Production, Manufacturing and Logistics

Proactive data envelopment analysis: Effective production and capacity expansion in stochastic environments

Chia-Yen Lee [a], Andrew L. Johnson [b,c,*]

[a] Institute of Manufacturing Information and Systems, National Cheng Kung University, Tainan City 701, Taiwan
[b] Department of Industrial and Systems Engineering, Texas A& M University, College Station, TX 77840, USA
[c] School of Business, Aalto University, Helsinki 00101, Finland

ABSTRACT

Demand fluctuations that cause variations in output levels will affect a firm's technical inefficiency. To assess this demand effect, a demand-truncated production function is developed and an "effectiveness" measure is proposed. Often a firm can adjust some input resources influencing the output level in an attempt to match demand. We propose a short-run capacity planning method, termed proactive data envelopment analysis, which quantifies the effectiveness of a firm's production system under demand uncertainty. Using a stochastic programming DEA approach, we improve upon short-run capacity expansion planning models by accounting for the decreasing marginal benefit of inputs and estimating the expected value of effectiveness, given demand. The law of diminishing marginal returns is an important property of production function; however, constant marginal productivity is usually assumed for capacity expansion problems resulting in biased capacity estimates. Applying the proposed model in an empirical study of convenience stores in Japan demonstrates the actionable advice the model provides about the levels of variable inputs in uncertain demand environments. We conclude that the method is most suitable for characterizing production systems with perishable goods or service systems that cannot store inventories.

© 2013 Elsevier B.V. All rights reserved.

1. Introduction

Data envelopment analysis (DEA) is a deterministic mathematical programming approach to productive efficiency analysis. Given the same input resources, a production process is called efficient if its output levels are at least as high as other observed production processes. However, in practice, reduced actual output can be caused by insufficient demand biasing productivity measures. Noting that DEA studies typically do not model demand effects, this paper describes "proactive DEA" which measures "effectiveness" (Lee & Johnson, 2012a), which distinguishes demand effects in productive efficiency analyses and identifies capacity adjustment levels in a planning process using stochastic programming (SP).

The literature on the demand effect in productivity and efficiency analysis is limited. Lee and Johnson (2011, 2012b) decompose a production process into capacity design, demand generation, operations, and demand consumption components,

and measure the productivity change of each component. They distinguish the production process from the demand generation/consumption process. The results indicate that demand fluctuations rather than production capabilities can cause technical regress. Further, the capacity design component generally has a significant effect on long-term productivity. Therefore, Lee and Johnson (2012a) propose a "demand-truncated production function" and estimate "effectiveness" to distinguish from efficient production where customer demand does not limit the levels of output. Typical of most DEA studies, Lee and Johnson (2012a) study only supports ex-post analysis of the data after production and sales. Therefore, this paper describes a planning model using DEA-based techniques that accounts for demand variations and ex-ante resource allocation to improve a firm's operational effectiveness. This model also supports planning functions by solving the short-run capacity expansion problem.

The classic capacity expansion problem (Luss, 1982; Manne, 1961) includes the following variables; let decision variable $\mathbf{x} \in \mathbb{R}_+^J$ be the vector of the input resource needed, $\mathbf{y} \in \mathbb{R}_+^Q$ be the vector of the products generated, $\mathbf{D} \in \mathbb{R}_+^Q$ be the demand vector with respect to each output, $\mathbf{A} \in \mathbb{R}_+^{J \times Q}$ be the vector representing the required resource per unit product, $\mathbf{C} \in \mathbb{R}_+^J$ be the vector of cost per unit resource, and $\mathbf{P} \in \mathbb{R}_+^Q$ be the vector of each product's

selling price. The solution indicates the resources required, $\mathbf{x}$, and the product quantity, $\mathbf{y}$, generated. The optimization formulation is as follows:

$$
\begin{aligned}
\text{Max} \quad & \mathbf{Py} - \mathbf{Cx} \\
s.t. \quad & \mathbf{Ay} \leqslant \mathbf{x} \\
& \mathbf{y} = \mathbf{D} \\
& \mathbf{x} \geqslant \mathbf{0}, \mathbf{y} \geqslant \mathbf{0}
\end{aligned}
\tag{1}
$$

The primary issues are determining the expansion sizes, expansion times, and expansion locations, and the objective is to minimize the discounted costs with respect to the expansion process (Luss, 1982). In general, all production factors are adjustable. The capacity expansion problem is a component of long-run production analyses. However, in the short run, the plant size, location, and capital stock for production are typically fixed, whereas the variable factors such as employment and material are adjustable (Marshall, 1920). Stigler (1939) argues that the variations in output can be described in terms of the law of diminishing returns and marginal productivity theory, holding all productive factors quantity constant, and adjusting the quantity of the remaining one. Wilson and Eckstein (1964) claim that long-run and short-run economic analyses represent different productivity behaviors and conclude that the long-run cost curve forms an envelope of the short-run cost. After distinguishing the characteristics of long-run and short-run production, our paper estimates the marginal product of variable inputs along the DEA frontier and identifies the adjustments in variable inputs that influence the output levels in response to the demand fluctuations. The relationship of marginal product between inputs and outputs is described in matrix $\mathbf{A}$ of problem (1) and the elements in matrix $\mathbf{A}$ are assumed constant in the typical capacity expansion literature. This assumption may violate the law of diminishing marginal returns and result in biased capacity estimates.

Considering the stochastic nature of demand, the typical capacity expansion problem can be extended to uncertain demand and formulated as (2), using a SP technique (Birge & Louveaux, 2011). Assuming a firm desires to maximize expected profit, let $\widetilde{\mathbf{D}}$ be a random variable of demand:

$$
\begin{aligned}
\text{Max} \quad & E\left[f(\mathbf{x}, \widetilde{\mathbf{D}})\right] - \mathbf{Cx} \\
s.t. \quad & \mathbf{x} \geqslant \mathbf{0}
\end{aligned}
\tag{2}
$$

where for a realization $\mathbf{d} \in \widetilde{\mathbf{D}}$

$$
\begin{aligned}
f(\mathbf{x}, \mathbf{d}) = \text{Max} \quad & \mathbf{Py} \\
s.t. \quad & \mathbf{Ay} \leqslant \mathbf{x} \\
& \mathbf{y} = \mathbf{d} \\
& \mathbf{y} \geqslant \mathbf{0}
\end{aligned}
$$

This paper makes three contributions to the literature. First, we use DEA to estimate production performance incorporating short-run capacity expansion decisions, thus creating a proactive DEA model. Using a production function estimated from observed production processes assures the feasibility of the recommended short-run capacity adjustment. Second, our definition of effective production, or the product generated from the production system to be consumed by realized demand, complements current efficiency analysis. Third, the proposed method considers the diminishing marginal benefits of inputs and estimates the marginal product which the short-run capacity expansion problem generally ignores; ignoring the diminishing marginal return of inputs, traditional capacity expansion models may severely underestimate the resources necessary to meet demand, resulting in lost sales, costly outcomes, etc. In other words, in some cases, it may be too costly to fill demand, thus reducing resources to decrease output levels may be preferred.

The remainder of this paper is organized as follows. Section 2 reviews the DEA literature on uncertainty. Section 3 introduces a demand-truncated production function and an effectiveness measure. Section 4 describes estimating marginal product to support capacity adjustment, considering the law of diminishing marginal returns. Section 5 discusses the SP models and performance indices assessing the solution quality. Section 6 gives a numerical example, and Section 7 describes an empirical study of Japanese convenience stores. Section 8 concludes.

2. Literature review of DEA under uncertainty

The literature has proposed a variety of DEA models under uncertainty primarily focusing on ex-post analysis. The Convex Nonparametric Least Squares (CNLS) and Stochastic semi-Nonparametric Envelopment of Data (StoNED) literature focuses on unifying the Stochastic Frontier Analysis (SFA) and DEA approaches to productivity analysis (Kuosmanen, 2008; Kuosmanen & Johnson, 2010; Kuosmanen & Kortelainen, 2012). These estimators impose axioms such as convexity and monotonicity while including a symmetric random component to capture the typical sources of noise: omitted variables, errors in measurement, differences between modeling variables and data. StoNED allows for inefficiency. Other models available in the DEA literature include chance-constrained DEA (Bruni, Conforti, Beraldi, & Tundis, 2009; Cooper, Huang, & Li, 1996; Land, Lovell, & Thore, 1993; Olesen & Petersen, 1995) which evaluates efficiency when uncertainty exists with respect to the constraints. It builds a production frontier that envelops all of the observations "most of the time", and a few firms may perform above the frontier with an a priori set probability. Sengupta (1987, 2000) transforms DEA models into equivalent models incorporating stochastic variations and uses chance-constrained programming solves the equivalent models, which explore how the stochastic data affects the efficiency measurement. Imprecise DEA, or fuzzy DEA (Cooper, Park, & Yu, 1999; Kao & Liu, 2000), which states that certain variables or data are known to exist within a given range or an ordinal relationship estimates efficiency intervals to characterize the imprecision. However, the efficiency distribution or its characteristics, such as the mean, are still unknown. Simulated-based DEA (Kao & Liu, 2009; Premachandra, Powell, & Watson, 2000) addresses imprecise data with a known probability distribution. The results are used to support statistical inference regarding the efficiency distribution. Ruggiero (2004), who proposes using data averaged across a time horizon to estimate the average efficiency for the time horizon, argues that the estimates are less biased than estimating efficiency in each time period and averaging the cross-sectional efficiencies over the entire horizon. Kuosmanen (2004), Branda (2013), Lamb and Tee (2012), and Beraldi and Bruni (2012) discuss risk ratios, conditional value-at-risk (CVaR), stochastic dominance (SD), and SP with stochastic data in financial applications. For addition methods not discussed, see for example Dyson and Shale (2010).

Typically, DEA methods under uncertainty focus on generating robust efficiency estimates considering variations in data. However, our proposed SPDEA also provides resource re-allocation recommendations for maximizing efficiency or effectiveness in uncertain environments. Table 1 summarizes and compares several DEA approaches used to address uncertainty.

3. Effective production

Efficiency is commonly measured in the productivity literature to quantify the observed productivity level relative to the best possible productivity level. However, if products or services are produced but not consumed, resources are wasted in storage.

Table 1
DEA under uncertainty.

Approach	Stochastic in: efficiency (E) or data (D)	Probability distribution of data: known (K) or unknown (U)	Purpose: robust efficiency (E) or resource reallocation (R)	Reference
Chance-constrained programming[a]	E	K	E	Land et al. (1993) Cooper et al. (1996) Olesen and Petersen (1995) Bruni et al. (2009)
Imprecise, fuzzy, and interval data	E + D	U	E	Triantis and Girod (1998) Cooper et al. (1999) Kao and Liu (2000)
Averaged data	D	U	E	Ruggiero (2004)
Simulation	D	K	E	Premachandra et al. (2000) Kao and Liu (2000)
Equivalent transformation	E + D	K	E	Sengupta (1987, 2000)
Risk measure, CVaR, SD, SP	D	K	E	Kuosmanen (2004) Branda (2013) Lamb and Tee (2012) Beraldi and Bruni (2012)
SP supporting resource allocation	D	K	E + R	This paper

[a] Chance-constrained problems may be considered as a subclass of SP problems; the former emphasizes the stochastic property on constraints whereas the latter focuses on stochastic coefficients in mathematical programming.

Section 3.1 defines the concept of effective production and introduces the truncated production function combining a classic production function and demand information. Section 3.2 describes an effectiveness estimator and its implementation using DEA. Section 3.3 illustrates how the effectiveness estimator complements the efficiency measure and defines the strategic position with respect to efficiency and effectiveness.

3.1. Demand-truncated production function

The production function f defines the maximum output that can be produced, given the quantities of input resources. Let $\boldsymbol{x}^F \in \mathbb{R}^I_+$ be a vector of the fixed input resources, $\boldsymbol{x}^V \in \mathbb{R}^J_+$ be a vector of the variable input resources, and $y^{PF} \in \mathbb{R}_+$ be the output level of a single-output production function. The standard production function with a single output shown in Eq. (3) satisfies the properties of nonegativity, weak essentiality, monotonicity, and concavity (Coelli, Rao, O'Donnell, & Battese, 2005).

$$y^{PF} = f(\boldsymbol{x}^F, \boldsymbol{x}^V) \tag{3}$$

We define the demand-truncated production function (DTPF) as the maximum demand for a product or service that can be fulfilled, given the input resource quantities. The DTPF incorporates demand and production function information. To maintain generality, demand is firm-specific. Let $y_k^{DTPF} \in \mathbb{R}_+$ be the output level of the production function truncated by the demand of firm k and $D_k \in \mathbb{R}_+$ be the potential realized demand of firm k. The DTPF with output level y_k^{DTPF} is the smaller of the two variables: the production function output y^{PF} and realized demand D_k as Eq. (4).

$$y_k^{DTPF} = \min(y^{PF}, D_k) = \min(f(\boldsymbol{x}^F, \boldsymbol{x}^V), D_k) \tag{4}$$

Let $Y_k \in \mathbb{R}_+$ be the output level of firm k. The effective output $Y_k^E \in \mathbb{R}_+$ describes the product or service generated by the production system to be consumed via customer demand of firm k and we define $Y_k^E = \min(Y_k, D_k)$. We say that a firm achieves effective production if the effective output level as identified by the DTPF is generated, i.e., $y_k^{DTPF} = Y_k^E$.

In the short-run, the fixed input levels cannot be adjusted, so the production function is a function of variable input. Fig. 1 defines the boundary of our analysis and illustrates the DTPF. Data is gathered regarding demand and the effective output level is calculated prior to our analysis and is thus data. Within our analysis,
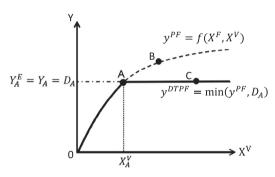

Fig. 1. Demand-truncated production function (DTPF).

production output which is a function of variable input and fixed input variable vectors is estimated using observed production data. The DTPF is estimated as the minimum of the firm specific demand and the production function. Firm A presents a supply–demand equilibrium where $D_A = Y_A = Y_A^E = f\left(X_A^F, X_A^V\right)$.[1] That is, firm A can produce the optimal output level without unfulfilled demand or excessive inventory. Firm A achieves effective production. We also show that Firm B is efficient but not effective unit and Firm C is effective but not efficient unit in Fig. 1. In addition, it is straightforward to validate the properties – nonegativity, weak essentiality, monotonicity, and concavity of truncated production function – since the minimum function of a production function and a constant demand level is a convex polyhedral.

3.2. Effectiveness measure

Let $\boldsymbol{x} \in \mathbb{R}^{I+J}_+$ denote the input vector and $\boldsymbol{y} \in \mathbb{R}^Q_+$ denote the output vector of the production system. Define the production possibility set as $T = \{(\boldsymbol{x}, \boldsymbol{y}): \boldsymbol{x}$ can produce $\boldsymbol{y}\}$ and estimate it by a piece-wise linear concave function enveloping all observations shown in (5). Let $i = \{1, 2, \ldots, I\}$ be the indexed set of fixed inputs, $j = \{1, 2, \ldots, J\}$ be the indexed set of variable inputs, $q = \{1, 2, \ldots, Q\}$ be the indexed set of output, and $k = \{1, 2, \ldots, K\}$ be the indexed set of firms. X_{ik}^F is the ith fixed input resource, X_{jk}^V is the jth variable

---

[1] Capital letters refer to data and lower-case letter refers to variable.

input resource, Y_{qk} is the amount of the qth production output, and λ_k is the multiplier of the kth firm. The model (5) defines the empirically estimated variable-return-to-scale (VRS) production possibility set $\widetilde{T}$.

$$\widetilde{T} = \left\{ (\boldsymbol{x},\boldsymbol{y}) \middle| \sum_{k=1}^{K} \lambda_k Y_{qk} \geqslant y_q, \quad \forall q; \quad \sum_{k=1}^{K} \lambda_k X_{ik}^F \leqslant x_i^F, \right.$$
$$\left. \forall i; \quad \sum_{k=1}^{K} \lambda_k X_{jk}^V \leqslant x_j^V, \quad \forall j; \quad \sum_{k=1}^{K} \lambda_k = 1; \quad \lambda_k \geqslant 0, \quad \forall k \right\} \quad (5)$$

Measure the efficiency θ using the DEA estimator. Define the output-oriented technical efficiency (OTE) as θ, where $D_y(\boldsymbol{x},\boldsymbol{y}) = \sup\{\theta | (\boldsymbol{x}, \theta\boldsymbol{y}) \in \widetilde{T}\}$, where $\theta \geqslant 1$, and a firm is efficient if $\theta = 1$.

Similarly, let $\boldsymbol{y}^E \in \mathbb{R}_+^Q$ denote an effective output vector. Define the effective production possibility set as $T^E = \{(\boldsymbol{x}, \boldsymbol{y}^E): \boldsymbol{x}$ can produce $\boldsymbol{y}^E$ that will be consumed in current period$\}$ and $T^E \subset T$. T^E illustrates the feasible region of the effective production possibility set which can be estimated by piece-wise linear concave envelopment truncated by demand level shown as model (6), where $\widetilde{T}^E$ is an estimated set of T^E.

$$\widetilde{T}^E = \left\{ (\boldsymbol{x},\boldsymbol{y}^E) \middle| \sum_{k=1}^{K} \lambda_k Y_{qk} \geqslant y_q^E, \quad \forall q; \quad D_q \geqslant y_q^E, \quad \forall q; \quad \sum_{k=1}^{K} \lambda_k X_{ik}^F \leqslant x_i^F, \right.$$
$$\left. \forall i; \quad \sum_{k=1}^{K} \lambda_k X_{jk}^V \leqslant x_j^V, \quad \forall j; \quad \sum_{k=1}^{K} \lambda_k = 1; \quad \lambda_k \geqslant 0, \quad \forall k \right\} \quad (6)$$

However, a generalized effectiveness measure is needed, because capacity shortage or capacity surplus leads to the profit loss. Y_{qk}^P is the amount of the qth penalized output with respect to demand D_{qk}. Measure effectiveness, θ^E, with respect to the DTPF. If there is a capacity shortage $Y_{qk} \leqslant D_{qk}$, then set $Y_{qk}^P = Y_{qk} - \alpha_{qk}(D_{qk} - Y_{qk}) \geqslant 0$, and the opportunity to sell $D_{qk} - Y_{qk}$ additional units is lost; otherwise $Y_{qk}^P = D_{qk} - \delta_{qk}(Y_{qk} - D_{qk}) \geqslant 0$, indicating a capacity surplus $Y_{qk} > D_{qk}$ and inventory $Y_{qk} - D_{qk}$ are generated, both of which are costly. Note that in calculating Y_{kj}^P, the penalty parameters $\alpha_{kj} \geqslant 0$ and $\delta_{kj} \geqslant 0$ are used to quantify the effect of lost sales and inventories, respectively, on effectiveness. If the cost of lost sales or holding inventory is insignificant, then $\alpha_{qk} = \delta_{qk} = 0$. Thus, we state the following definition

Definition 1. Production Effectiveness θ^E is the largest possible expansion of the penalized output at the given input level within the effective production possibility set, or mathematically, $D_y(x, y^P) = \sup\{\theta^E | (x, \theta^E y^P) \in \widetilde{T}^E\}$.[2]

In service industries, when a service is generated, it needs to be consumed immediately. Therefore, the surplus of services generated and not consumed represents a lost opportunity, $\delta_{qk} > 0$. In manufacturing industries, a surplus of capacity means lower utilization or machine idleness. However, if the product can be stored at a lower holding cost, the penalty for capacity surplus may be insignificant. On the other hand, if demand is not satisfied, both service and manufacturing firms reduce revenues by the size of the lost sales, and also lose market share, or damage their relationships with customers. When the total cost of missed sales is significant, $\alpha_{qk} > 0$.

This paper's definition of an effectiveness measure implies some notable issues. Because inventory holding costs are typically much smaller than the cost of lost sales, in many cases excess production will not result in significantly worse effectiveness scores.
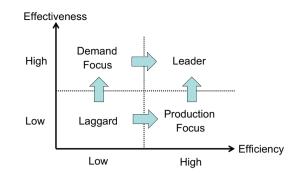

Fig. 2. Two-dimensional strategic position between efficiency and effectiveness.

However, when holding inventory is not possible, the effectiveness measure will significantly differ from the efficiency measure. These issues suggest that the method is most suitable for characterizing production systems with perishable goods, make-to-order production systems (pull systems), or service systems where services or inventories cannot be stored. In this paper, the proposed effectiveness measure is applied to Japanese convenience store. This industry is a business with high turn-over commodities and high product substitution. Each shop typically has a limited space for storing inventory, because of high rental costs, and the portion of daily-supplied foods is over 30% (Japan Franchise Association, 2010). The high ratio of perishable goods and inability to hold significant inventories justify the use of an effectiveness analysis.

3.3. Efficiency vs. effectiveness

Although efficiency and effectiveness are related metrics, they have different managerial implications. Efficiency measures the relative return on inputs used, whereas effectiveness indicates the ability to match demand given an existing production technology. High effectiveness generates revenues by providing products and services to customers; low effectiveness implies either poor resource management or an inability to meet customers' demands. Fig. 2 illustrates two-dimensional strategic position between efficiency and effectiveness. If the firm's efficiency and effectiveness are low, we use the term, "Laggard", since the firm tries to adopt its competitor's superior strategy to avoid being driven out of the industry. If the firm performs well in terms of efficiency and poorly in terms of effectiveness, we use the term, "Production Focus", since the firm tries to make the best use of its input resources and technology. If the firm performs poorly in terms of efficiency and well in terms of effectiveness, we use the term, "Demand Focus", since the firm tries to employ a market-oriented strategy to generate demand and maintain or expand market share. Finally, if the firm is both efficient and effective, we use the term, "Leader".

Assuming $\alpha_{qk} = \delta_{qk} = 0$, a significant gap between efficiency and effectiveness exists if demand is low, whereas efficiency and effectiveness are identical measures if demand is high. These observations show the measure of effectiveness is particularly important during economic downturns and lead to the following proposition.

Proposition 1. If there is no penalty for ineffective production, i.e. $\alpha_{qk} = 0$, then the effectiveness estimate converges to an efficiency as demand increases.

Proof 1. Based on the definition of effectiveness and model (6), for all output q we have $Y_q^P = Y_q$, and if $D_j \to \infty$, then the constraint $D_q \geqslant Y_q^P$ in model (6) is redundant. Thus, $\lim_{D_q \to \infty} \theta^E = \theta$. □

---

[2] In the capacity surplus case, if demand is low, the penalty will lead to Y_{qk}^P located outside of T^E (outside of the positive orthant). In this case the penalty is truncated by the x-axis (or $Y = 0$). Alternatively, a super efficiency measure could be used (Lovell & Rousem, 2003).

4. Capacity expansion and marginal product

Capacity is the maximal output level of a production process. The output is a result of the total productive capability of a firm's resources including workforce, machinery, and utilities. Capacity adjustment is the ability to adjust output levels to handle uncertainty by controlling variable resources in the short run (Alp & Tan, 2008). In production theory, capacity adjustment can be interpreted as the marginal product of the production function, i.e. the extra output generated by one more unit of an input. Fig. 3 shows the marginal product and production frontier with a single output. The production function's output level is defined as $y^{PF} = f(\mathbf{x}^F, \mathbf{x}^V)$. The marginal product at point B with respect to specific variable input j is $MP_{jB} = \frac{\partial f(\mathbf{x}^F, \mathbf{x}^V)}{\partial x_j^V}\bigg|_{X_{jB}^V}$.

There are two ways to estimate marginal product. SFA estimates the production function with a given functional form, e.g., a simple case of a linear function estimated by ordinary least squares (OLS) and the coefficients associated with the independent factors provide estimates of the marginal product. DEA constructs a piece-wise linear production function approximating the true production function, and the shadow prices of input and output characterize the marginal product. Both methods have drawbacks: SFA requires defining a functional form and risks potential misspecification, whereas DEA's observations on the production frontier do not have unique shadow prices.

Podinovski and Førsund (2010) propose a directional derivative technique to assess the marginal product of a nondifferential efficient frontier constructed by the DEA estimator. Let v_i^F, v_j^V, u_q and u_0 be the multipliers of fixed inputs, variables inputs, outputs, and the intercept term, respectively. Since marginal product is a characteristic of the frontier, for one specific efficient firm r, the following revised formulation calculates the marginal rate $\beta_{j^*q^*r}^{V+}$ approaching from the right side with respect to one particular variable input, j^*, and one output, q^*.

$$\frac{\partial^+ Y_{q^*r}}{\partial X_{j^*r}^V} = \beta_{j^*q^*r}^{V+} = \text{Min } v_{j^*}^V$$

$$s.t. \quad \sum_{i=1}^{I} v_i^F X_{ir}^F + \sum_{j=1}^{J} v_j^V X_{jr}^V - \sum_{q=1}^{Q} u_q Y_{qr} + u_0 = 0$$

$$\sum_{i=1}^{I} v_i^F X_{ik}^F + \sum_{j=1}^{J} v_j^V X_{jk}^V - \sum_{q=1}^{Q} u_q Y_{qk} + u_0 \geqslant 0 \quad (7)$$

$$u_{q^*} = 1$$

$$v_i^F, v_j^V, u_q \geqslant 0, \; u_0 \text{ is free}$$

To measure the marginal rate approaching from the left side, the objective function is replaced by the following equation.
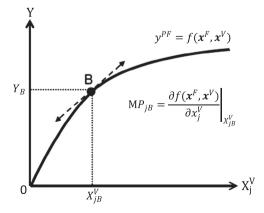

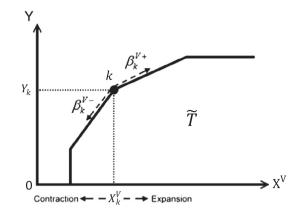**Fig. 3.** Marginal product of production function.

Fig. 4. Marginal product regarding short-run capacity expansion or contraction.

$$\frac{\partial^- Y_{q^*r}}{\partial X_{j^*r}^V} = \beta_{j^*q^*r}^{V-} = \text{Max } v_{j^*}^V \quad (8)$$

This paper uses $\beta_k^{V+} = \beta_{j^*q^*k}^{V+}$ and $\beta_k^{V-} = \beta_{j^*q^*k}^{V-}$ to denote the simplified notations of marginal products of short-run capacity expansion and contraction. Fig. 4 illustrates the marginal product β_k^{V+} or β_k^{V-} in the short run. Note that we do not define the marginal product for inefficient firms operating inside of the production frontier. However, to estimate how the consumed output expands with an increase of variable input, we assume that the marginal increase in output is the same as the marginal products of the reference firm on the frontier via output-oriented expansion.

5. Stochastic programming model

It is possible to formulate the capacity adjustment question as an if-then statement: if a particular demand scenario is realized, then how is short-run capacity adjusted? Section 5.1 discusses a scenario-based approach assuming that demand is a random variable. Defining a scenario by its demand level, an individual scenario analysis specifies an adjustment to variable input level for each potential outcome. Section 5.2 discusses a two-stage recourse problem (RP), since the scenario-based approach does not consider the robustness of solutions under varying scenarios.

5.1. Scenario-based approach

We use a scenario-based approach when we define a scenario as a particular realization of demand. Given a description of alternative potential scenarios, the model suggests suitable decisions regarding short-run capacity expansion or contraction under demand fluctuations. Note that the range of adjustment of variable input levels is limited to model-realistic limitations on management. We estimate the production possibility set by using the observed production data; the set does not change regardless of the assumptions about short-run capacity expansion.

There are two kinds of solutions: an individual scenario analysis, and an expected value (EV) solution (Birge & Louveaux, 2011). An individual scenario analysis represents the variable adjustment for one specific demand scenario; however, the EV solution is obtained by solving the model using the expected value of the random variable(s). We define the expected value of demand as $\overline{D}_r = \sum_{s=1}^{S} p_s D_{rs}$, where p_s represents the probability of sth scenario occurring and D_{rs} be the realized demand of rth firm in sth scenario.

Assuming fixed inputs cannot be adjusted in the short run, we develop a capacity expansion model based on the output-oriented variable returns to scale (VRS) DEA formulation with single output

as a revised dual form of effectiveness measure $D_y(\boldsymbol{x}, \boldsymbol{y}^P) = \sup\left\{\theta^E | (\boldsymbol{x}, \theta^E \boldsymbol{y}^P) \in \widetilde{T}^E\right\}$. Let Y_{ks} be the actual output, β_{jr}^{V+} be the marginal product characterized by β_{jr}^{V+} and β_{jr}^{V-} with respect to jth variable input of firm r, and R_{jr} be the parameter of adjustable range. Then the decision variables u_s, w_s, v_{is}^F, v_{js}^V, v_{0s} are the multipliers associated with outputs, demand, fixed inputs, variable inputs, and the VRS constraints of typical DEA formulation, respectively, d_{jrs} is the additional adjustment of variable input composed of positive $\left(d_{jrs}^+\right)$ or negative $\left(d_{jrs}^-\right)$ adjustments, y_{rs} and y_{rs}^P are the adjusted and penalized output, respectively, and θ_{rs}^E measures production effectiveness. To maintain the identical production possibility set after capacity expansion or contraction of firm r, index k includes a firm r' which contains the data for firm r before capacity expansion.

$$\text{Min} \quad M\theta_{rs}^E + \sum_{j=1}^{J}\left(d_{jrs}^+ + d_{jrs}^-\right) \tag{9.1}$$

$$s.t. \quad \theta_{rs}^E = \sum_{i=1}^{I} v_{is}^F X_{ir}^F + \sum_{j=1}^{J} v_{js}^V\left(X_{jr}^V + d_{jrs}\right) + w_s D_{rs} + v_{0s} \tag{9.2}$$

$$u_s y_{rs}^P + w_s y_{rs}^P = 1 \tag{9.3}$$

$$\sum_{i=1}^{I} v_{is}^F X_{ik}^F + \sum_{j=1}^{J} v_{js}^V X_{jk}^V - u_s Y_{ks} + v_{0s} \geq 0, \quad \forall k \setminus r \tag{9.4}$$

$$\sum_{i=1}^{I} v_{is}^F X_{ir}^F + \sum_{j=1}^{J} v_{js}^V\left(X_{jr}^V + d_{jrs}\right) - u_s y_{rs}^P + v_{0s} \geq 0 \tag{9.5}$$

$$y_{rs}^P = [y_{rs} - \alpha_{rs}(D_{rs} - y_{rs})](1 - z1_{rs}) \tag{9.6}$$

$$+ [D_{rs} - \delta_{rs}(y_{rs} - D_{rs})]z1_{rs} + \varepsilon \tag{9.7}$$

$$y_{rs} - D_{rs} < Mz1_{rs} \tag{9.8}$$

$$y_{rs} - D_{rs} \geq -M(1 - z1_{rs}) \tag{9.9}$$

$$y_{rs} = Y_r + \sum_{j=1}^{J} \beta_{jr}^V d_{jrs} \tag{9.10}$$

$$\beta_{jr}^V = \beta_{jr}^{V+} z2_{jrs} + \beta_{jr}^{V-}(1 - z2_{jrs}), \quad \forall j \tag{9.11}$$

$$d_{jrs} < Mz2_{jrs}, \qquad \forall j \tag{9.12}$$

$$d_{jrs} \geq -M(1 - z2_{jrs}), \quad \forall j \tag{9.13}$$

$$d_{jrs} = d_{jrs}^+ - d_{jrs}^-, \quad \forall j \tag{9.14}$$

$$-R_{jr}X_{jr}^V \leq d_{jrs} \leq R_{jr}X_{jr}^V, \quad \forall j \tag{9.15}$$

$$z1_{rs}, z2_{jrs} \in \{0, 1\}, \quad \forall j \tag{9.16}$$

$$y_{rs}^P, y_{rs}, u_s, w_s, v_{is}^F, v_{js}^V \geq 0, \quad \forall i, \forall j \tag{9.17}$$

$$d_{jrs}^+, d_{jrs}^- \geq 0 \text{ as integers}, \quad \forall j \tag{9.18}$$

The objective function Eq. (9.1) minimizes the product of the estimated effectiveness θ_{rs}^E with a large number M and a secondary objective of minimizing the variation in input adjustment. Note that we use an output-oriented DEA model. Eqs. (9.2)–(9.4) and (9.5) are the envelope constraints to build the production possibility set. ε is a small number to maintain feasibility when the penalized output $y_{rs}^P = 0$. Constraints (9.6), (9.7) and (9.8) calculate the effective output level for the two cases, $Y_r \leq D_{rs}$ and $Y_r > D_{rs}$. Eq. (9.9) determines the actual output level of firm r through capacity expansion. Constraints (9.10), (9.11) and (9.12) calculate the marginal output for short-run capacity expansion, i.e., $d_{jrs} \geq 0$, then $\beta_{jr}^V = \beta_{jr}^{V+}$; otherwise $\beta_{jr}^V = \beta_{jr}^{V-}$. Constraint (9.13) shows the adjustment of variable input via goal programming, $d_{jrs} \geq 0$ if and only if capacity expansion
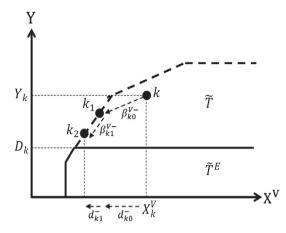

Fig. 5. Contraction example of the proposed algorithm.

with $d_{jrs}^+ \geq 0$; otherwise capacity contracts with $d_{jrs} \leq 0$ if and only if $d_{jrs}^- \geq 0$. The adjustment range is restricted in Eq. (9.14). Eq. (9.15) defines $z1_{rs}$ and $z2_{jrs}$ as binary variables. Eq. (9.16) includes nonnegative constraints. d_{jrs}^+ and d_{jrs}^- are integers because we are considering labor as a variable input. We drop constraint (9.17) if the variable input is not required to be an integer.[3] In fact, the model (9.1)–(9.16) and (9.17) is a mixed-integer nonlinear problem which is difficult to solve. A linearization and reformulation technique is used to solve the proposed model (see Appendix).

Note that the estimation of the increase in output is conservative if two or more variable inputs are expanded simultaneously, i.e. we estimate the marginal production of each variable input separately and then take the dot product of the marginal product vector. If there is a synergistic effect between the different variable inputs, this is not captured. However, because the production frontier limits the output level, the benefits of increasing multiple variables inputs leads to a resulting production vector within the production possibility set.

In order to maintain feasibility, meaning that a firm remains within the original production possibility set after taking the adjustment action, we calculate the effectiveness and resource adjustments using the proposed algorithm. Fig. 5 shows a contraction example. The proposed algorithm leads to Observation 1:

Observation 1. A firm that expands or contracts short-run capacity via the proposed algorithm with marginal products β_{jk}^{V+} or β_{jk}^{V-} remains feasible in the original production possibility set.

<u>Proposed Algorithm</u>

1. For one specific demand scenario D_r, start from specific firm $r = 1$.
2. For $r = 1$ to number of firms.
   - 2.1 Set step $t = 0$, $X_{jrt}^V = X_{jr}^V$ and $Y_{rt} = Y_r$.
   - 2.2 Calculate marginal products β_{jrt}^{V+} and β_{jrt}^{V-}.
   - 2.3 Run scenario-based approach (9.1)–(9.16) and (9.17) to calculate $X_{jrt}^V + d_{jrt}$, $\forall j$ and $Y_{rt} + \sum_{j=1}^{J} \beta_{jrt}^V d_{jrt}$.
   - 2.4 If $d_{jr} = \sum_t d_{jrt} = 0$, then go to step 2.8; otherwise run output-oriented DEA estimator to calculate efficiency θ_{rt}^{DEA}.[4]

2.5 If $\theta_{rt}^{DEA} \geqslant 1$, then get θ_{rt}^E, $d_{jr} = \sum_t d_{jrt}$, $\forall j$, and $Y_r = Y_{rt} + \sum_{j=1}^J \beta_{jrt}^V d_{jrt}$. Go to step 2.8.

Else if $\theta_{rt}^{DEA} < 1$ and $d_{jr} = \sum_t d_{jrt} < 0$, then run (10) to hold back the input adjustment on the efficient frontier

$$\text{Min} \quad \sum_{j=1}^J \left(X_{jrt}^V + d_{jrt} \right)$$

$$s.t. \quad \sum_{k=1}^K \lambda_k Y_k = Y_{rt} + \sum_{j=1}^J \beta_{jrt}^{V-} d_{jrt}$$

$$\sum_{k=1}^K \lambda_k X_{ik}^F \leqslant X_{ir}^F, \ \forall i$$

$$\sum_{k=1}^K \lambda_k X_{jk}^V \leqslant X_{jrt}^V + d_{jrt}, \ \forall j \tag{10}$$

$$\sum_{k=1}^K \lambda_k = 1$$

$$- R_{jr} X_{jr}^V \leqslant \left(X_{jrt}^V - X_{jr}^V \right) + d_{jrt}$$

$$\lambda_k \geqslant 0, \ \forall k$$

Else run the model (11)

$$\text{Max} \quad \sum_{j=1}^J \left(X_{jrt}^V + d_{jrt} \right)$$

$$s.t. \quad \sum_{k=1}^K \lambda_k Y_k = Y_{rt} + \sum_{j=1}^J \beta_{jrt}^{V+} d_{jrt}$$

$$\sum_{k=1}^K \lambda_k X_{ik}^F \leqslant X_{ir}^F, \ \forall i$$

$$\sum_{k=1}^K \lambda_k X_{jk}^V \leqslant X_{jrt}^V + d_{jrt}, \ \forall j \tag{11}$$

$$\sum_{k=1}^K \lambda_k = 1$$

$$R_{jr} X_{jr}^V \geqslant \left(X_{jrt}^V - X_{jr}^V \right) + d_{jrt}$$

$$\lambda_k \geqslant 0, \ \forall k$$

2.6 Set $X_{jr(t+1)}^V = X_{jrt}^V + d_{jrt}$, $\forall j$ and $Y_{r(t+1)} = Y_{rt} + \sum_{j=1}^J \beta_{jrt}^V d_{jrt}$.

2.7 Set $t = t + 1$ and go to step 2.2.

2.8 Set $r = r + 1$ and go to step 2.1.

5.2. Two-stage recourse approach

The two-stage recourse approach illustrated in Fig. 6 provides a robust solution to all scenarios. In our model, the first-stage decision, i.e. the here-and-now decision, corresponds to selecting the variable input level based on the demand forecasts. After demand is realized, the second-stage decision, i.e. the wait-and-see decision, corresponds to the measurement of effectiveness.

The two-stage recourse approach introduces an *expected recourse function* which characterizes the performance of the second stage, which in our formulation is the effectiveness estimates based on the short-run capacity decisions for each specific

Fig. 6. Two-stage recourse problem.

scenario. This approach generates robust solutions to the first-stage decisions. The two-stage recourse DEA model with capacity adjustment and uncertain demand is shown below.

$$\text{Min} \quad M \sum_{s=1}^S p_s \theta_{rs}^E + \sum_{j=1}^J \left(d_{jr}^+ + d_{jr}^- \right) \tag{12.1}$$

$$s.t. \quad \theta_{rs}^E = \sum_{i=1}^I v_{is}^F X_{ir}^F + \sum_{j=1}^J v_{js}^V \left(X_{jr}^V + d_{jr} \right) + w_s D_{rs} + v_{0s}, \ \forall s \tag{12.2}$$

$$u_s y_{rs}^P + w_s y_{rs}^P = 1, \ \forall s \tag{12.3}$$

$$\sum_{i=1}^I v_{is}^F X_{ik}^F + \sum_{j=1}^J v_{js}^V X_{jk}^V - u_s Y_{ks} + v_{0s} \geqslant 0, \ \forall k \setminus r, \forall s \tag{12.4}$$

$$\sum_{i=1}^I v_{is}^F X_{ir}^F + \sum_{j=1}^J v_{js}^V \left(X_{jr}^V + d_{jr} \right) - u_s y_{rs}^P + v_{0s} \geqslant 0, \ \forall s \tag{12.5}$$

$$y_{rs}^P = [y_r - \alpha_r(D_{rs} - y_r)](1 - z1_{rs}) + [D_{rs} - \delta_r(y_r - D_{rs})]z1_{rs} + \varepsilon, \ \forall s \tag{12.6}$$

$$y_r - D_{rs} < Mz1_{rs}, \ \forall s \tag{12.7}$$

$$y_r - D_{rs} \geqslant -M(1 - z1_{rs}), \ \forall s \tag{12.8}$$

$$y_r = Y_r + \sum_{j=1}^J \beta_{jr}^V d_{jr} \tag{12.9}$$

$$\beta_{jr}^V = \beta_{jr}^{V+} z2_{jr} + \beta_{jr}^{V-}(1 - z2_{jr}), \ \forall j \tag{12.10}$$

$$d_{jr} < Mz2_{jr}, \ \forall j \tag{12.11}$$

$$d_{jr} \geqslant -M(1 - z2_{jr}), \ \forall j \tag{12.12}$$

$$d_{jr} = d_{jr}^+ - d_{jr}^-, \ \forall j \tag{12.13}$$

$$- R_{jr} X_{jr}^V \leqslant d_{jr} \leqslant R_{jr} X_{jr}^V, \ \forall j \tag{12.14}$$

$$z2_{jr} \in \{0, 1\}, \ \forall j \tag{12.15}$$

$$z1_{rs} \in \{0, 1\}, \ \forall s \tag{12.16}$$

$$y_{rs}^P, y_r, d_{jr}^+, d_{jr}^-, u_s, w_s, v_{is}^F, v_{js}^V \geqslant 0, \ \forall i, \forall j, \forall s \tag{12.17}$$

$$d_{jr}^+, d_{jr}^- \geqslant 0 \text{ as integers}, \ \forall j \tag{12.18}$$

The objective function (12.1) is to minimize the effectiveness estimate θ_{rs}^E aggregated over scenario s with probability measures p_s and to minimize the adjustment variation for firm r. The two-stage recourse problem (RP) provides an adjustment to the variable input level d_{jr}, d_{jr}^+, d_{jr}^- and output level y_r in the first stage and attempts to minimize expected effectiveness by optimizing the variables y_{rs}^P, w_s, u_s, v_{is}^F, v_{js}^V, and v_{0s} in the second stage after the demand is realized. We define the recourse function as $g(d_{jr}, D_{rs}) = \min_{u, v \geqslant 0} \{\theta_{rs}^E | (12.2)-(12.8), (12.16)-(12.18)\}$ where d_{jr} is the first-stage decision and D_{rs} is a realized outcome of random variable $\widetilde{D}_{rs}$. The expected value of the recourse function is $E[g(d_{jr}, \widetilde{D}_{rs})] = \sum_s p_s \theta_{rs}^E$ under discrete scenarios. In the proposed algorithm we replace Step 1 by the realized outcome of demand distribution D_{rs} and Step 2.3 by using the two-stage recourse approach (12.1)–(12.12), (12.12)–(12.17) and (12.18).

5.3. Value of information and stochastic solution

With the solutions generated from model (9) and model (12), we now investigate the *quality* of the solutions. The most common metrics are the expected value of perfect information (EVPI), and the value of the stochastic solution (VSS) (Birge & Louveaux, 2011). EVPI measures the maximum amount the decision-maker is willing to pay in return for complete information about the future. By defining the effectiveness measures from the wait-and-see (WS) problem and the recourse problem (RP) as $WS = E_{\widetilde{D}}[\text{Max } g(d, \widetilde{D})]$ and $RP = \text{Max } E_{\widetilde{D}}[g(d, \widetilde{D})]$, we can then define EVPI as Eq. (13).

$$EVPI = WS - RP = E_{\widetilde{D}}[\text{Max } g(d, \widetilde{D})] - \text{Max } E_{\widetilde{D}}[g(d, \widetilde{D})] \qquad (13)$$

VSS is a measure of the quality of the expected value decision in terms of the recourse problem. Namely, it tells us the cost of ignoring uncertainty. Let $\bar{d}(\bar{D})$ be an EV solution and define the *expected result of using the EV solution* (EEV) as $EEV = E_{\widetilde{D}}[g(\bar{d}(\bar{D}), \widetilde{D})]$. Define VSS as follows.

$$VSS = RP - EEV = \text{Max } E_{\widetilde{D}}[g(d, \widetilde{D})] - E_{\widetilde{D}}[g(\bar{d}(\bar{D}), \widetilde{D})] \qquad (14)$$

6. Numerical example

This section discusses a numerical example without capacity expansion ($d_{jr} = 0$) to illustrate the effectiveness measure with $\alpha_k = 0$ and $\delta_k = 1$. Table 2 shows the data for 12 firms, including a fixed input, variable input, actual output, and three demand forecasts: pessimistic (PE), most-likely (ML), and optimistic (OP). This example postulates the probability of realizing each of the demand scenarios as 1/6 for the PE, 4/6 for the ML and 1/6 for the OP demand forecasts. Table 3 shows the efficiency and effectiveness measures. Note that PE, ML, OP, and EV are deterministic (and use the scenario-based approach), and SP uses the two-stage recourse problem (RP). Specifically, the difference between the EV solution and the SP solution is that the former calculates effectiveness using expected demand $\bar{D}_k$, whereas the latter estimates the expected value of effectiveness of the three demand scenarios. For illustrative purposes, we use the inverses of efficiency and effectiveness to restrict the scores between 0 and 1 in Sections 6 and 7.

Fig. 7 maps the efficiency and the effectiveness levels with the different demand scenarios on a two-dimensional coordinate graph whose four quadrants indicate the strategic position which

Table 2
Data for numerical example.

Firms	Fixed input	Var. input	Actual output	Pessimistic demand	Most-likely demand	Optimistic demand
A	9	5	10	6	9	12
B	4	7	8	5	6	9
C	4	9	11	6	8	13
D	5	9	9	7	8	10
E	7	7	10	7	9	13
F	6	7	7	4	6	9
G	10	8	10	7	8	11
H	8	6	7	7	8	9
I	5	6	11	6	7	12
J	4.5	10	10	8	10	12
K	4	8	12	7	8	12
L	10	7	5	3	5	8

Table 3
Efficiency, effectiveness, and EVPI.

Firm	Efficiency	Effectiveness						
		PE	ML	OP	EV	SP (RP)	WS	EVPI
A	1.00	0.33	0.89	1.00	0.89	0.82	0.82	0.00
B	1.00	0.40	0.67	1.00	0.74	0.68	0.68	0.00
C	0.92	0.17	0.63	0.92	0.71	0.60	0.60	0.00
D	0.75	0.71	0.88	0.90	0.90	0.85	0.85	0.00
E	0.87	0.57	0.89	0.87	0.93	0.83	0.83	0.00
F	0.61	0.25	0.83	0.78	0.87	0.73	0.73	0.00
G	0.83	0.57	0.75	0.91	0.80	0.75	0.75	0.00
H	0.64	1.00	0.88	0.78	0.88	0.88	0.88	0.00
I	1.00	0.17	0.43	1.00	0.57	0.48	0.48	0.00
J	0.83	0.75	1.00	0.83	1.00	0.93	0.93	0.00
K	1.00	0.29	0.50	1.00	0.59	0.55	0.55	0.00
L	0.44	0.33	1.00	0.63	0.97	0.83	0.83	0.00
Avg.	0.855	0.449	0.755	0.903	0.802	0.729	0.729	

provides guidelines of productivity improvement. The intersection of two axes describes the performance of industry level which is the average of the 12 firms weighted by actual output. For the PE-demand case, firms F and L are the laggards (the terminology introduced in Fig. 2), firms D, G, H, and J belong to demand focus, firms A, B, C, I, and K are attributed to production focus, and firm E is the leader. When demand is high, note that it does not limit the production possibility set and that effectiveness closely correlates with efficiency, i.e. the 12 firms tend to a diagonal line. This result is consistent with Proposition 1.

Fig. 7 shows the convergence process from the pessimistic to the optimistic demand scenario, i.e. effectiveness provides additional information beyond an efficiency measure during economic downturns. Note that the results show that all firms prefer to under-produce rather than over-produce, because inventory implies ineffective production. For example, firm C is highly ineffective in the PE and the ML cases, but the production function forms a demand-supporting limitation in the OP case. Thus, firm C must reduce its output level if it wants to become more effective.

Finally, Fig. 8 shows that the results of EV and SP are similar. When the variable defining scenarios is continuous, as is the case with demand, we would expect the results to be similar.[5] In addition, the last column of Table 3 indicates that all EVPI values equal zero; thus, there is no value in paying for perfect information because capacity adjustment is not considered.

7. Empirical study

This section discusses an application in an empirical study of the performance of 25 convenience stores in Japan under demand uncertainty in the first half of 2003. Since the opening of the first convenience stores in 1969, today about 42,889 stores serve 1.1 billion customers annually (Japan Franchise Association, 2010). Most convenience stores have a floor area of 100 square meters and carry about 3000 product types. Even though their prices are generally higher than supermarket prices, they are popular due to the range of services offered, e.g., courier and postal services, touch-screen monitors for finding jobs and ordering tickets, utility payments, ATMs, online shopping, and 24/7 convenience (Nipponia, 2001). Section 7.1 presents the dataset. Section 7.2 describes the results of estimating effectiveness using the scenario-based and the two-stage recourse approaches. We solve the proposed model using LINGO 9.0 Solver.
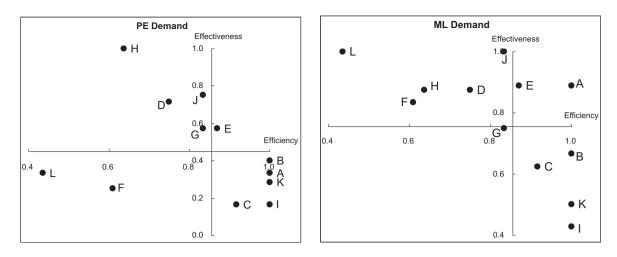
7.1. Data description

Table 4 summarizes the data from Sueyoshi (2003). In general, the convenience stores industry production process can be characterized by three input resources: capital, branch size (fixed inputs), and employee (variable input). The forecasted demand and actual output are measured in units; actual output is estimated by realized revenue divided by the average price per item.[6] Similarly, demand is calculated by estimating revenue divided by average price per item. Efficiency is estimated by using the actual output and effectiveness is estimated by using the proposed DTPF.

We use the following definitions: capital is the net worth of the equipment used to operate the convenience stores; branch size and employee are the number of branches and manpower headcount in the first half of 2003. Firm-specific demand is the estimated number of goods sold and characterized by the pessimistic, most-likely, and optimistic estimates provided by managers and chief

---

[5] However, when the scenarios are discrete, the results estimated using EV typically will not be realized. For SP the results are estimated for each scenario; thus the interpretation of the results is more direct.

[6] Data from the Japan Franchise Association (2010).

Fig. 7. Efficiency vs. effectiveness.

Fig. 8. Expected value solution vs. stochastic programming.

executive officers; see Sueyoshi (2003) for more details. The marginal product defined in section 4 is estimated by the Podinovski and Førsund (2010) method. Each convenience store can adjust capacity over some limited range by hiring/laying off employees. For simplicity, we limit the positive and negative adjustment in manpower resources to 15%; however, testing other values for this parameter produces proportionally similar results.

7.2. Efficiency and effectiveness analysis

This section illustrates efficiency and effectiveness estimation in a demand-dependent context. A demand-dependent context means that all firms consistently realize the same single demand state (such as pessimistic, most-likely, or optimistic). We postulate that demand $\bar{D}_k$ follows a beta distribution. The expected value can

Table 4
Descriptive statistics of raw data.

Stat.	Fixed input		Var. input	Output	Demand		
	Capital	Branch	Employee	Goods	Pessimistic	Most-likely	Optimistic
	Yen (10^6)	Unit	Headcount	Unit (10^6)			
Average	3452.5	1421.4	683.5	794.4	795.5	857.9	920.4
Std. dev.	7465.8	2291.5	1068.3	1330.2	1492.9	1558.6	1625.9
Max	30876.0	7780.0	4126.0	6191.3	6358.8	6586.6	6814.4
Min	3.0	22.0	8.0	12.1	11.7	13.1	14.7

Table 5
Efficiency, effectiveness, EVPI, and VSS.

CVS	Firm	Efficiency			Effectiveness															WS	EVPI	EEV	VSS
					PE			ML			OP			EV			RP						
		N	Y	Exp.	N	Y	Exp.	N	Y	Exp.	N	Y	Exp.	N	Y	Exp.	N	Y	Exp.				
am/pm	A	0.91	0.91	0	0.97	0.97	0	0.91	0.91	0	0.91	0.91	0	0.91	0.91	0	0.92	0.92	0	0.92	0.00	0.92	0.00
Heart in	B	0.82	0.82	0	0.97	0.97	0	0.95	0.95	0	0.88	0.88	0	0.95	0.95	0	0.94	0.94	0	0.94	0.00	0.94	0.00
HOTSPAR	C	0.54	0.55	−48	0.81	0.81	0	0.96	0.96	0	0.80	0.80	0	0.96	0.96	0	0.91	0.91	0	0.91	0.00	0.91	0.00
Apple Mart	D	1.00	1.00	0	0.89	0.89	0	1.00	1.00	0	1.00	1.00	0	1.00	1.00	0	0.98	0.98	0	0.98	0.00	0.98	0.00
Everyone	E	0.85	0.85	0	0.82	0.82	0	0.85	0.85	0	0.85	0.85	0	0.85	0.85	0	0.85	0.85	0	0.85	0.00	0.85	0.00
Caramel Mart	F	0.72	0.72	0	0.71	0.71	0	0.89	0.89	0	0.72	0.72	0	0.89	0.89	0	0.83	0.83	0	0.83	0.00	0.83	0.00
Coco Store	G	1.00	1.00	0	0.95	1.00	−13	1.00	1.00	0	1.00	1.00	0	1.00	1.00	0	0.99	1.00	−13	1.00	0.00	0.99	0.01
Community Store	H	1.00	1.00	0	0.87	0.92	−23	1.00	1.00	0	1.00	1.00	0	1.00	1.00	0	0.98	0.99	−23	0.99	0.00	0.98	0.01
Circle K	I	0.77	0.83	134	0.96	1.00	33	0.90	1.00	84	0.85	1.00	134	0.90	1.00	84	0.90	0.98	84	1.00	0.02	0.90	0.08
Sunkus	J	0.76	0.78	−141	0.96	1.00	67	0.92	1.00	121	0.89	0.98	141	0.92	1.00	121	0.92	0.99	121	1.00	0.01	0.92	0.07
Shop and Life	K	0.61	0.65	5	0.86	0.86	0	0.98	1.00	−2	0.94	1.00	4	0.97	1.00	−2	0.95	0.97	−2	0.98	0.01	0.95	0.02
Spar	L	0.91	0.91	0	0.87	0.87	0	0.92	0.92	0	0.91	0.91	0	0.92	0.92	0	0.91	0.91	0	0.91	0.00	0.91	0.00
Three F	M	0.76	0.76	0	0.95	0.95	0	0.91	0.91	0	0.88	0.88	0	0.91	0.91	0	0.92	0.92	0	0.92	0.00	0.92	0.00
Seikatsu Train	N	0.94	0.94	0	0.98	0.98	0	0.94	0.94	0	0.94	0.94	0	0.94	0.94	0	0.95	0.95	0	0.95	0.00	0.95	0.00
Seicomart	O	1.00	1.00	0	0.85	0.92	−40	1.00	1.00	0	1.00	1.00	0	1.00	1.00	0	0.98	0.99	−40	0.99	0.00	0.98	0.01
Seven Eleven	P	1.00	1.00	0	1.00	1.00	0	1.00	1.00	0	1.00	1.00	0	1.00	1.00	0	1.00	1.00	0	1.00	0.00	1.00	0.00
Daily Yamazaki	Q	1.00	1.00	0	0.86	1.00	−152	1.00	1.00	0	1.00	1.00	0	1.00	1.00	0	0.98	1.00	−152	1.00	0.00	0.98	0.02
Hiromaru Chain	R	0.47	0.47	0	0.79	0.79	0	1.00	1.00	0	0.84	0.84	0	1.00	1.00	0	0.94	0.94	0	0.94	0.00	0.94	0.00
Family Mart	S	0.76	0.76	0	0.97	1.00	44	0.90	1.00	152	0.85	1.00	260	0.90	1.00	152	0.90	0.98	152	1.00	0.02	0.90	0.08
My Shop Chain	T	1.00	1.00	0	0.70	0.70	0	1.00	1.00	0	1.00	1.00	0	1.00	1.00	0	0.95	0.95	0	0.95	0.00	0.95	0.00
Monpellie	U	1.00	1.00	0	1.00	1.00	0	1.00	1.00	0	1.00	1.00	0	1.00	1.00	0	1.00	1.00	0	1.00	0.00	1.00	0.00
Mon Mart	V	1.00	1.00	0	0.91	0.91	0	0.98	0.98	0	1.00	1.00	0	0.98	0.98	0	0.97	0.97	0	0.97	0.00	0.97	0.00
Lics	W	1.00	1.00	0	0.97	0.97	0	1.00	1.00	0	1.00	1.00	0	1.00	1.00	0	1.00	1.00	0	1.00	0.00	1.00	0.00
Little Star	X	0.74	0.74	0	0.83	0.83	0	0.89	0.89	0	0.74	0.74	0	0.90	0.90	0	0.86	0.86	0	0.86	0.00	0.86	0.00
Lawson	Y	0.69	0.69	0	0.98	0.98	0	0.92	0.92	0	0.87	0.87	0	0.92	0.92	0	0.93	0.93	0	0.93	0.00	0.93	0.00
Avg.		0.857	0.863		0.960	0.983		0.955	0.980		0.927	0.965		0.955	0.980		0.951	0.972					

be estimated as $\overline{D}_k = (PE_k + 4ML_k + OP_k)/6$.[7] Table 5 presents the efficiency, effectiveness, EVPI, and VSS estimations. N means that short-run capacity expansion is not allowed and Y means that manpower expansion is allowed. The "Exp." column indicates the change in variable input where positive values indicate short-run capacity expansion and negative values indicate contraction. The penalty is only for capacity surplus, i.e. due to the service industry of convenience stores. Table 5 shows the manpower adjustment for productivity improvement, for example, Family Mart needs to hire additional 152 employees to increase its effectiveness score from 0.9 to 1.00 in the ML demand case.

We note that the convenience stores empirical study aligns with the numerical study, in that higher demand increases the correlation between effectiveness and efficiency. The correlation coefficients are 0.26 and 0.76 in the pessimistic and the optimistic demand cases without capacity expansion, respectively. In addition, EVPI and VSS in most of the convenience stores chains are close to or equal to zero; thus, there is almost no difference between the two-stage RP and the expected value of EV, as shown

in Table 5. This result is driven by the limited demand fluctuations in the convenience stores industry during the first half of 2003. However, we prefer the EV method because it is computationally easier and provides similar results.

Concerning the ex-ante and ex-post analysis of short-run capacity expansion, Fig. 9 maps the efficiency and effectiveness under pessimistic demand on a two-dimensional coordinate and identifies the four quadrants representing different strategies similar to Fig. 2. The intersection of two coordinate axes indicates the industry's weighted average performance. For the ex-post analysis, we calculate the efficiency after applying the expansion of variable input recommended by the model (9) which may reduce efficiency to increase effectiveness. Again, the results are almost identical.

8. Conclusion

Proactive DEA embedded with stochastic programming technique estimates a production function and make short-run capacity decisions under demand uncertainty. The results of an empirical study of Japanese convenience stores show that the proposed model improves upon previous models by considering varying marginal product rates for expansion and contraction and the limitations due to diminishing returns. In addition, efficiency and

---

[7] The beta distribution is widely used in project management with PERT (Program Evaluation and Review Technique) and CPM (Critical Path Method) to plan activity times and scheduling (Hillier & Lieberman, 2002).
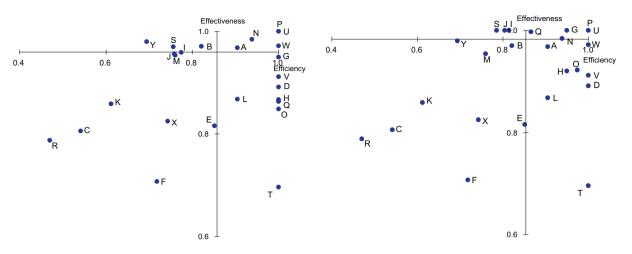
Fig. 9. Efficiency vs. effectiveness with pessimistic demand before and after expansion.

effectiveness estimates easily identify the influence of demand on productivity analysis.

Effectiveness complements efficiency measure. They are not mutually independent since DTPF is generated based on production function, but they have different managerial implications. Efficiency measures the relative return on inputs used, whereas effectiveness indicates the ability to match demand given an existing production technology. The revenue is directly associated with effectiveness measure. Two-dimensional strategy position supports the development of productivity improvement.

This paper makes three contributions to the DEA literature. First, we extend ex-post DEA efficiency estimation to aid in future planning, i.e. proactive DEA. Second, we propose an effectiveness measure and a demand-truncated production function to address demand, which is rarely discussed in the literature. Third, by considering the diminishing marginal benefits of inputs via marginal production estimation we ensure that the capacity adjustment remains inside of the production possibility set.

We note that this paper only considers the case of the change of one input contributing to the change of one output. In fact, the adjustment of one input can contribute to multiple outputs due to the law of marginal rate of technical substitution. Future development of a multi-output model with price information would be a valuable contribution. Moreover, the panel data and a dynamic analysis will be useful for supporting a sequential control of resource.

Acknowledgment

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.ejor.2013.07.043.

References

Alp, O., & Tan, R. (2008). Tactical capacity management under capacity flexibility. *IIE Transactions, 40*, 221–237.

Beraldi, P., & Bruni, M. E. (2012). Data envelopment analysis under uncertainty and risk. *World Academy of Science, Engineering and Technology, 66*, 837–842.

Birge, J. R., & Louveaux, F. (2011). *Introduction to stochastic programming* (2nd ed.). New York: Springer Verlag.

Branda, M. (2013). Diversification-consistent data envelopment analysis with general deviation measures. *European Journal of Operational Research, 226*(3), 626–635.

Bruni, M. E., Conforti, D., Beraldi, P., & Tundis, E. (2009). Probabilistically constrained models for efficiency and dominance in DEA. *International Journal of Production Economics, 117*(1), 219–228.

Coelli, T. J., Rao, D. S. P., O'Donnell, C. J., & Battese, G. (2005). *An introduction to efficiency and productivity analysis* (2nd ed.). New York: Springer.

Cooper, W. W., Huang, Z. M., & Li, S. X. (1996). Satisficing DEA models under chance constraints. *Annals of Operations Research, 66*, 279–295.

Cooper, W. W., Park, K. S., & Yu, G. (1999). IDEA and AR-IDEA: Models for dealing with imprecise data in DEA. *Management Science, 45*, 597–607.

Dyson, R. G., & Shale, E. (2010). Data envelopment analysis, operational research and uncertainty. *Journal of the Operational Research Society, 61*(1), 25–34.

Hillier, F. S., & Lieberman, G. J. (2002). *Introduction to operations research* (7th ed.). New York: McGraw-Hill. Chapter 10.

Japan Franchise Association, 2010. *Convenience store statistics report* <http://www.jfa-fc.or.jp/particle/42.html>.

Kao, C., & Liu, S.-T. (2000). Fuzzy efficiency measures in data envelopment analysis. *Fuzzy Sets and Systems, 113*, 427–437.

Kao, C., & Liu, S.-T. (2009). Stochastic data envelopment analysis in measuring the efficiency of Taiwan commercial banks. *European Journal of Operational Research, 196*, 312–322.

Kuosmanen, T. (2004). Efficient diversification according to stochastic dominance criteria. *Management Science, 50*(10), 1390–1406.

Kuosmanen, T. (2008). Representation theorem for convex nonparametric least squares. *The Econometrics Journal, 11*, 308–325.

Kuosmanen, T., & Johnson, A. (2010). Data envelopment analysis as nonparametric least squares regression. *Operations Research, 58*(1), 149–160.

Kuosmanen, T., & Kortelainen, M. (2012). Stochastic non-smooth envelopment of data: Semi-parametric frontier estimation subject to shape constraints. *Journal of Productivity Analysis, 38*(1), 11–28.

Lamb, J. D., & Tee, K.-H. (2012). Data envelopment analysis models of investment funds. *European Journal of Operational Research, 216*(3), 687–696.

Land, K. C., Lovell, C. A. K., & Thore, S. (1993). Chance-constrained data envelopment analysis. *Managerial and Decision Economics, 14*(6), 541–554.

Lee, C.-Y., & Johnson, A.L., 2012a. *Effectiveness: A measure of demand effect in productivity analysis* <http://ssrn.com/abstract=2070290>.

Lee, C.-Y., & Johnson, A. L. (2011). A decomposition of productivity change in the semiconductor manufacturing industry. *International Journal of Production Research, 49*(16), 4761–4785.

Lee, C.-Y., & Johnson, A. L. (2012b). Two-dimensional efficiency decomposition to measure the demand effect in productivity analysis. *European Journal of Operational Research, 216*(3), 584–593.

Lovell, C. A. K., & Rousem, A. P. B. (2003). Equivalent standard DEA models to provide super-efficiency scores. *Journal of the Operational Research Society, 54*, 101–108.

Luss, H. (1982). Operations research and capacity expansion problems: a survey. *Operations Research, 30*(5), 907–947.

Manne, A. S. (1961). Capacity expansion and probabilistic growth. *Econometrica, 29*, 632–649.

Marshall, A. (1920). *Principles of economics* (8th ed.). London: Macmillan and Co., Ltd.

Nipponia, 2001. *The evolving convenience store* <http://web-japan.org/nipponia/nipponia19/en/feature/feature06.html>.

Olesen, O. B., & Petersen, N. C. (1995). Chance constrained efficiency evaluation. *Management Science, 41*, 442–457.

Podinovski, V. V., & Førsund, F. R. (2010). Differential characteristics of efficient frontiers in data envelopment analysis. *Operations Research, 58*(6), 1743–1754.

Premachandra, I. M., Powell, J. G., & Watson, J. (2000). A simulation approach for stochastic data envelopment analysis. *Information and Management Sciences, 11*(1), 11–31.

Ruggiero, J. (2004). Data envelopment analysis with stochastic data. *Journal of the Operational Research Society, 55*, 1008–1012.

Sengupta, J. K. (1987). Data envelopment analysis for efficiency measurement in the stochastic case. *Computers and Operations Research, 14*(2), 117–129.

Sengupta, J. K. (2000). Efficiency analysis by stochastic data envelopment analysis. *Applied Economics Letters, 7*, 379–383.

Stigler, G. (1939). Production and distribution in the short run. *The Journal of Political Economy, 47*(3), 305–327.

Sueyoshi, T. (2003). Production frontier analysis. In C. Kao, S.-N. Hwang, & T. Sueyoshi (Eds.), *Management performance evaluation: data envelopment analysis* (pp. 97–117). Taipei: HwaTai Publishing Co.

Triantis, K., & Girod, O. (1998). A mathematical programming approach for measuring technical efficiency in a fuzzy environment. *Journal of Productivity Analysis, 10*(1), 85–102.

Wilson, T. A., & Eckstein, O. (1964). Short-run productivity behavior in US manufacturing. *The Review of Economics and Statistics, 46*(1), 41–54.