Capacity Dilemma: Economic Scale Size versus Demand Fulfillment

Chia-Yen Lee (cylee@mail.ncku.edu.tw)
Institute of Manufacturing Information and Systems,
National Cheng Kung University

Abstract
A firm tends to associate capacity planning with economic scale size and demand fulfillment for profit maximization. However, it is troublesome capacity dilemmas to achieve both of them simultaneously in stochastic environment. We propose a multi-objective stochastic programming with data envelopment analysis (DEA) constraints to find a compromise efficient benchmark.

Keywords: Data envelopment analysis, Most productive scale size, Stochastic programming

Introduction

Capacity planning plays an important role in a firm's strategy development and core competence enhancement. To determine potential cost advantages, microeconomics tends to use economies of scale that reduce the cost per output, since the fixed costs are spread out over more units of output when increasing scale size. Banker (1984) defined MPSS as the production scale size at which the average productivity of a production unit is maximized. Demand fluctuations produce a mismatch between capacity level and realized demand. Clearly, demand fulfillment leads to revenue maximization since selling more products both creates profits and conserves resources. In particular, demand forecasting and scale size decisions provide some guidelines to employment planning and inventory management. These two issues motivate this study.

Thus, a firm on one hand would like to identify the economic scale size for cost minimization per unit of product but on the other would like to satisfy the demand level for maximizing revenue. However, in practice, a firm may not achieve these two objectives simultaneously and faces a capacity dilemma. Figure 1 illustrates Firm A's capacity dilemma on a single-input and single-output production function. The S-shaped curve describes the production function and the dashed line is the constant returns-to-scale frontier that identifies the MPSS benchmark on the production function (Banker, 1984). D_A indicates Firm A's demand forecast and truncates the production function by the point D^F (i.e. demand fulfillment). Located below the production function, Firm A represents inefficiency. Facing a capacity dilemma between MPSS and demand fulfillment, Firm A needs to formulate the multi-objective decision analysis (MODA) problem. Based on the forecasts and estimates of scale size, cost structure, and expected revenue a firm need to move towards the target which shows a tradeoff between MPSS benchmark and demand fulfillment D^F.

This study makes three contributions to the literature. First, we push a typical "ex-post" data envelopment analysis (DEA) study of production function estimation towards an "ex-ante" DEA analysis for capacity planning. In fact, DEA is not only a method for estimating productive efficiency but also an approach finding direction for productivity improvement. We focus on how to set a target on the production function and move towards it can guide capacity planning. Second, we address the capacity dilemma between the MPSS and demand fulfillment by using MODA to develop a compromise solution. The trade-off between an MPSS (cost-oriented) strategy and a demand-chasing (revenue-oriented) strategy shows the risk preference of the decision-maker. Third, the solution comparison of MMR model and SP technique when addressing decision under strict uncertainty supplies useful managerial insights.

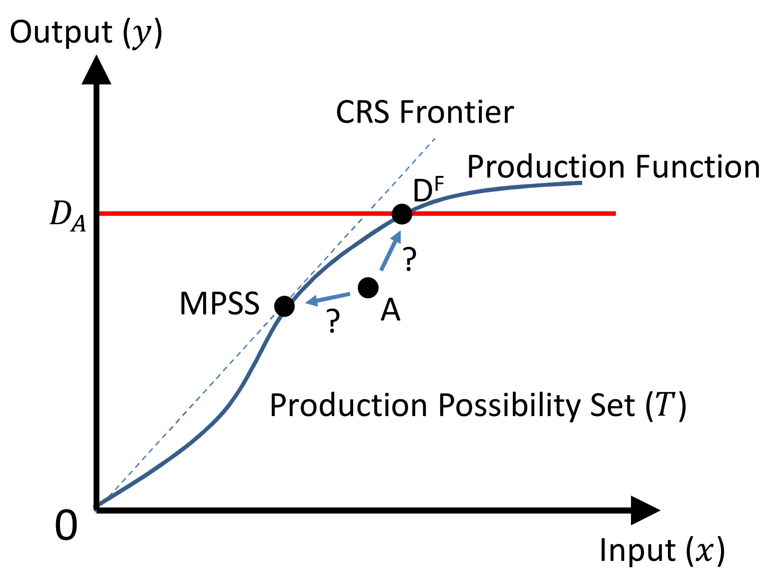

Figure 1 – Capacity dilemma

Most Productive Scale Size

An empirical production function characterizing production possibility set (PPS) can be estimated by data envelopment analysis (DEA) (Banker et al., 1984). Let $x \in \mathbb{R}_{+}^{I}$ denote the input vector and $y \in \mathbb{R}_{+}^{J}$ denote the output vector of the production system. Define the production possibility set as $T = \{(x, y): x \text{ can produce } y\}$ and estimate it by a piece-wise linear concave function enveloping all observations shown in (1), i.e. DEA formulation. Let $i = \{1,2, \dots, I\}$ be the indexed set of inputs, $j = \{1,2, \dots, J\}$ be the indexed set of output, and $k = \{1,2, \dots, K\}$ be the indexed set of firms. Index r is an alias of index k for a specific firm. X_{ik} is the i^{th} input resource, Y_{jk} is the amount of the j^{th} production output, and λ_k is the convex-combination multiplier of the k^{th} firm. The model (1) defines the empirically estimated variable-returns-to-scale (VRS) production possibility set $\tilde{T}^{VRS}$.

$$\tilde{T}^{VRS} = \{(x, y) | \textstyle\sum_{k=1}^{K} \lambda_k Y_{jk} \geq y_j, \forall j; \sum_{k=1}^{K} \lambda_k X_{ik} \leq x_i, \forall i; \sum_{k=1}^{K} \lambda_k = 1; \lambda_k \geq 0, \forall k \quad (1)$$

Measure the efficiency θ^{VRS} using the DEA estimator. For a specific firm r, we can measure the VRS input-oriented technical efficiency (ITE) θ_r^{VRS} by $D_I^{VRS}(x_r, y_r) = \sup\{\theta_r^{VRS} | (\theta_r^{VRS} x_r, y_r) \in \tilde{T}^{VRS}\}$, where $\theta_r^{VRS} \leq 1$, and a firm is technically efficient if $\theta_r^{VRS} = 1$.

Banker (1984) shows that MPSS is equivalent to the efficient benchmark on a constant-returns-to-scale (CRS) frontier. The idea of returns-to-scale is directly related to the estimation of MPSS since returns-to-scale illustrates the change of marginal product (MP). Banker (1984) claimed that this efficiency measure with respect to CRS frontier not only captures the technically inefficiency of a firm, but also any inefficiency due to the difference of its actual scale size from MPSS. Model (2) defines the empirically estimated CRS production possibility set $\tilde{T}^{CRS}$.

$$\tilde{T}^{CRS} = \{(x, y) | \textstyle\sum_{k=1}^{K} \lambda_k Y_{jk} \geq y_j, \forall j; \sum_{k=1}^{K} \lambda_k X_{ik} \leq x_i, \forall i; \lambda_k \geq 0, \forall k\} \quad (2)$$

Measure the efficiency θ^{CRS} using the DEA estimator. For a specific firm r, we can measure the CRS input-oriented overall efficiency (IOE) θ_r^{CRS} by $D_I^{CRS}(x_r, y_r) = \sup\{\theta_r^{CRS} | (\theta_r^{CRS} x_r, y_r) \in \tilde{T}^{CRS}\}$, where $\theta_r^{CRS} \leq 1$, and a firm is overall efficient if $\theta_r^{CRS} = 1$.

To identify the MPSS, based on VRS efficiency measure, project all of the inefficient firms to their efficient benchmarks on the VRS frontier via efficiency measure θ^{VRS}. That is, denote the efficient benchmarks by $(\theta_r^{VRS*} x_r, y_r)$, where θ_r^{VRS*} is the optimal solution generated from $D_I^{VRS}(x_r, y_r)$. Next, let these efficient benchmarks be new observations and estimate their CRS efficiency measure θ^{CRS}. Thus, we can use optimal solution λ_k^* of $D_I^{CRS}(\theta_r^{VRS*}x_r, y_r)$ to identify the MPSS if $\sum_{k=1}^{K}\lambda_k^* = 1$.

Figure 2 for a two-input and one-output case. For illustrative purposes, the VRS DEA frontier in Figure 2 does not need to be generated from origin due to DEA's free disposability property (also called strong disposability) of outputs.

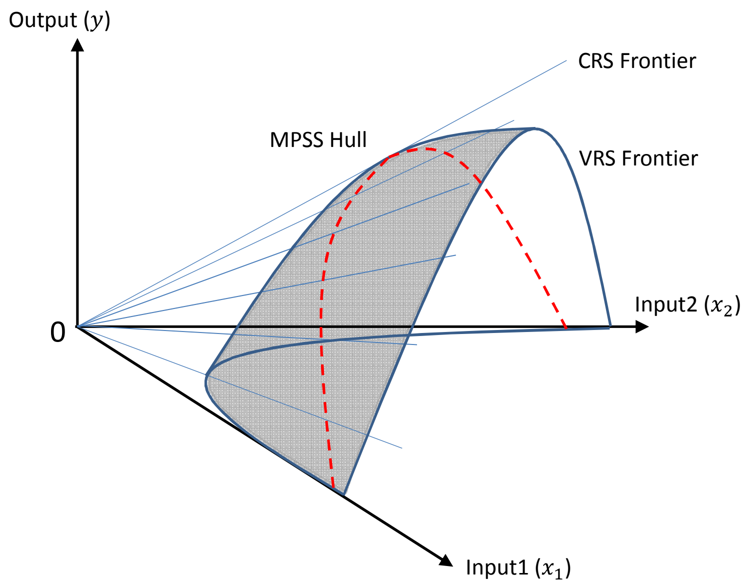

Figure 2 – MPSS on VRS frontier

Compromise Target between MPSS and Demand

This section presents a compromise solution according to MPSS and demand forecast. For simplicity, we limit discussion to the production function with multiple inputs and single output and apply the proposed solution to three separate cases (Case1: Demand level less than or equal to $\text{MPSS}^{\max}$; Case2: Demand level larger than $\text{MPSS}^{\max}$ and smaller than $Y^{\max}$; Case3: Demand level larger than $Y^{\max}$). The notations and definitions are described.

Let $(X_{ik'}^{MPSS}, Y_{k'}^{MPSS})$ be MPSS observation k' on frontier, where $k' = \{1,2,\dots,K'\}$. Denote $\text{MPSS}^{\max}$ as peak output of MPSS hull and define $\text{MPSS}^{\max} = \max\{y|y \in \text{MPSS Hull}\}$ and $\text{MPSS}^{\max} = (X_i^M, Y^M)$. Model (3) calculates $\text{MPSS}^{\max}$.

$$(X_i^M, Y^M) = \text{argMax}\left\{ y \left| \begin{array}{c} \sum_{k'=1}^{K'} \lambda_{k'} X_{ik'}^{MPSS} = x_i, \forall i \\ \sum_{k'=1}^{K'} \lambda_{k'} Y_{k'}^{MPSS} \geq y; \\ \sum_{k'=1}^{K'} \lambda_{k'} = 1; \\ \lambda_{k'} \geq 0, \forall k';\ x_i \geq 0, \forall i;\ y \geq 0 \end{array} \right. \right\} \qquad (3)$$

Note that there are potential multiple MPSS^{max} benchmarks; however, they do not affect our illustration. If the price of input is available, let c_i be the cost of the input resource i, and use model (3) to obtain an unique MPSS^{max} when replacing the objective function of model (3) by Max $My - \sum_i c_i x_i$, where M is a large positive number.

Denote Y^{max} as the peak output of the production function. Note that the peak output is limited since PPS is estimated by DEA. Let Y^{max} = (X_i^P, Y^P). Then calculate the Y^{max} using model (4).

$$\left(X_i^P, Y^P\right) = \text{argMax}\left\{ y \left| \begin{array}{c} \sum_{k=1}^{K} \lambda_k X_{ik} = x_i, \forall i; \\ \sum_{k=1}^{K} \lambda_k Y_k \geq y; \\ \sum_{k=1}^{K} \lambda_k = 1; \\ \lambda_k \geq 0, \forall k;\ x_i \geq 0, \forall i;\ y \geq 0 \end{array} \right. \right\} \tag{4}$$

Similarly, obtain a unique solution by replacing objective function with Max $My - \sum_i c_i x_i$. The unique solution of Y^{max} is an anchor point in DEA.

Based on MPSS^{max} and Y^{max}, separate the demand scenario into three categories for setting compromise targets between MPSS and demand fulfillment, respectively. Figure 3, which shows a one-input and one-output case, illustrates demand scenario regarding the three cases.

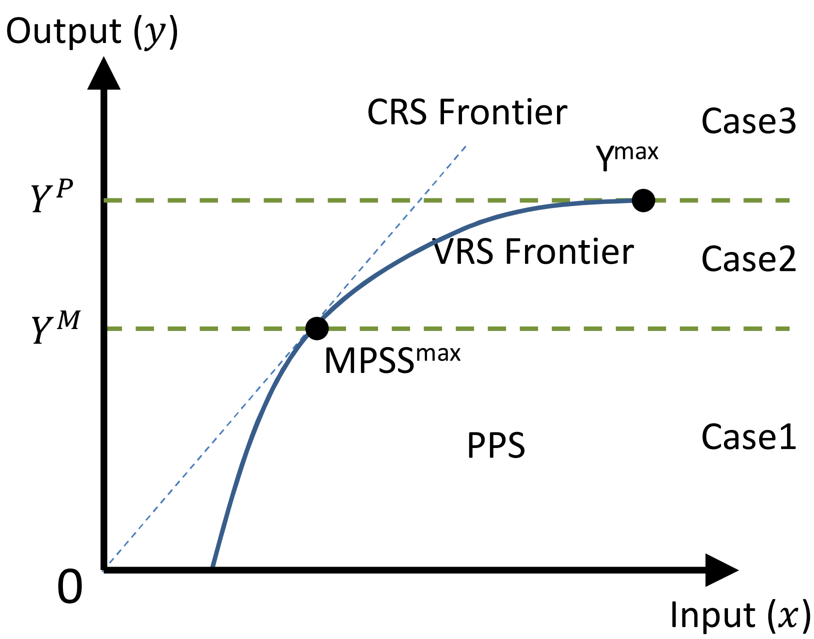

Figure 3 – Demand scenario with respect to three cases

Aggregated Target under Demand Uncertainty

This section proposes three models to address the compromise target in an environment characterized by highly volatile demand. In practice, however, continuous distribution of demand is difficult to build, and therefore we present an alternative approach, i.e. decision under strict uncertainty (i.e. no information about distribution): minimax regret approach (MMR) and the principle of indifference with respect to discrete demand scenario. For the three proposed models described below, $\text{Target}_s = (X_{is}^T, Y_s^T)$ is the compromise target obtained from Section 3 given one specific demand scenario s, and $\text{ATarget} = (X_i^{AT}, Y^{AT})$ is the aggregated compromise target under demand uncertainty. The widely used minimax regret approach (MMR) (Savage, 1951) addresses decision under strict uncertainty by quantifying the "regret" when a stochastic event occurs after decision-making. This approach is conservative, e.g., a decision-maker wants to avoid the worst case yet ensure a guaranteed minimum possible payoff. The regret refers to the "distance" from the aggregated target to each target under different demand scenarios. The MMR approach to minimize the maximal distance from aggregated target (X_i^{AT}, Y^{AT}) to each individual target (X_{is}^T, Y_s^T) generated by each corresponding demand scenario, i.e., our MMR model minimizes the maximum regret in terms of all scenarios.

To identify (X_i^{AT}, Y^{AT}) on the VRS frontier, we apply a sign-constrained convex nonparametric least squares (CNLS) approach to estimate the VRS frontier (Lee et al., 2013). In this study, a typical DEA formulation does not successfully support finding the benchmark on the frontier because DEA requires a predetermined orientation and our MMR model minimizes the distance (d) in the objective function rather than an efficiency term (θ_r^{VRS}) in DEA. Therefore, we employ an alternative DEA model, sign-constrained CNLS. Let α_k and β_{ik} be the decision variables of intercept and slope of ith input of firm k with respect to the linear regression, respectively. ε_k is the decision variable representing the inefficiency term. Index h is an alias of index k representing firm. Formulate the sign-constrained CNLS as model (5) and α_k^* and β_{ik}^* are optimal solutions. In model (5), the first constraint represents linear hyperplane, the second constraint impose concavity and the monotonicity on the underlying unknown production function is imposed by the third constraint, and the sign constraint $\varepsilon_k \leq 0$ denotes the inefficiency term.

$$(\alpha_k^*, \beta_{ik}^*) = \text{argMin} \left\{ \sum_k \varepsilon_k^2 \left| \begin{array}{c} Y_k = \alpha_k + \sum_i \beta_{ik} X_{ik} + \varepsilon_k, \forall k; \\ \alpha_k + \sum_i \beta_{ik} X_{ik} \leq \alpha_h + \sum_i \beta_{ih} X_{ik}, \forall k, \forall h \text{ and } k \neq h; \\ \beta_{ik} \geq 0, \forall i, \forall k; \\ \varepsilon_k \leq 0, \forall k; \end{array} \right. \right\} \quad (5)$$

Next, minimize the maximum distance (our MMR strategy) to each target (X_{is}^{T}, Y_{s}^{T}) and find the aggregated target (X_{i}^{AT}, Y^{AT}) by following model (6). If the input price and output price are available, formulate the first constraint $p|Y_{s}^{T} - y| + \sum_{i} c_{i}|X_{is}^{T} - x_{i}| \leq d, \forall s$, and then measure the distance in dollars. Let z_{k} be the binary variable of firm k for choosing one linear hyperplane in CNLS. The model (6) with objective function $Md - (py - \sum_{i} c_{i}x_{i})$ can be formulated to obtain the unique aggregated target.

$$(X_{i}^{AT}, Y^{AT}) = \text{argMin}$$

$$\left\{ Md - (py - \sum_{i} c_{i}x_{i}) \middle| \begin{array}{c} p(y_{s}^{+} + y_{s}^{-}) + \sum_{i} c_{i}(x_{is}^{+} + x_{is}^{-}) \leq d, \forall s; \\ x_{i} = X_{is}^{T} - x_{is}^{+} + x_{is}^{-}, \forall i, \forall s; \\ y = Y_{s}^{T} - y_{s}^{+} + y_{s}^{-}, \forall s; \\ y \leq \alpha_{k}^{*} + \sum_{i} \beta_{ik}^{*} x_{i} + M(1 - z_{k}), \forall k; \\ y \geq \alpha_{k}^{*} + \sum_{i} \beta_{ik}^{*} x_{i} - M(1 - z_{k}), \forall k; \\ \sum_{k} z_{k} = 1; \\ \sum_{k=1}^{K} \lambda_{k} X_{ik} \leq x_{i}, \forall i; \\ \sum_{k=1}^{K} \lambda_{k} Y_{k} \geq y; \\ \sum_{k=1}^{K} \lambda_{k} = 1; \\ z_{k} \in \{0,1\}, \forall k;\ \lambda_{k} \geq 0, \forall k; \\ x_{i}, y, x_{is}^{+}, x_{is}^{-}, y_{s}^{+}, y_{s}^{-} \geq 0, \forall i, \forall s \end{array} \right\} \quad (6)$$

The second proposed model, the "EV model", assumes that all demands are equally likely to occur only when there is no knowledge indicating unequal probabilities, and calculates the expected value of demand scenarios with equal probability. Let D_{rs} be the demand of scenario s of rth firm. Define the expected value of demand defined as $\bar{D}_{r} = \frac{1}{S}\sum_{s=1}^{S} D_{rs},\ \forall j$. Using the case identification described in the previous section, apply the EV model to generate the compromise target with respect to this expected-value scenario.

The third proposed model, the "SP model", applies stochastic programming (SP) with two-stage recourse problem to set the aggregated target (X_{i}^{AT}, Y^{AT}). The first-stage decision, i.e. the here-and-now, corresponds to selecting the aggregated target based on the demand forecasts. After demand is realized, the second-stage decision, i.e. the wait-and-see, corresponds to the minimization of the expected recourse function (i.e. expected regret) with respect to all scenarios with equal probability. Replace the decision variable d by d_{s} associated with each scenario s and change the objective function as $M(\frac{1}{S}\sum_{s} d_{s}) - (py - \sum_{i} c_{i}x_{i})$ in model (6).

Computational Study

This section describes the experimental setup and the performance of the three solution approaches addressing uncertainty. Firm-level data motivates this study. We introduce sufficient variation in the data sets to evaluate the solution approaches under different conditions and to study the impact of the problem parameters on capacity decisions. The sets of values for the various parameters in the model are as follows.

‘

Production function. We use the data generating process (DGP) for the two inputs and single output production function, and generate 50 efficient observations on the frontier given by $Y_k = (X_{1k})^{0.4}(X_{2k})^{0.4}$, where X_{1k} and X_{2k} independently follow uniform distribution $Uniform(1,10)$.

Price of input and outputs. We assume price availability inputs and outputs for allocatively efficient benchmark and uniqueness. Let $p = 15$, $c_1 = 6$, and $c_2 = 4$.

Inefficient firm r. We assume one inefficient firm r for finding the compromise target and randomly generate it as $(X_{1r}, X_{1r}, Y_r\ ) = (4.836, 6.315, 3.638)$.

We randomly generate 50 VRS efficient observations via DGP and find that 7 observations are MPSS. In particular, the peak output of MPSS hull is $\mathrm{MPSS}^{\max}$= $(X_1^M, X_2^M, Y^M) = (2.002, 6.107, 2.722)$. The peak output of production function is $\mathrm{Y}^{\max} = (X_1^P, X_2^P, Y^P) = (8.979, 8.868, 5.760)$. For simplicity, we discuss three demand forecasts generated from $Normal(Y_r, 1)$ to represent the Case1 with demand level 2.479, Case2 with demand 4.075, and Case3 with demand 6.027.

For the first demand scenario with lower demand level (i.e. Case1), we identify MPSS-DF as MPSS-DF= $(2.007, 4.839, 2.479)$. Thus, the inefficient firm can drive productivity towards MPSS-DF without a tradeoff between MPSS and demand.

For the second demand scenario with larger demand level than $\mathrm{MPSS}^{\max}$ and less than $\mathrm{Y}^{\max}$ (i.e. Case2), we identify the DF-$\mathrm{cMPSS}^{\max} = (X_1^D, X_2^D, Y^D) = (4.640, 7.226, 4.075)$. Given different tradeoff parameter w with bin size equal to 0.1, we formulate a convex combination between $\mathrm{MPSS}^{\max}$ and DF-$\mathrm{cMPSS}^{\max}$, and calculate ConMD as in Table 1. The target approximate to profit maximization is (3.433, 6.834, 3.521) with profit 4.882.

For the third demand scenario, with larger demand level than $\mathrm{Y}^{\max}$ (i.e. Case3), we formulate a convex combination between $\mathrm{MPSS}^{\max}$ and $\mathrm{Y}^{\max}$, and calculate ConMY as in Table 2. The target approximate to profit maximization benchmark is $\mathrm{MPSS}^{\max}$ = (3.433, 6.834, 3.521) with profit 4.393. The result shows that it is not worthy to catch the higher demand level when too far from MPSS by considering the

DMR effect of the production function.

Table 1 – Tradeoff between MPSS^{max} and DF-cMPSS^{max}

w	ConMD= $\left(X_i^{MD}, Y^{MD}\right)$	$\hat{\theta}^*$	Target $= (X_1, X_2, Y)$	Profit
0	DF-cMPSS^{max} (4.640, 7.226, 4.075)	1	(4.640, 7.226, 4.075)	4.38
0.1	(4.376, 7.114, 3.940)	1.034	(4.360, 7.141, 3.950)	4.522
0.2	(4.112, 7.002, 3.804)	1.063	(4.067, 7.045, 3.815)	4.639
0.3	(3.848, 6.890, 3.669)	1.092	(3.757, 6.943, 3.672)	4.762
0.4	**(3.585, 6.778, 3.534)**	**1.121**	**(3.433, 6.834, 3.521)**	**4.882**
0.5	(3.321, 6.667, 3.398)	1.115	(3.146, 6.707, 3.371)	4.859
0.6	(3.057, 6.554, 3.263)	1.097	(2.884, 6.578, 3.227)	4.786
0.7	(2.793, 6.442, 3.128)	1.080	(2.631, 6.453, 3.087)	4.715
0.8	(2.529, 6.330, 2.993)	1.062	(2.386, 6.331, 2.952)	4.646
0.9	(2.266, 6.219, 2.857)	1.030	(2.188, 6.216, 2.834)	4.516
1.0	MPSS^{max} (2.002, 6.107, 2.722)	1	(2.002, 6.107, 2.722)	4.393

Table 2 – Tradeoff between MPSS^{max} and Y^{max}

w	ConMY= $\left(X_i^{MY}, Y^{MY}\right)$	$\hat{\theta}^*$	Target $= (X_1, X_2, Y)$	Profit
0	Y^{max} (8.979, 8.868, 5.760)	1	(8.979, 8.868, 5.760)	-2.946
0.1	(8.281, 8.592, 5.456)	1.142	(8.770, 8.916, 5.714)	-2.568
0.2	(7.583, 8.316, 5.152	1.331	(8.493, 8.979, 5.654)	-2.064
0.3	(6.886, 8.040, 4.849)	1.595	(8.106, 9.067, 5.569)	-1.361
0.4	(6.188, 7.764, 4.545)	1.968	(7.497, 9.166, 5.423)	-0.303
0.5	(5.490, 7.487, 4.241)	2.328	(6.360, 9.045, 5.042)	1.296
0.6	(4.793, 7.211, 3.937)	2.259	(4.738, 8.340, 4.314)	2.924
0.7	(4.095, 6.935, 3.633)	1.973	(3.374, 7.539, 3.629)	4.039
0.8	(3.397, 6.659, 3.330)	1.531	(2.633, 6.842, 3.166)	4.324
0.9	(2.700, 6.383, 3.026)	1.214	(2.243, 6.397, 2.895)	4.378
1.0	**MPSS^{max} (2.002, 6.107, 2.722)**	**1**	**(2.002, 6.107, 2.722)**	**4.393**

Again, we use the three scenarios (i.e. demand forecasts) above and use the three proposed models to calculate the aggregated target under demand uncertainty. In the MMR model, we define regret as the distance measure from the aggregated target to each individual target found by each demand scenario separately. MMR shows the aggregated target is $\left(X_1^{AT}, X_2^{AT}, Y^{AT}\right) =$ (3.210, 5.021, 3.021) using model (11). In the EV model, based on the principle of indifference, the expected demand $\overline{D}_r =$ 4.193 corresponds to Case2. Thus, we calculate the DF-cMPSS^{max} $= (X_1^D, X_2^D, Y^D) =$ $(4.826, 7.476, 4.193)$, and then the aggregated target is (3.216, 6.858, 3.433). In the SP model, the two-stage recourse problem generates the aggregated target (2.002, 6.107, 2.722) as MPSS^{max}.

To summarize, the MMR model assessing the regret and minimizing the maximal regret leads to the worst case analysis, where the target is generated by an outlier demand scenario. Thus, the MMR model is sensitive to the outlier and provides a conservative target. In the EV model, the principle of indifference supports the equal probability of each scenario and the EV model calculates the expected value of demand scenarios as only one dummy scenario. While the method is straightforward, it may smooth out the outlier and eliminate the worst-case effect. The SP model with two-stage resource problem considers all demand scenarios including the outlier effect. Because it minimizes the expected recourse function, the SP model provides the most robust solution like $MPSS^{max}$ to address demand fluctuation. As mentioned, in the long run, $MPSS^{max}$ provides a cost-minimization benchmark addressing uncertainty due to an economic scale size. The result of SP model gives insights to justify that capacity installation should refer to economic scale size to address future uncertainty.

Conclusion

This study described a capacity dilemma- MPSS versus demand fulfillment. A multi-objective decision analysis (MODA) is required in the case of a firm that on one hand would like to identify the economic scale size for cost minimization per unit of product but on the other would like to satisfy the demand level for minimizing inventory loss (i.e. capacity surplus) or maximizing profit by reducing loss of sales (i.e. capacity shortage). Models were proposed to identify the MPSS benchmark and the benchmark close to demand forecast on the production function. The target set by firm provided the tradeoffs between MPSS and demand fulfillment.

Acknowledgments

This research was funded by Ministry of Science and Technology, Taiwan (MOST103-2221-E-006-122-MY3).

Reference

1. Banker, R. D. 1984. Estimating most productive scale size using data envelopment analysis. *European Journal of Operational Research* **17**(1): 35–44.
2. Savage, L. J. 1951. The theory of statistical decision. *Journal of the American Statistical Association* **46**: 55–67.
3. Lee, C.-Y., A. L. Johnson, E. Moreno-Centeno, T. Kuosmanen. 2013. A more efficient algorithm for convex nonparametric least squares. *European Journal of Operational Research* **227**(2): 391–400.