

Modification of Variable Returns to Scale Stochastic Data Envelopment Analysis (DEA) Models

Seyed Davood Forghani

Submitted to the
Institute of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

Master of Science
in
Industrial Engineering

Eastern Mediterranean University
February 2021
Gazimağusa, North Cyprus

Approval of the Institute of Graduate Studies and Research

Prof. Dr. Ali Hakan Ulusoy
Director

I certify that this thesis satisfies all the requirements as a thesis for the degree of Master of Science in Industrial Engineering.

Assoc. Prof. Dr. Gökhan İzbrak
Chair, Department of Industrial
Engineering

We certify that we have read this thesis and that in our opinion it is fully adequate in scope and quality as a thesis for the degree of Master of Science in Industrial Engineering.

Asst. Prof. Dr. Sahand Daneshvar
Supervisor

Examining Committee

1. Assoc. Prof. Dr. Adham Makkie

2. Asst. Prof. Dr. Sahand Daneshvar

3. Asst. Prof. Dr. Mazyar Ghadiri Nejad

ABSTRACT

Data Envelopment Analysis (DEA) was introduced under the name of a deterministic model assuming all the deviations from the estimated production frontier were one-sided indicating technical inefficiency. Biased estimations of inefficiency and production are provided by the model when deviations do not originate only from inefficiency but also from measurement errors. In 1988, Banker developed Data Envelopment Analysis as a stochastic model to reflect inefficiency and statistical noise simultaneously. However, from deterministic to stochastic, the problem with weak efficient frontiers and related biased results stayed the same.

This dissertation proposes a modification over Banker's stochastic DEA (SDEA) model by applying a limitation on the coefficients of inputs in the original model in order to change weak efficient hyperplane(s) while keeps general assumptions behind production function unaffected. This can change the production possibility set (PPS) while the frontier has the potential to give a better representation of the true production frontier. Comparing the results from the stochastic model and suggested modified model shows that the achieved model is providing a new benchmark for relative efficiency evaluation and production frontier estimation.

Keywords: Data Envelopment Analysis (DEA), Stochastic Data Envelopment Analysis (SDEA), Modified Model, Weak Efficient Frontier

ÖZ

Veri Zarflama Analizi (VZA), tahmin edilen üretim sınırından tüm sapmaların teknik verimsizliği gösteren tek taraflı olduğu varsayılarak deterministik bir model adı altında tanıtıldı. Sapmaların yalnızca verimsizlikten değil aynı zamanda ölçüm hatalarından da kaynaklandığında, model tarafından sapmaların verimsizlik ve üretim tahminleri sağlanır. 1988'de Banker, verimsizliği ve istatistiksel gürültüyü eş zamanlı yansıtmak için stokastik bir model olarak Veri Zarflama Analizini geliştirdi. Bununla birlikte, deterministikten stokastiklere Zayıf verimli sınırlar ve ilgili önyargılı sonuçlar ile ilgili sorun aynı kaldı.

Bu tez, Banker'in stokastik DEA (SDEA) modeli üzerinde, zayıf etkin hiper düzlem (ler) i değiştirmek için orijinal modeldeki girdi katsayılarına bir sınırlama uygulayarak, üretim fonksiyonunun arkasındaki genel varsayımları etkilenmeden koruyarak bir değişiklik önermektedir. Bu, üretim olasılık kümesini (PPS) değiştirebilirken, sınır, gerçek üretim sınırının daha iyi bir temsilini verme potansiyeline sahiptir. Stokastik model ve önerilen modifiye modelden elde edilen sonuçların karşılaştırılması, elde edilen modelin göreceli verimlilik değerlendirmesi ve üretim sınırı tahmini için yeni bir kriter sağladığını göstermektedir.

Anahtar Kelimeler: Veri Zarflama Analizi (VZA), Stokastik Veri Zarflama Analizi (SDEA), Değiştirilmiş Model, Zayıf Verimli sınır

TO MY ONLY REAL FORTUNE,
MY FAMILY

ACKNOWLEDGMENT

Throughout my graduate program and the writing of this dissertation, I received a great deal of support and assistance. I would first like to thank my dear supervisor, Asst. Prof. Dr. Sahand Daneshvar, whose expertise, wisdom, and diligence lit up all the way from the first day to the last. It would simply be impossible to make it without you.

I would like to kindly thank our department's staff members. Especially, I would like to single out our chair of department, Assoc. Prof. Dr. Gökhan İzbrak, for all the opportunities that I have been given. It has been a privilege to work under your chairmanship. I would also like to express my deepest gratitude to my respected advisor, Assoc. Prof. Dr. Adham Makkie and my dear professors, Prof. Dr. Bela Vizvari, and Assoc. Prof. Dr. Hüseyin Güden for their valuable guidance. You provided me with the tools that I needed to choose the right direction and successfully complete my Master's program.

Last but not least, I am the most grateful of my family, my parents, my brothers, Danial and Dara, My grandparents, and my uncles for all the unconditional love and support. Thank you for always being there for me.

TABLE OF CONTENTS

ABSTRACT	iii
ÖZ	iv
DEDICATION	v
ACKNOWLEDGMENT	vi
LIST OF TABLES	ix
LIST OF FIGURES	x
LIST OF ABBREVIATIONS	xi
1 INTRODUCTION	1
1.1 Problem Description.....	1
1.2 Thesis Structure.....	3
2 REVIEW OF THE LITERATURE.....	4
2.1 Frontier Estimation.....	4
2.2 Data Envelopment Analysis Models	9
2.3 Stochastic Aspects of DEA	11
2.4 Proposed Solutions to Stochastic DEA	13
3 METHODOLOGY.....	19
3.1 Traditional DEA Models	19
3.1.1 CCR Model.....	20
3.1.2 BCC Model.....	22
3.2 Stochastic DEA	24
3.3 Banker’s Stochastic DEA (SDEA) Model	25
3.4 Modification for Better Estimation	28
3.4.1 Weak Efficient Frontier	28

3.5 Illustrative Application.....	33
3.5.1 One Input and One Output.....	33
3.5.2 Two Input and One Output.....	37
3.6 Developing DEA and quantile regression to Stochastic Frontier Analysis (SFA)	40
3.6.1 Modification for a better approximation	42
4 CONCLUSION	45
4.1 Summary of Methods	45
4.2 Future Study	45
REFERENCES.....	47
APPENDICES	58
Appendix A: 1 input 1 output illustrated application of Banker’s SDEA and suggested modified model for different values of the weight “c”.....	59
Appendix B: 2 input 1 output illustrated application of Banker’s SDEA and suggested modified model for different values of the weight “c”.....	63
Appendix C: Combined DEA and quantile regression model compared with the suggested modified model (dataset 1).....	67
Appendix D: Combined DEA and quantile regression model compared with the suggested modified model (dataset 2).....	68
Appendix E: Sensitivity analysis report for the weight “c” using the data in section 3.5.2.....	69

LIST OF TABLES

Table 1: Different forms of the CCR model	21
Table 2: Different forms of the BCC model	23
Table 3: One input and one output data	34
Table 4: Two input and one output data	40

LIST OF FIGURES

Figure 1: The efficiency structure of a production process.....	20
Figure 2: Estimated production frontier of CCR model.....	22
Figure 3: Estimated production frontier of BCC model.....	24
Figure 4: Stochastic DEA frontier.....	29
Figure 5: Possible manipulations of weak efficient frontier	32
Figure 6: SDEA estimated production frontier	35
Figure 7: Modified SDEA estimated production frontier	36
Figure 8: Production function, stochastic and modified model (dataset 1)	43
Figure 9: Production function, stochastic and modified model (dataset 2)	43

LIST OF ABBREVIATIONS

BCC	Banker, Charnes, and Cooper
CCR	Charnes, Cooper, and Rhodes
COLS	Corrected Ordinary Least Square
DEA	Data Envelopment Analysis
DMU	Decision Making Unit
MAD	Minimum Absolute Deviation
MSE	Mean Square Error
PPS	Production Possibility Set
SDEA	Stochastic Data Envelopment Analysis

Chapter 1

INTRODUCTION

1.1 Problem Description

Since Farrell (1957) presented the terminology of inefficiency and its basic structure of measuring, a race for developing mathematical methods and econometric approaches to estimate inefficiency and optimal frontier started. Following Farrell's intentions, Charnes et al. (1978) introduced the initial DEA model (CCR) and years later, Banker et al. (1984) provided the BCC model under variable returns to scale assumption. From there on, DEA has been used in a broad range of applications and played a significant role in operation research studies.

Data envelopment analysis (DEA) is an effective technique for evaluating the relative efficiency of similar decision-making units (DMUs) with multiple inputs and outputs. according to observed data and a set of assumptions, DEA creates a reference technology set named as production possibility set (PPS), in which frontier, distinguishes the comparatively most efficient DMUs. A DMU is categorized as efficient or inefficient dependent on its location relative to the mentioned frontier.

The PPS frontier is constructed of two distinct kinds of facets:

- 1- FDEFs (Full Dimensional Efficient Facets)
- 2- FDWFs (Full Dimensional Weak Facets)

Ever since the DEA method was introduced, many researches have been carried out on identifying efficient frontiers and DMUs under their assessment. However, less attention has been paid to the weak efficient frontier and the corresponding DMUs. The weak efficient frontier which is created to satisfy convexity constraint could be counted as a drawback in basic DEA models since in sensitivity analysis, eliminating weak efficient DMUs and efficient DMUs on the intersection of efficient and weak efficient facets has no effect on the stability region. In addition to this, as in many production functions such as translog or logarithmic production functions like Cobb-Douglas, it is more realistic not to have a PPS in which by changing Inputs, Output doesn't change or vice versa.

Daneshvar et al. (2014) introduced a modified BCC model obtained by facet analysis, resulting in a developed stability region and a new benchmark for scoring formerly weak efficient DMUs and inefficient DMUs which were compared to the weak efficient frontier.

Daneshvar et al. (2014)'s paper became a motivation to examine the possibility of enhancing the PPS region in the stochastic DEA area. Certainly, Banker's stochastic DEA model (Banker (1988)) as a fundamental model for estimating production function in the stochastic field, could be an excellent choice.

In this thesis, Banker's SDEA model has been considered to be modified by adding a new constraint in order to push the frontier to get close to the true underlying production function. This adjustment has been applied to two SDEA models introduced by Banker and numerical examples have been presented to have a better illustration.

1.2 Thesis Structure

This work is divided into four chapters. In the first chapter, a general idea about the goal of this paper is given. The second chapter is devoted to the history of the related researches and the review of proposed methods. The methodology of the suggested modification and also corresponding numerical examples are discussed in chapter 3. Lastly, the conclusion and possible future studies are given in chapter 4.

Chapter 2

REVIEW OF THE LITERATURE

2.1 Frontier Estimation

In decades, the effort for estimating production functions stating the highest amount of output given a set of inputs with determined technology has been progressed by econometricians and engineers. However, still, the gap between the theoretical approaches and empirical reality exists. Farrell (1957), started debating about the possibility of approximating production function. Since then, a variety of approaches toward the true efficient frontier were presented which can be grouped into two basic categories: parametric and nonparametric methods. The parametric method needs a functional mathematical frame for the frontier. This can be considered not only as its feature but also as a drawback (Seiford & Thrall (1990)) since predicting the behavior of production to specify its form is not easy. On the other hand, a nonparametric method needs no prior information about the production frontier (Bauer (1990)). As DEA has no need for assumptions to indicate functional form, it is categorized as a nonparametric approach.

Efficient frontier analysis was classified by Lovell & Schmidt (1988) into four main classes:

1. Pure programming method
2. Modified programming method
3. Deterministic statistical frontier method

4. Stochastic frontier method.

These methods differ from one another in the use of statistical methods or mathematical programming to approximate the frontier, whether the frontier is imposed to be parametric or nonparametric, and the assumption on the stochastic or deterministic frontier.

Farrell (1957) proposed pure programming method for the first time. In this method, mathematical programming is used to create a frontier out of the reference set to examine the efficiency relative to the mentioned frontier. This method was offered by Charnes et al. (1978) in the shape of data envelopment analysis and also Färe et al. (1985). More efforts have been done by Varian (1984) and Banker & Maindiratta (1988) on production possibility set.

Pure programming method has a nonparametric characteristic allowing it to adapt without considering the inherent format of the actual efficient frontier. This approach envelops the data by constructing facets only by utilizing observations in the reference set. Lovell & Schmidt (1988) know this characteristic as the most reason to prefer pure programming over the rest approaches:

Perhaps the most appealing characteristic of the pure programming approach is that the input set it constructs is the smallest well-behaved set containing all the data. Such a set is piecewise linear and the construction process achieves considerable flexibility because the breaks among the pieces are determined endogenously so as to fit the data as closely as possible.

Lovell & Schmidt (1988), however, denotes the disadvantage of pure programming to be the fact that the frontier which envelopes the data is deterministic.

Modified programming method, the second type of efficient frontier analysis, was proposed by Farrell (1957). It also utilizes a series of arithmetical programs to build an efficient frontier. But the difference from pure programming is that in this method the purpose of using the programs is to estimate parameters of an anticipated functional form of the frontier. Therefore, it is considered as a parametric technique.

Lovell & Schmidt (1988) express three drawbacks of the modified programming as follows: firstly, the modified programming method, similar to the pure programming, presents a deterministic frontier that only reflects deviations due to technical inefficiency. Secondly, this method has its complexity working with multiple outputs situation. Lastly, being parametric, forces the need for a correctly chosen functional form to obtain suitable parameters.

Lovell and Schmidt's third category of efficient frontier analysis is the deterministic method which Afriat (1972) proposed and authors like Richmond (1974) and Greene (1980a), (1980b) developed the idea. This approach, as Lovell & Schmidt (1988) explain, envelops the data by a deterministic frontier, however, rather than mathematical programming, statistical techniques are used to estimate the frontier. Methods such as maximum likelihood and corrected ordinary least squares (COLS) are applied for frontier estimation as well as inefficiency density parameters associated with a supposed functional form. To confirm one sided deviations from the frontier, corrections are applied to the COLS method. Hypothesis tests about the frontier parameters are allowed when using statistical estimation.

Similar to previously discussed methods, being deterministic is a disadvantage to deterministic method. Moreover, although Kopp & Diewert (1982) and Zieschang

(1983) adapted the modified programming method to deal with multiple outputs, but still handling this situation is a difficulty in deterministic method. In addition, hypothesis tests and parameter estimations are variant based on the frontier's functional form and the distribution of the inefficiency terms. Also, increase in the requirements of sample size could be another burden.

Last class in Lovell & Schmidt (1988)'s classification of frontier estimation methods is the stochastic approach. It was first suggested by Aigner et al. (1977) and Meeusen et al. (1977) where the frontier is allowed to be literally stochastic. There are other contributions in this area such as Schmidt & Lovell (1979) and Jondrow et al. (1982).

In the stochastic frontier method, the error is constructed of two parts: inefficiency and random deviation. Like deterministic statistical and modified programming frontiers, Inefficiency variation is taken into account as a one-sided deviation term. Deviations from this nature are assumed to be caused by technical inefficiency in operations. On the other hand, random error is a two-sided deviation triggered from measurement errors, random perturbations and other stochastic sources of the system. Basically, stochastic frontier method can be considered as the extended deterministic statistical frontier method combined with a two-sided random deviation term.

Since stochastic frontier method uses statistical devices such as corrected ordinary least squares or maximum likelihood, is counted as a parametric approach. Jondrow et al. (1982), under particular distribution assumptions, for the two error components obtained separate estimations based on conditional distributions. Although, the estimations don't succeed to be statistically consistent, but using a distinct deviation for random factors allows DMUs to be inefficient due to reasons except system's

inefficiency. This combination of random deviations into the model could be considered as the main contribution of this method over the rest three methods.

As Lovell & Schmidt (1988) mention, the stochastic frontier method has weak points. The frontier and also error terms must be correctly assigned by related functional forms. In addition, sample sizes need to be increased in order to use statistical estimation techniques. Lastly, the multiple output inclusion in the model has been problematic.

In more recent years, DEA has been compared with parametric methods in many researches. Three distinct deterministic methods were used by Bjurek et al. (1990) to examine the insurance offices' efficiency. The observed efficiency scores obtained by using DEA, a quadratic production function, and a Cobb-Douglas function were slightly different. Banker et al. (1986) studied data envelopment analysis and translog estimation, to analyze hospital production. This research was continued by Banker et al. (1988) through simulated data analysis from a known parametric function. In terms of efficient or inefficient classification, they realized that DEA surpass translog estimation. Another result was that for larger samples, DEA efficiency estimates improved. Using Monte Carlo simulation, the performance of DEA and corrected ordinary least squares under the existence of measurement errors were investigated by Banker et al. (1993). By means of Cobb-Douglas frontiers, even in the presence of considerable measurement errors, DEA worked better than corrected ordinary least square (COLS) when nonclassical inefficiency distributions were presented. COLS acted better in classical distributions only if the sample size was fairly large while with relatively small sample sizes, DEA worked well in all considered cases. Like other studies, a single output case was considered.

2.2 Data Envelopment Analysis Models

Data envelopment analysis evaluates the efficiency of separate DMUs by comparing them to a frontier or boundary made from the reference set. The most important benefit of utilizing DEA is that there is no need to assume a functional form for the efficient frontier. Hence, the efficiency scores as well as their reference points are conditional since they are dependent on the organizations that are included in the sample. As Cooper & Tone (1997) declare:

The performance of a decision-making unit is considered to be fully DEA efficient if and only if the performance of other decision-making units does not provide evidence that some of its inputs or outputs could be improved without worsening some of its other inputs or outputs.

considering this essential structure as a basis, the initial model originated from Charnes et al. (1978) (CCR) and soon extended into many other forms. The BCC model was created by Banker et al. (1984) (BCC) as an alternative model to isolate the scale efficiency concept. Charnes et al. (1985) suggested the additive model and multiplicative model was created by Charnes et al. (1982). More extensions of the mentioned models, their applications, and capability evaluation of DEA have directed to thousands of articles regarding its application.

The goal of designing BCC was to particularly solve the problems concerning scale and technical efficiency as in CCR model they were aggregated. By adding a convexity constraint involving summation of the weightings of the under evaluation DMUs to equal one, the virtual DMUs with comparable scales to the real DMUs were created. Thus, by taking the CCR aggregate efficiency ratio to the merely technical efficiency of the BCC model, the scale efficiency could be defined. The only case that the aggregate efficiency of the CCR and BCC technical efficiency scores for a specific

DMU become equal is when DMU is of the optimal size. It should be noted that like all the DEA models, inputs and outputs in BCC and CCR are assumed positive in value.

Seiford & Thrall (1990) modified the scale derivations of the CCR model by grouping weight limitations based on the allowed returns to scale. They precisely suggested four models by adding a constraint on the summation of the weights according to the measurement of a specific returns to scale (constant returns to scale (original CCR model), increasing returns to scale, decreasing returns to scale, variable returns to scale (original BCC model)).

The multiplicative form of DEA was proposed by Charnes et al. (1982) in which a more classical Cobb-Douglas form was given to the efficient frontier. In this model, a log-linear envelopment is provided by utilizing the proportion of multiplication of exponentially weighted outputs over inputs while that of CCR model is a fraction of linear combinations. The multiplicative formulation can be altered to a linear programming using a logarithmic transformation. This model creates a piecewise log-linear frontier which is different from traditional linear frontiers. Charnes et al. (1985) introduced a unique additive programming model by entering a convexity constraint for the weights into the dual form of the multiplicative model.

Several other forms of DEA have been introduced during the years. Some of them are numerated here: Assurance Region approach by Thompson et al. (1990) and the Cone Ratio model by Charnes et al. (1989) and also Charnes et al. (1990), in which the effort was on restricting the efficient DMUs by controlling the possible multipliers. Concerning restrictions of weight flexibility, Dyson & Thanassoulis (1988) as well as

Wong & Beasley (1990) contributed by their researches. Banker & Morey (1986b), Kamakura (1988), and Rousseau & Semple (1993) have integrated DEA with categorical variables. Another integration was done by Banker & Morey (1986a) with nondiscretionary variables. Ordinal variables were included in DEA by Golany (1988), Cook & Kress (1991), and Cook et al. (1993), (1996). Data envelopment analysis with ordinal relationships was applied by Ali et al. (1991) to model time lag effects.

2.3 Stochastic Aspects of DEA

Generally, in most studies related to efficiency evaluation using DEA, the collected data is considered to be deterministic. Consequently, efficiency measurements are not assumed stochastic. What separates efficient DMUs from inefficient DMUs in a standard data envelopment analysis, is whether the efficiency value is one or not. In this way, any difference from an efficient score is only due to technical inefficiency. Huang & Li (1996) state that there is no place for stochastic variations in the data when using DEA. Deterministic approaches to random errors, like sensitivity analysis, still consider observations non-stochastic. Nonetheless, there are strong discussions in support of considering the data and in a result the frontier to be stochastic.

The set of analyzed organizations by DEA mostly represent a sample of population Sengupta (1995). Alone, the sample data gives only limited knowledge about the frontier. Efficiency score derived from sample understandings may or may not reveal the true concealed efficiency of the comparable DMUs rather than just those in the analyzed set. Likewise, statistics computed from the sample, for instance ratio of efficient DMUs, are fundamentally sample statistics and accordingly obey some unidentified sampling distribution.

Sengupta (1995) was not the only one who mentioned that the frontier is sensitive to the sample size. If the number of observations in the sample increase, the technical efficiency of the DMUs often decrease because there is a possibility that the DMUs which are placed close to the frontier increase. Banker (1993) claimed that under certain circumstances, the estimated frontier by DEA can reach the true frontier asymptotically. With inserting observations into the sample, efficiency scores either decrease or stay constant. Therefore, while many other statistical methods are affected by sample size, frontier estimation methods that work only based on extremal points are affected even more severely.

Another factor that can highly affect the efficiency scores in DEA is the choice of inputs and outputs. Epstein & Henderson (1989), along with others, have remarked that data envelopment analysis is sensitive to the selecting of variable. The fact that excluding critical factors can affect the outcomes of a DEA analysis was suggested in a research by Ahn & Seiford (1993). This conclusion not only holds for DEA but also is true about all the frontier analysis forms.

Epstein & Henderson (1989), Huang & Li (1996), among others, implied that measurement error could have major impact on the efficiency score of a DMU. As previously described, frontier estimation methods are in general, based on just extremal points. Outliers as a result of measurement error can influence the frontier estimates. A method was suggested by Sengupta (1995) in which outliers could be recognized and removed from data. His suggestion for frontier estimation, however, is centered on central location measures and emphasizes that almost all of the deviation is because of measurement errors.

The DEA selected model can also alter the results of efficiency scores. The ratio, additive, and multiplicative models and their alternatives each measure dissimilar aspects of efficiency. Ahn et al. (1988) and Ahn & Seiford (1993) performed researches to study the sensitivity of efficiency values based on model specifications. To study the sensitivity of the efficiency to the model selection, some researchers like Alshare (1998), have assessed efficiency scores of certain organizations using various DEA models. Model specification as well as selecting proper factors are probably best determined in study-by-study basis and there is no general research to solve it.

The notion of a severely deterministic frontier in which efficiency scores are fixed, is only logical under specific conditions. Effects of sampling, the observations quantity, the number of variables, choice of variable, and other error forms make the interpretation of deterministic efficiency scores extremely challenging. Nevertheless, many methods to stochastic DEA are proposed to deal with this problem.

2.4 Proposed Solutions to Stochastic DEA

Generally, for frontier analysis, two formations of error models are considered (Sengupta (1995)). For single output situation, the model with one error term is

$$y = f(x_1, \dots, x_m) - \varepsilon$$

$$\varepsilon \geq 0$$

here ε is a stochastic term with one-sided error. In this state, all observations vary randomly and in the same direction from the frontier. This model was suggested by Schmidt (1976) to bring statistical properties into frontier estimation employing mathematical programming methods. Later, Aigner et al. (1977) presented the composed error model containing two error elements: a one-sided error ω and a symmetric error ξ .

$$y = f(x, \beta) - \varepsilon$$

$$\varepsilon = \omega + \xi$$

$$\omega \geq 0 \text{ and } \xi \text{ free}$$

ω is defined as the deviation from the frontier created by factors that are controllable by the organization. On the other hand, component ξ denotes uncontrollable factors related to things like, as they call them, luck, machine performance, climate, measurement and observation errors, plus other kinds of random disturbance. Researches on DEA stochastic methods mostly tend to consider composed error.

Some of the stochastic DEA approaches are pointed out here in this literature including Banker's stochastic model, chance-constrained programming method and the "efficiency distribution approach" along with others. Moreover, statistical basics for DEA have been covered.

The first stochastic data envelopment analysis (SDEA) model was offered by Banker (1988) presenting a linear programming model in which a piecewise linear frontier similar to DEA shape passes within the middle of the data. By restricting production to meet the famous Afriat constraints Afriat (1972), The SDEA model minimized absolute deviations from observations to a nonparametric frontier. Banker et al. (1991) used the model to analyze the significance of contextual variables and to compare the outcome with the stochastic frontier model. Banker & Maindiratta (1992) realized how to estimate the model by employing maximum likelihood with presumed functional form.

In addition, the connections of stochastic DEA to quantile regression were discussed by Banker (1988). Specifically, a person can produce any practical quantile of the

inherent probability distribution by selecting the proper weights on the negative and positive residuals. In the event of equal weights, the median regression model is estimated by stochastic DEA without any required assumption on the production function. In another case when the underlying error distribution is normal distribution, then it is possible to use least squares rather than least absolute deviation to find maximum likelihood estimates. This extension was presented by Kuosmanen (2008) and continued by Kuosmanen & Johnson (2010). Wang et al. (2014) investigated the stochastic DEA and its connections to the nonparametric quantile regression. Moreover, in order to prevent crossing quantiles, the authors established an associated quantile function estimator and also studied estimation in a stochastic frontier framework.

Although Banker (1988) linked stochastic DEA model with quantile regression, the article did not attempt to investigate the decision on the optimal quantile. According to Kuosmanen & Kortelainen (2012), the median regression could be applied and be corrected for the intercept like in COLS. Following the effort of Azzalini & Capitanio (2013), Jradi & Ruggiero (2019) provided comparisons between SDEA, quantile regression with correct and incorrect functional form and true production function. Then they discussed corresponding optimal weight considering specific known distribution for deviations.

Banker et al. (2015) studied the sensitivity and stability of Banker's SDEA. Three cases regarding to perturbations of all inputs, perturbations of output and simultaneous perturbations of inputs and output was studied.

Charnes & Cooper (1959) introduced chance-constrained programming and the concept was developed by Kall (1976) and other researchers. Applications to the model were examined by Sengupta (1987), Land et al. (1993), and Olesen & Petersen (1995). The approach was expanded by Huang & Li (1996) applying joint chance constraints.

Fundamentally, in chance constrained programming, it is supposed that the linear constraints take place with some probability less than one. This method considered nonparametric since it does not require enforcing a functional form to the efficient frontier. As Olesen & Petersen (1995) stated, because there is a statistical basis in this model, random disturbances including measurement error are allowed.

According to Land et al. (1993), chance constraint programming approach generates a frontier enveloping the sample observations of DMUs “*most of the time*”. This model is in consistent with composed error method, specially integrates the random error into the calculation of the frontier (Olesen & Petersen (1995)).

Some questions were raised by Land et al. (1993) regarding using of chance constraint programming in DEA. The formulation of chance constrained give the permission to DMUs to go above the frontier with efficiency scores more than one. Land, Lovell, and Thore named these observations chance constrained hyper-efficient DMUs. Moreover, about setting the value of c , which is an essential component of the standard deviation in the probability constraints, they mentioned:

. . . the greater the stochastic variability of outputs (the greater the coefficient c), the greater is the band of output territory that is permitted outside the envelope. The envelope then moves successively closer to any given observation and the efficiency ratio approaches 1.0. For data with large amounts of uncertainty, the efficiency scores will automatically be close to

unity. The empirical usefulness of the chance constrained DEA calculation may then be questioned.

In continue, it is noted that the efficiency scores will increase as well, for given flexibility, under the condition that the tolerance level of the chance constraints is reduced.

Land et al. (1993), however, acknowledged the advantages of merging a stochastic frontier with data envelopment analysis model. By creating buffers, deterministically inefficient observations turn to chance constrained efficient DMUs. In this way, the frontier is allowed to give permission to some random errors without supposing that the DMU is technically inefficient. They stated that this addition of random error allows DMUs to be ranked as efficient while there is slack in some outputs.

An extended form of chance constrained approach was suggested by Olesen & Petersen (1995). Their model allows to evaluate DMUs in the existence of random multiple inputs and multiple outputs. Another stochastic DEA model exercising joint chance constraints was introduced by Huang & Li (1996). They showed required and enough conditions for a DMU in order to be stochastically non-dominant efficient.

The efficiency distribution method introduced by Sengupta (1996b) is based upon comparing of the output distributions of efficient and inefficient observations. Two approaches are considered regarding this method. The first is developed to estimate a single efficient frontier function considering only the efficient DMUs. A provable distribution of error components is produced from the deviation between real and DEA efficient output amounts for all DMUs. Methods like simple or generalized form of

moments as well as maximum likelihood could be employed to determine an estimation of the error probability density function.

The second approach introduced by Sengupta (1996b) is constructed on making “peer group” of efficient DMUs and comparing them with inefficient DMUs utilizing dummy regression. The mentioned “peer group” is derived from the notion of modal efficiency. Dummy variable regression works using a Boolean variable for classification to determine if a DMU is a member of the modal efficiency group.

Banker (1993) provided an imperative statistical foundation for single output case in DEA. He presented that the DEA estimators are equivalent to maximum likelihood estimators if the deviation of real output from the corresponding efficient output is counted as a stochastic variable with a probability density function that is monotonically decreasing. DEA is hence robust compared with parametric frontier estimation methods presented by Aigner & Chu (1968). These parametric methods have been indicated to be maximum likelihood estimators just when the inefficiency obeys exponential or half-normal distributions.

Several other approaches have been proposed for the application of statistical approaches to DEA. Canonical correlation was used by Sengupta (1990) to specify statistically significant elements of the input-output mixture that characterized an individual facet of the efficient frontier. Also, the integration of time varying efficiency results has been tackled by Sengupta (1996a), (1996b) and Charnes et al. (1984).

Chapter 3

METHODOLOGY

The methodology used in this research is based on linear algebra and modifications applied to the models are examined by the GAMS software package. To simulate real production observations, random values are generated by Office Excel. The goal in this dissertation is to manipulate Banker's stochastic DEA model to replace the weak efficient frontier by a better estimation of production function.

3.1 Traditional DEA Models

When DEA was proposed the idea was to determine weights for each input and output of a DMU in order to form virtual input and output as following:

$$\textit{Virtual input} = v_1 x_{1o} + \dots + v_m x_{mo}$$

$$\textit{Virtual output} = u_1 y_{1o} + \dots + u_s y_{so}$$

Hence, the goal is to maximize the undermentioned ratio by using linear programming:

$$\frac{\textit{Virtual output}}{\textit{Virtual input}}$$

In m inputs and s outputs case, the general efficiency structure of n DMUs are defined as below:

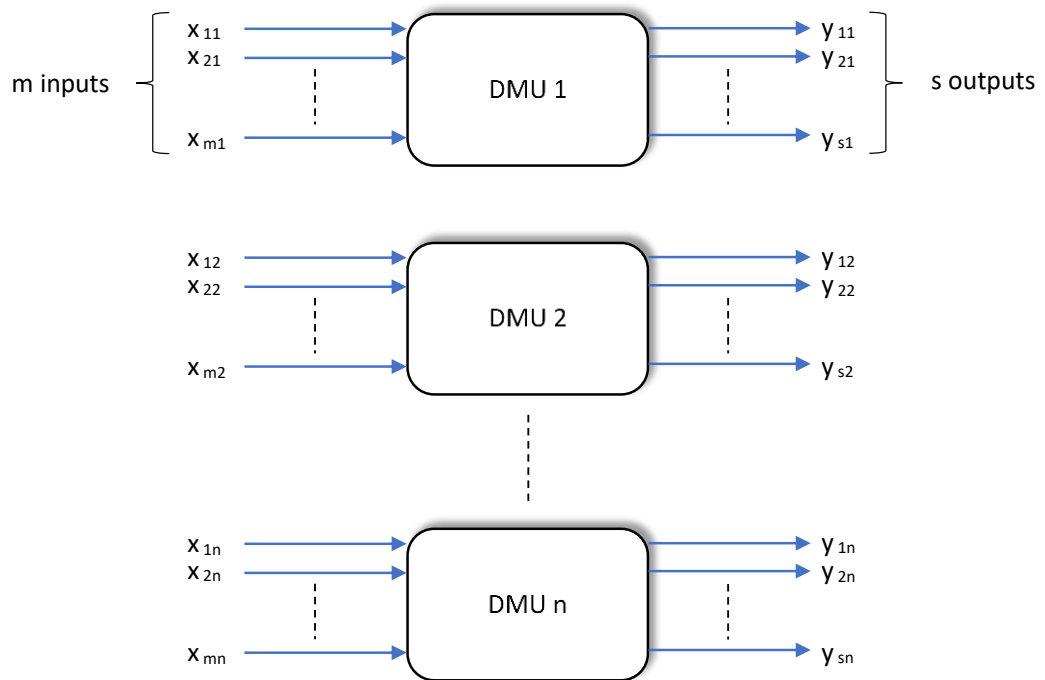


Figure 1: The efficiency structure of a production process

Therefore, one of the most basic DEA models, the CCR model, was introduced by Charnes et al. (1978) to show the efficient DMU(s) as well as the relative efficiency of the other DMUs.

3.1.1 CCR Model

The mathematical representation of the input-oriented and output-oriented primal CCR model for evaluating DMU "o" as well as the corresponding dual forms are given in the following table:

Table 1: Different forms of the CCR model

	Input-Oriented	Output-Oriented
Primary	$(LP_o) \max_{u,v} \quad u y_o$ <i>subject to</i> $vx_o = 1$ $-vX + uY \leq 0$ $v \geq 0, u \geq 0$	$(LPO_o) \min_{p,q} \quad px_o$ <i>subject to</i> $qy_o = 1$ $-pX + qY \leq 0$ $p \geq 0, q \geq 0$
Dual	$(DLP_o) \min_{\theta,\lambda} \quad \theta$ <i>subject to</i> $\theta x_o - X\lambda \geq 0$ $Y\lambda \geq y_o$ $\lambda \geq 0$	$(DLPO_o) \max_{\eta,\mu} \quad \eta$ <i>subject to</i> $x_o - X\mu \geq 0$ $\eta y_o - Y\mu \leq 0$ $\mu \geq 0$

The Production Possibility Set (PPS) of the CCR model is stated over five assumptions:

1. Input and output vectors (x_j, y_j) ($j = 1, \dots, n$) of n DMUs are semipositive ($x_j \geq 0, x_j \neq 0$ and $y_j \geq 0, y_j \neq 0$ for $j = 1, \dots, n$).
2. The observed activities (x_j, y_j) ($j = 1, \dots, n$) belongs to PPS.
3. Constant returns-to-scale (CRS) assumption: If an activity (x, y) belongs to PPS, the activity (tx, ty) belongs to PPS for any positive scalar t .
4. For an activity (x, y) in PPS, any semipositive activity (\bar{x}, \bar{y}) with $\bar{x} \geq x$ and $\bar{y} \leq y$ is included in PPS.
5. All the semipositive linear combinations of activities in PPS, falls within PPS boundaries.

To sum up, what is said in the five assumptions above, we can define PPS as following considering λ as a semipositive vector in R^n :

$$PPS = \{(x, y) \mid x \geq X\lambda, y \leq Y\lambda, \lambda \geq 0\},$$

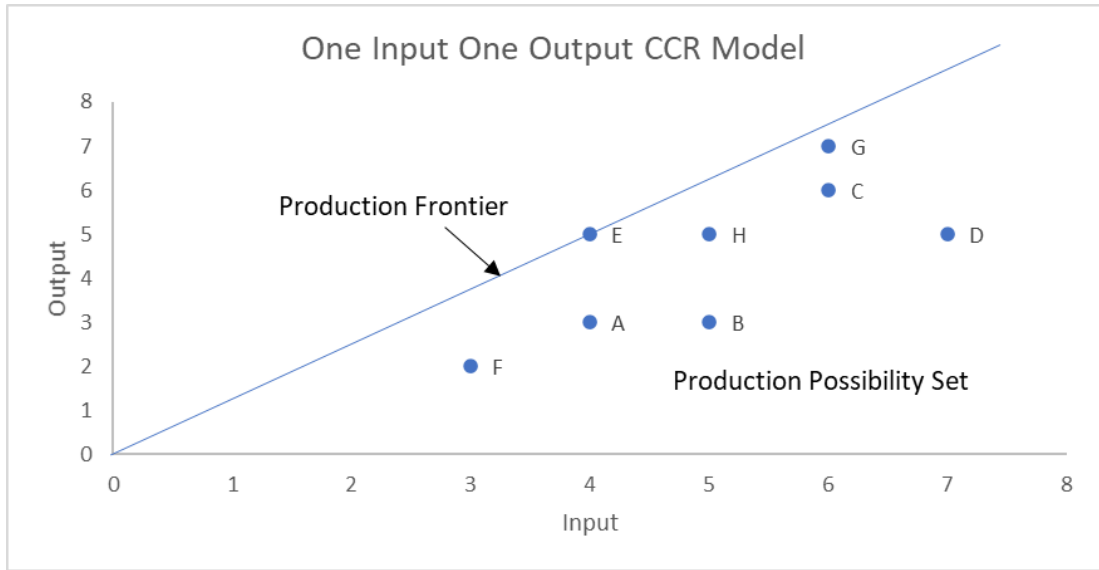


Figure 2: Estimated production frontier of CCR model

3.1.2 BCC Model

The well-known BCC model presented by Banker et al. (1984) is different from the CCR model in the sense that the production frontier is shaped by the convex hull of the studied DMUs. The Frontier has piecewise linear and concave features which leads to the variable returns-to-scale (VRS) assumption which differs from CCR in the convexity condition $\sum_{j=1}^n \lambda_j = 1, \lambda_j \geq 0, \forall j$ in its constraints.

The production possibility set of a BCC model is denoted by:

$$PPS = \{(x, y) \mid x \geq X\lambda, y \leq Y\lambda, e\lambda = 1, \lambda \geq 0\},$$

Where $\lambda \in R^n$, $X = (x_j) \in R^{m \times n}$ and $Y = (y_j) \in R^{s \times n}$ are given as data set, and e is a row vector with all elements equal to 1.

The mathematical representation of the input-oriented and output-oriented primal BCC model for evaluating DMU “o” as well as the corresponding dual forms are given in the following table:

Table 2: Different forms of the BCC model

	Input-Oriented	Output-Oriented
Primary	$\max_{u,v,u_0} \quad z = uy_o - u_0$ <p>subject to</p> $vx_o = 1$ $-vX + uY - u_0e \leq 0$ $v \geq 0, u \geq 0, u_0 \text{ free in sign.}$	$\min_{v,u,v_0} \quad z = vx_o - v_0$ <p>subject to</p> $uy_o = 1$ $vX - uY - v_0e \geq 0$ $v \geq 0, u \geq 0, v_0 \text{ free in sign}$
Dual	$\min_{\theta,\lambda} \quad \theta$ <p>subject to</p> $\theta x_o - X\lambda \geq 0$ $Y\lambda \geq y_o$ $e\lambda = 1$ $\lambda \geq 0$	$\max_{\eta,\lambda} \quad \eta$ <p>subject to</p> $X\lambda \leq x_o$ $\eta y_o - Y\lambda \leq 0$ $e\lambda = 1$ $\mu \geq 0$

The under evaluation $DMU_o (o \in (1, \dots, n))$ which can be shown as a point (X_o, Y_o) in an input-output coordinate system, is considered efficient DMU if the optimal value of the objective function of the dual model is equal to one. Considering figure 3, as a result of the formation of the BCC model, efficient points could be divided into three groups:

1. The strong efficient points like F, E & G
2. The Efficient points like all the points on the line FE and also EG
3. The weak efficient points like the points located on the vertical frontier connected to the point F (efficient in input-oriented and inefficient in output-oriented) plus horizontal frontier connected to the point G (efficient in output-oriented and inefficient in input-oriented)

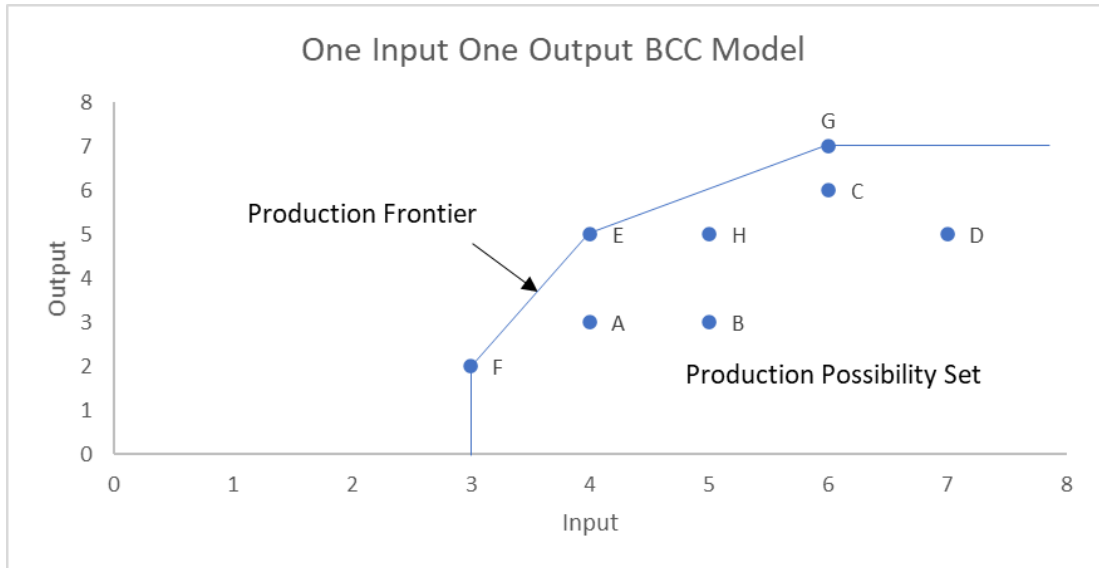


Figure 3: Estimated production frontier of BCC model

3.2 Stochastic DEA

Data envelopment analysis (DEA) is a suitable non-parametric method to assess the relative efficiency of multi-input and multi-output units based on practical data. Overall, observed data have inherent uncertainty, however, it is not easy to treat the stochastic data in the traditional DEA model. Therefore, the development of stochastic DEA is inevitable, where the uncertainty like measurement error should be integrated.

There are different notions of stochastic DEA comprising a variety of proposed methodologies that expands the idea or framework of conventional DEA in several directions. Generally, these directions could be grouped into three concepts or a combination of them:

1. The first direction develops DEA to be able to handle estimated deviations from frontier practice as random deviations.
2. The second direction develops DEA to be able to handle random noise in the form of either measurement errors or specification errors.

3. The third direction develops DEA to be able to regard or conceive the PPS as a random PPS, based on the random variation in data.

In this research, the first and second directions are evaluated and proposed modification has been applied to them to get a more realistic frontier that has less deviation from the production function.

3.3 Banker's Stochastic DEA (SDEA) Model

Banker (1988) introduced a basic model of SDEA. In this model, a symmetric two-sided deviation term peculiar to random factors (such as model specification and measurement errors) in company with the one-sided deviation term related to DMU's inefficiency is developed. As a result, only the single output case is considered because the multiple output case can be handled via nonlinear programming.

To express the relationship between this model and conventional DEA, consider the assumptions of BCC for estimating the PPS from observed data on output vectors y_j and input vectors $x_j, j = 1, \dots, n$. For the single output case the postulates for estimating the production frontier correspondence $y = f(x)$ relating the single output y to the input vector $x, f: X \rightarrow R$ where X is the convex hull of x_j , may be specified as:

Postulate 1: Monotonicity of Production Frontier

$$\text{if } y = f(x), y' = f(x') \text{ and } x \geq x', \text{ then } y \geq y'$$

Postulate 2: Concavity of Production Frontier

$$\begin{aligned} &\text{if } y = f(x), y' = f(x') \text{ and } 0 \leq \lambda \leq 1, \\ &\text{then } (1 - \lambda)y + \lambda y' \leq f((1 - \lambda)x + \lambda x') \end{aligned}$$

Postulate 3: Envelopment of Observed Data

for each observation $j = 1, \dots, n$, $y_j \leq f(x_j)$

Postulate 4: Minimum Extrapolation

if $g: X \rightarrow R$ satisfies postulates 1, 2 and 3 then $g(x) \geq f(x)$ for all $x \in X$

By considering these four postulates, the state of monotone increasing and concave production frontier is satisfied. Now for estimating this production frontier by stochastic DEA, the possible effect of uncontrolled random factors must be incorporated. Hence, such deviations caused by random factors and their stochasticity impact in the specification of the model is represented by the term u_j which may be expressed as:

$$u_j = u_j^+ - u_j^- \text{ with } u_j^+, u_j^- \geq 0$$

The random deviations u_j are supposed to be symmetric. Therefore, it is captured in the constraint below:

$$\sum_{j=1}^n u_j^+ = \sum_{j=1}^n u_j^-$$

Along with deviations due to random factors, as in traditional DEA, the inefficiency of the DMU may cause a shortfall in output compared to the predicted output level. Such deviations due to DMU inefficiency are shown by a nonnegative term v_j . To sum up, the actual output level could be represented as follows:

$$y_j = f(x_j) - u_j^+ + u_j^- - v_j$$

Then, the production frontier values are approximated by minimizing a weighted sum of the two deviations subject to the following constraints:

$$\text{Minimize } \sum_{j=1}^n (u_j^+ + u_j^- + cv_j)$$

subject to

for each $j = 1, \dots, n$ and for all $k = 1, \dots, n$ and $k \neq j$

$$(x_k - x_j)w_j + (v_j - v_k) + (u_j^+ - u_j^- - u_k^+ + u_k^-) \geq y_k - y_j$$

$$\sum_{j=1}^n (u_j^+ - u_j^-) = 0$$

$$w_j \geq 0, \quad v_j, u_j^+, u_j^- \geq 0$$

This model was first introduced by Banker in 1988. The SDEA model with m inputs and n observations, has $(m + 3)n$ variables and $n^2 - n + 1$ constraints and obviously is not infeasible since a feasible solution can be obtained from basic DEA when $u_j^+ = u_j^- = 0$.

The weight $c > 0$ in the objective function is a pre-specified constant which by giving different values, different estimates of the production function may be obtained. Therefore, the model represents a combination of the minimum absolute deviation (MAD) model (due to random factors) and the basic DEA model (due to inefficiencies).

Based on the next theorem, the correlation between the constant c and the contributions of the MAD and the DEA models are formalized.

Theorem: for any given data set $\{(x_j, y_j) \mid j = 1, \dots, n\}$, there exist c^M and c^D with $1/n \leq c^M \leq c^D \leq 2$ such that the model reduces to a minimum absolute deviation model (i.e. $v_j^* = 0$ for all j) for all $c > c^D$, and to the basic DEA formulation (i.e. $u^{+*} = u^{-*} = 0$ for all j) for all $c < c^M$.

Lemma 1: if $c > 2 \rightarrow v_j^* = 0$ for all $j = 1, \dots, n$

Lemma 2: if $c < 1/n \rightarrow u^{+*} = u^{-*} = 0$ for all $j = 1, \dots, n$

3.4 Modification for Better Estimation

the integrated model explained above consists of two different tools for estimating production function (MAD & DEA). While each of them has their own disadvantages when they are used separately, they can cover up each other's weak points by their advantages. For instance, MAD is a deterministic method that estimates only the average performance but DEA evaluates efficiency relative to a production frontier that measures the best obtainable performance. Furthermore, the problem with the regression-based parametric methods is that by specifying a particular parametric form, a considerable arbitrary and restrictive structure is imposed on the input-output correspondence. In contrast, DEA imposes a minimal structure of monotonicity and convexity on the PPS. On the other hand, DEA method only allows for one-sided inefficiency deviations whereas regression gives the possibility of having two-sided deviations component due to random errors. With that being said, there is another shortcoming with DEA that regression can't compensate in the explained stochastic model and in this paper, it is tried to be addressed.

3.4.1 Weak Efficient Frontier

In Banker's stochastic model, since all the deviations are measured vertically (direction of the deviations' vector due to MAD and DEA has to be the same so they can be summed for total deviation), starting from each DMU to the frontier line, the combined DEA model should be considered as output-oriented BCC.

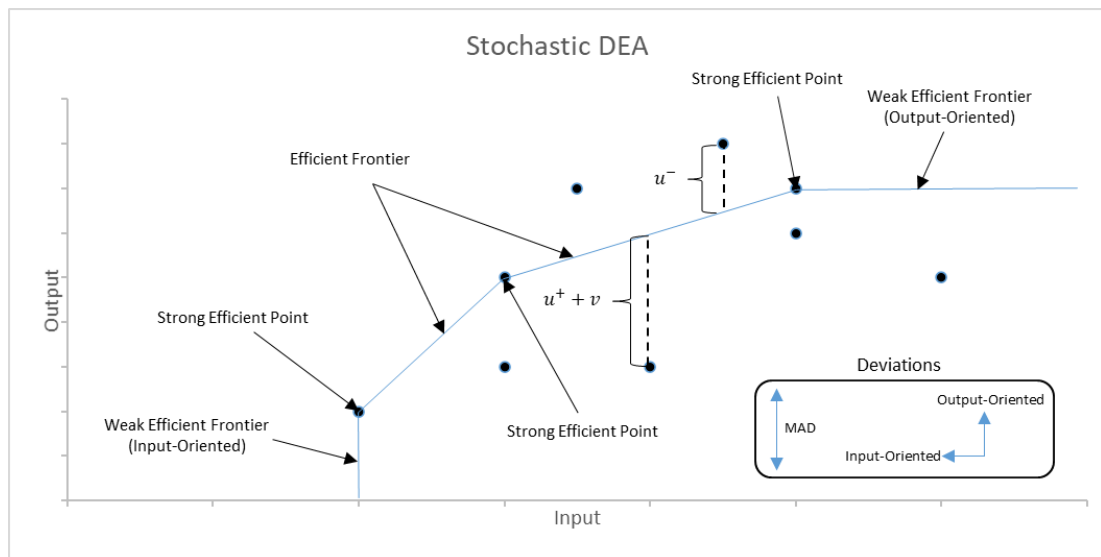


Figure 4: Stochastic DEA frontier

In output-oriented BCC, as it is shown in figure 4, the weak efficient frontier is the horizontal line continued from the efficient DMU with the greatest amount of input. The problem with the weak efficient frontier is that all the DMUs within the PPS below this line will be relatively compared to a frontier with constant value of output. This means that no matter what is the observed input, as long as the DMU is in this area, the relative efficiency will be evaluated only by its output rather than the ratio of output over input. In this paper, it has tried to suggest a modification on the constraints of the stochastic model to eliminate the weak efficient frontier in order to obtain a better estimation of production function. In the presented method, one output and multiple inputs case is considered. The reason is that although the presented stochastic model by Banker was suggested to be extended to multiple outputs in a similar way, the first set of constraints includes nonlinear terms. Hence, the calculations will be less controllable than for the single output.

what's more, initially the model is studied for one input and one output. Later, the extended multiple inputs is discussed.

To start with, it is worth to take a precise look at the Banker's stochastic Model. There are four variables in the following model:

$$\begin{aligned}
 & \text{Minimize } \sum_{j=1}^n (u_j^+ + u_j^- + cv_j) \\
 & \text{subject to} \\
 & \text{for each } j = 1, \dots, n \text{ and for all } k = 1, \dots, n \text{ and } k \neq j \\
 & (x_k - x_j)w_j + (v_j - v_k) + (u_j^+ - u_j^- - u_k^+ + u_k^-) \geq y_k - y_j \\
 & \sum_{j=1}^n (u_j^+ - u_j^-) = 0 \\
 & w_j \geq 0, \quad v_j, u_j^+, u_j^- \geq 0
 \end{aligned}$$

u_j^+ : the deviation of DMU_j from the frontier due to error (negative residual)

u_j^- : the deviation of DMU_j from the frontier due to error (positive residual)

v_j : the deviation of DMU_j from the frontier due to technical inefficiency (negative residual)

w_j : the slope of the estimated monotone increasing concave frontier at the point of the efficient output-oriented DMU_j

Most studies formerly done regarding this topic, have had the least attention to the variable w_j , which shows the slope of the frontier in an one input and one output stochastic model (2D space) or the normal vector of the frontier hyperplane in a multiple dimensional space. Nevertheless, in this paper, the effort is to concentrate on this variable and its potential to improve the model.

The suggested method for enhancing the weak efficient frontier toward the best practice frontier as a piecewise linear monotonically increasing and concave production function, could be applied in steps as given below:

Step 1: Solving the model as it is given above with observed inputs and outputs in GAMS software to obtain values for variables.

Note: the constant c could be defined by the examiner considering the importance of the type of deviations as well as its limits given before (*stochastic possibility*: $\frac{1}{n} < c < 2$, *Definitely MAD*: $c \geq 2$, *Definitely Classic DEA* $\leq 1/n$). In this research, a sensitivity analysis has been done using GAMS, to examine different amounts of this constant so that the possible effects could be investigated. The report can be found in appendix E.

Step 2: searching and finding the smallest non-zero value for w_j and naming it β .

Note: if there are any w_j with value zero, it shows that the corresponding DMU_j is placed on or compared with the weak efficient frontier.

Step 3: Adding a constraint stating that w_j have to be greater than or equal to β .

Note: for changing the weak efficient frontier in an output-oriented approach, corresponding w_j which is zero should be manipulated in a way that does not violate the assumptions and constraints made beforehand. Moreover, it should be able to resemble characteristics of a new efficient frontier as close as possible to the production function. In figure 5, possible outcomes are shown. However, only one of them is feasible to be applied. As it is illustrated in the figure, line OT is the weak efficient frontier with $w_j^* = w_{OT}^* = 0$. To modify the slope, different options are available:

1. \overline{OC} : it violates the assumption of monotonically increasing function.
2. \overline{OA} : it violates the assumption of concavity.
3. \overline{OB} : it doesn't violate any assumption but it's subjective to decide which value between zero and w_j^* .
4. $\overline{OO'}$: It doesn't violate any assumption. Also, it's vital to the consistency of the method.

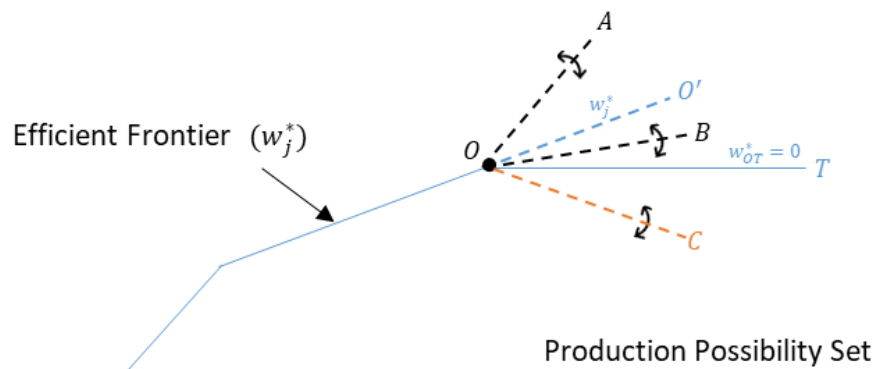


Figure 5: Possible manipulations of weak efficient frontier

By applying this condition, it will be guaranteed that the frontier line will not be horizontal. Also, since the monotonically increasing concave frontier is already satisfied by other constraints, it is definite that the so-called weak efficient frontier will be modified to continue with the same slope as the closest efficient frontier with minimum slope.

Step 4: Solving the following model by the GAMS software.

Note: the non-negativity constraint for w_j is redundant. This means that the number of constraints stays the same.

$$\text{Minimize } \sum_{j=1}^n (u_j^+ + u_j^- + cv_j)$$

subject to

for each $j = 1, \dots, n$ and for all $k = 1, \dots, n$ and $k \neq j$

$$(x_k - x_j)w_j + (v_j - v_k) + (u_j^+ - u_j^- - u_k^+ + u_k^-) \geq y_k - y_j$$

$$\sum_{j=1}^n (u_j^+ - u_j^-) = 0$$

$$w_j \geq \beta \text{ for each } j = 1, \dots, n$$

$$v_j, u_j^+, u_j^- \geq 0$$

3.5 Illustrative Application

To show the application, Banker has applied his method to a production situation, although, he didn't consider any DMU in the weak efficient area. Hence, to illustrate the modified method and its comparison with the classic effort, Banker's example is presented with some added values to the production data.

3.5.1 One Input and One Output

The monthly units produced in a company are given in the next table as the input while the single output is the labor hours. The first twelve months were given by Banker and the last three are added to examine the modified model.

Table 3: One input and one output data

Month	Output	Input
1	416.092	2089.51
2	349.785	1919.35
3	399.403	1974.43
4	455.73	2117.16
5	360.803	1792.88
6	396.241	1818.82
7	272.435	1537.66
8	312.949	1598.41
9	314.229	1701.87
10	416.09	1868.95
11	290.686	1554.14
12	260.762	1436.05
13	500	2117.16
14	520	2000
15	475	2050

First, the classic stochastic model is used to determine the best months in terms of labor productivity. The weight c is selected in the range of mere DEA to mere MAD and the corresponding graphs are presented for a better illustration of the sensitivity of the estimates on deviations caused by inefficiency or random factors.

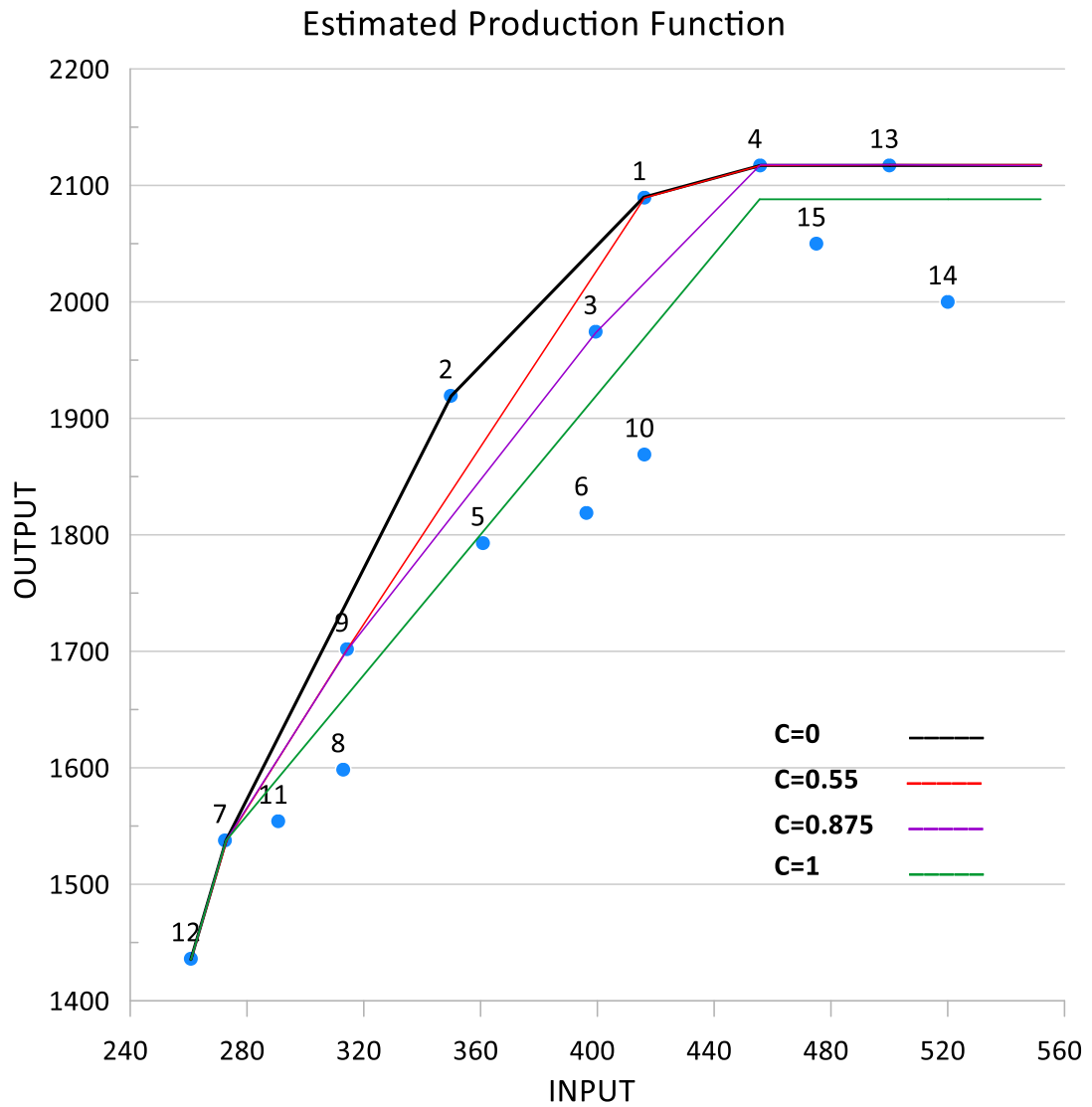


Figure 6: SDEA estimated production frontier

As it can be seen, by increasing the weight c , the estimated production frontier tends to move further down and as a result, symmetric deviation due to random factors appears. Concavity and monotonically increasing postulates are reasonably considered since by increasing the labor hour, the labor productivity tends to decrease due to the traffic and confusion at upper capacity operation, but it doesn't always mean that the productivity level stops at a point (in this example point number 4) and from that point by increasing the labor, production doesn't increase (points 13 & 14 & 15 are evaluated by weak frontier). In the presented modified model, the estimated production function

will continue growing with the least possible slope giving a more realistic approximation by turning the weak efficient frontier into an efficient frontier. It is obvious that by increasing the number of DMUs and finding an appropriate corresponding c value, the possibility of having a smooth frontier similar to logarithmic production functions rises.

By doing the explained steps toward the suggested modified model, the new frontiers will be obtained which are shown in the following graph.

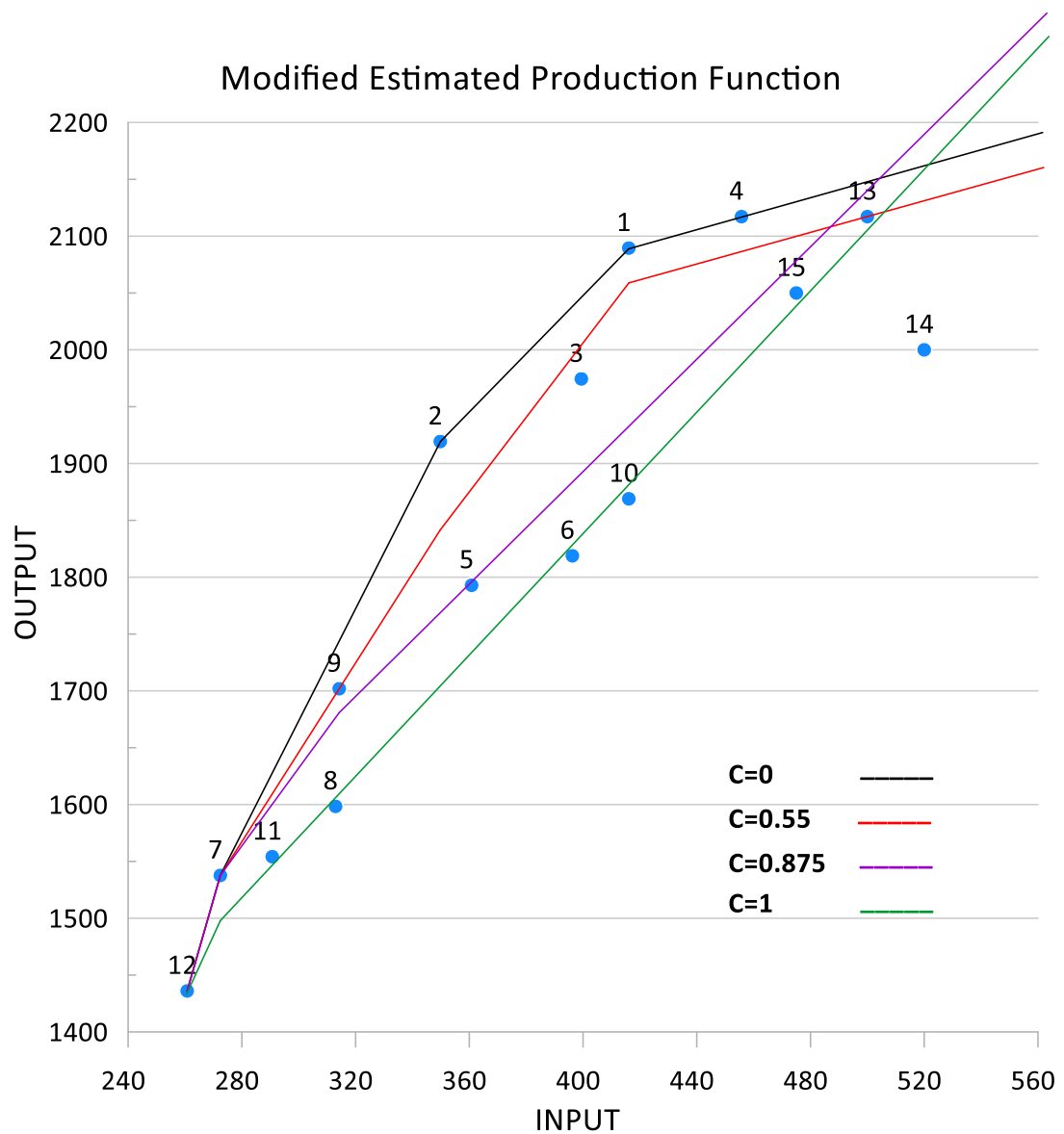


Figure 7: Modified SDEA estimated production frontier

Points 13, 14 and 15 are not compared to a weak frontier anymore. Moreover, the evaluation criteria have changed in many cases. For example, in the case $c = 0.55$, after solving the banker's traditional SDEA model by GAMS software, minimum nonzero w_j equals 0.7. Hence, $\beta = 0.7$ and by adding the new constraint to the model, new values for the slope and the deviations will be obtained. As it can be seen in figures 6 and 7, and also appendix A, DMUs 12, 7 and 9 that were strong efficient points, stay the same in the modified model. However, DMUs 1 and 4 that were strong efficient points before, are efficient but with error deviation in modified model. Therefore, it seems that they were located on the frontier because they had some error in their measurement and the actual frontier passes under the points 1 and 4. Although it implies that PPS is shrunk, yet for DMUs 14 and 15, which were compared with weak efficient frontier, the PPS has become larger. This assigns them smaller efficiency scores, but better opportunity to be improved. Weak efficient DMU 13 turns to a strong efficient point and the other DMUs can be compared with the new frontier. This means that another criterion has been obtained for comparing or improving the DMUs. The related results for the classic and modified model are given in appendix A. The deviation ratio $\frac{y_j}{\hat{y}_j} = \frac{y_j}{y_j + u^+ + v - u^-}$ and the efficiency estimates $1 - \left(\frac{v_j}{y_j}\right)$ are also reported.

3.5.2 Two Input and One Output

The classic stochastic model can also work perfectly for DMUs with more than one input. However, in the multiple output case, the behavior of the model becomes nonlinear. Hence, here the try is only to show how classic and modified models work with two inputs and one output as an example for the multiple input state.

To begin with, the stochastic model presented by Banker could be developed for two inputs case as bellow:

$$\begin{aligned}
 & \text{Minimize } \sum_{j=1}^n (u_j^+ + u_j^- + cv_j) \\
 & \text{subject to} \\
 & \text{for each } j = 1, \dots, n \text{ and for all } k = 1, \dots, n \text{ and } k \neq j \\
 & \sum_{i=1}^2 [(x_{ik} - x_{ij})w_{ij}] + (v_j - v_k) + (u_j^+ - u_j^- - u_k^+ + u_k^-) \geq y_k - y_j \\
 & \sum_{j=1}^n (u_j^+ - u_j^-) = 0 \\
 & w_{ij} \geq 0, \quad v_j, u_j^+, u_j^- \geq 0
 \end{aligned}$$

u_j^+ : the deviation of DMU_j from the frontier due to error (negative residual)

u_j^- : the deviation of DMU_j from the frontier due to error (positive residual)

v_j : the deviation of DMU_j from the frontier due to technical inefficiency (negative residual)

w_{ij} : the element of the normal vector to the estimated monotone increasing concave frontier plane at the point of the efficient output-oriented DMU_j

i : index of inputs ($i = 1, 2$)

j : index of the DMUs ($j = 1, 2, \dots, n$)

The steps to be taken from this model towards the modified model are close to the one with one input. However, since here we have two w for each DMU_j (w_{1j}, w_{2j}), two constraints will be added to the base model.

Step 1: Solving the model as it is given above with observed inputs and outputs in GAMS software to obtain values for variables.

Step 2: searching and finding the smallest non-zero values for w_{1j} and w_{2j} and naming them respectively β_1 and β_2 .

Note: if there are any DMU with $w_{1j} = w_{2j} = 0$, it shows that the corresponding DMU_j is placed on or compared with the weak efficient frontier (parallel to the Inputs surface in the coordinate system).

Step 3: Adding two following constraints to the basic model:

$$w_{1j} \geq \beta_1$$

$$w_{2j} \geq \beta_2$$

Step 4: Solving the following model by the GAMS software.

$$\text{Minimize } \sum_{j=1}^n (u_j^+ + u_j^- + cv_j)$$

subject to

for each $j = 1, \dots, n$ and for all $k = 1, \dots, n$ and $k \neq j$

$$\sum_{i=1}^2 [(x_{ik} - x_{ij})w_{ij}] + (v_j - v_k) + (u_j^+ - u_j^- - u_k^+ + u_k^-) \geq y_k - y_j$$

$$\sum_{j=1}^n (u_j^+ - u_j^-) = 0$$

$$w_{1j} \geq \beta_1$$

$$w_{2j} \geq \beta_2$$

$$v_j, u_j^+, u_j^- \geq 0$$

To demonstrate the model better, a numerical example is given below. An imagined table of data containing two inputs and one output for thirteen DMUs. First, the classic model is applied considering four different values for the weight c . Then, the modified

model is used to find the variables. The related tables of results are constructed in appendix B. The deviation ratio $\frac{y_j}{\hat{y}_j} = \frac{y_j}{y_j + u^+ + v - u^-}$ and the efficiency estimates $1 - \left(\frac{v_j}{y_j}\right)$ are also reported.

Table 4: Two input and one output data

	Input 1	Input 2	Output
DMU 1	1	7	2
DMU 2	3	4	2.5
DMU 3	3.5	4.3	3.32
DMU 4	8	1	1
DMU 5	1.5	10	3
DMU 6	4.5	6	5.75
DMU 7	7.5	9	6
DMU 8	12	1.5	2
DMU 9	8	10	6
DMU 10	1.6	6	1.8
DMU 11	8	7	2
DMU 12	6.4	6.7	4
DMU 13	10	11.6	5
DMU 14	4	4.5	3.21
DMU 15	1.5	8	2.12

3.6 Developing DEA and quantile regression to Stochastic Frontier Analysis (SFA)

Banker also extended DEA with an alternative stochastic model combining the deviations u_j^+ and v_j . The model was specified as:

$$\text{Minimize } \sum_{j=1}^n (\tau u_j^- + (1 - \tau)v_j^+)$$

subject to

for each $j = 1, \dots, n$ and for all $k = 1, \dots, n$ and $k \neq j$

$$(x_k - x_j)w_j - (v_k^+ - v_j^+) + (u_k^- - u_j^-) \geq y_k - y_j$$

$$w_j, u_j^-, v_j^+ \geq 0$$

The model can be considered as a nonparametric quantile regression with convexity and monotonicity assumptions that is able to estimate a piecewise linear production function. Hence, there is no need for specifying a theoretical functional form for production. Although, τ needs to be specified a priori. According to quantile regression notion, 100τ is the percentage of the observations appearing beneath the estimated production frontier. Therefore, the effectiveness of this SDEA model depends highly on choosing the appropriate value for τ .

If we assume that there are N DMUs producing a single output y using M inputs $x = (x_1, x_2, \dots, x_M)$, the frontier and observed production functions can be produced as follow:

$$y^f = f(x_1, x_2, \dots, x_M)$$

and

$$y = f(x_1, x_2, \dots, x_M) + \varepsilon$$

Here, ε is the error term demonstrating the deviation from the actual production frontier. Various distributional assumptions on ε result in different estimated production models. Negative ε represents a deterministic model in which the deviation is only due to inefficiency. On the other hand, if ε is distributed two-sided with mean zero, the model reflects only the statistical noise. As a result, any combination of these two leads to a stochastic frontier. In the mentioned model, the composition of inefficiency and statistical noise happens when $0.5 \leq \tau \leq 1$. $\tau = 1$ leads to absolute DEA with inefficiency deviations while $\tau = 0.5$ denotes nonparametric median regression taking into account only statistical errors.

3.6.1 Modification for a better approximation

The same steps as before are applied to the model to eliminate the weak efficiency frontier. First, the model is examined with GAMS software to obtain values for variables. Considering the smallest nonzero w_j and assigning it to the constant β , allows us to add a constraint to the model. The improved model will be as the following:

$$\text{Minimize } \sum_{j=1}^n (\tau u_j^- + (1 - \tau)v_j^+)$$

subject to

for each $j = 1, \dots, n$ and for all $k = 1, \dots, n$ and $k \neq j$

$$(x_k - x_j)w_j - (v_k^+ - v_j^+) + (u_k^- - u_j^-) \geq y_k - y_j$$

$$w_j \geq \beta$$

$$u_j^-, v_j^+ \geq 0$$

Now to illustrate this case, we suppose one output is produced using one input. We generate 50 random observations with inputs $x \sim U(1,10)$ and production function $y^f = x^{0.5}$ (Cobb-Douglas production function with productivity and labor equal to one, capital x and capital's share of 0.5). Therefore, related observed DMUs could be found by inserting $\varepsilon = u - v$ to the production function:

$$y = \exp(\varepsilon) y^f$$

To have a stochastic frontier, we specify half normal-normal ε with $v \sim |N(0, \sigma_v)|$ and $u \sim N(0, \sigma_u)$ resulting skew-normal distributional form. As it is presented in Azzalini & Capitanio (2013) and Jradi & Ruggiero (2019) have mentioned in their paper, the optimal value τ^* can be obtained from this formula:

$$\tau^* = 0.5 + \frac{1}{\pi} \arctan(\lambda)$$

Where $\lambda = \frac{\sigma_v}{\sigma_u}$.

In this paper we assumed $\sigma_v = 0.2$ and $\sigma_u = 0.1$, consequently $\tau^* = 0.8524$. For simulation, generated observation points, production function, stochastic model, and modified model have been graphed on a diagram for two sets of data to indicate their differences.

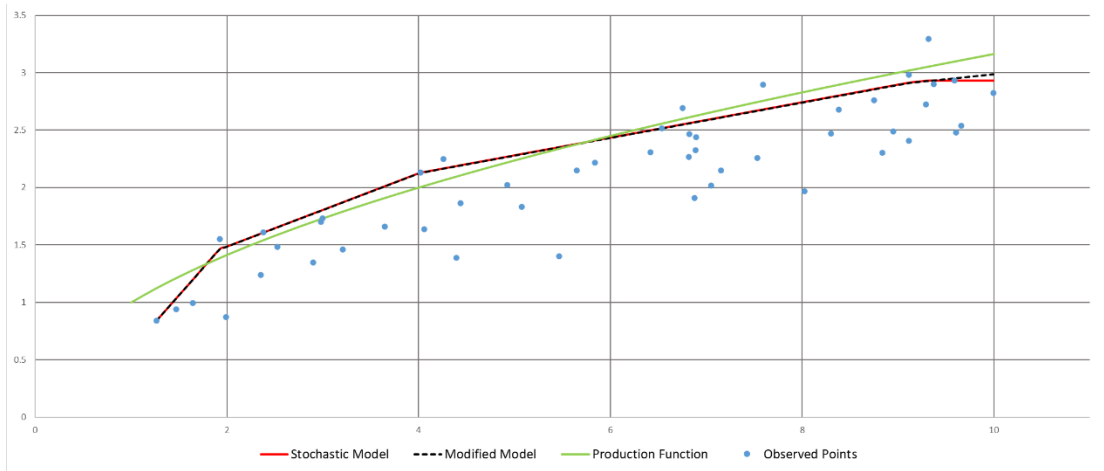


Figure 8: Production function, stochastic and modified model (dataset 1)

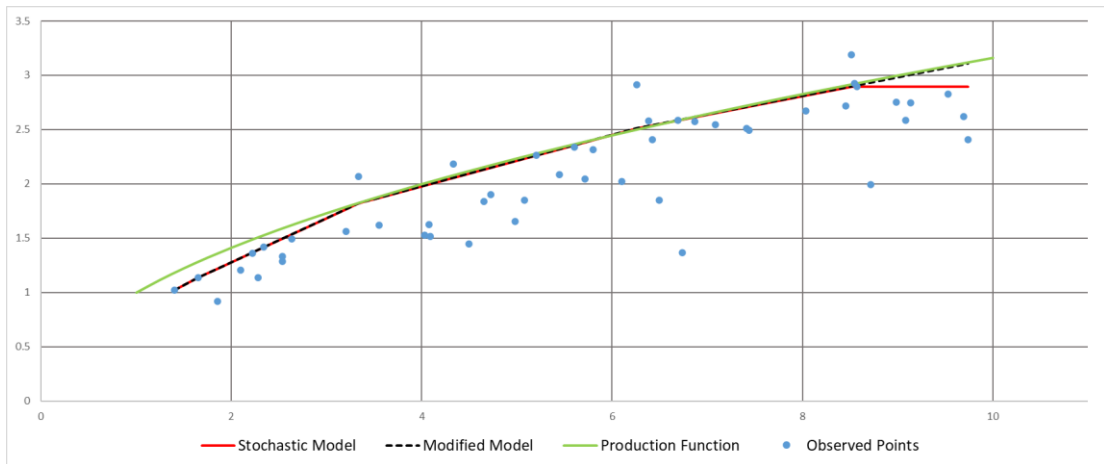


Figure 9: Production function, stochastic and modified model (dataset 2)

To be precise, the first input is generated 50 times randomly by MS Excel. Then, deviation components are produced under normal distribution. Accordingly,

production function output is obtained by replacing input in y^f and observed output by $y = \exp(\varepsilon) y^f$. Then, using GAMS the model was solved and after specifying β , the modified model was examined to get the new u_j^- and v_j^+ . The mean squared error (MSE) between production function and both stochastic and modified frontiers are calculated. In the first example by applying modification, MSE has reduced from 0.0108 to 0.0099 while in the second example, the change is even more significant. From 0.00685 to less than half (0.00339). All the related data is attached in appendices C & D.

Chapter 4

CONCLUSION

4.1 Summary of Methods

In DEA methodology, an important part is SDEA in which various stochastic models are studied according to the possibility of random deviations in observations. Among different approaches to SDEA, Banker's SDEA model was chosen to be the foundation of the presented investigation. The main goal was to address the weak efficient frontier considering the production function postulates in order to imitate the underlying true frontier. For this purpose, firstly, Banker's model including the integration of MAD and DEA is considered and after applying the suggested modification, corresponding numeric examples for one and two inputs with single output are presented. Secondly, Banker's model composed of quantile regression and DEA assuming optimal quantile is examined with suggested change in the constraints and numerical examples are provided to give a better illustration. In both cases, a new benchmark for estimating relative efficiency is obtained that not only follows general production function properties such as nonnegativity, monotonicity (nondecreasing), and concavity but also seems to be able to show a smoother frontier, especially in the weak efficient area.

4.2 Future Study

In this study, the modeling is done using GAMS software. Since for applying the suggested modification, we need to first, solve the original model, then, find the lower bound for the new constraint, and again solve the modified model, the procedure is

time-consuming. Hence, it can be improved using programming languages such as MATLAB and Python to do all the steps in one iteration.

A suggested area to focus on could be the case of multiple outputs. Banker (1988) introduced alternative models and extensions to deal with multiple outputs case which can be examined by applying the modification method. However, the first model used in this study becomes a nonlinear programming problem when multiple outputs are considered. Thus, the linearization methods could be helpful to address this issue.

Another potential area of investigation is sensitivity and stability analysis. The situation of perturbations of all inputs, output, and simultaneously all inputs and output could be studied on the modified models.

Although in all discussed cases an illustrative example is provided, the application to a real-world case is worthy for future research.

Moreover, finding a method for specifying the weight “c” based on the importance of the deviations and a specific desired estimate of the frontier could be studied in the future.

Finally, for the second model which was an integration of DEA and quantile regression, assigning other distributions rather than normal distribution to the error term could be a potential research. Also, trying different sets of data obtained from different production functions could be discussed elsewhere.

REFERENCES

- Afriat, S. N. (1972). Efficiency Estimation of Production Functions. *International Economic Review*, 13(3), 568. <https://doi.org/10.2307/2525845>
- Ahn, T., Charnes, A., & Cooper, W. W. (1988). Efficiency characterizations in different DEA models. *Socio-Economic Planning Sciences*, 22(6), 253–257. [https://doi.org/10.1016/0038-0121\(88\)90007-9](https://doi.org/10.1016/0038-0121(88)90007-9)
- Ahn, T., & Seiford, L. M. (1993). Sensitivity of DEA to Models and Variable Sets in a Hypothesis Test Setting: The Efficiency of University Operations. In *Creative and Innovative Approaches to the Science of Management*.
- Aigner, D., & Chu, S. F. (1968). On estimating the industry production function. *American Economic Review*, 58(4), 826–839.
- Aigner, D., Lovell, C. A. K., & Schmidt, P. (1977). Formulation and estimation of stochastic frontier production function models. *Journal of Econometrics*, 6(1), 21–37. [https://doi.org/10.1016/0304-4076\(77\)90052-5](https://doi.org/10.1016/0304-4076(77)90052-5)
- Ali, A. I., Cook, W. D., & Seiford, L. M. (1991). Strict vs. Weak Ordinal Relations for Multipliers in Data Envelopment Analysis. *Management Science*, 37(6), 733–738. <https://doi.org/10.1287/mnsc.37.6.733>
- Alshare, K. (1998). *Robustness of Data Envelopment (DEA) Efficiency Classifications: An Empirical Study of Jordanian Hospitals*. The University of

Texas at Arlington.

Azzalini, A., & Capitanio, A. (2013). *The Skew-Normal and Related Families*. Cambridge University Press. <https://doi.org/10.1017/CBO9781139248891>

Banker, R. D. (1988). Stochastic data envelopment analysis. *Working Paper, Carnegie Mellon University*.

Banker, R. D. (1993). Maximum likelihood, consistency and data envelopment analysis. A statistical foundation. *Management Science*, 39(10), 1265–1273. <https://doi.org/10.1287/mnsc.39.10.1265>

Banker, R. D., Charnes, A., & Cooper, W. W. (1984). Some Models for Estimating Technical and Scale Inefficiencies in Data Envelopment Analysis. *Management Science*, 30(9), 1078–1092. <https://doi.org/10.1287/mnsc.30.9.1078>

Banker, R. D., Charnes, A., Cooper, W. W., & Maindiratta, A. (1988). A Comparison of DEA and Translog Estimates of Production Frontiers Using Simulated Observations from a Known Technology. In *Applications of Modern Production Theory: Efficiency and Productivity* (pp. 33–55). Springer Netherlands. https://doi.org/10.1007/978-94-009-3253-1_2

Banker, R. D., Conrad, R., & Strauss, R. (1986). A Comparative Application of Data Envelopment Analysis and Translog Methods: An Illustrative Study of Hospital Production. *Management Science*, 32(1), 30–44. <https://doi.org/10.1287/mnsc.32.1.30>

- Banker, R. D., Datar, S. M., & Kemerer, C. F. (1991). Model to evaluate variables impacting the productivity of software maintenance projects. In *Management Science* (Vol. 37, Issue 1, pp. 1–18). <https://doi.org/10.1287/mnsc.37.1.1>
- Banker, R. D., Gadh, V., & Gorr, W. (1993). A Monte Carlo comparison of two production frontier estimation methods: Corrected ordinary least squares and data envelopment analysis. *European Journal of Operational Research*, 67(3), 332–343. [https://doi.org/10.1016/0377-2217\(93\)90289-Y](https://doi.org/10.1016/0377-2217(93)90289-Y)
- Banker, R. D., Kotarac, K., & Neralić, L. (2015). Sensitivity and stability in stochastic data envelopment analysis. *Journal of the Operational Research Society*, 66(1), 134–147. <https://doi.org/10.1057/jors.2012.182>
- Banker, R. D., & Maindiratta, A. (1988). Nonparametric Analysis of Technical and Allocative Efficiencies in Production. *Econometrica*, 56(6), 1315. <https://doi.org/10.2307/1913100>
- Banker, R. D., & Maindiratta, A. (1992). Maximum likelihood estimation of monotone and concave production frontiers. *Journal of Productivity Analysis*, 3(4), 401–415. <https://doi.org/10.1007/BF00163435>
- Banker, R. D., & Morey, R. C. (1986a). Efficiency Analysis for Exogenously Fixed Inputs and Outputs. *Operations Research*, 34(4), 513–521. <https://doi.org/10.1287/opre.34.4.513>
- Banker, R. D., & Morey, R. C. (1986b). USE OF CATEGORICAL VARIABLES IN

DATA ENVELOPMENT ANALYSIS. *Management Science*, 32(12), 1613–1627. <https://doi.org/10.1287/mnsc.32.12.1613>

Bauer, P. (1990). Recent developments in the econometric estimation of frontiers. *Journal of Econometrics*, 46(1–2), 39–56. [https://doi.org/10.1016/0304-4076\(90\)90046-V](https://doi.org/10.1016/0304-4076(90)90046-V)

Bjurek, H., Hjalmarsson, L., & Forsund, F. (1990). Deterministic parametric and nonparametric estimation of efficiency in service production. *Journal of Econometrics*, 46(1–2), 213–227. [https://doi.org/10.1016/0304-4076\(90\)90056-Y](https://doi.org/10.1016/0304-4076(90)90056-Y)

Charnes, A., Clark, C. T., Cooper, W. W., & Golany, B. (1984). A developmental study of data envelopment analysis in measuring the efficiency of maintenance units in the U.S. air forces. *Annals of Operations Research*, 2(1), 95–112. <https://doi.org/10.1007/BF01874734>

Charnes, A., & Cooper, W. W. (1959). Chance-Constrained Programming. *Management Science*, 6(1), 73–79. <https://doi.org/10.1287/mnsc.6.1.73>

Charnes, A., Cooper, W. W., Golany, B., Seiford, L. M., & Stutz, J. (1985). Foundations of data envelopment analysis for Pareto-Koopmans efficient empirical production functions. *Journal of Econometrics*, 30(1–2), 91–107. [https://doi.org/10.1016/0304-4076\(85\)90133-2](https://doi.org/10.1016/0304-4076(85)90133-2)

Charnes, A., Cooper, W. W., Huang, Z. M., & Sun, D. B. (1990). Polyhedral Cone-

Ratio DEA Models with an illustrative application to large commercial banks. *Journal of Econometrics*, 46(1–2), 73–91. [https://doi.org/10.1016/0304-4076\(90\)90048-X](https://doi.org/10.1016/0304-4076(90)90048-X)

Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research*, 2(6), 429–444. [https://doi.org/10.1016/0377-2217\(78\)90138-8](https://doi.org/10.1016/0377-2217(78)90138-8)

Charnes, A., Cooper, W. W., Seiford, L. M., & Stutz, J. (1982). A multiplicative model for efficiency analysis. *Socio-Economic Planning Sciences*, 16(5), 223–224. [https://doi.org/10.1016/0038-0121\(82\)90029-5](https://doi.org/10.1016/0038-0121(82)90029-5)

Charnes, A., Cooper, W. W., Wei, Q. L., & Huang, Z. M. (1989). Cone ratio data envelopment analysis and multi-objective programming. *International Journal of Systems Science*, 20(7), 1099–1118. <https://doi.org/10.1080/00207728908910197>

Cook, W. D., & Kress, M. (1991). A multiple criteria decision model with ordinal preference data. *European Journal of Operational Research*, 54(2), 191–198. [https://doi.org/10.1016/0377-2217\(91\)90297-9](https://doi.org/10.1016/0377-2217(91)90297-9)

Cook, W. D., Kress, M., & Seiford, L. M. (1993). On the Use of Ordinal Data in Data Envelopment Analysis. *The Journal of the Operational Research Society*, 44(2), 133. <https://doi.org/10.2307/2584361>

Cook, W. D., Kress, M., & Seiford, L. M. (1996). Data Envelopment Analysis in the

- Presence of Both Quantitative and Qualitative Factors. *Journal of the Operational Research Society*, 47(7), 945–953. <https://doi.org/10.1057/palgrave.jors.0470711>
- Cooper, W. W., & Tone, K. (1997). Measures of inefficiency in data envelopment analysis and stochastic frontier estimation. *European Journal of Operational Research*, 99(1), 72–88. [https://doi.org/10.1016/S0377-2217\(96\)00384-0](https://doi.org/10.1016/S0377-2217(96)00384-0)
- Daneshvar, S., Izbirak, G., & Javadi, A. (2014). Sensitivity analysis on modified variable returns to scale model in Data Envelopment Analysis using facet analysis. *Computers & Industrial Engineering*, 76, 32–39. <https://doi.org/10.1016/j.cie.2014.07.016>
- Dyson, R. G., & Thanassoulis, E. (1988). Reducing weight flexibility in data envelopment analysis. *Journal of the Operational Research Society*, 39(6), 563–576. <https://doi.org/10.1057/jors.1988.96>
- Epstein, M. K., & Henderson, J. C. (1989). Data Envelopment Analysis for Managerial Control and Diagnosis. *Decision Sciences*, 20(1), 90–119. <https://doi.org/10.1111/j.1540-5915.1989.tb01399.x>
- Färe, R., Grosskopf, S., Logan, J., & Lovell, C. A. K. (1985). Measuring Efficiency in Production: With an Application to Electric Utilities. In *Managerial Issues in Productivity Analysis* (pp. 185–214). Springer Netherlands. https://doi.org/10.1007/978-94-009-4982-9_8
- Farrell, M. J. (1957). The Measurement of Productive Efficiency. *Journal of the Royal*

Statistical Society. Series A (General), 120(3), 253.
<https://doi.org/10.2307/2343100>

Golany, B. (1988). Note-A Note on Including Ordinal Relations Among Multipliers in Data Envelopment Analysis. *Pubsonline.Informs.Org*, 34(8), 1029–1033.
<https://doi.org/10.1287/mnsc.34.8.1029>

Greene, W. H. (1980a). On the estimation of a flexible frontier production model. *Journal of Econometrics*, 13(1), 101–115. [https://doi.org/10.1016/0304-4076\(80\)90045-7](https://doi.org/10.1016/0304-4076(80)90045-7)

Greene, W. H. (1980b). Maximum likelihood estimation of econometric frontier functions. *Journal of Econometrics*, 13(1), 27–56. [https://doi.org/10.1016/0304-4076\(80\)90041-X](https://doi.org/10.1016/0304-4076(80)90041-X)

Huang, Z. M., & Li, S. (1996). Dominance stochastic models in data envelopment analysis. *European Journal of Operational Research*, 95(2), 390–403.
[https://doi.org/10.1016/0377-2217\(95\)00293-6](https://doi.org/10.1016/0377-2217(95)00293-6)

Jondrow, J., Lovell, C. A. K., Materov, I., & Schmidt, P. (1982). On the estimation of technical inefficiency in the stochastic frontier production function model. *Journal of Econometrics*, 19(2–3), 233–238. [https://doi.org/10.1016/0304-4076\(82\)90004-5](https://doi.org/10.1016/0304-4076(82)90004-5)

Jradi, S., & Ruggiero, J. (2019). Stochastic data envelopment analysis: A quantile regression approach to estimate the production frontier. *European Journal of*

Operational Research, 278(2), 385–393.

<https://doi.org/10.1016/j.ejor.2018.11.017>

Kall, P. (1976). *Stochastic Linear Programming* (Vol. 21). Springer Berlin Heidelberg. <https://doi.org/10.1007/978-3-642-66252-2>

Kamakura, W. A. (1988). Note-A Note on "The Use of Categorical Variables in Data Envelopment Analysis. *Management Science*, 34(10), 1273–1276. <https://doi.org/10.1287/mnsc.34.10.1273>

Kopp, R. J., & Diewert, W. E. (1982). The decomposition of frontier cost function deviations into measures of technical and allocative efficiency. *Journal of Econometrics*, 19(2–3), 319–331. [https://doi.org/10.1016/0304-4076\(82\)90008-2](https://doi.org/10.1016/0304-4076(82)90008-2)

Kuosmanen, T. (2008). Representation theorem for convex nonparametric least squares. *Econometrics Journal*, 11(2), 308–325. <https://doi.org/10.1111/j.1368-423X.2008.00239.x>

Kuosmanen, T., & Johnson, A. L. (2010). Data envelopment analysis as nonparametric least-squares regression. *Operations Research*, 58(1), 149–160. <https://doi.org/10.1287/opre.1090.0722>

Kuosmanen, T., & Kortelainen, M. (2012). Stochastic non-smooth envelopment of data: Semi-parametric frontier estimation subject to shape constraints. *Journal of Productivity Analysis*, 38(1), 11–28. <https://doi.org/10.1007/s11123-010-0201-3>

- Land, K. C., Lovell, C. A. K., & Thore, S. (1993). Chance-constrained Data Envelopment Analysis. *Managerial and Decision Economics*, 14(6), 541–554.
- Lovell, C. A. K., & Schmidt, P. (1988). A Comparison of Alternative Approaches to the Measurement of Productive Efficiency. *Springer*, 3–32.
https://doi.org/10.1007/978-94-009-3253-1_1
- Meeusen, W., Van, J., & Broeck, D. (1977). Efficiency Estimation from Cobb-Douglas Production Functions with Composed Error. In *International Economic Review* (Vol. 18, Issue 2).
- Olesen, O. B., & Petersen, N. C. (1995). Chance Constrained Efficiency Evaluation. *Management Science*, 41(3), 442–457. <https://doi.org/10.1287/mnsc.41.3.442>
- Richmond, J. (1974). Estimating the Efficiency of Production. In *International Economic Review* (Vol. 15, Issue 2).
- Rousseau, J. J., & Semple, J. H. (1993). Notes: Categorical Outputs in Data Envelopment Analysis. *Management Science*, 39(3), 384–386.
<https://doi.org/10.1287/mnsc.39.3.384>
- Schmidt, P. (1976). On the Statistical Estimation of Parametric Frontier Production Functions. *The Review of Economics and Statistics*, 58(2), 238.
<https://doi.org/10.2307/1924032>
- Schmidt, P., & Lovell, C. A. K. (1979). Estimating technical and allocative

- inefficiency relative to stochastic production and cost frontiers. *Journal of Econometrics*, 9(3), 343–366. [https://doi.org/10.1016/0304-4076\(79\)90078-2](https://doi.org/10.1016/0304-4076(79)90078-2)
- Seiford, L. M., & Thrall, R. M. (1990). Recent developments in DEA. *Journal of Econometrics*, 46(1–2), 7–38. [https://doi.org/10.1016/0304-4076\(90\)90045-U](https://doi.org/10.1016/0304-4076(90)90045-U)
- Sengupta, J. K. (1987). Data envelopment analysis for efficiency measurement in the stochastic case. *Computers and Operations Research*, 14(2), 117–129. [https://doi.org/10.1016/0305-0548\(87\)90004-9](https://doi.org/10.1016/0305-0548(87)90004-9)
- Sengupta, J. K. (1995). *Dynamics of Data Envelopment Analysis*. Springer Netherlands. <https://doi.org/10.1007/978-94-015-8506-4>
- Sengupta, J. K. (1996a). Dynamic data envelopment analysis. *International Journal of Systems Science*, 27(3), 277–284. <https://doi.org/10.1080/00207729608929214>
- Sengupta, J. K. (1996b). The efficiency distribution approach in data envelopment analysis: An application. *Journal of the Operational Research Society*, 47(11), 1387–1397. <https://doi.org/10.1057/jors.1996.174>
- Thompson, R. G., Langemeier, L. N., Lee, C. T., Lee, E., & Thrall, R. M. (1990). The role of multiplier bounds in efficiency analysis with application to Kansas farming. *Journal of Econometrics*, 46(1–2), 93–108. [https://doi.org/10.1016/0304-4076\(90\)90049-Y](https://doi.org/10.1016/0304-4076(90)90049-Y)
- Varian, H. R. (1984). *The Nonparametric Approach to Production Analysis*.

Econometrica, 52(3), 579. <https://doi.org/10.2307/1913466>

Wang, Y., Wang, S., Dang, C., & Ge, W. (2014). Nonparametric quantile frontier estimation under shape restriction. *European Journal of Operational Research*, 232(3), 671–678. <https://doi.org/10.1016/j.ejor.2013.06.049>

Wong, Y. H. B., & Beasley, J. E. (1990). Restricting weight flexibility in data envelopment analysis. *Journal of the Operational Research Society*, 41(9), 829–835. <https://doi.org/10.1057/jors.1990.120>

Zieschang, K. D. (1983). A note on the decomposition of cost efficiency into technical and allocative components. *Journal of Econometrics*, 23(3), 401–405. [https://doi.org/10.1016/0304-4076\(83\)90067-2](https://doi.org/10.1016/0304-4076(83)90067-2)

APPENDICES

Appendix A: 1 input 1 output illustrated application of Banker’s SDEA and suggested modified model for different values of the weight “c”

Observations			Classic Model								Modified Model							
DMU	X	Y	w	u+	u-	v	c	z	Deviation Ratio	Efficiency Estimates	β	w	u+	u-	v	z	Deviation Ratio	Efficiency Estimates
1	416.09	2089.51	0.70	0.00	0.00	0.00	0.00	0.00	1.000	1.000	0.70	2.57	0.00	0.00	0.00	0.00	1.000	1.000
2	349.79	1919.35	2.57	0.00	0.00	0.00			1.000	1.000		2.57	0.00	0.00	0.00		1.000	1.000
3	399.40	1974.43	2.57	0.00	0.00	72.25			0.965	0.963		2.57	0.00	0.00	72.25		0.965	0.963
4	455.73	2117.16	0.00	0.00	0.00	0.00			1.000	1.000		0.70	0.00	0.00	0.00		1.000	1.000
5	360.80	1792.88	2.57	0.00	0.00	154.75			0.921	0.914		2.57	0.00	0.00	154.75		0.921	0.914
6	396.24	1818.82	2.57	0.00	0.00	219.75			0.892	0.879		2.57	0.00	0.00	219.75		0.892	0.879
7	272.44	1537.66	4.94	0.00	0.00	0.00			1.000	1.000		4.94	0.00	0.00	0.00		1.000	1.000
8	312.95	1598.41	4.94	0.00	0.00	139.17			0.920	0.913		4.94	0.00	0.00	139.17		0.920	0.913
9	314.23	1701.87	4.94	0.00	0.00	42.03			0.976	0.975		4.94	0.00	0.00	42.03		0.976	0.975
10	416.09	1868.95	2.57	0.00	0.00	220.56			0.894	0.882		2.57	0.00	0.00	220.56		0.894	0.882
11	290.69	1554.14	4.94	0.00	0.00	73.58			0.955	0.953		4.94	0.00	0.00	73.58		0.955	0.953
12	260.76	1436.05	8.71	0.00	0.00	0.00			1.000	1.000		8.71	0.00	0.00	0.00		1.000	1.000
13	500.00	2117.16	0.00	0.00	0.00	0.00			1.000	1.000		0.70	0.00	0.00	30.92		0.986	0.985
14	520.00	2000.00	0.00	0.00	0.00	117.16			0.945	0.941		0.70	0.00	0.00	162.04		0.925	0.919
15	475.00	2050.00	0.00	0.00	0.00	67.16			0.968	0.967		0.70	0.00	0.00	80.63		0.962	0.961

Observations			Classic Model								Modified Model							
DMU	X	Y	w	u+	u-	v	c	z	Deviation Ratio	Efficiency Estimates	β	w	u+	u-	v	z	Deviation Ratio	Efficiency Estimates
1	416.09	2089.51	0.70	0.00	0.00	0.00	0.55	606.44	1.000	1.000	0.70	0.70	0.00	30.92	0.00	649.69	1.015	1.000
2	349.79	1919.35	3.74	0.00	77.78	0.00			1.042	1.000		3.27	0.00	77.78	0.00		1.042	1.000
3	399.40	1974.43	3.74	0.00	0.00	52.68			0.974	0.973		3.27	0.00	0.00	29.54		0.985	0.985
4	455.73	2117.16	0.00	0.00	0.00	0.00			1.000	1.000		0.70	0.00	30.90	0.00		1.015	1.000
5	360.80	1792.88	3.74	0.00	0.00	89.89			0.952	0.950		3.27	0.00	0.00	84.75		0.955	0.953
6	396.24	1818.82	3.74	77.78	0.00	118.68			0.903	0.935		3.27	0.00	0.00	174.80		0.912	0.904
7	272.44	1537.66	3.93	0.00	0.00	0.00			1.000	1.000		3.93	0.00	0.00	0.00		1.000	1.000
8	312.95	1598.41	3.93	0.00	0.00	98.43			0.942	0.938		3.93	0.00	0.00	98.43		0.942	0.938
9	314.23	1701.87	3.93	0.00	0.00	0.00			1.000	1.000		3.93	0.00	0.00	0.00		1.000	1.000
10	416.09	1868.95	3.74	0.00	0.00	220.55			0.894	0.882		3.27	139.60	0.00	50.04		0.908	0.973
11	290.69	1554.14	3.93	0.00	0.00	55.23			0.966	0.964		3.93	0.00	0.00	55.23		0.966	0.964
12	260.76	1436.05	8.71	0.00	0.00	0.00			1.000	1.000		8.71	0.00	0.00	0.00		1.000	1.000
13	500.00	2117.16	0.00	0.00	0.00	0.00			1.000	1.000		0.70	0.00	0.00	0.00		1.000	1.000
14	520.00	2000.00	0.00	0.00	0.00	117.16			0.945	0.941		0.70	0.00	0.00	131.12		0.938	0.934
15	475.00	2050.00	0.00	0.00	0.00	67.16			0.968	0.967		0.70	0.00	0.00	49.71		0.976	0.976

Observations			Classic Model								Modified Model							
DMU	X	Y	w	u+	u-	v	c	z	Deviation Ratio	Efficiency Estimates	β	w	u+	u-	v	z	Deviation Ratio	Efficiency Estimates
1	416.09	2089.51	2.47	0.00	70.39	0.00	0.875	790.99	1.035	1.000	2.47	2.47	0.00	156.71	0.00	994.92	1.081	1.000
2	349.79	1919.35	2.68	0.00	77.78	0.00			1.042	1.000		2.47	0.00	150.53	0.00		1.085	1.000
3	399.40	1974.43	2.68	0.00	0.00	0.00			1.000	1.000		2.47	0.00	82.90	0.00		1.044	1.000
4	455.73	2117.16	0.00	0.00	0.00	0.00			1.000	1.000		2.47	0.00	86.34	0.00		1.043	1.000
5	360.80	1792.88	2.68	0.00	0.00	78.19			0.958	0.956		2.47	3.19	0.00	0.00		0.998	1.000
6	396.24	1818.82	2.68	147.14	0.00	0.00			0.925	1.000		2.47	64.89	0.00	0.00		0.966	1.000
7	272.44	1537.66	3.93	0.00	0.00	0.00			1.000	1.000		3.43	0.00	0.00	0.00		1.000	1.000
8	312.95	1598.41	3.93	0.00	0.00	98.43			0.942	0.938		3.43	78.09	0.00	0.00		0.953	1.000
9	314.23	1701.87	3.93	0.00	0.00	0.00			1.000	1.000		3.43	0.00	20.98	0.00		1.012	1.000
10	416.09	1868.95	2.68	1.03	0.00	149.13			0.926	0.920		2.47	63.84	0.00	0.00		0.967	1.000
11	290.69	1554.14	3.93	0.00	0.00	55.23			0.966	0.964		3.43	46.07	0.00	0.00		0.971	1.000
12	260.76	1436.05	8.71	0.00	0.00	0.00			1.000	1.000		8.71	0.00	0.00	0.00		1.000	1.000
13	500.00	2117.16	0.00	0.00	0.00	0.00			1.000	1.000		2.47	23.14	0.00	0.00		0.989	1.000
14	520.00	2000.00	0.00	0.00	0.00	117.16			0.945	0.941		2.47	189.76	0.00	0.00		0.913	1.000
15	475.00	2050.00	0.00	0.00	0.00	67.16			0.968	0.967		2.47	28.48	0.00	0.00		0.986	1.000

Observations			Classic Model								Modified Model							
DMU	X	Y	w	u+	u-	v	c	z	Deviation Ratio	Efficiency Estimates	β	w	u+	u-	v	z	Deviation Ratio	Efficiency Estimates
1	416.09	2089.51	3.00	0.00	120.84	0.00	1.00	848.42	1.061	1.000	3.00	3.00	0.00	148.23	0.00	1152.85	1.076	1.000
2	349.79	1919.35	3.00	0.00	149.62	0.00			1.085	1.000		3.00	0.00	176.99	0.00		1.102	1.000
3	399.40	1974.43	3.00	0.00	55.83	0.00			1.029	1.000		3.00	0.00	83.22	0.00		1.044	1.000
4	455.73	2117.16	0.00	0.00	29.56	0.00			1.014	1.000		3.00	0.00	56.70	0.00		1.028	1.000
5	360.80	1792.88	3.00	9.91	0.00	0.00			0.995	1.000		3.00	0.00	17.47	0.00		1.010	1.000
6	396.24	1818.82	3.00	90.30	0.00	0.00			0.953	1.000		3.00	62.90	0.00	0.00		0.967	1.000
7	272.44	1537.66	3.00	0.00	0.00	0.00			1.000	1.000		3.00	0.00	27.35	0.00		1.018	1.000
8	312.95	1598.41	3.00	60.80	0.00	0.00			0.963	1.000		3.00	33.44	0.00	0.00		0.980	1.000
9	314.23	1701.87	3.00	0.00	38.82	0.00			1.023	1.000		3.00	0.00	66.18	0.00		1.040	1.000
10	416.09	1868.95	3.00	99.72	0.00	0.00			0.949	1.000		3.00	72.32	0.00	0.00		0.963	1.000
11	290.69	1554.14	3.00	38.28	0.00	0.00			0.976	1.000		3.00	10.92	0.00	0.00		0.993	1.000
12	260.76	1436.05	8.71	0.00	0.00	0.00			1.000	1.000		6.36	0.00	0.00	0.00		1.000	1.000
13	500.00	2117.16	0.00	0.00	29.56	0.00			1.014	1.000		3.00	75.84	0.00	0.00		0.965	1.000
14	520.00	2000.00	0.00	87.60	0.00	0.00			0.958	1.000		3.00	253.00	0.00	0.00		0.888	1.000
15	475.00	2050.00	0.00	37.60	0.00	0.00			0.982	1.000		3.00	68.00	0.00	0.00		0.968	1.000

Appendix B: 2 inputs 1 output illustrated application of Banker’s SDEA and suggested modified model for different values of the weight “c”

Observations				Classical Model									Modified Model									
DMU	X1	X2	Y	w1j	w2j	u+	u-	v	c	z	Deviation Ratio	Efficiency Estimates	β_1	β_2	w1j	w2j	u+	u-	v	z	Deviation Ratio	Efficiency Estimates
1	1.00	7.00	2.00	2.00	0.00	0.00	0.00	0.00	0.000	0.000	1.000	1.000	0.400	0.083	1.50	0.08	0.00	0.00	0.00	0.000	1.000	1.000
2	3.00	4.00	2.50	0.67	1.62	0.00	0.00	0.00			1.000	1.000			0.67	1.62	0.00	0.00	0.00		1.000	1.000
3	3.50	4.30	3.32	0.40	1.25	0.00	0.00	0.00			1.000	1.000			0.40	1.19	0.00	0.00	0.00		1.000	1.000
4	8.00	1.00	1.00	0.00	2.00	0.00	0.00	0.00			1.000	1.000			0.40	1.25	0.00	0.00	0.00		1.000	1.000
5	1.50	10.00	3.00	2.00	0.00	0.00	0.00	0.00			1.000	1.000			1.50	0.08	0.00	0.00	0.00		1.000	1.000
6	4.50	6.00	5.75	0.00	0.08	0.00	0.00	0.00			1.000	1.000			0.40	0.08	0.00	0.00	0.00		1.000	1.000
7	7.50	9.00	6.00	0.00	0.00	0.00	0.00	0.00			1.000	1.000			0.40	0.08	0.00	0.00	1.20		0.833	0.800
8	12.00	1.50	2.00	0.00	2.00	0.00	0.00	0.00			0.620	0.388			0.40	1.25	0.00	0.00	1.22		0.620	0.388
9	8.00	10.00	6.00	0.00	0.00	0.00	0.00	0.00			0.802	0.753			0.40	0.08	0.00	0.00	1.48		0.802	0.753
10	1.60	6.00	1.80	5.50	3.50	0.00	0.00	0.00			1.000	1.000			5.50	3.50	0.00	0.00	0.00		1.000	1.000
11	8.00	7.00	4.60	0.00	0.08	0.00	0.00	1.23			0.789	0.732			0.40	0.08	0.00	0.00	2.63		0.636	0.428
12	6.40	6.70	4.00	0.00	0.08	0.00	0.00	1.38			0.743	0.654			0.40	1.19	0.00	0.00	1.34		0.750	0.666
13	10.00	11.60	5.00	0.00	0.00	0.00	0.00	1.00			0.833	0.800			0.40	0.08	0.00	0.00	3.42		0.594	0.317
14	4.00	4.50	3.21	0.40	1.25	0.00	0.00	0.56			0.851	0.826			0.40	1.19	0.00	0.00	0.55		0.854	0.829
15	1.50	8.00	2.12	2.00	0.00	0.00	0.00	0.88			0.707	0.585			1.50	0.08	0.00	0.00	0.71		0.748	0.663

Observations				Classical Model								Modified Model										
DMU	X1	X2	Y	w1j	w2j	u+	u-	v	c	z	Deviation Ratio	Efficiency Estimates	$\beta 1$	$\beta 2$	w1j	w2j	u+	u-	v	z	Deviation Ratio	Efficiency Estimates
1	1.00	7.00	2.00	2.00	0.00	0.00	0.00	0.00	0.750	3.393	1.000	1.000	0.139	0.183	0.90	0.18	0.00	0.00	0.00	4.375	1.000	1.000
2	3.00	4.00	2.50	0.67	1.62	0.00	0.00	0.00			1.000	1.000			0.67	1.62	0.00	0.00	0.00		1.000	1.000
3	3.50	4.30	3.32	0.14	0.89	0.00	0.00	0.00			1.000	1.000			0.14	0.78	0.00	0.00	0.00		1.000	1.000
4	8.00	1.00	1.00	0.00	2.00	0.00	0.00	0.00			1.000	1.000			0.14	0.89	0.00	0.00	0.00		1.000	1.000
5	1.50	10.00	3.00	2.00	0.00	0.00	0.00	0.00			1.000	1.000			0.90	0.18	0.00	0.00	0.00		1.000	1.000
6	4.50	6.00	5.75	0.00	0.34	0.00	0.78	0.00			1.156	1.000			0.14	0.18	0.00	0.97	0.00		1.202	1.000
7	7.50	9.00	6.00	0.00	0.00	0.00	0.00	0.00			1.000	1.000			0.14	0.18	0.00	0.25	0.00		1.044	1.000
8	12.00	1.50	2.00	0.00	2.00	0.00	0.00	0.00			1.000	1.000			0.14	0.89	0.00	0.00	0.00		0.999	0.999
9	8.00	10.00	6.00	0.00	0.00	0.00	0.00	0.00			1.000	1.000			0.14	0.18	0.00	0.00	0.00		1.000	1.000
10	1.60	6.00	1.80	5.50	3.50	0.00	0.00	0.00			1.000	1.000			5.50	3.50	0.00	0.00	0.00		1.000	1.000
11	8.00	7.00	4.60	0.00	0.34	0.13	0.00	0.59			0.865	0.873			0.14	0.18	0.85	0.00	0.00		0.844	1.000
12	6.40	6.70	4.00	0.14	0.89	0.64	0.00	0.00			0.861	1.000			0.14	0.78	0.37	0.00	0.17		0.880	0.957
13	10.00	11.60	5.00	0.00	0.00	0.00	0.00	1.00			0.833	0.800			0.14	0.18	0.00	0.00	1.57		0.761	0.686
14	4.00	4.50	3.21	0.14	0.89	0.00	0.00	0.36			0.900	0.888			0.14	0.78	0.00	0.00	0.34		0.906	0.896
15	1.50	8.00	2.12	0.90	0.18	0.00	0.00	0.51			0.805	0.758			0.85	0.19	0.00	0.00	0.50		0.810	0.766

Observations				Classical Model								Modified Model										
DMU	X1	X2	Y	w1j	w2j	u+	u-	v	c	z	Deviation Ratio	Efficiency Estimates	$\beta 1$	$\beta 2$	w1j	w2j	u+	u-	v	z	Deviation Ratio	Efficiency Estimates
1	1.00	7.00	2.00	2.00	0.00	0.00	0.00	0.00	0.875	3.524	1.000	1.000	0.045	0.208	0.75	0.21	0.00	0.00	0.00	4.149	1.000	1.000
2	3.00	4.00	2.50	0.67	1.62	0.00	0.00	0.00			1.000	1.000			0.49	1.32	0.00	0.00	0.00		1.000	1.000
3	3.50	4.30	3.32	0.05	0.61	0.00	0.00	0.00			1.000	1.000			0.05	0.54	0.00	0.18	0.00		1.058	1.000
4	8.00	1.00	1.00	0.00	2.00	0.00	0.00	0.00			1.000	1.000			0.05	1.64	0.00	0.00	0.00		1.000	1.000
5	1.50	10.00	3.00	2.00	0.00	0.00	0.00	0.00			1.000	1.000			0.75	0.21	0.00	0.00	0.00		1.000	1.000
6	4.50	6.00	5.75	0.00	0.53	0.00	1.35	0.00			1.307	1.000			0.05	0.51	0.00	1.64	0.00		1.400	1.000
7	7.50	9.00	6.00	0.00	0.00	0.00	0.00	0.00			1.000	1.000			0.05	0.21	0.00	0.23	0.00		1.040	1.000
8	12.00	1.50	2.00	0.00	2.00	0.00	0.00	0.00			1.000	1.000			0.05	1.64	0.00	0.00	0.00		1.000	1.000
9	8.00	10.00	6.00	0.00	0.00	0.00	0.00	0.00			1.000	1.000			0.05	0.21	0.00	0.00	0.00		1.000	1.000
10	1.60	6.00	1.80	5.50	3.50	0.00	0.00	0.00			1.000	1.000			5.50	3.50	0.00	0.00	0.00		1.000	1.000
11	8.00	7.00	4.60	0.00	0.53	0.00	0.00	0.33			0.932	0.928			0.05	0.51	0.17	0.00	0.00		0.964	1.000
12	6.40	6.70	4.00	0.05	0.61	0.24	0.00	0.00			0.943	1.000			0.05	0.54	0.00	0.00	0.00		1.000	1.000
13	10.00	11.60	5.00	0.00	0.00	0.86	0.00	0.15			0.833	0.971			0.05	0.21	1.42	0.00	0.00		0.778	1.000
14	4.00	4.50	3.21	0.05	0.61	0.25	0.00	0.00			0.927	1.000			0.05	0.54	0.01	0.00	0.05		0.982	0.985
15	1.50	8.00	2.12	0.75	0.21	0.00	0.00	0.46			0.820	0.781			0.69	0.22	0.44	0.00	0.00		0.827	1.000

Observations				Classical Model								Modified Model										
DMU	X1	X2	Y	w1j	w2j	u+	u-	v	c	z	Deviation Ratio	Efficiency Estimates	$\beta 1$	$\beta 2$	w1j	w2j	u+	u-	v	z	Deviation Ratio	Efficiency Estimates
1	1.00	7.00	2.00	2.00	0.00	0.00	0.00	0.00	1.000	3.616	1.000	1.000	0.050	0.220	0.68	0.22	0.00	0.00	0.00	4.208	1.000	1.000
2	3.00	4.00	2.50	0.45	1.26	0.00	0.00	0.00			1.000	1.000			0.46	1.26	0.00	0.00	0.00		1.000	1.000
3	3.50	4.30	3.32	0.06	0.58	0.00	0.22	0.00			1.070	1.000			0.06	0.56	0.00	0.21	0.00		1.069	1.000
4	8.00	1.00	1.00	0.00	2.00	0.00	0.00	0.00			1.000	1.000			0.05	1.60	0.00	0.00	0.00		1.000	1.000
5	1.50	10.00	3.00	2.00	0.00	0.00	0.00	0.00			1.000	1.000			0.68	0.22	0.00	0.00	0.00		1.000	1.000
6	4.50	6.00	5.75	0.05	0.56	0.00	1.59	0.00			1.383	1.000			0.05	0.50	0.00	1.66	0.00		1.407	1.000
7	7.50	9.00	6.00	0.00	0.00	0.00	0.00	0.00			1.000	1.000			0.05	0.22	0.00	0.25	0.00		1.043	1.000
8	12.00	1.50	2.00	0.00	0.00	0.00	0.00	0.00			1.000	1.000			0.05	1.60	0.00	0.00	0.00		1.000	1.000
9	8.00	10.00	6.00	0.00	2.00	0.00	0.00	0.00			1.000	1.000			0.05	0.22	0.00	0.00	0.00		1.000	1.000
10	1.60	6.00	1.80	5.50	3.50	0.00	0.00	0.00			1.000	1.000			5.50	3.50	0.00	0.00	0.00		1.000	1.000
11	8.00	7.00	4.60	0.05	0.56	0.30	0.00	0.00			0.939	1.000			0.05	0.50	0.19	0.00	0.00		0.961	1.000
12	6.40	6.70	4.00	0.06	0.58	0.03	0.00	0.00			0.993	1.000			0.06	0.56	0.00	0.00	0.00		1.000	1.000
13	10.00	11.60	5.00	0.00	0.00	1.00	0.00	0.00			0.833	1.000			0.05	0.22	1.43	0.00	0.02		0.775	0.995
14	4.00	4.50	3.21	0.06	0.58	0.04	0.00	0.00			0.987	1.000			0.06	0.56	0.04	0.00	0.00		0.989	1.000
15	1.50	8.00	2.12	0.68	0.22	0.44	0.00	0.00			0.828	1.000			0.68	0.22	0.44	0.00	0.00		0.828	1.000

Appendix C: Combined DEA and quantile regression model compared with the suggested modified model (dataset 1)

DMU	Generated Data							SDEA				Modified SDEA			
	$X \sim U(1,10)$	$u \sim N(0,0.1)$	$v \sim N(0,0.2)$	$\varepsilon = u - v$	$\exp(\varepsilon)$	$X^{0.5}$	$\exp(\varepsilon) * (X^{0.5})$	w	u	v	SE	w	u	v	SE
1	4.393	-0.257	0.156	-0.413	0.662	2.096	1.387	0.154	0.000	0.800	0.008	0.154	0.000	0.800	0.008
2	2.982	0.088	0.103	-0.015	0.985	1.727	1.702	0.318	0.000	0.098	0.005	0.318	0.000	0.098	0.005
3	1.472	0.031	0.285	-0.254	0.776	1.213	0.942	0.994	0.000	0.094	0.032	0.994	0.000	0.094	0.032
4	4.019	0.108	0.048	0.060	1.062	2.005	2.130	0.154	0.000	0.000	0.016	0.154	0.000	0.000	0.016
5	9.606	-0.115	0.108	-0.224	0.799	3.099	2.478	0.000	0.000	0.453	0.028	0.086	0.000	0.477	0.021
6	8.951	0.027	0.211	-0.184	0.832	2.992	2.490	0.154	0.000	0.398	0.011	0.154	0.000	0.398	0.011
7	5.648	0.100	0.200	-0.101	0.904	2.377	2.149	0.154	0.000	0.231	0.000	0.154	0.000	0.231	0.000
8	3.648	-0.127	0.014	-0.141	0.868	1.910	1.658	0.318	0.000	0.354	0.010	0.318	0.000	0.354	0.010
9	9.290	-0.088	0.025	-0.113	0.893	3.048	2.723	0.086	0.000	0.205	0.014	0.086	0.000	0.205	0.014
10	1.645	0.002	0.255	-0.253	0.777	1.282	0.996	0.944	0.000	0.202	0.007	0.944	0.000	0.202	0.007
11	2.900	-0.052	0.180	-0.232	0.793	1.703	1.350	0.318	0.000	0.424	0.005	0.318	0.000	0.424	0.005
12	6.892	0.140	0.261	-0.120	0.887	2.625	2.327	0.154	0.000	0.244	0.003	0.154	0.000	0.244	0.003
13	9.112	0.060	0.287	-0.227	0.797	3.019	2.405	0.154	0.000	0.507	0.011	0.154	0.000	0.507	0.011
14	9.371	-0.019	0.036	-0.054	0.947	3.061	2.899	0.000	0.000	0.031	0.017	0.086	0.000	0.035	0.016
15	2.384	0.083	0.042	0.041	1.042	1.544	1.609	0.318	0.000	0.000	0.004	0.318	0.000	0.000	0.004
16	9.116	0.047	0.060	-0.013	0.987	3.019	2.980	0.154	0.067	0.000	0.011	0.154	0.067	0.000	0.011
17	7.533	-0.016	0.179	-0.195	0.823	2.745	2.258	0.154	0.000	0.412	0.006	0.154	0.000	0.412	0.006
18	8.026	-0.071	0.292	-0.363	0.695	2.833	1.970	0.154	0.000	0.776	0.008	0.154	0.000	0.776	0.008
19	1.993	-0.023	0.460	-0.483	0.617	1.412	0.871	0.318	0.000	0.614	0.005	0.318	0.000	0.614	0.005
20	5.464	-0.041	0.470	-0.512	0.599	2.338	1.401	0.154	0.000	0.950	0.000	0.154	0.000	0.950	0.000
21	7.593	0.066	0.015	0.050	1.052	2.756	2.898	0.154	0.219	0.000	0.006	0.154	0.219	0.000	0.006
22	4.439	0.066	0.189	-0.123	0.884	2.107	1.862	0.154	0.000	0.332	0.008	0.154	0.000	0.332	0.008
23	4.259	0.089	0.004	0.085	1.089	2.064	2.247	0.154	0.081	0.000	0.011	0.154	0.081	0.000	0.011
24	8.304	-0.074	0.081	-0.155	0.857	2.882	2.468	0.154	0.000	0.320	0.009	0.154	0.000	0.320	0.009
25	5.841	-0.022	0.064	-0.087	0.917	2.417	2.216	0.154	0.000	0.194	0.000	0.154	0.000	0.194	0.000
26	4.060	-0.128	0.080	-0.207	0.813	2.015	1.638	0.154	0.000	0.498	0.015	0.154	0.000	0.498	0.015
27	3.211	-0.003	0.203	-0.206	0.814	1.792	1.459	0.318	0.000	0.414	0.007	0.318	0.000	0.414	0.007
28	9.658	0.029	0.231	-0.202	0.817	3.108	2.540	0.000	0.000	0.390	0.031	0.086	0.000	0.419	0.022
29	5.075	-0.077	0.128	-0.206	0.814	2.253	1.834	0.154	0.000	0.458	0.002	0.154	0.000	0.458	0.002
30	1.926	0.138	0.026	0.112	1.118	1.388	1.552	0.318	0.088	0.000	0.006	0.318	0.088	0.000	0.006
31	6.825	-0.023	0.035	-0.058	0.943	2.612	2.464	0.154	0.000	0.096	0.003	0.154	0.000	0.096	0.003
32	9.591	-0.042	0.013	-0.055	0.946	3.097	2.930	0.000	0.000	0.000	0.028	0.086	0.000	0.023	0.021
33	8.381	0.055	0.132	-0.077	0.926	2.895	2.680	0.154	0.000	0.120	0.009	0.154	0.000	0.120	0.009
34	7.155	-0.213	0.005	-0.218	0.804	2.675	2.151	0.154	0.000	0.461	0.004	0.154	0.000	0.461	0.004
35	7.050	-0.049	0.227	-0.275	0.759	2.655	2.016	0.154	0.000	0.579	0.004	0.154	0.000	0.579	0.004
36	2.998	0.068	0.068	0.000	1.000	1.732	1.732	0.318	0.000	0.073	0.005	0.318	0.000	0.073	0.005
37	6.420	0.058	0.152	-0.094	0.910	2.534	2.306	0.154	0.000	0.192	0.001	0.154	0.000	0.192	0.001
38	6.897	0.062	0.136	-0.074	0.929	2.626	2.439	0.154	0.000	0.133	0.003	0.154	0.000	0.133	0.003
39	9.319	0.121	0.044	0.077	1.080	3.053	3.296	0.000	0.366	0.000	0.015	0.086	0.366	0.000	0.015
40	4.926	-0.046	0.048	-0.094	0.910	2.219	2.021	0.154	0.000	0.248	0.002	0.154	0.000	0.248	0.002
41	2.352	-0.128	0.087	-0.215	0.807	1.534	1.237	0.318	0.000	0.362	0.004	0.318	0.000	0.362	0.004
42	6.820	0.118	0.260	-0.142	0.868	2.612	2.266	0.154	0.000	0.294	0.003	0.154	0.000	0.294	0.003
43	2.526	0.055	0.123	-0.068	0.934	1.589	1.485	0.318	0.000	0.170	0.004	0.318	0.000	0.170	0.004
44	8.840	-0.113	0.144	-0.256	0.774	2.973	2.301	0.154	0.000	0.570	0.010	0.154	0.000	0.570	0.010
45	6.756	0.079	0.043	0.036	1.036	2.599	2.693	0.154	0.143	0.000	0.002	0.154	0.143	0.000	0.002
46	9.997	-0.007	0.106	-0.113	0.893	3.162	2.825	0.000	0.000	0.106	0.053	0.086	0.000	0.164	0.030
47	8.749	0.082	0.152	-0.069	0.933	2.958	2.759	0.154	0.000	0.097	0.010	0.154	0.000	0.097	0.010
48	6.540	0.132	0.148	-0.016	0.984	2.557	2.517	0.154	0.000	0.000	0.002	0.154	0.000	0.000	0.002
49	1.268	-0.053	0.237	-0.290	0.748	1.126	0.843	0.944	0.000	0.000	0.080	0.944	0.000	0.000	0.080
50	6.879	-0.096	0.221	-0.317	0.728	2.623	1.910	0.154	0.000	0.659	0.003	0.154	0.000	0.659	0.003

Appendix D: Combined DEA and quantile regression model compared with the suggested modified model (dataset 2)

DMU	Generated Data							SDEA				Modified SDEA			
	$X \sim U(1,10)$	$u \sim N(0,0.1)$	$v \sim N(0,0.2) $	$\epsilon = u - v$	$\exp(\epsilon)$	$X^{0.5}$	$\exp(\epsilon) * (X^{0.5})$	w	u	v	SE	w	u	v	SE
1	2.283	0.112	0.398	-0.286	0.751	1.511	1.135	0.406	0.000	0.257	0.014	0.406	0.000	0.257	0.014
2	6.690	0.108	0.109	-0.001	0.999	2.586	2.583	0.171	0.000	0.002	0.000	0.171	0.000	0.002	0.000
3	2.223	-0.045	0.045	-0.090	0.914	1.491	1.362	0.406	0.000	0.005	0.015	0.406	0.000	0.005	0.015
4	5.716	-0.061	0.094	-0.155	0.856	2.391	2.048	0.237	0.000	0.336	0.000	0.237	0.000	0.336	0.000
5	6.737	-0.121	0.522	-0.643	0.526	2.596	1.365	0.171	0.000	1.229	0.000	0.171	0.000	1.229	0.000
6	7.085	-0.018	0.026	-0.045	0.956	2.662	2.545	0.171	0.000	0.108	0.000	0.171	0.000	0.108	0.000
7	8.036	0.199	0.259	-0.059	0.942	2.835	2.671	0.171	0.000	0.145	0.000	0.171	0.000	0.145	0.000
8	4.727	0.020	0.154	-0.134	0.875	2.174	1.902	0.237	0.000	0.248	0.001	0.237	0.000	0.248	0.001
9	6.422	0.009	0.062	-0.052	0.949	2.534	2.405	0.171	0.000	0.135	0.000	0.171	0.000	0.135	0.000
10	6.493	-0.276	0.046	-0.322	0.725	2.548	1.847	0.171	0.000	0.705	0.000	0.171	0.000	0.705	0.000
11	8.572	0.005	0.015	-0.010	0.990	2.928	2.898	0.000	0.000	0.000	0.001	0.171	0.000	0.010	0.000
12	9.695	-0.165	0.008	-0.173	0.841	3.114	2.618	0.000	0.000	0.280	0.047	0.171	0.000	0.482	0.000
13	8.457	0.155	0.222	-0.067	0.935	2.908	2.720	0.171	0.000	0.168	0.000	0.171	0.000	0.168	0.000
14	5.601	-0.011	0.000	-0.011	0.989	2.367	2.340	0.237	0.000	0.016	0.000	0.237	0.000	0.016	0.000
15	1.404	0.055	0.205	-0.150	0.861	1.185	1.020	0.463	0.000	0.000	0.027	0.463	0.000	0.000	0.027
16	4.080	0.027	0.246	-0.219	0.803	2.020	1.623	0.237	0.000	0.374	0.001	0.237	0.000	0.374	0.001
17	2.095	0.086	0.269	-0.183	0.833	1.447	1.206	0.406	0.000	0.109	0.018	0.406	0.000	0.109	0.018
18	1.653	-0.050	0.074	-0.124	0.883	1.286	1.136	0.406	0.000	0.000	0.023	0.406	0.000	0.000	0.023
19	4.982	-0.186	0.113	-0.299	0.742	2.232	1.656	0.237	0.000	0.554	0.000	0.237	0.000	0.554	0.000
20	5.799	0.008	0.048	-0.040	0.961	2.408	2.314	0.237	0.000	0.090	0.000	0.237	0.000	0.090	0.000
21	7.413	-0.074	0.007	-0.081	0.922	2.723	2.511	0.171	0.000	0.198	0.000	0.171	0.000	0.198	0.000
22	6.256	0.251	0.099	0.152	1.164	2.501	2.913	0.237	0.401	0.000	0.000	0.237	0.401	0.000	0.000
23	4.033	0.017	0.293	-0.275	0.759	2.008	1.525	0.237	0.000	0.460	0.001	0.237	0.000	0.460	0.001
24	7.437	0.000	0.090	-0.090	0.914	2.727	2.492	0.171	0.000	0.222	0.000	0.171	0.000	0.222	0.000
25	8.988	0.097	0.184	-0.086	0.917	2.998	2.751	0.000	0.000	0.147	0.010	0.171	0.000	0.228	0.000
26	5.206	0.204	0.212	-0.008	0.992	2.282	2.263	0.237	0.000	0.000	0.000	0.237	0.000	0.000	0.000
27	6.872	0.038	0.056	-0.018	0.982	2.621	2.575	0.171	0.000	0.041	0.000	0.171	0.000	0.041	0.000
28	4.651	-0.038	0.121	-0.159	0.853	2.157	1.840	0.237	0.000	0.292	0.001	0.237	0.000	0.292	0.001
29	4.330	0.160	0.111	0.049	1.050	2.081	2.186	0.237	0.130	0.000	0.001	0.237	0.130	0.000	0.001
30	2.535	-0.073	0.139	-0.213	0.808	1.592	1.287	0.406	0.000	0.207	0.010	0.406	0.000	0.207	0.010
31	2.343	0.033	0.111	-0.078	0.925	1.531	1.416	0.406	0.000	0.000	0.013	0.406	0.000	0.000	0.013
32	4.089	0.003	0.293	-0.290	0.748	2.022	1.513	0.237	0.000	0.485	0.001	0.237	0.000	0.485	0.001
33	2.537	-0.066	0.112	-0.177	0.838	1.593	1.334	0.406	0.000	0.161	0.010	0.406	0.000	0.161	0.010
34	5.445	-0.081	0.032	-0.113	0.893	2.333	2.083	0.237	0.000	0.236	0.000	0.237	0.000	0.236	0.000
35	9.137	0.146	0.242	-0.096	0.908	3.023	2.745	0.000	0.000	0.153	0.016	0.171	0.000	0.260	0.000
36	5.080	0.075	0.271	-0.197	0.822	2.254	1.852	0.237	0.000	0.382	0.000	0.237	0.000	0.382	0.000
37	2.632	0.035	0.120	-0.084	0.919	1.622	1.491	0.406	0.000	0.042	0.008	0.406	0.000	0.042	0.008
38	8.543	0.013	0.012	0.001	1.001	2.923	2.927	0.000	0.029	0.000	0.001	0.171	0.024	0.000	0.000
39	4.493	0.005	0.387	-0.382	0.683	2.120	1.447	0.237	0.000	0.647	0.001	0.237	0.000	0.647	0.001
40	8.513	0.169	0.079	0.090	1.094	2.918	3.192	0.171	0.295	0.000	0.000	0.171	0.295	0.000	0.000
41	3.339	0.163	0.040	0.124	1.132	1.827	2.068	0.237	0.247	0.000	0.000	0.237	0.247	0.000	0.000
42	6.385	0.081	0.060	0.021	1.021	2.527	2.579	0.171	0.046	0.000	0.000	0.171	0.046	0.000	0.000
43	3.550	0.026	0.178	-0.152	0.859	1.884	1.619	0.237	0.000	0.252	0.000	0.237	0.000	0.252	0.000
44	8.715	-0.059	0.333	-0.393	0.675	2.952	1.994	0.000	0.000	0.904	0.003	0.171	0.000	0.938	0.000
45	9.739	-0.162	0.099	-0.261	0.771	3.121	2.405	0.000	0.000	0.493	0.050	0.171	0.000	0.703	0.000
46	9.531	0.057	0.144	-0.088	0.916	3.087	2.828	0.000	0.000	0.069	0.036	0.171	0.000	0.243	0.000
47	6.101	-0.092	0.108	-0.200	0.819	2.470	2.022	0.237	0.000	0.453	0.000	0.237	0.000	0.453	0.000
48	3.208	0.015	0.150	-0.136	0.873	1.791	1.564	0.406	0.000	0.203	0.001	0.406	0.000	0.203	0.001
49	1.856	-0.085	0.311	-0.396	0.673	1.362	0.917	0.406	0.000	0.301	0.021	0.406	0.000	0.301	0.021
50	9.082	-0.113	0.041	-0.154	0.857	3.014	2.584	0.000	0.000	0.314	0.013	0.171	0.000	0.411	0.000

Appendix E: Sensitivity analysis report for the weight “c” using the data in section 3.5.2

W.L: w_{ij}

Up.L: u_j^+

Un.L: u_j^-

V.L: v_j

----- 54 PARAMETER result

	Value of C	Z
1	0.000	0.000
2	0.050	0.233
3	0.100	0.465
4	0.150	0.698
5	0.200	0.930
6	0.250	1.163
7	0.300	1.395
8	0.350	1.628
9	0.400	1.860
10	0.450	2.093
11	0.500	2.325
12	0.550	2.558
13	0.600	2.791
14	0.650	3.023
15	0.700	3.256
16	0.750	3.393
17	0.800	3.454
18	0.850	3.501
19	0.900	3.548
20	0.950	3.595
21	1.000	3.616
22	1.050	3.616
23	1.100	3.616
24	1.150	3.616
25	1.200	3.616
26	1.250	3.616

----- 52 PARAMETER C = 0.000

----- 52 VARIABLE W.L

	1	2	3	4	5	6
1	2.000	0.670	0.400		2.000	
2		1.617	1.248	2.000		0.083
+	8	10	11	12	14	15
1		5.500			0.400	2.000
2	2.000	3.500	0.083	0.833	1.248	

----- 52 VARIABLE Up.L

(ALL 0.000)

----- 52 VARIABLE Un.L

(ALL 0.000)

---- 52 VARIABLE V.L

11 1.233, 12 1.383, 13 1.000, 14 0.560, 15 0.880

---- 52 PARAMETER C = 0.050

---- 52 VARIABLE W.L

	1	2	3	4	5	6
1	2.000	0.670	0.372		2.000	
2		1.617	1.211	2.000		0.083
+	8	10	11	12	14	15
1		5.500		0.121	0.372	1.114
2	2.000	3.500	0.083	1.034	1.211	0.148

---- 52 VARIABLE Up.L

(ALL 0.000)

---- 52 VARIABLE Un.L

(ALL 0.000)

---- 52 VARIABLE V.L

11 1.233, 12 1.295, 13 1.000, 14 0.538, 15 0.585

---- 52 PARAMETER C = 0.100

---- 52 VARIABLE W.L

	1	2	3	4	5	6
1	2.000	0.670	0.372		2.000	
2		1.617	1.211	2.000		0.083
+	8	10	11	12	14	15
1		5.500		0.121	0.372	1.114
2	2.000	3.500	0.083	1.034	1.211	0.148

---- 52 VARIABLE Up.L

(ALL 0.000)

---- 52 VARIABLE Un.L

(ALL 0.000)

---- 52 VARIABLE V.L

11 1.233, 12 1.295, 13 1.000, 14 0.538, 15 0.585

---- 52 PARAMETER C = 0.150

---- 52 VARIABLE W.L

	1	2	3	4	5	6
1	2.000	0.670	0.372		2.000	
2		1.617	1.211	2.000		0.083
+	8	10	11	12	14	15
1		5.500		0.121	0.372	1.114

2 2.000 3.500 0.083 1.034 1.211 0.148

---- 52 VARIABLE Up.L
(ALL 0.000)

---- 52 VARIABLE Un.L
(ALL 0.000)

---- 52 VARIABLE V.L

11 1.233, 12 1.295, 13 1.000, 14 0.538, 15 0.585

---- 52 PARAMETER C = 0.200

---- 52 VARIABLE W.L

	1	2	3	4	5	6
1	2.000	0.670	0.372		2.000	
2		1.617	1.211	2.000		0.083
+	8	10	11	12	14	15
1		5.500		0.121	0.372	1.114
2	2.000	3.500	0.083	1.034	1.211	0.148

---- 52 VARIABLE Up.L
(ALL 0.000)

---- 52 VARIABLE Un.L
(ALL 0.000)

---- 52 VARIABLE V.L

11 1.233, 12 1.295, 13 1.000, 14 0.538, 15 0.585

---- 52 PARAMETER C = 0.250

---- 52 VARIABLE W.L

	1	2	3	4	5	6
1	2.000	0.670	0.372		2.000	
2		1.617	1.211	2.000		0.083
+	8	10	11	12	14	15
1		5.500		0.121	0.372	1.114
2	2.000	3.500	0.083	1.034	1.211	0.148

---- 52 VARIABLE Up.L
(ALL 0.000)

---- 52 VARIABLE Un.L
(ALL 0.000)

---- 52 VARIABLE V.L

11 1.233, 12 1.295, 13 1.000, 14 0.538, 15 0.585

---- 52 PARAMETER C = 0.300

```

----- 52 VARIABLE W.L
          1          2          3          4          5          6
1         2.000        0.670        0.372          2.000        2.000
2         1.617        1.211        2.000          0.083
+          8          10          11          12          14          15
1         5.500        0.121        0.372        1.114
2         2.000        3.500        0.083        1.034        1.211        0.148

```

```

----- 52 VARIABLE Up.L
          ( ALL          0.000 )

```

```

----- 52 VARIABLE Un.L
          ( ALL          0.000 )

```

```

----- 52 VARIABLE V.L
11 1.233,    12 1.295,    13 1.000,    14 0.538,    15 0.585

```

```

----- 52 PARAMETER C          =          0.350

```

```

----- 52 VARIABLE W.L
          1          2          3          4          5          6
1         2.000        0.670        0.372          2.000        2.000
2         1.617        1.211        2.000          0.083
+          8          10          11          12          14          15
1         5.500        0.121        0.372        1.114
2         2.000        3.500        0.083        1.034        1.211        0.148

```

```

----- 52 VARIABLE Up.L
          ( ALL          0.000 )

```

```

----- 52 VARIABLE Un.L
          ( ALL          0.000 )

```

```

----- 52 VARIABLE V.L
11 1.233,    12 1.295,    13 1.000,    14 0.538,    15 0.585

```

```

----- 52 PARAMETER C          =          0.400

```

```

----- 52 VARIABLE W.L
          1          2          3          4          5          6
1         2.000        0.670        0.372          2.000        2.000
2         1.617        1.211        2.000          0.083
+          8          10          11          12          14          15
1         5.500        0.121        0.372        1.114
2         2.000        3.500        0.083        1.034        1.211        0.148

```

```

----- 52 VARIABLE Up.L
          ( ALL          0.000 )

```

```

----- 52 VARIABLE Un.L

```

(ALL 0.000)

---- 52 VARIABLE V.L

11 1.233, 12 1.295, 13 1.000, 14 0.538, 15 0.585

---- 52 PARAMETER C = 0.450

---- 52 VARIABLE W.L

	1	2	3	4	5	6
1	2.000	0.670	0.372		2.000	
2		1.617	1.211	2.000		0.083
+	8	10	11	12	14	15
1		5.500		0.121	0.372	1.114
2	2.000	3.500	0.083	1.034	1.211	0.148

---- 52 VARIABLE Up.L

(ALL 0.000)

---- 52 VARIABLE Un.L

(ALL 0.000)

---- 52 VARIABLE V.L

11 1.233, 12 1.295, 13 1.000, 14 0.538, 15 0.585

---- 52 PARAMETER C = 0.500

---- 52 VARIABLE W.L

	1	2	3	4	5	6
1	2.000	0.670	0.372		2.000	
2		1.617	1.211	2.000		0.083
+	8	10	11	12	14	15
1		5.500		0.121	0.372	1.114
2	2.000	3.500	0.083	1.034	1.211	0.148

---- 52 VARIABLE Up.L

(ALL 0.000)

---- 52 VARIABLE Un.L

(ALL 0.000)

---- 52 VARIABLE V.L

11 1.233, 12 1.295, 13 1.000, 14 0.538, 15 0.585

---- 52 PARAMETER C = 0.550

---- 52 VARIABLE W.L

	1	2	3	4	5	6
1	2.000	0.670	0.372		2.000	
2		1.617	1.211	2.000		0.083
+	8	10	11	12	14	15
1		5.500		0.121	0.372	1.114

2 2.000 3.500 0.083 1.034 1.211 0.148

---- 52 VARIABLE Up.L
(ALL 0.000)

---- 52 VARIABLE Un.L
(ALL 0.000)

---- 52 VARIABLE V.L

11 1.233, 12 1.295, 13 1.000, 14 0.538, 15 0.585

---- 52 PARAMETER C = 0.600

---- 52 VARIABLE W.L

	1	2	3	4	5	6
1	2.000	0.670	0.372		2.000	
2		1.617	1.211	2.000		0.083
+	8	10	11	12	14	15
1		5.500		0.121	0.372	1.114
2	2.000	3.500	0.083	1.034	1.211	0.148

---- 52 VARIABLE Up.L
(ALL 0.000)

---- 52 VARIABLE Un.L
(ALL 0.000)

---- 52 VARIABLE V.L

11 1.233, 12 1.295, 13 1.000, 14 0.538, 15 0.585

---- 52 PARAMETER C = 0.650

---- 52 VARIABLE W.L

	1	2	3	4	5	6
1	2.000	0.670	0.372		2.000	
2		1.617	1.211	2.000		0.083
+	8	10	11	12	14	15
1		5.500		0.121	0.372	1.114
2	2.000	3.500	0.083	1.034	1.211	0.148

---- 52 VARIABLE Up.L
(ALL 0.000)

---- 52 VARIABLE Un.L
(ALL 0.000)

---- 52 VARIABLE V.L

11 1.233, 12 1.295, 13 1.000, 14 0.538, 15 0.585

---- 52 PARAMETER C = 0.700

```

----- 52 VARIABLE W.L
          1          2          3          4          5          6
1         2.000        0.670        0.372          2.000        2.000
2         1.617        1.211        2.000          0.083
+          8          10          11          12          14          15
1         5.500        0.121        0.372        1.114
2         2.000        3.500        0.083        1.034        1.211        0.148

```

```

----- 52 VARIABLE Up.L
          ( ALL      0.000 )

```

```

----- 52 VARIABLE Un.L
          ( ALL      0.000 )

```

```

----- 52 VARIABLE V.L
11 1.233,   12 1.295,   13 1.000,   14 0.538,   15 0.585

```

```

----- 52 PARAMETER C = 0.750

```

```

----- 52 VARIABLE W.L
          1          2          3          4          5          6
1         2.000        0.670        0.139          2.000
2         1.617        0.892        2.000          0.342
+          8          10          11          12          14          15
1         5.500        0.139        0.139        0.902
2         2.000        3.500        0.342        0.892        0.892        0.183

```

```

----- 52 VARIABLE Up.L
11 0.131,   12 0.644

```

```

----- 52 VARIABLE Un.L
6 0.775

```

```

----- 52 VARIABLE V.L
11 0.586,   13 1.000,   14 0.358,   15 0.514

```

```

----- 52 PARAMETER C = 0.800

```

```

----- 52 VARIABLE W.L
          1          2          3          4          5          6
1         2.000        0.670        0.045          2.000        2.000
2         1.617        0.609        2.000          0.533
+          8          10          11          12          14          15
1         5.500        0.045        0.045        0.753
2         2.000        3.500        0.533        0.609        0.609        0.208

```

```

----- 52 VARIABLE Up.L
12 0.240,   13 0.646,   15 0.464

```

```

----- 52 VARIABLE Un.L

```

6 1.350

---- 52 VARIABLE V.L

11 0.333, 13 0.354, 14 0.254

---- 52 PARAMETER C = 0.850

---- 52 VARIABLE W.L

	1	2	3	4	5	6
1	2.000	0.670	0.045		2.000	
2		1.617	0.609	2.000		0.533
+	8	10	11	12	14	15
1		5.500		0.045	0.045	0.753
2	2.000	3.500	0.533	0.609	0.609	0.208

---- 52 VARIABLE Up.L

12 0.240, 13 0.646, 15 0.464

---- 52 VARIABLE Un.L

6 1.350

---- 52 VARIABLE V.L

11 0.333, 13 0.354, 14 0.254

---- 52 PARAMETER C = 0.900

---- 52 VARIABLE W.L

	1	2	3	4	5	6
1	2.000	0.670	0.045		2.000	
2		1.617	0.609	2.000		0.533
+	8	10	11	12	14	15
1		5.500		0.045	0.045	0.753
2	2.000	3.500	0.533	0.609	0.609	0.208

---- 52 VARIABLE Up.L

12 0.240, 13 0.855, 14 0.254

---- 52 VARIABLE Un.L

6 1.350

---- 52 VARIABLE V.L

11 0.333, 13 0.145, 15 0.464

---- 52 PARAMETER C = 0.950

---- 52 VARIABLE W.L

	1	2	3	4	5	6
1	2.000	0.670	0.045		2.000	
2		1.617	0.609	2.000		0.533
+	8	10	11	12	14	15
1		5.500		0.045	0.045	0.753

2 2.000 3.500 0.533 0.609 0.609 0.208

---- 52 VARIABLE Up.L

12 0.240, 13 0.855, 14 0.254

---- 52 VARIABLE Un.L

6 1.350

---- 52 VARIABLE V.L

11 0.333, 13 0.145, 15 0.464

---- 52 PARAMETER C = 1.000

---- 52 VARIABLE W.L

	1	2	3	4	5	6
1	2.000	0.454	0.062		2.000	0.050
2		1.257	0.583	2.000		0.564
+	8	10	11	12	14	15
1		5.500	0.050	0.062	0.062	0.679
2	2.000	3.500	0.564	0.583	0.583	0.220

---- 52 VARIABLE Up.L

11 0.298, 12 0.028, 13 1.000, 14 0.042, 15 0.440

---- 52 VARIABLE Un.L

3 0.216, 6 1.592

---- 52 VARIABLE V.L

(ALL 0.000)

---- 52 PARAMETER C = 1.050

---- 52 VARIABLE W.L

	1	2	3	4	5	6
1	2.000	0.454	0.062		2.000	0.050
2		1.257	0.583	2.000		0.564
+	8	10	11	12	14	15
1		5.500	0.050	0.062	0.062	0.679
2	2.000	3.500	0.564	0.583	0.583	0.220

---- 52 VARIABLE Up.L

11 0.298, 12 0.028, 13 1.000, 14 0.042, 15 0.440

---- 52 VARIABLE Un.L

3 0.216, 6 1.592

---- 52 VARIABLE V.L

(ALL 0.000)

---- 52 PARAMETER C = 1.100


```

----      52 VARIABLE W.L
          1          2          3          4          5          6
1         2.000        0.454        0.062          2.000        2.000        0.050
2         1.257        0.583        2.000          0.062        0.564
+          8          10          11          12          14          15
1         5.500        0.050        0.062        0.062        0.679
2         2.000        3.500        0.564        0.583        0.220

```

```

----      52 VARIABLE Up.L
11 0.298,   12 0.028,   13 1.000,   14 0.042,   15 0.440

```

```

----      52 VARIABLE Un.L
3 0.216,   6 1.592

```

```

----      52 VARIABLE V.L
          ( ALL      0.000 )

```

```

----      52 PARAMETER C          =          1.150

```

```

----      52 VARIABLE W.L
          1          2          3          4          5          6
1         2.000        0.454        0.062          2.000        2.000        0.050
2         1.257        0.583        2.000          0.062        0.564
+          8          10          11          12          14          15
1         5.500        0.050        0.062        0.062        0.679
2         2.000        3.500        0.564        0.583        0.220

```

```

----      52 VARIABLE Up.L
11 0.298,   12 0.028,   13 1.000,   14 0.042,   15 0.440

```

```

----      52 VARIABLE Un.L
3 0.216,   6 1.592

```

```

----      52 VARIABLE V.L
          ( ALL      0.000 )

```

```

----      52 PARAMETER C          =          1.200

```

```

----      52 VARIABLE W.L
          1          2          3          4          5          6
1         2.000        0.454        0.062          2.000        2.000        0.050
2         1.257        0.583        2.000          0.062        0.564
+          8          10          11          12          14          15
1         5.500        0.050        0.062        0.062        0.679
2         2.000        3.500        0.564        0.583        0.220

```

```

----      52 VARIABLE Up.L
11 0.298,   12 0.028,   13 1.000,   14 0.042,   15 0.440

```

```

----      52 VARIABLE Un.L

```

```

3 0.216,      6 1.592

----          52 VARIABLE V.L
                ( ALL          0.000 )

-----
----          52 PARAMETER C                      =          1.250
----          52 VARIABLE W.L
                1          2          3          4          5          6
1          2.000          0.454          0.062          2.000          0.050
2          1.257          0.583          2.000          2.000          0.564
+          8          10          11          12          14          15
1          5.500          0.050          0.062          0.062          0.679
2          2.000          3.500          0.564          0.583          0.220

----          52 VARIABLE Up.L
11 0.298,      12 0.028,      13 1.000,      14 0.042,      15 0.440

----          52 VARIABLE Un.L
3 0.216,      6 1.592

----          52 VARIABLE V.L
                ( ALL          0.000 )

```

the coding in GAMS software for sensitivity analysis of the weight “c”

```

Sets
i index of input /1,2/
r index of output /1/
j index of DMUs /1*11/
Iteration /1*30/;
Alias (k,j);
Scalar C /0.95/;

Parameter
result(Iteration,*);

table
X(i,j) Inputs
      1      2      3      4      5      6      7      8      9      10      11
1      1      3      3.5      8      1.5      4.5      7.5      12      8      1.6      8
2      7      4      4.3      1      10      6      9      1.5      10      6      7 ;

table
Y(r,j) Output
      1      2      3      4      5      6      7      8      9      10      11
1      2      2.5      3.32      1      3      5.75      6      2      6      1.8      2 ;

Variables
Z
Positive variable W
Positive Variable Up
Positive Variable Un
Positive Variable V ;

Equations
ObjectiveFunction
Col(j,k)
Co2;
*Co3(j);

ObjectiveFunction .. Z =e= sum((j),Up(j)+Un(j)+C*V(j)) ;

```

```

Co1(j,k)          .. Sum((i), (X(i,k)-X(i,j))*W(i,j))+V(j)-
V(k)+Up(j)-Un(j)-Up(k)+Un(k) =g= Sum((r), Y(r,k)-Y(r,j)) ;
Co2               .. sum((j), Up(j)-Un(j)) =e= 0 ;
*Co3(j)          .. W("1",j) =g= 0.333;

```

```

Model StochasticDEA /all/;

```

```

Loop (Iteration,
C = C + 0.05
Solve StochasticDEA using LP Minimizing Z;
result(Iteration, "Value of C")= C;
result(Iteration, "Z") = Z.l;
*result(Iteration, "W") = W.l;
*result(Iteration, "Up") = Up.l;
*result(Iteration, "Un") = Un.l;
*result(Iteration, "V") = V.l;
);
Display result;
execute_unload 'FirstResult.gdx';

```