Manufacturing & Service Operations Management

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Dynamic Capacity Investment with Two Competing Technologies

Wenbin Wang, Mark E. Ferguson, Shanshan Hu, Gilvan C. Souza

M&SOM MANUFACTURING & SERVICE OPERATIONS MANAGEMENT
Vol. 15, No. 4, Fall 2013, pp. 616–629
ISSN 1523-4614 (print) | ISSN 1526-5498 (online)

http://dx.doi.org/10.1287/msom.2013.0438

Dynamic Capacity Investment with Two Competing Technologies

Wenbin Wang
Department of Operations Management, School of International Business Administration,
Shanghai University of Finance and Economics, 200433 Shanghai, China,
wang.wenbin@shufe.edu.cn

Mark E. Ferguson
Management Science Department, Moore School of Business, University of South Carolina, Columbia, South Carolina 29208,
mark.ferguson@moore.sc.edu

Shanshan Hu, Gilvan C. Souza
Department of Operations and Decision Technologies, Kelley School of Business, Indiana University, Bloomington, Indiana 47405
{hush@indiana.edu, gsouza@indiana.edu}

With the recent focus on sustainability, firms making adjustments to their production or distribution capacity levels often have the option of investing in newer technologies with lower carbon footprints and/or energy consumption. These more sustainable technologies typically require a higher up-front investment but have lower operating (fuel or energy) costs. What complicates this decision is the fact that the projected dollar savings from the more sustainable technologies fluctuate considerably due to uncertainty in fuel prices, and the total capacity may not be utilized at 100% because of fluctuations in the demand for the product. We consider the firm's capacity adjustments over time given a portfolio of technology options when the demand and the fuel costs are stochastic and possibly dependent. Our model also allows for usage-based capacity deterioration. We provide the analytical structure of the optimal policy, which assigns different control limits for investing, staying put, and disinvesting in the capacities of the competing technology choices for each realization of demand and fuel costs at each period. We also present an application of our model to the problem of designing a delivery truck fleet for a beverage distributor.

Key words: sustainable operations; dynamic capacity investment; technology choice
History: Received: October 15, 2011; accepted: February 18, 2013. Published online in *Articles in Advance* June 14, 2013.

1. Introduction

From an environmental perspective, a promising trend in recent years has been the number of firms who make investments in energy saving technologies for the production and distribution of their products. These technologies often require a higher up-front investment than traditional technologies, but promise lower operating costs over their life span. As one of many examples, in 2007 Coca-Cola Enterprises (CCE) started replacing some of their fleet of diesel-based delivery trucks with diesel-electric hybrid vehicle (HEV) trucks. A large firm such as CCE typically maintains a fleet of hundreds of delivery vehicles. A traditional diesel truck costs approximately $60,000, whereas an HEV truck costs $94,000. The HEV option, however, provides significant savings in fuel over its 12-year life span (for a typical CCE truck that drives 18,000 miles per year, the savings are about 100 gallons of diesel per month). The challenge for a manager here is how to dynamically make long-term capacity adjustment decisions among the available technologies.

One aspect of this problem that makes it challenging is that fuel prices fluctuate considerably. For example, the average price of a gallon of diesel in the United States varied between $2.84 and $4.73 between July 2007 and July 2008. A second complication is that, because of seasonal and random fluctuations in demand, the total capacity is rarely utilized at 100%. For example, CCE's demand for soft drinks can be two to three times higher in the summer months compared with the winter months. CCE does have the option of renting additional capacity during peak demand, which can be viewed as a penalty cost for demand in excess of capacity. In addition, the deterioration of a particular capacity type depends on its usage, which adds additional uncertainty to the projected future capacity needs. Thus, the projected future capacity needs and the projected yearly savings from the more sustainable technologies depend on the underlying assumptions about fuel prices, demand, and how the technologies will be utilized.

With two technology choices, the problem becomes even more difficult. A common method used in practice for choosing between a traditional and a sustainable technology is to use a net present value (NPV) comparison of the two competing technologies, in terms of initial investment and operating costs; and after the technology is selected, a newsvendor-type analysis can be used to determine the amount of capacity to acquire. The NPV calculation usually assumes 100% capacity utilization and ignores usage-based capacity deterioration, which results in an either/or solution: either all HEVs or all diesel trucks are purchased at any period. Given the realities previously mentioned, however, it is reasonable that a mixed portfolio investment should be adopted. This is because the business case for the sustainable technology depends on its expected utilization and, because of demand uncertainty and seasonality, capacity units will have different expected utilizations.

We contribute to the literature by providing a decision support model for making dynamic capacity adjustment decisions with two competing technologies available for meeting uncertain demand, where one technology has a lower operating cost but a higher acquisition cost than the other. Operating costs (fuel and maintenance costs) are uncertain, and can be correlated with demand, as well as with the penalty cost for demand in excess of capacity. The model allows for usage-based capacity deterioration. As we discuss in §2, the literature provides little guidance for decision making under these conditions, which are realistic for firms such as CCE. From a methodological standpoint, we contribute to the literature on dynamic capacity adjustment by providing the structure of the optimal policy, which assigns different control limits for investing, staying put, and disinvesting in the capacities of the competing technology choices. These control limits are different for each realization of demand and fuel costs in a period. Our methodological contribution is novel because existing research on capacity adjustment with two capacity choices (Dixit 1997, Eberly and Van Mieghem 1997) considers two complementary capacity choices (e.g., labor and materials), deterministic operating cost, and no capacity deterioration; in contrast, our capacity choices are substitutes, our operating cost is random, and there is capacity deterioration from usage. A key insight from our numerical study is that, with seasonal demand, the firm optimally maintains a balanced portfolio of capacity types—in the CCE example, about 54% HEV for the base case. If the firm heuristically maintains an HEV-only fleet (with capacity adjustment dynamically optimized), then its total discounted cost (acquisition and operating costs) is only 1.24% lower than optimal; an HEV-only fleet is optimal if there is no demand seasonality. Thus, environmentally friendly technologies are also economically attractive.

This paper is organized as follows. In §2, we position our contribution with respect to the literature. We state our assumptions and formulate our model in §3. We present the general analytical structure of the optimal policy in §4. In §5, we present the results of a numerical study that is representative of CCE's case. We discuss an extension to more than two technologies in §6, and we conclude in §7.

2. Literature Review

Our research is related to the literature on dynamic capacity adjustment as well as to a recent stream of literature on the adoption of sustainable technologies.

The literature on dynamic capacity adjustment is broad, and models that incorporate stochastic demand go back to Manne (1961); see Luss (1982) for a review of earlier research. More recently, economists have approached capacity expansion as a dynamic investment under uncertainty, exploring the embedded option values. Option pricing techniques are often used in continuous time, stochastic control models. Pindyck (1986) considers incremental capacity investments; later extensions are summarized by Dixit et al. (1994). In contrast, researchers in operations management (OM) generally approach capacity expansion as a discrete-time stochastic dynamic program (DP). A comprehensive review of this literature, from both economics and OM perspectives, is provided by Van Mieghem (2003). We review here only the most related research to our problem.

Most papers consider adjustment in capacity for a single technology and assume that capacity is irreversible (no capacity salvaging) and durable (no deterioration). A small subset of the literature considers capacity deterioration, as we do. In a continuous-time model, Dixit et al. (1994, Chap. 11.1.F) model capacity deterioration as events from a Poisson process and where the deterioration rate (applied to each unit of capacity) is constant over time. Rajagopalan (1998) and Rajagopalan et al. (1998) model deterioration in a capacity expansion problem with the possibility of equipment replacement. Both papers assume that demand is deterministic and increasing over time, which is not realistic in our setting where demand is seasonal and stochastic. Chao et al. (2009) consider a dynamic capacity expansion problem with random demand and where the capacity acquisition is instantaneous but constrained to a random supply; deterioration rates are random or deterministic, but not usage dependent as in our case.

In contrast to the stream of research above, our paper considers two alternative capacity technologies.

In a static setting, earlier work considering multiple technologies has focused on the capacity investment decision for electricity producers. Crew and Kleindorfer (1976) study the problem of optimal pricing and capacity planning when the available plant types (technologies) have different investment and generating costs, and when demand is stochastic and price dependent. Their model was later extended by including supply uncertainty (i.e., random availability of the installed capacity) in Chao (1983) and Kleindorfer and Fernando (1993). Drake et al. (2010) consider the capacity investment decision for two technologies with different emission intensities. They consider uncertain prices for emissions allowance, so that the stochastic operating costs for the two technologies are perfectly correlated. Similarly to our paper, this stream of research includes stochastic demand, multiple technologies, and a technology deployment policy based on increasing operating costs. The key difference is that we consider dynamic capacity investment and, as a result, our research is more related to the multifactor dynamic capacity expansion and contraction literature. Dixit (1997) and Eberly and Van Mieghem (1997) independently characterize the optimal solution as a control-limit policy. As explained later in §4, our optimal policy is also a control-limit type, but with a crucial difference: the two technologies considered in our research are substitutes (e.g., conventional diesel versus HEV trucks), whereas the two above referenced papers consider complementary production factors such as labor and materials. In addition, our model includes uncertain operating cost and capacity deterioration. Uncertain operating cost in production planning has been considered, in different settings, in papers by Ding et al. (2007), Guo et al. (2011), Huchzermeier and Cohen (1996), Kazaz et al. (2005), and Plambeck and Taylor (2011), among others. Contrary to our paper, in this stream of research the technology choice is absent in that only a single technology is considered. Kleindorfer et al. (2012) consider two technologies (electric and conventional) in a fleet renewal problem. In their setting, however, demand is deterministic and capacity is fully utilized, so that a technology's expected operating cost can be computed at the time of its acquisition. This results in an optimal policy where only one technology type is selected in each period. In contrast, the seasonal and stochastic nature of demand in our problem, coupled with our assumption that the total operating cost per unit (including deterioration) of the environmentally superior technology is lower than the conventional technologies, imply an optimal policy where the firm often invests in both technologies.

Finally, there is a stream of literature that studies investment in sustainable technologies; for a review, see Jaffe et al. (2003). Most of this literature, however, considers a deterministic setting and a single-technology choice. With multiple technologies, recent research investigates policy incentives for investing in renewable energy technologies (solar and wind), both of which have issues with supply uncertainty (intermittency). Ambec and Crampes (2010) and Garcia et al. (2012) approach this problem under deterministic demand, whereas Aflaki and Netessine (2011) assume stochastic demand. All three papers consider two technology choices: renewable/intermittent and conventional/reliable. Their primary focus is on the interaction of the two technologies (renewable and nonrenewable) and its impact on the capacity investment decisions in a single-period setting. Unlike these papers, we consider a dynamic setting.

3. Model

We consider a dynamic capacity adjustment problem. The planning horizon is infinite, comprised of periods of equal lengths.

3.1. Technology-Related Costs

For each period, the random variable X denotes demand, which has support $[\underline{x}, \bar{x}]$ and follows a discrete-time Markov process. Capacity to meet this demand can be acquired from two competing technologies. Technology c is the conventional technology, whereas technology e is the environmentally superior one. Technologies differ in their up-front unit acquisition cost v_i and a variable operating cost that is proportional to the capacity usage in the period as follows. For each unit of capacity for technology i, there is a stochastic fuel cost P_i and a maintenance cost m_i. The sum of these two usage-based variable costs $P_i + m_i$ is referred to as the *operating cost* per unit, and we assume it to be bounded. The conventional technology has a lower acquisition cost ($v_c < v_e$), but a significantly higher fuel cost so that its operating cost is higher: $P_c + m_c > P_e + m_e$. Denote $\mathbf{P} = (P_c, P_e)$. The fuel cost $\mathbf{P}$ is modeled as a discrete-time mean-reverting stochastic process. We allow for correlation between $\mathbf{P}$ and X: Their joint probability distribution in a period may depend on their realizations in the previous period. In the CCE example, X is the number of trucks required to meet demand in the current period (month), the unit of capacity is trucks, P_c is the monthly fuel cost for a conventional diesel truck (at 100% utilization), and P_e is the monthly fuel cost for an HEV, so that $P_e = \lambda P_c$ with $\lambda < 1$.

Throughout this paper, we use the convention that lowercase (uppercase) symbols denote deterministic (stochastic) variables and parameters. Hence, the realization of a random variable is expressed in lowercase, and the inequality of the two operating costs above holds in the probabilistic sense (i.e., almost surely, or a.s.). Our main results hold when v_i and m_i are random, but our focus is on the impact of stochastic demand and operating cost.

3.2. Sequence of Events

In each period, the firm observes its capacity $\mathbf{k} = (k_c, k_e)$, the demand x, and fuel cost $\mathbf{p}$ for that period. The firm then makes its capacity adjustment decision. Our model allows for capacity to be increased or decreased. Capacity reduction (salvaging) may occur due to trends such as diminishing demand. Salvaging one unit of capacity of type i has a salvage value $s_i < v_i$. The possibility of capacity salvaging distinguishes this paper from earlier work on irreversible capacity expansion. Instead, we use a model framework more aligned to research on multifactor dynamic capacity expansion and contraction (Dixit 1997, Eberly and Van Mieghem 1997). Denote by $y_i \in [0, \bar{x}]$ the capacity level after adjustment for technology i, and $\mathbf{y} = (y_c, y_e)$. The capacity adjustment cost is

$$V(\mathbf{k}, \mathbf{y}) := \sum_{i \in \{c, e\}} \left[v_i (y_i - k_i)^+ - s_i (k_i - y_i)^+ \right]. \qquad (1)$$

The firm then uses $\mathbf{y}$ to meet demand x. Any demand realization that exceeds the firm's capacity is either lost or met using an outsourced option and incurs a unit penalty cost Φ. The interpretation of Φ depends on the application context: Φ could represent the profit margin for losing one unit of demand, or, in the case of CCE, it represents the cost of renting additional capacity. In the latter case, Φ is likely to be correlated with $\mathbf{P}$; for generality, Φ is a random variable.

3.3. Capacity Deterioration

A novel feature of our model is the incorporation of usage-based capacity deterioration in a tractable manner. Ideally, one would track the usage of each capacity unit and optimize capacity investment decisions based on usage across all units. For a fleet of hundreds of trucks, as is the case of CCE, that would imply creating a state variable for each truck to track its usage, which is computationally intractable. We develop a computationally tractable model based on two simplifications: (i) capacity is assumed to be perfectly divisible, and (ii) capacity deterioration occurs at the aggregate level, for each distinct technology. The first simplification is standard in the dynamic capacity investment literature. Simplification (ii) has been used in the literature on dynamic capacity investment with a single technology; see Dixit et al. (1994, Chap. 11) and Chao et al. (2009). Similar to this literature, we assume that in each period, a fixed portion of the capacity is obsolete, and it is referred to as *natural deterioration*, but in contrast to this literature, our model allows for additional capacity loss due to its usage. Specifically, for technology i, denote by γ_i the *natural deterioration rate* and by β_i the *usage-based deterioration rate*. Define $\theta_i(\mathbf{y}, x)$ as technology i's capacity usage (a function of total capacity and demand, as discussed later). Given the capacity level after adjustment at the beginning of the period as y_i, the firm loses capacity in that period by an amount $\gamma_i y_i$ due to natural deterioration, and by an amount $\beta_i \theta_i(\mathbf{y}, x)$ due to usage. Hence, the initial capacity next period before any adjustment is

$$\kappa_i(\mathbf{y}, x) := y_i - \gamma_i y_i - \beta_i \theta_i(\mathbf{y}, x) \quad \text{for } i \in \{c, e\}. \qquad (2)$$

We impose the following assumptions:

Assumption 1. $\gamma_i + \beta_i \le 1$ *for both technologies* $i = c$ *and* e.

Assumption 2. $P_c + m_c + \beta_c v_c \le_{\text{a.s.}} \Phi$ *for any period*.

Assumption 3. $(P_c + m_c) - (P_e + m_e) \ge_{\text{a.s.}} \beta_e v_e - \beta_c s_c$ *for any period*.

Assumption 1 guarantees that the total capacity deterioration does not render obsolete the entire capacity in one period (i.e., it ensures that κ_i is nonnegative, since $\theta_i \le y_i$, as we see below). Assumption 2 ensures that it is never optimal to leave demand unmet and take the penalty cost when the conventional technology c is available—the cost of meeting one unit of demand with technology c is equal to the operating cost $(P_c + m_c)$ plus the usage-based deterioration cost $\beta_c v_c$, as it costs the firm v_c to replace one unit of capacity. Assumption 3 guarantees a fixed and intuitive priority ordering of the two technologies to meet demand, as described in Lemma 1 (all proofs are provided in the e-companion, available as supplemental material at http://dx.doi.org/10.1287/msom.2013.0438); we interpret its meaning next.

Lemma 1. *It is always optimal to use the environmentally superior technology e first to meet demand*. (*Technology c is only used if technology e is fully utilized*.)

In Assumption 3, the left-hand side is positive by definition, because the operating cost for technology c is higher than for technology e. The first term in the right-hand side is the depreciation cost of using technology e alone. By using technology e, however, one unit of capacity of technology c goes unused. The only value that one unit of unused capacity has when total supply exceeds total demand is the value of salvaging the excess capacity s_c. Thus, the true marginal depreciation cost of using one unit of technology e over technology c is $\beta_e v_e - \beta_c s_c$. Assumption 3 states that the operating cost savings of using technology e over technology c is no less than the marginal depreciation cost of doing so. Assumption 3 is not very restrictive: for the CCE example, it is satisfied as long as the annualized usage-based deterioration rates (β_e and β_c) are less than 7.5%.

3.4. Problem Formulation

Lemma 1 specifies a simple priority dispatching policy, which simplifies the problem formulation. The capacity usage in a period is

$$\theta_i(\mathbf{y}, x) := \begin{cases} \min\{(x-y_e)^+, y_c\} & \text{for } i = c, \\ \min\{x, y_e\} & \text{for } i = e. \end{cases} \tag{3}$$

The one-period cost, excluding capacity adjustment cost, is

$$C(\mathbf{y}, \mathbf{p}, x) := \sum_{i\in\{c,e\}} (p_i + m_i)\theta_i(\mathbf{y}, x) + (x - (y_c + y_e))^+\Phi. \tag{4}$$

The dynamic program can be formulated as follows. Let $G(\mathbf{y}, \mathbf{p}, x)$ represent the total expected discounted cost in an infinite horizon at state $(\mathbf{y}, \mathbf{p}, x)$ (after incurring adjustment cost), and $J(\mathbf{k}, \mathbf{p}, x)$ represent the total expected discounted cost in an infinite horizon at state $(\mathbf{k}, \mathbf{p}, x)$ (before capacity adjustment). The one-period cost functions C and V are continuous and bounded, given our assumptions of a bounded operating cost. We further assume that the evolution of $\mathbf{P}$ and X satisfies the Feller property.[1] Then, Stokey et al. (1989, Theorem 9.6) implies that there exists a unique stationary value function J satisfying

$$G(\mathbf{y}, \mathbf{p}, x) = C(\mathbf{y}, \mathbf{p}, x) + \alpha \mathsf{E}_{\mathbf{P}, X}\big[J(\boldsymbol{\kappa}(\mathbf{y}, x), \mathbf{P}, X) \mid \mathbf{p}, x\big] \tag{5}$$

and

$$J(\mathbf{k}, \mathbf{p}, x) = \min_{0\le \mathbf{y}\le \bar{x}} \big\{V(\mathbf{k}, \mathbf{y}) + G(\mathbf{y}, \mathbf{p}, x)\big\}. \tag{6}$$

where $\boldsymbol{\kappa} = (\kappa_c, \kappa_e)$ denotes the vector of functions defined in (2). (See the second item of the e-companion for details about the Bellman operator and the existence and uniqueness of J.)

4. Optimal Capacity Adjustment Policy

In this section, we characterize the structure of the optimal policy for the dynamic program. Our assumptions imply both a simple priority dispatching policy (Lemma 1), and convexity of $G(\mathbf{y}, \mathbf{p}, x)$ in $\mathbf{y}$ (Proposition 1(a)), which in turn drives the structure of the optimal policy specified in Proposition 1(c) and motivated as follows. In any period, let G represent the cost-to-go function excluding capacity adjustment costs. Convexity and continuity of G in $\mathbf{y}$ imply that its partial derivatives exist almost anywhere. Thus, $-\partial G/\partial y_i$, where it is well defined, can be interpreted as the *marginal benefit* (in the form of cost saving) from increasing the capacity of technology i. The firm increases capacity as long as this marginal benefit exceeds the acquisition cost. Similarly, the firm reduces capacity as long as this marginal benefit is less than the salvage value. In other words, the optimal capacity $\mathbf{y}^*(\mathbf{k}, \mathbf{p}, x)$ satisfies

$$s_i \le -\frac{\partial}{\partial y_i} G(\mathbf{y}^*, \mathbf{p}, x) \le v_i \quad \text{for } i = c, e. \tag{7}$$

Condition (7) gives rise to a control limit policy: If $\mathbf{k}$ satisfies (7) for technology i, then the optimal policy is to not adjust i's capacity, $y_i^* = k_i$. Otherwise, i's capacity is adjusted to a level that either sets the left inequality binding after capacity salvaging ($y_i^* < k_i$) or sets the right inequality binding after capacity investment ($y_i^* > k_i$). Note that these control limits depend on $(\mathbf{p}, x)$. This type of control limit policy is called an ISD (invest/stay put/disinvest) policy, first defined by Eberly and Van Mieghem (1997). We adapt their definition to our context:

Definition 1. A policy $\mathcal{Y}(\mathbf{p}, x)$, for a given prevailing fuel cost and demand observations $\mathbf{p}$ and x, is an *ISD policy* if, for each technology $i = c, e$, there exist two critical functions $k_i^L(k_j) \le k_i^R(k_j)$ with $j \ne i$. The two critical functions define three action regions: capacity i is increased to $y_i^* = k_i^L$ if $k_i < k_i^L$ (invest), decreased to $y_i^* = k_i^R$ if $k_i > k_i^R$ (disinvest), or remains the same $y_i^* = k_i$ otherwise (stay put).

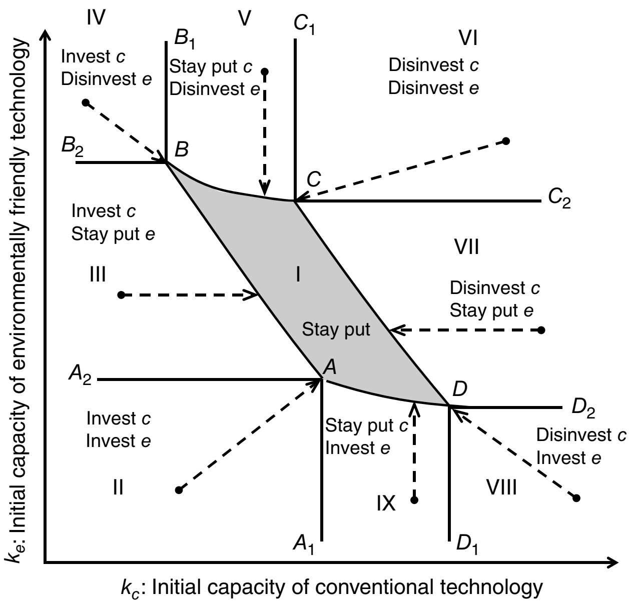
We formally state the above results in Proposition 1:

Proposition 1. (a) $G(\mathbf{y}, \mathbf{p}, x)$ *is convex in* $\mathbf{y}$;
(b) $J(\mathbf{k}, \mathbf{p}, x)$ *is convex in* $\mathbf{k}$; *and*
(c) *the optimal capacity adjustment follows an ISD policy* $\mathcal{Y}(\mathbf{p}, x)$.

Figure 1 illustrates the ISD policy for a given $(\mathbf{p}, x)$: For technology c, curves A_1ABB_1 and D_1DCC_1 correspond to the left and right critical functions $k_c^L(k_e)$ and $k_c^R(k_e)$, respectively. For technology e, curves A_2ADD_2 and B_2BCC_2 correspond to the left and right critical functions $k_e^L(k_c)$ and $k_e^R(k_c)$. These four curves segment the plane into nine zones. If the initial capacity pair (k_c, k_e) falls into the central zone I, then the firm makes no capacity adjustment. If the initial capacity pair is in zones II, IV, VI, and VIII, then the optimal decision is to adjust the capacities to the closest corner of zone I, namely, points A, B, C, and D, respectively. If the initial capacity pair is in zones III, V, VII, and IX, then the optimal decision is to keep capacity of one technology unchanged while increasing/decreasing the capacity of the other technology to the boundary of zone I. The structure shown in Figure 1 enables an efficient numerical solution of the dynamic program, as we discuss in §5.

[1] See Stokey et al. (1989, p. 220). A sufficient condition of the Feller property in our model is to assume that the joint probability density function of $\mathbf{P}$ and X is continuous and bounded.

Figure 1 An ISD Policy: Move to the Nearest Point in the Grey Region



Comparing Figure 1 with Dixit (1997, Figure 1) and Eberly and Van Mieghem (1997, Figure 2), we note an important distinction of our ISD policy: The boundaries of the stay-put zone I are downward sloping, so that the target capacity level for a technology is lower at a higher capacity level of the other technology. In contrast, in Dixit (1997) and Eberly and Van Mieghem (1997), the target capacity level of a technology moves in the same direction as the initial capacity level of the other technology. This difference is structural: In those two papers, the different capacity factors are used as complements because their profit functions (to be maximized) are supermodular, or the equivalent cost functions (to be minimized) are submodular. In contrast, in our setting the two competing technologies are used as substitutes. Our cost-to-go function after minimization, J, is supermodular, as formalized below:

Proposition 2. (a) $G(\mathbf{y}, \mathbf{p}, x)$ *is supermodular in* $\mathbf{y}$; (b) $J(\mathbf{k}, \mathbf{p}, x)$ *is supermodular in* $\mathbf{k}$; *and*
(c) *the ISD control limits* $k_i^L(k_j)$ *and* $k_i^R(k_j)$ *are decreasing in* k_j *for* $i = c, e$, *and* $j \neq i$.

5. Application: Coca-Cola Enterprises

In this section, we apply our model to the problem of designing a delivery truck fleet for the four largest distribution centers of CCE in North America.[2] The objective is to derive managerial insights from the application of our model to a real-world situation.

[2] These distribution centers are now part of the Coca-Cola Company.

Table 1 Technology Specification

	Conventional diesel (c)	HEV (e)
Monthly travel distance (miles/month)	1,500	1,500
Fuel efficiency (miles/gallon)	4.75	6.98
Monthly fuel consumption (gallons/month)	316	215
Acquisition cost v_i ($/truck)	60,000	94,000[a]
Maintenance cost ($/mile)	0.03	0.03
Maintenance cost m_i ($/month-truck)	45	45
Salvage value s_i (dollars/truck)	25,000	35,000

[a] This value includes the battery replacement cost through the life cycle.

We use secondary data from various sources to estimate and calibrate model parameters. A period here is one month.

5.1. Parameter Calibration and Setup

The parameters indicated in this section apply to the entire numerical study unless noted otherwise in §5.6. The four largest CCE distributors in North America deliver about 10 million cases of soft drinks per year. The unit of demand is *monthly truckload*, which is equivalent to the average number of soft drink cases that a truck delivers in one month ($30\frac{1}{3}$ days). The daily delivery load for a CCE driver is 450 cases,[3] so that one demand unit is equivalent to $450 \times 30\frac{1}{3} = 13{,}650$ cases per month. This means, average yearly demand is 732.5 monthly truckloads.

CCE uses two types of trucks, conventional diesel and HEV, whose parameter values are displayed in Table 1. We consider an annual natural deterioration rate of 2% and an annual used-based deterioration rate of 7%. The discount rate for future costs is 7.5% per year, which is CCE's approximate weighted average cost of capital.[4]

5.1.1. Demand Process. Soft drink sales exhibit strong seasonality. Thus, the demand process can be described as $X_t = \underline{X} + \sum_{i=1}^{11} a_i x_{it} + \gamma \varepsilon_t$, where a_i is a seasonal component corresponding to month i, and x_{it} is 1 if t is month i of the year and 0 otherwise. For computational simplicity, and to simulate correlation between fuel cost and demand, the random term ε_t was set equal to the diesel price per gallon D_t, whose process is described below, and $\gamma = 1$. We have obtained detailed demand data for a particular CCE market and used it to estimate CCE's seasonal parameters as follows. Using the model $X'_t = \underline{X}' + \sum_{i=1}^{11} a'_i x_{it}$, we estimate the parameters $\underline{X}'$ and a'_i from our data using regression. We then rescale $\underline{X}'$ and a'_i so that total yearly demand equals CCE's 732.5 monthly truckloads.

[3] CCE's careers, driver openings (http://www.cokecce.com/pages/allContent.asp?page_id=167; accessed October 1, 2011).

[4] http://www.wikiwealth.com/wacc-analysis:cce (accessed October 1, 2011.

5.1.2. Fuel Cost Processes. Mean-reverting processes are widely used in energy research to model spot fuel prices (e.g., Bessembinder et al. 1995, Clewlow and Strickland 2000). These processes capture the property that spot fuel prices, despite high volatility, tend to return to long-term stationary levels. We use a binomial lattice (tree) approximation (BLA) to simulate the mean-reverting process for the diesel price per gallon D_t. BLA is widely used in financial models (see Nelson and Ramaswamy 1990) and provides good accuracy while greatly reducing the number of fuel cost states. Specifically, we set $\ln D_{t+\Delta t} = \ln D_t \pm \sigma\sqrt{\Delta t}$, where σ is the process volatility parameter, $\Delta t = \frac{1}{12}$ (one month), and the probability of upward movement at each node is given as follows. Denote by η and $\bar{D}$ the parameters *mean-reverting coefficient* and *long-term diesel price*, respectively. Define $v_t \doteq \eta(\ln \bar{D} - \ln D_t) - \frac{1}{2}\sigma^2$ and $q_t \doteq \frac{1}{2} + (v_t/2\sigma)\sqrt{\Delta t}$. Then the probability of upward movement is q_t if $0 \le q_t \le 1$, 0 if $q_t \le 0$, and 1 if $q_t \ge 1$ (Hahn and Dyer 2008). Through a least-squares procedure, the historical diesel prices in the United States betweeen 2009 and 2012 are used for calibrating the mean-reverting process parameters: $\eta = 0.18$, $\sigma = 0.11$, $D_1 = 4.1$ per gallon, and $\bar{D} = 8.5$ per gallon. From Table 1, $\mathbf{P} = (316D_t, 215D_t)$.

5.1.3. Penalty Cost. We estimate the penalty cost by the monthly cost of renting capacity from a third party, $\underline{\Phi} = \$4{,}300$, which includes rental fees and insurance,[5] plus the average fuel cost across conventional diesel and HEV trucks: $\Phi = 4{,}300 + 0.5(P_c + P_e)$. We have experimented with alternative specifications $\Phi_1 = 4{,}300 + P_c$ and $\Phi_2 = 4{,}300 + P_e$, and our results are similar: for example, for our base case the optimal capacity of HEV and/or conventional trucks using either Φ_1 or Φ_2 is within 2% of the results that use Φ instead.

5.2. Computational Approach

We use a variant of value iteration to compute the infinite-horizon value functions J for the DP as follows. We use a corresponding finite-horizon version of the DP with a planning horizon of 120 periods (10 years). Denote by $J_t^j(\cdot)$ the value function for period t and iteration j at a given state. For iteration 1, we set the terminal values $J_{121}^1(\mathbf{k}, \mathbf{p}, x) = -s_c k_c - s_e k_e$ for all $(\mathbf{p}, x)$; the values of $J_1^1(\cdot)$, for each state, are then computed by backward induction. In subsequent iterations j, we set $J_{121}^j(\cdot) = J_1^{j-1}(\cdot)$ and compute $J_1^j(\cdot)$, again by backward induction. This procedure stops when there is convergence for J_1; we observe convergence (with an error of less than 0.01%) after 16 iterations.

In addition, considering the high dimensionality of the state space, we do not compute the value function J_t^j for every possible state $(\mathbf{k}, \mathbf{p}, x)$. Instead, we take advantage of the ISD policy structure, and only search and bookkeep the ISD policy limits for every possible $(\mathbf{p}, x)$; that is, we find, as explained below, the control limits $k_c^L(k_e)$, $k_e^L(k_c)$, $k_c^R(k_e)$, and $k_e^R(k_c)$ for each $(\mathbf{p}, x)$, from which the optimal policy $\mathbf{y}$ (and the corresponding new values of J_t^j) can be easily computed for a given $\mathbf{k}$. When computing the control limits for a particular state $(\mathbf{p}, x)$, less than 10% of the value functions of all possible states need to be visited. We therefore only compute and store those values, and by doing so, we reduce the computing time by more than 90%. To find the control limits for each possible $(\mathbf{p}, x)$, we first locate the four corner points (A, B, C, and D) from Figure 1 as follows. For instance, to find A, we use the fact that $\tilde{J}(\mathbf{y}; \mathbf{k}, \mathbf{p}, x) := v_c(y_c - k_c) + v_e(y_e - k_e) + G(\mathbf{y}, \mathbf{p}, x)$, which represents the total expected discounted cost under capacity expansion only, is convex in either y_c or y_e for a given $(\mathbf{k}, \mathbf{p}, x)$. Given an initial capacity guess A_0, we change (increase or decrease) the value of y_c and y_e, alternately, to reach the minimum of $\tilde{J}$ until there is convergence to point A.

Finally, for each 120-period iteration cycle, we truncate the BLA for D_t between the values of \$0.50 and \$15, resulting in 44 possible nodes (instead of 120).

5.3. Properties of the Optimal Policy

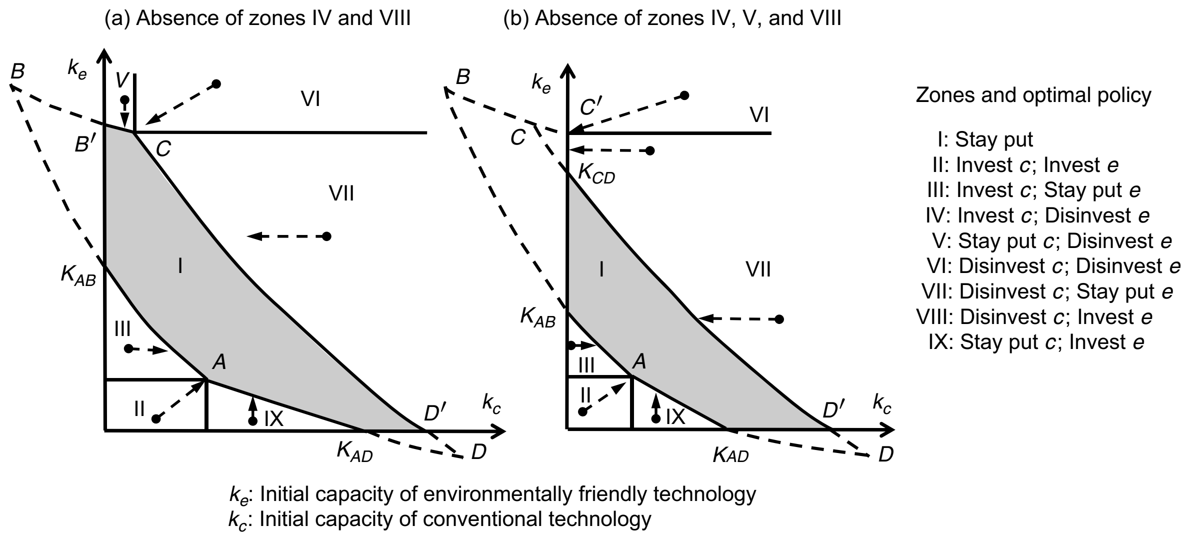
We first report on general properties of the optimal policy. All numerically derived results and insights in this paper are stated as observations. Observation 1 is related to the degeneration of the optimal policy—when certain policy zones are absent, as shown in Figure 2. Degeneration results from the nonnegativity constraints for the initial capacity ($\mathbf{k}$) and the capacity after optimal adjustment ($\mathbf{y}$), and has managerial implications.

Observation 1. (a) If there is a need to reduce the fleet size, then the firm often salvages conventional rather than HEV capacity (absence of zones IV and V), unless HEV capacity is very high (zone VI). This is because the fuel cost savings from retaining HEV capacity exceed the difference in salvage values between the two technologies.

(b) When the firm has low HEV capacity but high conventional capacity, and fuel costs are sufficiently low, the firm does not invest in HEV capacity (absence of zone VIII). With low fuel costs, fuel savings from HEV technology do not justify its higher acquisition costs if the firm has sufficient conventional capacity.

[5] We did not have access to CCE's rates for renting a delivery truck from a third party. Internet truck rental rates range between \$650 and \$900 per week, depending on the truck, with additional insurance of \$33–\$40 per day (http://reservations.ryder.com/; accessed October 1, 2011). Thus, a \$140 daily rate is a reasonable estimate.

Figure 2 Two Typical Degenerated ISD Policies



Observation 2 is related to the rate of substitution between the two technologies.

Observation 2. (a) Consider two firms in zone III (VII), where it is optimal to invest (disinvest) in conventional capacity only. The firms are identical, except that firm 2 has one fewer unit of HEV capacity than firm 1. Then firm 2 optimally invests in one more (disinvests one fewer) unit of conventional capacity than firm 1 in zone III (VII). This means, a 1:1 rate of substitution.

(b) Consider two firms in zone IX (V), where it is optimal to invest (disinvest) in HEV capacity only. The firms are identical, except that firm 2 has one fewer unit of conventional capacity than firm 1. Then firm 2 optimally invests in (disinvests) less than one unit of HEV capacity than firm 1 in zone IX (V).

Observation 2(a) results from the slopes of the boundaries AB and CD in Figure 1 being approximately -45 degrees, where Observation 2(b) results from the slopes of AD and BC being flatter than -45 degrees; this can be proved formally for the special case of no deterioration ($\beta_i = 0$). The intuition behind Observation 2(a) comes from Lemma 1, as HEV technology is optimally always dispatched first to meet demand, resulting in high utilization. So, if the firm has enough HEV capacity but wants to expand its conventional capacity, it needs one extra unit of conventional capacity if one unit of HEV capacity is lost because HEV capacity is almost always utilized. Similarly for Observation 2(b), if the firm has enough conventional capacity but wants to expand its HEV capacity, it needs less than one extra unit of HEV capacity if one unit of conventional capacity is lost because conventional capacity is not frequently utilized.

5.4. Sample Path Analysis of Dynamic Capacity Adjustment

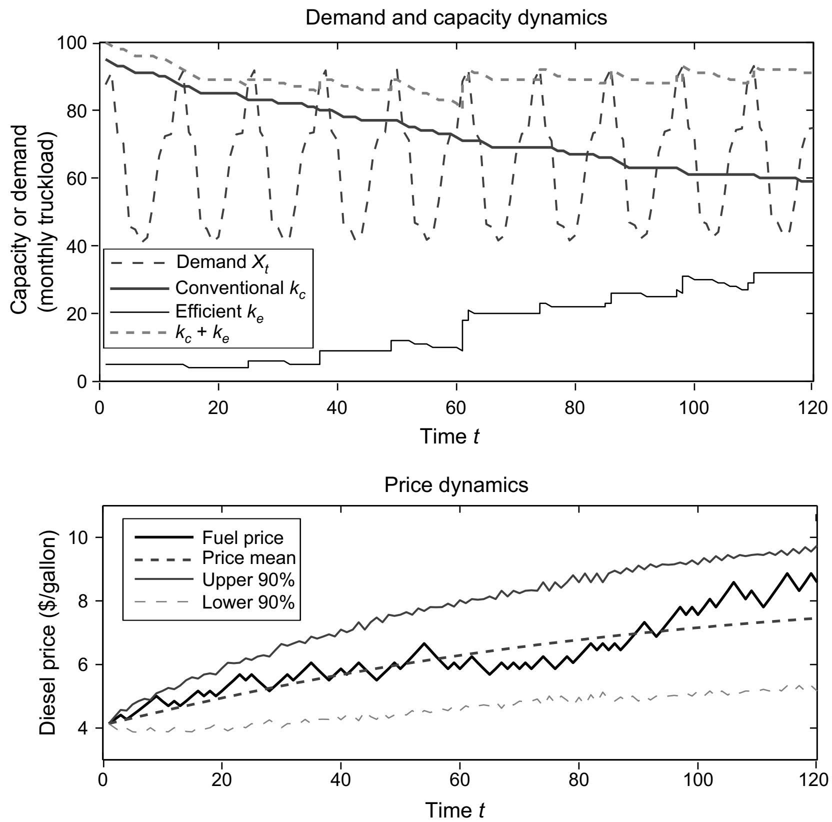
We illustrate in Figure 3 how the firm dynamically adjusts its capacity in response to random fuel cost and demand realizations. We use a representative sample path for D_t, where the firm starts with capacity portfolio $(k_c, k_e) = (95, 5)$ (the 5% fraction of HEV capacity is similar to CCE's current portfolio), and D_t trends upward. In general, more HEV capacity k_e is purchased as diesel prices increase. Many significant increases in HEV capacity occur just before seasonal demand reaches its peak, for instance, in periods 36 and 61. In contrast, the firm never acquires conventional diesel capacity, so that k_c decreases over time due to natural and usage-based deteriorations.

5.5. Optimal Capacity Portfolio and Cost Performance

We now report on the optimal portfolio and total discounted cost in an infinite horizon (i.e., the value function J) starting with zero capacity, fuel cost vector $\mathbf{p}_1 = (1{,}295.6,\ 881.5)$ corresponding to current diesel price $D_1 = \$4.1$, and realized demand for the first period (July) $x_1 = 88$. For brevity, henceforth we refer to $\mathbf{y}((0,0), \mathbf{p}_1, x_1)$ as "optimal portfolio," and $J((0,0), \mathbf{p}_1, x_1)$ as "total cost." We first compare this optimal capacity portfolio against two benchmarks: 100% conventional diesel capacity and 100% HEV capacity. The 100% HEV benchmark, for example, is obtained by constraining the decision in the DP to $y_c = 0$ for all states. The results are displayed in Table 2 and discussed below. We also compare our approach against a heuristic (NPV).

5.5.1. Optimal Capacity Portfolio. CCE's optimal portfolio includes 87 trucks, of which 40 (46%) are conventional diesel and 47 (54%) are HEV. CCE's total cost is \$31.15 million; this is 1.2% lower than the 100% HEV benchmark and 6.2% lower than the

Figure 3 Sample Path Analysis



100% conventional diesel benchmark. Also note that the total capacity for the 100% HEV benchmark is 78 trucks, compared with 87 for the optimal portfolio; this means that the 100% HEV benchmark relies more on third-party rentals to meet peak demand (i.e., the penalty cost) because of the higher acquisition cost for HEV trucks. Expected yearly fuel consumption in an infinite horizon and carbon footprint[6] (measured in terms of tons of CO_2-equivalent (tCO_2e)) are 3.0% lower and 41.7% higher than optimal for the 100% HEV and 100% conventional benchmarks, respectively. Note that an HEV-only fleet already results in significant total cost savings over a conventional diesel-only fleet.

Observation 3. Firms such as CCE should adopt environmentally friendly fuel-efficient vehicles even without additional government incentives. Conventional diesel vehicles, which have lower acquisition cost, can still be used as "safety" capacity, being deployed only when demand is high.

5.5.2. Comparison against an NPV Heuristic. Firms routinely use NPV analysis to guide investment decisions. Applied to our case, such a heuristic would involve two steps. In Step I, the heuristic decides which technology is better at a given state: The NPV of expected fuel savings over the life of an HEV truck (where in each month t, the diesel price used for calculation is the expected value of D_t from the BLA process) is compared against its additional initial investment vis-à-vis the conventional diesel truck, assuming full utilizations. After deciding on the technology, Step II estimates the total capacity via a newsvendor-type analysis, where the demand distribution is discrete and assigns a probability of 1/12 to each of the 12 monthly demand forecasts, the penalty cost for not meeting demand is Φ, and the overage cost is equal to the unit capacity acquisition cost divided by the truck's average life of 144 months. At any state, the firm acquires new capacity only if the total capacity amount found by the newsvendor analysis is larger than the firm's current capacity. We find (Table 2) that the total cost of the heuristic is $32.02 million, which is 2.8% higher than the optimal solution at $31.15 million, and that the heuristic acquires 100% HEVs. The optimality gap originates from each of the two steps (technology selection and capacity amount), where each roughly contributes 50%.

5.6. Sensitivity Analysis

We now provide the results of a sensitivity analysis for various model parameters, where in each study

[6] Carbon footprint is (tCO_2e) $= 0.010242 *$ fuel consumption in gallons of diesel (U.S. Environmental Protection Agency 2008).

Table 2 Optimal Portfolio, Single-Technology Benchmarks, and NPV Heuristic

		Single-technology benchmarks		
	Optimal portfolio	HEV	Conventional diesel	NPV heuristic
Conventional diesel (% of fleet)	40 (46%)	0 (0%)	88 (100%)	0 (0%)
HEV (% of fleet)	47 (54%)	78 (100%)	0 (0%)	87 (100%)
Total cost ($M)	31.15	31.53	33.07	32.03
Total cost (% of the optimal solution)	100	101.2	106.2	102.8
Expected fuel consumption (thousand gallons/year)	174.14	168.97	246.71	168.06
Carbon footprint (tCO_2e/year)	1,784	1,731	2,527	1,721
Carbon footprint (% of the optimal solution)	100	97.0	141.7	96.5

we vary one parameter at a time from the base values described in §5.1.

5.6.1. Capacity Deterioration. We compare the optimal solution under different deterioration rates of 5%, 7%, and 9%.[7] We find that as the deterioration rate increases from 5% to 7% and then to 9%, the firm's investment in HEV trucks decreases from 62 (71% of fleet) to 47 (54% of fleet), and then to 37 (42% of fleet), respectively, whereas the total capacity of the optimal portfolio is the same at 87 trucks across all deterioration rates.

Observation 4. Ceteris paribus, a higher deterioration rate for both technologies (which implies more frequent capacity replacements) makes HEV capacity less appealing because of its higher acquisition cost.

5.6.2. Demand Seasonality. Since demand seasonality drives several of our insights, we performed a sensitivity analysis on our demand pattern by scaling CCE's seasonal demand factors between 0% (no seasonality) and 100% (CCE's seasonality), while maintaining the same yearly mean demand of 732.5. Thus, 0% seasonality corresponds to a constant expected demand of $732.5/12 = 61$ per period. The results are displayed in Figure 4.

Observation 5. Ceteris paribus, higher seasonality implies a lower fraction of HEV capacity in the optimal portfolio, a higher total cost, and a higher carbon footprint. With zero seasonality, and the current outlook on diesel prices, the firm acquires 100% HEV.

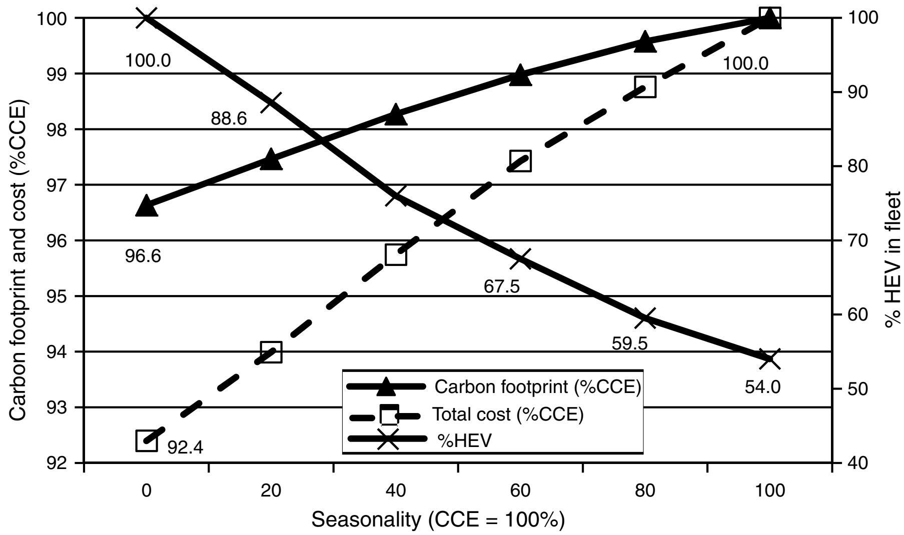
It is optimal to acquire 100% HEV when there is no seasonality because of our assumptions and the fact that capacity is almost always utilized. Carbon footprint is lower as a result.

5.6.3. Acquisition Cost (Tax Credit). Tax credits are a frequently used policy instrument for promoting green technologies. In its simplest form, a tax credit directly reduces the initial capacity investment and can be operationalized as a percentage reduction in the incremental unit acquisition cost of the HEV technology relative to the conventional diesel one. For example, a 10% tax credit means that the unit acquisition cost for an HEV truck (v_e) is reduced by $(\$94{,}000-\$60{,}000)(0.10)$ to $90,600. Figure 5(a) shows that the percentage of HEV capacity in CCE's fleet increases from 54% to 84% as the tax credit increases from 0 to 40%, with a corresponding decrease in carbon footprint, and a linear decrease in total cost as shown in Figure 5(b). In all cases, the optimal portfolio has the same total capacity of 87 trucks.

Tax credits for HEVs have been used mostly for stimulating technological development, as they increase HEV adoption rate, allowing auto manufacturers to spread out research and development (R&D) costs. We can also assess the efficiency of a tax credit toward short-term carbon footprint reduction through the ratio between dollars invested by the government in tax credits and the corresponding carbon footprint reduction. The measurement of government investment (in dollars) associated with a tax credit depends on the price of diesel (which impacts the firm's optimal portfolio), the period of time when the tax credit is in effect, and so forth. Define $Z(k)$ as the expected number of HEVs acquired per year in an infinite horizon for tax credit k; this is displayed in Figure 5(c). To measure the true impact of tax credit k on a firm's decision, we need to consider the difference $Z(k) - Z(0)$, as most firms start from a given fleet when the tax credit is implemented (including CCE). Thus, the government investment per tCO_2e can be estimated as $((Z(k)-Z(0))3{,}400k)/(CF(k)-CF(0))$, where $CF(k)$ is the annual carbon footprint for tax credit k. This is shown in Figure 5(c), and for CCE it varies from $\$35/tCO_2e$ for a 10% credit to $\$170/tCO_2e$ for a 40% tax credit. These numbers should be viewed as a lower bound because, in reality, the tax credit could be wasted on firms who would have purchased additional HEV vehicles without the credit due to aging fleets or general expansion. Estimates for the social cost of CO_2 varies widely—a meta-analysis of 28 studies by Tol (2005) found the median and mean to be $14 and $93 per tCO_2e, respectively. Thus, tax credits imply a significantly higher government investment

[7] Note that the 9% rate is higher than the 7.5% threshold discussed after Assumption 2. Our assumptions, however, provide only a sufficient condition for convexity, and our numerical analysis indicate that the 9% rate does not break the convexity of the cost function in this case.

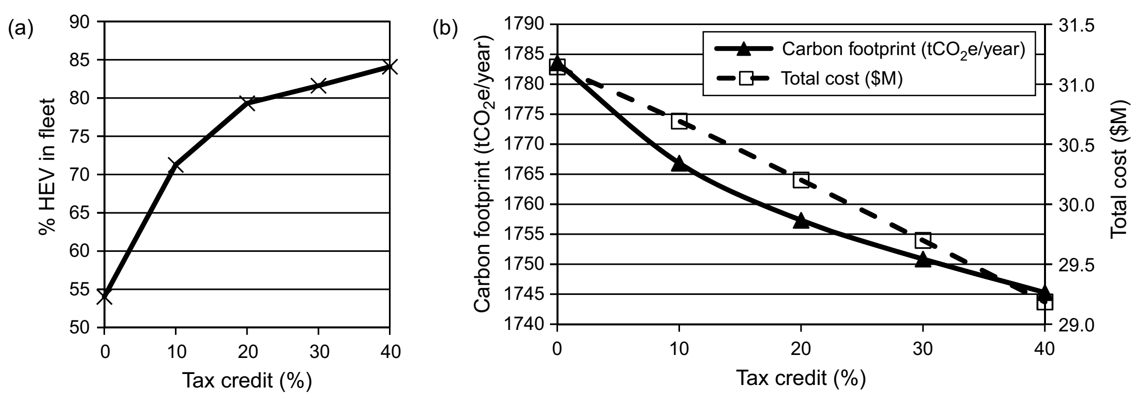
Figure 4 Impact of Demand Seasonality on the Optimal Policy



than the median estimated social cost of carbon. The numbers from Figure 5(c) apply to CCE, which has a highly seasonal demand. Recall that with zero seasonality, the firm already acquires 100% HEV without a tax credit, and thus a tax credit would have no impact on such a firm's fleet and carbon footprint.

Observation 6. Tax credits do not result in a socially optimal short-term reduction in carbon footprint for firms with highly seasonal demand such as CCE, given a median estimate of the social cost of carbon of $14/tCO_2e$. This is because, although tax credits increase the fraction of HEV in their fleet (replacing conventional technology), those additional HEV are mostly used to meet excess demand, which does not result in a major carbon footprint reduction. Tax credits are even less efficient at reducing carbon footprint for firms with lower demand seasonality, because such firms already prefer (without a

Figure 5 Impact of an HEV Tax Credit on the Optimal Policy and Carbon Footprint



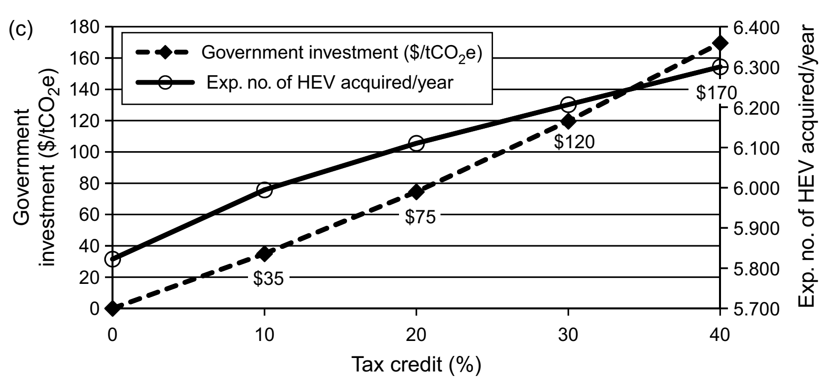
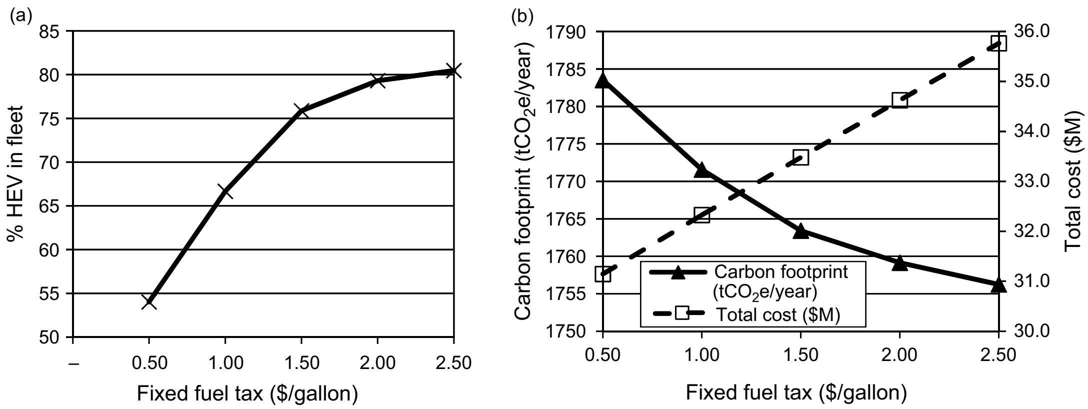


Figure 6 **Impact of a Fixed Tax on the Optimal Policy**



tax credit) to have a large portion of HEV capacity in their portfolio.

It should also be noted that a sustained government investment in tax credits could reduce the HEV acquisition cost in the long run, as higher HEV sales could help finance R&D that improves HEV technology. This "learning curve" effect is not captured in our model; hence the use of the qualifier "short-term" in Observation 6.

5.6.4. Mean of Diesel Price (Fixed Fuel Taxes). A fixed tax imposed on the price of diesel increases its mean price. A fixed tax is used in the United States at an average value of $0.54 per gallon as of July 2012.[8] As shown in Figure 6(a), the percentage of HEV capacity in CCE's fleet increases from 54% to 80% as the fixed tax increases from $0.50 to $2.50 per gallon, ceteris paribus. Figure 6(b) shows, however, that there is a significant increase in total cost relative to the corresponding decrease in carbon footprint. For example, as the government doubles the fixed tax from $0.50 to $1.00, average yearly fuel consumption decreases by only 1,166 gallons, from 174,137 to 172,971 gallons (yearly carbon footprint decreases from 1784 to 1772 tCO_2e), resulting in a fuel cost increase of $81,704 at a $4.10 diesel price, or $76,572 at a $8.50 diesel price. Thus, this is equivalent to a (roughly) $6,500 increase in operating cost per tCO_2 not emitted. This high number has two causes: (i) the fuel tax is applied uniformly to the entire fuel consumption of the firm, and (ii) because of CCE's seasonal demand pattern, additional HEVs substitute for diesel trucks, which are not fully utilized. These insights are preserved even if the fuel tax is not applied to the penalty cost, since the fuel tax is only a relatively small portion of the penalty cost. Notice that we are only stating the impact of fuel taxes as a policy mechanism to reduce carbon footprint—fuel taxes have a different policy objective in the United States, which is to fund transportation infrastructure.

Observation 7. Fixed fuel taxes are not an efficient policy mechanism to reduce carbon footprint for firms with highly seasonal demand such as CCE. This is because they result in a significant increase in operating cost, but not a significant reduction in fuel consumption. Fixed fuel taxes are even less efficient as demand seasonality decreases.

We have also performed a sensitivity analysis on the volatility of fuel cost (by modifying the σ parameter in the BLA) and on the correlation between fuel cost and demand (by modifying the parameter γ in the demand X_t), and found the results to be relatively insensitive to them.

6. Extension: Multiple Technologies

Our model considers two technologies, which differ in acquisition and operating costs. In many situations, firms may choose among more than two technologies. For instance, for delivery trucks, other technologies exist outside diesel and HEVs, such as plug-in HEVs or natural gas.

It is straightforward to generalize our model to more than two technologies. Suppose there are $H \geq 2$ technologies in the market, and for notational simplicity, we refer to technology 0 as the penalty cost, that is, $v_0 = 0$, $P_0 = \Phi$, and $m_0 = 0$. Similar to the assumptions made before with two technologies, we assume the following:

Assumption 4. (a) *The technologies are ordered according to increasing acquisition cost but decreasing total operating cost*: $v_h < v_{h+1}$ *and* $P_{h+1} + m_{h+1} <_{\text{a.s.}} P_h + m_h$, *for* $h = 0, 1, \ldots, H-1$.

(b) $\gamma_i + \beta_i < 1$, *for* $h = 0, 1, \ldots, H-1$.

(c) $P_{h+1} + m_{h+1} + \beta_{h+1} v_{h+1} \leq_{\text{a.s.}} P_h + m_h + \beta_h s_h$, *for* $h = 0, 1, \ldots, H-1$.

[8] http://www.api.org/statistics/fueltaxes/ (accessed October 1, 2011).

Assumptions 4(a)–4(c) require the technologies to be ordered throughout the infinite planning horizon. If major developments occur, such as a significant decrease in the fuel price for one technology (e.g., natural gas), then the problem needs to be resolved, because the order may change (e.g., natural gas may become preferred to HEV). Under Assumptions 4(a)–4(c), Lemma 1 and Proposition 1 can be generalized:

Proposition 3. *Suppose Assumption* 4(a) *holds for a dynamic capacity adjustment problem with* $H \geq 2$ *technologies. Then*

(a) *it is optimal to first use a more fuel-efficient technology* (*the one with lower operating cost*); *and*

(b) $G(\mathbf{y}, \mathbf{p}, x)$ *is convex in* $\mathbf{y}$, $J(\mathbf{k}, \mathbf{p}, x)$ *is convex in* $\mathbf{k}$, *and correspondingly, the optimal capacity adjustment follows an ISD policy.*

Proposition 3 indicates that convexity is maintained for the cost-to-go function and so is the ISD policy. However, the supermodularity property of Proposition 2 does not always hold if $H > 2$. This is not surprising because, in general, little structure exists for minimization of a supermodular function with three or more variables.[9]

7. Concluding Remarks

Firms routinely make capacity investment decisions when more energy efficient technologies are available. These technologies typically have a higher up-front acquisition cost but a lower operating cost than conventional technologies. This capacity adjustment decision is complicated by uncertain fuel costs, and the fact that capacity is not always 100% utilized due to seasonal demand and volatile fuel costs. Higher utilizations also imply a higher deterioration rate for capacity. Current methods for evaluating capacity investment decisions fall short in several of these dimensions.

Our paper attempts to fill this gap and provides a tractable decision support model for dynamic capacity adjustment decisions that incorporates all of these features, in the context of two technologies. We formulate the problem as an infinite-horizon stochastic dynamic program and find that the optimal policy is of the ISD type for each realization of demand and fuel costs in a period.

Our model assumes a constant maintenance cost. We have also studied a model with usage-based maintenance cost, which requires the categorization of each capacity type into a finite number of conditions, thus expanding the state space. This modification, however, breaks our structural result that it is always optimal to dispatch HEV technology first to meet demand: capacity dispatched first wears out quicker, incurring higher maintenance cost and increasing the need for future investment. As a result, there is no structure to the optimal policy.

We find that an HEV-only fleet performs well relative to the optimal policy. Firms, however, have been slow in adopting HEV technologies for their fleet. One reason might be that fuel-efficient technologies are developing rapidly, and thus firms may be delaying investment to wait for newer and better technologies. To address this question, we have studied a finite-horizon extension of our model, where there is the possibility of a newer, more efficient HEV technology in the future. This modification requires a new state variable representing the stochastic condition of the new technology in each period. Assuming a constant probability ρ of a new technology in each period (if it is not yet available), we find that as ρ increases, the firm indeed decreases the optimal HEV capacity (i.e., y_e decreases). The optimal initial capacity for the conventional technology, however, first increases and then decreases as ρ increases—for small values of ρ, the firm increases conventional capacity to make up for reduced HEV capacity; for ρ high enough, the firm postpones investments in both capacity types, relying more on third party capacity (the penalty cost) to meet demand.

In addition to providing the structure of the optimal policy for a challenging dynamic capacity adjustment problem with two technologies, we provide managerial insights through a numerical analysis performed using CCE data for its delivery truck fleet. A key insight is that a balanced portfolio is optimal with seasonal demand: CCE should dedicate 54% of its fleet to be HEV, which results in a 6.2% total cost savings over a diesel-only fleet. Another key insight is that a 100% HEV only fleet performs well, as long as capacity adjustment is dynamically optimized, with a total cost only 1.24% worse than optimal, and it is optimal if there is no demand seasonality.

Supplemental Material

Supplemental material to this paper is available at http://dx.doi.org/10.1287/msom.2013.0438.

Acknowledgments

The authors thank Erica Plambeck, an anonymous associate editor, and four referees for their helpful feedback, which significantly improved this paper from earlier versions.

[9] Consider a supermodular function of two variables (x, y). By defining $x' = -x$, one can change the function into a submodular one of (x', y) so that the structural results for minimizing a submodular function can be used. However, this "trick" cannot be used when there are more than two variables.

References

Aflaki S, Netessine S (2011) Strategic investment in renewable energy source. Working paper, INSEAD, Fontainebleau, France.

Ambec S, Crampes C (2010) Electricity production with intermittent sources of energy. Working paper, University of Toulouse, Toulouse, France.

Bessembinder H, Coughenour JF, Seguin PJ, Smoller MM (1995) Mean reversion in equilibrium asset prices: Evidence from the futures term structure. *J. Finance* 50(1):361–75.

Chao H (1983) Peak load pricing and capacity planning with demand and supply uncertainty. *Bell J. Econom.* 14(1):179–190.

Chao X, Chen H, Zheng S (2009) Dynamic capacity expansion for a service firm with capacity deterioration and supply uncertainty. *Oper. Res.* 57(1):82–93.

Clewlow L, Strickland C (2000) *Energy Derivatives: Pricing and Risk Management* (Lacima Group, Houston).

Crew MA, Kleindorfer PR (1976) Peak load pricing with a diverse technology. *Bell J. Econom.* 7(1):207–231.

Ding Q, Dong L, Kouvelis P (2007) On the integration of production and financial hedging decisions in global markets. *Oper. Res.* 55(3):470–489.

Dixit A (1997) Investment and employment dynamics in the short run and the long run. *Oxford Econom. Papers* 49(1):1–20.

Dixit AK, Pindyck RS, Davis GA (1994) *Investment Under Uncertainty* (Princeton University Press, Princeton, NJ).

Drake D, Kleindorfer PR, Van Wassenhove LN (2010) Technology choice and capacity investment under emissions regulation. Working paper, INSEAD, Fontainebleau, France.

Eberly JC, Van Mieghem JA (1997) Multi-factor dynamic investment under uncertainty. *J. Econom. Theory* 75(2):345–387.

Garcia A, Alzate JM, Barrera J (2012) Regulatory design and incentives for renewable energy. *J. Regulatory Econom.* 41(3):315–336.

Guo X, Kaminsky P, Tomecek P, Yuen M (2011) Optimal spot market inventory strategies in the presence of cost and price risk. *Math. Methods Oper. Res.* 73(1):109–137.

Hahn WJ, Dyer JS (2008) Discrete time modeling of mean-reverting stochastic processes for real option valuation. *Eur. J. Oper. Res.* 184(2):534–548.

Huchzermeier A, Cohen MA (1996) Valuing operational flexibility under exchange rate risk. *Oper. Res.* 44(1):100–113.

Jaffe AB, Newell RG, Stavins RN (2003) Technological change and the environment. Mäler K-G, Vincent JR, eds. *Handbook of Environmental Economics*, Vol. 1 (Elsevier Science B.V., Amsterdam), 461–516.

Kazaz B, Dada M, Moskowitz H (2005) Global production planning under exchange-rate uncertainty. *Management Sci.* 51(7): 1101–1119.

Kleindorfer P, Neboian A, Roset A, Spinler S (2012) Fleet renewal with electric vehicles at La Poste. *Interfaces* 42(5):465–477.

Kleindorfer PR, Fernando CS (1993) Peak-load pricing and reliability under uncertainty. *J. Regulatory Econom.* 5(1):5–23.

Luss H (1982) Operations research and capacity expansion problems: A survey. *Oper. Res.* 30(5):907–947.

Manne AS (1961) Capacity expansion and probabilistic growth. *Econometrica: J. Econometric Soc.* 29(4):632–649.

Nelson DB, Ramaswamy K (1990) Simple binomial processes as diffusion approximations in financial models. *Rev. Financial Stud.* 3(3):393–430.

Pindyck RS (1986) *Irreversible Investment, Capacity Choice, and the Value of the Firm* (National Bureau of Economic Research Cambridge, MA).

Plambeck EL, Taylor TA (2011) On the flexibility to adapt to co-fluctuations in feedstock costs and output prices. Working paper, Stanford University, Stanford, CA.

Rajagopalan S (1998) Capacity expansion and equipment replacement: A unified approach. *Oper. Res.* 46(6):846–857.

Rajagopalan S, Singh MR, Morton TE (1998) Capacity expansion and replacement in growing markets with uncertain technological breakthroughs. *Management Sci.* 44(1):12–30.

Stokey NL, Lucas RE, Prescott EC (1989) *Recursive Methods in Economic Dynamics* (Harvard University Press, Cambridge, MA).

Tol RSJ (2005) The marginal damage costs of carbon dioxide emissions: An assessment of the uncertainties. *Energy Policy* 33(16):2064–2074.

U.S. Environmental Protection Agency (2008) Direct emissions from mobile combustion sources. Climate Leaders: Greenhouse Gas Inventory Protocol Core Module Guidance EPA430-K-08-004, Office of Air and Radiation, Washington, DC.

Van Mieghem JA (2003) Capacity management, investment, and hedging: Review and recent developments. *Manufacturing Service Oper. Management* 5(4):269–302.