Policy Uncertainty and Real Options in Switching of Peak Generators

Stein-Erik Fleten[a,*], Marius Johansen[a], Alois Pichler[a], Carl J. Ullrich[b]

[a]*Norwegian University of Science and Technology, Norway*
[b]*James Madison University, Virginia, USA*

Abstract

This paper examines empirically how economic factors, government policy, and strategic interactions affect manager's decisions to switch between operating and stand-by states for peaking electric power generators. We model the switching decisions using a structural model of a dynamic optimal decision game. We focus on the power markets in the Northeastern United States, where annual observations of such decisions are available. The results indicate that regulatory uncertainty significantly increases firms' perception of switching costs, and that large power producers are noticeably more influenced by their economic environment during their decision-making than small firms.

Keywords: Structural Estimation, Stochastic Optimization, Clustering Analysis, Peak Capacity, Regulatory Uncertainty

1. Introduction

Reliability adequacy in the power grid is fundamental for most economic activity. Energy-only markets suffer from an inherent market failure because (i) electricity cannot be stored and must thus be produced at the same time as it is consumed, and (ii) demand-response mechanisms are not yet sufficiently in place so that consumers can act on, or even see, real prices (cf. Cramton and Stoft [7]). The result is that power systems can only ensure reliability adequacy by having more capacity available at any given time than what is reasonable to expect that the market will demand. This makes the existence of flexible peak generators with the ability to quickly ramp up and down production, such as gas-fired combustion turbine-based generators, vital for power markets.

Despite their importance, peak generators have traditionally received little or no compensation for staying in an operation-ready state, and only received revenues by competing in the spot market, cf. Joskow [18], Bowring [3]. As peak generators have the highest marginal costs among electricity generating sources (cf. Energy Information Administration [13]), they only produce during times of shortage when prices are high. And even in such instances, consumers are protected by price caps. This has led to the so-called *missing money problem*, where market participants have been unable to earn sufficient surplus to justify keeping generators in a continuous operating-ready state or building new capacity.

This has resulted in many Independent System Operators (ISOs), particularly in the United States, establishing separate capacity markets alongside the traditional spot and reserve markets. Examples include the Reliability Pricing Model (RPM) by Pennsylvania-New Jersey-Maryland (PJM) in 2007 and the

Forward Capacity Market (FCM) by the New England Independent System Operator (ISO-NE) in 2010. The intent of these markets is to ensure that peak generators are motivated to stay ready to generate electricity when needed, despite only actually producing a few hours a year.

Nevertheless, the recent advent of renewable energy sources (RES) such as solar photovoltaics and wind turbines are threatening to outcompete fossil fuel generation as they have a near-zero marginal cost. Particularly in Europe, many gas-fired combustion based power plants have either seen their values drop significantly or been abandoned altogether, cf. Caldecott and McDaniels [5]. Yet, the increased amounts of renewable energy causes additional fluctuations in the power grid due to the intermittent nature of weather conditions (Duic and Carvalho [11], Lund [23]), which increases the existing need for flexible units. Regulators thus face a significant challenge: as thermal peak generators are increasingly motivated to enter states of temporary or permanent shut-down, the regulators' need for them to remain operational increases.

This paper aims to gain further insight into this problem by investigating the operational decisions made by thermal peak generators to transition between being operational-ready or shut down, either temporarily or permanently. Using the terminology of the U.S. Energy Information Administration (EIA), we consider that a peak generator is always in one of three distinct *operating states*, defined below.

Operating state (OP): The generator has the *ability* to initiate production on short notice, but is not required to actually produce electricity. It holds the option of entering the stand-by state for an irreversible one-time investment cost.

Stand-by state (SB): This state is also known as mothballing or temporary shut-down state. The generator saves on maintenance costs but forgoes revenues as it cannot initi-

---

*Corresponding author. Ph: +47 73591296.
Email address: stein-erik.fleten@iot.ntnu.no (Stein-Erik Fleten)

ate production. It holds the options of entering the operating or retirement state for irreversible, one-time investment costs.

Retired state (RE): The generator is abandoned and cannot become operational again.

The operational decisions under consideration are to remain in the actual state (OP, SB or RE), or the specific action of *shutting down (OP→SB)*, *starting up (SB→OP)* and *abandoning (SB→RE)* individual generators. These are known collectively as *switching decisions*.

It is assumed that these acts of switching come with real irreversible costs.[1] As such, switching decisions are investment decisions where the intent is to maximize prospective profits given one's belief of the future. The decision-making therefore requires the consideration of strategic and forward-looking economic factors.

This paper empirically analyses the effects such factors have on the switching decisions for peak generators in the Northeastern United States in the period 2001–2011. The analysis is made possible by Form 860, which is collected annually by the EIA and includes observed operating states. The decision-making process is constructed as discrete-time and discrete-space dynamic decision game, where these economic factors are included as explanatory variables. Economic primitives such as maintenance and switching costs are retrieved using structural estimation.

The theoretical principles of dynamic decision models date back to Bellman [1] and the concepts of dynamic programming and optimal control policies. Such models characterize the optimal decision of the problem mathematically through a set of Bellman equations, conditional on an exogenous state process. The papers Brennan and Schwartz [4] and Dixit [9] extended this framework by assuming that the exogenous state process behaves stochastically, founding what is today known as *real option theory*.

The technique of structural estimation originates from Rust [26], who studies the problem in which bus companies must decide whether to replace or repair the bus engines in their fleet vehicles. In this case the external analyst is faced with incomplete information as he can observe the decisions made by the bus companies, but not the underlying primitives of the Bellman equations, which explain how the decision-making process is governed.

In this paper we develop a structural model based on these principles, upon which we use the observed switching decisions of the peak generators to estimate the economic primitives describing *how* firms make such decisions. We employ elements of the non-parametric structural estimation approach developed in Fleten et al. [15], because it is ideally suited for complex dynamic decision problems as it allows for the simultaneous comparison of an arbitrary number of options. An important departure from their approach, however, is the treatment of the stochastic expectation of the exogenous state process, where we employ clustering analysis in an effort to address the challenge of dimensionality. The results demonstrate that this is a useful technique in a high-dimensional state space.

This paper adds to the empirical analysis utilizing high-resolution data on the behavior of individual industrial agents. The papers by Frayer and Uludere [16] and Caldecott and McDaniels [5] provide estimates of the value of owning thermal peak generators, both in light of its inherent option value due to its production flexibility and the challenges related to the advent of renewables in the market. Others, e.g., Hiebert [17], Knittel [21] and Craig and Savage [6] explore how market restructuring motivates firms to improve the thermal efficiency of their generators. Lin and Thome [22] investigate the dynamic entry/exit game of ethanol plants using a structural model partly based on a semi-parametric estimation technique developed by Pakes et al. [25]. Their game-theoretical aspect serves as a basis for the construction of a competition variable in this paper, which takes into account how the *competitors* of an individual generator choose to switch between operating states. Fleten et al. [14] specifically addresses switching decisions of peak generators, and evaluates the empirical effect of a number of economic factors on these decisions using a reduced-form regression model.

The contributions of this paper include uncovering primitives linking the economic factors under consideration directly to the *costs* of switching, as opposed to estimating their effects on the *probability* of switching as previously done in the literature. Additionally, this paper investigates whether or not generators take into account the switching actions of their competitors when making their own decision, a strategic element not properly addressed in the literature. Finally, we evaluate whether or not payments from the young capacity markets affect the generators' susceptibility to these economic factors using temporal sample splits. Portfolio effects are similarly addressed.

Outline of the paper. Section 2 presents the theoretical model for the switching problem. Section 3 presents the case data and definitions of the state variables. Section 4 outlines the structural model while Section 5 provides the results and a discussion. We conclude in Section 6.

2. The Theoretical Switching Problem

This section formulates a mathematical model for peak generator switching under the assumption that the decision makers act rationally and aim at maximizing individual profits.

Assume that a generator-owner is allowed a decision at every discrete stage t to switch from operating state s in period t to the operating state u in the following period $t+1$. The switching decision is based on the economic and technical information available to the owner at that time.

The time-varying vector X_t represents the state process holding all publicly known exogenous factors describing the economic environment in which the generator and its competitors

---

[1]Examples of real costs of switching could be the process of hiring/letting go of/reassigning workers and re-negotiations of contracts with suppliers.

are operating. Such factors include relevant political regulations and measures for the uncertainty of the profitability of the generator with respect to competitiveness and prices. We assume that these variables evolve according to a controlled Markov chain.

Unobserved heterogeneity refers to how individual agents act differently when presented the same information. We involve unobserved heterogeneity by including an i.i.d. stochastic state process ϵ_t, that represents the private information available to each decision-maker, but inherently unobservable to an external analyst. Examples could include technical conditions of the generator, the particular cost structure of the firm, and the decision maker's willingness to take risks.

The profit for a generator is clearly dependent on its current operating state s and state process X_t, but also on the choice u regarding the operating state next period, as this decides potential profits for the coming year. Further, various maintenance and switching costs are important. The annual net profit can thus be formulated as

$$g(x_t, s_t, u_t, \epsilon_t) = g(x_t, s_t, u_t) + \epsilon_t, \qquad (1)$$

where the private information ϵ_t acts as a *profit shock*. The particular decomposition (1) is according Rust [26] and known as *additive separability*. Here ϵ_t is assumed to be conditionally independent from X_t. The value of owning a generator can thus be represented as the expectation of the sum of all discounted future net profits, provided that the owner makes the optimal decision u each period. This is given by

$$V(x, s, \epsilon) = \max_{u \in S} \left[\mathbb{E}\Big(\sum_{t=0}^{\infty} \beta^t g(X_t, s_t, u_t, \epsilon_t) \Big| X_0 = x \Big) \right], \qquad (2)$$

where the dynamically optimal decision policy is given by whichever action u that maximizes (2). This expression can be reformulated to

$$V(x, s, \epsilon) = \max_{u \in S} \left[g(x, s, u, \epsilon) + \beta \mathbb{E}\Big(\int V(X_1, u, \epsilon_1)\mathcal{E}(d\epsilon_1) \Big| X_0 = x \Big) \right], \qquad (3)$$

as demonstrated in Fleten et al. [15]. Eq. (3) represents the familiar Bellman equation and acts as a necessary condition for optimality in a dynamic programming problem. By defining the *expected value function* as

$$v(x, s) = \mathbb{E}\Big(\int V(X_1, s, \epsilon_1)\mathcal{E}(d\epsilon_1) | X_0 = x \Big), \qquad (4)$$

Eq. (3) becomes

$$V(x, s, \epsilon) = \max_{u \in S} \left[g(x, s, u, \epsilon) + \beta \, v(x, u) \right], \qquad (5)$$

and by taking the expectation of (5) we obtain an alternative expression for (4), i.e.,

$$v(x, s) = \mathbb{E}\left[\int \max_{u \in S} \left[g(X_1, \epsilon_1, s, u) + \beta v(X_1, u) \right] \mathcal{E}(d\epsilon_1) \Big| X_0 = x \right], \qquad (6)$$

which is a fixed point equation for the function v.

The integrand in (6) is a maximum of random variables. Recall that a maximum of random variables is an extreme value distribution which must be a Gumbel distribution in our case, as ε has positive and negative tails (the Fisher–Tippett–Gnedenko theorem[2]). A closed form expression for the Gumbel distribution under maximization is given by

$$\int \max_{u \in S}(\epsilon_u + c_u)\mathcal{E}(d\epsilon_u) = b \cdot \log\Big(\sum_{u \in S} \exp\Big(\frac{c_u}{b}\Big) \Big), \qquad (7)$$

(cf. Appendix B or Fleten et al. [15, Proposition 8] for further details), where b is the scale parameter of the limiting Gumbel variable. By specifying $c_u := g(X_1, s, u) + \beta v(X_1, u)$, Eq. (6) reduces to

$$v(x, s) =$$
$$\mathbb{E}\left[b \cdot \log\Big[\sum_{u \in S} \exp\Big(\frac{g(X_1, s, u) + \beta v(X_1, u)}{b}\Big) \Big] \Big| X_0 = x \right]. \qquad (8)$$

For notational convenience we introduce the operator

$$t_g(v)(x, s) :=$$
$$\mathbb{E}\left[b \cdot \log\Big[\sum_{u \in S} \exp\Big(\frac{g(X_1, s, u) + \beta v(X_1, u)}{b}\Big) \Big] \Big| X_0 = x \right], \qquad (9)$$

which allows expressing the Bellman equation abstractly as a fixed point equation as

$$v = t_g(v). \qquad (10)$$

Eq. (10) represent a constraint in the structural model in Section 4 below.

3. Data and State Variable Definitions

The primary data source used is Form 860, which is annually collected by the EIA and which contains generator-level specific information about existing and planned generators in the United States, including recorded operating states. Supply and demand of electricity is collected from the *Electricity Supply and Demand* (ES&D) database, which is annually published by the North American Electricity Reliability Corporation (NERC). Commodity prices are collected from the New York Mercantile Exchange (NYMEX) and the wholesale electricity market system operators' websites.

The following subsection outlines the temporal and geographical focus addressed in this paper. The efficiency of a unit, i.e. heat rate plays a dominant role in describing the performance of thermal generators (Section 3.2), it is part of the state variable X_t detailed in Section 3.3 below.

---

[2]The three extreme value distributions are the Gumbel (Type I), the Fréchet (Type II) and the Weibull (Type III) distributions.

3.1. Time Frame and Focus Region

This paper focuses on peak generators on the Northeastern United States over the period 2001 to 2011. Specifically, we include data from three major wholesale electricity markets, the *Pennsylvania-New Jersey-Maryland (PJM)*, the *New England Independent System Operator (ISO-NE)*, and the *New York Independent System Operator (NYISO)*, displayed in Figure A.3 in the Appendix. This includes generators in the following states: Connecticut, Delaware, Illinois, Indiana, Kentucky, Massachusetts, Maryland, Maine, Michigan, North Carolina, New Hampshire, New Jersey, New York, Ohio, Pennsylvania, Rhode Island, Tennessee, Virginia, Vermont and West Virginia. Additionally, the District of Columbia is included.

We focus on the time period starting in 2001 because this represents the majority of the available data on switching decisions after significant changes were made to Form 860 in 2001, making previous data unsuitable for comparison in this analysis.

With 21 U.S. states and 11 years of observations, there is a total of 13 078 generator-year observations of performed switching decisions.[3] This corresponds to 1 388 unique generators owned and managed by 332 companies.

3.2. Heat Rate

During times of peak demand, electricity is close to being a homogeneous product. As a result, the profit of a peak generator is heavily dependent on its marginal cost of production, which in turn is largely determined by its efficiency. Even in markets with no retail competition, where producers traditionally are able to offset high marginal costs by charging consumers more through regulated prices, efficiency is becoming increasingly important, as several U.S. states have introduced incentive regulation where utilities are rewarded for high efficiency.[4]

The industry measure of efficiency for thermal generators is the *heat rate*, which measures how much fuel in MMBtu (millions of British Thermal Units) is needed to produce one MWh of electricity. A high heat rate thus corresponds to a low efficiency. As there is 3.41275 MMBtu in one MWh, the relationship between efficiency (E) and heat rate (H) is given by

$$H = \frac{3.41275}{E}100\,\%.$$

Instead of using nameplate efficiency, this paper relies on heat rates calculated from actual data on consumed fuel input and electricity output. This is advantageous when modeling a strategic investment game, as the heat rates can change over time due to (i) deterioration due to attrition and age and (ii) improvements from technological upgrades.[5] The heat rates based on actual data capture the sum of these two effects. The literature indicates that thermal efficiency has increased by 2 %–50 %-points from the early 1980s to the late 1990s due to regulated heat programs and introduction of retail competition (cf. Hiebert [17], Knittel [21]). This is comparable to our own data, where a generator's maximum recorded heat rate is on average 25 % higher than its minimum.

3.3. State Variable Definition

The cash flow of a peak generator is determined by its spark spread, the difference between the price of the electricity sold and the cost of the fuel consumed to produce it. The spark spread $Y_{n,d,r}$ (in \$/MWh) for generator n, located in wholesale market r, for a particular day d, is

$$Y_{n,d,r} = p^e_{d,r} - H_n \cdot p^f_{d,r} - \upsilon_{n,d},$$

where $p^e_{d,r}$ and $p^f_{d,r}$ are the electricity price (\$/MWh) and fuel price (\$/MMBtu) for day d in region r, and $\upsilon_{n,d}$ is the variable non-fuel generation cost.

The first state variable under consideration is profitability per unit of capacity (\$/kW). It is given by

$$P_{n,t,r} = \sum_{d=1}^{T} \max(Y_{n,d,r}, 0) \cdot \left(\frac{16}{1000 \; kW/\,MW}\right), \qquad (11)$$

where T is the number of days in year t and 16 is the number of peak hours in a day.[6] Note that the max operation in (11) indicates that for days when the spark spread is negative, the plant chooses not operate and the profit is zero. This means that owning a peak generator is comparable to holding a collection of daily European call options on the spark spread, cf. Deng et al. [8].

The second state variable under consideration is related to the uncertainty of the spark spread. Real option theory states that investors are likely to hold potential investment decisions during times of uncertainty given the option to make the decision later (Dixit and Pindyck [10]). Hence, the spark spread standard deviation is likely to affect switching decisions as they involve irreversible investment costs. Using the definition of the spark spread given above, its standard deviation is[7]

$$S_{n,t,r} = \mathsf{std}(Y_{n,d,r}), \qquad d = 1, \ldots, T. \qquad (12)$$

---

[3]In the sample there were recorded 77 transitions directly from OP to RE. This behaviour is inconsistent with normal economic incentives for the peak generators, and the observations are likely caused by reasons other than economics, e.g., catastrophic failures or environmental concerns. Hence, they have been removed from the sample.

[4]For example, regulated prices have been set conditional on firm-level average efficiency, providing efficient utilities with a higher price-cost margin, cf. Craig and Savage [6], Knittel [21].

[5]An example is how refined components can lead to more accurate damper control of boiler temperature, which in turn helps optimize the combustion and hence the generator's heat rate [6].

[6]As the focus is on peaking generators, the electricity prices are taken from the peak hours of the day, i.e., 06:00–22:00 (HE7-HE22 in industry parlance). Daily peak prices are calculated as the simple average of the hourly spot prices during the peak period.

[7]It would be advantageous to rather use forward-looking data for estimating this variable, e.g. using implied numbers based on market-observed prices of spark spread options. Such an approach was used successfully by Kellogg [19] for oil volatility. However, as such options failed to attract volume on the exchanges, the available data is insufficient.

As switching decisions are in part bets on the future, firms are likely to consider forward-looking measures of expected profit. A suitable proxy for future profits is a measure of future demand for capacity. The capacity buffer required by system operators is formally known as *reserve margin*. If the reserve margin for the following period is projected to be low, the demand for the peak generators' capacity that period should be high, and consequently generators should be motivated to remain in, or switch to the operating state. The reserve margin is known by several definitions in the literature, though the one consistent with the peak generator literature (cf. Fleten et al. [14]) is

$$R_{t,r} \equiv \frac{(c_{t,r} - d_{t,r})}{d_{t,r}}, \qquad (13)$$

where $c_{t,r}$ represents the planned capacity and $d_{t,r}$ the projected peak demand for electricity in region r in year t.[8]

A power producer participating in the spot market is competing against other producers. As such, the profit for a single generator is dependent on whether its competitors are also participating in the same market (i.e., are in the operating state) at the same time, and how efficient the generator is compared to these active competitors. If a firm owns a generator that holds a competitive advantage over other power producers, the firm is likely to sell electricity even during times of low demand because it can underbid the competition. This is likely to affect switching decisions. For this reason, we involve a dynamic variable that measures the relative efficiency of a single generator compared to competing generators in the same U.S. state for a particular year. That is, the measure of *inverse competitive advantage* for generator n in year t is given by

$$C_{t,n} = \begin{cases} \frac{H_{t,n}}{H_A} & \text{if } n(A) > 0, \\ 0 & \text{if } n(A) = 0, \end{cases} \qquad (14)$$

where A is the set of all generators located in the same U.S. state as generator n that are in the operating state in year t, H_A is the average heat rate of all elements in A and $n(A)$ is the number of elements in A. $C_{t,n}$ is a dynamic variable describing the strategic game between the generators. It captures (i) the internal dynamics of generator n through the temporal changes in $H_{t,n}$ and (ii) the switching decisions of the neighboring generators as the set A only includes generators that are in the operating state in year t. When $C_{t,n} < 1$, generator n has a lower heat rate than its average competition and thus holds an advantage equivalent to a higher price-cost margin. When $C_{t,n} > 1$, the opposite holds true. As $C_{t,n}$ is a continuous variable, relative positions within these two cases are captured as well. The special case of $n(A) = 0$ occurs only one time in the data.[9] Setting the variable to zero in this case implies that competitive concerns do not influence switching decisions, which makes sense seeing as there were no competitors present at that time.

The introduction of retail competition in many U.S. power markets was one of the most significant regulatory changes in the 1990s and 2000s, with significant impact on individual utilities and generators. Retail competition means that consumers can choose the cheapest supplier of electricity available. This is beneficial for efficient generators as they are likely to experience increased demand, whereas less efficient generators are going to face the opposite effect. Hence, this regulatory change

---

[8]Since the annual U.S demand peaks during the summer, peak hour summer demand is used to calculate the reserve margin.

[9]The 16 generators located in the District of Columbia in 2001 were all in the SB state.

Table 1: Summary statistics of state variables observed

Transition type	Owner type	Num. of observ.	P_i[\$/kW] μ	σ	U_i μ	σ	R_i μ	σ	S_i μ	σ	C_i μ	σ
	All	10798	5.148	(5.544)	0.026	(0.151)	0.181	(0.056)	0.030	(0.016)	0.998	(0.299)
$OP \to OP$	Small	4796	5.670	(5.834)	0.023	(0.150)	0.185	(0.047)	0.028	(0.014)	0.970	(0.234)
	Large	6002	4.731	(5.487)	0.028	(0.165)	0.178	(0.067)	0.031	(0.018)	1.021	(0.340)
	All	97	2.740	(4.157)	0	(0)	0.175	(0.048)	0.023	(0.009)	1.031	(0.319)
$OP \to SB$	Small	61	3.586	(4.841)	0	(0)	0.175	(0.053)	0.023	(0.010)	0.992	(0.384)
	Large	36	1.305	(1.964)	0	(0)	0.174	(0.039)	0.024	(0.007)	1.099	(0.137)
	All	234	7.131	(7.267)	0.004	(0.065)	0.177	(0.047)	0.037	(0.026)	1.001	(0.476)
$SB \to OP$	Small	167	8.118	(7.698)	0	(0)	0.180	(0.041)	0.039	(0.029)	0.923	(0.389)
	Large	67	4.669	(5.363)	0.015	(0.122)	0.169	(0.059)	0.031	(0.013)	1.196	(0.603)
	All	1868	4.033	(4.980)	0.026	(0.158)	0.176	(0.048)	0.030	(0.019)	1.132	(0.470)
$SB \to SB$	Small	1340	3.925	(4.983)	0.028	(0.166)	0.176	(0.045)	0.030	(0.019)	1.075	(0.456)
	Large	528	4.309	(4.967)	0.019	(0.136)	0.178	(0.056)	0.031	(0.021)	1.274	(0.476)
	All	81	1.268	(3.016)	0	(0)	0.174	(0.030)	0.021	(0.010)	1.240	(0.157)
$SB \to RE$	Small	18	5.417	(4.380)	0	(0)	0.145	(0.054)	0.033	(0.013)	1.074	(0.162)
	Large	63	0.082	(0.221)	0	(0)	0.182	(0.008)	0.018	(0.006)	1.287	(0.119)

Notes: μ is the mean and σ is the standard deviation. *OP* is the operating state, *SB* the stand-by state and *RE* the retired state. State variables: P_i is the profitability measure, U_i the regulatory uncertainty dummy, R_i the reserve margin, S_i the spark spread standard deviation and C_i the inverse measure of competitive advantage. Results are shown for (i) the entire sample, as well as the sub-samples of observations where the generator is managed by (ii) a small firm and (iii) a large firm. A firm's size is measured by the accumulated summer peak capacity of all its generators. Section 5.2 outlines details of this sample split.

has the potential to immensely alter the competitive situation among power producers. Retail competition is a policy at the U.S. state level. Fleten et al. [14] employ a binary variable in their regression analysis of the switching problem which describes if there is *uncertainty* in the market about whether or not this regulatory change will take effect in a particular state. This indicator is based on a retail competition index developed by Billingsley and Ullrich [2], which translates qualitative data published by the EIA about state-level deregulation into discretized levels describing how far this political process has proceeded in a given state for a particular year. Levels corresponding to there being an political investigation or recommendation for deregulation active in the state Senate, prior to any actual decisions having been made, indicate that power producers will be uncertain with regards to their regulatory environment.

We employ this same variable in our structural model. As an uncertainty measure it is expected to carry real option effects. Thus, the final state variable is

$$U_{t,s} \in \{0, 1\}, \quad (15)$$

a binary variable equal to 1 if there is regulatory uncertainty present in U.S. state s in year t, and zero otherwise.

The subscript i is used to denote a generator-year observation. As this captures all temporal, geographical and generator-specific characteristics, we use this subscript to simplify the notation of the state variables.

The complete state process is finally

$$X_i = \left(P_i, S_i, R_i, C_i, U_i \right), \quad (16)$$

defined by (11)–(15). Table 1 presents summary statistics of these variables.
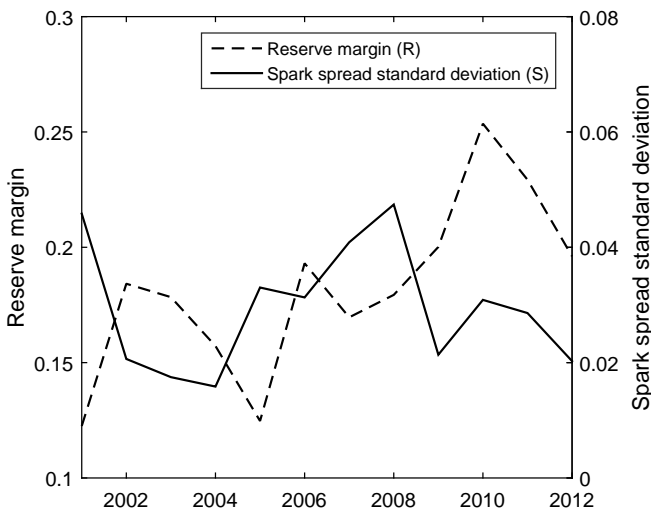
4. Structural Model

4.1. Choice of Model

Structural estimation aims at estimating parameters which are hidden inside a dynamic decision model. Such parameters represent the fundamental economic primitives which we are interested in.

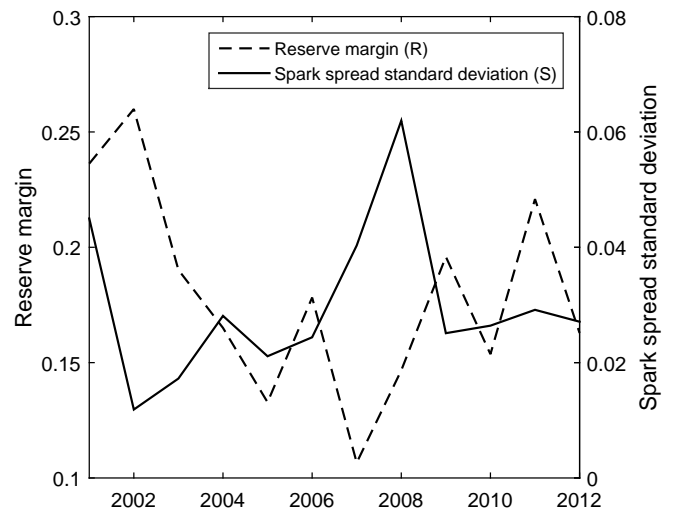Our model relates the three switching decisions directly to economic elements:

  i) the payoff the generator will receive for being in the operating state,

 ii) maintenance costs for being in either the operating state (M_{OP}) or stand-by state (M_{SB}), and

iii) switching costs associated with either shut-down ($K_{OP \to SB}$), start-up ($K_{SB \to OP}$) or abandonment ($K_{SB \to RE}$) decisions.

Understanding the costs of switching is the key to understanding the observed behaviour in the data.

Figure 2 displays the switching decisions of plant managers in our database. Only 3.3% of the observed decisions are switches between different operating states, indicating the presence of significant barriers for such actions. A central principle is that even though there are monetary costs to switching, the power producers' *perception* of these costs is likely to be affected by the economic environment. That is, the switching costs can be interpreted as a combination of monetary cost and risk. Imagine that a plant manager finds it likely that profits will increase next year. This will lower the manager's perception of the risk associated with switching to the operating state, thus reducing his perception of the costs of making such an action.

(a) PJM

(b) ISO-NE

Figure 1: Projected reserve margin (R) and the spark spread standard deviation (S) for the two largest power markets in the sample. Note: The spark spread is calculated using average values for the heat rate.

Conversely, knowing that his generator is very inefficient compared to the other active generators in the area might increase his impression of this cost.

Incorporating this principle into the structural model necessitates distinguishing between two sub-groups in X_i, as the state variables hold different economic interpretations. The first group consists of the profitability measure P_i. As a calculated value it has an inherent economic meaning and can therefore be placed explicitly in the structural model as the payoff a generator receives for being in the operating state.[10] The second group consists of the remaining four variables, which we refer to collectively as the set X_i', i.e.,

$$X_i = (P_i, \underbrace{S_i, R_i, C_i, U_i}_{X_i'}) \tag{17}$$

(cf. (17)). The state variables X_i' are not calculated and can only be incorporated into the model empirically. These can be perceived as the economic environment of the generator, and thereby also to describe the risks associated with switching decisions.
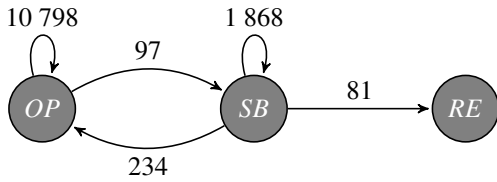

Figure 2: State Diagram for Peak Generator Switching. Included is the number of observations of each switching decision in the full sample.

A plant manager's perception of the switching costs can thus be represented by expressing them as linear empirical functions of the state variables. That is,

$$K_{OP \to SB}(X') = \gamma_0 + \gamma^\top X',$$
$$K_{SB \to OP}(X') = \lambda_0 + \lambda^\top X',$$
$$K_{SB \to RE}(X') = \eta_0 + \eta^\top X',$$

where the constants γ_0, λ_0 and η_0 are the monetary cost of switching, whereas γ, λ and η are four-dimensional vectors of coefficients of the strategic considerations.

Note that when a firm abandons a generator, needed space might be freed and the generator could be sold on the second-hand market. Therefore, the decision of abandoning the generator is likely to cause a cash inflow rather than a cash outflow;

that is, $K_{SB \to RE} < 0$. Parts of the abandonment cost can hence be interpreted as a salvage value.

The annual profit function g (cf. (1), or (2)) thus is

$$g_\theta(X|s, u) = \begin{cases} P - M_{OP} & \text{if } s{=}OP, u{=}OP, \\ \frac{P}{2} - K_{OP \to SB}(X') - \frac{(M_{OP}+M_{SB})}{2} & \text{if } s{=}OP, u{=}SB, \\ \frac{P}{2} - K_{SB \to OP}(X') - \frac{(M_{OP}+M_{SB})}{2} & \text{if } s{=}SB, u{=}OP, \\ -M_{SB} & \text{if } s{=}SB, u{=}SB, \\ -K_{SB \to RE}(X') - \frac{M_{SB}}{2} & \text{if } s{=}SB, u{=}RE, \end{cases} \tag{18}$$

where the decision $s{\to}u$ decides the combination of profit, maintenance and switching costs applicable to the generator that year.[11]

The EIA only records operating states on an annual frequency. Since we cannot know at which time during the year switches are made, we assume that they take place half-way through the year. The profits and maintenance costs are distributed accordingly in (18).

The structural parameters to be estimated are held in θ, i.e.,

$$\theta := (M_{OP}, M_{SB}, \gamma_0, \lambda_0, \eta_0, \gamma, \lambda, \eta).$$

4.2. Optimization Problem

Retrieving the parameter θ can be done by solving a constrained non-linear optimization problem of the maximum likelihood of the observed data, cf. Su and Judd [27]. Since the parameters are only present in $g_\theta(X|s, u)$, the maximum likelihood is centered on this function. That is, the structural estimation procedure becomes solving the problem

$$\begin{aligned} \text{maximize} \quad & \mathcal{L}(g, v_g, (X_i, s_i, u_i)_{i=1}^N), \\ \text{subject to} \quad & v_g = t_g(v_g), \\ & g \in \mathcal{G}, \end{aligned} \tag{19}$$

where $\mathcal{L}$ is the likelihood function of the observed data $(X_i, s_i, u_i)_{i=1}^N$ given a particular profit function g_θ, N is the number of generator-year observations i, and the set of profit functions $\mathcal{G}$ is given in (18). The constraint $v_g = t_g(v_g)$ is the Bellman fixed-point equation (10).

The structural formulation (19) requires a maximum likelihood estimator for the objective function, as well as a treatment of the conditional expectation in the Bellman constraint. The Gumbel variable for the process ϵ allows an explicit expression for the likelihood estimator. The explicit *probability of choice* of the Gumbel variable is

$$P_v(u|x, s) = \frac{\exp\left(\frac{g(x, s; u) + \beta v(x, u)}{b}\right)}{\sum_{u' \in S} \exp\left(\frac{g(x, s; u') + \beta v(x, u')}{b}\right)}$$

(cf. (7)–(9)) as detailed, e.g., in Fleten et al. [15, Proposition 9]. It follows that the corresponding likelihood function is

$$\mathcal{L}\big(g, v_g, (X_i, s_i, u_i)_{i=1}^N\big) = \prod_{i=1}^N P_v(u_i|X_i, s_i). \qquad (20)$$

As evident from Figure 2, some actions occur more frequently in the data than others. Left unchecked, this imbalance causes biases in the optimization as (20) treats every observation equally. For this reason, and consistent with the literature (cf. King and Zeng [20]), the likelihood function is augmented such that each observation is weighted appropriately so that the importance of each type of decision is equated in the estimation. Thus, the augmented and final (log-)likelihood function is

$$\log \mathcal{L}(g, v_g, (X_i, s_i, u_i)_{i=1}^N) = \sum_{i=1}^N \frac{1}{N_i} \log\big(P_v(u_i|X_i, s_i)\big),$$

where

$$N_i \in \{10\,798;\ 97;\ 234;\ 1\,868;\ 81\},$$

are the weights attributed to observation i according to which sub-set in Figure 2 the observation belongs to.

4.3. Clustering Analysis

It is common in the literature to discretize the individual state variables and construct a state space as the combinations of these. This approach poses challenges for a high number of state variables, as the size space tends to become larger than the data can accommodate to obtain statistically meaningful estimates of the transition probabilities. This is known as the curse of dimensionality.

We address this problem by discretizing the data into clusters using k-means clustering analysis. This technique groups the observations (X_i) into k clusters by minimizing the squared Euclidean distance between the observations and the cluster centroids $\bar{x}_c$, where the centroids also serve as reference points for all observations in the cluster. All state variables are properly scaled such that their dimensions are comparable. This avoids unintended favoritism of a particular state variable in the clustering process.

The conditional expectation in (8) is treated by estimating the transition probability of each pair of subsequent observations of a given generator,

$$(X_i, X_{i+1}) \qquad \text{for } i = 1, 2, \ldots N-1.$$

The state process is assumed to behave according to a Markov chain. The transitional probabilities of the discretized state space are estimated empirically. That is, the probability that a state process X_i that is currently in cluster c will pass over to cluster j is given by

$$M_{cj} = \Pr(X_{i+1} \in j | X_i \in c),$$

which constitute the entries in a Markov transition matrix. This formulation implicitly captures any dependency between the individual state variables.

By redefining the Bellman equation in (8) on the discretized state space given by the cluster centroids $\bar{x}_c$, a proper estimator for the constraint is given by the weighted probability of transitioning from cluster c to any other cluster j, i.e.,

$$\hat{t}_g(v)(\bar{x}_c, s) = \sum_{j=1}^k M_{cj} \cdot b \cdot \log\Big[\sum_{u \in S} \exp\Big(\frac{g(\bar{X}_j, s, u) + \beta\, v(\bar{X}_j, u)}{b}\Big)\Big], \quad (21)$$

where $\bar{X}_j$ are the centroid values at which the profit function and discounted expected value functions are evaluated, before being properly weighted by the probabilities in M_{cj}.

Note that the clustering automatically serves as an interpolation of the observations. This has crucial computational advantages, as this limits the number of variables in the optimization problem.

This estimator can be further improved by evaluating the constraint not for the centroid values $\bar{X}_j$ but rather as the average of the actually observed state variables for observations transitioning to cluster j, originating from cluster c. The average is achieved by replacing the outermost sum in (21) with a sum over every observation originating in cluster c, including the weights $w_{c,j}$ which denotes the number of observed transitions from c to j. Thus, the final estimator becomes

$$\hat{\hat{t}}_g(v)(\bar{x}_c, s) = \sum_{i \in Z} \frac{M_{c,j(i+1)}}{w_{c,j(i+1)}} \cdot b \cdot \log\Big[\sum_{u \in S} \exp\Big(\frac{g(X_{i+1}, s, u) + \beta v(X_{i+1}, u)}{b}\Big)\Big], \quad (22)$$

where Z is the set of observations located in cluster c, and $j(i+1)$ denotes the cluster of i's subsequent observation.

The conditional expectation operator $\hat{\hat{t}}$ in (22) completes the mathematical description of the structural estimation problem (19).

4.4. Uncertainty Estimates

The uncertainty of the parameters in θ is found by non-parametric bootstrapping. Observations are randomly drawn with replacement to generate 21 independent data panels. The structural estimation is applied on each panel, and the standard deviation of each parameter is taken across the results from the random samples. Significance is determined by the two-sided student's t-test.

4.5. Identification

Separate identification of the parameters in θ requires that the state variables exhibit independent variation. For the problem under consideration, this is ensured by (i) great cross-sectional variation in the individual state processes as they are defined at either regional (i.e., for each ISO), U.S. state or generator level and (ii) little inter-dependency between the state processes.

The variables most likely to violate this assumption are the projected reserve margin and the spark spread standard deviation. To see the intuition behind this, remember that the spark

8

spread is dependent on the wholesale electricity price. This in turn points to the equilibrium between capacity and demand for electricity, which are the two elements constituting the reserve margin. That is, the spark spread standard deviation is potentially highly correlated with the *present* reserve margin, which could be considered antecedent to the *projected* reserve margin (R). As the reserve margin is a mean-reverting process and can be considered auto-regressive, one would expect the spark spread to be negatively correlated with the projected reserve margin.

A quick glance at Figure 1 shows that there are signs of a negative correlation between these variables. However, a closer inspection reveals that the magnitude of the correlation is small, ranging from -0.27 to -0.54 between the three power markets in the sample, and in only one of the markets is the correlation significant at the 10 %-level. The low correlation can be explained by the influence of the fuel prices (natural gas or crude oil, depending on the generator) included in the spark spread. Also, the projected demands for electricity used in estimating the projected reserve margin rely on standardized temperature values, whereas the present reserve margin uses the actual temperature and weather conditions. This leads to a greater source of variation. Also, the statistics above are calculated assuming average values for the heat rate, which means that the actual data exhibit even more variation than indicated by Figure 1, as the spark spread is generator-specific.

5. Results

In what follows we set the discount factor β set to 0.91, which corresponds to an interest rate of 10%. Further, in consistency with the literature (cf. Rust [26] and Su and Judd [27]) we normalize the scale parameter in the Gumbel distribution by $b = 1$.

5.1. Model Formulations

We address a variant of the model outlined in Section 4 first: the *base model* is used to get an estimate of maintenance and switching costs, and only the profitability measure P is considered in this formulation. More specifically, the base model addresses the available observations with $X = (P, -, \ldots, -)$ in (17). Due to the lower number of parameters it is possible to capture additional unobserved heterogeneity by treating the most important of these, the maintenance cost in the operating state M_{OP} and the start-up cost $K_{SB \to OP}$, as random parameters with a given distribution (cf. Train [28]).[12]

In this way we obtain monetary estimates for the maintenance and switching costs for two sub-samples sorted by the size of the firm managing the generator. M_{OP} and $K_{SB \to OP}$ (in [\$/kW]) are treated as normal random variables, whose mean and standard deviation are displayed in Table 2.

The estimated parameter values θ are the monetary values of maintenance and switching costs with units in [\$/kW]. Such estimates are surprisingly difficult to obtain in the industry due to a wide variety of maintenance policies and managerial priorities across plants, cf. Fleten et al. [15]. Therefore, these estimates hold great value for, e.g., regulators wishing to find suitable levels for capacity payments or other compensations for keeping peak generators in the operating state.

The base model is validated using Monte Carlo simulations.[13] Table 2 shows the result for the base model.

The *augmented model* includes all additional state variables in X. The parameters from this formulation explain how the plant managers' perception of the switching costs is affected by their economic environment. Results of the estimation for the augmented model are given in Table 3.

5.2. Sample Splits

We perform two types of sample splits in order to answer the following questions:

i) Does the firm's size (i.e., portfolio) influence the plant manager's reaction to the economic environment?

ii) Has the recent introduction of capacity payments in the U.S. power markets altered this behaviour?

We capture the portfolio effects by distinguishing the generators by the size of the firms managing them. As a measure of size we use the accumulated peak capacity of all the generators belonging to the firm. The observations are split into two groups; one with 7 398 observations belonging to the 56 largest firms and another with the remaining 7 081 observations of generators belonging to the 276 significantly smaller firms in the total sample.[14] When compared in terms of market shares, these groups hold 56 % (44 %, resp.) of the total peak capacity in the sample.

The effect of the capacity payments is inferred by a temporal sample split. We compare the full sample over the period 2001–2011 with the sub-sample from 2001–2008, which excludes the data after the capacity markets in the PJM and ISO-NE became active.[15] Ensuring identification of this effect is difficult as the results could be influenced by other unknown temporal events. Though, as capacity markets have been considered economic lifelines to peak generators it constitutes a reasonable approximation.

---

[12]The distribution used is discretized and thus managed using representative points (quantizers) with optimal weights such that the distance to the genuine distribution is minimized (cf. Pagès [24]).

[13]Simulated data panels are created using fictional values for the structural parameters in θ. The structural estimation procedure retrieves back the initial parameter values.

[14]The differences in sizes are significant. All of the "small" firms are located in the bottom 3.3% of the total range of accumulated capacity used to measure the size of the firms.

[15]It would be advantageous to perform the estimation on only the sub-sample 2009–2011. However, the amount of data available is insufficient for significant parameter estimation.

5.3. Discussion of Results

The maintenance and switching costs are shown to differ depending on the size of the firms, as seen in Table 2. The maintenance cost in the operating state ($E[M_{OP}]$) is 34% lower for large firms, pointing to significant scale benefits. According to the Norwegian system operator Statnett, two of their thermal peak plants had an average maintenance cost in the operating state of 13.4 \$/kW.[16] This is more than our estimate, an indicator for lower operating costs in the U.S. Maintenance costs in the stand-by state (M_{SB}) are low and of comparable values for small and large firms, demonstrating that when generators are shut-down little resources are spent on their preservation. The cost of abandonment ($K_{SB \to RE}$) is negative and similar in both groups. This is likely because economies of scale have reduced effect on the second-hand market of used generators. The costs of shutting down ($K_{OP \to SB}$) is noticeably lower than the costs of starting-up ($E[K_{SB \to OP}]$), and extremely low for small firms. The sizes of the error estimates ($\sigma[M_{OP}]$ and $\sigma[K_{SB \to OP}]$) points to considerable heterogeneity within the groups.

Moving attention to Table 3; the presence of regulatory uncertainty (U) associated with the introduction of retail competition in the electricity market has a consistent and large positive effect on shut-down and abandonment costs. The uncertainty is thus acting as a barrier to switching. As these decisions involve (dis-)investments, this also implies substantial real option effects.

The reserve margin (R) is significant and holds a consistent and negative effect on shut-down and abandonment costs for large firms, prior to the advent of effective capacity markets, effectively lowering the plant manager's perception of the switching costs when the prospect of future profits is reduced. Conversely, the reserve margin is almost always insignificant for the same firms after the introduction of the capacity markets. This could indicate that the presence of capacity payments overshadows the companies' attention to the reserve margin. The reserve margin is similarly not significant for small firms in any of the time periods considered, which suggests that small firms are less receptive to such projections.

The reserve margin's effect on start-up decisions is more difficult to interpret. The variable is consistently significant in the period 2001–2008, but has also here negative signs, meaning that the projected reserve margin is lowering the cost of starting up a generator. This is counter to intuition, as an expectation of lower future profits (i.e., a high reserve margin) should act as a deterrent against entering the operation-ready state and represent a higher start-up cost. Furthermore, for large firms evaluated over entire sample period the projected reserve margin is significant with a positive sign (i.e., low expected reserve margin leads to higher start-up costs), though this disputes the notion that capacity payments averts the firms' attention from the reserve margin as seen for the other switching decisions. For this reason, the reserve margin's effect on start-up decisions remains ambiguous.

The spark spread standard deviation (S) primarily has a significant and positive effect on start-up and abandonment costs. This uncertainty measure therefore acts as a barrier to switching in a similar manner as the regulatory uncertainty, thus reinforcing the notion that these firms make decisions in accordance with real option theory. Interestingly, the monetary equivalent of S's effect on the start-up cost (when using the respective means of S) is roughly half of its effect on the abandonment decisions. In comparison, the regulatory uncertainty has a monetary effect of 2–4 times that of the spark spread standard deviation.

For the shut-down costs, the spark spread standard deviation is extremely low and insignificant for large firms but surprisingly negative and significant for small firms. Negative coefficients for S represent behavior which is inconsistent with the real option interpretation of this variable, as it would imply that irreversible investments were encouraged under uncertainty. Since owning a generator is similar to owning call options on the spark spread, one could claim that the spark spread standard deviation should increase the profit for the generator as increased uncertainty inherently increases the value of holding such options.[17] However, even with this interpretation one would expect the spark spread standard deviation to increase the perception of the shut-down cost, not to decrease it. It is therefore more likely that the economic environment has an inherently low influence on small firms' shut-down decisions as their monetary costs of this action are very low (cf. Table 2). This could invalidate the structural model for this sub-sample, thus prompting unpredictable results.

The inverse measure of competitive advantage (C) is signif-

[16]See http://e24.no/energi/kostet-en-milliard-ble-aldri-tatt-i-bruk/23437044.

[17]Attempts were made to treat S as an empirical profit measure in conjunction with P. The results, however, gave negative coefficients for S which is incompatible with both interpretations of the variable.

Table 2: Structural parameter estimates for the *base model* formulation.

Parameter type	M_{OP} (heterogeneous)		M_{SB}	$K_{OP \to SB}$	$K_{SB \to OP}$ (heterogeneous)		$K_{SB \to RE}$
Sample type	$E[M_{OP}]$	$\sigma[M_{OP}]$			$E[K_{SB \to OP}]$	$\sigma[K_{SB \to OP}]$	
Small firms	7.14	4.14	1.50	0.00	3.80	1.78	-8.00
	(0.47)	(0.27)	(0.36)	(1.52)	(3.34)	(1.77)	(1.54)
Large firms	4.70	2.72	1.67	2.48	3.84	2.23	-8.42
	(0.47)	(0.27)	(0.39)	(1.75)	(2.09)	(1.21)	(3.16)

Notes: Standard deviations from bootstrapping are in parentheses.

Table 3: Structural parameter estimates for the *augmented model* formulation.

Firm type: Period:	Large 2001–2008	Small 2001–2008	Large 2001–2011	Small 2001–2011
Part of *g(·)-function*	*Coefficient or parameter values*			
M_{OP}				
constant	3.000***	4.316***	2.386***	5.782**
	(0.000)	(0.432)	(0.210)	(1.314)
M_{SB}				
constant	2.088***	0.123	0.000	0.859
	(0.606)	(0.566)	(0.000)	(2.260)
$K_{SB \to OP}$				
constant	0.000	0.917	0.001	0.047
	(0.000)	(1.904)	(0.007)	(0.213)
Regulatory uncertainty (U)	4.150	14.519	6.414	15.966**
	(6.747)	(10.135)	(6.444)	(4.869)
Projected reserve margin (R)	-6.480**	-44.926**	23.210**	-19.902
	(3.010)	(13.731)	(5.642)	(10.398))
Spark spread standard deviation (S)	137.708***	99.822***	104.451**	114.710***
	(25.910)	(21.332)	(32.254)	(22.595)
Competitiveness measure (C)	0.600	2.530	-1.402	1.592
	(0.859)	(1.493)	(1.158)	(1.344)
$K_{OP \to SB}$				
constant	0.000	0.205	0.001	0.176
	(0.000)	(0.326)	(0.003)	(0.725)
Regulatory uncertainty (U)	16.330***	17.907**	15.784**	14.651
	(1.585)	(4.687)	(4.799)	(9.458)
Projected reserve margin (R)	-10.221***	20.981	-9.210	13.831
	(2.297)	(11.576)	(5.665)	(11.075)
Spark spread standard deviation (S)	-5.839	-120.112***	-8.425	-139.506***
	(22.537)	(25.700)	(22.179)	(22.820)
Competitiveness measure (C)	-1.480**	1.593	-2.896*	-1.696
	(0.662)	(2.101)	(1.325)	(3.628)
$K_{SB \to RE}$				
constant	-10.732***	-19.221***	-30.235***	-21.648*
	(1.037)	(4.080)	(2.118)	(9.399)
Regulatory uncertainty (U)	14.758***	18.966***	18.487***	18.413*
	(1.622)	(2.519)	(3.284)	(7.774)
Projected reserve margin (R)	-12.447**	13.138	-7.093	13.003
	(4.768)	(13.746)	(5.142)	(10.262)
Spark spread standard deviation (S)	213.833***	-22.904	309.075***	-45.525*
	(59.608)	(11.349)	(59.748)	(17.863)
Competitiveness measure (C)	-0.193	-0.102	-1.692	-0.810
	(0.637)	(0.887)	(0.905)	(1.118)

Notes: Standard deviations from bootstrapping are in parentheses.
Significance codes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

icant only for the shut-down costs of large firms. With negative coefficients, it shows that large firms are more likely to shut down their generator if the current nearby competing generators are more efficient. This follows the a priori hypothesis.

The lack of its significance for abandonment-decisions could be explained by the dual role this variable holds for such actions. When interpreting C as a measure the generator's inverse competitive advantage in the market, the owner of an inefficient generator should be more likely to abandon it, i.e., it should reduce the value of this switching cost. However, inefficient generators are also worth less on the second-hand market and should thus represent a lower salvage value for the owner, which is equivalent to a higher cost of abandonment. Its insignificant parameter values suggest that these two opposing effects are of comparable sizes.

The small firms are generally less influenced by the state variables than larger firms. This is likely because portions of this group have other priorities than profit maximization, which is a key assumption in the decision-model presented in this paper. Almost 16% of the agents in the sub-sample with small firms are non-utilities such as hospitals and other industrial participants that are not primarily engaging in the energy sector.

Such firms employ the generators to ensure their own emergency reliability in case of blackouts or to facilitate their own in-house need of electricity. The decision-making behaviour of these firms should thus be far more dependent on the individual company itself rather than the larger marked trends (i.e., heterogeneity should be far greater in this group), and is likely the cause of the lack of significant structural parameters for this group.

6. Conclusion

Thermal peak generators play a vital role in maintaining reliability adequacy in power systems. Their importance has recently increased by the escalating proportions of renewable electricity sources (RES) in the market. Yet, due to the low marginal costs of the RES they are also being forced out competitively, which prompts the generators to enter states of temporary and permanently shutdown. This is highly undesired by regulators. This paper examines empirically how economic factors, government policy and strategic interactions affect thermal peak generator's decisions to switch between operating and stand-by states. Specifically, it shows that the economic environment influence a plant manager's perception of the costs associated with such switching decisions.

The results show that the decision-making of power producers is greatly influenced by real option waiting effects. Particularly, uncertainty related to whether or not a U.S. state will implement retail competition in their electricity markets are shown to dramatically increase the costs of switching between operating states. Uncertainty of the profits of the peak generator, the spark spread standard deviation, is shown to have similar impact. These results are particularly relevant for present-day Europe, where substantial investment projects are currently aimed at increasing the transmission capacity between its many individual electricity markets. In order for these markets to become an efficient and coherent place for trading energy, substantial regulation reform will be required, and subsequently uncertainty is likely to hit the markets.

Several capacity markets have been established in the Northeastern United States post 2008. Their intent has been to motivate firms financially to keep their generators operational and able to generate electricity on short notice if needed. Such capacity payments have been believed to significantly alter the way firms respond to economic factors and strategic considerations. We only find such effects for the projected reserve margin, which ceases to be a significant influence on the switching costs after the capacity markets were introduced.

There are also evidence of strong portfolio effects, in which a small group of large utilities are found to be far more responsive to economic factors in their decision-making than smaller firms. Large firms are shown to strategically consider the switching actions of nearby competitors when making their own investment decisions, particularly when considering shutdown decisions. Similarly, the same utilities are found to be influenced by projections of next year's reserve margin when deciding to shut-down or abandon its generators. None of these effects are found for small firms. This implies that a significant portion of

the generating capacity in the market is controlled by firms that might not necessarily produce electricity out of financial motivation. This is relevant information for regulators wishing to create incentives for generators to stay operational.
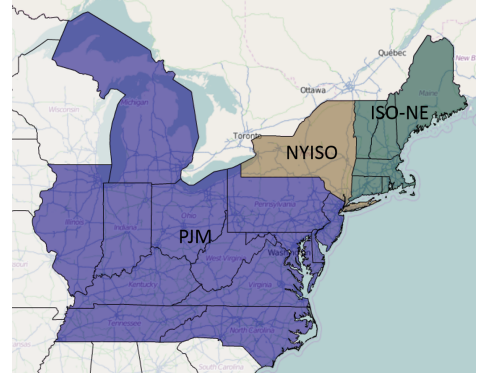
Appendix A. Regions

Figure A.3: The three power markets under consideration. The *Pennsylvania-New Jersey-Maryland (PJM)*, the *New England Independent System Operator (ISO-NE)*, and the *New York Independent System Operator (NYISO)*

Appendix B. The Gumbel variable

For the sake of completeness we state two major properties of Gumbel variables, which are used in the text. A comprehensive discussion of extreme value distributions can be found in Embrechts et al. [12].

The cumulative distribution function (cdf) of a Gumbel distribution is $F(z) = \exp\left(-e^{-\frac{z-\mu}{b}-\gamma}\right)$, where $\gamma = 0.57721566\ldots$ is the Euler–Mascheroni constant. Its mean is μ, and the variance is $b^2 \frac{\pi^2}{6}$.

Proposition Appendix B.1 (The extreme value distribution is closed under maximization). *Let* $(\varepsilon_i)_{i=1}^n$ *be independent Gumbel variables with mean* μ_i *and common scale parameter* $b > 0$. *Then the maximum* $\varepsilon := \max\{\varepsilon_i + c_i : i = 1, \ldots n\}$ *of the shifted variables is again Gumbel distributed with mean*

$$\mathbb{E}(\varepsilon) = \mu := b \cdot \log\left(\sum_{i=1}^n \exp\left(\frac{\mu_i + c_i}{b}\right)\right)$$

and the same scale parameter b, *where* $c_i \in \mathbb{R}$ *are arbitrary constants.*

The following proposition addresses the probability of choice. Again, an explicit formula is available for shifted Gumbel variables.

Proposition Appendix B.2 (Choice probabilities for shifted Gumbel variables). *Let* $(\varepsilon_i)_{i=1}^n$ *be independent Gumbel distributed random variables with individual mean* μ_i *and common scale*

parameter $b > 0$. Then the probability of choice for the variables shifted by c_i has the explicit representation

$$P\left(\varepsilon_1 + c_1 = \max_{i \in \{1,2,...n\}} \varepsilon_i + c_i\right) = \frac{\exp\left(\frac{c_1 + \mu_1}{b}\right)}{\exp\left(\frac{c_1 + \mu_1}{b}\right) + \cdots + \exp\left(\frac{c_n + \mu_n}{b}\right)}. \tag{B.1}$$

References

[1] R. Bellman. Dynamic programming and Lagrange multipliers. *Proceedings of the National Academy of Sciences of the United States of America*, 42(10):767–769, 1956.

[2] R. S. Billingsley and C. J. Ullrich. Regulatory Uncertainty, Corporate Expectations, and the Postponement of Investment: The Case of Electricity Market Deregulation. *SSRN Journal*, 2011. ISSN 1556-5068. doi:10.2139/ssrn.1944217.

[3] J. Bowring. Capacity Markets in PJM. *Economics of Energy & Environmental Policy*, 2(2):47–64, 2013. doi:10.5547/2160-5890.2.2.3.

[4] M. J. Brennan and E. S. Schwartz. Evaluating natural resource investments. *The Journal of Business*, 58(2):135–157, 1985. URL http://www.jstor.org/stable/2352967.

[5] B. Caldecott and J. McDaniels. Stranded generation assets: Implications for European capacity mechanisms, energy markets and climate policy. Technical report, Smith School of Enterprise and the Environment, University of Oxford (Oxford, UK), 2014. URL http://www.smithschool.ox.ac.uk/research-programmes/stranded-assets/Stranded%20Generation%20Assets%20-%20Working%20Paper%20-%20Final%20Version.pdf.

[6] J. D. Craig and S. J. Savage. Market Restructuring, Competition and the Efficiency of Electricity Generation: Plant-level Evidence from the United States 1996 to 2006. *The Energy Journal*, 34(1), 2013. doi:10.5547/01956574.34.1.1.

[7] P. Cramton and S. Stoft. A Capacity Market that Makes Sense. *The Electricity Journal*, 18(7):1040–6190, 2005. doi:10.1016/j.tej.2005.07.003.

[8] S.-J. Deng, B. Johnson, and A. Sogomonian. Exotic electricity options and the valuation of electricity generation and transmission assets. *Decision Support Systems*, 30(3):383–392, 2001. doi:10.1016/S0167-9236(00)00112-3.

[9] A. Dixit. Entry and exit decisions under uncertainty. *Journal of Political Economy*, 97(3):620–638, 1989. URL http://www.jstor.org/stable/1830458.

[10] A. K. Dixit and R. S. Pindyck. *Investment under uncertainty*. Princeton University Press, 1994.

[11] N. Duic and M. Carvalho. Increasing renewable energy sources in Island energy supply: case study Porto Santo. *Renewable Sustainable Energy Review*, 8(4):383–399, 2004. doi:10.1016/j.rser.2003.11.004.

[12] P. Embrechts, C. Klüppelberg, and T. Mikosch. *Modelling Extremal Events*, volume 33 of *Stochastic Modelling and Applied Probability*. Springer, 1997. URL http://books.google.com/books?id=BXOI2pICfJUC.

[13] Energy Information Administration. Levelized Cost and Levelized Avoided Cost of New Generation Resources in the Annual Energy Outlook 2015. Report, Energy Information Administration, 2015. URL http://www.eia.gov/forecasts/aeo/pdf/electricity_generation.pdf.

[14] S.-E. Fleten, E. Haugom, and C. J. Ullrich. Keeping the lights on until the regulator makes up his mind. 2012. URL http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1928618.

[15] S.-E. Fleten, E. Haugom, A. Pichler, and C. J. Ullrich. Structural Estimation of Switching Costs for Peak Power Plants. *Under submission*, 2015.

[16] J. Frayer and N. Z. Uludere. What is it worth? Application of real options theory to the valuation of generation assets. *The Electricity Journal*, 14(8):40–51, 2001. doi:10.1016/S1040-6190(01)00237-8.

[17] D. L. Hiebert. The Determinants of the Cost Efficiency of Electric Generating Plants: A Stochastic Frontier Approach. *The Southern Economic Journal*, 68(4):935–946, 2002. doi:10.2307/1061501.

[18] P. L. Joskow. Capacity payments in imperfect electricity markets: Need and design. *Utilities Policy*, 16(3):159–170, 2008. doi:10.1016/j.jup.2007.10.003.

[19] R. Kellogg. The Effect of Uncertainty on Investment: Evidence from Texas Oil Drilling. *American Economic Review*, 104(6):1698–1734, 2014. doi:10.1257/aer.104.6.1698.

[20] G. King and L. Zeng. Logistic regression in rare events data. *Political Analysis*, 9(2):137–163, 2001.

[21] C. R. Knittel. Alternative Regulatory Methods and Firm Efficiency: Stochastic Frontier Evidence from the U.S. Electricity Industry. *The Review of Economics and Statistics*, 84(3):530–540, 2002. doi:10.1162/003465302320259529.

[22] C.-Y. C. Lin and K. E. Thome. Investment in corn-ethanol plants in the Midwestern United States: An analysis using reduced-form and structural models. 2010. URL www.des.ucdavis.edu/faculty/Lin/ethanol_US_paper.pdf.

[23] H. Lund. Large-scale integration of optimal combinations of PV, wind and wave power into the electricity supply. *Renewable Energy*, 31(4):503–515, 2006. doi:10.1016/j.renene.2005.04.008.

[24] G. Pagès. A space quantization method for numerical integration. *Journal of Computational and Applied Mathematics*, 89(1):1–38, 1998. doi:10.1016/S0377-0427(97)00190-8.

[25] A. Pakes, M. Ostrovsky, and S. Berry. Simple estimators for the parameters of discrete dynamic games (with entry/exit examples). *RAND Journal of Economics*, 38(2):373–399, 2007. doi:10.1111/j.1756-2171.2007.tb00073.x.

[26] J. Rust. Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher. *Econometrica*, 55(5):999–1033, 1987. doi:10.2307/1911259.

[27] C.-L. Su and K. L. Judd. Constrained Optimization Approaches to Estimation of Structural Models. *Econometrica*, 80(5):2213–2230, 2012. doi:10.3982/ECTA7925.

[28] K. E. Train. *Discrete choice methods with simulation*. Cambridge University Press, 2002. doi:10.1017/CBO9780511753930.