

Authors are encouraged to submit new papers to INFORMS journals by means of a style file template, which includes the journal title. However, use of a template does not certify that the paper has been accepted for publication in the named journal. INFORMS journal templates are for the exclusive purpose of submitting to an INFORMS journal and should not be used to distribute the papers in print or online or to submit the papers to another publication.

# Operating Peak Power Plants as Real Options

Stein-Erik Fleten

Norwegian University of Science and Technology, NO-7491, Trondheim, Norway, Contact: [stein-erik.fleten@iot.ntnu.no](mailto:stein-erik.fleten@iot.ntnu.no)

Erik Haugom

Lillehammer University College, Lillehammer, Norway, Norwegian University of Science and Technology, NO-7491, Trondheim, Norway, Contact: [erik.haugom@hil.no](mailto:erik.haugom@hil.no)

Alois Pichler

Norwegian University of Science and Technology, Actuary of the Actuarial Association of Austria, Contact: [aloisp@ntnu.no](mailto:aloisp@ntnu.no)

Carl J. Ullrich

James Madison University, Department of Finance and Business Law, Harrisonburg, VA, 22807, Contact: [ullriccj@jmu.edu](mailto:ullriccj@jmu.edu)

Peak power plants operate only a few hours per year. In order to avoid ongoing maintenance costs managers may shutdown operating-ready plants, putting them into standby mode. Plants in the standby mode may be made ready for operation again (started up) or abandoned permanently. We analyze decisions to switch between operating modes and describe a method within structural estimation for estimating maintenance costs and irreversible switching costs. Based on U.S. data for 2001-2009, we arrive at useful cost estimates for the case of peak power plants. The method is applicable to estimation of switching costs for industrial facilities more generally.

*Key words:* Structural estimation, Hamilton–Jacobi–Bellman equations, value function, stochastic optimization, switching options.

---

## 1. Introduction

Penetration of renewable energy generation technology such as solar photovoltaics and wind turbines continues to increase in the United States. The output of solar and wind plants varies with ambient weather conditions. Large real-time variations in the output of solar and wind plants can make it harder to maintain reliability of the electrical grid. A recent Wall Street Journal article

by Smith and Cook (2015) explores the problems currently facing Hawaii, a state with a relatively large amount of renewable generation. According to the authors, “... *sudden swings in the output of solar and wind ... force the state’s main utility to scramble to try to keep overall supply of power steady.*” Flexible peak power plants - mainly gas-fired combustion turbines - help to ensure reliability of the grid by providing quick-start and load-following (i.e., varying output in real-time in order to match demand) capabilities.

Solar and wind power-generating technologies have zero or near-zero operating costs. As penetration of renewables increases owners of existing fossil fuel power plants are likely to see the value of their assets erode. As detailed in Caldecott and McDaniels (2014), the value of gas-fired combustion turbine-based power plants in Europe has dropped significantly. The owner of an operating-ready plant holds the real option to shutdown the plant, i.e., put the plant into standby mode, as its value declines.<sup>1</sup> The introduction of generating technologies with low operating costs should make it more likely that an operating-ready plant will be shutdown.

On the other hand, many electricity systems operate capacity markets, for example the Reliability Pricing Mechanism in the Pennsylvania-New Jersey-Maryland (PJM) market. The goal of a capacity market is to provide incentive to build and/or maintain plants needed for system reliability, even though these plants are expected to actually generate power only a small fraction of the year. The owner of a plant which was previously shutdown holds the real options to either restart or abandon the plant. Capacity markets should make it more likely that a plant is already in the standby mode will be restarted and less likely that the plant will be abandoned.

Shutdown, start-up, and abandonment entail one-time switching costs. The seminal article by Brennan and Schwartz (1985) provides the basic framework for the real options to switch operating modes. Moel and Tufano (2002) examine switching in the case of gold mines. In their regressions Moel and Tufano (2002) proxy for the one-time cost of shutting down a mine (a “closing cost” in their language) with capitalized costs and an indicator variable for underground mines.

Estimates of switching costs for peak power plants are surprisingly difficult to obtain in practice. Discussions with industry professionals indicate that shutdown and start-up costs depend upon

maintenance policies, ambient weather conditions, managerial priorities, and other factors which vary across plants.<sup>2</sup> Firms can essentially “walk away” from an operating-ready peak plant and incur little or no shutdown cost. Start-up costs depend upon the amount of maintenance the plant received while it was in the standby mode.

We use nonparametric structural estimation to determine the one-time switching costs associated with shutting down, restarting, and abandoning peak power plants. We also estimate the continuing costs of maintenance. Our case study is made possible by the availability of detailed data from the United States. Each year the owners of existing plants must file Form 860 with the Energy Information Administration. Included in Form 860 is the so-called *status* of the plant. From these data it is possible to determine whether an existing plant was shutdown, restarted, or abandoned during the previous year. That is, we can observe actual shutdowns, start-ups, and abandonments. Our sample includes 8,189 plant-year observations from the period 2001-2009.

Structural estimation of dynamic discrete choice models was introduced in the paper Rust (1987) by Rust, in which he studies replacing or repairing bus engines in fleet vehicles. The papers by Rust (1987), Gamba and Tesser (2009), and Su and Judd (2012) consider parametric stochastic processes (exponential distributions or geometric Brownian motions) to model the underlying state variable. Our estimation method is based on a nonparametric representation of the dynamics of the exogenous state variable. The transitions are modeled by a nonparametric Markovian model and we estimate the transition operator using kernel functions, which are well-known from nonparametric density estimation and nonparametric regression.

Previous work that uses nonparametric methods in structural estimation of dynamic models include Bansal et al. (1995), who nevertheless specify an autoregressive structure for the exogenous state variables. Newey et al. (1999), Bontemps et al. (2000), Guerre et al. (2000) and Li et al. (2002) employ nonparametric estimation for structural estimation of static models within simultaneous equations, auctions, and labor search. Strebulaev and Whited (2012) review dynamic models (including contingent claims and real options) and structural estimation in finance.

Deng et al. (2001), Tseng and Barz (2002), and Thompson et al. (2004), among others, apply real options methods to the operation and valuation of electricity assets. Structural estimation applications to electricity related problems include Rust and Rothwell (1995) and Rothwell and Rust (1997), who study the effect of a regulatory shift and plant optimal lifetime, respectively, for nuclear power plants. In energy economics more broadly, recent work on structural estimation include analysis of timing; Rapson (2014) studies the timing of appliance investment, Kellogg (2014) well drilling, Muehlenbachs (2015) well decommissioning, Burr (2012) solar PV investment, and Lin and Thome (2013) corn-ethanol plant investment. Potential applications outside the realm of power plants include oil rigs (Kenyon and Tompaidis 2001), oil properties (Smith and McCardle 1998), mines (Moel and Tufano 2002), LNG terminals (Lai et al. 2011), and industrial facilities with investment and exit options more generally (Kwon 2010). The respective articles cited all feature dynamic decision processes subject to observable uncertain exogenous profitability drivers. Brekke and Øksendal (1994) offer an analysis of the optimal solution of such switching problems quite generally, noting that it belongs to the class of generalized impulse control problems studied, e.g., by Bensoussan and Lions (1984).

*Outline of the paper.* Section 2 describes structural estimation and provides motivation. Section 3 addresses the new nonparametric approach to model the transitions. Section 4 describes the case study and details the data. Results are presented in Section 5. In Section 6 we discuss the relevance of the results and conclude.

## 2. Theoretical Framework

### 2.1. The framework for structural estimation

In this section we define the nomenclature and introduce the structural estimation method. The notation as well as the outline closely follow the literature, for example in Gamba and Tesser (2009).

#### Nomenclature

$k$  Time index; the unit time period is a year.

$X_k$  The state process; in our specific case the state process is an indicator of profitability per unit of capacity expressed in units of dollars per kilowatt,  $\frac{\$}{kW}$ .

$(X_k, \varepsilon_k)$  The augmented state process; the second process,  $\varepsilon_k$ , is not accessible to observation.

$s, u \in S \equiv \{\text{operating, standby, retired}\}$  are operating states of the power plants and decided by the plant manager.

$(X_i, s_i, u_i)$  An observation consists of a profitability  $X_i$  during the current year, the state  $s_i$  of the system in the current year, and  $u_i$ , the state of the system in the following year after the managers decision.<sup>3</sup>

$g(x, s; u)$  The payoff during a single period. The payoff function  $g(\cdot)$  comprises the expected cash flow for the next period and all costs associated with the transition from  $s$  to  $u$ .

$V(x, s)$  Value function — the accumulated discounted future payoffs achieved from an optimal policy.

$v(x, s)$  Expected (or  $s$ -alternative-specific) value function. The function  $v$  is the average of the different value functions  $V$  among all agents operating a power plant in the market.

$\beta \in (0, 1)$  Discount factor  $\beta = \frac{1}{(1+\text{interest rate})}$ .

## 2.2. Bellman equation of an individual decision maker

The state process  $(X_0, X_1, \dots)$  is observed over a sequence of years. A decision  $s_k$  is allowed at every stage  $k$  ( $k = 0, 1, \dots$ ). Assume the development of the process depends only on its history, that is the process is nonanticipative and the decision is based on the history of the system. Employing a payoff function  $g(\cdot)$  which describes the profitability in each period, the investor will thus maximize the function

$$V(x, s) := \max_{s_k \in \sigma(X_k)} \mathbb{E} \left( \sum_{k=0}^{\infty} \beta^k g(X_k, s_k; s_{k+1}) \middle| X_0 = x \right);$$

the maximum here is among all decision processes  $(s_0, s_1, \dots)$ , with  $s_0 = s$ , the current state. It is natural to assume that the process  $X$  is Markovian and independent of time. In this Markovian framework the decision depends on the current stage only, i.e.,  $s_k = s_k(X_k)$  for some measurable

function  $s_k$  (the Doob–Dynkin lemma, cf. Kallenberg (2002, Lemma 1.13) or Shiryaev (Shiryaev 1996, Theorem II.4.3)), which is made explicit by writing  $s_k \triangleleft \sigma(X_k)$ . It follows that

$$\begin{aligned}
V(x, s) &= \max_{s_k \triangleleft \sigma(X_k)} \mathbb{E} \left( \sum_{k=0}^{\infty} \beta^k g(X_k, s_k; s_{k+1}) \middle| X_0 = x \right) \\
&= \max_{s_k \triangleleft \sigma(X_k)} \mathbb{E} \left( g(X_0, s_0; s_1) + \beta \cdot \sum_{k=0}^{\infty} \beta^k g(X_{k+1}, s_{k+1}; s_{k+2}) \middle| X_0 = x \right) \\
&= \max_{s_k \triangleleft \sigma(X_k)} \mathbb{E} \left( g(X_0, s_0; s_1) + \beta \cdot \mathbb{E} \left( \sum_{k=0}^{\infty} \beta^k g(X_{k+1}, s_{k+1}; s_{k+2}) \middle| X_1 = x_1 \right) \middle| X_0 = x \right) \\
&= \max_{s_k \triangleleft \sigma(X_k)} \mathbb{E} (g(X_0, s_0; s_1) + \beta \cdot V(X_1, s_1) | X_0 = x) \\
&= \max_{u \in S} g(x, s; u) + \beta \cdot \mathbb{E} (V(X_{k+1}, u) | X_k = x), \tag{1}
\end{aligned}$$

which is the usual Bellman equation, cf. Fleming and Soner (2006, Chapter IX). Note that  $X_{k+1}$  is distributed, conditional on  $X_k$ , in the same way as  $X_1$  is distributed conditional on  $X_0$ . For this reason we shall express the conditional expectations based on  $X_0$  in what follows.

It follows from (1) that the value function  $V(\cdot, \cdot)$  is a fixed point for

$$T(V)(x, s) := \max_{u \in S} g(x, s; u) + \beta \cdot \mathbb{E} (V(X_1, u) | X_0 = x), \tag{2}$$

that is,

$$V = T(V). \tag{3}$$

The operator  $T$  is Lipschitz continuous, if considered on the Banach space  $\ell^\infty([0, \infty) \times S)$  of uniformly bounded functions or  $C([0, \infty) \times S)$ , the linear space of continuous functions on  $[0, \infty) \times S$  equipped with the sup-norm  $\|f\|_\infty := \sup_{x \geq 0, s \in S} |f(x, s)|$ .  $T$  is monotone and a contraction by Blackwell's sufficient conditions (for every individual  $s \in S$ , the discount factor  $\beta < 1$  being the Lipschitz constant). The Banach fixed point theorem thus ensures that (3) has a unique solution, and further that  $|V| \leq \frac{\|g\|_\infty}{1-\beta}$ . The value function is thus uniformly bounded, provided that  $g(\cdot)$  is uniformly bounded, and  $V$  is continuous, provided that  $g$  is continuous.

### 2.3. The dynamic programming equation for a decision maker in an uncertain economic environment

In an economic environment several decision makers are observed, each of whom makes decisions individually. We assume that each individual decision maker acts rationally, but has more information than the observing economist. To capture this situation we consider the augmented process  $(X_k, \varepsilon_k)_{k=0}$ , with the process  $\varepsilon$  independent of  $X$ . This idiosyncratic shock ( $\varepsilon$ ) will affect the payoff function of the decision makers, as seen below. Every  $\varepsilon_k = (\varepsilon_{k,u})_{u \in S}$  is a vector carrying the additional information which is associated with the actions  $u \in S$ . This information is hidden from the economist who observes only the state  $X_k$ .

Every individual decision maker bases his decision on the augmented state space process  $(X_k, \varepsilon_k)$  and the payoff function  $g(x, \varepsilon, s; u)$ . In analogy to (1) it follows that

$$V(x, \varepsilon, s) = \max_{u \in S} g(x, \varepsilon, s; u) + \beta \cdot \mathbb{E} \left( \int V(X_1, \varepsilon_1, u) \mathcal{E}(d\varepsilon_1 | X_1) \middle| X_0 = x \right) \quad (4)$$

has to hold for the value function. The transition of the idiosyncratic shock  $\varepsilon$ , which is described by the distribution  $\mathcal{E}$ , is independent from  $X$  by assumption and one may integrate on every fiber  $\{X_0 = x\}$  separately.  $V$  is the value function, and  $\mathcal{E}$  is the probability measure for  $\varepsilon$  (conditional independence, cf. Rust (1987)).

Define now the *expected value function*, or  $s$ -alternative-specific value function

$$v(x, s) := \mathbb{E} \left( \int V(X_1, \varepsilon_1, s) \mathcal{E}(d\varepsilon_1) \middle| X_0 = x \right). \quad (5)$$

Then the Bellman equation (4) becomes

$$V(x, \varepsilon, s) = \max_{u \in S} g(x, \varepsilon, s; u) + \beta \cdot v(x, u) \quad (6)$$

and, by taking expectations of (6),

$$\begin{aligned} v(x, s) &= \mathbb{E} \left( \int V(X_1, \varepsilon_1, s) \mathcal{E}(d\varepsilon_1) \middle| X_0 = x \right) \\ &= \mathbb{E} \left( \int \max_{u \in S} g(X_1, \varepsilon_1, s; u) + \beta \cdot v(X_1, u) \mathcal{E}(d\varepsilon_1) \middle| X_0 = x \right). \end{aligned} \quad (7)$$

The latter equation is a fixed point equation for  $v$ , but in contrast to (1) the maximization and expectation are interchanged.

To manage the inner integral of a maximum one may specify the payoff function  $g$  by accounting for the shock term in a linear way according to

$$g(x, \varepsilon, s, u) = g(x, s, u) + \varepsilon_u, \quad (8)$$

and by specifying the distribution of  $\mathcal{E}$ . Rust (1987) uses the term additive separability for the particular decomposition (8) and observes the maximum operation in (8). It is well-known from extreme value theory that the normalized maximum converges to an extreme value distribution. The Gumbel variable is the only extreme value distribution with two-sided support (no other limit is possible by the Fisher–Tippett–Gnedenko theorem, cf. Embrechts et al. (1997)), and it is thus natural to specify  $\varepsilon$  as a process of mutually independent Gumbel variables, which is independent from  $X$ . This allows an important simplification of formula (7). Indeed, the Gumbel distribution is closed under maximization, and in this case a closed form formula for expectations is available and given by

$$\int \max_{u \in S} (\varepsilon_u + c_u) \mathcal{E}(\mathrm{d}\varepsilon_u) = b \cdot \log \left( \sum_{u \in S} \exp \frac{c_u}{b} \right), \quad (9)$$

as is detailed in Proposition 1 in the appendix. Specifying  $c_u := g(X_1, s; u) + \beta \cdot v(X_1, u)$  and applying (9) to

$$v(x, s) = \mathbb{E} \left( \int \max_{u \in S} (g(X_1, s; u) + \varepsilon_{1,u} + \beta \cdot v(X_1, u)) \mathcal{E}(\mathrm{d}\varepsilon_{1,u}) \middle| X_0 = x \right) \quad (10)$$

reduces the inner integral. The fixed point equation (10) simplifies to

$$v(x, s) = \mathbb{E} \left( b \cdot \log \left( \sum_{u \in S} \exp \left( \frac{g(X_1, s; u) + \beta \cdot v(X_1, u)}{b} \right) \right) \middle| X_0 = x \right). \quad (11)$$

REMARK 1. Notice, that by assuming a Gumbel distribution the inner integral in (9) simplifies to a simple expression involving the logarithm of a sum of exponentials, but the integrals with respect to  $\varepsilon$  disappear. The Gumbel distribution is an *extreme value type I* distribution. A proof that



Gumbel's distribution is closed under maximization is provided in Proposition 1 in the Appendix, such that (11) is justified. The Gumbel distribution was first incorporated in the model by Rust, while the Generalized Extreme Value Models with conditional logit choice date back to a series of papers by McFadden (cf., for example, McFadden (1974)). A particular advantage of the Gumbel distribution is the simple expression (11) and the explicit formula for the conditional choice probability (Proposition 2 in the Appendix), which will be exploited as well in what follows.

REMARK 2 (THE SPECIAL CASE  $b = 0$ ). The parameter  $b$  smoothes the kink in the function  $(y_u)_{u \in S} \mapsto \max \{y_u : u \in S\}$ . It holds that

$$\max \{y_u : u \in S\} \leq b \cdot \log \sum_{u \in S} e^{\frac{y_u}{b}} \xrightarrow{b \rightarrow 0} \max \{y_u : u \in S\}$$

uniformly and the approximation quality increases whenever  $b$  decreases to 0. In particular (11) results in

$$v(x, s) = \mathbb{E} \left( \max_{u \in S} g(X_1, s; u) + \beta \cdot v(X_1, u) \middle| X_0 = x \right)$$

in the limit whenever  $b = 0$ . This is a recursion for  $v(x, s) = \mathbb{E}(V(X_1, s) | X_0 = x)$ , where  $V$  is the usual value function in Section 2.2. This recovers the classical Bellman theory, but involving the Gumbel distribution leads to a more general, still tractable model. In this situation ( $b = 0$ ) the genuine value function can be recovered, by (2), as  $V(x, s) = \max_{u \in S} g(x, s; u) + \beta \cdot v(x, u)$ .

The additional parameter  $b$  can be interpreted as a degree of uncertainty, as the standard deviation of a Gumbel distribution is  $b \frac{\pi}{\sqrt{6}} \simeq 1.28b$ , where  $b$  is the scale parameter. In particular, the choice  $b = 0$  represents decisions without deviations: this degenerate case describes the classical situation in which all managers decide in the same way.

In analogy to (2) it is convenient to introduce the operator

$$t_g(v)(x, s) := \mathbb{E} \left( b \cdot \log \sum_{u \in S} \exp \frac{g(X_1, s; u) + \beta \cdot v(X_1, u)}{b} \middle| X_0 = x \right). \quad (12)$$

Eq. (11) then rewrites as a fixed-point equation on  $\ell^\infty([0, \infty) \times S)$  as

$$v = t_g(v),$$

which is in analogy with (3).  $t_g$  is again a contraction with Lipschitz constant  $\beta < 1$  and Banach's fixed point theorem ensures that (12) has a unique solution (which we call  $v_g$ ) in the proper space.

**REMARK 3 (THE OPTIMAL POLICY).** Eq. (2) provides the optimal strategy for the individual manager by maximizing  $u \in S$ . This is not the case for the expected value function (11). The reason for this difference is because an optimal decision cannot be specified in a random environment. The function  $v$  is an average over all decision makers, per (5). The concept of a single *optimal decision* cannot be based on  $v$ , as  $v$  is the average over all decision makers.

### 3. Structural estimation

Structural estimation is a technique to uncover parameters of an economic model, which are hidden deeply in the model. In our case study (cf. Section 4 below) the parameters  $\theta$  we want to estimate are transition costs and maintenance costs of peak power plants. The switching costs we seek to estimate appear only in the payoff function  $g_\theta(\cdot)$ . The best model can be selected by a maximum likelihood approach, that is by solving the problem (cf. Su and Judd (2012))

$$\begin{aligned} & \text{maximize } \mathcal{L}(g, v_g, (X_i, s_i, u_i)_{i=1}^N) \\ & \text{subject to } v_g = t_g(v_g), \\ & g \in \mathcal{G}, \end{aligned} \tag{13}$$

where  $N$  is the number of observations and  $\mathcal{L}$  is the likelihood of observing data  $(X_i, s_i, u_i)_{i=1}^N$  conditional on the payoff function  $g(\cdot) \in \mathcal{G} = \{g_\theta(\cdot) : \theta \in \Theta\}$ . The payoff function  $g \in \mathcal{G}$  is chosen from a set  $\mathcal{G}$  of potential candidate functions.  $t_g$  is the operator (12) for the specified function  $g(\cdot)$ .  $v_g$  is the expected value function corresponding to the payoff  $g(\cdot)$  satisfying the constraint  $v_g = t_g(v_g)$ . The constraint  $v_g = t_g(v_g)$  is notably the fixed-point equation, which is an equation to be satisfied by  $v_g$  everywhere on  $[0, \infty) \times S$ .

To make a maximum likelihood estimator available we need an expression for the probability of choice. We explicitly use the probability of choice formula

$$P_v(u|x, s) = \frac{\exp\left(\frac{g(x, s; u) + \beta v(x, u)}{b}\right)}{\sum_{u' \in D} \exp\left(\frac{g(x, s; u') + \beta v(x, u')}{b}\right)}, \quad (14)$$

which follows from the assumption that the process  $\varepsilon$  follows a Gumbel distribution as detailed in Proposition 2 in the Appendix. The likelihood function  $\mathcal{L}$  is thus

$$\mathcal{L}(g, v, (X_i, s_i, u_i)_{i=1}^n) = \prod_{i=1}^N P_v(u_i | X_i, s_i).$$

In the language of Su and Judd (2012), the objective  $\mathcal{L}(g, v, X)$  in (13) is an *augmented* likelihood function because  $\mathcal{L}$  involves the function  $v$  as an auxiliary variable. It is well-known that the maximum-likelihood estimator is consistent (i.e., the estimates converge in probability to the value being estimated as  $n \rightarrow \infty$ ) and efficient (no consistent estimator has lower asymptotic mean square error).

REMARK 4. The classical approach to solving (13) is the nested fixed point (NFXP) algorithm (cf. Rust (1987)). In the NFXP, for every choice  $g \in \mathcal{G}$  the fixed point equation  $v_g = t_g(v_g)$  has to be solved, as the function  $v_g$  enters the objective in the maximization (13) or (18) below. This is the most expensive part of the computational problem in NFXP. In contrast, in the Su and Judd (2012) approach, the solution  $v_g$  maximizing (13) is a by-product of the optimization process and has the statistical interpretation of a nuisance parameter.

The approach described below, building on the ideas of Su and Judd (2012), addresses this problem. It introduces a direct estimator for  $t_g$  which is free of parameters. Moreover, the approach described ensures convergence to the continuous solution  $v_g$ . Approximations of the solution are constructed by fixing a grid of supporting points on the positive real line for every  $s \in S$  and by linear interpolation of the functions  $v(\cdot, s)$ ,  $s \in S$  in between. The supporting points are refined successively to a dense set in  $\mathbb{R}_{\geq 0}$  for every  $s \in S$ , which ensures pointwise convergence of the approximations to  $v_g$ .

*Estimation of conditional expectation.* The probability in the maximum likelihood estimator (13) involves the operator  $t_g$  which is an expectation, conditional on  $\{X_0 = x\}$ . To evaluate  $t_g(v)(x, \cdot)$  at a specified point  $x$  (cf. (12)) it is necessary to evaluate a conditional expectation. To estimate the conditional expectation of  $f(X_1)$  relative to  $X_0$ , that is  $\mathbb{E}(f(X_{k+1}) | X_k)$ , we pair subsequent observations and consider

$$(X_i, X_{i+1}) \quad \text{for } i = 1, 2, \dots, N-1. \quad (15)$$

Then the Nadaraya–Watson estimator (well-known from kernel regression) for the operator

$$t_g(v)(x, s) = \mathbb{E} \left( b \cdot \log \sum_{u \in S} \exp \frac{g(X_1, s; u) + \beta \cdot v(X_1, u)}{b} \middle| X_0 = x \right)$$

is

$$\hat{t}_g(v)(x, s) := \sum_{i=1}^{N-1} \frac{K\left(\frac{x-X_i}{h}\right)}{\sum_{i'=1}^{N-1} K\left(\frac{x-X_{i'}}{h}\right)} \cdot b \cdot \log \sum_{u \in S} \exp \frac{g(X_{i+1}, s; u) + \beta \cdot v(X_{i+1}, u)}{b}, \quad (16)$$

where  $K(\cdot)$  is an appropriate kernel function and  $h > 0$  a suitable bandwidth. Consistency of this estimator, in an even broader context, is justified in Atuncar et al. (2008).

The estimator  $\hat{t}_g$  maintains all properties of the original operator  $t_g$ , as the following lemma reveals.

**LEMMA 1.** *For the choice  $\beta < 1$  the mapping  $v \mapsto \hat{t}_g(v)$  is a contraction on  $\ell^\infty([0, \infty) \times S)$  (the linear space of bounded function on  $[0, \infty) \times S$ ), and  $v = \hat{t}_g(v)$  has a unique fixed point.*

*Proof of Lemma* Observe first that the sample  $(X_i)_{i=1}^N$  and  $S$  are finite, such that  $g(X_{i+1}, s; u)$  is uniformly bounded. Moreover the mapping

$$v \mapsto b \cdot \log \sum_{u \in S} \exp \frac{g_u + \beta \cdot v}{b} = b \cdot \log \left( \exp \left( \frac{\beta \cdot v}{b} \right) \cdot \sum_{u \in S} \exp \left( \frac{g_u}{b} \right) \right) = \beta \cdot v + b \cdot \log \sum_{u \in S} \exp \left( \frac{g_u}{b} \right)$$

is an affine linear function in  $v$  with slope  $\beta < 1$ . Due to the construction of the operator  $\hat{t}_g$  in (16) with nonnegative weights summing to 1 it follows that  $\hat{t}_g$  is a contraction with Lipschitz constant  $\beta < 1$ . Hence Banach's fixed-point theorem applies and guarantees a unique fixed-point  $v_g$  in  $\ell^\infty([0, \infty) \times S)$ , that is,  $v_g = \hat{t}_g(v_g)$ .  $\square$

The estimator  $\hat{t}_g$  maintains essential properties on functions which are piecewise linear. This observation is important for numerical treatments as it allows us to consider linear spline functions in implementations. The following corollary is immediate.

**COROLLARY 1 (Interpolation).** *For  $d+1$  fixed numbers  $x_0 < x_1 < \dots < x_d$  in  $\mathbb{R}$  let  $I$  denote the linear interpolation operator, such that*

$$I(v_0, \dots, v_d)(x) = \begin{cases} v_0 & \text{if } x \leq x_0, \\ v_j \frac{x_{j+1}-x}{x_{j+1}-x_j} + v_{j+1} \frac{x-x_j}{x_{j+1}-x_j} & \text{if } x_j \leq x \leq x_{j+1}, \\ v_d & \text{if } x \geq x_d. \end{cases}$$

Then

$$\hat{t}_g((v_0^s, \dots, v_d^s)_{s \in S}) := \left( \sum_{i=1}^{N-1} \frac{K\left(\frac{x_j - X_i}{h}\right)}{\sum_{i'=1}^{N-1} K\left(\frac{x_j - X_{i'}}{h}\right)} \cdot b \log \sum_{u \in S} \exp \frac{g(X_{i+1}, s; u) + \beta \cdot I(v_0^u, \dots, v_d^u)(X_{i+1})}{b} \right)_{j=0}^d \quad (17)$$

is a contraction on  $\mathbb{R}^{(d+1) \cdot |S|}$  with a unique fixed point. ( $|S|$  is the cardinality of different state modes;  $|S| = 3$  in our case).

*The choice of the kernel and bandwidth.* For our set of data and our particular purposes we find the logistic kernel

$$K(x) = \frac{1}{4} \frac{1}{\left(\cosh \frac{x}{2}\right)^2} = \frac{1}{e^x + 2 + e^{-x}}$$

convenient, because

- it allows for all moments, and
- its tails are fat enough to include more distant observations as well.

The choice of this particular logistic kernel is not restrictive, other kernels provide reasonable results as well. For the bandwidth we chose Silverman's rule of thumb (cf. Silverman (1998)), that is

$$h_N = \text{std}(X_i) \cdot \left( \frac{4}{(m+2)N} \right)^{\frac{1}{m+4}} \approx \text{std}(X_i) \cdot N^{\frac{-1}{(m+4)}},$$

where  $\text{std}(\cdot)$  is the standard deviation of the sample,  $m = 2$  is the dimension of each individual pair of the samples (cf. (15)) and  $N$  the sample size.

## 4. Case Study

The specific application we consider is the case of shutting down, restarting, and abandoning peak power plants. We limit our sample to simple cycle combustion turbine power plants (hereafter CTs) located in the northeastern part of the United States. See Fleten, Haugom, and Ullrich (2013) for further details about the choice of plants and geographic locations. The sample period is 2001–2009. In this section we describe the data and define the payoff function  $g(\cdot)$  used in the optimization exercise.

### 4.1. Data

The owners of power plants in the United States must each year file Form 860 with the Energy Information Administration (EIA). Included in Form 860 is the status of each power plant. For our purposes, the relevant statuses are as follows:

- OP — operating,
- SB — standby, and,
- RE — retired.

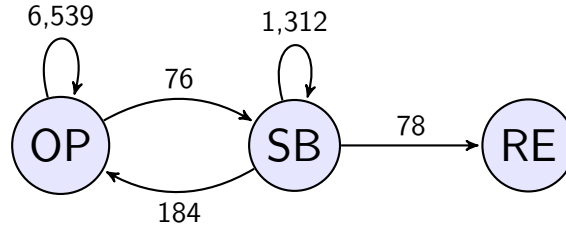
A plant in state OP is available for operation. A plant in state SB has been shutdown and cannot be made ready for operation in the short term.<sup>4</sup> A plant which in state RE has been abandoned and cannot return to service.

- We define a *shutdown* to occur when a plant moves from state OP in year  $k$  to state SB in year  $k + 1$  and label this transition  $OP \rightarrow SB$ .

- We define a *start-up* to occur when plant moves from state SB in year  $k$  to state OP in year  $k + 1$  and label this transition  $SB \rightarrow OP$ .

- We define an *abandonment* to occur when a plant moves from state SB in year  $k$  to state RE in year  $k + 1$  and label this transition  $SB \rightarrow RE$ .

Other possible (non)transitions include  $OP \rightarrow OP$  and  $SB \rightarrow SB$ . Figure 1 summarizes the status changes in our dataset.



**Figure 1** Transitions between the three states *operating* (OP), *standby* (SB) and *retired* (RE). Included is the number of transitions.

*Augmented likelihood function.* Our data consist of distinct groups. There are 6,539 occurrences of operating plants continuing to operate ( $OP \rightarrow OP$ ), but only 76 occurrences of shutdown ( $OP \rightarrow SB$ ). This is not accounted for in the log-likelihood as (13) treats each observation in the same way.

Following the literature (cf., for example, King and Zeng (2001)) it is natural to augment the likelihood to reflect the different sample sizes of the groups. This is accomplished by the *augmented likelihood*

$$\begin{aligned}
 & \text{maximize} \quad \sum_{i=1}^N \frac{1}{N_i} \log P_{v_g}(u_i | X_i, s_i) \\
 & \text{subject to } v_g = t_g(v_g), \\
 & \quad g \in \mathcal{G}.
 \end{aligned} \tag{18}$$

Here,

$$N_i \in \{6,539; 76; 184; 1,312; 78\}$$

is the sample size of the group to which the observation  $(X_i, s_i, u_i)$  belongs to as in Figure 1 (instead of  $N = 8,189$  in (13)) ( $i \in \{OP \rightarrow OP, OP \rightarrow SB, SB \rightarrow OP, SB \rightarrow SB, SB \rightarrow RE\}$ ). As a consequence the five groups provided in Figure 1 are equally weighted, and the weights within the group are chosen to reflect the individual importance of each group.

## 4.2. Observation

An observation in our estimation exercise is a triple  $(X_i, s_i, u_i)$  consisting of the following ingredients:

1. the profitability  $X_i$  for the current year,
2. the operating state of the power plant  $s_i \in S$  in the current year, and,
3. the decision of the manager regarding the operating state  $u_i \in S$  of the power plant in the upcoming year.

### 4.3. Profitability

*Spread Options.* The cash flow for a power plant is determined by the spark spread, the difference between the price of electricity and the cost of fuel used to produce it. A peak plant consists of a series of daily European call options on the spark spread. The *heat rate* of a power plant is the amount of fuel, measured in millions of British thermal units ( $MMBtu$ ), required to generate one megawatt hour ( $MWh$ ) of electricity. A lower number therefore indicates greater efficiency.

Consider a plant which has heat rate  $H$  ( $MMBtu/MWh$ ). We calculate the plant-specific spark spread ( $S_n$ ) expressed in units of dollars per megawatt hour ( $$/MWh$ ), for day  $n$  as

$$S_n = P_n^e - H * P_n^f - V,$$

where  $P_n^e$  is the day  $n$  electricity price ( $$/MWh$ ),  $P_n^f$  is the day  $n$  fuel price ( $$/MMBtu$ ), and  $V$  ( $$/MWh$ ) is the variable non-fuel generation cost.

Daily spot prices for New York Harbor No. 2 Oil and NYMEX Henry Hub natural gas are taken from the EIA website. Electricity prices come from the Independent System Operator (ISO) websites. Consistent with our focus on peaking plants, we use electricity prices for the peak period of the day, defined as the industry standard 16 hour period from 06:00 (“hour ending” 7, or HE7) through 22:00 (HE22). We obtain daily peak prices by taking the simple average of the hourly spot prices during the peak period.

Profitability per unit of capacity ( $$/kW$ ) is the state variable  $X$  in our optimization. The profitability per unit of capacity ( $$/kW$ ) for year  $k$  is given by

$$X_k = \sum_{n=1}^{T_k} \max(S_n, 0) * \left( \frac{16}{1000 \text{ kW/MW}} \right),$$



Transition	$OP \rightarrow OP$	$OP \rightarrow SB$	$SB \rightarrow OP$	$SB \rightarrow SB$	$SB \rightarrow RE$	Total
Observations	6,539	76	184	1,312	78	8,189
State variable current year						
Average in \$/ kW	12.6	5.4	17.7	12.3	2.1	12.5
Standard deviation	14.0	9.9	16.4	14.1	5.0	14.0
State variable upcoming year						
Average \$/ kW	9.5	4.4	6.3	7.1	N/A	9.0
Standard deviation	10.5	7.6	8.1	7.6	N/A	10.0

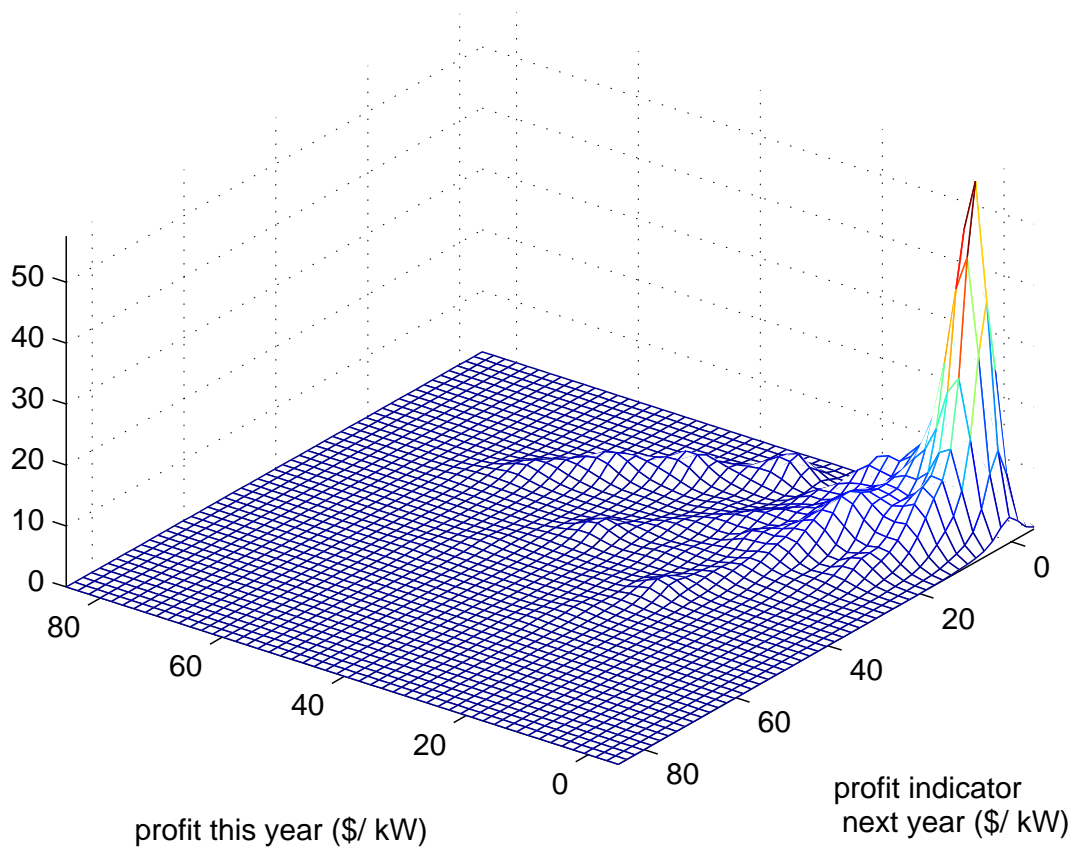
**Table 1** Summary statistics profitability  $X$ , the state variable, in units of \$/kW.

where 16 is the number of peak hours in a day and  $T_k$  is the number of days in year  $k$ . The max function captures the optionality of the plant. On days for which the spread is negative, the plant does not operate and the profit is zero.

*Observation Pairs.* The optimization relies on pairs of observations  $(X_k, X_{k+1})$ , the profitability in the current year  $k$  and in the upcoming year  $k + 1$ . At the time of the decision, the profitability for the upcoming year is not yet known.<sup>5</sup>

We calculate profitability in both the current year  $k$  and the upcoming year  $k + 1$  using actual electricity prices and fuel prices. Because we use actual fuel and electricity prices, together with plant-specific heat rate information, we can calculate profitability for all plants in the sample, operational or otherwise. For those plants which have status  $SB$ , the profitability is hypothetical. In this case  $X_k$  is the profitability which would have obtained if the plant had been in state  $OP$  in year  $k$ . We assume that if the plant had been operating the effect on electricity prices and fuel prices would have been negligible.

Table 1 presents summary statistics for profitability. Notice that the total average profitability for the current year (\$12.5/kW, from the final column) is roughly one third higher than the average profitability for the upcoming year (\$9.0/kW). This is because the first year of our sample, 2001, was the single most profitable year during the sample period, with average profitability \$37.7/kW,



**Figure 2** Bivariate density of the observed transition  $(X_k, X_{k+1})$  of the annual profitability indicator.

not shown in the table. Because 2001 is the first year of our sample, it is not included in the averages for the upcoming year.

Figure 2 presents the evolution of profitability from one year to the next. The density in Figure 2 is estimated based on the pairs  $(X_k, X_{k+1})$  (cf. (15)), which are available from the observations: one ordinate represents the profitability indicator of this year,  $X_k$ , the other ordinate the profitability indicator in the subsequent year,  $X_{k+1}$ . The density thus describes the Markov kernel, which is used in the expectation to compute the value function, for example in (11).

**REMARK 5 (LACK OF EVIDENCE FOR A PARAMETRIZED MODEL).** Figure 2 gives evidence that there is no common pattern of transitions of the profitability from a year to the next. This is a clear indicator that the model cannot be treated by employing standard tools by assuming, e.g.,

a (geometric) Brownian motion. If the transitions were described by a GBM, then a slice of the density plot would be lognormal.

#### 4.4. Unobserved heterogeneity

Heterogeneity means that different decision makers will react differently, even if the state variable  $X_k$  is the same. This reflects different cost structures, strategic positions, and switching opportunities, all of which are unobserved by the economist. As discussed above, we capture the difference between the (within-model) expected immediate payoff and the real payoff observed by the decision maker, i.e., the so-called idiosyncratic shock, by Gumbel variables with a common scale parameter  $b$ . Corollary 2 in the Appendix elaborates that the difference of Gumbel variables has logistic distribution, such that assuming a logistic distribution with parameters is consistent with (24) (in the Appendix), and thus consistent with the general approach to structural estimation. Consistent with the literature (cf. Rust (1987)) we normalize  $b$  to 1.

In addition to the (limited) heterogeneity induced by the idiosyncratic shock we capture heterogeneity by allowing some of the payoff function parameters to be random variables with a given distribution (cf. Train (2002)). The distribution is discretized for the implementation. The discretization chosen employs the representative points (quantizers) of the distribution with the corresponding optimal weights, such that the distance to the genuine distribution is minimized (cf. Pagès (1998), who uses the Wasserstein distance). Specifically, we let two of the parameters be (discretized versions of) Gaussian, i.e., we estimate the mean and standard deviation of two of the cost parameters.<sup>6</sup> Graf and Luschgy (2000, Table 5.1) state the representative points and the corresponding weights explicitly for this distribution. This approach reflects differences across plants: some plant owners will have more efficient maintenance and switching operations than others. However, if startup costs are similar across plants, then the estimated standard deviation for this startup cost will be low.

#### 4.5. The payoff function $g(\cdot)$

The payoff function  $g(\cdot)$  describes the expected cash flow for the year. The profitability indicator  $X$  for the year is a calculated value, as discussed above.

We also must include two other costs, (i) the costs of continuing maintenance  $M_u$  given that the generator is in state  $u$ , and, (ii) the costs associated with the transitions themselves,  $K_{s \rightarrow u}$ . Before adding the cross-plant heterogeneity, the payoff function is given by

$$g_\theta(X, s; u) = \begin{cases} X - M_{OP} & \text{if } s = \text{operating and } u = \text{operating,} \\ \frac{1}{2}(X - M_{OP} - M_{SB}) - K_{OP \rightarrow SB} & \text{if } s = \text{operating and } u = \text{standby,} \\ \frac{1}{2}(X - M_{SB} - M_{OP}) - K_{SB \rightarrow OP} & \text{if } s = \text{standby and } u = \text{operating,} \\ -M_{SB} & \text{if } s = \text{standby and } u = \text{standby,} \\ -\frac{1}{2}M_{SB} - K_{SB \rightarrow RE} & \text{if } s = \text{standby and } u = \text{retired,} \\ -\infty & \text{else,} \end{cases} \quad (19)$$

where the parameter  $\theta$  carries the costs, i.e.,

$$\theta = (M_{OP}, M_{SB}, K_{OP \rightarrow SB}, K_{SB \rightarrow OP}, K_{SB \rightarrow RE}).$$

To exclude other transitions the value  $-\infty$  is included in the payoff function. Notice that for plants which are either shutdown or started up, we include only half of the profit, as well as half of the maintenance cost for both the operational ( $OP$ ) and shutdown ( $SB$ ) states,  $\frac{1}{2}(X - M_{OP} - M_{SB})$ . Because our data are observed at the annual frequency, we do not know when during status changes occur. We assume that shutdowns and start-ups happen mid-year so that in both cases the plant is assumed operation for half of the year. Similarly we include half of the maintenance cost for the  $SB$  state. The parameters of interest are the switching costs  $K_{OP \rightarrow SB}$ ,  $K_{SB \rightarrow OP}$  and  $K_{SB \rightarrow RE}$ . While start-ups ( $OP \rightarrow SB$ ) and shutdowns ( $SB \rightarrow OP$ ) involve cash outflows (switching costs), retirements ( $SB \rightarrow RE$ ) may result in a cash inflow. That is, it may be that  $K_{SB \rightarrow RE} < 0$ . A negative value for  $K_{SB \rightarrow RE}$  would then be interpreted as a resale (or salvage) value.

A generator which has been retired has no value beyond any potential salvage value as described above, that is

$$v(\cdot, \text{retired}) = 0.$$

What remains to be computed is

$$v(\cdot, s), \quad \text{for } s \in \{\text{operating, standby}\}.$$

As justified in Corollary 1, it is possible to employ linear interpolation by fixing supporting points  $x_0 < x_1 < \dots < x_d$ . The problem is

$$\begin{aligned} & \text{maximize } \frac{1}{N} \sum_{i=1}^N \log P_{I(v_g)}(u_i | X_i, s_i) \\ & \text{subject to } v_g = \hat{t}_g(v_g), \\ & g \in \mathcal{G}, \end{aligned} \tag{20}$$

where  $g_\theta \in \mathcal{G}$ , in view of (19), means that

$$\theta := (M_{OP}, K_{OP \rightarrow SB}, K_{SB \rightarrow OP}, M_{SB}, K_{SB \rightarrow RE})$$

are the variables in the optimization procedure (20). We impose the constraints  $M_{OP} \geq 0$ ,  $K_{OP \rightarrow SB} \geq 0$ ,  $K_{SB \rightarrow OP} \geq 0$  and  $M_{SB} \geq 0$  on our optimization procedure, and this is reflected in the functions  $g \in \mathcal{G}$  as well.

Irrespective of the supporting points  $x_0 < \dots < x_d$  this problem has a solution, the problem is always feasible for every choice of

$$\theta = (M_{OP}, K_{OP \rightarrow SB}, K_{SB \rightarrow OP}, M_{SB}, K_{SB \rightarrow RE}).$$

By augmenting the sequence  $x_0 < \dots < x_d$  by additional points a net is obtained, which converges finally to the value function  $v$ , the solution of (13).

## 5. Results

Table 2 contains our results. The first column presents the estimated maintenance costs ( $M_{OP}$ ) for a plant which is in the operating state. The estimates for the mean of this parameter range from

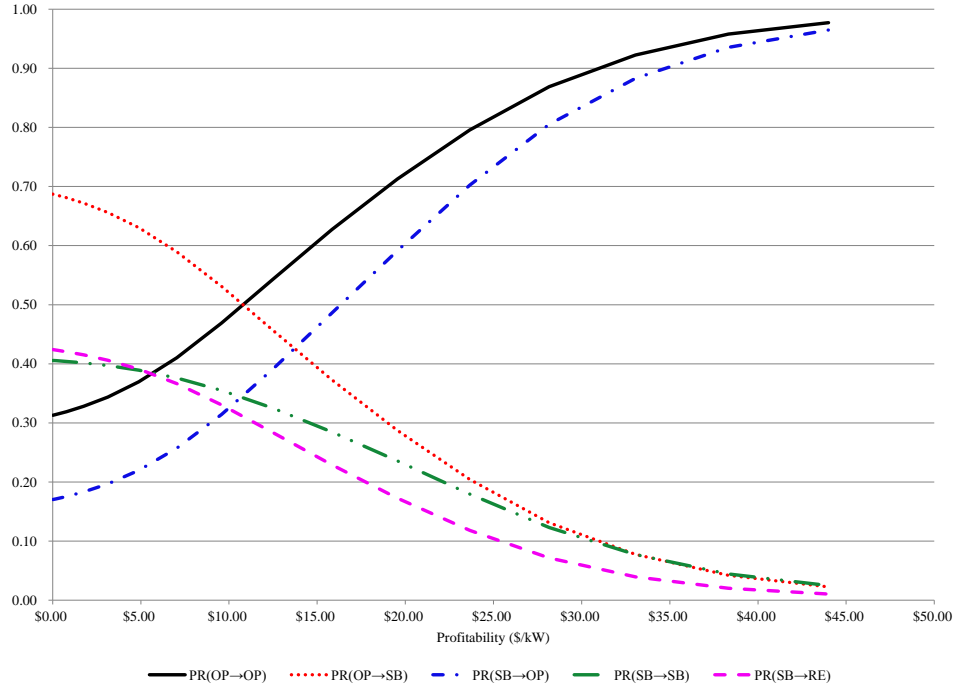
	$E(M_{OP})$	$\sigma_{M_{OP}}$	$M_{SB}$	$K_{OP \rightarrow SB}$	$E(K_{SB \rightarrow OP})$	$\sigma_{K_{SB \rightarrow OP}}$	$K_{SB \rightarrow RE}$
$\beta = 0.91$	12.4	7.18	0.00	0.00	12.3	4.90	-11.3
	(0.82)	(0.48)	(0.07)	(0.52)	(2.17)	(1.02)	(0.75)
$\beta = 0.95$	14.1	6.81	0.38	3.71	6.60	3.79	-15.7
	(0.43)	(0.26)	(0.09)	(0.86)	(0.94)	(0.53)	(0.73)

**Table 2** Maintenance and switching cost estimates for peak power plants, in units of \$/kW.  $M_{OP}$  is the ongoing cost of maintenance for an operating plant.  $M_{SB}$  is the ongoing cost of maintenance for a plant which has been shutdown.  $K_{OP \rightarrow SB}$  is the one-time cost of shutdown.  $K_{SB \rightarrow OP}$  is the one-time cost of start-up.  $K_{SB \rightarrow RE}$  is the one-time cost of abandonment. The numbers in parentheses are standard deviations of the estimates, calculated by (parametric) bootstrapping. Results are shown for two different discount factors  $\beta = 0.91$ , and  $\beta = 0.95$ .

\$12.4/kW to \$14.1/kW. The standard deviation in the maintenance costs ( $\sigma_{M_{OP}}$ ) is estimated from \$6.81/kW to \$7.18/kW, indicating wide variation in maintenance costs (and efforts) across plants. Unlike switching costs, estimates of ongoing maintenance costs are available from the EIA. These estimates provide a point of comparison for our results. According to the assumptions in the 2009 Annual Energy Outlook prepared by the EIA, the annual fixed maintenance costs for a conventional combustion turbine are \$12.11/kW. See Table 8.2 - Cost and Performance Characteristics of New Central Station Electricity Generating Technologies in EIA Energy Information Administration (2009), on page 89.

The third and fourth column presents estimated maintenance costs in the shutdown state ( $M_{SB}$ ) and shutdown costs ( $K_{OP \rightarrow SB}$ ) respectively. These numbers are very low, from \$0/kW to \$3.71/kW, and reflect the real world reality that the owner of a peaking plant can essentially turn the machine off and ignore it. Further, once the generator is shutdown management often chooses to invest little to nothing in its upkeep.

The fifth and sixth column of results in Table 2 indicate that the cost of restarting ( $K_{SB \rightarrow OP}$ ) are relatively large and variable, comparable with the cost of maintaining a unit ready for operation. Start-up costs depend greatly upon the maintenance performed on the plant while it was shutdown.



**Figure 3** The figure displays the probability of switching modes as a function of profitability (\$/kW).

Finally, the switching costs associated with an abandonment ( $K_{SB \rightarrow RE}$ ) are negative, that is the owner of the plant can expect to receive a cash inflow upon retirement of the plant. This result is consistent with the existence of a secondary market for used CTs. (These figures could also reflect the value of freeing the space for a replacement plant.) The salvage values reported in Table 2 range from \$11.3/kW to \$15.7/kW. According the the 2009 AEO, the cost for a brand new (conventional) CT is \$638/kW. Therefore our estimated salvage values range from approximately 1.8% to 2.5% of the cost of a brand new CT.

Figure 3 plots the transition probabilities as a function of profitability. The probabilities of a plant remaining in the operating state ( $OP \rightarrow OP$ ) and starting up ( $SB \rightarrow OP$ ) each increase in profitability, while the probabilities for remaining in the shutdown state ( $SB \rightarrow SB$ ), shutting down ( $OP \rightarrow SB$ ), and abandonment ( $SB \rightarrow RE$ ) each decrease in profitability.

Consider a generator in the operating state (*OP*). The probability of a plant remaining in the operating state (solid line) and the probability of shutting down (dotted line) cross (each equal 50%) at slightly above \$10/kW. Also at this level, the three possible decisions in the shutdown state (*SB*) are at about 32–34% probability. This means that at profitability indicator levels of around \$10/kW, it is hard for the model to predict which decision will be chosen, for both operating and standby plants.

### 5.1. The value function $v_g$

Our value function, which is estimated structurally, is calculated as an expectation over all decision makers equation (11). It is sometimes called “specific value function” in the economics literature. The value function is plotted in Figure 4 for three values of the interest rate, 5%, 10%, and 15%. Consider the 10% interest rate. The value function for the operating and standby states cross at approximately \$14/kW. Below this value, generators have higher value in the shutdown state, while above this value, the generators are worth more if it is operational.

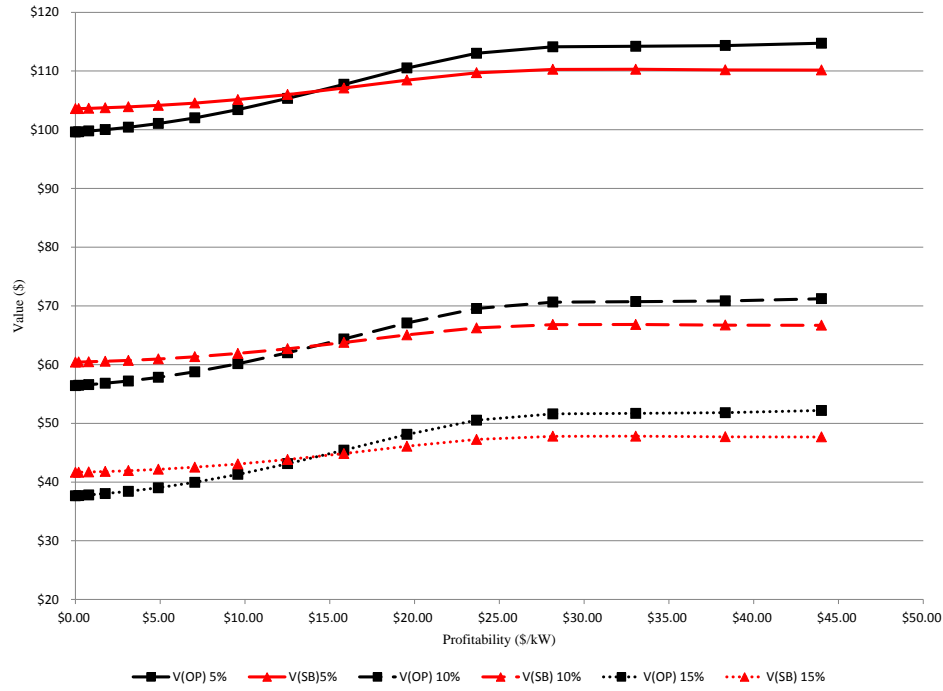
The specific levels at which the value functions cross in Figure 4 are not trigger levels as in a traditional real options model. Because we introduce unobserved heterogeneity, the crossing points are *averages* taken across all decision makers (cf. (5)). We can interpret the crossing point at \$14/kW by saying that plant managers, on average, have their machines operating if the returns exceed \$14/kW.

## 6. Conclusion

The owner of a peak power plant holds the real option to shutdown the plant. The owner of a plant which was previously shutdown holds the real options to either restart or retire the plant. Shutdown, start-up, and retirement entail one-time costs. Estimates of these switching costs are difficult to obtain in practice.

We use nonparametric structural estimation to obtain estimates of switching costs. Our results contribute by providing realistic estimates of these switching costs which can be accounted for by





**Figure 4** Value  $V(\cdot)$  as a function of profitability (\$/kW) for three values of the interest rate, 5%, 10%, and 15%.

both plant owners and for regulators concerned with ensuring system reliability. Comparison of our estimates of ongoing maintenance costs with published estimates gives us confidence in the results.

## Appendix

For the sake of completeness we state properties of Gumbel variables, which are used in the text. It is shown that the Gumbel distribution is closed under maximization (indeed, this is the essential property of any extreme value distribution). Further, a closed form formula for the probability of choice is provided for the Gumbel distribution. A very comprehensive discussion of extreme value distributions can be found in Embrechts et al. (1997).

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

**PROPOSITION 1 (The extreme value distribution is closed under maximization).** *Let  $(\varepsilon_i)_{i=1}^n$  be independent random variables which are Gumbel distributed with mean  $\mu_i$  and common*

scale parameter  $b > 0$ . Then the maximum  $\varepsilon := \max \{\varepsilon_i + c_i : i = 1, \dots, n\}$  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.

*Proof of Proposition 1* From the cumulative distribution function of the Gumbel distributions with respective means it follows that

$$\begin{aligned} P \left( \max_{i \in \{1, \dots, n\}} \varepsilon_i + c_i \leq z \right) &= P(\varepsilon_1 + c_1 \leq z, \varepsilon_2 + c_2 \leq z, \dots, \varepsilon_n + c_n \leq z) \\ &= \prod_{i=1}^n P(\varepsilon_i \leq z - c_i) = \prod_{i=1}^n \exp \left( -e^{-\frac{z - c_i - \mu_i}{b} - \gamma} \right) \\ &= \exp \left( -\sum_{i=1}^n e^{-\frac{z - c_i - \mu_i}{b} - \gamma} \right) = \exp \left( -e^{-\frac{z}{b} - \gamma} \cdot \sum_{i=1}^n e^{\frac{\mu_i + c_i}{b}} \right) \\ &= \exp \left( -e^{-\frac{z}{b} - \gamma} \cdot e^{\frac{\mu}{b}} \right) = \exp \left( -e^{-\frac{z - \mu}{b} - \gamma} \right), \end{aligned}$$

because  $\sum_{i=1}^n e^{\frac{\mu_i + c_i}{b}} = e^{\frac{\mu}{b}}$ . This reveals the assertion.  $\square$

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

**PROPOSITION 2 (Choice probabilities for shifted Gumbel variables).** *Let  $(\varepsilon_i)_{i=1}^n$  be independent Gumbel distributed random variables with individual mean  $\mu_i$  and common scale parameter  $b > 0$ . Then the probability of choice for the variables shifted by  $c_i$  is*

$$P \left( \varepsilon_1 + c_1 = \max_{i \in \{1, 2, \dots, n\}} \varepsilon_i + c_i \right) = \frac{\exp \left( \frac{c_1 + \mu_1}{b} \right)}{\exp \left( \frac{c_1 + \mu_1}{b} \right) + \dots + \exp \left( \frac{c_n + \mu_n}{b} \right)}. \quad (21)$$

*Proof of Proposition 2* Without loss of generality one may consider a pair  $(\varepsilon_1, \varepsilon_2)$  of independent Gumbel variables with location parameter 0, because the maximum in (21) itself is Gumbel distributed by Proposition 1.

Thus

$$\begin{aligned} P(\varepsilon_1 + c_1 \geq \varepsilon_2 + c_2) &= P(\varepsilon_2 \leq \varepsilon_1 + c_1 - c_2) \\ &= \int_{-\infty}^{\infty} f(x_1) \int_{-\infty}^{x_1 + c_1 - c_2} f(x_2) dx_2 dx_1 \\ &= \int_{-\infty}^{\infty} f(x_1) \exp \left( -e^{-\frac{x_1 + c_1 - c_2}{b}} \right) dx_1, \end{aligned} \quad (22)$$

where the cdf of the Gumbel distribution has been substituted. By substituting the probability density function (pdf)  $f$ , (22) continues as

$$\begin{aligned}
 P(\varepsilon_1 + c_1 \geq \varepsilon_2 + c_2) &= \int_{-\infty}^{\infty} \frac{1}{b} \exp\left(-\frac{x_1}{b} - e^{-\frac{x_1}{b}}\right) \exp\left(-e^{-\frac{x_1+c_1-c_2}{b}}\right) dx_1 \\
 &= \int_{-\infty}^{\infty} \frac{1}{b} e^{-\frac{x_1}{b}} \exp\left(-e^{-\frac{x_1}{b}} \left(1 + e^{-\frac{c_1-c_2}{b}}\right)\right) dx_1 \\
 &= \left[ \frac{\exp\left(-e^{-\frac{x_1}{b}} \left(1 + e^{-\frac{c_1-c_2}{b}}\right)\right)}{1 + e^{-\frac{c_1-c_2}{b}}} \right]_{x_1=-\infty}^{\infty} \\
 &= \frac{1}{1 + e^{-\frac{c_1-c_2}{b}}} = \frac{e^{\frac{c_1}{b}}}{e^{\frac{c_1}{b}} + e^{\frac{c_2}{b}}}.
 \end{aligned} \tag{23}$$

This completes the proof.  $\square$

Finally we give provide a proof that the difference of Gumbel variables enjoys a logistic distribution (cf. Nadarajah (2007)).

**COROLLARY 2.** *If  $\varepsilon_1$  and  $\varepsilon_2$  are Gumbel distributed with mean  $\mu_1$  and  $\mu_2$  and common scale parameter  $b > 0$ . Then the difference  $\varepsilon := \varepsilon_2 - \varepsilon_1$  follows a logistic distribution with mean  $\mu = \mu_2 - \mu_1$  and cumulative distribution function*

$$F_{\varepsilon}(z) = \frac{1}{1 + \exp\left(-\frac{z-\mu}{b}\right)}, \tag{24}$$

which is the distribution function of a logistic variable.

*Proof of Corollary 2* It follows from (23) in the proof of the preceding theorem that

$$F_{\varepsilon}(z) = P(\varepsilon_2 - \varepsilon_1 \leq z) = P(\varepsilon_1 + z \geq \varepsilon_2) = \frac{1}{1 + e^{-\frac{z-(\mu_2-\mu_1)}{b}}},$$

which completes the proof.  $\square$

## Endnotes

1. Note that we use the term *shutdown* to refer to what is sometimes called mothballing or laying up in the real options literature. We do not use shutdown to refer to overnight cycling of plants.
2. We thank Steve Marshall of Lakeland Electric and Paul D. Clark II of the City of Tallahassee for sharing their insights and experience. See Appendix B in Fleten et al. (2013) for a discussion of the real world problems associated with determining the costs of shutting down and starting up a peak power plant.

3. The index  $i$  refers to the observation ( $i = 1, \dots, 8, 189$ ) while the index  $k$  refers to the year. In what follows,  $X_k$  is the profitability in year  $k$  and is a random variable.  $X_i$  is the realization of the random variable and is one component of a plant-year observation. The other two components each observation are  $s_i$ , the state of the plant in the current year, and  $u_i$ , the state of the plant in the upcoming year. The state of the plant in the upcoming year,  $u \equiv s_{k+1}$ , therefore represents the *decision* of the plant manager.

4. The EIA provides variable definitions in a *Layout* file accompanying the EIA 860 data. The 2000 *Layout* file defines SB as “*Cold Standby (Reserve): deactivated (mothballed), in long-term storage and cannot be made available for service in a short period of time, usually requires three to six months to reactivate.*”

5. In practice the plant manager might rely upon production costing software (e.g., PROSYM, UPLAN, EGEAS, PowrSym) to simulate the operation of the regional electric system and therefrom derive an estimate of profitability for the upcoming year  $k + 1$ .

6. Computations were too burdensome when increasing the number of parameters affected by such heterogeneity, from two to three. We tried most combinations of parameters and ended up as indicated in the next section; with cross-plant heterogeneity in startup costs and for the maintenance cost of an plant ready for operation.

## Acknowledgments

We recognize the Norwegian research centre CenSES, Centre for Sustainable Energy Studies, and acknowledge support from the Research Council of Norway (RCN) through project 228811. A. Pichler acknowledges RCN project 207690.

## References

- Atuncar, G.S., C.C.Y. Dorea, C.R. Gonçalves. 2008. Strong consistency of Kernel density estimates for Markov chains failure rates. *Statistical Inference for Stochastic Processes* **11**(1) 1–10. doi:10.1007/s11203-006-9003-7.
- Bansal, Ravi, A Ronald Gallant, Robert Hussey, George Tauchen. 1995. Nonparametric estimation of structural models for high-frequency currency market data. *Journal of Econometrics* **66**(1) 251–287.

- Bensoussan, Alain, Jacques Louis Lions. 1984. *Impulse control and quasi-variational inequalities*. Gaunthier-Villars.
- Bontemps, Christian, Jean-Marc Robin, Gerard J Van den Berg. 2000. Equilibrium search with continuous productivity dispersion: Theory and nonparametric estimation. *International Economic Review* **41**(2) 305–358.
- Brekke, K., B. Øksendal. 1994. Optimal switching in an economic activity under uncertainty. *SIAM Journal on Control and Optimization* **32**(4) 1021–1036. doi:10.1137/S0363012992229835.
- Brennan, M.J., E.S. Schwartz. 1985. Evaluating natural resource investments. *The Journal of Business* **58**(2) 135–157.
- Burr, Chrystie. 2012. Subsidies, tariffs and investments in the solar power market. *Working Paper, University of Arizona*.
- Caldecott, Ben, Jeremy McDaniels. 2014. Stranded generation assets: Implications for European capacity mechanisms, energy markets and climate policy. Working paper, University of Oxford.
- Deng, Shi-Jie, Blake Johnson, Aram Sogomonian. 2001. Exotic electricity options and the valuation of electricity generation and transmission assets. *Decision Support Systems* **30**(3) 383–392.
- Embrechts, Paul, Claudia Klüppelberg, Thomas Mikosch. 1997. *Modelling Extremal Events, Stochastic Modelling and Applied Probability*, vol. 33. Springer. URL [books.google.com/books?id=BX0I2pICfJUC](http://books.google.com/books?id=BX0I2pICfJUC).
- Energy Information Administration. 2009. Assumptions to the Annual Energy Outlook 2009 (DOE/EIA-0554(2009)).
- Fleming, Wendell Helms, Halil Mete Soner. 2006. *Controlled Markov Processes and Viscosity Solutions*. Springer-Verlag.
- Fleten, Stein-Erik, Erik Haugom, Carl J. Ullrich. 2013. Keeping the lights on until the regulator makes up his mind. Working paper, Available at SSRN: <http://ssrn.com/abstract=192861>.
- Gamba, Andrea, Matteo Tesser. 2009. Structural estimation of real options models. *Journal of Economic Dynamics & Control* **33** 798–816.
- Graf, Siegfried, Harald Luschgy. 2000. *Foundations of Quantization for Probability Distributions, Lecture Notes in Mathematics*, vol. 1730. Springer-Verlag Berlin Heidelberg.

- Guerre, Emmanuel, Isabelle Perrigne, Quang Vuong. 2000. Optimal nonparametric estimation of first-price auctions. *Econometrica* **68**(3) 525–574.
- Kallenberg, Olav. 2002. *Foundations of Modern Probability*. Springer.
- Kellogg, Ryan. 2014. The effect of uncertainty on investment: Evidence from Texas oil drilling. *American Economic Review* **104** 1698–1734.
- Kenyon, Chris, Stathis Tompaidis. 2001. Real options in leasing: The effect of idle time. *Operations Research* **49**(5) 675–689.
- King, Gary, Langche Zeng. 2001. Logistic regression in rare events data. *Political Analysis* **9**(2) 137–163.
- Kwon, H Dharma. 2010. Invest or exit? optimal decisions in the face of a declining profit stream. *Operations Research* **58**(3) 638–649.
- Lai, Guoming, Mulan X Wang, Sunder Kekre, Alan Scheller-Wolf, Nicola Secomandi. 2011. Valuation of storage at a liquefied natural gas terminal. *Operations Research* **59**(3) 602–616.
- Li, Tong, Isabelle Perrigne, Quang Vuong. 2002. Structural estimation of the affiliated private value auction model. *RAND Journal of Economics* 171–193.
- Lin, C-Y Cynthia, K Thome. 2013. Investment in corn-ethanol plants in the Midwestern United States: An analysis using reduced-form and structural models. Working Paper, University of California at Davis, Davis.
- McFadden, Daniel. 1974. *Conditional logit analysis and qualitative choice behavior*, chap. 4. Academic Press, New York, 105–142.
- Moel, Alberto, Peter Tufano. 2002. When are real options exercised? an empirical study of mine closings. *Review of Financial Studies* **15**(1) 35–64.
- Muehlenbachs, Lucija. 2015. A dynamic model of cleanup: Estimating sunk costs in oil and gas production. *International Economic Review* **56**(1) 155–185.
- Nadarajah, Saralees. 2007. Linear combination of gumbel random variables. *Stochastic Environmental Research and Risk Assessment* **21**(3) 283–286. doi:10.1007/s00477-006-0063-4.
- Newey, Whitney K, James L Powell, Francis Vella. 1999. Nonparametric estimation of triangular simultaneous equations models. *Econometrica* **67**(3) 565–603.

- Pageès, Gilles. 1998. A space quantization method for numerical integration. *Journal of Computational and Applied Mathematics* **89**(1) 1 – 38. doi:10.1016/S0377-0427(97)00190-8.
- Rapson, David. 2014. Durable goods and long-run electricity demand: Evidence from air conditioner purchase behavior. *Journal of Environmental Economics and Management* **68**(1) 141–160.
- Rothwell, Geoffrey, John Rust. 1997. On the optimal lifetime of nuclear power plants. *Journal of Business & Economic Statistics* **15**(2) 195–208.
- Rust, John. 1987. Optimal replacement of GMC bus engines: An empirical model of Harold Zurcher. *Econometrica* **55**(5) 999–1033. URL <http://www.jstor.org/stable/1911259>.
- Rust, John, Geoffrey Rothwell. 1995. Optimal response to a shift in regulatory regime: The case of the US nuclear power industry. *Journal of Applied Econometrics* **10** S75–S118.
- Shiryaev, Albert Nikolayevich. 1996. *Probability*. Springer, New York.
- Silverman, Bernard W. 1998. *Density Estimation for Statistics and Data Analysis*. London: Chapman & Hall/CRC.
- Smith, James E, Kevin F McCardle. 1998. Valuing oil properties: integrating option pricing and decision analysis approaches. *Operations Research* **46**(2) 198–217.
- Smith, Rebecca, Lynn Cook. 2015. Hawaii wrestles with vagaries of a green grid. *The Wall Street Journal* **June 29, 2015** B1.
- Strebulaev, Ilya A., Toni M. Whited. 2012. Dynamic models and structural estimation in finance. *Foundations and Trends in Finance* **6** 1–163.
- Su, Che-Lin, Kenneth L. Judd. 2012. Constrained optimization approaches to estimation of structural models. *Econometrica* **80** 2213–2230.
- Thompson, M., M. Davison, H. Rasmussen. 2004. Valuation and optimal operation of electric power plants in competitive markets. *Oper. Res.* **52**(4) 546–562.
- Train, Kenneth. 2002. *Discrete choice methods with simulation*. Cambridge University Press.
- Tseng, C.L., G. Barz. 2002. Short-term generation asset valuation: A real options approach. *Operations Research* **50**(2) 297–310.