Efficiency Predictions by Fuzzy Piecewise Auto-regression

Bo Hsiao
Department of Information Management, National Taiwan University

Ching-Chin Chern
Department of Information Management, National Taiwan University

Ming-Miin Yu
Department of Transportation Science, National Taiwan Ocean University

Gwo-Hshiung Tzeng
Institute of Management of Technology, National Chiao Tung University

Abstract

Predicting productivity and efficiency during the transfer of input to output is a key issue in many manufacturing and service operation applications. Operation research and econometrics acknowledge that efficiency analysis is a major research issue. Data envelopment analysis (DEA) has substantially reshaped the result of information economics in previous years. However, the extent to which asymmetric information is relevant for efficiency prediction has rarely been sought empirically. Efficiency prediction plays a crucial role in many state-of-the-art applications and planning. Forecasting methodologies that can accurately predict efficiency scores can help in strategic decision-making. This study focuses on fuzzy piecewise auto-regression and the catching-up efficiency index (CIE), which supports efficiency prediction. In this study, two regression models were formulated by utilizing data from commercial banks in Taiwan from 2002 to 2005. These models were used to validate banking efficiency scores for 2005 and 2006, and to predict the banking efficiency scores for 2007. The results of a thorough computational analysis provide a range indicating the prediction value for each bank based on fuzzy regression characteristics.

Key words : Data Eenvelopment Analysis (DEA), Fuzzy Piecewise Auto-Regression, Catching-up Efficiency Index (CIE), Banking Performances.

應用模糊分段自我迴歸於效率預測

蕭鉢
國立台灣大學資訊管理學系

陳靜枝
國立台灣大學資訊管理學系

游明敏
國立台灣海洋大學運輸科學系

曾國雄
國立交通大學科技管理研究所

摘要

轉換為輸入與輸出關係來預測生產力與效率已廣範應用於許多生產與服務作業管理領域，因此，作業研究與計量經濟等領域普遍認為效率分析與預測為其主要的研究議題。在過去幾年，當資料包絡分析已經成功形塑資訊經濟議題時，效率預測卻顯少有學者進行研究。效率預測在當今許多規畫應用中扮演著關鍵角色。一個好的預測方法可以有效輔助決策者做效率預測。因此，不同於過去傳統效率預測方法，本研究採用模糊分段自我迴歸方法與效率追趕指標去處理效率預測問題。藉由可能性與必然性回歸模式應用於台灣22 家商業銀行效率預測，透過 2002 到 2005 年三期資料去驗證2005年到2006年效率值，並且以此為基礎去預測 2007 年各銀行的效率。基於過去各商業銀行的績效表現，透過這個分析架構，我們可以提供給各銀行未來效率值可能區間。經由指示性研究發現與比較，透過模糊分段自我迴歸與效率追趕指標處理，可以有效預測相對效率議題。

關鍵字：資料包絡分析(DEA)、 模糊分段自我迴歸、效率追趕指標 (CIE)、銀行績效評估

1. INTRODUCTION

Forecast methodologies for efficiency are rarely applied to predict productivity and efficiency for real-world applications, although an analytic framework has been proposed for business functions such as production, marketing, research and development, and finance. Most existing forecast methodologies focused on predicting output from input (Troutt et al. 2005). However, when forecasting methodology is applied to relative efficiency, selection data become more difficult because earlier approaches used absolute historical data or efficiency scores. As such, conventional forecasting approaches cannot be used for relative concepts such as time series of efficiency.

There are two competing paradigms on efficiency analysis. The first uses mathematical programming techniques, such as data envelopment analysis (DEA), which are popular in operation research. The other employs the regression approach, such as stochastic frontier analysis (SFA), which is widely accepted in econometrics. These two methodologies have specific characteristics and limitations. DEA does not require explicit assumptions regarding the function structure of the stochastic frontier. SFA imposes an explicit, possibly overly restrictive, frontier function on the models. DEA is based on non-parametric approaches, while SFA is based on parametric ones. Therefore, DEA, unlike SFA, cannot provide mechanisms for predictions; however, it is difficult to define parametric and frontier functions in SFA.

To accomplish the efficiency prediction objective, a new hybrid approach comprising catching-up efficiency index (CIE) and fuzzy piecewise auto-regression analyses is proposed to predict efficiency and reinforce the prediction ability of DEA. CIE is a measure of technical efficiency change during the analyzed period (catching-up effect or movement toward the frontier). The CIE index ignores input-versus-output relationships. The fuzzy piecewise regression analysis developed by Yu et al. (Yu et al. 1999; Yu et al. 2001) provides information to understand the dynamics of variable data and forecast efficiency when two specific regression estimation models are used simultaneously. A two-stage process is used to predict efficiency. The CIE is calculated with efficiency evaluation in the first stage, while validation and/or prediction are done in the second stage. In the first stage, DEA techniques are used to evaluate the efficiency scores for some periods and transfer efficiency scores to CIE indices. In the second stage, fuzzy piecewise auto-piecewise regression is followed to calculate the CIE index data and forecast the value, which falls into two ranges. The first is the possibility estimation model, which suggests that predicted values should be included in the regression ranges; the second is the necessity estimation model, which proposes that the predicted values should be excluded in the regression ranges.

The rest of the paper is organized as follows. Section 2 reviews the related literature. Section 3 describes the problem and presents the proposed theoretical framework. Section 4 demonstrates the results of our study for the data from 22 Taiwanese commercial banks and offers our conclusions and suggestions for future research. The conclusion and future work are given in Section 5.

2. LITERATURE REVIEW

The DEA approach is suitable for analyzing institutional data such as those from the government (Yunos & Hawdon 1997; Hsu & Hsueh 2009), schools (Soteriou et al. 1998; Tyagi et al. 2009), hospitals (Chilingerian & Sherman 1990; Ozcan 1995), and banks (Chen & Yeh 2000). While it can be suitable in evaluating efficiency, it is not applicable for prediction and/or forecast. Traditional DEA studies focus on "one-shot state" efficiency analyses. Few approaches (e.g., SFA) predict efficiency either by modeling the production relationship or by using soft computing techniques. However, modeling the production/frontier function or framing an analyzed environment structure has many limitations and is difficult to achieve.

2.1 Efficiency Predictions

Generally, econometricians tend to favor regression-based or sophistication approaches; management scientists favor DEA approaches to evaluate performance issues (Thanassoulis 1993; Bowlin et al. 1985; Schmidt 1986; Cubbin & Zamani 1996). Thanassoulis (Thanassoulis 1993) found that DEA is suitable in regression analysis. Meanwhile, Schmidt (Schmidt 1986) proposed a major drawback of DEA is the lack of a statistical basis. It is difficult to decide on either the relevance or the credibility of these conflicting results, and the fundamental difference between regression analysis and DEA is not clearly understood. To help understand DEA characteristics, two main advantages need to be demonstrated: (1) DEA is based on ratio concepts and not on absolute input versus output relationships; and (2) the efficiency score is relative to the frontier, and not on anyone's own scores (Golany & Roll 1989).

Ratio provides scale invariance characteristics such that they can ignore scale influences on the performance result. Therefore, they can be extrapolated for evaluation. Despite some limitations, many techniques, such as key business performance measurements, are applied for ratio analysis to evaluate income and balance sheet financial statements simultaneously. Each projected metric in ratio analysis has its own unique goal value tied to the business strategic vision. For example, financial ratio analysis is used for performance evaluation (Caves et al. 1982; Megginson et al. 1994); it simultaneously measures one input and one output. Challenges in financial ratio analysis include the lack of accredited financial ratio models and weight selection. Therefore, for more flexible analysis, rules on ratio analysis may be constructed using complicated computations by higher-order equation. DEA can work with sample rules (i.e., input/output) and allow for the evaluation of multiple outputs and inputs (Golany & Roll 1989; Caves et al. 1982). Ratio analysis needs complex data requirements to make it suitable for evaluation, unlike in DEA, which does not need a large sample size (Golany & Roll 1989).

Ratio analysis refocuses the resources to converge on "the goal" (i.e., efficiency will be 1) and does not reflect actual scenarios (i.e., compared with other decision-making units or DMUs). DEA is a relative concept in that specific DMU efficiency is dependent on best practices or frontiers, not on itself. As such, conventional evaluation techniques cannot fully fit the requirement from the inherent characteristics of DEA. Under such conditions, if conventional approaches need to be implemented, they may need to combine many relationship constraints to satisfy such characteristics.

DEA can simultaneously deal with radio and ordinal scale data, but regression analysis is hard to implement. DEA approaches lack requirements on assumptions of any pre-specified functional form of the production function, and tend to avoid the problem of parameter measures (Golany & Roll 1989).

Unfortunately, these advantages also cause disadvantages on the lack of frontier functions of DEA. The absence of requirements on assumptions of any pre-specified functional form of the production/frontier function implies that a major drawback of DEA in forecasting stems from the lack of prediction capability. This is apparent in other mathematical models, such as in regression analysis and prediction approaches. Models should be able to estimate efficiency predictions over time. The efficiency prediction of DEA does not have such capability because the DEA model cannot simultaneously handle both negative values (e.g., data representing decay) in the data set and a frontier shift over time (Cook & Seiford 2009).

Some studies enhanced the predicting efficiency of DEA by combining it with other predicting techniques. Productivity change, explained in terms of technical change recently became widely accepted in predicting efficiency change. It can be simplified, to some degree, to become an uncomplicated forecasting functionality (frontier shift). The Malmquist index, which is used to predict productivity change, plays an import role in supporting such discussion; it was first introduced by Caves et al. (Caves et al. 1982) in productivity change. Färe et al. (Färe et al. 1989) decomposed productivity change into efficiency change and technical change, and constructed a non-parametric mathematical programming model to arrive at a solution. Caves et al. (Caves et al. 1982) and Färe et al. (Färe et al. 1989) showed that under certain conditions, the Malmquist index approximated the Törnqvist (Törnqvist 1936) and Fisher indices (Caves et al. 1982; Fisher 1922), which are easy to compute and are generally accurate, but may be biased to estimate productivity in the presence of inefficiency (Coelli et al.1998; Färe et al. 1995). However, the Malmquist index may not provide the complete picture because it only considers productivity change in two periods, although it can be extended to multiple periods through index multiplication. The Malmquist index is based on only two adjacent periods and may ignore past performances over more than two previous periods.

Sueyoshi (Sueyoshi 2000) proposed stochastic DEA, a model formulated by chance constraint programming and estimated using the program evaluation and review technique or critical path method, for a restructuring strategy applied by the Japanese Petroleum Company. He used stochastic efficiency (or "aspiration level") and conventional efficiency (or "risk criterion") to decide future efficiency. However, he made several assumptions on the stochastic variables of output for computation conveniences, and the normal distribution is assumed for a stochastic variable when conducting a statistical test. Moreover, stochastic DEA predicts efficiency based on data from only one previous period; thus, its prediction ability relies heavily on its required assumptions (e.g., error terms standard deviation are equal to zero).

Unlike Sueyoshi (Sueyoshi 2000), Kao and Liu (Kao & Liu 2004) introduced fuzzy concepts to forecast efficiency based on uncertain data represented in a range instead of a single value. The result of the prediction is presented as a range. They adapted fuzzy concepts in DEA, relaxed the assumptions of Sueyoshi (Sueyoshi 2000) on the error term variances of output variables (equal to zero), and assumed the output probability as beta distribution. Meanwhile, similar to the model of Sueyoshi (Sueyoshi 2000), the model of Kao and Liu (Kao & Liu 2004) only considers a single state; it does not base its predictions on past performances of DMU. The model treats uncertainty data evaluation rather than forecasting.

Yeh et al. (Yeh et al. 2010) proposed a novel model to integrate rough set theory (RST) with support vector machine (SVM) techniques to increase the accuracy of predicting business failure. In their model, DEA is employed to evaluate input/output efficiency, remove redundant attributes in an RST approach, and reduce the number of independent variables without losing important information. They used such information as a preprocessor to improve the accuracy of business failure prediction through SVM.

Wu et al. (Wu et al. 2006) integrated DEA and neural networks (NNs) to examine/forecast the relative efficiency of each branch office of Canadian banks. Tsai (Tsai et al. 2009) constructed the consumer loan default prediction model by conducting DEA-discriminate analysis (DA) and NNs. However, their model needs longer computing time and larger computing resources. It classifies data into two patterns during the training phases: good examples (positive data) and failed examples (negative data). These methods provide regression results determined either by structure error minimization (Yeh et al. 2010) or by empirical error minimization (Wu et al. 2006; Tsai et al. 2009). However, if a specific DMU outperforms its previous performances, this specific DMU may be viewed as an outlier, and its performance will be ignored in these two models.

Edirisinghe and Zhang (Edirisinghe & Zhang 2007) proposed a complicated multi-step heuristic algorithm with random sampling and local search. It automatically selects a combination of inputs and outputs, in which the emerging DEA measure of financial strength is maximally correlated with stock performance. They generated a relative financial strength indicator, and demonstrated it to be predictive of stock returns. The major contribution of their method is flexibility and automation in selecting input and output parameters to maximize the predictive ability of emerging DEA estimation on stock performance. Although this approach uncovered the black box of the forecasting mechanize (e.g., NNs), it is difficult for this approach to decide the "suitable" solution beforehand; it easily falls into the local solution.

Efficiency evaluation through DEA has been widely applied in numerous empirical cases. However, it does not determine the extent to which asymmetric information is relevant for efficiency prediction, which has been rarely questioned empirically. Previous approaches in efficiency prediction did not account for the appropriate forecasting method and prediction variables, and consequently suffered from influences of variable variance (e.g., Sueyoshi 2000), computing resource/efficiency (e.g., Edirisinghe & Zhang 2007; Yeh et al. 2010), and data challenge (e.g., Malmquist index extending to forecasting problem). This study proposes a model to solve efficiency prediction by using fuzzy piecewise auto-regression and the catching-up index, as developed by Yu et al. (Yu et al. 1999; Yu et al. 2001). The efficiency forecasting of the commercial bank was also reviewed, after which pertinent inputs and outputs were applied in our study.

2.2 Commercial Bank Evaluation

The banking sector, based on the applications of Miller and Noulas (1996), is regarded as an intermediary to bank transfers or deposits, even in the investment market. This approach reflects the way of evaluating the efficiency of commercial banks, which takes commercial banks as entities that use labor and capital to transform deposits into loans and securities. For the intermediation approach, three inputs and outputs each are chosen for each commercial bank. We used data such as amount of money deposited, employment expenditures, and banking assets as inputs. The amount of loans, investments, and commission revenue of loans were used as outputs to evaluate the performances of commercial banks. Table 1 summarizes the measures used as inputs and outputs of the application and reference papers.

Table 1: Measures of Inputs and Outputs of the Taiwan Commercial Banks

Dimension	Measures	References
Inputs	Amount of money deposited	Yue(1992) ;Lin(2002) ; Tortosa-Ausina (2003)
	Employment expenditures	Ausina (2003) Barr et al.(2002);Haslem et al.(1999); Grabowski et al.(1994)
	Banking assets	Grabowski et al.(1994) ;Lin(2002) Barr et al.(2002)
Outputs	Amount of loans	Yue(1992);Lin(2002);Kao et al. (2004) Tortosa-Ausina (2003); Grabowski et al.(1994)
	Investments	Miller & Noulas (1996);Haslem et al.(1999)
	Commission revenue of loans	Yue(1992) ;Kao et al. (2004) Barr et al.(2002); Tortosa-Ausina (2003); Grabowski et al.(1994)

3. MODELING AND FORMULATION

This study proposes a forecasting method comprising fuzzy piecewise auto-regression and CIE to predict efficiency and provide help in strategic decision making. This section introduces the modeling concepts used for efficiency prediction/forecasting, including those focusing on fuzzy piecewise regression and catching-up index.

The methodologies are described as follows. First, any measurement technique of DEA can evaluate the efficiency performance of each DMU in each period. As such, the efficiency score of each DMU at each period will be computed. To calculate the efficiency score improvement or decay, we determine the catching-up index of two adjacent periods. If the CIE of a specific DMU is larger than 1, it represents the improvement of the specific DMU’ s efficiency at the calculation period compared with the base period. Otherwise, it represents the decay of the specific DMU efficiency performance. Afterward, the CIE of each DMU in these periods will be forecast based on past CIE efficiency performance. These CIE data sets will be the inputs of the fuzzy piecewise auto-regression. Fuzzy piecewise auto-regression will find two ranges for the future forecast by using two specific regression models for each DMU. The possibility estimation model suggests that the predicted values should be included in the regression ranges. The necessity estimation model proposes that predicted values should be excluded in the regression ranges. After calculating these two ranges from these two regression models, we can obtain four CIE coefficients within the two ranges for each DMU. Using the four coefficients of each DMU, we can forecast the efficiency performance for each DMU for future periods.

3.1 Notations

Following the previous description, the proposed approach can be implemented in four phases. In the first phase, DEA is used to evaluate the efficiency score (e.g., to vest amount of deposits (input) to produce investment (output)) of each DMU in each period. The second phase involves applying the efficiency score of each DMU to calculate the catching-up index. The third phase builds two regression models using fuzzy piecewise/non-piecewise auto-regression, while the fourth phase focuses on validating the calculations and forecasting, which were applied in the regression model of the third phase. Table 2 shows the notations used in the proposed model.

Table 2: Description of Notations

Variable /Notation	Definition/Item
N	Number of DMUs.
T	Number of periods.
P	Number of change points.
n_a	Number of input variables.
n_b	Number of output variables.
$X_{aj}^{(t)}$	Vector of *a-th* specific input variables of the *j-th* DMU at *t-th* period.
$Y_{bj}^{(t)}$	The *b-th* specific outputs variables of the *j-th* DMU at all *t-th* period.
$j\ (j=1,\ldots,N)$	Indexes for DMUs.
$k\ (k=1,\ldots,N)$	Indexes for DMUs.
$t\ (t=1,..,T)$	Indexes for periods.
$a\ (a=1,..,n_a)$	Indexes for input variables.
$b\ (b=1,..,n_b)$	Indexes for output variables.
$p\ (p=1,..,P)$	Indexes for change points.
$\theta_j^{(t)}$	Input efficiency scores of *j-th* DMU at *t-th* period.
$\pi_j^{(t)}$	Output efficiency scores of *j-th* DMU at *t-th* period.
$p_j^{(t)}$	Overall efficiency scores of *j-th* DMU at *t-th* period.
$\lambda_j^{(t)}$	Vector for projecting *j-th* DMU at *t-th* period.
$\delta_j^{t,t\text{-}1}$	Catching-up index s of *j-th* DMU at *t-th* period and *t*-1 period.
ρ_j^U	The upper bound of possibility regression prediction CIE values of *j-th* DMU after fuzzy piecewise regression is completed.
ρ_j^L	The lower bound of possibility regression prediction CIE values of *j-th* DMU after fuzzy piecewise regression is completed.
π_j^U	The upper bound of necessity regression prediction CIE values of *j-th* DMU after fuzzy piecewise regression is completed.
π_j^L	The lower bound of necessity regression prediction CIE values of *j-th* DMU after fuzzy piecewise regression is completed.
$\xi_{j,t}^{(U)}$	The upper bound of possibility regression prediction *t-th* period's efficiency values of *j-th* DMU.
$\xi_{j,t}^{(U)}$	The lower bound of possibility regression prediction period's *t-th* efficiency values of *j-th* DMU.
$\psi_{j,t}^{(U)}$	The upper bound of necessity regression prediction period's *t-th* efficiency values of *j-th* DMU.
$\psi_{j,t}^{(L)}$	The lower bound of necessity regression prediction period's *t-th* efficiency values of *j-th* DMU.

3.2 Phase I: Efficiency Evaluations (Graphic Hyperbolic Measure)

The graphic hyperbolic measure for the introduced proposition is illustrated in this study, in which any measurement technique of DEA can evaluate efficiency performance (Zofío & Lovell 2001). This concept was updated in research conducted by Färe et al. (Färe et al., 1989), who introduced the graph hyperbolic measure. Based on the notations in Section 3.1, assuming there are T periods, N DMUs, with each DMU (DMU$_j$, $j \in R_{+}^{N}$) with $X_{aj}^{(t)}$ input ($a \in R_{+}^{n_a}$) in t period ($t \in R_{+}^{T}$) and $Y_{bj}^{(t)}$ output ($b \in R_{+}^{n_b}$) in t period, the technology can then be described generally by the output sets, as shown by Model (1).

$$\begin{aligned} \mathbf{P}^{(t)}(x) = \{(x^{(t)}, y^{(t)}) : &\sum_{j=1}^{N} \lambda_{j}^{(t)} Y_{bj}^{(t)} \geq Y_{br}^{(t)}, \quad b = 1,\ldots,b_b, \\ &\sum_{j=1}^{N} \lambda_{j}^{(t)} X_{aj}^{(t)} \leq X_{ar}^{(t)}, \quad a = 1\ldots,n_a, \\ &\sum_{j=1}^{N} \lambda_{j}^{(t)} = 1, \\ &\lambda_{j}^{(t)} \geq 0, \qquad j = 1,\ldots.,N, t = 1,\ldots,T\} \end{aligned} \tag{1}$$

where ($X_{ar}^{(t)}, Y_{br}^{(t)}$) represents *r-th* DMU, *a-th* input, and *b-th* output in *t-th* period. The value $\lambda_{j}^{(t)}$ is the intensity variable to contract or expand the individually observed activities of *j-th* DMU to construct the convex combinations of the observed input and output in *t-th* period. Model (1) is assumed closed and bounded, satisfying the conditions of the strong disposability of the desirable output and input. The inequality constraints in Model (1) on input and output reflect that output and input are freely disposable. The graphic hyperbolic measure based on the definitions set by Färe et al. (Färe et al. 1989) is depicted in Model (2).

min $\rho_{k}^{(t)}$
subject to.

$$\begin{aligned} &\sum_{j=1}^{N} \lambda_{j}^{(t)} Y_{bj}^{(t)} \geq \frac{Y_{bk}^{(t)}}{\rho_{k}^{(t)}}, \qquad b = 1,\ldots,n_b, \\ &\sum_{j=1}^{N} \lambda_{j}^{(t)} X_{aj}^{(t)} \leq \rho_{k}^{(t)} X_{ak}^{(t)}, \quad a = 1,\ldots,n_a, \\ &\sum_{j=1}^{N} \lambda_{j}^{(t)} = 1, \qquad j = 1,\ldots,N, \\ &\forall \lambda_{j}^{(t)} \geq 0, \qquad t = 1,\ldots,T \end{aligned} \tag{2}$$

If a specific efficient *k-th* DMU in *t-th* period measures its efficiency by Model (2), it follows that inputs and outputs use the same scalar to adjust to the frontier. As such, they form a hyperbolic-adjusting path for evaluation. Chang (Chang 1999) extended the Färe et al. (Färe et al. 1994) model by minimizing the arithmetic mean of proportional reduction in inputs and proportional expansion in outputs. Chang's (Chang 1999) model was thus modified as follows:

$$\min \quad \rho_k^{(t)} = \frac{(\theta_k^{(t)} + \pi_k^{(t)})}{2}$$

subject to.

$$\begin{aligned}
&\sum_{j=1}^{N} \lambda_j^{(t)} Y_{bj}^{(t)} \geq \frac{Y_{bk}^{(t)}}{\theta_k^{(t)}}, && b = 1, \ldots, n_b, \\
&\sum_{j=1}^{N} \lambda_j^{(t)} X_{aj}^{(t)} \leq \pi_k^{(t)} X_{ak}^{(t)}, && a = 1, \ldots, n_a, \\
&\sum_{j=1}^{N} \lambda_j^{(t)} = 1, && j = 1, \ldots, N, \\
&\forall \lambda_j^{(t)} \geq 0, && t = 1, \ldots, P
\end{aligned} \tag{3}$$

where $\pi_k^{(t)}$ represents the *k-th* DMU maximum expansion of outputs in *t-th* period, while $\theta_k^{(t)}$ represents the *k-th* DMU inputs efficiency score in *t-th* period. Then, $\pi_k^{(t)}$ and $\theta_k^{(t)}$ are equal to one; thus, their overall efficiency scores can be efficient in the *t-th* period.

3.3 Phase II: CIE

After Phase I, we can calculate the periods from 1 to *t+1* using Model (3), and use two periods to demonstrate. Two periods, *t* and *t+1*, are defined after Phase I to measure the productivity change of *k-th* DMU. Based on Model (3), two-period efficiency scores of *k-th* DMU, ρ_k^t and ρ_k^{t+1}, can be obtained. We then calculate the catching-up index between periods *t* and *t+1*, as shown in Eq. (4). CIE is the ratio of period t and period *t+1* efficiency scores, indicating a measure of technical efficiency changes for the analyzed periods (catching-up effect or movement toward the frontier). CIE>1 represents efficiency improvement; otherwise, it represents efficiency regression.

$$\delta_k^{t,t+1} = \frac{\rho_k^{(t+1)}}{\rho_k^{(t)}} \tag{4}$$

After Phase II, we have number *T* CIE data. The number of *T-1* data will form independent variables of fuzzy piecewise regression, and the *T-th* data will be dependent variables.

3.4 Phase III: Fuzzy Piecewise/Non-piecewise Regression

Fuzzy regression analysis can be interpreted as an interval estimation of dependent variables (Yu et al. 1999; Tanaka & Ishibuchi 1992; Tanaka 1987; Tanaka et al. 1989; Tanaka & Watada 1988; Huang & Tzeng 2008). Generally, an interval covering all training data is calculated, and a membership function is constructed based on this interval. The effect of a quadratic function is the same as a linear one, and we adopt the linear form instead of quadratic programming (QP) in Phase III to illustrate. After Phase II, let $\delta_k^{t-1,t}$ represent the dependent variables and $\delta_k^{t-2,t-1}, \delta_k^{t-3,t-2}, ..., \delta_k^{1,2}$ the independent variables of a forecasting function for *k-th* DMU. From the data set of period *P*, we have one dependent variable $\delta_k^{t-1,t}$ and *P-2* independent variables. The interval linear regression model for *k-th* DMU with an output (dependent variables) calculated from all given data (independent variables) is represented as Eq. (5):

$$\delta_k^{t-1,t} = A_0 + A_1\delta_k^{t-2,t-1} + ... + A_{P-3}\delta_k^{1,2} \tag{5}$$

where $\delta_k^{t-1,t}$ is *k-th* DMU, the predicted interval corresponding to the input vector ($\delta_k^{t-2,t-1}$, $\delta_k^{t-3,t-2}$,.., $\delta_k^{1,2}$) and t is the index for time (t=1,...,P). Thus, ($\delta_k^{t-2,t-1}$, $\delta_k^{t-3,t-2}$,.., $\delta_k^{1,2}$) is a one-dimensional input vector for *k-th* DMU, representing the CIE of two adjacent periods. An interval defined by the ordered pair in brackets is represented as Eq. (6):

$$A = [a_L, a_R] = [a : a_L < a < a_R] \tag{6}$$

where a_L denotes the left limit and a_R denotes the right limit of A. Interval A is likewise denoted by its center and width (radius) as shown in Eq. (7):

$$A = (a_c, a_w) = \{a : a_c - a_w \le a \le a_c + a_w\} \tag{7}$$

where a_c denotes the center and a_w denotes the width (e.g., radius of $a_w \geq 0$, such as half of the width of A). The linear model of Eq. (5) is shown in Eqs. (8)– (10):

$$\begin{aligned}\delta_k^{t-1,t} &= A_0 + A_1\delta_k^{t-2,t-1} + ... + A_{P-2}\delta_k^{1,2} \\ &= (a_{0c,k}, a_{0w,k}) + (a_{1c,k}, a_{1w,k})\delta_k^{t-2,t-1} + ... + (a_{P-2c,k,} a_{P-2w,k})\delta_k^{1,2} \\ &= (Y_{kc}, Y_{kw})\end{aligned} \tag{8}$$

$$Y_{kc} = a_{0c,k} + a_{1c,k}\delta_k^{t-2,t-1} + ... + a_{P-2c,k}\delta_k^{1,2} \tag{9}$$

$$Y_{kw} = a_{0w,k} + a_{1w,k}\left|\delta_k^{t-2,t-1}\right| + ... + a_{P-2w,k}\left|\delta_k^{1,2}\right| \tag{10}$$

where Y_{kc} represents the center and Y_{kw} the width of predicted interval $\delta_k^{t-1,t}$ of *k-th* DMU. The specific two-estimation models (i.e., possible estimation and necessity model) are considered for the input-output data $(\delta_k^{t-2,t-1}, \delta_k^{t-3,t-2}, ..., \delta_k^{1,2}; \delta_k^{t-1,t})$. First, the possibility estimation model can be represented as Eq. (11),

$$\begin{aligned}(\delta_k^{t-1,t})^* &= A_0^* + A_1^*(\delta_k^{t-2,t-1})^* + ... + A_{P-2}^*(\delta_k^{1,2})^* \\ &= (a_{0c,k}^*, a_{0w,k}^*) + (a_{1c,k}^*, a_{1w,k}^*)\delta_k^{t-2,t-1} + ... + (a_{P-2c,k}^*, a_{P-2w,k}^*)\delta_k^{1,2} \\ &= (Y_{kc}^*, Y_{kw}^*)\end{aligned} \tag{11}$$

which satisfies Model (12) conditions.

$$\delta_k^{t-1,t} \subseteq (\delta_k^{t-1,t})^* \quad , \quad t = 1,..,T \tag{12}$$

The estimated interval $(\delta_k^{t-1,t})^*$ by the possibility model always includes the observed interval $\delta_k^{t-1.t}$. Second, the necessary estimate model can be represented as Eq. (13),

$$\begin{aligned}(\delta_k^{t-1,t})_* &= A_{0*} + A_{1*}\delta_k^{t-2,t-1} + ... + A_{P-2*}\delta_k^{1,2} \\ &= (a_{0c*,k}, a_{0w*,k}) + (a_{1c*,k}, a_{1w*,k})\delta_k^{t-2,t-1} + ... + (a_{P-2c*,k}, a_{P-2w*,k})\delta_k^{1,2} \\ &= (Y_{kc*}, Y_{kw*})\end{aligned} \tag{13}$$

which satisfies Model (14) conditions.

$$(\delta_k^{t-1,t})_* \subseteq \delta_k^{t-1,t} \quad , \quad t = 1,..,T \tag{14}$$

The estimated interval $(\delta_k^{t-1,t})_*$ by necessity should include the observed interval, $\delta_k^{t-1.t}$. The possibility model and necessity model relationship can be expressed as Model (15), as shown in Figure 1.

$$(\delta_k^{t-1,t})_* \subseteq \delta_k^{t-1,t} \subseteq (\delta_k^{t-1,t})^* \tag{15}$$

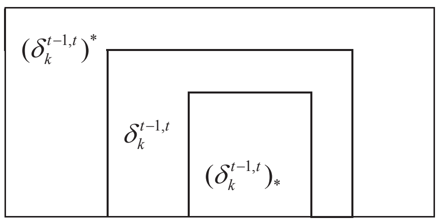

Figure 1: The Possibility and Necessity Model Relationships

The fuzzy regression is then extended to fuzzy piecewise regression and adapted to our framework. Fuzzy piecewise regression analysis was first developed and validated by Yu et al. (Yu et al. 1999; Yu et al. 2001; Yu et al. 2005). We follow the previous descriptions and notations discussed in Section 3.1. A linear programming formulation used to determine the necessity area by the piecewise linear interval model is presented in this subsection for linear piecewise regression, commonly observed in forecasting, as shown in Eq. (16). In contrast, the possibility of the piecewise linear model could be the same as Eq. (16) by substituting B_p^* to B_{p*}.

$$(\delta_k^{t-1,t-2})_* = h(\delta_k^{t-2,t-1}) + \sum_{p=1}^{P-1}\left\{\frac{B_p^*}{2}\left(\left|\delta_k^{t-2-p,t-1-p} - P_p\right| + \delta_k^{t-2-p,t-1-p} - P_p\right)\right\} \tag{16}$$

where $h(\delta_k^{t-2,t-1}) = a_{0*} + a_{1*}\delta_k^{t-2,t-1}$, while B_{p*} is the interval of the necessity model of B_p, and B_p^* is the interval of the necessity model of B_p. $B_p = (B_{pc}, B_{pw})$ represents the center and radius of B_p.

If B_p is a change-point, then

$$\frac{\left(\left|\delta_k^{t-2-p,t-1-p} - P_p\right| + \delta_k^{t-2-p,t-1-p} - P_p\right)}{2} = \begin{cases} \delta_k^{t-2-p,t-1-p} - P_p, & \text{if } \delta_k^{t-2-p,t-1-p} \geq P_p \\ 0 & \text{, if } \delta_k^{t-2-p,t-1-p} < P_p \end{cases} \tag{17}$$

where $P_p = \{P_1, .., P_p, ., P_{N-2}\}$ $(p \in R_+^{N-2})$ are the values of variables $\delta_k^{t-2-p,t-1-p}$ and are subject to an ordering constant $P_1 < P_2 < ... < P_p$ $(p \leq N-2)$.

Following the previous discussion, the fuzzy piecewise auto-regression QP formula for analysis is represented as Model (18).

$$\textbf{min} \ \sum_{K=1}^{N}\left\{a_{0w*} + a_{1w*}\delta_k^{t-2,t-1} + \sum_{p=1}^{P-1}\left\{B_{pw*}\frac{(|\delta_k^{t-2-p,t-1-p} - P_p| + \delta_k^{t-2-p,t-1-p} - P_p)}{2}\right\}\right\}^2$$

subject to.

(Possibility constraints)

$$a_{0c}^* + a_{1c}^*\delta_k^{t-2,t-1} + \sum_{p=1}^{P-1}\left\{\frac{B_{pc}^*}{2}\left(\left|\delta_k^{t-2-p,t-1-p} - P_p\right| + \delta_k^{t-2-p,t-1-p} - P_p\right)\right\}$$
$$-\left\{a_{0w}^* + a_{1w}^*\delta_k^{t-2,t-1} + \sum_{p=1}^{P-1}\left\{\frac{B_{pw}^*}{2}\left(\left|\delta_k^{t-2-p,t-1-p} - P_p\right| + \delta_k^{t-2-p,t-1-p} - P_p\right)\right\}\right\} \leq \delta_k^{t-1,t-2} + \varepsilon,$$
$$a_{0c}^* + a_{1c}^*\delta_k^{t-2,t-1} + \sum_{p=1}^{P-1}\left\{\frac{B_{pc}^*}{2}\left(\left|\delta_k^{t-2-p,t-1-p} - P_p\right| + \delta_k^{t-2-p,t-1-p} - P_p\right)\right\}$$
$$+\left\{a_{0w}^* + a_{1w}^*\delta_k^{t-2,t-1} + \sum_{p=1}^{P-1}\left\{\frac{B_{pw}^*}{2}\left(\left|\delta_k^{t-2-p,t-1-p} - P_p\right| + \delta_k^{t-2-p,t-1-p} - P_p\right)\right\}\right\} \geq \delta_k^{t-1,t-2} - \varepsilon,$$
$$P \leq N-2, \forall\ k = 1, ..., N$$

(Necessity constraints)

$$a_{0c*}+a_{1c*}\delta_k^{t-2,t-1}+\sum_{p=1}^{P-1}\left\{\frac{B_{pc*}}{2}(\left|\delta_k^{t-2-p,t-1-p}-P_p\right|+\delta_k^{t-2-p,t-1-p}-P_p)\right\}$$
$$-\left\{a_{0w*}+a_{1w*}\delta_k^{t-2,t-1}+\sum_{p=1}^{P-1}\left\{\frac{B_{pw*}}{2}(\left|\delta_k^{t-2-p,t-1-p}-P_p\right|+\delta_k^{t-2-p,t-1-p}-P_p)\right\}\right\}\geq\delta_k^{t-1,t}+\varepsilon,$$
$$a_{0c*}+a_{1c*}\delta_k^{t-2,t-1}+\sum_{p=1}^{P-1}\left\{\frac{B_{pc*}}{2}(\left|\delta_k^{t-2-p,t-1-p}-P_p\right|+\delta_k^{t-2-p,t-1-p}-P_p)\right\}$$
$$+\left\{a_{0w*}+a_{1w*}\delta_k^{t-2,t-1}+\sum_{p=1}^{P-1}\left\{\frac{B_{pw*}}{2}(\left|\delta_k^{t-2-p,t-1-p}-P_p\right|+\delta_k^{t-2-p,t-1-p}-P_p)\right\}\right\}\leq\delta_k^{t-1,t}-\varepsilon,$$
$$P\leq N-2,\forall\ k=1,\ldots,N \tag{18}$$

3.5 Phase IV: Validation and Forecasting

We then calculate $a_{0c}^*, a_{1c}^*, B_{pc}^*$, and B_{pw}^* $(p\in R_+^P, P\leq N-1)$. By substitution, using Eqs. (19) and (20), we determine two values for *k-th* DMU, the upper bound ρ_k^U (denoted as PRY) and the lower bound ρ_k^L (denoted as PLY). Any $\delta_k^{t-1,t-2}$ will lie on $[\rho_k^U,\rho_k^L]$

$$\rho_k^U=(a_{0c}^*+a_{1c}^*\delta_k^{t-2,t-1}+\sum_{p=1}^{P-1}B_{pc}^*\delta_k^{t-2-p,t-1-p})+\left\{a_{0w}^*+a_{1w}^*\delta_k^{t-2,t-1}+\sum_{p=1}^{P-1}B_{pw}^*\delta_k^{t-2-p,t-1-p}\right\} \tag{19}$$

$$\rho_k^L=(a_{0c}^*+a_{1c}^*\delta_k^{t-2,t-1}+\sum_{p=1}^{P-1}B_{pc}^*\delta_k^{t-2-p,t-1-p})-\left\{a_{0w}^*+a_{1w}^*\delta_k^{t-2,t-1}+\sum_{p=1}^{P-1}B_{pw}^*\delta_k^{t-2-p,t-1-p}\right\} \tag{20}$$

Similarly, we calculate a_{0c*}, a_{1c*}, B_{pc*} and B_{pw*} ($p\in R_+^P$, $P\leq N-1$). By substitution, using Eqs. (21) and (22), we determine two values for *k-th* DMU, the upper bound π_k^U (denoted as NRY) and the other lower bound π_k^L (denoted as NLY). For any $\delta_k^{t-1,t-2}, \delta_k^{t-1,t-2}\notin[\pi_k^U,\pi_k^L]$

$$\pi_k^U=a_{0c*}+a_{1c*}\delta_k^{t-2,t-1}+\sum_{p=1}^{P-1}B_{pc*}\delta_k^{t-2-p,t-1-p}-\left\{a_{0w*}+a_{1w*}\delta_k^{t-2,t-1}+\sum_{p=1}^{P-1}B_{pw*}\delta_k^{t-2-p,t-1-p}\right\} \tag{21}$$

$$\pi_k^L=a_{0c*}+a_{1c*}\delta_k^{t-2,t-1}+\sum_{p=1}^{P-1}B_{pc*}\delta_k^{t-2-p,t-1-p}+\left\{a_{0w*}+a_{1w*}\delta_k^{t-2,t-1}+\sum_{p=1}^{P-1}B_{pw*}\delta_k^{t-2-p,t-1-p}\right\} \tag{22}$$

For any *k-th* DMU, we check if the four values, $\rho_k^U\geq\pi_k^U\geq\pi_k^L\geq\rho_k^L$, are satisfied. If satisfied, the *p-1* period efficiency values, ($p_k^{(t-1)}$) to ρ_k^U, π_k^U, π_k^L, and ρ_k^L, are multiplied to obtain the four efficiency values ($\xi_{k,t}^{(U)}$, $\psi_{k,t}^{(U)}$, $\psi_{k,t}^{(L)}$, and $\xi_{k,t}^{(L)}$) of the *t-th* periods. After the four values are obtained, we check $p_k^{(t)}\in[\xi_{k,t}^{(U)},\psi_{k,t}^{(U)}]$ or $p_k^{(t)}\in[\psi_{k,t}^{(L)},\xi_{k,t}^{(L)}]$. After validation, shift the time horizon from *t* to *t+1* period to forecast the efficiency of each DMU.

4. ILLUSTRATION AND DISCUSSION

4.1 Input and Output Variables

The data set obtained from the database of Taiwan Economics Journals (TEJ) from 2002 to 2007 consists of 22 observations ($N=22$). All observations could be referred to as the biggest possibility in Taiwan's commercial banks. Appendix A shows the number of DMUs and a map of commercial bank names. Section 2.2 shows that three inputs ($n_a=3$) and three outputs ($n_b=3$) are chosen. The three output variables are the commission revenue amounts of loans (y_1, in 10^3 New Taiwan dollars (NTD)), amount of loans (y_2 , in 10^3 NTD), and investment (y_3 , in 10^3). The three inputs are assets (x_1, in 10^3 NTD), employee expenditures (x_2, in 10^3 NTD), and amount of deposits (x_3, in 10^3 NTD). We assume that the 2003–2007 data have been processed to pass through gross national product deflators based on 2002 price levels to illustrate the approach. Descriptive statistics of the used variables are presented in Appendix B.

4.2 Efficiency Predictions

The efficiency scores could be calculated by the Section 3.2 evaluation of Model (3). Only DMUs 5, 6, 12, and 22 are efficient from 2002 to 2007. After applying Eq. (4), the CIE index could be calculated from 2002 to 2007. If the cell values are larger than 1, they represent an adjacent period of efficiency improvement; otherwise, they show a case of decay.

The validating CIE ranges are summarized in Table 3. The ranges for $\delta_k^{2003,2002}$, $\delta_k^{2004,2003}$, and $\delta_k^{2005,2004}$ $(\forall k=1,..,22)$ are denoted as independent variables, and $\delta_k^{2005,2006}$ are dependent variables $(\forall k=1,...,22)$. The possibility estimation model and necessity estimation model based on Model (18) can be obtained, as shown in Eq. (23):

$$\delta_k^{2006.2005}=[1.815,0,0.104]-0.6278615*\delta_k^{2003,2002}-0.2207748*\delta_k^{2004,2003} \tag{23}$$

where [1.815, 0, 0.0104] represents the center located at 1.814. The necessity estimation model radius is equal to zero, and the possibility estimation model radius is equal to 0.0104. In Eq. (23), the necessity does not provide ranges but crisp values, which means the upper and lower bound necessities are the same. The first four rows of Table 3 report the regression variables except the first row (i.e., $\delta_k^{2003,2002}$, $\delta_k^{2004,2003}$, $\delta_k^{2005,2004}$); the fifth row ($\delta_k^{2006,2005}$) shows the independent variable; the sixth row shows actual data.

Table 3: Validating CIE index, (PRY, NRY, NRL, PLY)

DMU(k)	ρ_k^U	π_k^U	π_k^L	ρ_k^L	$\delta_k^{2006,2005}$
1	1.061833	0.9541342	0.9541342	0.8464355	0.9942
2	1.057949	0.9467880	0.9467880	0.8356266	0.9771
3	1.064589	0.9531994	0.9531994	0.8418101	0.9701
4	1.072987	0.9609873	0.9609873	0.8489873	0.9050
5	1.069892	0.9661136	0.9661136	0.8623356	1.0000
6	1.069892	0.9661136	0.9661136	0.8623356	1.0000
7	1.082653	0.9759561	0.9759561	0.8692591	1.0267
8	1.077701	0.9739233	0.9739233	0.8701453	1.0124
9	1.066209	0.9575539	0.9575539	0.8488982	0.9871
10	1.084549	0.9693631	0.9693631	0.8541767	0.9593
11	1.091144	0.9626179	0.9626179	0.8340916	1.0017
12	1.069892	0.9661136	0.9661136	0.8623356	1.0000
13	1.069892	0.9661136	0.9661136	0.8623356	0.9738
14	1.085427	0.9678546	0.9678546	0.8502828	1.0147
15	1.050965	0.9389646	0.9389646	0.8269646	0.9950
16	1.084823	0.9728231	0.9728231	0.8608231	1.0288
17	1.110073	1.006295	1.006295	0.9025174	1.0028
18	1.088236	0.9720209	0.9720209	0.8558057	1.0153
19	1.052119	0.9360543	0.9360543	0.8199894	0.9396
20	1.042399	0.9338922	0.9338922	0.8253852	0.9656
21	1.098591	0.9900062	0.9900062	0.8814217	0.9979
22	1.069892	0.966136	0.966136	0.8623356	1.0000

In predicting the efficiency of 2006, Eq. (23) was used as the base and multiple of $p_k^{(2005)}$, as shown in Eq. (24).

$$p_k^{(2006)} = ([1.815, 0, 0.104] - 0.6278615 * \delta_k^{2003,2002} - 0.2207748 * \delta_k^{2004,2003}) * p_k^{(2005)} \tag{24}$$

Forecasting the 2007 efficiency scores should be prepared after validation. The period from 2006 to 2007, based on Eq. (24), was moved such that its forecast functions as Eq. (25).

$$p_k^{(2007)} = ([1.815, 0, 0.104] - 0.6278615 * \delta_k^{2004,2003} - 0.2207748 * \delta_k^{2005,2004}) * p_k^{(2006)} \tag{25}$$

Table 4 shows the accuracy rate as approximately 87% (DMUs 18, 20, 21 were failed to predict). Figure 2 shows the forecasting results.

Table 4: Forecasting 2007 Efficiency, PRY, NRY, NRY, NLY

DMU k	$\xi^{(U)}_{k,2007}$	$\psi^{(U)}_{k,2007}$	$\psi^{(L)}_{k,2007}$	$\xi^{(L)}_{k,2007}$	$p^{(2007)}_{k}$
1	1.0468	0.9436	0.9436	0.8405	0.9754
2	1.0020	0.9022	0.9022	0.8023	0.9530
3	1.0061	0.9054	0.9054	0.8047	0.9808
4	0.9139	0.8250	0.8250	0.7360	0.9140
5	1.0699	0.9661	0.9661	0.8623	1.0000
6	1.0699	0.9661	0.9661	0.8623	1.0000
7	1.0813	0.9790	0.9790	0.8767	0.9923
8	1.0549	0.9549	0.9549	0.8549	1.0000
9	0.9892	0.8925	0.8925	0.7957	0.9596
10	0.9790	0.8795	0.8795	0.7799	0.9794
11	0.9973	0.8938	0.8938	0.7903	0.9738
12	1.0699	0.9661	0.9661	0.8623	1.0000
13	1.0419	0.9408	0.9408	0.8398	1.0000
14	1.0722	0.9700	0.9700	0.8678	1.0000
15	1.0220	0.9262	0.9262	0.8303	0.9514
16	0.9753	0.8780	0.8780	0.7807	0.9438
17	0.9944	0.8988	0.8988	0.8031	0.9789
18	0.9183	0.8254	0.8254	0.7325	0.9667
19	0.9490	0.8583	0.8583	0.7677	0.9081
20	0.9397	0.8471	0.8471	0.7545	0.9431
21	0.9705	0.8750	0.8750	0.7795	0.9808
22	1.0699	0.9661	0.9661	0.8623	1.0000

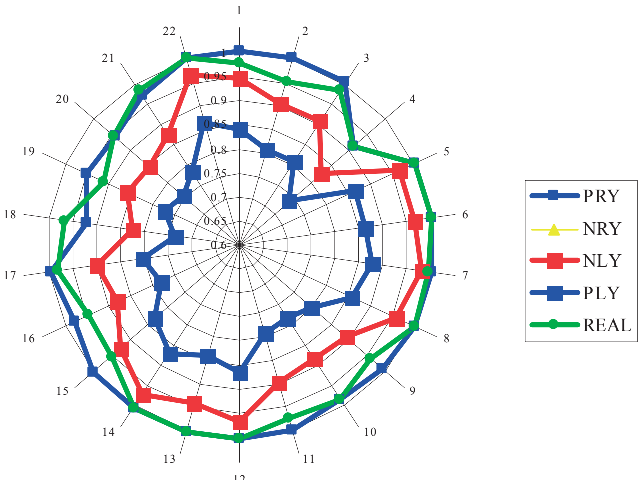

Figure 2: $p_k^{(2007)}$ PRY, NRY, NLY, and PLY by Fuzzy Piecewise Auto-regression

4.3 Disscussion

A forecasting method by hybrid CIE and fuzzy piecewise auto-regression is presented in solving two issues: variables and methodologies selection. The CIE in the variable selections was used as the dependent/independent variable to forecast actual scenarios in place of absolute variables. In 2007, $\xi_{k,2007}^{(U)}$ are the optimal efficiency scores of *k-th* DMU, while $\xi_{k,2007}^{(L)}$ are the pessimistic efficiency scores of *k-th* DMU, and $\psi_{k,2007}^{(L)}$ $(\text{or}\, \psi_{k,2007}^{(U)})$ are the most possible efficiency scores of *k-th* DMU. Table 4 shows that our approaches have 12 DMUs reaching the frontier, but 7 DMUs reached the frontier in the real world (DMUs 1, 2, 3, 7 are different in our analysis). In view of these DMUs, their efficiency scores are closer to our $\psi_{k,2007}^{(L)}$ $(\text{or}\, \psi_{k,2007}^{(U)})$ values.

Three DMUs (DMU 18, 20, 21) failed the forecast. In the case of DMU 18, efficiency scores ranged from 0.8492 to 0.9218 between 2002 and 2006, but the 2007 efficiency scores of DMU 18 were up to 0.966. Fuzzy piecewise auto-regression refers to past efficiency to regress the most possible efficiency scores. DMU 18 historical data does not prove that it can reach over 0.9218 such that the possibility regression predicts its optimal values at 0.9183. In addition, forecasting DMU 5 is always equal to one if the conventional regression approach is applied. DMU 5 historical data presents the efficiency of DMU 5, which is always equal to one. However, fuzzy piecewise auto-regression provides a range from 0.8623 to 1, with values closer to 0.9663. This was caused by the catching-up effect. Although the CIE does not draw input and output relationships, prior concepts about efficiency scores were compared with frontiers or best practices. Thus, the efficiency of DMU 5 may or may not be frontier. As such, the two ranges elaborate that efficiency scores were compared with the frontier. The fuzzy ranges (i.e., between $\xi_{k,t}^{(U)}$ and $\xi_{k,t}^{(L)}$) provided relative concepts and the highest possible efficiency $(\psi_{k,t}^{(U)}, \psi_{k,t}^{(L)})$.

This study attempts to compare non-piecewise fuzzy regression to evaluate the data of 22 Taiwan commercial banks from 2002 to 2007 in the selection of methodologies. Models (11) and (13), similar to piecewise conditions, were used to calculate the possibility and necessity estimation models. The efficiency of the 2006 validation can be calculated through Eq. (26)

$$Eff_{2006} = (0.9089242 + 0.05872535 * CIE_{2003}^{2004} + [0,0,0.1259782] * CIE_{2004}^{2005}) * Eff_{2005} \quad (26)$$

This non-piecewise approach (Eq. 26) uses a larger range to cover predicting and real data, unlike the piecewise approach, which uses the loose range. Using prediction ranges from zero to one implies that we do not need any predicting approaches. As such, we could understand that the predicting functionality relies on not only accuracy but also precision (i.e., predicting ranges are good for providing small ranges). Therefore, the piecewise approach could be better than the non-piecewise approach.

5. ANALYSIS

This paper discusses two hybrid methodological developments to show how the efficiency of DEA can be used in forecasting. This proposed method has two advantages. First, CIE shows the relative efficiency of two adjacent periods and avoids the direct use of input and output variables. Therefore, CIE provides not only a priori relative concepts on frontiers and best practices, but also shows possible efficiency. Second, historical data were used to regress the possibility and necessity estimation model in place of the random error-type regression model. The four ranges provide decision makers with specific DMU suggestions that the specific DMU stands for (i.e., if the specific DMU does not frontier in the current, such that it can make an effort to reach the frontier).

However, the analyses in the present study have certain limitations. First, efficiency prediction can be divided two parts: efficiency shift and efficiency movement. The former is mainly caused by technique change but the latter is caused by changes in the input and output ratio. Our analysis solved the efficiency shift issues, but did not solve those for efficiency moment. This issue will be evaluated in our future work. Second, we do not explain external effects that influence the evaluation result, such as government power. Third, we excluded incomplete data, newcomers, and mergers and acquisitions of commercial banks from 2002 to 2007. The DEA method can be applied to evaluation and planning techniques. Further research can be conducted regardless of the method, on other possible concepts.

REFERENCE

1. Barr, R.S., Killgo, K.A., Siems, T.F., and Zimmel, S. “Evaluating the Productive Efficiency and Performance of U.S. Commercial Banks,” *Financial Industry Studies working paper 99-3*, Federal Reserve Bank of Dallas, 2002.
2. Bowlin, W.F., Charnes, A., Cooper, W.W., and Sherman, H.D. “Data envelopment and regression approaches to efficiency estimation and evaluation,” *Annals of Operational Research* (2:1), 1985, pp.113–138.
3. Chang, C.C. “The Nonparametric Risk-Adjusted Efficiency Measurement: An Application of Taiwan’s Credit Department of Farmers’Associations.” *American Journal of Agricultural Economics* (81:4), 1999, pp. 902-913.
4. Chen, T.Y., and Yeh, T. L. “A Measurement of Bank Efficiency, Ownership and Productivity Changes in Taiwan,” *Service Industries Journal* (20), 2000, pp. 95-109.
5. Chilingerian, J.A., and Sherman, H.D. “Managing physician efficiency and effectiveness in providing hospital services,” *Health Service Management Research* (3:1), 1990, pp. 3-15.

6. Caves, D.W., Christensen, L.R., and Diewert, W.E. "The economic theory of index numbers and the measurement of input, output and productivity," *Econometric* (50:6), 1982, pp. 1393–1414.
7. Coelli, T., Prasada Rao, D.S., and Battese, G.E. *An Introduction to Efficiency and Productivity Analysis*, Kluwer Academic Publishers, Boston, 1998.
8. Cubbin, J., and Zamani, H. "A comparison of performance indicators for training and enterprise councils," *Annals of Public and Cooperative Economics*, 1996.
9. Cook, W.D., and Seiford, L.M. "Data envelopment analysis (DEA)-thirty years on," *European Journal of Operational Research* (192:1), 2009, pp. 1-17.
10. Edirisinghe, N.C.P., and Zhang, X. "Generalized DEA model of fundamental analysis and its application to portfolio optimization," *Journal of Banking & Finance* (31:11), 2007, pp. 3311–3335.
11. Färe, R., Grosskopf, S., Lindgren, B., and Roos, P. "Productivity developments in Swedish hospitals: A Malmquist output index approach," *Discussion Paper No. 89-3*, Southern Illinois University, 1989.
12. Färe, R., Grosskopf, S., and Lovell, C.A.K. *Production Frontiers*, Cambridge University Press, Cambridge, 1994.
13. Färe, R., Grosskopf, S., and Roos, P. "Productivity and quality changes in Swedish pharmacies," *Internal J. Production Economics* (39:1-2), 1995, pp. 137-144.
14. Fisher, I. *The Making of Index Numbers*, Houghton-Mifflin Co, Boston, 1922.
15. Grabowski, R., Rangan, N., and Rezvanian, R. "The Effect of Deregulation on the Efficiency of U.S. Banking Firms," *Journal of Econometrics and Business* (46:1), 1994, pp. 39-54.
16. Golany, B., and Roll, Y. "An Application Procedure for DEA," *Omega* (17:3), 1989, pp. 237-250.
17. Haslem, J.A., Scheraga, C.A., and Bedingfield, J. P. "DEA Efficiency Profiles of U.S. Banks Operating Internationally," *International Review of Economics and Finance* (8:2), 1999, pp. 165-182.
18. Huang, C.Y., and Tzeng, G.H. "Multiple generation product life cycle predictions using a novel two-stage fuzzy piecewise regression analysis method," *Technological Forecasting and Social Change* (75:1), 2008,pp. 12-31.
19. Hsu, F.M., and Hsueh, C.C. "Measuring relative efficiency of government sponsored R&D projects: A three-stage approach," *Evaluation and Program Planning* (32:2), 2009, pp. 178-186.
20. Miller, S.M., and Noulas, A.G. "The technical efficiency of large bank production," *Journal of Banking and Finance* (20:3), 1996, pp. 495-509.
21. Megginson, W.L., Nash, R.C., and Randenborgh, M.V. "The Financial and Operating Performance of Newly Privatized Firms: An International Empirical Analysis," *Journal of*

Finance (49:2), 1994, pp. 403-452.

22. Kao, C., and Liu, S.T. "Predicting bank performance with financial forecasts: A case of Taiwan commercial banks." *Journal of Banking and Finance* (28:10), 2004, pp. 2353-2368.
23. Lin, P.W. "An efficiency analysis of commercial bank mergers in Taiwan: Data envelopment analysis," *Taiwan Academy of Management Journal* (1), 2002, pp. 341-355.
24. Ozcan, Y.A. "Efficiency of hospital service production in local markets: the balance sheet of U.S. medical armament," *Social Economic Planning Sciences* (29:2), 1995, pp. 139-150.
25. Schmidt, P. "Frontier production functions," *Econometric Reviews* (2), 1986, pp. 289–328.
26. Soteriou, A.C., Karahanna, E., Papanastasiou, C., and Diakourakis, M.S. "Using DEA to evaluate the efficiency of secondary school: the case of Cyprus," *International Journal of Educational Management* (12:2), 1998, pp. 65-73.
27. Sueyoshi, T. "Stochastic DEA for restructure strategy: an application to a Japanese petroleum company," *Omega* (28:4), 2000, pp 385-398.
28. Tanaka, H. "Fuzzy data analysis by possibilistic linear models," *Fuzzy Sets and Systems* (24:3), 1987, pp. 363-375.
29. Tanaka, H., Hayashi, I., and Watada, J. "Possibilistic linear regression analysis for fuzzy data," *European Journal of Operational Research* (40:3), 1989, pp.389-396.
30. Tanaka, H., and Ishibuchi, H. Possibilistic regression analysis based on linear programming in J. Kacprzyk, M. Fedrizzi (Eds.), *Fuzzy Regression Analysis*, Omnitech Press, Warsaw, Poland, 1992, pp. 47-60.
31. Tanaka, H., and Watada, J. "Possibilistic linear systems and their application to the linear regression model," *Fuzzy Sets and Systems* (27:3), 1988, pp. 275-289.
32. Tsai, M.C, Lin, S.P, Cheng, C.C., and Lin, Y.P. "The consumer loan default predicting model- An application of DEA-DA and neural network," *Expert systems with application* (36:9), 2009, pp. 11682-11690.
33. Tortosa-Ausina, E. "Nontraditional activities and bank efficiency revisited: A distributional analysis for Spanish financial institution," *Journal of Economics and Business* (55:4), 2003, pp.371-395.
34. Thanassoulis, E. "A comparison of regression analysis and data envelopment analysis as alternative methods for performance assessments," *Journal of the perational Research Society* (44:11), 1993, pp. 1129–1144.
35. Tyagi, P., Yadav, S.P., and Singh, S.P. "Relative performance of academic departments using DEA with sensitivity analysis," *Evaluation and Program Planning* (32:2), 2009, pp. 168-177.
36. Troutt, M.D., Hu, M.Y., and Shanker, M.S. "A distribution-free approach to estimating

best response values with application to mutual fund performance modeling," *European Journal of Operational Research* (166:2), 2005, pp. 520-527.

37. Törnqvist, L. "The bank of Finland's consumption price index," *Bank of Finland Monthly Bulletin* (10), 1936, pp. 1–8.
38. Wu, D., Yang, Z., and Liang, L. "Using DEA-neural network approach to evaluate branch efficiency of a large canadian bank," *Expert Systems with Applications* (31:1), 2006, pp. 108-115.
39. Yeh, C.C., Chi, D.J., and Hsu, M.F. "A hybrid approach of DEA, rough set and support vector machines for business failure prediction," *Expert system with applications* (37:2), 2010,pp. 1535-1541.
40. Yu, J.R., Tzeng, G.H., and Li, H.L. "A general piecewise necessity regression analysis based on linear programming," *Fuzzy Sets and Systems* (105:3), 1999, pp. 429-436.
41. Yu, J.R., Tzeng, G.H., and Li, H.L. "Interval piecewise regression model with automatic change-point detection by quadratic programming," *Internation Journal of Uncentainty Fuzziness and Knowledge-Based System* (13:3), 2005, pp. 347-361.
42. Yu, J.R., Tzeng, G.H., and Li, H.L. "General Fuzzy Piecewise Regression Analysis with Auto Change Point Detection," *Fuzzy Sets and Systems* (119:2), 2001, pp. 247-257.
43. Yue, P. "Data Envelopment Analysis and Commercial Bank Performance: A Primer with Applications to Missouri Banks," *Federal Reserve Bank of St. Louis Review* (74), 1992, pp. 31-45.
44. Yunos, J. M., and Hawdon, D. "The efficiency of the national electricity board in Malaysia: an inter countrycomparison using DEA," *Energy Economics* (19), 1997, pp. 255-269.
45. Zofío, J.L., and Lovell, C. A. K. "Graph Efficiency and Productivity Measures: An Application to US Agriculture," *Applied Economics* (33), 2001, pp. 1433-1442.19.

Appendix A: The 22 Taiwan commercial banks, their respective numbers, and commercial bank names

DMU	Bank	DMU	Bank	DMU	Bank	DMU	Bank
1	ChangHwa Bank	7	ChinaTrust Bank	13	Union Bank of Taiwan	19	TC Bank
2	First Bank	8	Cathay Bank	14	Bank of SinoPac	20	Entie Bank
3	Hua Nan Bank	9	Fubon Bank	15	E.SUN Bank	21	JIH SUN Bank
4	Mega Bank	10	Taiwan Business Bank	16	YuanTa Bank	22	Taiwan Cooperative Bank
5	King's Town Bank	11	Bank of Kaohsiung	17	TaiShin Bank		
6	TaiChung Bank	12	COSMOS Bank	18	Far Eastern Bank		

Appendix B: Taiwan Commerical Banks Descriptive Data from 2002 to 2007

Period	Variable	Mean	Stdev	Max	Min
2002	Commission revenue	10753525.05	9243351.32	26902952.00	1375074.00
	Amount of loans	3616.45	2927.48	9793.00	932.00
	Investment	467917.41	402384.71	1548830.00	117027.00
	Asset	2053.45	2272.26	10035.00	186.00
	Employee expenditures	355658.82	294983.18	1127292.00	88803.00
	Amount of deposit	61017177.91	66461982.33	252734211.00	6477928.00
2003	Commission revenue	11495825.50	9798462.43	33424599.00	2393015.00
	Amount of loans	3977.91	2974.29	10514.00	987.00
	Investment	512122.05	427232.73	1640739.00	12187.00
	Asset	2595.77	2671.52	12432.00	301.00
	Employee expenditures	389112.05	310228.62	1168354.00	93959.00
	Amount of deposit	111311167.36	117341836.63	333353270.00	2520892.00
2004	Commission revenue	11715734.95	9608655.15	33978243.00	2345382.00
	Amount of loans	4500.50	3330.18	11965.00	1043.00
	Investment	552297.82	437430.53	1681279.00	125098.00
	Asset	3659.18	3780.32	17806.00	335.00
	Employee expenditures	424049.91	322145.36	1220995.00	102397.00
	Amount of deposit	140642055.32	143383968.37	398996916.00	2478795.00
2005	Commission revenue	12438270.64	9694251.83	33429366.00	2373280.00
	Amount of loans	4754.82	3447.33	12459.00	1061.00
	Investment	603705.36	440713.06	1608747.00	127303.00
	Asset	3850.73	4131.05	18620.00	282.00
	Employee expenditures	468763.91	332844.09	1258451.00	100867.00
	Amount of deposit	136170532.18	130640406.93	390606990.00	4736937.00
2006	Commission revenue	13157930.59	10091337.20	33591776.00	2344018.00
	Amount of loans	4841.23	3397.60	12160.00	842.00
	Investment	667680.73	487253.49	1917281.00	123279.00
	Asset	3844.36	4181.17	18782.00	325.00
	Employee expenditures	538125.45	414045.91	1703126.00	98120.00
	Amount of deposit	143565546.77	132704623.16	384000000.00	4771517.00
2007	Commission revenue	13142148.36	10374880.60	33866474.00	2265556.00
	Amount of loans	5213.32	3447.35	12143.00	1043.00
	Investment	690772.41	515772.77	1998654.00	130526.00
	Asset	4733.82	5236.30	24661.00	362.00
	Employee expenditures	563528.50	437382.17	1719370.00	110744.00
	Amount of deposit	133674192.27	126465809.61	371145484.00	2773119.00