HE
CER
Helsinki
Center
of
Economic
Research

Discussion Papers

Uncertainty and energy saving investments

Matti Liski
Aalto University School of Economics, HECER and MIT-CEEPR

and

Pauli Murto
Aalto University School of Economics and HECER

Uncertainty and energy saving investments*

Abstract

Energy costs are notoriously uncertain but what is the effect of this on energy saving investments? We find that real-option frictions imply a novel equilibrium response to increasing but uncertain energy costs: early investments are cautious but ultimately real-option frictions endogenously vanish, and the activity affected by higher energy costs fully recovers. We use electricity market data for counterfactual analysis of the real-option mark-ups and policy experiments. Uncertainty alone implies that the early compensation to new technologies exceeds entry costs by multiple factors, and that uncertainty-reducing subsidies to green energy can benefit the consumer side at the expense of the old capital rents, even in the absence of externalities from energy use.

JEL Classification: D81, D41, L1, Q42

Keywords: energy cost, investment, real options, competitive equilibrium, electricity.

Matti Liski

Department of Economics
Aalto University School of Economics
P.O. Box 21240
FI-00076 AALTO
FINLAND

e-mail: matti.liski@hse.fi

Pauli Murto

Department of Economics
Aalto University School of Economics
P.O. Box 21240
FI-00076 AALTO
FINLAND

e-mail: pauli.murto@hse.fi

1 Introduction

Figure 1 shows the real price of crude oil over the past 150 years. The defining feature of this key determinant of energy costs is persistent uncertainty: the prediction for future energy costs depends on the current level of the price, and the scope for potential deviations from the prediction increases, the farther to the future one is looking.[1] It is an obvious fact that uncertainty of this type and magnitude is relevant for long-lived investments targeted at saving on energy use in capital intensive sectors such as electricity, transportation, housing, and manufacturing. Yet, the implications of uncertainty for capital replacement and energy use in an equilibrium context have not been scrutinized.

To capture the dynamic response to increasing but uncertain energy costs, we develop an equilibrium model of investments in energy savings. We find that uncertainty gives rise to a novel capital adjustment pattern which has implications for policies expediting the transition to green technologies. The uncertainty of costs implies uncertain returns to green investments, generating equilibrium real-option mark-ups over entry costs. As will be shown, the mark-ups are first increasing over time as entry makes the market more risky for later entrants. But, ultimately, the real-option frictions will endogenously vanish in equilibrium: investments in green technologies are uncertainty reducing, as they close the transmission channel of the persistent uncertainty. New technologies may bring about uncertainties of their own such as those of wind and solar power but, as long as these are idiosyncratic (or non-persistent) by nature, the source of investor caution is eliminated during the technology transition. As a result, the model predicts final output expansion and consumer price decline, despite the increasing primary energy costs.

While our framework fits multiple contexts, the electricity sector is of particular importance.[2] Primary energy costs in this sector are uncertain, and the existing production capacity is capital intensive and heterogeneous.[3] In comparison with many other sectors

---

[1]Pindyck (1999), after studying a price history of similar length, found that for practical (investment) purposes one may take the price process as given by a random walk. Similarly, Hamilton (2009) concluded "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". Dvir and Rogoff (2009) find evidence for interesting regime changes between the early industrialization and the recent past, but persistence and volatility remain relevant for investment decisions.

[2]For example, in the US the sector uses 42 per cent of primary energy, 34 per cent of fossil fuels, and produces about 40 per cent of carbon dioxide emissions. See Joskow (2008).

[3]We may rank the technologies in the electricity sector in the order of increasing dependence on primary energy fuels as follows: hydro, wind, nuclear, coal, gas, and oil. Technologies represent various combinations of *ex ante* sunk costs and *ex post* variable costs such that the firms may have been indifferent
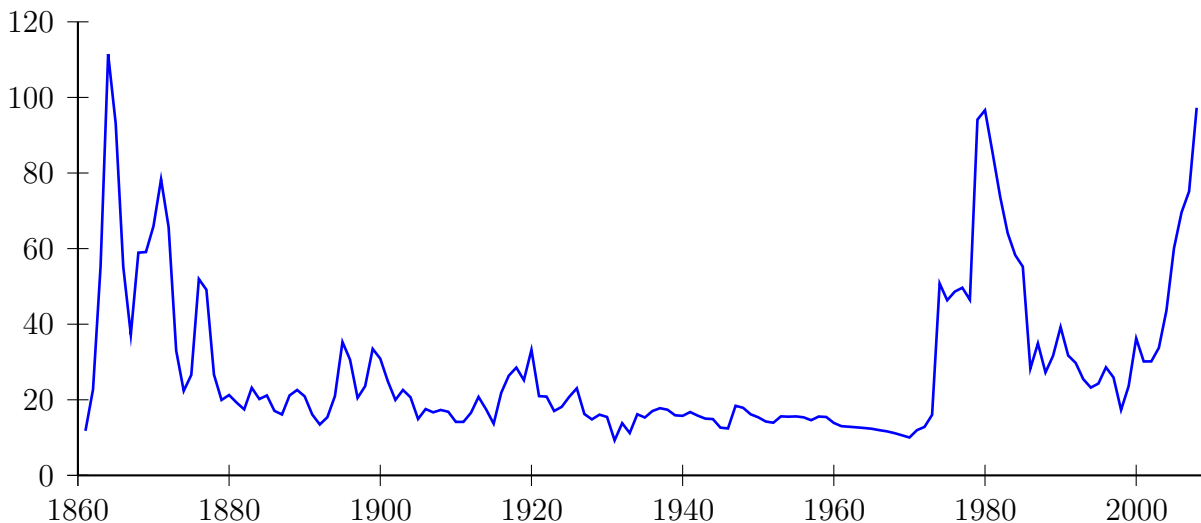
Figure 1: Crude oil prices 1861-2008 measured in 2008 US Dollars. Source: BP Statistical Review of World Energy, 2009.

such as housing and transportation, there are relatively mature technologies in electricity generation that are substitutes for the existing fossil fuel dependent technologies. In this sector, it is also transparent why the persistent uncertainty of fuels is transmitted to the returns of green technologies: in deregulated electricity markets the fossil-fuel plants are marginal producers so that the variation of their fuel costs is reflected in the electricity price, thereby making the entrants' payoffs uncertain in a way that generates real-option mark-ups. Moreover, in electricity markets the data on market fundamentals is exceptionally precise, allowing for a precise estimation of the revenues for entering technologies, and thus a relatively realistic counterfactual analysis of the equilibrium mark-ups needed for entry at each conceivable future level of the fuel cost.

For the quantitative assessment of the investor caution coming from uncertainty, we develop a general representation of the investment equilibrium, lending itself to the analysis of the estimated industry structure. We take data for the analysis from the Nordic electricity market.[4] While the quantitative magnitudes are likely to differ across

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. Heterogeneity of variable costs implies that units are heterogeneously hit by higher energy costs. It is a standard practice in electricity market studies to evaluate the rents from heterogeneity to isolate them from rents arising from market power, for example. See Wolfram (2000) and Borenstein et al. (2002). These rents are important for adjustment delays in capital replacement in our model.

[4]The Nordic market is among the largest deregulated wholesale electricity markets in the world. It has the interesting feature that more than half of the production comes from renewable energy sources.

markets, the lessons seem applicable to deregulated electricity markets in general: they share the common feature that marginal producing units are fossil-fuel plants through which the uncertainty is transmitted to the returns for the green technologies, which is the main pre-condition for the results.

To isolate the investor caution coming from the uncertainty, we intentionally make assumptions that are favorable to entry. We find that the uncertainty of conventional energy costs alone is a significant source of investor caution. Under relatively optimistic scenarios, the price-cost margins exceed investors' costs by multiple factors during the transition. As a result, the transition towards green technologies in the electricity sector is likely to imply temporarily high consumer prices. We also perform policy experiments. In the absence of external benefits such as pollution damages, uncertainty-reducing green energy subsidies distort the overall welfare, and imply a remarkably large net transfer of wealth from producers to consumers: the cost of the intervention falls to a large extent on the old capital rents. We provide quantitative assessment of the social losses, reduction in the entry mark-ups, and change in the speed of transition following from the uncertainty-reducing subsidies.

After the oil price shock of the 70's, the 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; and, e.g., Thompson and Taylor 1995). Capital replacement approaches for understanding the dynamic price response are based on capital adjustment costs coming either exogenously (Pindyck and Rotemberg, 1983) or from a "putty-clay" structure (Atkeson and Kehoe, 1999).[5] Our results provide another perspective on adjustment frictions in capital replacement and on the implied energy demand response. It should be emphasized that we do not consider an economy-level response to prices; rather, the framework is well suited for gauging the consumer price increase needed for the capital replacement to take place in particular industries.

Our main result —the output contraction-expansion patter generating the equilibrium real options— is in spirit similar with that implied by the putty-clay model (Atkeson and Kehoe, 1999). The mechanism behind our result is quite different, however. The short-run output contracts in order to create a mark-up for early green entrants. The mark-

We exploit the exogenous variation in the availability of renewable energy to identify the supply curve for the fuel-dependent power. We discuss the institutional details and the estimation in the Application Section; for a description of the Nordic market, see M. von der Fehr et al. (2005) and Amundsen and Bergman (2006).

[5]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.

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. Moreover, our model implies a distinct feature of capital replacement costs: consumers face a temporary increase in uncertainty of the final good price. We also find that the greater is the heterogeneity of the existing structure with respect to fuel efficiency, the larger is the output contraction needed, and the associated increase of the consumer price level and its volatility. Finally, the long-run output expands, because eventually the energy market becomes disconnected from the output market, and this shuts down the transmission channel for uncertainty to the output market and thus to the entrants' profits. The decline in the transmitted uncertainty reduces the required mark-ups and boosts investment, leading to the recovery of output.[6] These features of the capital replacement are socially optimal.

We focus on capital replacement of existing technologies with mature alternative technologies to communicate the significance of investor caution coming from uncertainty alone, and its implications for welfare and policies. We therefore rule out additional sources of inertia such as risk aversion, scarcity of investment opportunities (e.g., nuclear or wind power sites), incomplete capital markets, and, most fundamentally, the non-existence of mature substitute technologies. The higher energy costs may need to induce innovations first and capital replacement only thereafter; see Jaffe and Stavins (1995), Newell et al. (1999), and Popp (2001, 2002). These sources of inertia are distinct from and additional to those identified by our model.

Our equilibrium formulation and computations utilize the real-options theory (see, e.g., Dixit and Pindyck, 1994) and, in addition, results developed in Leahy (1993) and further developed in Baldurson and Karatzas (1997). To generate the values of future income streams for entering plants, we take the fuel price uncertainty as characterized by Pindyck (1999), and Hamilton (2009).

In Section 2, we describe a simple version of the model, and use figures to explain the basic mechanism (the analytical solution is in a supplementary Appendix). In Section

---

[6]The green technologies may have uncertainties of they own, and one may think that this jeopardizes the generality of our conclusions. To fix ideas, consider wind power, which is subject to considerable variation in the availability of wind. From the investor's point of view, this uncertainty looks the same independently of the year, and therefore this uncertainty is not a source of investor caution in the sense discussed in this paper. While the wind availability is clearly correlated across locations, it should not be correlated across years in a way that generates persistence and real-options for entrants. This same argument seems to apply other renewable sources of energy.

3, we describe the equilibrium more generally and connect the equilibrium to Leahy's results. In Section 4, we estimate the revenues for entrants using data from the Nordic electricity market, and perform the policy experiments. The computational Appendix and the program for simulations (including data) is available on the authors' webpage.

2 A simple model

To fix ideas we first present the energy use response to prices in a simple final-good demand and supply framework, and then consider a more general version that allows multiple interpretations (in Section 3). For now, assume an inverse demand of the final good $p = D(q) = A - Bq$, where p is the price and q is the final good demand, and the parameters A and B are strictly positive. We can think of any final good market whose supply side uses energy inputs, or alternatively, energy-saving technologies.[7]

There are two basic sources of supply, namely the old energy-intensive technology and the new energy-saving technology.[8] There is 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 is $x_t + Cq_t^f$, where $C > 0$. The marginal cost is thus strictly increasing in q^f, and x is the intercept of the old technology supply curve — x can be thought of as the direct purchase cost of primary energy (fuel).

The new technology firms produce the same output but use no energy input. We denote the measure of these firms by k, i.e., the existing energy-saving capital stock is k. We will introduce the entry problem of a new firm, but for a moment we take 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 set its variable cost to zero.[9] The combined total

---

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

[8]The new technology may still use energy but this energy is not coming from the fuel inputs considered. 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

[9]The exact level of this cost does not matter as long as the new capital is the least-cost option, once in place. Under this assumption, any positive flow cost can eliminated and incorporated into the initial investment cost of a new entrant.

inverse supply can then be written as

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

Given k and x, the usage of old structure $q^f = q - k$ clears the final-good market,[10]

$$p = D(q) = S(q, k, x) > 0. \quad (2)$$

Let us introduce time and uncertainty. Time is continuous, and the energy cost $x_t \geqslant 0$ follows a Geometric Brownian Motion (GBM) with drift $\alpha > 0$ and standard deviation $\sigma \geqslant 0$,

$$dx_t = \alpha x_t dt + \sigma x_t dz_t. \quad (3)$$

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 long run (α is strictly positive). The assumptions on x_t introduce persistence into the fuel prices.[11]

For a given k, the stochastic fuel price implies that the output price is a stochastic process too. That is, condition (2) implicitly defines $q = q(k, x_t)$ and thus the output price,

$$p = D(q(k, x_t)) = P(k, x_t) > 0. \quad (4)$$

Because an entrant supplies one marginal unit of output, the output price process $P(k, x_t)$ is the revenue process for an entrant, given existing k. 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 marginal capital unit by paying an irreversible up-front 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. Finally, we assume $r > \alpha$.

2.1 Transition without uncertainty

To isolate the heterogeneity of old units as a source of gradualism in the build-up of the energy-saving capital stock, let us eliminate uncertainty and assume 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 (3).

---

[10]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 positive, $D(k) > 0$.

[11]Pindyck (1999) concludes "...for purposes of making investment decisions, one could just as well treat the price of oil as Geometric Brownian Motion, or related random walk process."
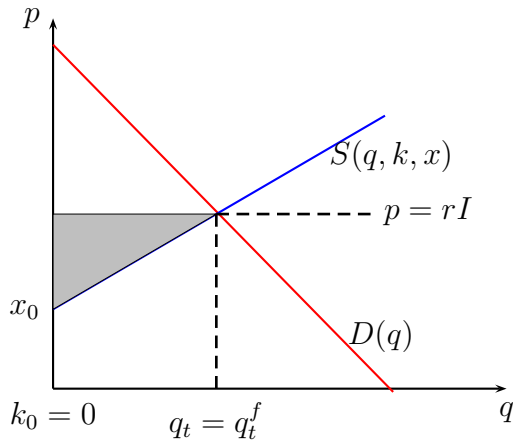
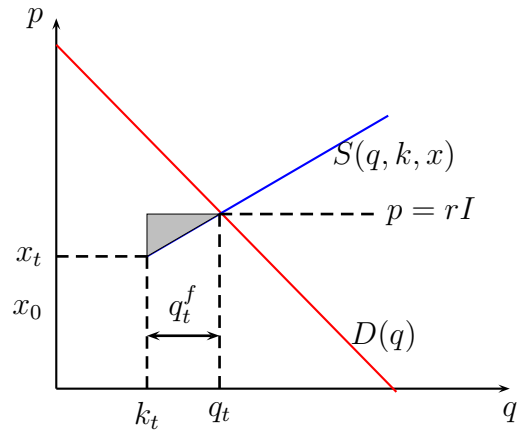Figure 2: Entry of the first energy-saving capital unit under certainty.

Figure 3: Entry of additional energy saving capital at k_t under certainty.

In Figure 2, the entry cost of one marginal 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 cannot enter, and the old technology satisfies the full demand. But since the energy price is on an upward trend shifting the supply curve up, the output price must meet rI at some point, and then the first new technology unit enters the market, as its present-value revenue p/r covers the investment cost I. This is the situation depicted in Figure 2.

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 3 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. Note that the "speed" of the transition depends on the slope of the supply (heterogeneity of old units) but not on the slope of the demand.

2.2 Output contraction and the mark-up

We assume now that there is volatility in the energy cost process, i.e., $\sigma > 0$ in (3). 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 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 4, the first entry takes place 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 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 $P_L(\hat{x}_0)$.[12] For the mark-up to arise, the output must contract below the level that we saw without uncertainty.

From the standard real-options analysis (Dixit and Pindyck, 1994), we know that the first entering unit requires a mark-up above its deterministic entry cost, when facing the above described uncertainty.[13] Consider now how this mark-up develops in equilibrium when more capital enters. We argue that the mark-up and thus the consumer price must increase.

Figure 5 depicts time point $t > 0$ where there are k_t new units in place. We have drawn Figure 5 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

---

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

[13]For a moment, we put aside the issue of how this firm conceives the actions of later entrants 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.
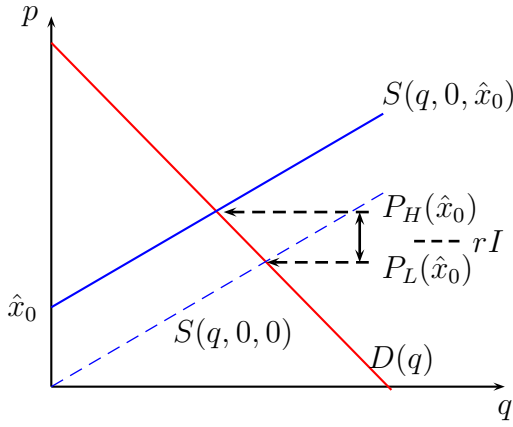
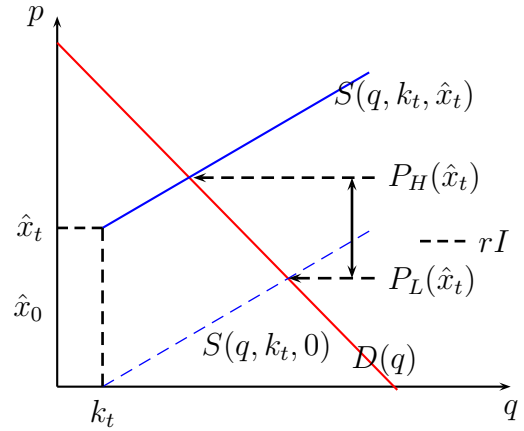Figure 4: Entry of the first energy-saving capital unit under uncertainty.

Figure 5: Entry of additional energy saving capital at k_t under uncertainty.

technology supply curve remain the same. 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 and output contraction follow from the combination of two elements in the model. First, heterogeneity ensures that the old production structure is replaced only gradually. Second, the old structure can benefit from the potential downside development in the fuel market.[14] This downside risk for the new entrant implies that the input-saving substitute can only enter when not only the energy but also the consumer price reaches record high levels. We see that uncertainty protects the rents of the old supply structure, as more extreme energy and consumer prices are needed to induce additional entry.

2.3 Output expansion and the mark-up decline

We argue 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.

---

[14]One may ask if this property arises only because the old structure remains existing by assumption. We show in Section 3.2 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 on the insights provided by this simpler model.
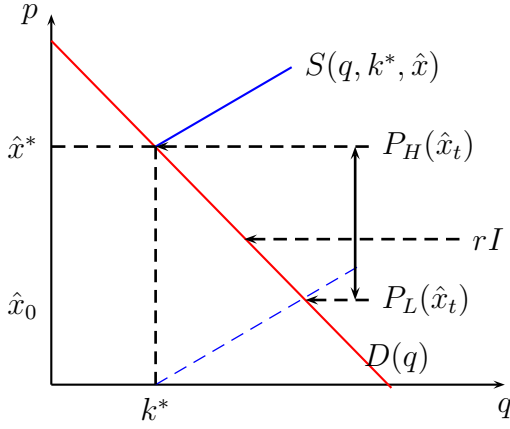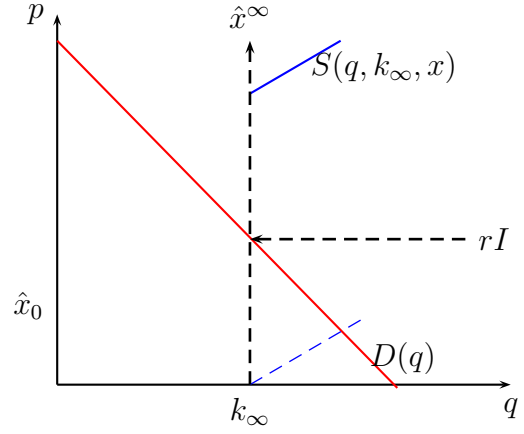
10

Figure 6: The consumer price peak   Figure 7: The long-run equilibrium

Consider 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 6 for this situation. The market environment cannot become more risky for a new entrant than the situation described here, and therefore 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 immediate future: the input price must decline by the discrete amount $\hat{x}_t - \hat{x}^* > 0$ for old production to take place and, by the persistence of x, this cannot happen next moment. 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 output price declines in $\hat{x}_t$ after peaking at $P_H(\hat{x}^*)$. The output recovers and the required entry mark-up declines.

For sufficiently high $\hat{x}_t$, the old structure produces with negligible probability 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 output at all previous entry points. In this sense, the adjustment friction of capital replacement vanishes. See Figure 7, where k_∞ denotes the limit of the capital stock that is approached asymptotically as $\hat{x} \to \infty$.

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*

11

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 σ;*

- *if $C \to 0$, $P_H(\hat{x}^*)$ approaches the first entry price;*

- *investor mark-up disappears in the long run: $P_H(\hat{x}_t) \to rI$ as $\hat{x} \to \infty$.*

Proof. Supplementary material, available at the authors' webpage. ∎

The proof is straightforward real-options analysis, once the connection to Leahy (1993) is understood (we explain the connection in detail in the next Section). The solution gives an explicit upper bound on uncertainty measured by σ^* such that entry starts before the market shuts down.[15]

The resistance to entry depends on the heterogeneity measured by the slope of the supply curve, C. The greater is slope C, the lower is the responsiveness of entry to higher energy prices. When C is close to zero, there is a large one-time entry soon as the output price reaches $P_H(\hat{x}^*)$, followed by gradual output expansion and entry.[16] Recall that in the absence of uncertainty, the "speed" of the transition depends on the slope of the supply only, for a given input price development. Here, with uncertainty, the speed depends on the slope of the demand too. For example, if $C = 0$, the demand slope is important for gradual output expansion and entry; highly elastic demand tends to imply more gradual transition for a given input price process. Finally, if we increase the energy cost uncertainty, captured by σ, the old technology is more protected from entry in the following sense: its replacement requires higher mark-ups and thus prices.

As we will see in our Application, policies reducing the uncertainty that the entrants face will cut down entry mark-ups, speed up the transition, and transfer wealth from old producers to consumers.

For illustration, see Figures 8-10 depicting equilibrium paths based on the simple model. In Figure 8, we show a sample path for the energy price (solid line) and the historical maximum (dotted line). Figure 9 depicts the path for the energy saving capital. Figure 10 depicts the output price path, declining to the deterministic investment cost

---

[15]If uncertainty exceeds this level, the peak output price would be the first entry price.

[16]When $C = 0$, the output contraction phase does not exist since the downside risk for the entrants is not affected by entry: as long as the old capacity is producing, the output price is just x. Therefore, when the first investment trigger is reached, all producing units are replaced. From this point onwards, the old units are idle at investment points, and the reasoning for output expansion is the same as for $C > 0$.
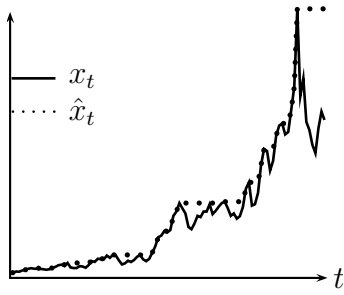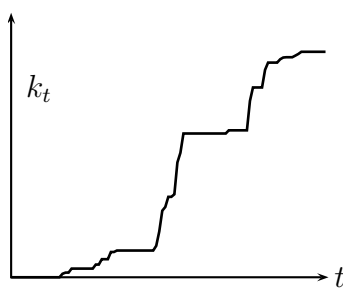
Figure 8: Sample path for the energy price

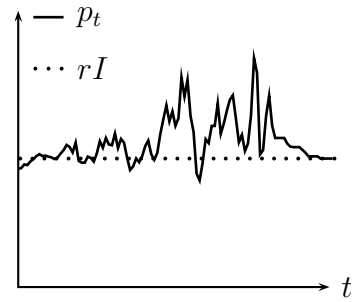Figure 9: The energy saving capital stock

Figure 10: The consumer price path

towards the end of the path. Figure 10 thus shows that the temporary increase and the final decline of the final-good uncertainty is a novel feature of the capital replacement path implied by this model.[17]

2.4   Discussion

Before moving to the more general framework that prepares the ground for empirical applications, let us discuss the potential interpretations and extensions of the results. By being a simple supply and demand framework, the model should fit any context where these concepts can be applied. What is essential for the results is the payoff dependence of the old and new technologies. In electricity or manufacturing, this dependence is achieved by the fact that all technologies are serving the same final-good demand — the consumer does not care how much each producer spends energy as long as the final good is the same. In housing and transportation, the payoff dependence must be understood slightly differently, although the principle is the same. In housing, the final good is the service from heating or cooling the houses, and the capital structure is embodied in the fleet of houses. Old houses depend more on energy providers (e.g., gas, oil, electricity, co-generation) than newly insulated or otherwise restructured houses. The investments in heterogenous houses are made by owners, and they thus destroy demand of the energy providers. The better is the market-level insulation of houses, the lower is the demand for energy providers and thus the lower are the consumers' energy prices, for a given primary energy cost. Note that in this description the energy providers' marginal cost is

---

[17]Note that the overall volatility of consumer prices increases temporarily during the transition because (i) for geometric Brownian motion larger x means higher absolute volatility for x, and (ii) the domain for conceivable consumer prices increases for reasons explained above.

increasing in the energy provided. Similar description applies to transportation, where consumers make investments in energy-using consumer durables (automobiles) providing the transportation service. Automobiles are heterogenous in their fuel efficiency, and consumers switch to more efficient automobiles when direct fuel costs increase. Here, the fuel cost to consumers is the price at the pump, which depends on refinery, transportation, and other costs of serving the gasoline, in addition to the primary energy cost. More investments mean lower demand for gasoline and thus lower prices, leading to the payoff dependence between technologies.[18]

Note that the supply curve $S(q, k, x)$ need not depend on k and x in the additive way as assumed for the illustration; the next section sets up a framework where the relationship between the equilibrium price, capital k, and cost x is general. This allows for multiple interpretations of how investments save energy, e.g., including piecemeal upgrades of old plants and structures, or new investments that are explicitly additional to the old ones. Another generalization regarding investments that we could easily include is heterogeneity of investment costs. Rather than assuming a fixed constant I we could work with an increasing investment cost function $C(I)$, reflecting underlying scarcities limiting the overall entry to the market and creating rents to early entrants. For example, the quality differences of wind and nuclear power sites, or the limited availability of special materials needed for alternative technologies may represent sources of increasing investment costs. Increasing costs, if considered reasonable, can have substantial implications in the sense that the overall entry can be limited to the extent that output contraction is not followed by recovery. However, for most of the paper we keep the assumption of unlimited entry at a given investment cost I. This is because we like to make assumptions that favor entry to isolate transition delays coming from investor caution only.[19]

3 General model

3.1 Equilibrium

In this section we generalize the setting, and define the equilibrium more formally by building on the results of Leahy (1993). We do not have to be specific about the process

---

[18]For studies considering energy efficiency in these sectors, see Linn (2008) for manufacturing, Fischer et al. (2007) for analysis of automobile fuel efficiency, and Jaffe and Stavins (1995) fo energy savings in housing.

[19]In the empirical application, we discuss elements that tend to further increase "adjustment costs" such as the investor heterogeneity.

describing the energy costs but we can assume a general diffusion process

$$dx = \alpha(x)dt + \sigma(x)dw, \tag{5}$$

where w is a Wiener process, and x is restricted to positive values.[20] This formulation admits the commonly used specifications used in the analysis of irreversible investments under uncertainty.

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 output 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 (k_t, x_t). It is therefore natural to restrict to Markovian strategies. Moreover, in the current context we can restrict further to strategies that can be expressed in cut-off form.

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

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

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

---

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

We use $\mathbf{x}^*(k) = \infty$ to indicate that the firm does not enter at any level of x_t, and $\mathbf{x}^*(k) = 0$ to indicate that the firm enters immediately for any value of $x_t \geq 0$.[21]

We will see that there is a symmetric equilibrium, where all firms adopt the same cut-off strategy x^*. To this end, we must derive the capacity path that such a strategy profile induces. Let us denote by $\hat{x}_t$ the historical maximum value of x_t up to time t:

$$\hat{x}_t \equiv \sup_{\tau \leq t}\{x_\tau\}. \tag{6}$$

We can now define the aggregate capacity path as a function of $\hat{x}_t$.

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

$$\mathbf{k}^*(\hat{x}_t) = \inf\{k \geq 0 \,|\, \mathbf{x}^*(k) > \hat{x}_t \ \}. \tag{7}$$

Note that $\{\mathbf{k}^*(\hat{x}_t)\}$ is an increasing stochastic process, and its value at time t is fully specified by the development of x_t up to time t. For interpretation, recall that there is a continuum of entrants, each with a strategy commanding them to enter as soon as x_t hits $\mathbf{x}^*(k)$. Of course, it would make no sense to assume that all the firms actually enter at the same time but, 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 $\mathbf{x}^*(k_t)$, 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.[22]

Equation (7) together with (5) defines the law of motion of k for a given symmetric entry strategy $\mathbf{x}^*$. To find an equilibrium, we must also check the optimality of the entry strategy against 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

---

[21]Leahy (1993) expresses strategies as a cut-off level for the output 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.

[22]This "rationing" among entrants could be alternatively modeled by (i) giving names to entrants and assuming asymmetric strategies, or by (ii) introducing heterogeneity in entry costs. In the limit where the heterogeneity vanishes, the aggregate outcomes are identical in all of these formulations.

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:

$$F(x_t, \hat{x}_t; \mathbf{k}) = \sup_{\tau^* \geq t} E\left[\int_{\tau^*}^{\infty} P(\mathbf{k}(\hat{x}_\tau), x_\tau)e^{-r(\tau-t)}d\tau - Ie^{-r(\tau^*-t)}\right], \tag{8}$$

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. 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 $\{k_t\} \equiv \{\mathbf{k}(\hat{x}_t)\}$. 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

$$F(x_t, \hat{x}_t; \mathbf{k}) = 0. \tag{9}$$

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

$$E\left[\int_t^{\infty} P(\mathbf{k}(\hat{x}_\tau), x_\tau)e^{-r(\tau-t)}d\tau\right] - I = 0. \tag{10}$$

Definition 3 *The equilibrium consists of a Markov cut-off strategy profile $\mathbf{x}^*$ and corresponding capacity path $\{k_t\} = \{\mathbf{k}^*(\hat{x}_t)\}$ such that*

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

- *(10) holds whenever $x_t = x^*(\mathbf{k}^*(\hat{x}_t))$*

- *$\mathbf{k}^*(\hat{x}_t)$ is given by (7)*

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) and further developed by Baldursson and Karatzas (1997), can be formalized as follows. An entering firm that thinks the current capacity $\mathbf{k}^*(\hat{x}_t) = k_t$ remains unchanged in the future solves the entry time from

$$F^m(x_t; k) = \sup_{\tau^m \geq t} E\left[\int_{\tau^m}^{\infty} P(k, x_\tau)e^{-r(\tau-t)}d\tau - Ie^{-r(\tau^m-t)}\right]. \tag{11}$$

Note that the solution to (11) can be expressed as a cut-off rule.

17

Lemma 1 *The optimal solution to (11) can be expressed as a cut-off rule $x^m(k)$, so that the optimal stopping time τ^m is the first moment when x hits $x^m(k)$ from below:*

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

Proof. The problem (11) is a standard exercise problem of a perpetual call option, where the value of the underlying asset at time t is given by

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

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

The following proposition states that the model has an equilibrium that can be found by solving the myopic problem (11) for all fixed values of k.

Proposition 2 *The model has an equilibrium where the entry threshold $\mathbf{x}^*(k)$ is given by the solution to (11). The corresponding capacity path is given by (7). The entry threshold $\mathbf{x}^*(k)$ is increasing in k.*

Proof. By Lemma 1, the solution to (11) for any $k \geq 0$ is a cutoff policy, which we can denote $\mathbf{x}^*(k)$. 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 $\mathbf{x}^*(k)$ is increasing in k. The proof that the solution to (11) constitutes a competitive equilibrium can be constructed 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. ∎

3.2 Extension

One may argue that the capital-replacement frictions in our description depend on the assumption that the new technology is built to co-exist with the old production structure. This is not the case – qualitative results do not depend on the assumption that old units remain in the market.[23] It is easy to see why. Assume that there is a continuum of

---

[23]Note that the new entrants could also leave the market, but given the assumed cost advantage of existing new plants, exit is a more relevant option for old units.

infinitesimal firms, and each active firm has one unit of capital of either type. The new units operate as before but at given time t each fuel-dependent firm 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 incurs a fuel-dependent production cost, as before, plus the flow cost $c > 0$. An idle unit pays just c. An exiting unit pays a one-time cost $I_f > 0$ and, of course, avoids any future costs. Assume further that

$$I_f < \frac{c}{r} < I_f + I,$$

where the first inequality implies that exit saves on unavoidable costs for an old capacity unit, and the second inequality that replacing an old unit by a new unit is costly.

Now, if c is sufficiently low, violating the first inequality above, the old units would never exit in equilibrium, implying that the exit option does not change the equilibrium description of this paper. When the above inequalities are satisfied, then exit will take place in equilibrium but, under sufficient uncertainty, the replacement of old units with new ones is not one-to-one but aggregate capacity must temporarily increase during the transition. This follows since there are real-option frictions also on the exit-side of the market, implying that old units are willing to accept potentially a long period of idleness before leaving the market. This feature is the key for the result that increasing mark-ups and consumer prices are needed for the transition to take place.[24] We refer to our working paper Liski and Murto (2006) for the precise conditions and characterization.[25]

4   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. We believe that the main lessons apply to electricity

---

[24]We discuss this issue further after the application, in Section 4.4.

[25]Methodologically, a distinct feature in comparison to the current paper is that there is a two-dimensional state space due to the capital stocks associated with the two technologies. In particular, the equilibrium concept and the technique for solving it can be seen as generalizations of Leahy (1993) to multiple dimensions.

markets in general because one key property of electricity generation is common across markets: persistent uncertainty comes from fuel prices which is transmitted to entrants profits through thermal electricity generation. Uncertainties from other sources such as wind and hydro power availability, or demand are likely to be idiosyncratic by nature, as they look the same next year or after ten years.

The data and its sources are in the data Appendix available at the authors' webpage.

4.1 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. By several measures, it is relatively tightly integrated cross-border wholesale electricity market, serving majority of the ca. 400 Twh annual demand in the Nordic region (see Amundsen and Bergman, 2006).

Wholesale electricity trade is organized through a common pool, Nord Pool, which is a power exchange owned by the national transmission system operators.[26] 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 a 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, the system price captures well the relevant market price.[27]

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. The solid line in Figure 11 depicts the weekly system price in years 2000-05.[28]

---

[26]For a description of the trading institution and market clearing data, see www.nordpool.com.

[27]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. See the Nordic Master Grid Plan 2008 published by the Organization for Nordic Transmission System Operators (Nordel) for discussion of the bottlenecks and descriptive statistics regarding the degree of market integration. Juselius and Stenbacka (2008) conclude that Finland, Sweden and part of Norway unambiguously pass the competition law test for an integrated market.

[28]To recap the market development, we may call the years 2000-01 as years of abundant availability of

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 seasonal shifts, which we also seek to capture in our estimation procedure detailed later. Table 1 provides the breakdown of the capacity in 2008.

	GW				
	Denmark	Finland	Norway	Sweden	Total
Installed capacity	12.6	17.0	30.8	34.2	94.6
Nuclear	-	2.6	-	8.9	11.5
Fossil fuels	8.8	9.2	.7	5.1	23.9
Renewable	3.8	5.2	30.1	20.6	59.1
-hydro	.01	3.1	29	16.2	48.9
-bio	.03	1.9	.01	2.7	5.0
-wind	3.2	.01	.04	1.0	4.8

Table 1: Installed capacity in gigawatts (GW) by energy source 2008. Annual statistics, Nordel 2008.

---

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) for a detailed explanation and analysis of the price spike.

4.2 Empirical implementation

We estimate an aggregate supply function for the non-hydro generation from data on the weekly system price and total non-hydro output in 2000-05. We regress the weekly non-hydro supply on the price of electricity, the weekly prices of fossil fuels and the time of the year (month). The thermal generation costs vary within the year for reasons related to heating demand and also possibly due to maintenance. To capture these seasonal patterns we include month dummies d_t in the regression equation,

$$q_t^f = \beta_0 + \beta_1 \ln p_t^{elec} + \delta x_t + \gamma d_t + \varepsilon_t, \tag{12}$$

where q_t^f is the thermal supply, x_t is the vector of fuel prices, and t is week. Note that since dummies are defined for months, the estimation effectively produces a monthly supply curve. The generation q^f in this estimation is composed of all other production than hydro, including wind power and the net imports of electricity to the Nordic region.[29]

Kauppi and Liski (2008, section 4.3) estimate the same equation for a different purpose; the detailed estimation results are reported in their Table 2. The exogenous variation in the availability of hydroelectricity allows for a relatively efficient identification of the non-hydro supply; the level of hydro reservoirs are strongly correlated with the price of electricity but not with the cost of thermoelectricity. The hydro reservoirs are thus used as instruments in the estimation. Figure 11 illustrates the fit with observed prices, when actual non-hydro supply and oil prices are inserted to the estimated equation (12) to produce an estimate for the weekly electricity price.[30]

Our strategy is to generate the expected annual revenue for entrants using the estimated supply curve (12). To this end, let q_i^h and q_i^d denote the realized monthly hydro production and final demand, respectively. The difference $D_i = q_i^d - q_i^h$ is the realized residual demand that non-hydro production must meet in the absence of new capital. For k units of new capital, non-hydro production is $D_i - k$, and the implied market price is given by the estimated supply curve (12). We assume that D_i follows a normal distribution with the first and second moments taken from historical data.[31] We denote

---

[29]We also run the supply regression for thermal only, but this has a very small effect on the results; probably so because the marginal plant is a thermal plant in both cases and dummies capture well the seasonal utilization of other capacity types. For clarity, we therefore make the distinction only between hydro and non-hydro capacity.

[30]We drop all other fuel prices than oil from the regression, as this has a minimal effect on the results (the fuel prices are highly correlated).

[31]We construct observations for D_i using data for demand and hydro production over the years 2000-
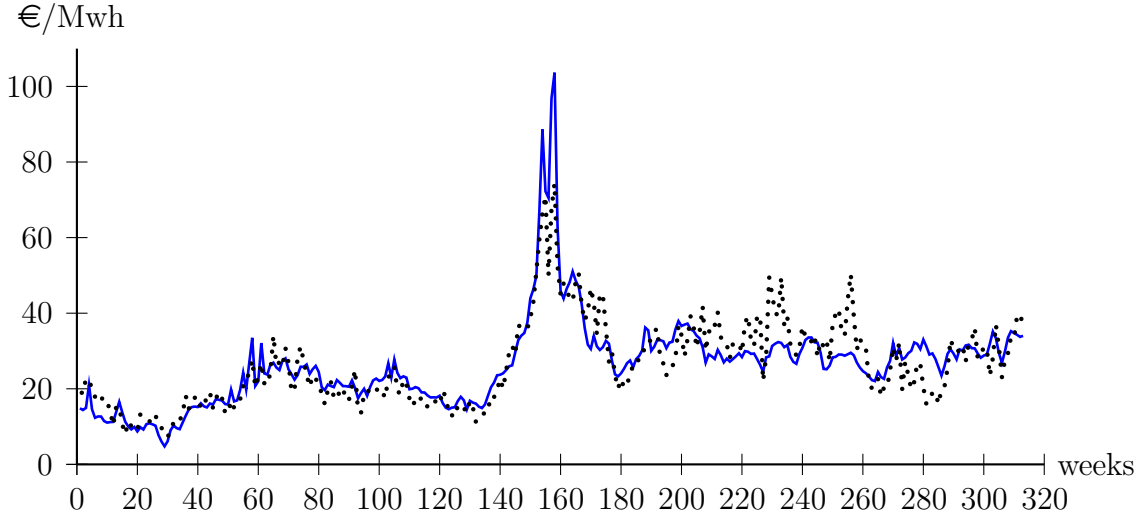
Figure 11: The market price (solid curve), and the price implied by the estimated supply curve (dotted curve). Weekly observations from six years 2000-05.

by M and H the number of months and hours in a year, respectively.[32] The expected annual average price can be expressed as follows:

$$P(k, x) = \frac{1}{M} \sum_{i=1}^{M} \int \Pi(i, x, D_i - k) dF_i, \tag{13}$$

where the monthly price $p_i = \Pi(.)$ is implied by the inverse of (12), and F_i is the cumulative distribution function for D_i in month i. Note that for the expected annual revenue, we must multiply (13) by the number of hours per year, H. Because the uncertainty coming from monthly demand for thermal is idiosyncratic, we can apply $\widetilde{P}(k, x) \equiv H \cdot P(k, x)$ in the subsequent analysis exactly as we used $P(k, x)$ before — all the assumptions made for $P(k, x)$ in Section 3.1 hold for $\widetilde{P}(k, x)$.

Figure 12 depicts the basic properties of the estimated price function $P(k, x)$. 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,

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.

[32]In our computations, one month is exactly 4 weeks, so we have $M = 13$. The number of hours per year is $H = 8760$.
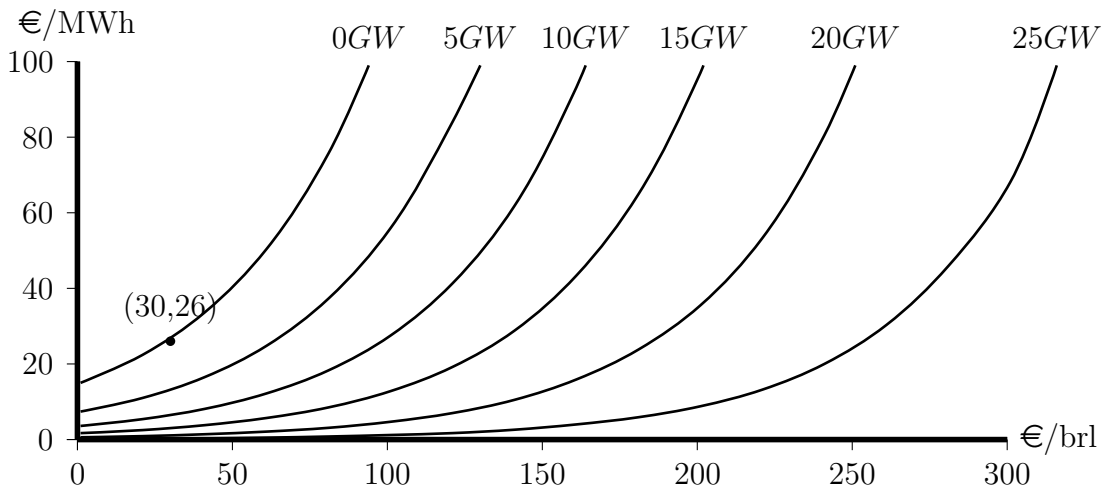
Figure 12: Average annual electricity price (€/MWh) as given by estimated $P(x, k)$. Oil price €/barrel on the horizontal axis. Installed new capacity in GigaWatts indicated for each curve.

which is quite precisely what the case $k = 0$ indicates. In the Figure, we add new capacity in 5000 MW chunks ending at 30000 MW. [33]

The equilibrium is computed by setting up a discrete time version of the continuous-time model, where the fuel price follows a binomial approximation of the geometric Brownian motion with a short time interval between periods (we use period length $\Delta = 1/50$ years). The capacity is also added in small discrete units. The details of the computations, including the Matlab programs and the data files, are available at the authors' web page.

4.3   Simulation results

For the counterfactual simulations, we take the fuel price as given by a Geometric Brownian Motion. Based on Pindyck (1999) and Hamilton (2008), there is no clear trend in long-term oil prices, so we set $\alpha = 0$. Pindyck concludes that the volatility is quite stable and close to $\sigma = .2$ (see his Table 3).[34] We have no single estimate for the investment

---

[33]The total capacity of fossil-fuel fired capacity in this market is 23000 MW; adding this much new capacity does not entirely eliminate dependence on fuel prices because of the idiosyncratic uncertainty in demand. We choose 30000 MW as the upper bound for new capacity in illustrations, because this is close to theoretical maximum entry in computations below.

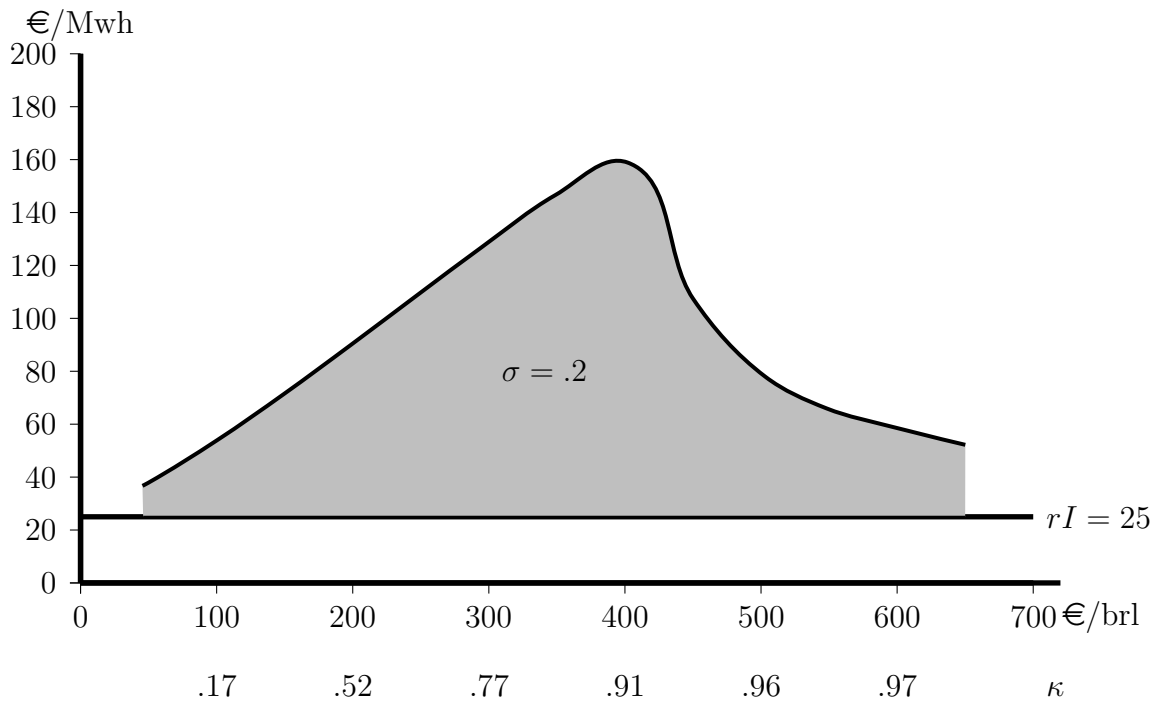[34]See, however, Dvir and Rogoff (2009).

Figure 13: Equilibrium price-cost margin when $rI = 25$ and $\sigma = .2$. κ=fraction of capacity replaced (1=30 000 Mw).

cost rI,[35] but the subsidy levels applied in practice imply that new green capacity can enter when receiving a fixed-price in the range 25 €/Mwh to 80 €/Mwh.[36] We set the risk-free interest rate at 4 per cent.

See first Figure 13 which shows the equilibrium mark-up for new entrants as a function of the fuel price record. This Figure depicts the mark-up at entry points only. In this Figure, we assume $rI = 25$ and $\sigma = .2$. Note that the assumed entry cost is at the low end of the empirical support, and the uncertainty is lower than that characterizing the recent historical fuel prices. Yet, the peak electricity price of 160 € implies more than 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

---

[35]Since we want to express investment cost in units comparable with hourly prices, we must interpret I here as the investment cost for a capacity unit that yields one MWh per year, that is, constant output flow of 1/8760 MW.

[36]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).

25

electricity price is still in the domain of historical observations.

Figure 14 is otherwise the same but the investment cost is taken to be an increasing linear function of the capacity installed such that the first investment cost is $rI = 25$, and the cost after 30 Gigawatts of new capacity is $rI = 80$. The Figure depicts the equilibrium cost as a function of the oil price record; this relationship is not linear. Not surprisingly, higher costs imply higher prices, but otherwise the pattern remains unaltered.

4.4 Discussion

We argue that the results are quite robust in the following sense: the mark-ups for the first entrants, capturing the early friction, are not sensitive to changes in the environment due to our inability to model future events precisely. First, our revenue function may not be precise when oil prices and capacities are far off the empirical support, and therefore we may not describe the profitability of the later entry precisely. This potential mistake has no effect on early entry: 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. This follows from the Leahy's myopia result, as explained in Section 3.

Second, we may not have a good idea of the future development of the entry cost. For example, there can be a sharper than anticipated increase in investment costs as more entry takes place, but this would leave the mark-up required by the first entrant untouched. This follows again from the myopia result. In this sense, the degree of early friction in entry is unaffected by changes in costs of future entry, while the future costs will of course influence the long-run price levels.[37]

Third, we may consider the extended model discussed in Section 3.2, where the old capacity will leave the market gradually over time. If the old capital is costly to maintain, it will be scrapped at some point; the peak consumer price reached during the transition will be lower if fraction of the old capital leaves the market for good. However, the mark-up for the first entrant is independent of the future price path, viewed at the point of entry, and thus it is independent on how exactly the two capital goods interact in equilibrium, provided the old capital is not so unproductive that it decides to exit before new entry takes place.

---

[37]It is perhaps surprising that we can even think of declining entry cost, and the implied friction in early entry increases. Suppose there is a constant investment cost but Poisson arrival rate for a permanent cost reduction. It is straightforward to show that the early mark-up will increase if such a possibility is included.
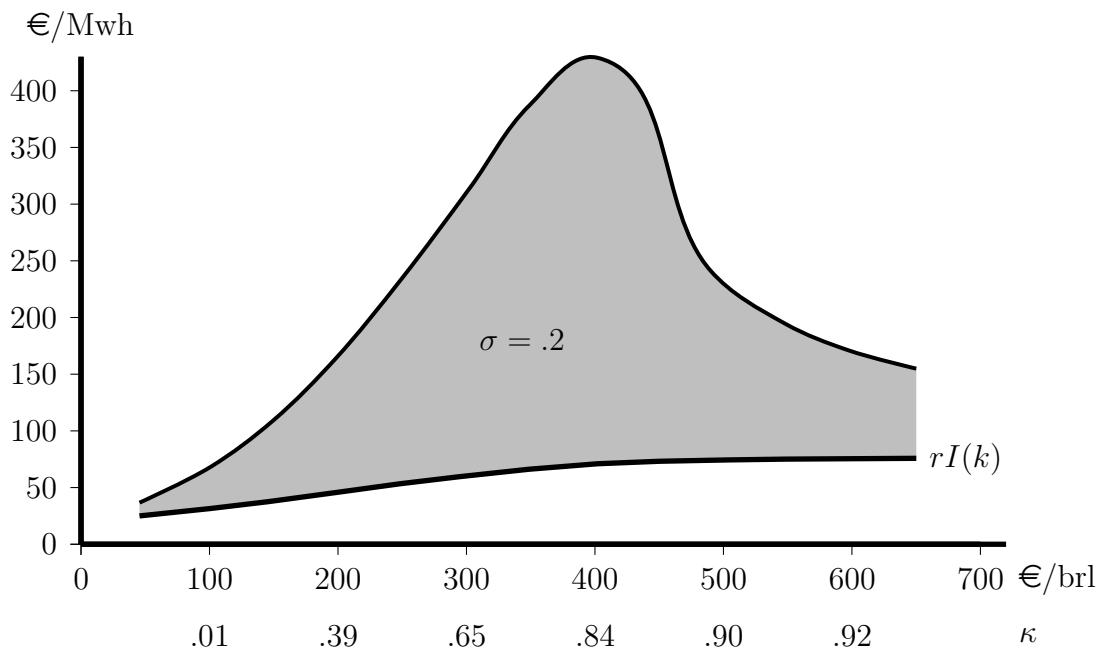
Figure 14: Equilibrium price-cost margin when $rI = 50$ and $\sigma = .2$. κ=fraction of capacity replaced (1=30 000 Mw).

In addition to these elements, it is clear that trends in demand, fuel prices, capital depreciation, or risk aversion of investors can shape how exactly the future prices increase and then finally decline.

Finally, let us discuss a feature specific to the electricity application. Recall that the output should contract in order to create the real-options mark-ups for early entrants. How does this pattern arise in electricity markets where demand and thus total output are relatively inelastic? Obviously, no output contraction is necessarily needed, if the demand is totally inelastic and the supply curve just shifts up with the increasing fuel costs. However, in this market, there can be contraction of the demand without change in total production since large industrial consumers are on both sides of the wholesale market. 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.[38]

---

[38]This description applies well to the paper and pulp industry, for example. 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.

4.5 Policy experiments

Primary energy inputs are often imposing external costs to the society, when their use releases unabated pollutants. If the social cost is fully internalized through a first-best penalty on the use of the inputs, the model description remains valid, provided the source of persistent uncertainty remains. The problem is that the social cost is in most cases not exactly known but, nevertheless, 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 for energy. [39]

Perhaps the most visible 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, 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.[40] We are interested in understanding how this type of uncertainty-reducing instrument influences the entry, the overall surplus, and its division between producers and consumers. To focus on these issues only, we ignore the potential external benefits from a faster transition to greener technologies. The analysis applies to multiple forms of subsidies, but we frame the subsidy as a feed-in tariff to fix ideas. 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

$$\tau < rI.$$

Whenever the final-good price falls below τ, all new producers are compensated for the difference $\tau - p_t$. We assume that the tariff cost is collected from the consumers in a non-distorting manner. 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 enters. For τ sufficiently close to rI, the new technology faces no downside risk, and for τ sufficiently low, the tariff is not effectively influencing the risk and entry. A tariff that is between these extreme levels has interesting implications for the transition.

The change in the consumer surplus due to a tariff is captured by the implied change

---

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

[40]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).

in the total cost of procuring electricity for the consumers. Recall that monthly demand D_i is drawn from distribution F_i. For fixed k and x, the expected total annual cost of procuring the electricity needed to satisfy the demand is given by

$$C(k, x) = H \cdot \frac{1}{M} \sum_{i=1}^{M} \int \Pi(i, x, D_i - k) \cdot D_i \cdot dF_i.$$

In addition to this cost, the consumers must also pay for the subsidies that accrue to the new capacity units through the feed-in tariff. At given k and x, the expected annual subsidy can be expressed as

$$S_\tau(k, x) = H \cdot \frac{1}{M} \sum_{i=1}^{M} \int \max\left[0; \tau - \Pi(i, x, D_i - k)\right] \cdot k \cdot dF_i.$$

Given tariff τ, the total expected cost to consumers at time t is given by

$$\mathbf{C}_\tau(x_t, \hat{x}_t) = \mathbb{E} \int_{s=t}^{\infty} e^{-r(s-t)} \left[C\left(\mathbf{k}_\tau(\hat{x}_s), x_s\right) + S_\tau\left(\mathbf{k}_\tau(\hat{x}_s), x_s\right)\right] ds,$$

where x_t and $\hat{x}_t$ are the levels of the fuel price and its current record at t. The change in consumer surplus due to tariff τ is then

$$\Delta_\tau \mathbf{C}(t) = \mathbf{C}_0(x_t, \hat{x}_t) - \mathbf{C}_\tau(x_t, \hat{x}_t),$$

where $\mathbf{C}_0$ is the expected cost of procurement for $\tau = 0$. For producers, let $W(k, x)$ denote the expected annual profits accruing to the old capacity units at given k and x. This can be written as

$$W(k, x) = H \cdot \frac{1}{M} \sum_{i=1}^{M} \int \left[\int_{q=k}^{D_i} \Pi(i, x, D_i - k) - \Pi(i, x, D_i - q) dq\right] dF_i.$$

The total expected profits for tariff τ evaluated at time t are

$$\mathbf{W}_\tau(x_t, \hat{x}_t) = \mathbb{E} \int_{s=t}^{\infty} e^{-r(s-t)} W\left(\mathbf{k}_\tau(\hat{x}_s), x_s\right) ds.$$

The change in producers' surplus due to tariff τ is then

$$\Delta_\tau \mathbf{W}(t) = \mathbf{W}_0(x_t, \hat{x}_t) - \mathbf{W}_\tau(x_t, \hat{x}_t),$$

and the change in the total surplus is

$$\Delta_\tau \mathbf{C}^{soc}(t) = \Delta_\tau \mathbf{C}(t) + \Delta_\tau \mathbf{W}(t).$$
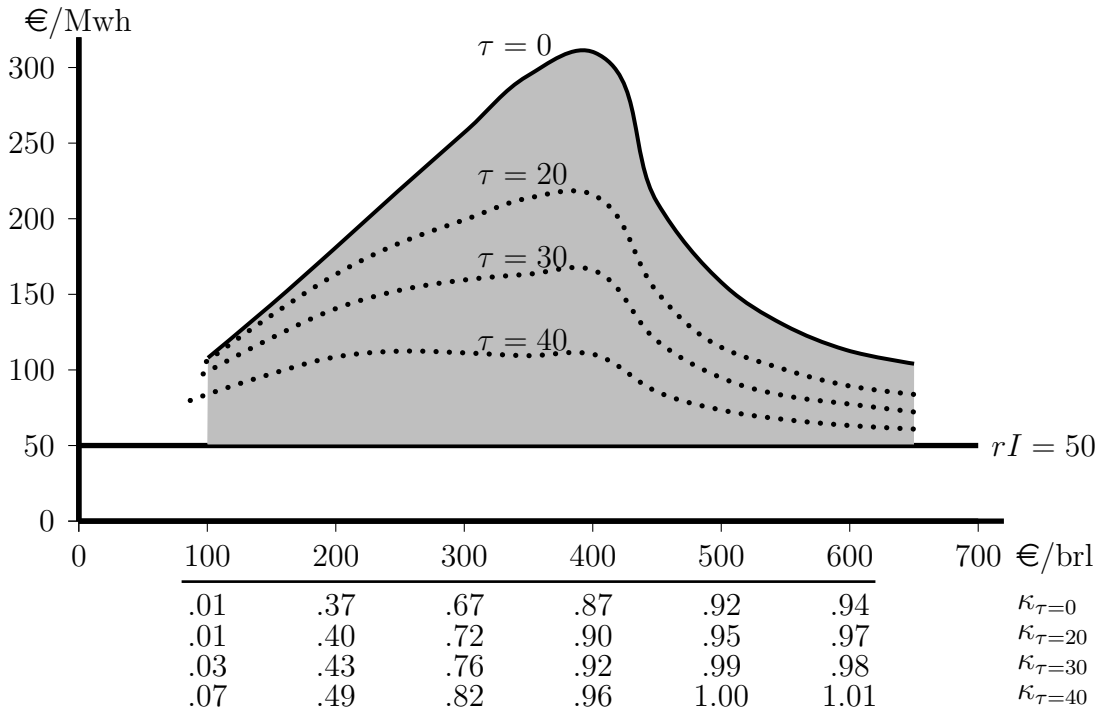
29

Figure 15: 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).

This last expression must be negative since it gives the social cost due to the distorted investment path that the tariff induces, calculated at some t.[41] The details of the simulations below are in the computational Appendix (available at the authors' webpage, as well as the Matlab code).

Consider first how the tariff changes the investment path and the development of the real-options mark-up. In Figure 15, we depict the equilibrium entry mark-up for different tariff levels when the investment cost and uncertainty are $rI = 50, \sigma = .2$, resp. 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 (indicated by the numbers under the Figure).

Let us then consider the changes in the consumer surplus $\Delta_\tau C$, the old producers'

---

[41]Note that the tariff is only one way of implementing a faster replacement of capital. We could choose any distorted capacity path such that $\widetilde{k}(\hat{x}) \geq k_0(\hat{x})$, and compute the implied subsidy from the requirement that the entrants make exactly zero expected profit with the subsidy.
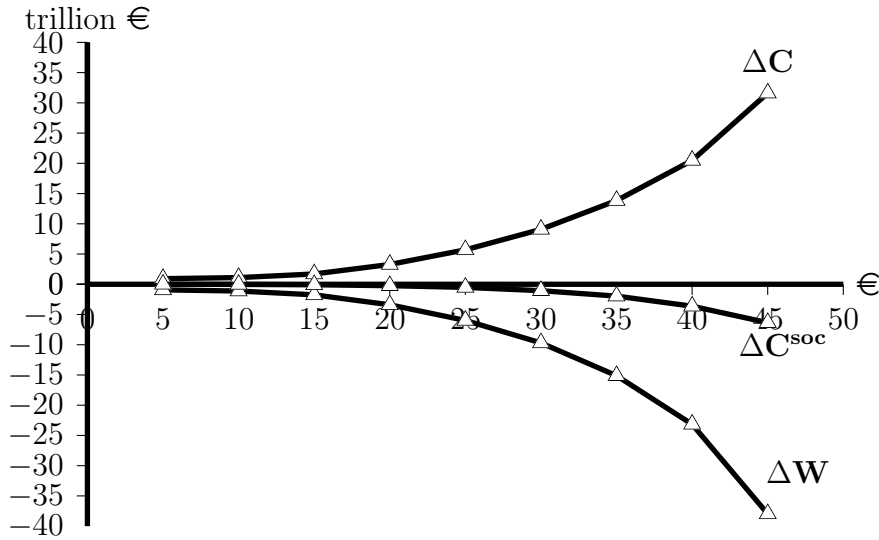
Figure 16: Consumers' gain ($\Delta\mathbf{C}$), producers' loss ($\Delta\mathbf{W}$), and the total welfare loss ($\Delta\mathbf{C^{soc}}$) for tariff levels $\tau = 5, 10, 15, ..., 45$. All expressions evaluated at the fuel price of the first entry when $\tau = 45$.

surplus $\Delta_\tau\mathbf{W}$, and the social loss $\Delta_\tau\mathbf{C}^{soc}$. In Figure 16 we depict these changes as a function of the tariff level $\tau = 5, 10, 15, ..., 45$. All expressions are calculated at $t = 0$, where $x_0 = \hat{x}_0$ is chosen to be the first entry point of the most front-loaded capacity path induced by the largest tariff $\tau = 45$.

The subsidization of new entry must obviously destroy a fraction of the rents of the existing production structure. However, the overall social loss is remarkably small relative to the transfers between consumers and producers, while the absolute sums are large (trillions €). The main result is that consumers can appropriate the old capital rents by destroying relatively little of the total surplus, even in the absence of un-modeled benefits of front-loading the investments.[42] While it is not unheard of that tariff protection or subsidies can benefit either consumer or producer side at the expense of overall welfare, the transfer of wealth is relatively sensitive to uncertainty-reducing policies in this dynamic context.[43]

---

[42]Consumers can also 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).

[43]To better understand the magnitudes, let us transform the consumers' gain to annuity, i.e., to a number that gives the annual average saving in the cost of procuring the electricity. Dividing this

5 Concluding remarks

Our research 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.[44] 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.[45]

Yet another step towards general equilibrium relates to the macroeconomic effects of the energy demand change. Macroeconomists have found it puzzling that the oil prices 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 induce. 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 interesting to have a quantitative assessment of the effects identified in this

---

sum by the annual expected consumption indicates the reduction in the price of MWh. This saving is $.2, .3, .5, ..., .5.6, 8.6$ € per MWh for tariff levels $\tau = 5, 10, 15, ...40, 45$, respectively. Thus, the largest saving, 8.6 € is about 17 per cent of the investment cost level $rI = 50$.

[44]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 transition, capturing similar features, but arising from persistent uncertainty in the energy input supply.

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

paper in a macroeconomic context.[46]

Finally, from the more applied perspective it would be interesting to analyze the real-option frictions simultaneously in several deregulated wholesale electicity markets such the California, New England, and the Pennsylvania,New Jersey, and Maryland (PJM) markets. Our model and computational procedures provide a unified framework for studying implications of different market structures for the green technology transition. The data for such an undertaking is readily available (see Bushnell et al., 2008).

References

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

[2] Atkeson, A., and 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] Baldurson, F., and I. Karatzas, 1997. Irreversible Investment and Industry Equilibrium. Finance and Stochastics 1, 69-89.

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

[5] 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.

[6] Berndt E.R., 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.

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

[8] Bushnell, J.B., E. T. Mansur, and C. S. Saravia, 2008. Vertical Arrangements, Market Structure, and Competition: An Analysis of Restructured US Electricity Markets. American Economic Review, Vol. 98, No. 1, pp. 237-266.

---

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

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

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

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

[12] Dvir, E., and K.S., Rogoff, 2009. Three Epochs of Oil, NBER Working Paper No. 14927.

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

[14] M. von der Fehr, N.-H., E. Amundsen, and L. Bergman, The Nordic market: signs of stress? The Energy Journal, Special Edition on European Electricity Liberalisation, July 2005.

[15] 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.

[16] Hamilton, J. D. 2008. Oil and the Macroeconomy, in New Palgrave Dictionary of Economics, 2nd edition, eds. Steven Durlauf and Lawrence Blume, Palgrave McMillan Ltd.

[17] Hamilton, J. D. 2009. Understanding Crude Oil Prices, Energy Journal, vol 30, no. 2, pp. 179-206.

[18] 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.

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

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

[21] Jaffe A.B., and R.N. Stavins, 1995. Dynamic Incentives of Environmental Regulation: The Effects of Alternative Policy Instruments on Technology Diffusion. Journal of Environmental Economics and Management 29, 43-63.

[22] 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.

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

[24] 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.

[25] 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.

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

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

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

[29] Newell, R. G, A. B. Jaffe, and R. N. Stavins, 1999. The Induced Innovation Hypothesis and Energy-Saving Technological Change. The Quarterly Journal of Economics, vol. 114, No. 3, 941-975.

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

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

[32] Pindyck, R. S., and J.J. Rotemberg, 1983. Dynamic Factor Demands and the Effects of Energy Price Shocks. American Economic Review, vol. 73, issue 5, 1066-1079.

[33] Popp,.D.C., 2001. The Effect of New Technology on Energy Consumption. Resource and Energy Economics, vol. 23, issue. 3, 215-39.

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

[35] 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.

[36] 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.

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

[38] Wolfram, C. 1999. Measuring Duopoly Power in the British Electricity Spot Market. American Economic Review, 89, 805–26.