Murto, 2006. Backstop Technology Adoption. HECER WP-Paper No 94.

[23] European Commission, 2005. Interactions of the EU ETS with Green and White Certificate Schemes, report prepared by NERA for Directorate-General Environment.

[24] Nordhaus, W., 1973. The Allocation of Energy Reserves. Brookings Papers 3, 529-570.

[25] Nordhaus, W.D., 2007. WhoÕs Afraid of a Big Bad Oil Shock? The Brookings Panel on Economic Activity.

[26] Pindyck, R. S., 1978. The Optimal Exploration and Production of Nonrenewable Resources, Journal of Political Economy, vol. 86, issue 5, 841-61.

[27] Pindyck, R. S., 1993. A Note on Competitive Investment under Uncertainty, The American Economic Review, Vol. 83, No. 1., 273-277.

[28] Popp,.D.C., 2002. The Effect of New Technology on Energy Consumption. Resource and Energy Economics, vol. 23, issue. 3, 215-39.

[29] Popp, D.C., 2002. Induced Innovation and Energy Prices. American Economic Review, vol. 92, issue. 1, 160-80.

[30] Roques F.A., W.J., Nuttall, D.M, Newbery, R. de Neufville, and S. Connors, 2006. Nuclear Power: A Hedge against Uncertain Gas and Carbon Prices? Energy Journal, vol. 27, issue. 4, 1-23.

[31] Thompson, P., and T. G., Taylor, 1995. The Capital-Energy Substitutability Debate: A New Look. Review of Economics and Statistics, vol. 77, issue. 3, 565-69.

[32] Wei, C., 2003. Energy, the Stock Market, and the Putty-Clay Investment Model. American Economic Review, vol. 93, issue, 311-23

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