Stochastic Data Envelopment Analysis: Oriented and Linearized Models*

František Brázdik†

January 2004

Abstract

In this paper the chance constrained problems for DEA analysis are constructed. The goal is to construct oriented DEA models that account for stochastic noise in the analyzed data. The noise in the form of single factor symmetric error is incorporated into the model and the corresponding stochastic programming problem is created. The stochastic models are transformed into their deterministic equivalents and then linearized. The linearized form of model allows to use the interior point methods for solving the linear programming problems.

Keywords: stochastic dea, linear programming, efficiency
JEL classification: C14, D24, C61

*This research was supported by the World Bank Fellowship.

†A joint workplace of the Center for Economic Research and Graduate Education, Charles University, Prague, and the Economics Institute of the Academy of Sciences of the Czech Republic. Address: CERGE-EI, P.O. Box 882, Politických vězňu 7, Prague 1, 111 21, Czech Republic; E-Email: frantisek.brazdik@cerge.cuni.cz

1 Introduction

Data envelopment analysis (DEA) involves an alternative principle for extracting information about a population of observations, so called decision making units (DMUs), that are described by the same quantitative characteristics. This is reflected by assumption that each DMU uses the same set of inputs to produce the same set of outputs, but the inputs are consumed and outputs are produced in various amounts.

DEA and Stochastic Frontier Analysis (SFA) models have been developed for the purpose of production frontier search. DEA involves an alternative approach to SFA for information extraction from the population observations of decision processes. The DEA approach is a nonparametric approach to estimating the production frontier and therefore DEA does not require specification of the production function form. DMUs are directly compared against a peer or combination of peers. In contrast to parametric approaches for information extraction, the objective of the DEA is to calculate a linear (piecewise linear) frontier determined by a set of Pareto-efficient DMUs. This frontier is used to calculate the relative measure (among the elements of analyzed DMU set) of DMU's technical efficiency.

The measurement of input and output values is the subject to errors and noise. Also, the analyzed production sector may face different shocks. The noise in data usually leads to mistakes in production frontier specification and efficiency scores. The dilemma of efficiency evaluation approach choice depends on the trade off between the minimal specification that favors DEA and handling of stochastic error in measuring DMU efficiency that favors SFA. To compete with SFA in error handling, the stochastic data envelopment analysis (SDEA) approach was developed by considering the value of inputs and outputs as random variables in the SDEA approach.

Present SDEA approaches lead to chance constrained optimization problems that consume extensive amount of computational time for the optimal solution search even for the simple stochastic models. I continue the chance constrained programming tradition in which random disturbances are incorporated in inputs and outputs on the assumption that the probabilistic distribution of disturbances is known. In this paper, two classes of efficiency dominance are defined. According to the dominance definitions the two classes of stochastic models are derived. In the further model development proportional input reduction and output augmentation is introduced in the models and by use of the simplified error structure and linearization methods the linear deterministic models are derived. The linearization allows to use the interior point methods for linear programming problems that are capable to solve large size problems using significantly reduced amount of computational time in comparison with solving chance constrained optimization problems.

The following section reviews the literature on DEA and SDEA. In the third section usual notation is introduced. The fourth section defines production possibility set properties that will be used to construct stochastic models in the following sections. Two model classes will be derived according to different approaches of stochastic error inclusion. Subsequently, the model's error structure is presented and incorporated in the model and the derivation of linearized models is described. This is followed by the introduction of the efficiency measure that is used to construct the oriented models that are linearized. The last section proposes numerical methods that can be used to solve such models and introduces my solver that can be used to solve these models when applying these models for specific analysis.

2 Literature review

As Charnes, Cooper, Lewin and Seiford (1994) explained in their introduction, the story of data envelopment analysis begun with Edwardo Rhodes's dissertation, which was the basis for the later published paper by Charnes, Cooper and Rhodes (1978). In his dissertation, E. Rhodes analyzed the educational program for disadvantaged students in the USA. He compared the performance of students from participating and not participating schools in the program. The performance was recorded in terms of inputs and outputs, e.g.:"Increased self-esteem" (measured by psychology tests) as output and the time spent by mother reading with child as input. The following work on efficiency evaluation of multiple inputs and outputs technology led to the paper by Charnes et al. (1978), where the CCR model for DEA was formulated.

The presented CCR model is capable of handling only the technology with constant returns to scale. This fact is reflected in the shape of the production possibility frontier when the frontier is formed by a single ray. The DMU is evaluated as efficient if it is an element of production possibility frontier. To handle the variable returns to scale the CCR model was extended by Banker, Charnes and Cooper (1984). Since the BCC model's frontier is a piecewise linear set, Banker et al. (1984) defined weak efficiency (weak efficient DMU has nonzero slacks) and efficiency (efficient DMU has zero slacks). Further, only the efficient DMUs are elements of the estimated production possibility set frontier in the framework of the BCC model.

Since 1978, over 1000 articles, books and dissertations have been published[1] and DEA theory and applications have rapidly extended. As many applications suggest, DEA can be a powerful tool when used wisely. Two capabilities that make DEA a powerful tool are capability of handling multiple inputs and outputs models and that these inputs and outputs can have different measurement units. For example, input could be in units of lives saved or it could be in units of dollars without requiring an a priori tradeoff between the two. This property allows expansion of DEA methodology into very different production sectors. To examine the efficiency of hospitals or health care centers is one very popular application of DEA, e.g. a recent study by Halme and Korhonen (1999) examines the dental care units or the study by Byrnes and Valdmanis (1989) where 123 US hospitals were covered. Other applications of DEA methodology cover industries like air transportation (Land, Lovell and Thore 1993), fishing (Walden and Kirkley 2000) and banking (Ševčovič, Halická and Brunovský 2001).

The study by Byrnes and Valdmanis (1989) continued the expanding interest in health care applications of DEA that occurred in the beginning of the 1980's. Many earlier studies of cost efficiency calculated only the technical-efficiency of DMUs and this study also examined the decomposition of overall efficiency into its component parts. This means that Byrnes and Valdmanis (1989) ascertained how efficiently hospitals are using each of the inputs or outputs in comparison to other competitors from the DMUs set. Authors also mentioned how the managers can utilize the information. Their approach shows the variability of information that can be gathered using the DEA model.

The expanding number of papers devoted to DEA helped to identify the limita-

[1] According to Emrouznejad (1995-2001) homepage.

tions of the DEA approach. An analyst should keep these limitations in mind when choosing whether or not to use DEA. DEA is good at estimating the "relative" efficiency of a DMU but it converges very slowly to "absolute" efficiency. In other words, DEA reveals how well DMU is doing compared to other DMU but not compared to a "theoretical maximum." This is the result of the analyst's limitation in knowledge of the true production function. Figure (1) shows the difference between the true production frontier and the estimated production frontier.

Since DEA is an extreme point technique, noise (even symmetrical noise with zero mean) such as measurement error can cause significant problems, because the frontier is sensitive to these errors.

As the consequence of this, theoretical attempts to incorporate these errors were made. The SDEA works are based on the theoretical paper by Land et al. (1993), where the authors used their new models to examine the efficiency of the same schooling program for disabled scholars as in Charnes et al. (1978). In Land et al. (1993), the authors offered the prospect of stochastic data envelopment analysis and constructed their own model (LLT model). They introduced the stochastic component to DEA and created chance constrained problems by introducing the variability to outputs that is conditional on inputs. This simply means that only outputs were taken as normally distributed random variables. After the stochastic optimization problems were created, Land et al. (1993) transformed these problems to their deterministic equivalents, which allowed them to determine the efficient DMUs.

Olesen and Petersen (1995) presented a different approach to incorporate the stochastic component into DEA. Olesen and Petersen (1995) assumed that inefficiency of DMU can be decomposed into true inefficiency and disturbance term. The approaches of Land et al. (1993) and Olesen and Petersen (1995) to SDEA are compared by Olesen (2002) and the weaknesses of both approaches are identified. The LLT model is criticized because it does not account for all the correlations that can occur in disturbances. Olesen (2002) critique the OP-model proposed by Olesen and Petersen (1995) because the OP-model ignores the fact that a convex combination of, e.g., two i.d. random input output vectors from two DMUs has a lower variance than the random vectors themselves, except for the case where the input output vectors are perfectly correlated. After Olesen (2002) stressed the weaknesses of both the models, he proposed a model that combines attractive features of the LLT and OP models. Straightforward remedy for the OP model is to take the union of confidence regions for any linear combination of the stochastic vectors themselves rather than using a piecewise linear envelopment of the confidence regions. Olesen (2002) implemented this idea and derived the combined chance constrained model in his paper.

The theoretical paper by Huang and Li (n.d.) sketches stochastic models with the possibility of variations in inputs and outputs. Huang and Li (n.d.) defined the efficiency measure of a DMU via joint probabilistic comparisons of inputs and outputs with other DMUs which can be evaluated by solving a chance constrained programming problem. By utilizing the theory of chance constrained programming, deterministic equivalents are obtained for both situations of multivariate symmetric random disturbances and a single random factor in the production relationships. The linear deterministic equivalent and its dual form are obtained via the programming theory under the assumption of the single random factor. An analysis of stochastic variable returns to scale is developed using the idea of stochastic supporting hyperplanes. The relationships of the presented SDEA models with some conventional DEA models are also discussed.

The paper by Gstach (1998) shows that there are research directions in which the future developments on DEA and SDEA can be driven. Gstach (1998) addresses the issue that the outcome of a production process might not only deviate from a theoretical maximum due to inefficiency, but also due to non-controllable errors. As is often the case, this raises the issue of reliability of DEA in noisy environments. The author proposes to assume an i.i.d. data generating process with a bounded noise component. This assumption makes the approach that mixes the parametric and non-parametric approach to production frontier estimation feasible. Gstach (1998) propose using DEA to estimate a pseudo frontier (nonparametric shape estimation) and then apply a maximum likelihood-technique to the DEA-estimated efficiencies to estimate the scalar value by which this pseudo-frontier must be shifted downward to get the true production frontier (location estimation).

At the end of this review, the paper that is devoted to DEA models computational problems is presented. The major problems associated with solving the DEA models are the analysis of large set of DMUs and the solutions with zero elements. The analysis of large data set leads to large size optimization problems that can be costly to solve. The solutions with zero elements cause problems when these solutions are interpreted as inputs and outputs shadow prices. Gonzales-Lima, Tapia and Thrall (1996) present the primal-dual interior-points computational methods as the methods that significantly improve the reliability of solution in comparison to simplex methods. The interior-points methods maximize the product of the positive components among solutions, which means that the number of zero components of the optimal solution is minimized. Due to this solution's property it is easier to interpret the DEA models results.

3 Notation

In this section, the notation and definitions are introduced which are introduced will be used to construct the stochastic DEA models in the next section. The notation in this paper coincides with the notation used in papers by Li (1998) and Huang and Li (n.d.). The task is to analyze the set of DMU_j where $1 \leq j \leq n$. Each of the DMUs is described by the vector $\tilde{x}_j \in \mathbb{R}^m_+, \tilde{x}_j = (\tilde{x}_{1j}, \ldots, \tilde{x}_{mj})^T$ of m input amounts that are used to produce s outputs in amounts described by vector $\tilde{y}_j \in \mathbb{R}^s_+$, $\tilde{y}_j = (\tilde{y}_{1j}, \ldots, \tilde{y}_{sj})^T$. These vectors are aggregated to matrices of inputs and outputs and the following matrix notation will be used:

matrix of inputs	$\tilde{X} = (\tilde{x}_1, \ldots, \tilde{x}_n)$
i^{th} row of 'input' matrix $\tilde{X}$	${}_i\tilde{x} = (\tilde{x}_{i1}, \ldots, \tilde{x}_{in})^T,\ i = 1, \ldots, m$
matrix of expected inputs	$\bar{X} = (\bar{x}_1, \ldots, \bar{x}_n)$
i^{th} row of expected 'input' matrix $\bar{X}$	${}_i\bar{x} = (\bar{x}_{i1}, \ldots, \bar{x}_{in})^T,\ i = 1, \ldots, m$
matrix of outputs	$\tilde{Y} = (\tilde{y}_1, \ldots, \tilde{y}_n)$
r^{th} row of 'output' matrix $\tilde{Y}$	${}_r\tilde{x} = (\tilde{y}_{r1}, \ldots, \tilde{y}_{rn})^T,\ r = 1, \ldots, s$
matrix of expected outputs	$\bar{Y} = (\bar{y}_1, \ldots, \bar{y}_n)$
r^{th} row of expected 'output' matrix $\bar{Y}$	${}_r\bar{y} = (\bar{y}_{r1}, \ldots, \bar{y}_{rn})^T,\ r = 1, \ldots, s$

The additional notation is introduced in the following section where the error structure is described.

4 Stochastic efficiency dominance

In the stochastic framework the DMUs are characterized by distributions moments. The production possibility set is defined in terms of inputs and outputs means. In further model development, the second moments will be used in the transformation from the chance constrained problem to its deterministic equivalent.

The true production possibility set can be constructed using the true production function. In nonparametric DEA methodology, this function is not known, therefore the general production possibility set is defined and the set of properties that the production possibility set should fulfill is postulated.

Definition 1. General stochastic production possibility set $T \subset \mathbb{R}_+^{m+s}$ is defined as follows: $T = \{(\tilde{x}, \tilde{y}) \mid \text{ using inputs } \tilde{x} \text{ outputs } \tilde{y} \text{ can be produced}\}$

When constructing the stochastic production possibility set T using the DMU_j , $j = 1, \ldots, n$, T should have these properties:

Property 1. Convexity: If $(\tilde{x}_j, \tilde{y}_j) \in T$, $j = 1, \ldots, n$ and $\lambda \in \mathbb{R}_+^n$, $\sum_{j=1}^n \lambda_j = 1 \Rightarrow (\tilde{X}\lambda, \tilde{Y}\lambda) \in T$.

Property 2. Inefficiency property: If $(\bar{x}, \bar{y}) \in T$ and $x \geq \bar{x}$, then $(x, \bar{y}) \in T$.
If $(\bar{x}, \bar{y}) \in T$ and $y \leq \bar{y}$ then $(\bar{x}, y) \in T$.

The second property of production possibility set means that less output can be produced with the same inputs. It reflects the situation when some amount of inputs is wasted in production production process.

Property 3. Minimum extrapolation: T is the intersection of all sets satisfying convexity and inefficiency property and subject to that each of the observed vectors $(\tilde{x}_j, \tilde{y}_j) \in T, j = 1, \ldots, n$.

Set $T_0 = \{(\tilde{x}, \tilde{y}) \mid \tilde{x} \geq \tilde{X}\lambda, \tilde{y} \leq \tilde{Y}\lambda, \lambda \geq 0\}$ satisfies the aforementioned properties. T_0 is the stochastic generalization of the production possibility set of constant returns to scale production function as it was used by Charnes et al. (1978) in derivation of the CCR model. Similarly, the set $T_1 = \{(\tilde{x}, \tilde{y}) \mid \tilde{x} \geq \tilde{X}\lambda, \tilde{y} \leq \tilde{Y}\lambda, e^T\lambda = 1, \lambda \geq 0\}$ also satisfies the properties postulated for the stochastic production set. Set T_1 is the production possibility set appropriate for production technology with variable returns to scale. Further, set T_1 and its modifications will be used to derive models with variable returns to scale.

Generally, the parameterized production possibility set T_φ can be defined as follows: $T_\varphi = \{(\tilde{x}, \tilde{y}) \mid \tilde{x} \geq \tilde{X}\lambda, \tilde{y} \leq \tilde{Y}\lambda, \varphi(e^T\lambda) = \varphi, \lambda \geq 0\}$. T_φ covers the cases of T_1 and T_0 for choice of parameter φ, $\varphi = 0, 1$.

Next, the concept of efficiency used in DEA is based on following efficiency definition:

Definition 2. Relative Efficiency: A DMU is to be rated as efficient on the basis of available evidence if and only if the performances of other DMUs does not show that some of its inputs or outputs can be improved without worsening some of its other inputs or outputs.

The efficient DMUs identification in the set of all DMUs is done through the efficiency dominance relation. In this section, two classes of efficiency dominance will be defined. The first definition will be used to create the almost 100% chance constrained models. The second dominance definition will be used to define the chance constrained models. These models will be derived in the following sections, where the theorems that relate the efficiency dominance definition and constructed models will be mentioned.

The point in the production possibility set T_φ is called the efficient point of the production possibility set if there is not another production point that produces more of output without consuming more input, or consumes less of input without producing less output. This leads to the following efficiency domination definition among the DMUs:

Definition 3. Efficiency dominance: DMU_j is not dominated in the sense of efficiency if $\nexists (x^*, y^*) \in T_\varphi$ such that $x^* \leq x_j$ or $y^* \geq y_j$ with at least one strict inequality for input or output components.

This definition demonstrate the efficiency concept of DEA and is used to derive the deterministic models and there is no possibility of the violating of the production possibility set properties. In the deterministic environment, the non-dominated elements of the production possibility set frontier set up the production envelopment surface. The Figure (1) shows the set of DMUs divided into the efficient and inefficient DMUs. The efficient DMUs are used to set up the estimate of the production possibility frontier. The elements of the the production possibility envelopment estimate dominate the other elements of the production possibility set.

In the stochastic framework, the efficiency domination violations are allowed with the probability α, $0 \leq \alpha \leq 1$.[2] In the chance constrained programming methodology the term $1-\alpha$ is interpreted as the modeler's confidence level and α is interpreted as the modeler's risk. The risk equals to the probability measure of the extent to which specific conditions are violated.

In the almost 100% confidence approach, the efficiency dominance can be violated with probability α and the production possibility constraints are almost certainly not violated. In the chance constrained model, it is allowed to violate the constraints specifying the production possibility set with probability α. For the case of the almost 100% confidence chance constrained approach, let's consider the case where Pareto-efficiency can be violated due to random errors and therefore the stochastic efficiency of point is defined as follows:

Definition 4. Stochastic efficiency of point in set T_φ: $(\tilde{x}^*, \tilde{y}^*) \in T_\varphi$ is called α–stochastically efficient point associated with $T_\varphi \Leftrightarrow$ if the analyst is confident that $(\tilde{x}^*, \tilde{y}^*)$ is Pareto-efficient with probability $1-\alpha$ in the set T_φ.

This definition means that point $(\tilde{x}^*, \tilde{y}^*)$, considered as efficient, is dominated (in the sense of efficiency dominance) by any other point in T_φ with a probability less or equal to α. Using the definition of efficient point the efficiency of DMU_j is defined as follows:

Definition 5. Stochastic efficiency of DMU_j : DMU_j is α–stochastically efficient in set $T_\varphi \Leftrightarrow (\tilde{x}_j, \tilde{y}_j)$ is point associated with DMU_j and $(\tilde{x}_j, \tilde{y}_j) \in T_\varphi$ is α–stochastically efficient production point associated with T_φ.

These definitions and aforementioned properties of the set T_φ straightforwardly imply that for efficient DMU_j and for any λ such that $\varphi(e^T\lambda) = \varphi, \lambda \geq 0$ the

[2] In the following text it is assumed $0.5 \leq \alpha \leq 1$.

expression

$$Prob(\tilde{X}\lambda \le \tilde{x}_j, \tilde{Y}\lambda \ge \tilde{y}_j) \quad \le \quad \alpha$$

holds with at least one strict inequality in input–output constraints.

The deterministic and stochastic approaches are compared from the view of the shape of efficient points set to demonstrate the difference between them. Figure (1) shows the deterministic frontier and compares it to the true production possibility frontier. We observe two inefficient DMUs, namely DMU #4 and DMU #5. The solid piecewise linear line is the unknown true production possibility frontier and the dashed line is DEA estimate of the production possibility frontier. At Figure (2) the same DMUs are pictured and the set of efficiency dominant DMUs is pictured as the grey shaded area. Comparison of Figures (1) and (2) shows that the deterministic production possibility set frontier is a subset of the stochastic possibility set frontier. Due to this fact more DMUs can be evaluated as efficiency dominant in the stochastic framework than in the deterministic. At Figure (2) this is the case for the DMU #4 that became an efficient unit in the stochastic framework.

The second class of stochastic models considered in this paper is the case where production possibility set constraints can be violated with prescribed probability. In this approach, the modeler's risk relates to the risk that the dominating point is not element of the production possibility set due to the presence of noise in the data. This means that the production point can be dominated by the point that violates the production possibility set constraints. For the purpose of chance constrained models derivation, the efficiency dominance is defined in the following way:

Definition 6. Chance constrained efficiency: Point (x_{ij}, y_{rj}) is not dominated in the sense of chance constrained efficiency if $\nexists\, (x^*, y^*)$ such that $Prob(x_i^* \le x_{ij}) \ge 1-\alpha$ or $Prob(y_r^* \ge y_{rj}) \ge 1-\alpha$, $0.5 \le \alpha < 1$, $j = 1, \ldots, n$;
$i = 1, \ldots, m$; $r = 1, \ldots, s$ holds with at least one strict inequality for input or output components.

Similarly as in the case of α-stochastic dominance, the production point is efficient in the sense of chance constrained efficiency defined by definition (6) when this point is not dominated by any other point that fulfills each constraint of the set T_φ with probability at least $1-\alpha$. To guarantee the convexity of the set that is used for search of dominating points the constraint $0.5 \leq \alpha < 1$ is applied, therefore the chance constrained efficiency of DMU_j is defined as follows:

Definition 7. Chance constrained efficiency of DMU_j : DMU_j is chance constrained efficient with probability α, $0.5 \leq \alpha < 1$ in set $T_\varphi \Leftrightarrow (\tilde{x}_j, \tilde{y}_j)$ is point associated with DMU_j and $(\tilde{x}_j, \tilde{y}_j) \in T_\varphi$ is chance constrained efficient production point associated with T_φ.

In the following sections, the α-stochastic efficiency definition will be used to derive almost 100% chance constrained problems. By the specifications of projection direction on the envelopment surface these models will then be oriented and linearized. The same process can be applied to models according to chance constrained efficiency definition, therefore only final version of chance constrained models will be presented.

4.1 Stochastic model

In this section, the derivation of the almost 100% confidence chance constrained problem is reviewed. The derived model is the stochastic equivalent of the additive DEA model and will be the basis for the further theoretical development of SDEA models. In the following subsection, the specific assumptions about the error structure in the data are made and the stochastic model is transformed to its deterministic equivalent.

Now, from the set properties it follows that:

$$\{\tilde{X}\lambda \leq \tilde{x}_j, \tilde{Y}\lambda \geq \tilde{y}_j\} \subset \{e^T(\tilde{X}\lambda - \tilde{x}_j) + e^T(\tilde{y}_j - \tilde{Y}\lambda) < 0\}^3$$

and using probability properties following inequality is derived:

$$Prob(\tilde{X}\lambda \leq \tilde{x}_j, \tilde{Y}\lambda \geq \tilde{y}_j) \leq Prob(e^T(\tilde{X}\lambda - \tilde{x}_j) + e^T(\tilde{y}_j - \tilde{Y}\lambda) < 0).$$

Therefore for λ such that $\varphi(e^T\lambda) = \varphi$ and $\lambda \geq 0$ the condition $Prob(e^T(\tilde{X}\lambda - \tilde{x}_j) + e^T(\tilde{y}_j - \tilde{Y}\lambda) < 0) \leq \alpha$ is a sufficient condition for DMU_j to be α–stochastically efficient.

[3]The inequality type change is due to the additional restriction that $\{\tilde{X}\lambda \leq \tilde{x}_j, \tilde{Y}\lambda \geq \tilde{y}_j\}$ holds with at least one strict inequality.

Using the sufficient condition for α–stochastical efficiency of DMU_j, Cooper, Huang, Lelas, Li and Olesen (1998) constructed the following almost 100% confidence chance constrained problem (in matrix notation) for evaluating the efficiency of DMU_j , $j = 1, \ldots, n$:

$$
\begin{aligned}
\max_{\lambda} \quad & Prob(e^T(\tilde{X}\lambda - \tilde{x}_j) + e^T(\tilde{y}_j - \tilde{Y}\lambda) < 0) - \alpha \qquad (1) \\
s.t. \quad & \\
& Prob({}_i\tilde{x}\lambda < \tilde{x}_{ij}) \geq 1 - \epsilon, \qquad i = 1, \ldots, m; \\
& Prob({}_r\tilde{y}\lambda > \tilde{y}_{rj}) \geq 1 - \epsilon, \qquad r = 1, \ldots, s; \\
& \varphi(e^T\lambda) = \varphi, \\
& \lambda \geq 0,
\end{aligned}
$$

where ϵ is non–Archimedean infinitesimal quantity.[4] The optimal solution of problem (1) is related with stochastic efficient point by using the following theorems:

Theorem 1. Let DMU_j be α-stochastically efficient. The optimal value of objective function in the chance constrained programming problem (1) is less than of equal zero.

Theorem 2. If the optimal value objective functional of problem (1) is greater than zero, then DMU_j in not α-stochastically efficient.[5]

The Theorem (2) implies that if the maximum value of the chance functional $Prob(e^T(\tilde{X}\lambda - \tilde{x}_j) + e^T(\tilde{y}_j - \tilde{Y}\lambda) < 0)$ exceeds α, then DMU_j is not α–stochastically efficient. The value of chance functional of additive model represented by the problem (1) can be used as the simplest efficiency measure when interpreted as the sum of input excess and output slack. In the following sections, the sophisticated efficiency measures will be introduced.

[4]This means that ϵ is less than any positive real number. According to Charnes et al. (1994): "Computational Aspects of DEA", $\epsilon < \min_{j=1,\ldots,n} 1/\sum_{i=1}^{m} x_{ij}$ is selected in the calculations of these models.

[5]These theorems are corollaries of Theorem (3) by Cooper et al. (1998).

4.2 Symmetric shock model

In this subsection, the single symmetric error structure that allows to transform model from chance constrained problem to the linear programming problem is introduced. In the single factor symmetric model, errors in all variables are driven by single symmetric shock ε. Let's consider the structure of m inputs and s outputs of DMU_j in the following form:

$$\tilde{x}_{ij} = \bar{x}_{ij} + a_{ij}\varepsilon \qquad i = 1, \ldots, m;$$
$$\tilde{y}_{ij} = \bar{y}_{ij} + b_{ij}\varepsilon, \qquad r = 1, \ldots, s;$$

where ε follows normal distribution with $E(\varepsilon) = 0, Var(\varepsilon) = \sigma_\varepsilon^2$.[6] To simplify the matrix description of the shock effects to DMUs the following vector notation is introduced:

$$a_j = (a_{1j}, \ldots, a_{mj})^T, \quad b_j = (b_{1j}, \ldots, b_{sj})^T, \quad j = 1, \ldots, n;$$
$${}_ia = (a_{i1}, \ldots, a_{in}), \qquad {}_rb = (b_{r1}, \ldots, b_{rn}), \quad i = 1, \ldots, m,\ r = 1, \ldots, s;$$

and these vectors are aggregated to construct the following matrices of input and output variations:

$$A = (a_1, a_2, \ldots, a_n), B = (b_1, \ldots, b_n).$$

Using properties of normal distribution it is derived that ${}_i\tilde{x}\lambda - \tilde{x}_{ij}$ is distributed $N({}_i\bar{x}\lambda - \bar{x}_{ij}; (({}_ia\lambda - a_{ij})\sigma_\varepsilon)^2)$ and $({}_r\tilde{y}\lambda - \tilde{y}_{rj}) \sim N({}_r\bar{y}\lambda - \bar{y}_{rj}; ((b_{rj} - {}_rb\lambda)\sigma_\varepsilon)^2)$. Applying the inverse of cumulative distribution function $\Phi(\alpha)$ the constraints and objective function in the almost 100% confidence chance constrained problem (1) are rewritten and the following non-stochastic equivalent is derived:

$$\begin{array}{lll}
\min_\lambda & e^T(\bar{X}\lambda - \bar{x}_j) + e^T(\bar{y}_j - \bar{Y}\lambda) + & \\
 & + \mid e^T(A\lambda - a_j) + e^T(\bar{b}_j - \bar{B}\lambda) \mid \sigma_\varepsilon \Phi^{-1}(\alpha) & \\
s.t. & & \\
 & {}_i\bar{x}\lambda \leq \bar{x}_{ij} + \mid {}_ia\lambda - a_{ij} \mid \sigma_\varepsilon \Phi^{-1}(\epsilon), & i = 1, \ldots, m, \\
 & \bar{y}_{rj} \leq {}_r\bar{y}\lambda + \mid b_{rj} - {}_rb\lambda \mid \sigma_\varepsilon \Phi^{-1}(\epsilon), & r = 1, \ldots, s, \\
 & \varphi(e^T\lambda) = \varphi, & \\
 & \lambda \geq 0. &
\end{array} \tag{2}$$

[6] σ_ε^2 is used for purpose of units of measurement scaling. The units of measurement rescaling is used when numerical problems with tiny diagonals of the input-output variance matrices occurs.

To linearize constraints the absolute value terms from constraints in problem (2) are removed. The absolute values term in constraint are decomposed into the difference of two positive numbers and into the objective function the cumulative term $\epsilon(\sum_{r=1}^{s}(q_{1r}+q_{2r})+\sum_{i=1}^{m}(h_{1i}+h_{2i}))$ is added. This constraint modification does not affect the optimal solutions of problem (2). Therefore, problem (2) is equivalent to the following problem with linear constraints:

$$
\begin{array}{lll}
\min_{\lambda,q_{kr},h_{ki}} & e^T(\bar{X}\lambda-\bar{x}_j)+e^T(\bar{y}_j-\bar{Y}\lambda)+ & \\
 & +\mid e^T(A\lambda-a_j)+e^T(b_j-\bar{B}\lambda)\mid\sigma_\varepsilon\Phi^{-1}(\alpha)+ & \\
 & +\epsilon(\sum_{r=1}^{s}(q_{1r}+q_{2r})+\sum_{i=1}^{m}(h_{1i}+h_{2i})) & \\
s.t. & & \\
 & {}_i\bar{x}\lambda\leq\bar{x}_{ij}+(h_{1i}+h_{2i})\sigma_\varepsilon\Phi^{-1}(\epsilon), & \\
 & {}_ia\lambda-a_{ij}=h_{1i}-h_{2i}, & i=1,\ldots,m, \\
 & \bar{y}_{rj}\leq{}_r\bar{y}\lambda+(q_{1r}+q_{2r})\sigma_\varepsilon\Phi^{-1}(\epsilon), & \\
 & b_{rj}-{}_rb\lambda=q_{1r}-q_{2r}, & r=1,\ldots,s, \\
 & \varphi(e^T\lambda)=\varphi, & \\
 & \lambda\geq0,q_{kr}\geq0,h_{ki}\geq0, & k=1,2.
\end{array}
\tag{3}
$$

In the next step, the absolute value from the objective function is removed. The inverse of cumulative distribution function $\Phi(\alpha)$ takes positive or negative value, to account for this factor let's define δ such that:

$$
\delta=\begin{cases}-1 & \text{if } \alpha<0.5;\\ 0 & \text{if } \alpha=0.5;\\ 1 & \text{if } \alpha>0.5.\end{cases}
$$

The absolute value term in the objective function is the sum of the absolute value terms in constraints; therefore, the decomposition that was used in constraints is just substituted in the objective function. Thus as in Li (1998), the absolute value terms are eliminated from objective function and the following problem with linear objective function is obtained:

$$
\begin{aligned}
\min_{\lambda,q_{kr},h_{ki}} \quad & e^T(\bar{X}\lambda - \bar{x}_j) + e^T(\bar{y}_j - \bar{Y}\lambda) + \\
& +\delta(e^T(A\lambda - a_j) + e^T(b_j - \bar{B}\lambda))\sigma_\varepsilon \Phi^{-1}(\alpha) + \\
& +\epsilon(\textstyle\sum_{r=1}^{s}(q_{1r} + q_{2r}) + \sum_{i=1}^{m}(h_{1i} + h_{2i})) \\
s.t. \quad & \\
& {}_i\bar{x}\lambda \le \bar{x}_{ij} + (h_{1i} + h_{2i})\sigma_\varepsilon \Phi^{-1}(\epsilon), \\
& {}_i a\lambda - a_{ij} = h_{1i} - h_{2i}, \qquad i = 1, \ldots, m, \\
& \bar{y}_{rj} \le {}_r\bar{y}\lambda + (q_{1r} + q_{2r})\sigma_\varepsilon \Phi^{-1}(\epsilon), \\
& b_{rj} - {}_r b\lambda = q_{1r} - q_{2r}, \qquad r = 1, \ldots, s, \\
& \varphi(e^T\lambda) = \varphi, \\
& \lambda \ge 0, q_{kr} \ge 0, h_{ki} \ge 0, \qquad k = 1, 2.
\end{aligned}
\tag{4}
$$

Problem (4) is known as the envelopment formulation of the DEA model, because the optimal solution identifies the projected point on to envelopment surface for DMU_j. Now, define e as the vector of ones and the dual problem (5) to primal problem (4), which is a modification of Li (1998) is stated as follows:

$$
\begin{aligned}
\max_{\mu,\nu,\eta,\omega,\psi_j} \quad & \mu^T\bar{y}_j - \nu^T\bar{x}_j - \eta^T b_j - \omega^T a_j - \varphi\psi_j \\
s.t. \quad & \\
& \mu^T\bar{y}_l - \nu^T\bar{x}_l - \eta^T b_l - \omega^T a_l - \varphi\psi_j \le 0, \qquad l = 1, \ldots, n; \\
& -\sigma_\varepsilon \Phi^{-1}(\varepsilon)\mu + \eta \ge -\sigma_\varepsilon(\Phi^{-1}(\varepsilon) + \varepsilon)e - \delta\sigma_\varepsilon \Phi^{-1}(\alpha)e, \\
& -\sigma_\varepsilon \Phi^{-1}(\varepsilon)\mu - \eta \ge -\sigma_\varepsilon(\Phi^{-1}(\varepsilon) + \varepsilon)e + \delta\sigma_\varepsilon \Phi^{-1}(\alpha)e, \\
& -\sigma_\varepsilon \Phi^{-1}(\varepsilon)\nu - \omega \ge -\sigma_\varepsilon(\Phi^{-1}(\varepsilon) + \varepsilon)e - \delta\sigma_\varepsilon \Phi^{-1}(\alpha)e \\
& -\sigma_\varepsilon \Phi^{-1}(\varepsilon)\nu + \omega \ge -\sigma_\varepsilon(\Phi^{-1}(\varepsilon) + \varepsilon)e + \delta\sigma_\varepsilon \Phi^{-1}(\alpha)e, \\
& \mu \ge e, \\
& \nu \ge e, \\
& \eta, \omega, \psi_j \text{ unconstrained.}
\end{aligned}
\tag{5}
$$

For the DMU_j represented by point $(\tilde{x}_j, \tilde{y}_j)$, the stochastic hyperplane:

$$Prob(c^T\tilde{x}_j + d^T\tilde{y}_j + f_j \le 0) = 1 - \epsilon$$

is supporting hyperplane for T_φ at $(\tilde{x}_j, \tilde{y}_j)$ if and only if

$$c^T\tilde{x}_j + d^T\tilde{y}_j + f_j + \Phi^{-1}(\epsilon)\sigma_\varepsilon \mid c^T a_j + d^T b_j \mid = 0 \tag{6}$$

$$\text{and for } \forall\, (\tilde{x}, \tilde{y}) \in T_\varphi :\ c^T\tilde{x} + d^T\tilde{y} + f_j + \Phi^{-1}(\epsilon)\sigma_\varepsilon \mid c^T a_j + d^T b_j \mid \geq 0. \qquad (7)$$

The derived problem (5) is known as the multiplier problem because the optimal solutions $(\mu_j^*, \nu_j^*, \eta_j^*, \omega_j^*, \psi_j^*)$, $j = 1, \ldots, n$, for the set of problems (5) set up the supporting hyperplanes used for production possibility set frontier estimation. The unique optimal solution $(\mu_j^*, \nu_j^*, \eta_j^*, \omega_j^*, \psi_j^*)$ of problem (5) such that

$$\mu_j^{*T}(b_j - b_k) + \nu_j^{*T}(a_j - a_k) - \Phi^{-1}(\epsilon)\sigma_\varepsilon(\mid \mu_j^{*T} b_j - \nu_j^{*T} a_j \mid - \mid \mu_j^{*T} b_k - \nu_j^{*T} b_k \mid) \geq 0,$$

for $k = 1, \ldots, n$, identifies the almost 100% confidence hyperplane
$Prob(\mu_j^{*T}\tilde{y}_j - \nu_j^{*T}\tilde{x}_j + f_j^* \leq 0) = 1 - \epsilon$, where
$f_j^* = -\eta_j^{*T} b_j - \omega_j^{*T} a_j - \varphi\psi_j^* + \Phi^{-1}(\epsilon)\sigma_\varepsilon \mid \mu_j^{*T} b_j - \nu_j^{*T} a_j \mid$.
The indentified almost 100% confidence hyperplane is supporting hyperplane to T_φ at DMU_j. In case without the unique solution of problem (5) the supporting hyperplane for T_φ at $(\tilde{x}_j, \tilde{y}_j)$ is not uniquely identified. The supporting hyperplanes for analyzed DMUs are used to construct the estimate of the production possibility frontier is constructed. In the following sections, the sign of f_j will be related with the returns to scale.

5 Efficiency measure

The solution of problems (4) and (5) consists of a projected point and the supporting hyperplane. The optimal solution of envelopment problem (4) identifies the projected point $(\hat{x}_j, \hat{y}_j) = (\bar{X}\lambda_j^*, \bar{Y}\lambda_j^*)$ on the supporting hyperplane adjacent to DMU_j. The point $(\bar{X}\lambda_j^*, \bar{Y}\lambda_j^*)$ is the optimal solution of the envelopment problem (4) and it is element of the hyperplane obtained as the solution of the multipliers problem (5). In this section, the projected point and envelopment defining hyperplane will be utilized to create the inefficiency measure for the DMU_j.

For the purpose of a simple inefficiency measure creation, the distance measure of a discrepancy between the expected and projected point can be used. Therefore the simple inefficiency measure is $(\hat{x}_j, \hat{y}_j) - (\bar{x}_j, \bar{y}_j)$. This discrepancy measure expresses the difference between the efficient frontier represented by the projected point $(\hat{x}_j, \hat{y}_j)$ and the present position of DMU_j. Starting from $(\bar{x}_j, \bar{y}_j)$ different projection paths on the corresponding part of envelopment surface can be followed. Therefore various projected points and inefficiency measures can be created. The choice of projection path means specification of the type of strategy that must to be used to enhance efficiency. In the following derivation of efficiency measure, directions of lowering inputs and increasing in outputs will be distinguished.

First, for inputs of DMU_j let's denote $\bar{x}_j - \bar{X}\lambda = e_j$ and following the path $-e_j$ the inputs can be decreased and the projected point is moved towards the production possibility frontier. This projection direction is pictured in Figure (5) as the input reduction direction. The point DMU5i is the input oriented projection of DMU#5. When the path of input reduction is used in the construction of point projected on the production possibility frontier the input oriented DEA model is created.

Similarly, the DEA output oriented model is created when for projection on the production possibility frontier the path $s_j = \bar{Y}\lambda - \bar{y}_j$ is used. The path s_j projects DMU_j on to production possibility frontier in outputs augmenting direction. This projection point is shown on Figure (5) as the point DMU5o.

Next, to determine the maximal scale effects in inputs reduction or outputs augmentation, the projection paths s_j, e_j are decomposed to a proportional increase (decrease) of output (input) and residual as follows:
$s_j = \rho\bar{y}_j + \delta_s^j$, $e_j = \gamma\bar{x}_j + \delta_e^j$, where proportional increase of outputs ρ and proportional decrease of inputs γ are defined as follows:

$$\begin{aligned} \rho = &\quad \min_{r=1,\ldots,s} \quad \frac{\hat{y}_{rj} - \bar{y}_{rj}}{\bar{y}_{rj}} \geq 0, \\ \gamma = &\quad \min_{i=1,\ldots,m} \quad \frac{\bar{x}_{ij} - \hat{x}_{ij}}{\bar{x}_{ij}} \geq 0, \end{aligned} \tag{8}$$

and $\delta_e^j \geq 0, \delta_s^j \geq 0,\ j = 1, \ldots, n.$[7]

Next, the notation of Ali and Seiford (1993) is used to define new variables for the proportional part of projection. The new variable for output oriented model are defined as $\phi = 1 + \rho$ and for input oriented model $\theta = 1 - \gamma$. From construction of the scaling parameters, the optimal value of θ (ϕ for input output problem) satisfies

[7]Note that at least one component of each δ is zero, because of projection on to the production possibility frontier.

$\theta \leq 1$ ($\phi \geq 1$). The maximal output scale effect is identified by maximal ϕ and the maximal input reduction the minimal θ.

For the identification of scale effects and efficiency evaluation two stage models are constructed. In the first model stage, the maximal ϕ or minimal θ is found to identify the maximal proportional effect. In the second stage of modelling, the identified scale effect is utilized to optimize for envelope distance and the DMU's efficiency is evaluated. Thus, in the first stage the $\bar{x}_j$ is replaced with $\theta\bar{x}_j$ in constraints and the maximal input reduction is identified by the minimal value of θ. The optimal solution to the first stage for DMU_j is denoted as $\hat{\theta}_j$. Then the $\hat{\theta}_j$ is used in the second stage where the $\bar{x}_j$ is replaced by $\hat{\theta}_j\bar{x}_j$ in problem (1) and the efficiency score is evaluated. Similarly, for the input oriented model the $\bar{y}_j$ is replaced by $\phi\bar{y}_j$ and the maximal output augmentation is identified by $\hat{\phi}_j$. As in the input oriented case, the efficiency of the DMU_j is evaluated by model where the $\bar{y}_j$ is replaced by $\hat{\phi}_j\bar{y}_j$ in problem (1). These two stage models are summarized in the Table (2).

In these models, the values of $\hat{\phi}_j^{-1}$ and $\hat{\theta}_j$ can be used as the inefficiency measures. For $\hat{\phi}_j = 1$ ($\hat{\theta}_j = 1$) the point is the boundary point of T_φ but not necessarily it represents the efficient point. The DMU_j is identified as efficient if for proportional scaling parameter $\hat{\phi}_j = 1$ ($\hat{\theta}_j = 1$) and the second stage models identifies DMU_j as α-stochastically efficient. The additional condition is interpreted as the sum of slacks and for α-stochastic efficiency it is required that it holds with probability $1-\alpha$ and using this condition for efficiency evaluation a class of weakly efficient points is defined. Where the as the weakly efficient point is considered such point that fulfills only the condition on optimal value of the proportional scaling parameter $\hat{\phi}_j = 1$ ($\hat{\theta}_j = 1$).

6 Oriented SDEA models

Two stage oriented models described in Table (2) can be collapsed to one stage model. In both stages the objective function optimization is subject to the same constraints, the only difference being the objective function. To merge these stages in one optimization problem, the non-Archimedean ϵ is used as weight for second stage objective function. The presence of non-Archimedean ϵ in the objective function allows such proportional movement towards the frontier that it drive out the slacks optimization. In the following sections almost 100% confidence chance constrained oriented models will be defined and their linearization will be presented.

Output oriented model In this section, the almost 100% confidence chance constrained output oriented problem is linearized. The one stage model, derived from two stage optimization model presented in Table (2) is stated as follows:

$$
\begin{aligned}
\max_{\lambda,\phi} \quad & \phi + \epsilon(Prob(e^T(\tilde{X}\lambda - \tilde{x}_j) + e^T(\phi\tilde{y}_j - \tilde{Y}\lambda)) - \alpha) && && (9)\\
s.t. \quad & Prob({}_i\tilde{x}\lambda < \tilde{x}_{ij}) \geq 1 - \epsilon, && i = 1, \ldots, m; \\
& Prob({}_r\tilde{y}\lambda > \phi\tilde{y}_{rj}) \geq 1 - \epsilon, && r = 1, \ldots, s; \\
& \varphi(e^T\lambda) = \varphi \\
& \lambda \geq 0;
\end{aligned}
$$

After, the same linearization procedure, applied to problem (1) and described in the previous section is used, the following model is derived:

$$
\begin{aligned}
\max_{\lambda,q_{kr},h_{ki},\phi} \quad & \phi - \epsilon[e^T(\bar{X}\lambda - \bar{x}_j) + e^T(\phi\bar{y}_j - \bar{Y}\lambda) + \\
& +\delta(e^T(A\lambda - a_j) + e^T(\phi\bar{b}_j - \bar{B}\lambda))\sigma_\varepsilon\Phi^{-1}(\alpha)] + \\
& +\epsilon(\textstyle\sum_{r=1}^{s}(q_{1r} + q_{2r}) + \sum_{i=1}^{m}(h_{1i} + h_{2i})) \\
s.t. \quad & {}_i\bar{x}\lambda \leq \bar{x}_{ij} + (h_{1i} + h_{2i})\sigma_\varepsilon\Phi^{-1}(\epsilon), \\
& {}_i a\lambda - a_{ij} = h_{1i} - h_{2i}, && i = 1, \ldots, m, && (10) \\
& \phi\bar{y}_{rj} \leq {}_r\bar{y}\lambda + (q_{1r} + q_{2r})\sigma_\varepsilon\Phi^{-1}(\epsilon), \\
& \phi b_{rj} - {}_r b\lambda = q_{1r} - q_{2r}, && r = 1, \ldots, s, \\
& \varphi(e^T\lambda) = \varphi, \\
& \lambda \geq 0, q_{kr} \geq 0, h_{ki} \geq 0, && k = 1, 2,
\end{aligned}
$$

Input oriented model Similarly, as for the output oriented model represented by problem (9), the almost 100% confidence chance constrained input oriented model will be derived. The one stage model is created by merging the two stage optimization process presented in Table (2). The one stage model is stated as follows:

$$
\begin{aligned}
\min_{\lambda,\theta} \quad & \theta - \epsilon(Prob(e^T(\tilde{X}\lambda - \theta\tilde{x}_j) + e^T(\tilde{y}_j - \tilde{Y}\lambda)) - \alpha) && && (11)\\
s.t. \quad & Prob({}_i\tilde{x}\lambda < \theta\tilde{x}_{ij}) \geq 1 - \epsilon, && i = 1, \ldots, m; &&\\
& Prob({}_r\tilde{y}\lambda > \tilde{y}_{rj}) \geq 1 - \epsilon, && r = 1, \ldots, s; &&\\
& \varphi(e^T\lambda) = \varphi && &&\\
& \lambda \geq 0; && &&
\end{aligned}
$$

Applying the same linearization procedure as for the output oriented model the following linear deterministic equivalent of the input oriented almost 100% chance constrained model is derived:

$$
\begin{aligned}
\min_{\lambda,q_{kr},h_{ki},\theta} \quad & \theta + \epsilon[e^T(\bar{X}\lambda - \theta\bar{x}_j) + e^T(\bar{y}_j - \bar{Y}\lambda)+ && &&\\
& +\delta(e^T(A\lambda - \theta a_j) + e^T(\bar{b}_j - \bar{B}\lambda))\sigma_\varepsilon\Phi^{-1}(\alpha)]+ && &&\\
& +\epsilon(\textstyle\sum_{r=1}^{s}(q_{1r} + q_{2r}) + \sum_{i=1}^{m}(h_{1i} + h_{2i})) && &&\\
s.t. \quad & {}_i\bar{x}\lambda \leq \theta\bar{x}_{ij} + (h_{1i} + h_{2i})\sigma_\varepsilon\Phi^{-1}(\epsilon), && &&\\
& {}_ia\lambda - \theta a_{ij} = h_{1i} - h_{2i}, && i = 1, \ldots, m, && (12)\\
& \bar{y}_j\lambda \leq {}_r\bar{y} + (q_{1r} + q_{2r})\sigma_\varepsilon\Phi^{-1}(\epsilon), && &&\\
& b_{rj} - {}_rb\lambda = q_{1r} - q_{2r}, && r = 1, \ldots, s, &&\\
& \varphi(e^T\lambda) = \varphi, && &&\\
& \lambda \geq 0, q_{kr} \geq 0, h_{ki} \geq 0, && k = 1, 2, &&
\end{aligned}
$$

The optimal solution $(\lambda^*, q^*_{kr}, h^*_{ki}, \phi^*)$ of problem (10) ($(\lambda^*, q^*_{kr}, h^*_{ki}, \theta^*)$ for problem (12)) is used to evaluate the efficiency of DMU_j . When the output (input) oriented model is used the DMU_j is α-stochastically efficient if the following two conditions are satisfied:

1. $\phi^* = 1$ ($\theta^* = 1$);
2. $e^T(\bar{X}\lambda^* - \bar{x}_j) + e^T(\phi^*\bar{y}_j - \bar{Y}\lambda^*) + |e^T(A\lambda^* - a_j) + e^T(\phi^*\bar{b}_j - \bar{B}\lambda^*)|\sigma_\varepsilon\Phi^{-1}(\alpha) \geq 0$ $(e^T(\bar{X}\lambda^* - \theta^*\bar{x}_j) + e^T(\bar{y}_j - \bar{Y}\lambda^*) + |e^T(A\lambda^* - \theta^*a_j) + e^T(\bar{b}_j - \bar{B}\lambda^*)|\sigma_\varepsilon\Phi^{-1}(\alpha) \geq 0)$.

As it was mentioned in the section on efficiency measure introduction a class of weak efficient DMUs can be defined. The analyzed DMU_j is identified as weak efficient when the optimal solution of associated problem satisfies $\phi^* = 1$ or $\theta^* = 1$.

7 Chance constrained DEA model

As in the section on almost 100% chance constrained models, inputs and outputs are considered to be jointly normally distributed and the following chance constrained version of DEA model is constructed:

$$\begin{aligned}
\max_{\lambda} \quad & e^T(\bar{X}\lambda - \bar{x}_j) + e^T(\bar{y}_j - \bar{Y}\lambda) \\
s.t. \quad & \\
& Prob({}_i\tilde{x}\lambda < \tilde{x}_{ij}) \geq 1 - \alpha, \quad i = 1, \ldots, m; \\
& Prob({}_r\tilde{y}\lambda > \tilde{y}_{rj}) \geq 1 - \alpha, \quad r = 1, \ldots, s; \\
& \varphi(e^T\lambda) = \varphi \\
& \lambda \geq 0;
\end{aligned} \tag{13}$$

This problem relates to the definition of chance constrained efficiency domination in the fourth section. Applying the same procedure as for almost 100% chance constrained problems the two stage problems are derived. These two stage chance constrained problems are summarized in the Table (3).

As for problem (1) the dual problem to problem (13) can be derived and the optimal solutions used to identify the stochastic supporting hyperplanes to analyzed DMUs. Using the identified stochastic supporting hyperplanes the production possibility frontier can be estimated.

The same linearization procedure as was used to linearize problem (1) and described in previous section, is applied after the two stage problem is merged in one optimization problem. The following oriented and linearized chance constrained models are derived:

Output oriented model

$$\begin{array}{lll}
\max_{\lambda,q_{kr},h_{ki},\phi} & \phi - \epsilon(e^T(\bar{X}\lambda - \bar{x}_j) + e^T(\phi\bar{y}_j - \bar{Y}\lambda) + & \\
& +\epsilon(\sum_{r=1}^{s}(q_{1r}+q_{2r}) + \sum_{i=1}^{m}(h_{1i}+h_{2i})) & \\
s.t. & {}_i\bar{x}\lambda \leq \bar{x}_{ij} + (h_{1i}+h_{2i})\sigma_\varepsilon\Phi^{-1}(\alpha), & i=1,\dots,m, \\
& {}_ia\lambda - a_{ij} = h_{1i} - h_{2i}, & i=1,\dots,m, \\
& \bar{y}_j\lambda \leq \phi_r\bar{y} + (q_{1r}+q_{2r})\sigma_\varepsilon\Phi^{-1}(\alpha), & r=1,\dots,s, \\
& \phi b_{rj} -_r b\lambda = q_{1r} - q_{2r}, & r=1,\dots,s, \\
& \varphi(e^T\lambda) = \varphi, & \\
& \lambda \geq 0, q_{kr} \geq 0, h_{ki} \geq 0, & k=1,2, \\
& & i=1,\dots,m, \\
& & r=1,\dots,s.
\end{array} \tag{14}$$

Input oriented model

$$\begin{array}{lll}
\min_{\lambda,q_{kr},h_{ki},\theta} & \theta + \epsilon(e^T(\bar{X}\lambda - \theta\bar{x}_j) + e^T(\bar{y}_j - \bar{Y}\lambda)) + & \\
& +\epsilon(\sum_{r=1}^{s}(q_{1r}+q_{2r}) + \sum_{i=1}^{m}(h_{1i}+h_{2i})) & \\
s.t. & {}_i\bar{x}\lambda \leq \theta\bar{x}_{ij} + (h_{1i}+h_{2i})\sigma_\varepsilon\Phi^{-1}(\alpha), & i=1,\dots,m, \\
& {}_ia\lambda - \theta a_{ij} = h_{1i} - h_{2i}, & i=1,\dots,m, \\
& \bar{y}_j\lambda \leq_r \bar{y} + (q_{1r}+q_{2r})\sigma_\varepsilon\Phi^{-1}(\alpha), & r=1,\dots,s, \\
& b_{rj} -_r b\lambda = q_{1r} - q_{2r}, & r=1,\dots,s, \\
& \varphi(e^T\lambda) = \varphi, & \\
& \lambda \geq 0, q_{kr} \geq 0, h_{ki} \geq 0, & k=1,2, \\
& & i=1,\dots,m, \\
& & r=1,\dots,s.
\end{array} \tag{15}$$

Similarly, as for Problems (12) and (10), the optimal solution $(\lambda^*, q^*_{kr}, h^*_{ki}, \phi^*)$ of problem (14) ($(\lambda^*, q^*_{kr}, h^*_{ki}, \theta^*)$ for problem (15)) can be used to evaluate the efficiency of DMU_j as in the previous section.

The DMU_j is chance constrained efficient if the following two conditions are satisfied:

1. $\phi^* = 1$ $(\theta = 1)$;

2. All expected values of slacks are zero,
   $e^T(\bar{X}\lambda - \bar{x}_j) = 0$ and $e^T(\phi\bar{y}_j - \bar{Y}\lambda) = 0$
   $(e^T(\bar{X}\lambda - \theta\bar{x}_j) = 0$ and $e^T(\bar{y}_j - \bar{Y}\lambda) = 0)$

To simplify the evaluation of efficiency score the following two efficiency measures for stochastic models which are stochastic equivalents for measures introduced by Tone (1993), are proposed:

$$\begin{aligned}
\text{Input oriented:} \quad \chi_j &= \left(\theta^* + \frac{e^T(\bar{X}\lambda^* - \theta^*\bar{x}_j)}{e^T\bar{x}_j}\right)\frac{e^T\bar{y}_j}{e^T\bar{y}_j - e^T(\bar{y}_j - \bar{Y}\lambda^*)} \\
\text{Output oriented:} \quad \tau_j^{-1} &= \left(\phi^* - \frac{e^T(\phi^*\bar{y}_j - \bar{Y}\lambda^*))}{e^T\bar{y}_j}\right)\frac{e^T\bar{x}_j}{e^T\bar{x}_j + e^T(\phi^*\bar{y}_j - \bar{Y}\lambda^*)}
\end{aligned}$$

The proposed efficiency measures τ and χ have following properties:

1. $0 \leq \tau_j, \chi_j \leq 1$

2. $\chi_j = 1,\ \tau_j = 1 \Leftrightarrow DMU_j$ is chance constrained efficient

3. τ_j and χ_j are units invariant measures

4. τ_j and χ_j are monotonic increasing in inputs and outputs

5. τ_j and χ_j are decreasing in the relative values of the slacks

6. $\tau_j = \phi^*,\ \chi_j = \theta^* \Leftrightarrow$ the expected values of all slacks are zero.

These measures make it easier to evaluate the efficiency score of DMU_j because they take into account the values of maximal proportional increase and the slacks (residuals) values.

8 Introducing returns to scale

Further development of aforementioned DEA models requires the inclusion of the returns to scale type to the model specification. As it is mentioned in the literature review section the CCR model was used to analyze the set of DMUs that were using the production function with constant returns to scale. The BCC model and its variations were developed to analyze the production function with variable returns to scale. The same methodology will be used here to develop the stochastic DEA models with variable (VRS), non-increasing and non-decreasing returns to scale. To introduce the returns to scale in to the stochastic the concept by Banker et al. (1984) is used in terms of expected values in the following definition:

Definition 8. Returns to scale. Let DMU_j is stochastically efficient and the point $Z_\delta = ((1+\delta)\bar{x}_j, (1+\delta)\bar{y}_j)$ is point in δ–neighborhood of $(\bar{x}_j, \bar{y}_j)$:

- The Non-Decreasing returns to scale are present $\Leftrightarrow \exists\, \delta^* > 0$ such that $Z_\delta \in T_\varphi$ for $\delta^* > \delta \geq 0$ and $Z_\delta \notin T_\varphi$ for $-\delta^* < \delta < 0$
- The Constant returns to scale are present $\Leftrightarrow \exists\, \delta^* > 0$ such that $Z_\delta \in T_\varphi$ for $\mid \delta \mid < \delta^*$
- The Non-Increasing returns to scale are present $\Leftrightarrow \exists\, \delta^* > 0$ such that $Z_\delta \notin T_\varphi$ for $\delta^* > \delta \geq 0$ and $Z_\delta \in T_\varphi$ for $-\delta^* < \delta < 0$

The various types of returns to scale are reflected by different shapes of the production possibility set frontier that is set up by intersection of supporting hyperplanes identified by solutions of multiplier problems. In the case of constant returns to scale (CCR model by Charnes et al. (1978)) the envelopment surface consist of a single half line that passes through origin as it is shown in Figure (3). Figure (3) also presents the production possibility frontier of the model with the variable returns to scale that is referred to as the BCC model. In the case of VRS presence the frontier is piecewise linear set. These types of returns to scale of production possibility set are parameterized via the selection of φ and constraint type associated with the φ as follows:

$$\varphi = \begin{cases} 0 & \text{Constant returns to scale (CCR model)} \\ 1 & \text{Variable returns to scale (BCC model)} \end{cases}$$

Since the α-stochastically efficient point $(\tilde{x}_j, \tilde{y}_j)$ satisfies condition (6) for the point $Z_\delta = ((1+\delta)\bar{x}_j, (1+\delta)\bar{y}_j)$ can be derived:

$$\begin{aligned} c^T(1+\delta)\tilde{x}_j + d^T(1+\delta)\tilde{y}_j + f_j + (1+\delta)\Phi^{-1}(\epsilon)\sigma_\varepsilon \mid c^T a_j + d^T b_j \mid \; &= \\ = (1+\delta)(c^T\tilde{x}_j + d^T\tilde{y}_j + f_j + \Phi^{-1}(\epsilon)\sigma_\varepsilon \mid c^T a_j + d^T b_j \mid) - \delta f_j = -\delta f_j. \qquad (16) \end{aligned}$$

Therefore the point $Z_\delta \in T_\varphi$ only if and only if the $-\delta f_j \geq 0$. Employing definition (8) the relation between returns to scale and the sign of f_j reveals. The returns to scale type is reflected in the type of the constraint $\varphi(e^T\lambda) = \varphi$ for $\varphi = 1$. The Table (1) summarizes the constraint variations and relates the constraint type on λ, the returns to scale type and frontier hyperplane characteristic.

9 Summary of models

In the previous sections the various models were derived and related to definitions of α-stochastic and chance constrained efficiency dominance. Therefore, the Table (4) that summarizes these models is presented. It is evident that the models based on different efficiency dominance definition lead to different evaluation of efficiency. It should be stressed that even the models using the same efficiency dominance definition but with different orientation choice lead to different results. Therefore, the choice of the efficiency dominance type, returns to scale and projection path to the envelopment surface (the set of dominating points in the production possibility set) is the crucial choice for the efficiency analysis. The important choice for efficiency score evaluation is the choice of efficiency measure.

The returns to scale choice affects the shape of the production possibility set envelopment. The restrictions on returns to scale are related to four types of the envelopment surface shape through the geometry of the production possibility set and these restrictions interpreted as the restriction on λ in the envelopment problem or a restriction on supporting hyperplanes in the multiplier problem.

The evaluation of efficiency score is based on distance measurement between the point that represent DMU and the associated point on the envelopment surface. This distance measure used in additive models is the most simple efficiency measure. More sophisticated efficiency measure is created using the measure of maximal proportional inputs reduction (output augmentation) while keeping the levels of outputs (inputs) fixed. This proportional input (output) scaling approach is interpreted as the selection of projection path towards the envelopment surface and results to creation of oriented SDEA models.

The use of Non-Archimedean infinitesimal, ϵ is closely related to unit invariance property of the objective function values of derived models, because the result of multiplication by ϵ is not unit dependent and its role is to distinguish between the efficient and inefficient DMUs which are elements of the production possibility set boundary. The use of unit invariant models also delivers the possibility of units of measurement change when numerical problems (e.g.: tiny diagonal matrices) are expected to arise when the models are solved.

In the Table (4), presented SDEA are compared only with the most popular DEA models that appear in the present literature. The additional SDEA models can be created as the extensions of models covered in this paper using the extensions procedures for DEA models.

10 Method for SDEA model solving

To solve the linear problems associated with SDEA models the interior point method (IPM) is used. IPM is used because it is less computational costly than the simplex methods for large sized problems. The second reason for IPM use is that the IPM solutions satisfy the strong complementarity slackness condition (SCSC). The SCSC solution is solution with maximal product of the positive components of the optimal solution. Therefore, SCSC solutions are optimal solutions with minimal number of zero components. In the case of not unique optimal solution, the SCSC solution is not the vertex of the optimal solution set as it is in the case when the simplex algorithm is used. Simplex type algorithm searches for the optimal solution among the vertices of the feasibility set, therefore the optimal solution is a vertex (vertices). The IPM generates the infinite sequence of points that converges to a optimal solution and the iteration process stops when the iterates are sufficiently close to the optimal solution.

For the purpose of IPM use the linearized problems can be easily transformed to the standard linear programming form:

$$\begin{array}{llll} \text{Primal:} & \min \mathbf{c}^T \mathbf{x} & \text{Dual:} & \max \mathbf{b}^T \mathbf{y} \\ & \text{s.t. } \mathbf{A}\mathbf{x} = \mathbf{b}, \mathbf{x} \geq 0 & & \text{s.t. } \mathbf{A}^T \mathbf{y} + \mathbf{z} = \mathbf{c}, \mathbf{z} \geq 0. \end{array} \tag{17}$$

In the case of linearized stochastic problems, vectors $\mathbf{x}$, $\mathbf{c}$, $\mathbf{z} \in \mathbb{R}^{n+3(m+s)+1}$; vectors $\mathbf{y}$, $\mathbf{b} \in \mathbb{R}^{2(m+s)+1}$ and matrix $\mathbf{A} \in \mathbb{R}^{(2(m+s)+1)\times(n+3(m+s)+1)}$. When using the primal-dual version of interior point method the primal and dual problem (17) are solved simultaneously. The optimal solution of linear programming problem associated with DEA model consists of envelopment and multiplier problem solutions and these solutions are utilized to evaluate the efficiency of analyzed DMU and identification of the adjacent supporting hyperplane.

Using the complementarity constraint $\mathbf{z}^T\mathbf{x} = 0$ (equivalent to duality gap condition $\mathbf{c}^T\mathbf{x} - \mathbf{b}^T\mathbf{y} = 0$) together with the feasibility constraints the following optimality condition for problem (17) is stated:

$$\begin{pmatrix} \mathbf{Ax} - \mathbf{b} \\ \mathbf{A}^T\mathbf{y} + \mathbf{z} - \mathbf{c} \\ \mathbf{z}^T\mathbf{x} \end{pmatrix} = \begin{pmatrix} \mathbf{0} \\ \mathbf{0} \\ 0 \end{pmatrix}, \tag{18}$$

where $\mathbf{z}, \mathbf{x} \geq 0$. The strict complementarity solution to problem (18) is such solution $(\mathbf{x}, \mathbf{z}, \mathbf{y})$ that $x_i + z_i > 0, \forall i = 1, \ldots, n+3(m+s)+1$. In the case that the problem (18) solution set is not a singleton set (the problem (17) does not have unique solution) the set of strict complementarity solutions is equal to the set of SCSC solutions. In the case of unique solution of problem (17), the unique solution to problem (18) is the SCSC solution.

To solve problem (18) usually a variant of Newton's method is used. The created solver uses the Mehrotra's predictor-corrector algorithm that belongs to the class of path following IPM algorithms. The algorithm uses the combination of Newton's direction (gap decreasing direction) and centering direction to solve the sequence of Problems (18), where the complementarity constraint is modified to $\mathbf{x}_k^T\mathbf{z}_k = \mu_k$ and sequence μ_k converges to 0 for $k \to \infty$. The iteration process stops if the problem is solved with desired accuracy or the limit for number of iterations is reached.

The solver for proposed SDEA models is constructed using the procedures package known as PCx linear solver obtained from Optimization Technology Center at Argonne National Laboratory and Northwestern University.

11 Conclusions

The major contribution of this theoretical paper is the development of four oriented stochastic DEA models and description of their properties. Using the techniques of stochastic problems linearization the proposed stochastic models were linearized, so the interior point methods for linear problems can be used to solve linear programming problems associated with the models.

The created solver for problems associated with the SDEA models, uses primal-dual interior point method algorithm and both primal and dual solution are utilized in efficiency evaluation and estimation of production possibility frontier.

The planed application of SDEA methodology is to compare the productivity analysis results obtained by parametric methods to results obtained by SDEA.

References

Ali, Agha Iqbal and Lawrence M. Seiford, *The measurement of productive efficiency: Techniques and Applications*, New York: Oxford University Press,

Banker, R. D., Abraham Charnes, and William W. Cooper, "Some Models for Estimating Technical and Scale Inefficiencies in Data Envelopment Analysis," *Management Science*, 1984, *30*, 1078–192.

Byrnes, Patricia and Vivian Valdmanis, *Cost Analysis Applications Of Economics and Operation Research*, Berlin: Springer Verlag, 1989.

Charnes, Abraham, William W. Cooper, and E. Rhodes, "Measuring the efficiency of decision making units.," *European Journal of Operational Research*, 1978, *2*, 429–444.

___, ___, Arie Y. Lewin, and Lawrence M. Seiford, *Data Envelopment Analysis: Theory, Methodology and Applications*, Kluwer Academic Publishers, 1994.

Cooper, William W., Zhimin Huang, Vedran Lelas, Susan X. Li, and Ole B. Olesen, "Chance Constrained Programming Formulations for Stochastic Characterizations of Efficiency and Dominance in DEA," *Journal of Productivity Analysis*, 1998, *9*, 53–79.

Emrouznejad, Ali, "Ali Emrouznejad's DEA HomePage," web page 1995-2001. Warwick Business School, Coventry CV4 7AL, UK.

Gonzales-Lima, Maria D., Richard A. Tapia, and Robert M. Thrall, "On the construction of strong complementarity slackness solutions for DEA linear programming problems using a primal–dual interior–point method," *Annals of Operations Research*, 1996, *66*, 139–162.

Gstach, Dieter, "Another approach to data envelopment analysis in noisy environments: DEA+," *Journal of Productivity Analysis*, 1998, *9*, 161–176.

Halme, Merja and Pekka Korhonen, "Restricting Weights In Value Efficiency Analysis," October 1999.

Huang, Zhimin and Susan X. Li, "Chance Constrained Programming And Stochastic Dea Models." School of Business, Adelphi University, Garden City, New York.

Land, K.C., C.A.K Lovell, and S. Thore, "Chance-constrained Data Envelopment Analysis," *Managerial and Decision Economics*, 1993, *14*, 541–554.

Li, Susan X., "Stochastic models and variable returns to scales in data envelopmnet analysis," *European Journal of Operational Research*, 1998, *104*, 532–548.

Olesen, O. B., "Comparing and Combining Two Approaches for Chance Constrained DEA," Technical Report, The University of Southern Denmark December 2002.

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

Tone, Kaoru, "An Epsilon-Free DEA and a New Measure of Efficiency.," *Journal Of The Operations Research Society Of Japan*, 1993, *36* (3), 167–174.

Ševčovič, Daniel, Margaréta Halická, and Pavol Brunovský, "DEA analysis for a large structured bank branch network," *Central European Journal of Operations Research*, 2001, *9* (4), 329–343.

Walden, John B. and James E. Kirkley, "Measuring Technical Efficiency and Capacity in Fisheries by Data Envelopment Analysis Using the General Algebraic Modeling System (GAMS): A Workbook," report, National Marine Fisheries Service, Woods Hole, Massachusetts October 2000.

A Figures and Tables

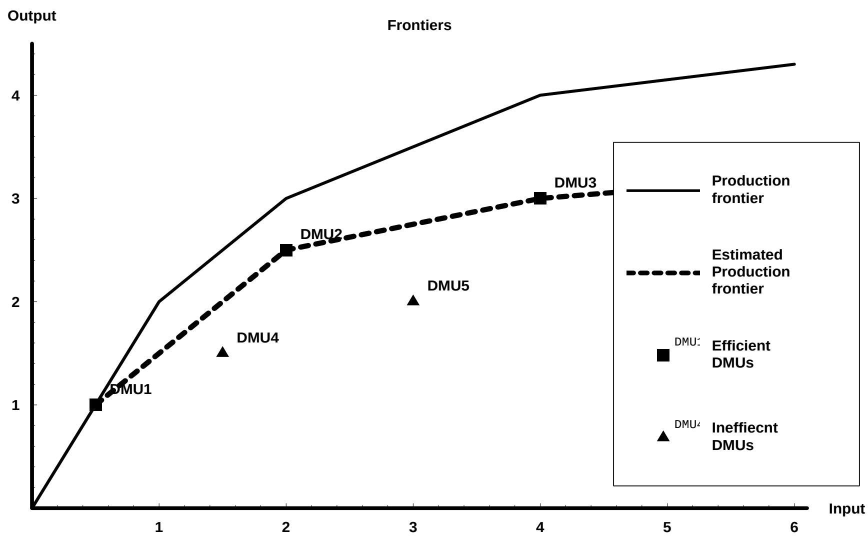

Figure 1: DEA production possibility frontier

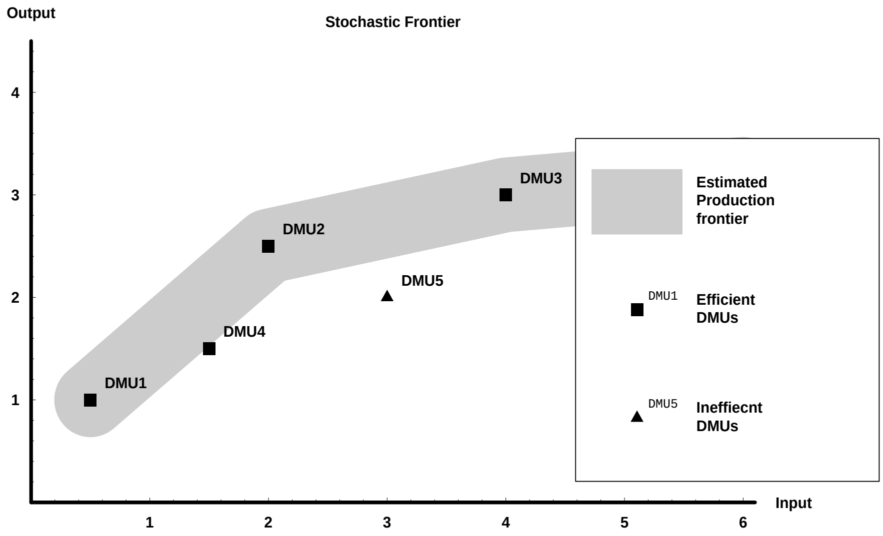

Figure 2: Stochastic production possibility frontier

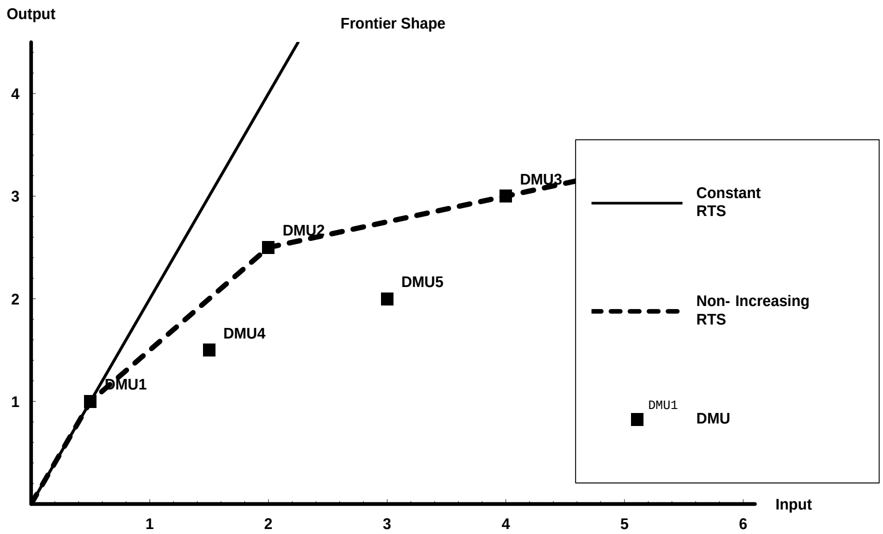

Figure 3: Returns to scale - Constant, Non-Increasing

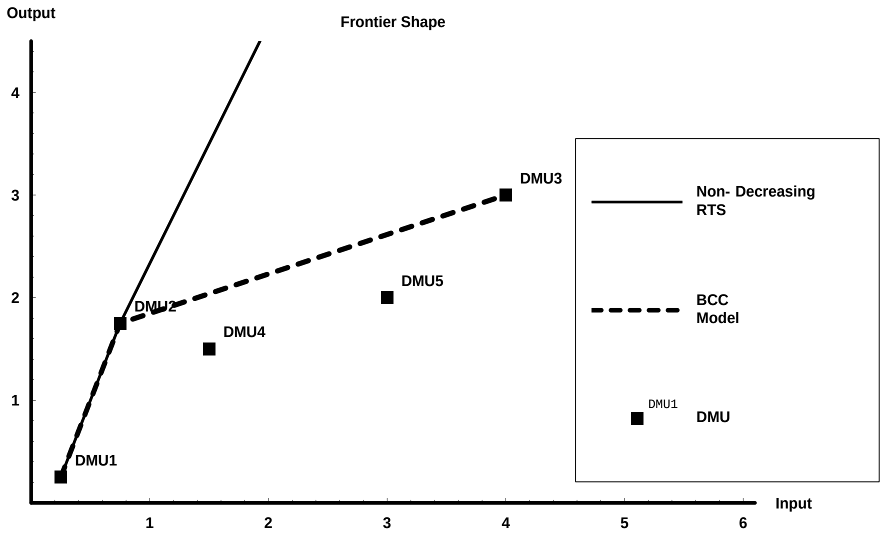

Figure 4: Returns to scale - Non-Decreasing, BCC model

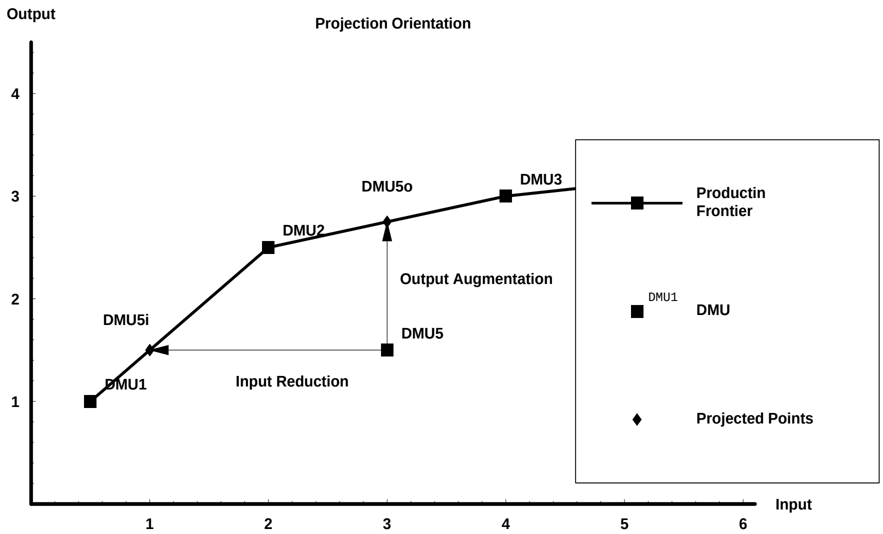

Figure 5: Projection on the production possibility frontier

Model (Orientation)	Returns to scale	Constraint	Hyperplane
CCR model			
(Input, Output)	Constant	None, $\varphi = 0$	Passes trough origin
BCC model			
(Input, Output)	Variable	$e^T\lambda = 1$	Not constrained
SDEA models			
(Input)	Non-Decreasing	$e^T\lambda \geq 1$	$f_j^* \geq 0$
(Input)	Non-Increasing	$e^T\lambda \leq 1$	$f_j^* \leq 0$
(Input)	Constant	None	$f_j^* = 0$
(Output)	Non-Decreasing	$e^T\lambda \geq 1$	$f_j^* \leq 0$
(Output)	Non-Increasing	$e^T\lambda \leq 1$	$f_j^* \geq 0$
(Output)	Constant	None	$f_j^* = 0$

Table 1: Returns to scale

Output oriented model

First stage	Second stage
$\max_{\lambda,\phi} \phi$	$\max_{\lambda} Prob(e^T(\tilde{X}\lambda - \tilde{x}_j) + e^T(\hat{\phi}_j \tilde{y}_j - \tilde{Y}\lambda)) - \alpha$
s.t. $Prob({}_i\tilde{x}\lambda < \tilde{x}_{ij}) \geq 1 - \epsilon$	s.t. $Prob({}_i\tilde{x}\lambda < \tilde{x}_{ij}) \geq 1 - \epsilon$
$Prob({}_r\tilde{y}\lambda > \phi\tilde{y}_{rj}) \geq 1 - \epsilon$	$Prob({}_r\tilde{y}\lambda > \hat{\phi}_j \tilde{y}_{rj}) \geq 1 - \epsilon$
$\varphi(e^T\lambda) = \varphi$	$\varphi(e^T\lambda) = \varphi$
$\lambda \geq 0$	$\lambda \geq 0$
	$i = 1, \ldots, m;\ r = 1, \ldots, s.$

Input oriented model

First stage	Second stage
$\min_{\lambda,\theta} \theta$	$\max_{\lambda} Prob(e^T(\tilde{X}\lambda - \hat{\theta}_j \tilde{x}_j) + e^T(\tilde{y}_j - \tilde{Y}\lambda)) - \alpha$
s.t. $Prob({}_i\tilde{x}\lambda < \theta_j \tilde{x}_{ij}) \geq 1 - \epsilon$	s.t. $Prob({}_i\tilde{x}\lambda < \hat{\theta}_j \tilde{x}_{ij}) \geq 1 - \epsilon$
$Prob({}_r\tilde{y}\lambda > \tilde{y}_{rj}) \geq 1 - \epsilon$	$Prob({}_r\tilde{y}\lambda > \tilde{y}_{rj}) \geq 1 - \epsilon$
$\varphi(e^T\lambda) = \varphi$	$\varphi(e^T\lambda) = \varphi$
$\lambda \geq 0$	$\lambda \geq 0$
	$i = 1, \ldots, m;\ r = 1, \ldots, s.$

Table 2: Two stages of oriented almost 100% confidence chance constrained models

Output oriented model

First stage	Second stage
$\max_{\lambda,\phi} \phi$	$\max_{\lambda} e^T(\tilde{X}\lambda - \tilde{x}_j) + e^T(\hat{\phi}_j\tilde{y}_j - \tilde{Y}\lambda)$
s.t. $Prob({}_i\tilde{x}\lambda < \tilde{x}_{ij}) \geq 1 - \alpha$	s.t. $Prob({}_i\tilde{x}\lambda < \tilde{x}_{ij}) \geq 1 - \alpha$
$Prob({}_r\tilde{y}\lambda > \phi\tilde{y}_{rj}) \geq 1 - \alpha$	$Prob({}_r\tilde{y}\lambda > \hat{\phi}_j\tilde{y}_{rj}) \geq 1 - \alpha$
$\varphi(e^T\lambda) = \varphi$	$\varphi(e^T\lambda) = \varphi$
$\lambda \geq 0$	$\lambda \geq 0$
	$i = 1, \ldots, m;\ r = 1, \ldots, s.$

Input oriented model

First stage	Second stage
$\min_{\lambda,\theta} \theta$	$\max_{\lambda} e^T(\tilde{X}\lambda - \hat{\theta}_j\tilde{x}_j) + e^T(\tilde{y}_j - \tilde{Y}\lambda)$
s.t. $Prob({}_i\tilde{x}\lambda < \theta_j\tilde{x}_{ij}) \geq 1 - \alpha$	s.t. $Prob({}_i\tilde{x}\lambda < \hat{\theta}_j\tilde{x}_{ij}) \geq 1 - \alpha$
$Prob({}_r\tilde{y}\lambda > \tilde{y}_{rj}) \geq 1 - \alpha$	$Prob({}_r\tilde{y}\lambda > \tilde{y}_{rj}) \geq 1 - \alpha$
$\varphi(e^T\lambda) = \varphi$	$\varphi(e^T\lambda) = \varphi$
$\lambda \geq 0$	$\lambda \geq 0$
	$i = 1, \ldots, m;\ r = 1, \ldots, s.$

Table 3: Two stages chance constrained models

Model (Orientation)	Returns to Scale	Envelopment Type	Range	Units Invariant	Involves Non-Archimedean
Additive	Variable	Piecewise linear	objective value≤ 0	No	No
	Constant	Piecewise linear		No	No
Almost 100% confidence	Constant	St. Hyperplane	objective value$\leq \sigma_\varepsilon \Phi^{-1}(\epsilon)$	No	Yes
additive model; Problem (4)	Variable	St. Hyperplanes	$\mid e^T(A\lambda - a_j) + e^T(\bar{b}_j - \bar{B}\lambda) \mid$	No	Yes
BCC model (input)	Variable	Piecewise linear	$0 < \theta \leq 1$	Yes	Yes
BCC model (output)	Variable	Piecewise linear	$1 \leq \phi$	Yes	Yes
CCR model (input)	Constant	Piecewise linear	$0 < \theta \leq 1$	Yes	Yes
CCR model (output)	Constant	Piecewise linear	$1 \leq \phi$	Yes	Yes
Almost 100% confidence					
oriented models,	Variable	St. Hyperplanes	$0 < \theta \leq 1,\ 1 \leq \phi$	Yes	Yes
Problems (12),(10)	Constant	St. Hyperplane	$0 < \theta \leq 1,\ 1 \leq \phi$	Yes	Yes
(input, output)					
Chance constrained					
oriented models	Variable	St. Hyperplanes	$0 < \theta \leq 1,\ 1 \leq \phi$	Yes	Yes
Problems (15),(14)	Constant	St. Hyperplane	$0 < \theta \leq 1,\ 1 \leq \phi$	Yes	Yes
(input, output)					

Table 4: Comparison of models